Contribution to the recent history of archaeology by using some digital humanities methods and techniques applied to field recording documents of an archaeological site excavated in 1970s

Christophe Tuffery

To cite this version:
Christophe Tuffery. Contribution to the recent history of archaeology by using some digital humanities methods and techniques applied to field recording documents of an archaeological site excavated in 1970s. Journal of Data Mining and Digital Humanities, In press, Atelier Digit_Hum, 10.46298/jd-mdh.10847. hal-03931854v3

HAL Id: hal-03931854
https://hal.science/hal-03931854v3
Submitted on 14 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Contribution to the recent history of archaeology that applies digital humanities methods and techniques to field recording documents from an archaeological site excavated in 1970s

Christophe Tufféry¹*

1 Institut national de recherches archéologiques préventives (Inrap)

*Corresponding author: Christophe Tufféry christophe.tuffery@culture.gouv.fr

Abstract
This article presents the results of an archaeological archive research project. Field recording documents from the Rivaux site in France which was excavated from the 1970s to the 1990s were exploited. After digitising a set of field notebook pages, the author developed an application called Archeotext which enables these documents to be transcribed and georeferenced. Some of the results obtained show new ways of exploiting this type of archive by using certain methods and techniques from the digital humanities.¹

Keywords
archaeology; archives; field notebooks; recording methods; digital humanities

INTRODUCTION

Many archaeologists have written about the recent history of archaeology [Giligny, 2011, Djindjian, 2016a and 2016b, Plutniak, 2017a and 2017b, Favory et al. 2018, Giligny and Desachy, 2019, Tufféry and Augry, 2021]. The use of archaeological archives has already been the subject of several research projects, notably on an international scale such as the European project Archives of European Archaeology [Schnapp et al., 2007]. These works are stimulating contributions to the history of archaeological science and also provide a useful overview of the methods used in the past and the results already obtained. Some of these projects use documentation from ancient archaeological excavations and rely on artificial intelligence for automatic transcription using an approach inspired by the principles of participatory science [Girard et al., 2019]. In some publications, researchers use their own memories and their own archives from several decades before the periods in which they themselves took part in excavations [Audouze, 2003, Djindjian, 2002, O'Brien and al., 2005]. In other cases, this kind of historical work could only be carried out after the disappearance of the people concerned and by people who were not involved either in the place, the excavations or in the period these works of recent history of the archaeological discipline focus on.

The proposed account covers a period and events in which I was personally involved and so might seem to lack a minimum of objectivity. However, the author's personal involvement in this type of research does not mean it is totally impossible to offer a distanced account. Of

¹ This research was carried out in the context of my thesis at CY Cergy Paris Université [Tufféry, 2023].
course, our memories may introduce biases and an *a posteriori* interpretation that ‘rewrites’ history. This is what constitutes the “biographical illusion” [Bourdieu, 2004]). However, as Gautier and Zancarini-Fournel state, “describing what is real supposes adopting a point of view from which to see the thing that is to be described (...) Adopting a particular point of view means admitting that other, differing, points of view may be possible. Thus it means positively accepting engaging in a confrontation of points of view to reach a more finely graduated description of something we claim to know, emerging from the plurality of these perspectives” [Gautier and Zancarini-Fournel, 2022:9] (author’s translation). The results presented here have the advantage of using original materials and well-established principles of analysis even if these are presented by a researcher involved in the research object concerned.

Memory inevitably has some limitations and indeed carries some risk of falsification. However it also has a number of advantages such as the possibility of compensating for the weaknesses of certain archives, at least in part. Taking stock of one’s own history by looking back at one’s trajectory and experience should therefore not be totally discounted from the outset. The exercise of reflexivity, self-analysis and ego-history have been theorised by a number of historians (Maurice Agulhon, Pierre Chaunu, Georges Duby, Raoul Girardet, Jacques le Goff, Michelle Perrot, René Rémond). For archaeology, a majority of authors offer more accounts of their own lives and work than researchers with more limited contributions on the reflexive dimension [Demoule and Jacob, 2012], [Guilaine, 2019] and [Clottes, 2015]. My contribution to the genre is much more modest and does not claim to be as important as that of these authors but it does attempt to add something by reflecting on my own career and my own work as a researcher.

Using a set of archives from an excavation I took part in the 1980s, I propose to demonstrate the possibilities of using digital methods and techniques to exploit and enhance archaeological field records including archives I produced over 40 years ago.

I DATA SOURCES USED

As an archaeologist, my experience in this field began in 1980, working for the first time as a volunteer on digs in France - one on a preventive dig in the centre of Clermont-Ferrand, the other on the scheduled dig at *Les Rivaux* in the Haute-Loire département [Daugas and Raynal, 1977; Daugas and al., 1979]². This site was excavated under the responsibility of Jean-Pierre Daugas (1946-2011) the archaeologist who passed on to us almost everything we know about the archaeological methods and techniques in use in the 1980s.

On this site, the pioneering use of an *Apple Ile* personal computer was experimented with during the digging carried out from 1985 to 1990. A dedicated computerised recording programme called *Rivalo* was also used that made it possible to record in .txt format all the information entered on paper field sheets on-site during the day for each square of the dig. In 2009, I was able to recuperate - albeit not without difficulty - the data produced by this programme after making a specific connector to link an *Apple Ile* to a PC running under Windows [Tufféry, 2022].

In 2019, I was able to consult a number of personal field notebooks (Figure 1) that had belonged to the afore-mentioned Jean-Pierre Daugas³.

---

² *Les Rivaux* site had been known about since the mid-nineteenth century but was mainly excavated starting in the second half of the twentieth century by Pierre Bout and then from the end of the 1960s in the form of prospecting followed by rescue intervention and finally a scheduled dig.

³ I must give my warmest thanks to Jean-Pierre’s wife, Chloë Daugas for agreeing to give me these archives. I would also like to thank Julie Charmoillaux and Christophe Gilabert for making it easier for us to consult *Les Rivaux* archives at the Centre de Conservation et d’Etudes in Les Martres-de-Veyre and for having digitised all the reports available for this dig for our benefit.
These notebooks take the form of daybooks and contain a considerable quantity of scientific observations in the field, lists of people met or to contact, things to be done or equipment to be acquired for current digs. Archives of this type are a hybrid genre, somewhere between a personal journal, scientific field notation and a laboratory notebook. Some of the pages in the notebooks contain several examples of sheets of squared paper filled in by hand by the author or by one of the other researchers with whom he collaborated (Figure 2).

Finally, the documentary sources used in my research include part of the archives of Les Rivaux excavation site already mentioned.

II CHANGES IN RECORDING METHODS: THE EXAMPLE OF THE DIG AT LES RIVAUX

Jean-Pierre Daugas deposited a large part of the archives of the dig at Les Rivaux with the CCE in Les Martres-de-Veyre during his lifetime. These dig archives are similar to many others of the same period that have been deposited with archive services and other Conservation and Study Centres (Centre de Conservation et d’Etudes). They comprise boxes of various types of document - paper sheets used for recording in the field, field notebooks, correspondence, photographs, drawings, sketches, plans, etc. (Figure 3).

The research for my thesis consisted of showing how the methods used for excavation and for recording data changed between the time work on the dig began in 1970 and ended at this site in 1990 through analysis of various types of first-hand sources of documents used on the site. Various models of paper record sheets were used at the site and the examples given here show some of the models used between the early 1970s and the late 1980s.
The various models comprise an upper part of the record sheet with sections for descriptions of the site and the area being excavated and other sections for the structures observed, the various information collected (topography, photography) and the samples taken. There are also handwritten notes in places and references to the field notebook. In another of the models used at *Les Rivaux*, there are two boxes for entering text (description of layers and structures) and one box for a rough map showing the location of the areas being excavated. Lastly, a final box at the foot of the page could be used to list the archaeological material found in the area excavated (Figure 4).

During the following excavation campaigns, other templates of record sheet were used. The following example was used for each square metre and each archaeological layer (Figure 5). My name appears as the author of the record. The date given is the date of the first record on the sheet, not the last. Detailed data on the archaeological material was recorded. Once the find was discovered and detached, its X and Y coordinates within the square were recorded using a tape measure. These values were used to record the exact position and dimensions of the find which subsequently allowed the distribution of the finds in each square and at each archaeological level to be mapped. The finds were also described in terms of orientation, dip, dimensions, basic sketch or shape, designation and nature and the archaeological layer or soil to which they belonged.
Figure 4: An example of a field record sheet used in the mid-1970s.

Figure 5: Another example of field record sheet for locating and describing archaeological finds.

As this example shows, the presence of traces of earth and dirty fingerprints indicate that these sheets were used in the field. As I remember, the forms were attached to an A4-sized piece of lightweight wood (such as balsa) by a black metal clip and were filled in using a pencil. During digging, pencils were either left on the wood backing, stuck in the metal clip, held between the excavator’s teeth or tucked behind the excavator’s ear. This type of form was very similar to many others used in other archaeological digs. The same type of form can be found in numerous publications in France and abroad concerning field recording process and documentation [Lyman L. R. & O’Brien, 2006, Ramirez Galicia, 2019, Plutniak, 2021].

Lastly, there was one other very important documentary source for this site, namely two field notebooks (Figure 6). Between 1972 and 1980, these documents were kept by a number of the excavators. While the first was filled in by a number of the excavators, the second was mainly filled in by the site manager. During the 1980s, these field notebooks gradually gave way to standardised field record sheets like those already presented.

All examples of the other type of archive, as previously presented, have all been deposited as public records at the Conservation and Study Centres where they may be quite easily consulted. However the person in charge of the archaeological site considered the field notebooks to be private documents and therefore he did not deposit them with the Conservation and Study Centres during his lifetime.

This example highlights the hybrid status of this type of documentation in practice. This was the case for many archaeologists who considered this type of document to be a private archive for a long time whereas today it is viewed as clearly public documentation from a legal point of view.
I used a variety of digital tools and methods to analyse the content of these documents and their links with the other documentary sources available for the Les Rivaux.

III DIGITAL TOOLS AND PROCESSING METHODS

3.1 Transcription of field notebooks

First, I decided to digitise and then digitally transcribe the notebooks. Digitising was carried out using a scanner at a resolution of 300 dpi. Displaying the files produced by this digitisation raised a number of problems. Firstly, entries in the notebooks were made in black, blue or sometimes red or green ballpoint pen for the annotations or alternatively in lead pencil that had often been erased and was thus less legible. Secondly, some of the phrases had been crossed out by their authors, making them difficult or even impossible to read. Moreover, the writing on the back of some of the pages shows through on the front because in places the ink used has seeped through the paper of some of the pages. Lastly, it is sometimes difficult to make out some of the sketches and diagrams because here again a number of different inks were used: some have remained clear, while others are now almost illegible.

In an attempt to resolve these problems, I carried out a number of processes using Adobe Illustrator software, particularly its B&W vectorisation function. Combining this function with Adobe Illustrator's ordinary vectorisation function, it was possible to make the inks on the back of the pages concerned disappear. I was thus able to obtain a version of the sketches and diagrams in vector format (Figure 7).
However for some of the sketches the ordinary vectorisation functions in Adobe Illustrator did not produce satisfactory results so I used the free software PhotoFiltre, testing a number of processing functions. I used the Gamma Contrast and Correction functions and achieved results that were sometimes better than those obtained with the ordinary vectorisation function in Adobe Illustrator (Figure 8).

Figure 8: Comparison of two types of processing carried out using two different software programmes on a sketch in a field notebook - PhotoFiltre’s ‘gamma correction’ function on the left, Adobe Illustrator’s ordinary vectorisation function on the right.
Once I had processed all the digitised pages in the field notebook to make the texts and sketches as legible as possible, I decided to develop a digital transcription application which I called *Archeotext*.

The application operates in conjunction with an Internet browser (Chrome, Opera or Safari). It uses a series of forms for entering information in HTML5 format (Figure 9). Alongside the pages in HTML5 format are some JavaScript code files and a CSS-type sheet.

![Figure 9: Screenshot of the Archeotext application. Display of a page in the field notebook (on the left) and the result of its manual transcription (on the right).](image)

The page number indicated by the first script for the notebook page to be displayed in *Archeotext* (on the left in .jpg format) and its transcription (on the right in text format) must be entered to browse the transcribed pages of the field notebook (select a page in the list). The reference for the square to which this page is connected can be found in the Square section in the top left-hand corner of the *Archeotext* application (Figure 9).

These sheets make it possible to enter the transcription data and other descriptive data as well as metadata in accordance with ISO15836 (“Dublin-Core”). All the data is stored in a *SQLite* database on the user’s device. Data can be exported as a single file in .XML format and can also be used directly from the *SQLite* database using computerised management system

---

4 *Archeotext* is not currently released under an open source licence as it was adapted from EDArc, a field recording system developed by the French National Institute for Preventive Archaeological Research (INRAP). It cannot therefore be reused but this restriction could be lifted if the Inrap agreed to open up the source code of the EDArc application which I used to develop *Archeotext*.

5 This application could easily be adapted and used for the archives of other excavation sites.
software such as QGIS without passing via an export file. This possibility also makes it possible to correlate the transcription data and the geometric data of the areas excavated on the site. I transcribed all 150 pages of the most comprehensive of the two field notebooks, covering the period from 1972 to 1982. The second notebook, comprising some fifty pages, covers the period from 1980 to 1990.

Using the eScriptorium on-line application and the LECTAUREP model, I was able to explore its possibilities for a different transcription of Les Rivaux field notebooks from the one I made using Archeotext. The eScriptorium application was developed for transcribing digitised handwritten medieval documents and uses a technique for automatic character recognition called Optical Character Recognition and Handwritten Text Recognition (OCR/HTR). This uses Kraken character recognition developed by Ben Kiessling. It is necessary to begin by training the engine, by manually transcribing the content of the first few pages of a handwritten document (Figure 10). Once the automatic transcription has been carried out and checked, the result can be downloaded in text or XML format. I therefore wanted to use this eScriptorium application as well as the Archeotext application and was able to assess its advantages and limits.

Figure 10: Screenshot of the eScriptorium application. Display of a page in the field notebook (on the left) and result of its transcription (on the right).

Both Archeotext and eScriptorium have their own value and limitations. For example, it is very difficult for scripts to be recognised when they were produced by more than one author because the spellings differ from one author to another. Neither could provide everything I required in researching the archive material represented by these field notebooks namely distinguishing the

---

6 I would like to thank Peter Stokes of Paris Sciences et Lettres (PSL) and Alix Chagué most warmly. The LECTAUREP (LECTure Automatique de REPertoires) project started in 2018. It is a joint initiative run by the Minutier Central des Notaires de Paris of the Archives Nationales, the ALMAAnaCH (Automatic Language Modeling and Analysis & Computational Humanities) team at the French National Institute for Research in Computer Science and Control (Inria) and the EPHE (École Pratique des Hautes Études) in partnership with the Ministry of Culture.

7 This project is part of the ‘Scripta’ Initiatives de Recherches Interdisciplinaires et Stratégiques (IRIS) at Paris Sciences et Lettres (PSL) http://ephenum.hypotheses.org/1412 and http://escripta.hypotheses.org/ consulted on 5 September 2023.
variety of spellings, recognising the content of sketches or identifying other various types of content than writing, etc.

3.2 Integration in computerised management system software

Using the QGIS geographical information system, I was able to retrace the perimeter of the dig from the 1970s to the 1990s by using various types of reference data (old georeferenced rough plan of the dig, the Institut Géographique National’s on-line large-scale cartographic services and the land registry plan).

Once I had digitised the old plan of the dig, I was able to adjust it using the georeferencing tool in QGIS. I then drew a grid of one-metre squares over the area covered by the old excavation plan. This shows that barely a quarter of all the squares in the total area had been investigated (Figure 11).

Figure 11: Display in QGIS of the grid, the limits of the excavation (in red), the southern limits showing the evolution of the erosion of the site, the squares excavated according to the first field notebook (in orange), the other squares excavated between 1980 and 1982 (in pink).

The excavated squares shown in the first field notebook transcribed are mainly in the southern part of the dig, an area that has gradually disappeared as the result of erosion caused by the Borne stream from the late 1970s and through the 1980s.

Data about the archaeological finds comes from the data recorded in .txt format using the Rivalo programme. I was able to recuperate this data and convert the information into .gpkg format after a quantity of processing. In QGIS, I displayed data of the Rivalo programme by square meters (Figure 12). The total material from these files only covers some of the squares, mainly in the central part of the squared area between two of the shoulders and slightly in the north-western part of the footprint of the dig.
The data on the archaeological finds comprises a range of descriptive information which can be accessed by selecting any one of them (Figure 13).

Lastly, the transcription data created using the Archeotext application which is stored in a SQLite database can be consulted directly from QGIS. Using a script, transcribed content can be displayed in a window by clinking on the corresponding square (Figure 14).
The possibility of consulting the pages in the field notebook and the results of their transcription directly using QGIS provides an integrated access to these archives from the cartographic representation of the dig, the location of the dig units (the squares) and the items of archaeological material represented resulting from the data in the Rivalo programme. This integration of multi-source data in computerised management system software is one of the major advantages of using several technologies such as those presented here from a common consultation interface – in this case the cartographic interface of the QGIS software.

Conclusion

Overall, this research has made it possible to show the importance of noting information in the field and the use of certain traditional media in relation to both the exercise of the scientific skill of archaeologists and the life of the social group constituted by a dig team on-site. Archives such as those we studied attest to a variety of practices in terms of noting information in the field and to the fact that the use of digital devices has increased in recent times. Observations at the Inrap since 2010 have shown that field notebooks seem to be losing their importance as a result of the rollout of the digital devices that are replacing them, at least in part. Collective daybooks for digs such as those used at Les Rivaux are not often used nowadays. This is particularly true during preventive archaeology operations which is probably because the time factor exerts greater pressure than is the case for scheduled archaeological operations. At best, the dig manager keeps his/her own daybook. It would seem that this type of documentation is not as popular among archaeologists as it once was, particularly those of the younger generation. The latter generally tend to record data in standard format, sometimes using digital applications on tablets and smartphones rather than on paper sheets. However, this does not mean the traditional media for taking notes in the field have totally disappeared and been replaced by an exclusive use of digital media as I observed the coexistence of both modes of recording information.

In his publications on places of knowledge, Christian Jacob refers to inscriptions as graphic places of knowledge. For [Jacob, 2014:1], both textual or graphic places of knowledge that such
inscriptions constitute “are not only the media for fixing and archiving pre-extant knowledge. They construct that knowledge; objectivise it, materialise it. They are part of the field of view, interpretation and social circulation. They are thus active devices and interfaces between the producer and anyone decoding and using them” (author’s translation). The interactions between knowledge and the related stakeholders in society (at the time of production or decades later) and between those who produce knowledge and those who access it constitute one of the fields of cataloguing, transcribing, decoding and publishing dig archives.

This work echoes that of [Bert and Lamy, 2021:29], revealing the often invisible dimension of the work of researchers in three areas that are hard to dissociate - the places, the objects and the gestures of science. For these authors, “knowledge is inseparable from the conditions of its expression” (author’s translation). I believe it is impossible to separate content from the authors of archaeological knowledge and therefore that archives should not be considered cold, disembodied objects that present the construction of knowledge in material form. These archives bear the traces of knowledge in the form of dynamic processes and human undertakings. The fact that they are now produced in native digital form or that it is now possible to convert them into digital form after they are produced in a non-digital form does not fundamentally alter the fact that they are archives. In this sense, the digital humanities do not make it possible to reach a new level of scientificness, to use Michel Foucault’s definition of the notion of episteme [Foucault, 1969]. The British archaeologist J. Huggett stresses that the application of information technology to archaeology is still relatively recent and that we are thus not in a position to imagine what the results of its effects on the production of archaeological knowledge will be in several decades’ time [Huggett, 2000]. However, the digital humanities already make it possible to renew our relationship with archaeological dig archives, giving them a second life and proposing new ways of using them and perhaps occasionally making it possible to make something visible that seemed hidden, as in the very act of digging.

References


* [Ces lieux de savoir graphiques au sens large] "ne sont pas seulement les supports de fixation et d’archivage d’un savoir qui leur préexisterait. Ils construisent ce savoir, ils l’objectivent, ils le matérialisent. Ils l’inscrivent dans le champ du regard, de l’interprétation et de la circulation sociale. Ce sont donc des dispositifs actifs, des interfaces entre celui qui les a produits et celui qui les déchiffre et les utilise".

9 "les connaissances sont solidaires de leurs conditions d’énonciation".


Gautier C. and Zancarini-Fournel M. De la défense des savoirs critiques. Quand le pouvoir s’en prend à l’autonomie de la recherche. La Découverte (Paris), 2022.


