SAA
SOCIETY FOR AMERICAN ARCHAEOLOGY

ARTICLE

Strategies for Archaeological Data Collection and Management: The Development of a Digital Method in the Lower Gallery of La Garma (Spain)

Carlos García-Noriega[1] (iD), Rodrigo Portero[2], Alba Ruiz-Cabanzón[1] (iD), Roberto Ontañón[3] (iD), and Pablo Arias[1]

[1]Instituto Internacional de Investigaciones Prehistóricas de Cantabria, Universidad de Cantabria-Gobierno de Cantabria-Grupo Santander, Santander, Spain; [2]Departamento de Prehistoria, Historia Antigua y Arqueología, Universidad de Salamanca, Salamanca, Spain; and [3]Museo de Prehistoria y Arqueología de Cantabria, Cuevas Prehistóricas de Cantabria, Gobierno de Cantabria, Santander, Spain

Corresponding author: Carlos García-Noriega; Email: carlos.garcianoriega@unican.es

Abstract

This article presents an innovative workflow for the acquisition and storage of archaeological data. The system is based on open-source software to enhance method replication and media accessibility. QGIS software is used as the central platform, connected to a spatial database developed in PostgreSQL and managed with the SQL and Python programming languages. The aim is to achieve an efficient, flexible, and reproducible digital method for data collection and management that can be applied to surface archaeological surveys. During the implementation and development of the method, we have recorded over 4,600 archaeological remains in two different structures with traces of Upper Paleolithic activity in the Lower Gallery of La Garma (Cantabria, Spain). After 18 months of continuous work, the results obtained demonstrate the usefulness and versatility of this procedure, which can be adapted to each context and to the specific needs of each researcher. Our goal is not simply to systematize archaeological documentation, as traditionally proposed, but to establish a simple and robust method for data collection and preservation, accessible to any user. Its fully open-source approach aims to promote a model that is nurtured by the use and contributions of the research community.

Resumen

En el siguiente trabajo exponemos el desarrollo de un novedoso flujo de trabajo para la captura y almacenamiento de datos. El sistema creado se basa en la utilización de programas de código abierto, que facilitan la reproducción del método y la accesibilidad a los medios, teniendo como objetivo la recolección de datos geoespaciales y la documentación arqueológica de su expresión física. Utilizamos el programa QGIS como plataforma central conectada una base de datos con características espaciales, basada en PostgreSQL, y gestionado con los lenguajes de programación SQL y Python. La premisa de este trabajo es conseguir un método digital de recolección y gestión de datos eficaz, flexible y reproducible para cualquier estudio de superficies de suelos arqueológicos. En la prueba y desarrollo de este método hemos documentado más de 4,600 restos arqueológicos dentro de una estructura de vivienda del Paleolítico superior en La Galería Inferior de La Garma y una estructura natural en el mismo lugar. En un proceso de trabajo que ha llevado 18 meses continuos, cuyos resultados demuestran el valor de esta herramienta de trabajo, caracterizada por una enorme adaptabilidad y flexibilidad a cada contexto particular o a las necesidades del investigador. No proponemos un método de sistematización para la documentación de restos arqueológicos, nuestro objetivo final es establecer un método sencillo y robusto de recolección de los datos, abriendo el conocimiento de estas técnicas a cualquier usuario. Íntegramente en código abierto para construir un modelo que se nutra del uso y avance de la propia comunidad investigadora.

Keywords: GIS; methodology; photogrammetry; programming; spatial databases; Upper Paleolithic

Palabras clave: SIG; metodología; fotogrametría; programación; bases de datos espaciales; Paleolítico superior

Since the 1980s, the computerization of archaeological projects has been pivotal in the evolution of the discipline. Initially, this process focused on the use of databases and statistical analysis. However, advancements in computing technology have led to the development of specialized workflows, encompassing everything from data collection to analysis (Dibble and McPherron 1988, 1997; McPherron and Dibble 1987, 2002). The adoption of free and open-source software (FOSS) has played a key role in this transformation, with initiatives such as Open-archaeo (Batist and Roe 2023) facilitating the reuse and adaptation of tools and code within archaeological practice. Nevertheless, the lack of standardization in data management remains a significant methodological challenge, particularly regarding the preservation and accessibility of archaeological data (Colleter et al. 2020; Gencheva 2023; Reed et al. 2015).

Technological advancements have significantly improved archaeological techniques, tools, and outcomes, steering the discipline toward noninvasive methods that safeguard heritage for future generations while ensuring research reproducibility (Arias and Ontañón 2014; Boyd et al. 2021; Ellis 2016; Galeazzi 2016; Rayne et al. 2017; Vandenbulcke et al. 2016; Wallrodt 2016). However, data management practices continue to vary considerably between research teams. Despite ongoing efforts to create and standardize these processes through FOSS-based systems (Reed et al. 2015), the absence of a widely adopted universal solution continues to pose challenges (Ducke 2012; Holdaway et al. 2019).

One of the main issues is that each research team tends to establish its own system for data collection, storage, description, and analysis. While this allows for adaptation to the specific needs of each team, it complicates the integration and long-term preservation of data for future use (Kansa and Kansa 2021; Mandolesi 2009; Trimmis 2018). Additionally, data digitization is often performed inconsistently or without consideration for long-term preservation, increasing the risk of losing valuable information (Reed et al. 2015).

Free and open-source software introduces a new form of collaboration that, as Ducke (2012) points out, promotes transparency, modification, and tool reuse. In today's era of data science, it is crucial for archaeology to adopt large-scale data management practices that prioritize transparency and collaboration (Kansa and Kansa 2021; Katsianis et al. 2008). The goal is to maximize data reuse and facilitate collaboration among researchers by adopting open-source practices that emphasize transparency and adaptability throughout the workflow (Dibble and McPherron 1988). Projects like Open-archaeo, which promote the use of free tools and software designed specifically for archaeologists, represent a key step toward democratizing data management (Batist and Roe 2023, 2024).

In this context, we propose a practical and accessible workflow that allows any researcher, regardless of technical expertise, to manage the digital data of their project. Our approach advocates for noninvasive archaeology and the use of free and open-source software, with the aim of establishing a standardized method for data collection and storage. This method seeks to minimize human error while prioritizing the conservation of archaeological remains. The proposed standardization does not aim to unify all procedures but rather to create a common foundation that facilitates interoperability and collaboration, as evidenced by recent FOSS initiatives (Batist and Roe 2024; Dibble and McPherron 1988).

We illustrate the effectiveness of this approach through its application in a specific case study: the Lower Gallery of La Garma (Cantabria, Spain). This site is ideal for the development of new digital techniques for the study, conservation, and dissemination of archaeological heritage. The case study provides an opportunity to create and test a standardized and accessible system. By applying our model at La Garma, we aim to evaluate its adaptability and applicability in a specific archaeological context, marking an important step toward fostering greater collaboration and advancement in this field.

The Lower Gallery of La Garma: Conservation and Noninvasive Methods

La Garma Hill, described as an elevation reaching 186 m asl, is located near the Bay of Santander, in the town of Omoño (Ribamontán al Monte, Cantabria). The karst complex in the hill is characterized by 13 sites that reveal a succession of occupations from the Lower Paleolithic to the Middle Ages (Arias and Ontañón 2012; Arias et al. 2000, 2011, 2018; Cueto et al. 2020). The record includes a set of rock art
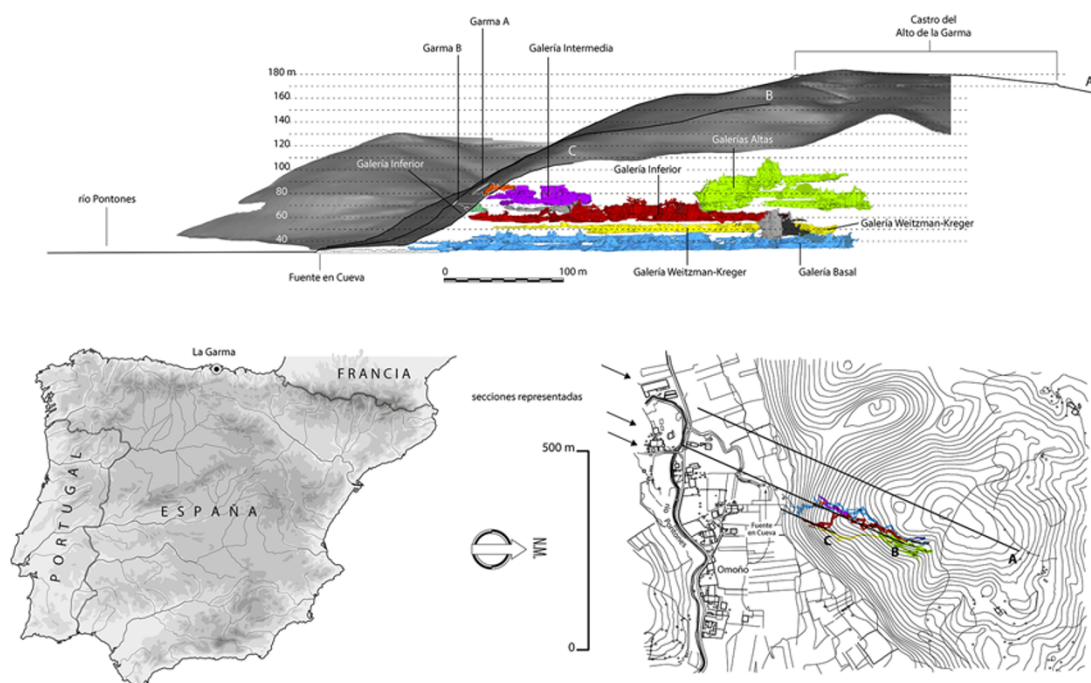
Figure 1. Location and mapping of the La Garma karst complex (Spain).

with more than 500 graphic expressions, floors full of archaeological artifacts, and remains of anthropic structures, the result of the activities carried out at the site. Given its scientific and cultural relevance, La Garma was included in the UNESCO World Heritage List in 2008 (Arias and Ontañón 2014; Ontañón et al. 2008).

The most outstanding part and central object of this methodological case study is a section of the karst system currently known as the Lower Gallery. This site, at an altitude of 55–58 m, is a cave passage 300 m long. The particularity of this site is that its original entrance was closed at the end of the Pleistocene, in about 14,500 cal BC in the Middle Magdalenian (Arias 2009; Arias et al. 2005; Ontañón 2003), so that all the prehistoric remains were left on the surface in an excellent state of preservation, thanks to the cessation of sedimentation and postdepositional processes (Arias and Ontañón 2014; Arias et al. 2000; Cueto et al. 2020; Ontañón 2003). At present, access to this part of the system is via two 7 and 15 m shafts, located in other sections of the karst complex called La Garma A and the Intermediate Gallery (Figure 1).

La Garma is an exceptional example of a unique conservation context. Since its discovery in 1996, the team has worked to protect the archaeological area, facing a double challenge: documenting an extensive set of archaeological remains and ensuring their conservation (Arias and Ontañón 2014; Arias et al. 2000, 2011; Cueto et al. 2020). This project integrates research, conservation, and dissemination, working with noninvasive techniques to minimize the impact on archaeological floors, remains, and structures (Arias and Ontañón 2014). The goal is to avoid remanipulating the archaeological remains, site, or its context once it has been properly documented.

A fundamental conservation protocol begins at the entrance to the site, where sterilized footwear is used to prevent the entry of organic matter or microorganisms. In addition, a narrow path, walkways, and other facilities have been marked to protect the prehistoric floors, and these measures are complemented by sensors for climate and environmental monitoring (Arias and Ontañón 2014; Arias et al. 2000; Martin-Pozas et al. 2024; Sanchez-Moral et al. 2021).
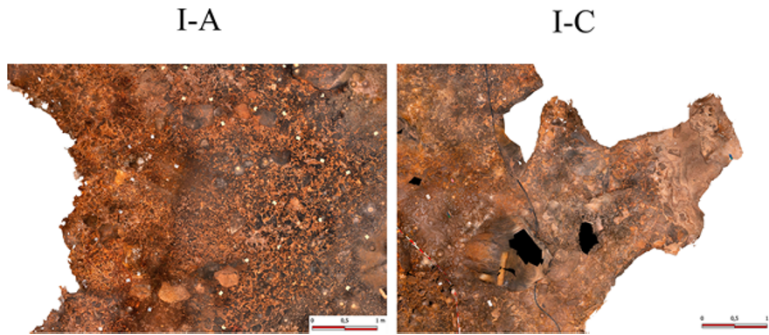
Figure 2. Orthomosaics of structures I-A (*left*) and I-C (*right*) for analysis.

Our documentation work prioritizes in situ study to minimize the manipulation or extraction of remains and protect their archaeological context, although this implies a slower process (Arias and Ontañón 2014). Therefore, the need to develop an efficient data collection, storage, and analysis system has led us to create an integrative model, which takes advantage of digital technologies such as photogrammetry, laser scanning, total station, and GPS, among others. The current challenge is to integrate these data into an accessible management system that facilitates their use by any researcher in the field or in the laboratory (Galeazzi 2016; Kansa and Kansa 2021), while maintaining our goal of minimizing intervention at the site once an archaeological element has been documented.

Digitization and Conservation in the Context of the Study

In the Lower Gallery of La Garma, our case study focuses on Structures I-A and I-C (Figure 2). Both are located in Zone 1, about 40 m from the original entrance to the cave. I-A, with a surface area of 5 m^2, is a Paleolithic habitation structure, with a subcircular floor plan, formed by speleothems and limestone blocks, located under an overhanging roof 1.6–1.7 m high. Its floor is covered by very dense concentrations of archaeological remains dating from the Magdalenian period. In turn, Area I-C is a natural chamber of about 11 m^2, which houses vestiges of activities, with a notable predominance of graphic expressions, such as paintings and engravings (Arias et al. 2011; Ontañón 2003).

The documentation of these spaces has been carried out cooperatively by three members of the team, using a shared server that allows multiple users to work simultaneously on the same project. In this sense, La Garma presents enormous technical challenges, as it is a cave. However, the problem of setting up a network connection that allows access to the server was solved with the installation of the GARMANET communications system in 2007–2008, facilitating the creation of a LAN network that allows connections both inside and outside the cave (Arias and Ontañón 2014).

Protocol

Our conservation procedures at this site include another change of footwear for accessing and moving around the site, as well as the use of nitrile gloves for handling the remains. Documentation begins with a photographic record of the remains to be studied, followed by a detailed study and photogrammetry when necessary, ensuring the relocation and conservation of the objects in their original location. Access to the materials is by means of a retractable aluminum walkway (with a maximum extension of 2 m and a width of 35 cm) with padded bases to protect the ground and the archaeological remains. This allows us to avoid direct contact with the floor and fully preserve the site (Figure 3). In addition, a strict protocol for time spent in the area, limited to four continuous hours of work three days a week, was in line with the environmental control criteria at La Garma (Arias and Ontañón 2014; Arias et al. 2000, 2011).

Microsoft Surface Pro-8 and Pro-9 tablets were used for the fieldwork; they provide perfect adaptability for fieldwork and robustness to draw vector-layer geometries and perform complex analyses with large files or datasets. The process of working with those devices consisted of identifying the element to be documented, drawing its geometric representation in the vector layer, and completing all the

Figure 3. Researcher working with the La Garma work protocols during our digital data collection system in the cave. Photograph courtesy of Carlos García-Noriega.
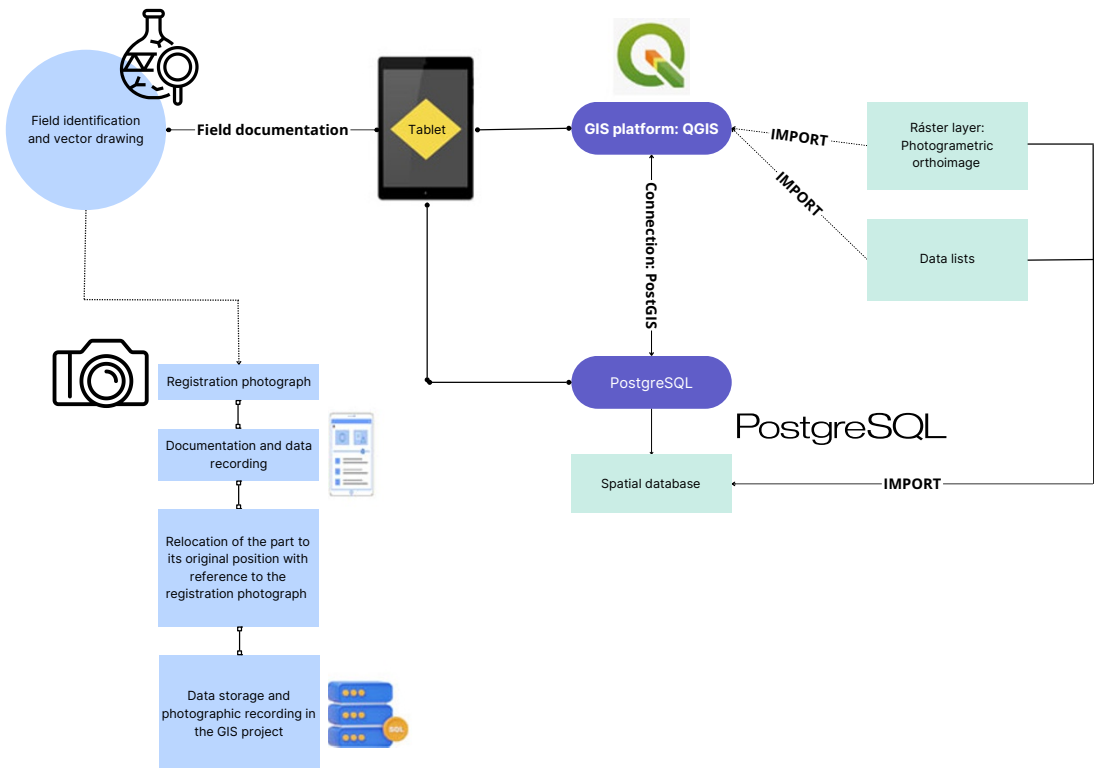

Figure 4. Field documentation process utilizing a system based on QGIS and PostgreSQL.

values or attributes for accurate documentation, before returning the archaeological object to its original location (Figure 4). At the same time, the record photographs were associated with the geographic information system (GIS) project through programmed actions, allowing access to them for each recorded element.

Digital Conservation

Finally, an enormous effort has been devoted to the creation of more than 500 photogrammetric models of the study area and its archaeological remains. These models allow subsequent analysis, for example in the reassembly of lithic pieces (both in digital format and in physical format through 3D printing). This procedure is crucial for the preservation of fragile remains and also facilitates future research without the need to physically access the site. The high-precision photogrammetric models include the Palaeolithic structure I-A, the adjacent floors, the cave walls near structure I-A, graphic expressions, remains of osseous industry, ornaments, and lithics. All of these are stored as supporting material for the data collected in our database.

Workflow

The main challenge at La Garma is the preservation of the archaeological context, which limits freedom of movement and complicates work and data collection. Our solution, which adopts noninvasive techniques, begins with the construction of a raster archive documenting the entire site. Since 2012, in collaboration with the firm GIM Geomatics, we have produced a high-resolution giga-orthoimage. This was achieved through a combination of photogrammetry and terrestrial laser scanning, using a FARO Photon 120 3D laser scanner and two cameras, a Hasselblad CF-39 and Sony NEX-7. The whole process was performed with a system of poles and cranes that avoided direct contact with the archaeological floors (Arias and Ontañón 2014).

The information was processed with our own software, developed by GIM Geomatics, which grants full freedom over the rights of our products. The result is a product of the highest quality for the documentation and study of the archaeological remains on the surface, reaching a resolution of 200 microns (Figure 5). This has allowed us to accurately characterize a vast extension of Paleolithic surfaces, resulting in a reliable document for the conservation and dissemination of the archaeological heritage in the cave (Arias and Ontañón 2014). The objective of this photogrammetry is to provide a reference raster layer for a GIS platform to identify the archaeological remains recorded in the field. On this layer, we draw the vector geometries and record each analyzed remain, obtaining a georeferenced product that allows precise field location. Additionally, this procedure accurately locates each piece and allows us to extract XYZ coordinates when integrated with other products, such as a digital elevation model (DEM).

Spatial Database Configuration

PostgreSQL version 12.2 is used for data management. This allows us to operate both with graphical user interface (GUI) applications, such as pgAdmin, which greatly facilitate the process, and with a programming language to predefine and streamline the replication process (Figure 6).

The creation of a PostgreSQL database through pgAdmin begins with setting up a "Server Group," to which we assign a name. This serves as a logical collection of servers, facilitating the organization and management of the servers and their databases. To do this, we select the option to create a Server Group in the "Object" tab.

Next, we register a "Server," defined as an instance of the database that provides access to our configuration, by assigning it a name in the general attributes window. Additionally, we establish the connection location. Since we are working locally, we enter "localhost" or, alternatively, the local IP in the "Host" field. Finally, in this same window, we enter a password for our server, with the option to save it, depending on our preference for controlled access to the server.

The next step is to create the database and its internal structure using "Schemas" and "Tables." We select "create database" on our server and start by entering the name of the database. We leave the default owner set to "postgres" and ensure that in the "Definition" section, the "Encoding" is set to UTF-8.

Once the database has been created, we begin its configuration. The first recommended step is to create an extension that will allow us to connect our database to a GIS platform. To do this, in the "Extension" section we enter "postgis" as the name. This enables us to connect our spatial database to a new GIS project.
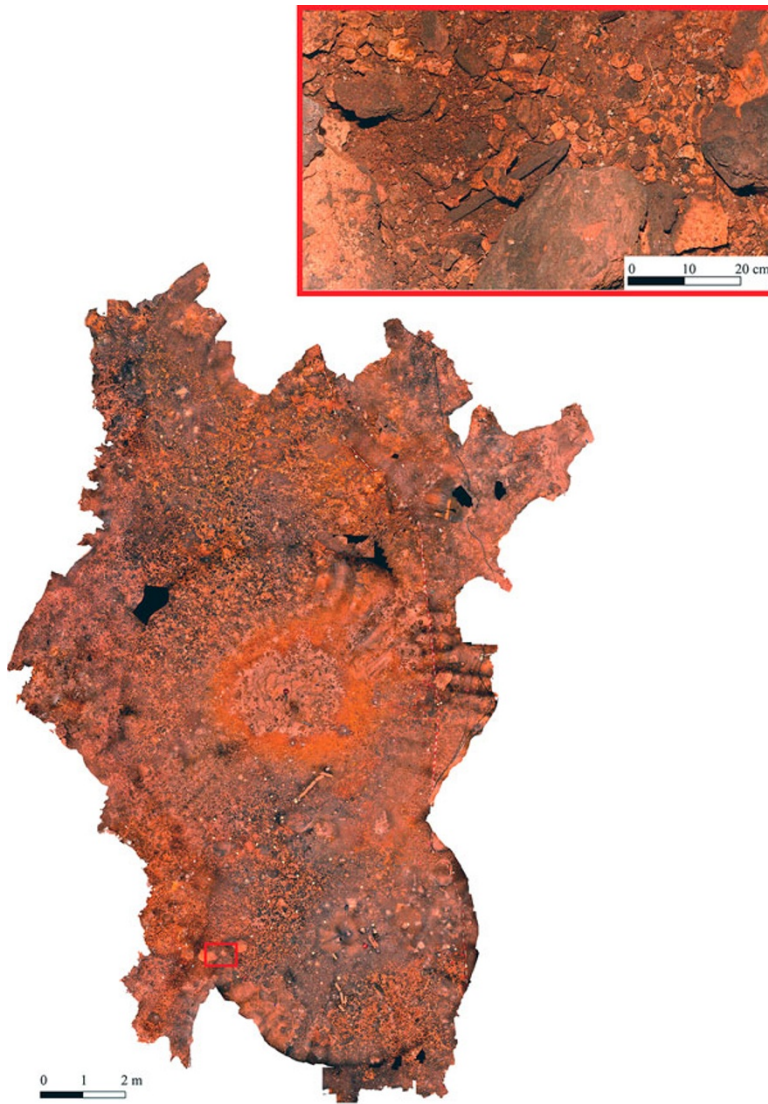
Figure 5. Giga-orthoimage with 200-micron resolution of Zone 1 in the Lower Gallery of La Garma.

We then proceed with the internal configuration by initially creating a schema or multiple schemas depending on our needs. The "Schema" functions as a logical collection of objects, grouping the internal components of a database such as tables, views, functions, et cetera, facilitating data access and management. To create it, we set a name and leave the rest of the configuration as default. With the "Schema" created, we proceed to create the "Tables."

These "Tables" can be defined as data structures that organize our information into rows and columns, with each row being an entry or record and each column representing a characteristic or attribute of that record. We create the tables by going to the corresponding section within our "Schema." In the "General" window we enter the name of the table, and in the "Columns" tab we create, one by one, the records we want to characterize for each archaeological remain (i.e., for the documentation of the data). The procedure consists of "add row" to generate a new attribute, where we enter the name and type of data to be recorded, with the ability to select from a multitude of formats according to our needs, such as various numerical values, alphanumeric, Boolean, and JSON, among others. It is also recommended
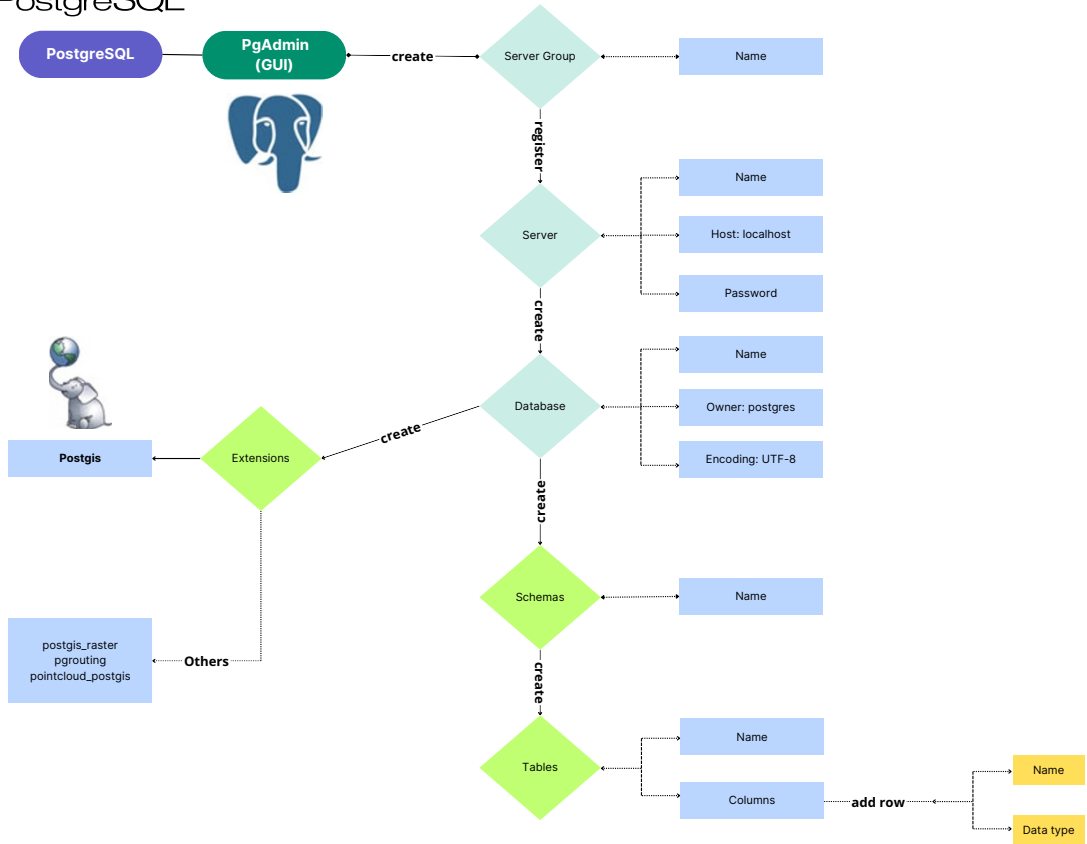
Figure 6. Workflow process for creating and configuring a PostgreSQL database.

to select "Primary key" if the values of a specific row must be unique in each case. For example, when creating an inventory number, this practice is highly recommended, as a number can never be repeated. Additionally, a spatial index for each table is also recommended as a key factor to speed up calculations or analysis with large datasets (Video 1).

In our project, we have created a table for each type of archaeological element, with geometries varying between points, lines, and polygons: Fauna, Lithic, Malacofauna, Osseous Industry, Ornaments, Art, Human Remains, Metallic Objects, Stratigraphic Units, Displaced Elements, Structural Elements, Natural Elements, Samples, and Dating. The choice of each geometry is adapted to our needs. For example, while archaeological remains represented as polygons (e.g., Fauna or Lithics) enable greater interaction in the analysis, other categories such as dating samples (point-based geometry) may require less interaction. Each table is formed by a series of attributes based on the specific survey methods for each type of remain. In other words, each of them acts as a specialized database that ensures the correct cataloging and documentation of the remains and the archaeological context.

This whole system defines the general basis for configuring and making a spatial database functional for connection to a GIS project. We have pre-established the entire process in structured query language (SQL) code to provide a base model to work with. This strategy aims to simplify the reproduction of the procedure and the creation of new models, given the flexibility of the design and its adaptation to each context and preference. The integration is performed in a GIS platform, which allows us to work with and capture vector and geospatial data. The software used is QGIS version 3.22 Białowieża, a robust and growing open-source project with a large community of developers (Video 2). However, it is worth mentioning that this same system can be applied to any version of QGIS 3.x onward.

Video 1. Creation of PostgreSQL database in pgAdmin.

Video 2. Connectiong PostgreSQL database to GIS system (QGIS).

Improvements: Lists and Records and Transects

In this stage of the process, we implemented several improvements to reduce human error and standardize a uniform method for data recording. Initially, typographical errors can occur during data entry. For example, if we have an attribute to record the physical state of an archaeological remain, which can be either "whole" or "broken," typographical errors may compromise the homogeneity of the data. This could lead to subsequent errors during analysis and slow down the documentation process by requiring manual correction of the same attribute in each record.

To facilitate this process, lists of predefined values are created in the GIS platform. These lists are written in a simple code that we will save in CSV format, representing a table composed of two columns with
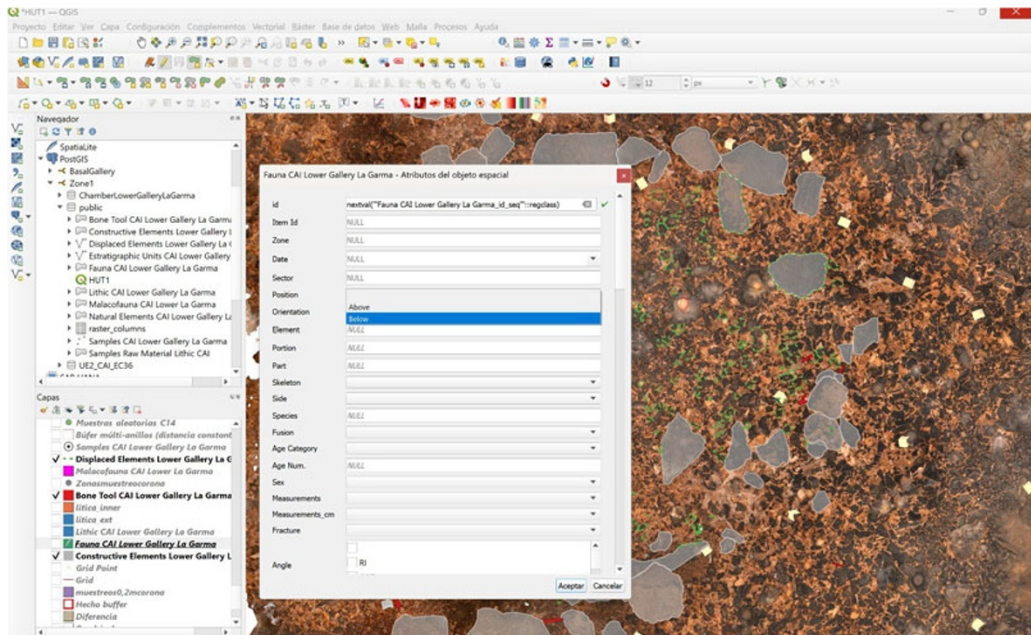
Figure 7. Functionality of the lists created for documentation.

an index value and its associated key value. Finally, these lists must be configured in both PostgreSQL and the GIS platform to ensure their correct functioning (Figure 7).

Creating these lists is a simple process. First, in a notepad, write the columns and rows, separating them with tabs, commas, or any other delimiters you prefer. You must include an index attribute and a value attribute. In the first column, enter the key value, which in this case is numeric, and in the second column, the corresponding attribute. Following the previous example, the second column will contain the attributes "whole" and "broken" in each row. Once these lists are created, they are integrated into the GIS project, which optimizes the cleanliness and efficiency of data collection.

First, we import each list in CSV format into QGIS. Then, we open the "Layer Properties" option and in the "Attribute Form" section, we select the control-type option "Value Relation." This allows us to select the previously imported CSV layers and set the corresponding field for the index and value attributes in the associated fields (Figure 8). This process must be repeated for each list we have (Video 3).

Another improvement in our procedures is a novel documentation system that involves creating points and transects to position the documented elements. During the process, these points will be connected to transects, allowing us to clearly track the geolocation of the documentation and study process of the archaeological remains. This new part of the procedure is included as a fundamental tool for fieldwork and may inspire future improvements based on our proposed procedure (Figure 9).

The process we call Records and Transects is as follows: in PostgreSQL, we create a table called "Record," which establishes the points that will record each documentation made with the tables created for this purpose, such as lithics, fauna, et cetera layers. In addition, we create a second layer called "Transects," with line features that will represent the connections of the remains as we process the data. For this to work, we must create a "Trigger Function," which contains the execution code for a "Trigger," defined for each data table. It is important to understand that a "Trigger" is a mechanism that allows us to execute an action automatically in response to events occurring in our tables, serving as a useful tool to automate processes and maintain data integrity.
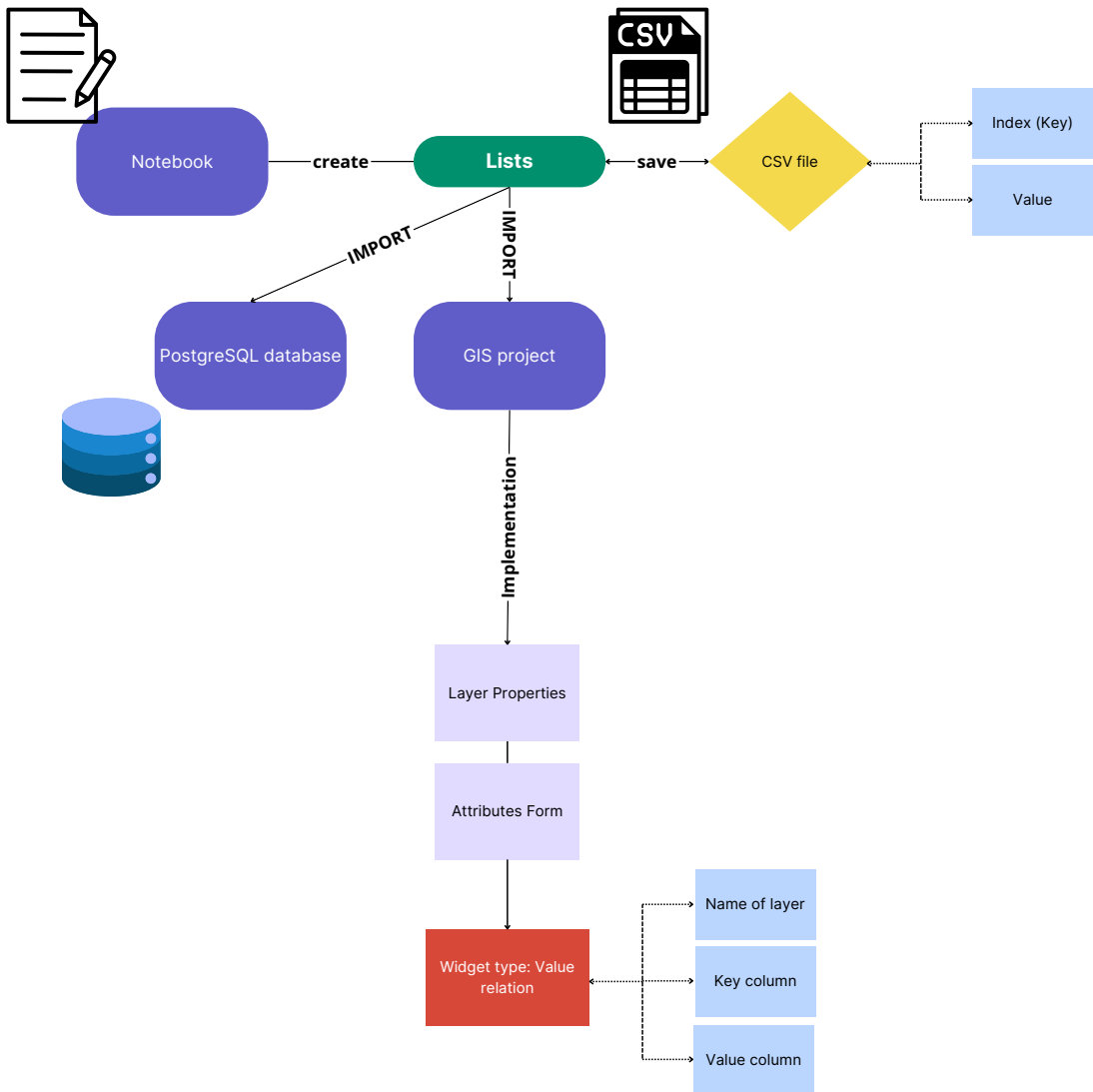
Figure 8. Workflow for creating lists.

The first function, called "insert_record," will automatically insert each record in point format once we have drawn in one of the vector layers of the archaeological remains, also saving the type of remains documented in its attributes. Then, for the transects, we create a second function called "generate_transect," which will systematically join the points created with lines, facilitating a better understanding of the locations where we are processing the data (Figure 10).

Once these triggers and functions are ready in our database, we can work on the GIS project with the aforementioned documentation system. As with the other tables, we only need to import the record and transect layers into our project and update the layer as we record data. In this way, we achieve an efficient system for controlling the geospatial documentation of our data (Video 4).

Multiple Users

All data are managed on a local server, ensuring the security of our data and simultaneous communication or work between multiple users, from a network connection, in our case GARMANET (Arias and Ontañón 2014). This network allows multiple users to work within the same local connection, but, for

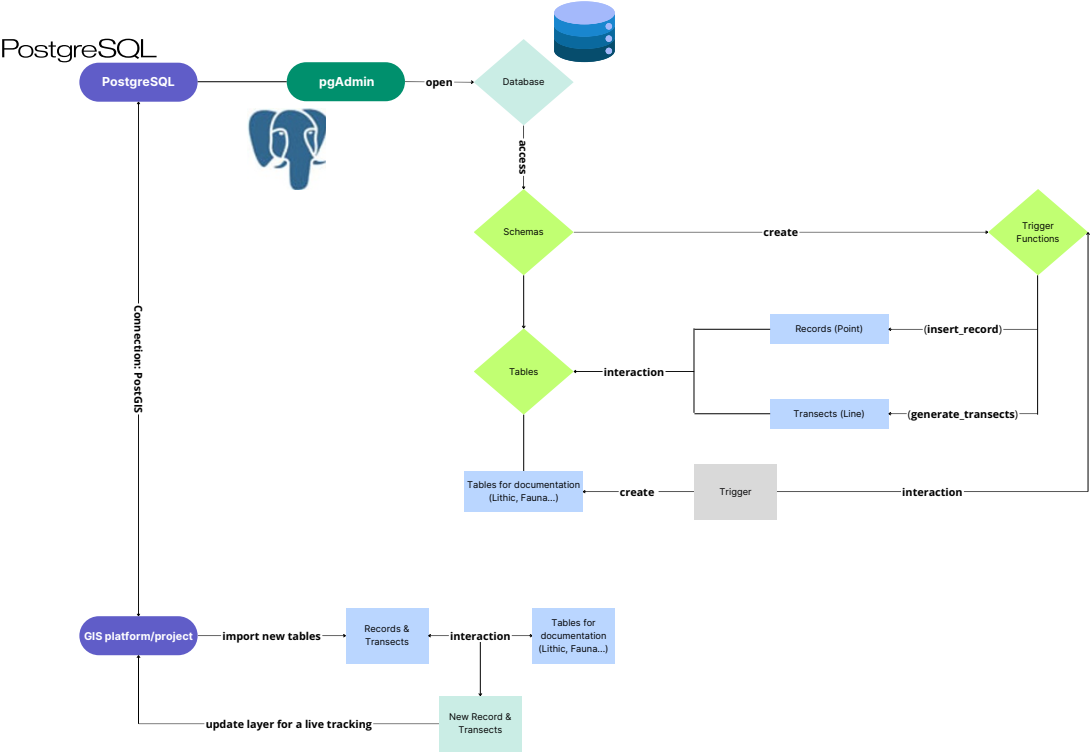**Video 3.** Creation and configuration dropdown lists for QGIS.

Figure 9. Workflow for creating triggers and trigger functions and their interaction.

this to function correctly, one device must be established as the host. The other users can then connect to the server and access the corresponding database. This procedure can be easily done in the pgAdmin application. To do this, the device that wants to connect must access its connection configuration and, in "Host," change the attribute "localhost" to the IP address of the device acting as the administrator. With
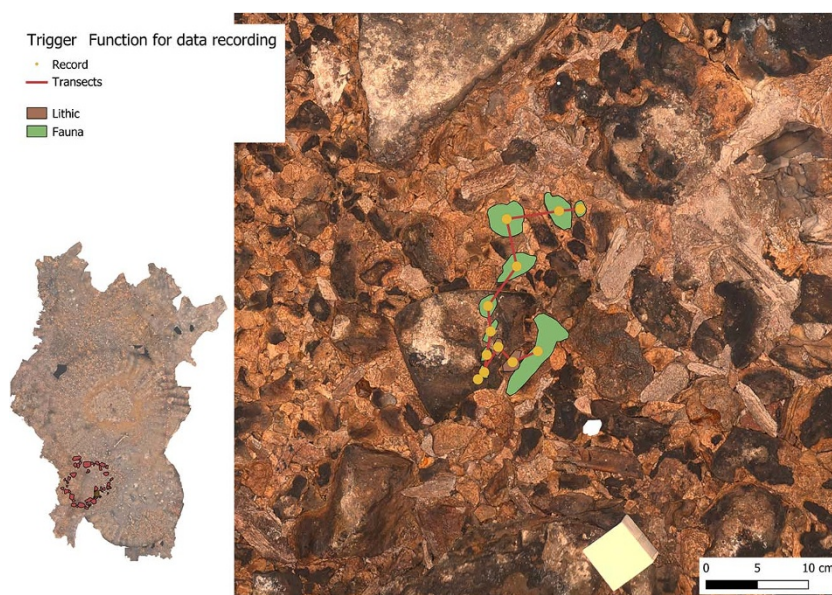
Figure 10. Process of creating points and transects for positioning documented elements in the archaeological site.

Video 4. Records and transects.

this simple change, users can connect to the database, start their QGIS projects, and work simultaneously with other users (Supplementary Text 1).

Python

Finally, our digital method has been developed using Python (Python 3 version), an Object Oriented Programming language. We chose this language because of its great potential for data analysis, specifically for geospatial and geostatistical analysis. In addition, it stands out as a language on the rise, with a
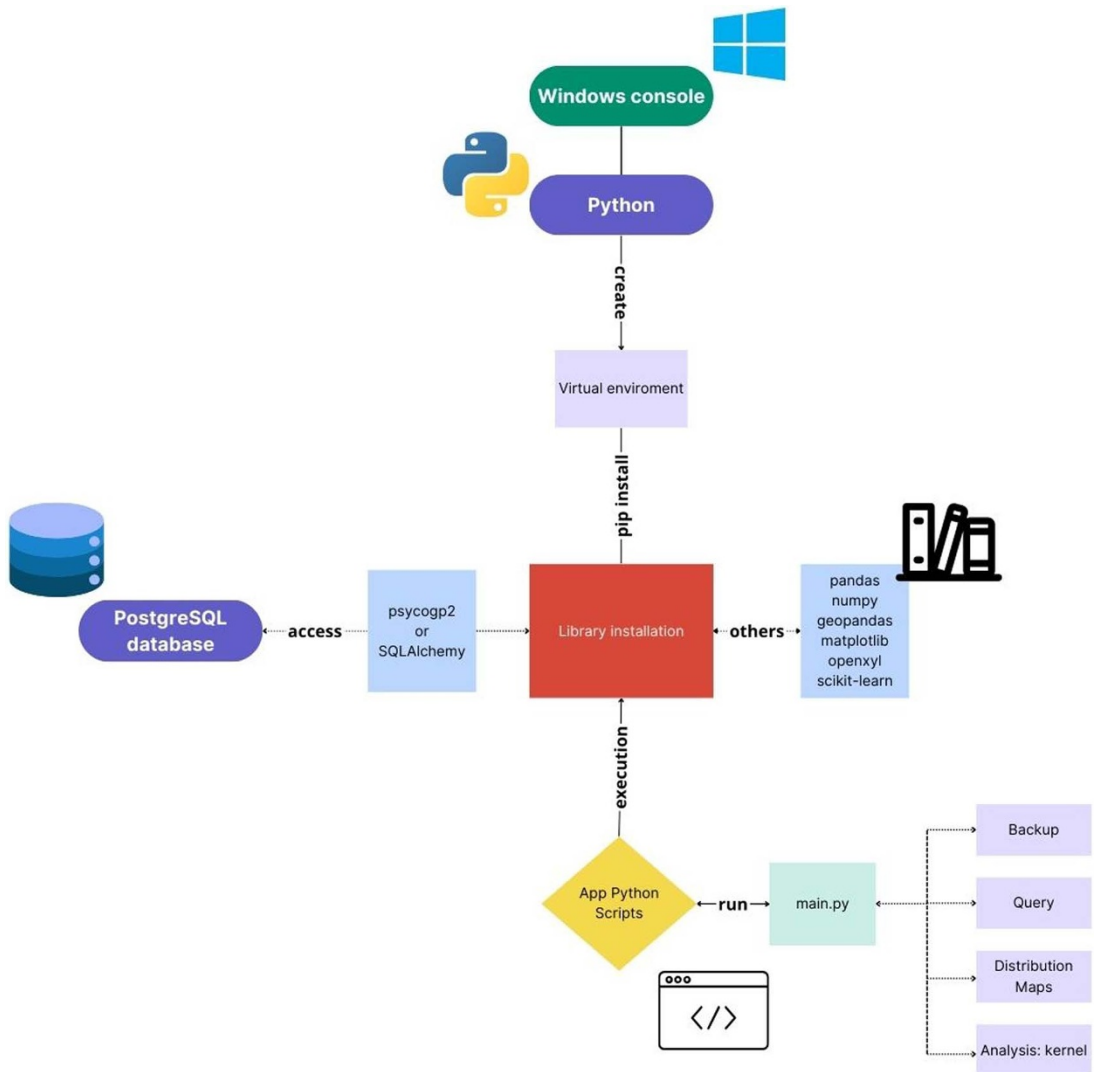
Figure 11. Working structure using Python with our data stored in a PostgreSQL database.

large community and a relatively easy learning curve. Its main advantages are that it allows code reuse and extension, as well as the creation of complex systems that facilitate teamwork (Figure 11).

We provide a base model of our work with Python, organized in an application for storing, saving, and basic analysis of our data. This set of scripts is openly available in a GitHub repository, allowing execution, modification, and adaptation to other contexts or needs.

To run the application, we use a command console. Since we work with a Windows operating system, we use the system's own console. For correct operation, we must first create a virtual environment. This allows us to establish a unique and secure project that is not affected by previous installations of packages or libraries from other projects. This practice is essential for the correct functioning of our scripts and is a common and practically obligatory practice in the programming world.

Once the virtual environment for our project has been created, we proceed to install the necessary libraries using the "pip install" command. The first library to install is psycopg2 or SQLAlchemy, which will allow us to connect to and access the data in our PostgreSQL database. Next, we install other libraries
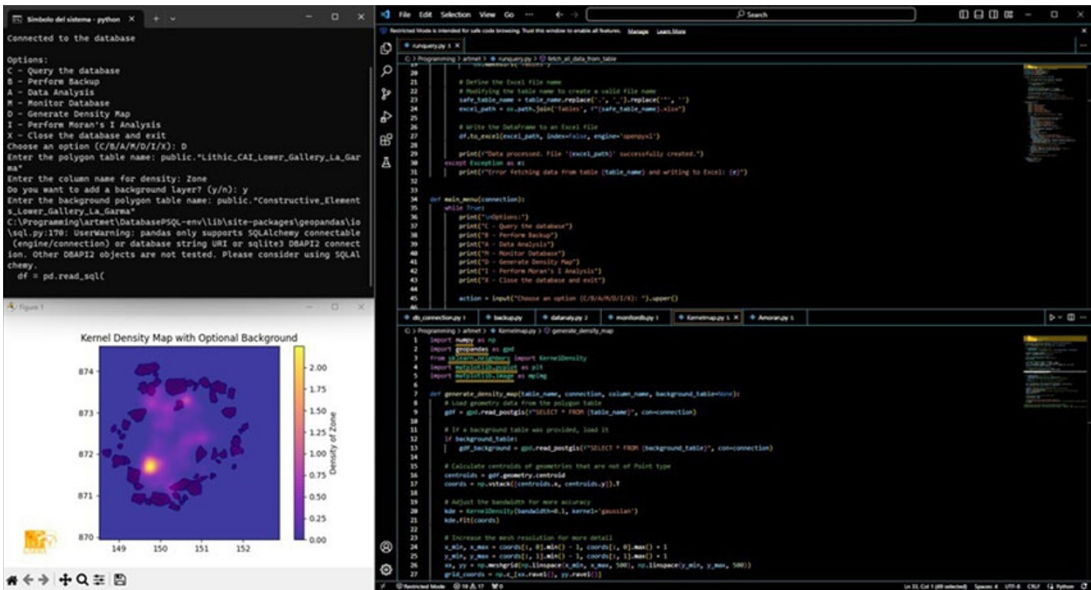
Figure 12. Connection to PostgreSQL database with Python and creation of a tool for data management.

necessary for data management and analysis. In our case, these libraries are: pandas, numpy, geopandas, matplotlib, openpyxl, and scikit-learn.

Now we can proceed to execute our application. To do this, we simply run the main script, which organizes and calls the rest of the functions. The predefined tasks include querying the database, creating and saving a backup, and monitoring the state of the database, as well as frequent and basic geospatial analysis, such as estimating Kernel densities and creating maps of distributions (Figure 12).

This model serves as a predefined configuration and starting point, but can be modified and adapted to each specific case. All lines of code are open-source, allowing flexibility, permanence, and improvement over time of the procedure discussed in this research work (video 5).

Toward Standardized Data Collection

Technological innovations have transformed archaeology, allowing for the storage and analysis of large datasets and prompting a rethink of both field and laboratory practices (Boyd et al. 2021; Calzado-Martínez et al. 2023; Stamnas et al. 2021). These advances have improved documentation and conservation management by offering a new approach to efficient data storage and collection, while also adapting tools from other disciplines to archaeological work (Batist and Roe 2024; Galeazzi 2016; Neubauer 2004; Rayne et al. 2017). Despite the progress, significant challenges remain regarding standardization and process integration within digital archaeology (Colleter et al. 2020; Reed et al. 2015).

One of the most significant efforts in this regard is the Open-archaeo project (Batist and Roe 2023, 2024), which has driven the creation and reuse of open-source tools and code specifically designed for archaeology. This approach fosters collaboration between developers and users, allowing these tools to be adapted to varied archaeological contexts. In this sense, FOSS solutions in archaeology offer an alternative development model that emphasizes the reuse and adaptation of tools to meet specific needs (Reed et al. 2015). Although progress has been made in creating applications and plug-ins that address particular needs (Brovelli and Magni 2003; Cosmas et al. 2003; Couillet et al. 2022; Eichert 2014; Gencheva 2023; Martínez-Carrillo et al. 2009; Nigst et al. 2010; Sauer 2023), the inherent lack of flexibility in many of these solutions can limit their long-term applicability.

A notable example of this issue is OpenDig (Vincent et al. 2013, 2014), a tool designed for the recording, editing, and publication of archaeological data. Despite its functionalities, its popularity has been

Video 5. Connecting PostgreSQL database with Python and execution scripts.

limited, probably due to a lack of updates and the emergence of more adaptive and modular alternatives. Similarly, pyarchinit (Cocca and Mandolesi 2016; Mandolesi 2009; Montagnetti and Mandolesi 2020), a plug-in that represents a methodological approach closely aligned with our proposal, addresses the problem of coordination in data management methods, promoting integration, stability, and ease of updating within a predefined model. While we recognize the value of this solution, unlike its authors we believe that the diversity of procedures should not be seen as a barrier but rather as an opportunity to enrich methodologies applied in different archaeological contexts. Each case study has unique needs, as do the researchers. Rather than imposing a unification of procedures, we propose developing a common base that allows for interoperability and data exchange (Batist and Roe 2023, 2024; McPherron et al. 2008). Our methodology therefore is based on the principles of flexibility and adaptability, as demonstrated by Batist and Roe (2024), enabling users to modify and expand tools according to emerging needs.

With this objective, we have developed a methodology that prioritizes collaborative work, allowing multiple researchers to participate simultaneously in the investigation. At the same time, human errors are minimized through the automation of data recording via lists, ensuring the homogenization and consistency of information for future use. The implementation of real-time documentation systems based on data geopositioning (Records&Transect) not only simplifies a complex problem but also allows for continuous improvements, such as the use of GPS positioning systems, which increases the accuracy and adaptability of the system over time (Dibble and McPherron 1988; Ducke 2012).

Instead of developing libraries or plug-ins that tend to become obsolete, we propose a protocol and an open-source application that facilitate the management and analysis of archaeological data. This approach ensures the reproducibility, extension, and permanence of the procedure over time, supported by an active community of users and developers (Batist and Roe 2024). The ultimate goal is to provide methodological tools that allow users to create their own systems, tailored to their specific needs, promoting the permeability of knowledge and the adoption of open digital techniques (Batist and Roe 2024; Holdaway et al. 2019). This simple yet flexible methodology guarantees the longevity and efficiency of the procedure (Colleter et al. 2020; Dibble and McPherron 1988), consolidating itself as a robust and adaptable alternative to proprietary and less flexible solutions. It offers an economically and technically accessible solution to any researcher, regardless of their level of expertise, allowing for the complete integration of existing data and its adaptation to new research contexts.

Conclusions

The process we have described in this work shows remarkable efficiency in the task of documentation and conservation of archaeological remains, obtaining accurate and fast results in the capture, storage, and analysis of data. Although we have used both SQL and Python in our project to streamline the processes of database creation and configuration, connection, access, and data analysis, these technologies are today accessible owing to the development of intuitive interfaces. This allows us to fully reproduce the project described here without the need for extensive computer or programming knowledge.

The increasing utility and cost-effectiveness of generating high-precision, high-quality raster products from photogrammetry or laser scanning makes this approach viable for any researcher or team. In addition, the availability of creating a database and connecting it to a GIS platform with open-source software significantly reduces costs, as there is no need for often expensive licences.

The fact that this approach is based on open-source software with strong and active communities, such as QGIS, PostgreSQL, or Python, creates sufficient confidence that our digital data will endure over time. It also allows for continuous improvement and adaptation of the method to different cases and contexts. This case study represents a first step toward an open-source methodological community in the archaeological discipline, focused on data collection and storage.

We conclude with an analogy that summarizes our work: the previous approaches use tools to generate a finished product that the user can consume, while we provide the tools necessary for users to create their own consumer products. This subtle difference underlines the importance of true permeability of knowledge and the application of digital techniques in archaeology. The essence of open source lies precisely in this democratization of knowledge, as well as in the ability of users to actively participate in the creation of content that contributes to collective growth.

Author Contributions.   Carlos García-Noriega (Conceptualization: Lead; Formal analysis: Lead; Investigation: Lead; Methodology: Lead; Software: Lead; Visualization: Lead; Writing – original draft: Lead; Writing – review & editing: Lead) Rodrigo Portero (Formal analysis: Equal; Investigation: Equal; Writing – review & editing: Supporting) Alba Ruiz-Cabanzón (Formal analysis: Equal; Investigation: Equal; Writing – review & editing: Supporting) Roberto Ontañón (Writing – review & editing: Supporting) Pablo Arias (Writing – review & editing: Supporting)

Data Availability Statement.   The code used for the development and implementation of the digital methods described in this article is currently hosted in a private GitHub repository: https://github.com/Garmaproject/Garma.git. The repository will be made publicly available upon publication of this article. Until then, access can be granted upon reasonable request to the corresponding author.

Competing Interests.   All the authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or nonfinancial interest in the subject matter or materials discussed in this manuscript.

Supplementary Material.   The supplementary material for this article can be found at https://doi.org/10.1017/aap.2024.44.
Supplementary Text 1. Details on setting up connection parameters, securing remote connections, and best practices for managing multiple users accessing the same PostgreSQL database.

References Cited

Arias, Pablo. 2009. Rites in the Dark? An Evaluation of the Current Evidence for Ritual Areas at Magdalenian Cave Sites. *World Archaeology* 41(2):262–294. https://doi.org/10.1080/00438240902843964.

Arias, Pablo, César González-Sainz, Alfonso Moure-Romanillo, and Roberto Ontañón. 2000. La Zona Arqueológica de La Garma (Cantabria): Investigación, conservación y uso social. *Trabajos de Prehistoria* 57(2):41–56. https://doi.org/10.3989/tp.2000.v57.i2.246.

Arias, Pablo, Enrique Gutierrez Cuenca, José A. Hierro Gárate, Francistco Etxeberria, Lourdes Herrasti, and Paloma Uzquiano. 2018. Hidden in the Depths, Far from People: Funerary Activities in the Lower Gallery of La Garma and the Use of Natural Caves as Burial Places in Early Medieval Cantabria, Northern Spain. In *Caves and Ritual in Medieval Europe, AD 500–150*, edited by Knut Andreas Bergsvik and Marion Dowd, pp. 133–151. Oxbow Books, Oxford.

Arias, Pablo, and Roberto Ontañón. 2012. La Garma (Spain): Long-Term Human Activity in a Karst System. In *Caves in Context: The Cultural Significance of Caves and Rockshelters in Europe*, edited by Knut A. Bergsvik and Robin Skeates, pp. 101–117. Oxbow Books, Oxford.

Arias, Pablo, and Roberto Ontañón. 2014. La Garma: Un proyecto orientado al estudio del arte paleolítico, su contexto y su conservación. In *Cien años de arte rupestre paleolítico: Centenario del descubrimiento de la Cueva de la Peña Cándamo (1914–2014)*, edited by Mario Menéndez and Maria S. Corchón, pp. 174–194. Acta Dalmanticensia. Estudios históricos y geográficos 160. Ediciones Universidad de Salamanca, Spain.

Arias, Pablo, Roberto Ontañón, Esteban Álvarez-Fernández, María Teresa Aparicio-Alonso, Adriana M. Chauvin-Grandela, Ignacio Conte, Marián Cueto, et al. 2005. La estructura magdaleniense de La Garma A: Aproximación a la organización espacial de un hábitat paleolítico. In *O Paleolítico: Actas do IV Congresso de Arqueologia Peninsular (Faro, 14 a 19 de Setembro de 2004)*, edited by Nuno Bicho and María Soledad Corchón Rodríguez, pp. 123–142. Centro de Estudos de Patrimonio, Departamento de Historia, Arqueologia e Patrimonio, Universidade do Algarve, Faro, Portugal.

Arias, Pablo, Roberto Ontañón, Esteban Álvarez-Fernández, Marián Cueto, Mikel Elorza, and Cristina García-Moncó. 2011. Magdalenian Floors in the Lower Gallery of La Garma: A Preliminary Approach. In *Site-Internal Spatial Organization of Hunter-Gatherer Societies: Case Studies from the European Palaeolithic and Mesolithic*, edited by Sabine Gaudzinski, Olaf Jöris, Martina Sensburg, Martin Street, and Elaine Turner, pp. 31–51. Verlag des Römisch-Germanischen Zentralmuseums, Mainz, Germany.

Batist, Zachary and Joe Roe. 2023. Open-Archaeo: A Resource for Documenting Archaeological Software Development Practices. *Journal of Open Archaeology Data* 11(9):1–7. https://doi.org/10.5334/joad.111.

Batist, Zachary and Joe Roe. 2024. Open Archaeology, Open Source? Collaborative Practices in an Emerging Community of Archaeological Software Engineers. *Internet Archaeology* 67. https://doi.org/10.11141/ia.67.13.

Boyd, Michael J., Rosie Campbell, Roger C. P. Doonan, Catherine Douglas, Georgios Gavalas, Myrsini Gkouma, Claire Halley, et al. 2021. Open Area, Open Data: Advances in Reflexive Archaeological Practice. *Journal of Field Archaeology* 46(2):62–80.

Brovelli, Maria, and Diego Magni. 2003. An Archaeological Web GIS Application Based on Mapserver and PostGIS. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 34(5/W12):89–94.

Calzado-Martínez, Alberto, Ángel Luis García-Fernández, Lidia M. Ortega-Alvarado, and Francisco R. Feito-Higueruela. 2023. Integrated Information System for 3D Interactive Reconstruction of an Archaeological Site. *Journal on Computing and Cultural Heritage* 16(3):44. https://doi.org/10.1145/3586077.

Cocca, Enzo, and Luca Mandolesi. 2016. Analisi statistiche e geostatistiche con pyarchinit: Prima sperimentazione. *Archaeologia e Calcolatori* Supp. 8:132–140.

Colleter, Rozenn, Jean-Baptiste Romain, and Jean-Baptiste Barreau. 2020. HumanOS: An Open Source Nomadic Software Database for Physical Anthropology and Archaeology. *Virtual Archaeology Review* 11(23):94–105. https://doi.org/10.4995/var.2020.13422.

Cosmas, John, Take Itegaki, Damian Green, Nisha Joseph, Luc Van Gool, Alexey Zalesny, Desi Vanrintel, et al. 2003. Providing Multimedia Tools for Recording, Reconstruction, Visualisation and Database Storage/Access of Archaeological Excavations. In *4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage, VAST (2003)*, edited by David Arnold, Alan Chalmers, and Franco Niccolucci, pp. 165–174. Eurographics, Eindhoven, Netherlands.

Couillet, Armelle, Hélène Rougier, Dominique Todisco, Josserand Marot, Olivier Gillet, and Isabelle Crevecoeur. 2022. New Visual Analytics Tool and Spatial Statistics to Explore Archeological Data: The Case of the Paleolithic Sequence of La Roche-à-Pierrot, Saint-Césaire, France. *Journal of Computer Applications in Archaeology* 5(1):19–34. https://doi.org/10.5334/jcaa.81.

Cueto, Marián, Edgar Camarós, Pedro M. Castaños, Roberto Ontañón, and Pablo Arias. 2020. Highlighting the Role of Carnivores as a Multifunctional Resource among the Middle Magdalenian: The Case of the Lower Galley of La Garma (Cantabria, Spain). *Journal of Archaeological Science: Reports* 30:102221. https://doi.org/10.1016/j.jasrep.2020.102221.

Dibble, Harold L., and Shannon P. McPherron. 1988. On the Computerization of Archaeological Projects. *Journal of Field Archaeology* 15(4):431–440. https://doi.org/10.1179/jfa.1988.15.4.431.

Dibble, Harold L., and Shannon P. McPherron. 1997. The Making of *Combe-Capelle on CD-ROM*. *Journal of Field Archaeology* 24(1):59–66. https://doi.org/10.1179/jfa.1997.24.1.59.

Ducke, Benjamin. 2012. Natives of a Connected World: Free and Open Source Software in Archaeology. *World Archaeology* 44(4):571–579. https://doi.org/10.1080/00438243.2012.743259.

Eichert, Stefan. 2014. *OpenATLAS – An Open Source Database Application for Archaeological, Historical, and Spatial Data*. In *Proceedings of the 18th Conference on Cultural Heritage and New Technologies CHNT, Vienna, Austria, 11–13 November 2013*, edited by Wolfgang Börner and Susanne Uhlirz, pp. 1–8. Museen der Stadt Wien – Stadtarchäologie, Vienna.

Ellis, Steven J. R. 2016. Are We Ready for New (Digital) Ways to Record Archaeological Fieldwork? A Case Study from Pompeii. In *Mobilizing the Past for a Digital Future: The Potential of Digital Archaeology*, edited by Erin Walcek Averett, Jody Michael Gordon, and Derek B. Counts, pp. 51–75. Digital Press at the University of North Dakota, Grand Forks.

Galeazzi, Fabrizio. 2016. Towards the Definition of Best 3D Practices in Archaeology: Assessing 3D Documentation Techniques for Intra-Site Data Recording. *Journal of Cultural Heritage* 17:159–169.

Gencheva, Veronika. 2023. From Data Collection to Analysis: Designing a Relational Database for Archaeological Research in the Eastern Rhodope Region. *Bulgarian E-Journal of Archaeology* 13(1):99–114.

Holdaway, Simon J., Joshua Emmitt, Rebecca Phillipps, and Sina Masoud-Ansari. 2019. A Minimalist Approach to Archaeological Data Management Design. *Journal of Archaeological Method Theory* 26:873–893. https://doi.org/10.1007/s10816-018-9399-6.

Kansa, Eric, and Sarah W. Kansa. 2021. Digital Data and Data Literacy in Archaeology Now and in the New Decade. *Advances in Archaeological Practice* 9(1):81–85.

Katsianis, Markos, Spyros Tsipidis, Kostas Kotsakis, and Alexandra Kousoulakou. 2008. A 3D Digital Workflow for Archaeological Intra-Site Research Using GIS. *Journal of Archaeological Science* 35(3):655–667. https://doi.org/10.1016/j.jas.2007.06.002.

Mandolesi, Luca. 2009. PyArchinit-Python, QGIS e PostgreSQL per la gestione dei dati di scavo. *Archaeologia e Calcolatori* Suppl. 2:209–222.

Martínez-Carrillo, Ana L., Arturo Ruiz-Rodríguez, Francisco Mozas-Martínez, and José M. Valderrama-Zafra. 2009. An Interactive System for Storage, Analysis, Query and Visualization of Archaeological Pottery. Paper presented at the 37th Annual International Conference on Computer Applications and Quantitative Methods in Archaeology, CAA, Williamsburg, Virginia.

Martin-Pozas, Tamara, Angel Fernandez-Cortes, Soledad Cuezva, Valme Jurado, Jose Luis Gonzalez-Pimentel, Bernardo Hermosin, Roberto Ontañon, et al. 2024. Microclimate, Airborne Particles, and Microbiological Monitoring Protocol for Conservation of Rock-Art Caves: The Case of the World-Heritage Site La Garma Cave (Spain). *Journal of Environmental Management* 351:119762. https://doi.org/10.1016/j.jenvman.2023.119762.

McPherron, Shannon P. and Harold L. Dibble. 1987. Hardware and Software Complexity in Computerizing Archaeological Projects. *Advances in Computer Archaeology* 4:25–40.

McPherron, Shannon P., and Harold L. Dibble. 2002. *Using Computers in Archaeology: A Practical Guide*. McGraw-Hill, Boston.

McPherron, Shannon P., Harold L. Dibble and Deborah Olszewski, 2008. GPS Surveying and On-Site Stone Tool Analysis: Equipping Teams for Landscape Analysis in the Egyptian High Desert. In *Layers of Perception: Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2–6, 2007*, edited by Posluschny, Andreas, Lambers, Kai, Herzog, Irmela, p. 9. Dr. Rudolf Habelt GmbH, Bonn, Germany.

Montagnetti, Roberto, and Luca Mandolesi. 2020. QGIS, Pyarchinit and Blender: Surveying and Management of Archaeological Data with Open Source Solutions. *Archeomatica* 10(4):30–41.

Neubauer, Wolfgang. 2004. GIS in Archaeology: The Interface between Prospection and Excavation. *Archaeological Prospection* 11(3):159–166.

Nigst, Philip R., Thomas B. Viola, and Walpurga Antl-Weiser. 2010. Digital Documentation of Paleolithic Excavations: A Case Study. In *New Aspects of the Central and Eastern European Upper Paleolithic—Methods, Chronology, Technology and Subsistence: Symposium by the Prehistoric Commission of the Austrian Academy of Sciences, Vienna, November 9–11, 2005*, edited by Christine Neugebauer-Maresch and Linda R. Owen, pp. 311–317. Verlag der Österreichischen Akademie der Wissenschaften, Vienna.

Ontañón, Roberto. 2003. Sols et structures d'habitat du Paléolithique Supérieur, nouvelles données depuis les Cantabres: La Galerie Inférieure de La Garma (Cantabrie, Espagne). *L'Anthropologie* 107(3):333–363.

Ontañón, Roberto (coordinator). 2008. Paleolithic Cave Art of Northern Spain (Extension to Altamira). Proposal of Inscription of Properties in the UNESCO World Heritage / El arte rupestre paleolítico de la Cornisa Cantábrica (extensión de Altamira). Propuesta de inscripción de bienes en la lista del Patrimonio Mundial de la UNESCO. Comisión de coordinación del bien "Arte rupestre paleolítico de la Cornisa Cantábrica", Santander, Spain.

Rayne, Louise, Jennie Bradbury, David Mattingly, Graham Philip, Robert Bewley, and Andrew Wilson. 2017. From Above and on the Ground: Geospatial Methods for Recording Endangered Archaeology in the Middle East and North Africa. *Geosciences* 7(4):100. https://doi.org/10.3390/geosciences7040100.

Reed, Denne, W. Andrew Barr, Shannon P. McPherron, Rné Bobe, Denis Geraads, Jonathan G. Wynn, and Zeresenay Alemseged. 2015. Digital Data Collection in Paleoanthropology. *Evolutionary Anthropology: Issues, News, and Reviews* 24(6):238–249. https://doi.org/10.1002/evan.21466.

Sanchez-Moral, Sergio, Valme Jurado, Angel Fernandez-Cortes, Soledad Cuezva, Tamara Martin-Pozas, Jose Luis Gonzalez-Pimentel, Roberto Ontañon, and Cesareo Saiz-Jimenez. 2021. Environment-Driven Control of Fungi in Subterranean Ecosystems: The Case of La Garma Cave (Northern Spain). *International Microbiology* 24:573–591. https://doi.org/10.1007/s10123-021-00193-x.

Sauer, Florian. 2023. A GIS-Based Digital Documentation Protocol for High-Resolution Documentation of Paleolithic Sites. *Journal of Paleolithic Archaeology* 6:21. https://doi.org/10.1007/s41982-023-00148-w.

Stamnas, Anastasios, Dimitrios Kaimaris, Charalampos Georgiadis, and Petros Patias. 2021. Comparing 3D Digital Technologies for Archaeological Fieldwork Documentation: The Case of Thessaloniki Toumba Excavation, Greece. *ISPRS—International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 46(1):713–720.

Trimmis, Konstantinos P. 2018. Paperless Mapping and Cave Archaeology: A Review on the Application of DistoX Survey Method in Archaeological Cave Sites. *Journal of Archaeological Science: Reports* 18:399–407.

Vandenbulcke, Annelies, Samuel Van Ackere, Michiel Decock, Cornelis Stal, and Alain De Wulf. 2016. Preservation of the Archaeological Heritage of the North Sea Using webGIS. In *Informatics, Geoinformatics, and Remote Sensing, Conference Proceedings, Vol. 3, Cartography & GIS. International Multidisciplinary Scientific GeoConference-SGEM*, pp. 375–382. Stef92 Technology, Sofia, Bulgaria.

Vincent, Matthew L., Falko Kuester, and Thomas E. Levy. 2013 Opendig: In-Field Data Recording for Archaeology and Cultural Heritage. In *2013 Digital Heritage International Congress (DigitalHeritage), Marseille, France, 2013*, pp. 539–542. Institute of Electrical and Electronics Engineers, New York. https://doi.org/10.1109/DigitalHeritage.2013.6744823

Vincent, Matthew L., Falko Kuester, and Thomas E. Levy. 2014. OpenDig: Digital Field Archaeology, Curation, Publication, and Dissemination. *Near Eastern Archaeology* 77(3):204–208. https://doi.org/10.5615/neareastarch.77.3.0204.

Wallrodt, John. 2016. Why Paperless: Technology and Changes in Archaeological Practice, 1996–2016. In *Mobilizing the Past for a Digital Future: The Potential of Digital Archaeology*, edited by Erin Walcek Averett, Jody Michael Gordon, and Derek B. Counts, pp. 33–50. Digital Press at the University of North Dakota, Grand Forks.