Virtual Archaeology Review, **10**(20): 17-27, 2019
https://doi.org/10.4995/var.2019.10377

Received: June 12, 2018
Accepted: November 6, 2018

ARCHAEOLOGICAL DOCUMENTATION AND DATA SHARING: DIGITAL SURVEYING AND OPEN DATA APPROACH APPLIED TO ARCHAEOLOGICAL FIELDWORKS

DOCUMENTACIÓN ARQUEOLÓGICA E INTERCAMBIO DE DATOS: TOPOGRAFÍA DIGITAL Y PROCEDIMIENTO CON DATOS ABIERTOS APLICADO A TRABAJOS DE CAMPO ARQUEOLÓGICO

Mattia Previtali*, Riccardo Valente

Politecnico di Milano, Department of Architecture, Built Environment and Construction Engineering, Via Ponzio 31, 20133 Milano, Italy.
mattia.previtali@polimi.it; riccardo.valente@polimi.it

Highlights:

- Open data will change the accessibility and disciplinary expertise in the archaeological field.
- The main aim of this paper is to present a workflow for archaeological data sharing as open data with a large level of interoperability.
- Digital acquisition techniques are used to document archaeological excavations and a Geographic Information System (GIS) is generated that is published as open data.

Abstract:

The open data paradigm is changing the research approach in many fields such as remote sensing and the social sciences. This is supported by governmental decisions and policies that are boosting the open data wave, and in this context archaeology is also affected by this new trend. In many countries, archaeological data are still protected or only limited access is allowed. However, the strong political and economic support for the publication of government data as open data will change the accessibility and disciplinary expertise in the archaeological field too. In order to maximize the impact of data, their technical openness is of primary importance. Indeed, since a spreadsheet is more usable than a PDF of a table, the availability of digital archaeological data, which is structured using standardised approaches, is of primary importance for the real usability of published data. In this context, the main aim of this paper is to present a workflow for archaeological data sharing as open data with a large level of technical usability and interoperability. Primary data is mainly acquired through the use of digital techniques (e.g. digital cameras and terrestrial laser scanning). The processing of this raw data is performed with commercial software for scan registration and image processing, allowing for a simple and semi-automated workflow. Outputs obtained from this step are then processed in modelling and drawing environments to generate digital models, both 2D and 3D. These crude geometrical data are then enriched with further information to generate a Geographic Information System (GIS) which is finally published as open data using Open Geospatial Consortium (OGC) standards to maximise interoperability.

Keywords: Virtual archaeology; 3D reconstruction; open data; Geographic Information System (GIS)

Resumen:

El paradigma de los datos abiertos está cambiando el enfoque de investigación en muchos campos de estudio como son la teledetección y las ciencias sociales. Está respaldado por decisiones gubernamentales y políticas que están impulsando la ola de los datos abiertos y en este contexto también la arqueología se ve afectada por esta nueva tendencia. En muchos países, los datos arqueológicos todavía están protegidos o se permite un acceso limitado. Sin embargo, el fuerte apoyo político y económico hacia la publicación de datos gubernamentales como datos abiertos también cambiará el panorama en el campo arqueológico. Para maximizar el efecto de los datos, su apertura técnica es de primordial importancia. De hecho, dado que una hoja de cálculo es más útil que la digitalización PDF de una tabla, de manera similar, la disponibilidad de los datos arqueológicos digitales y su estructuración mediante enfoques estandarizados es de importancia primordial para una utilización real de los datos publicados. En este contexto, el objetivo principal de este documento es presentar un flujo de trabajo para compartir datos arqueológicos como datos abiertos con un gran nivel de usabilidad técnica e interoperabilidad. La adquisición de datos primarios se realiza principalmente mediante el uso de técnicas de adquisición digital (por ejemplo, cámaras digitales y escaneado láser terrestre). El procesamiento de los datos crudos se realiza con software comercial para el registro de los escaneados y el procesamiento de imágenes que permite un procesamiento simple y semiautomático. Los resultados del área arqueológica obtenidos de este paso se procesan en modelos y entornos de dibujo que permiten generar modelos digitales, 2D y 3D. Esos datos geométricos crudos se enriquecen luego con información adicional para generar un Sistema de Información Geográfica (SIG) que finalmente se publica con datos abiertos usando los estándares del Consorcio Geoespacial Abierto (OGC) que maximizan la interoperabilidad.

Palabras clave: arqueología virtual; reconstrucción 3D; información abierta; Sistema de Información Geográfica (SIG)

*Corresponding author: Mattia Previtali, mattia.previtali@polimi.it

1. Introduction

Economists define data as non-rivalrous goods. It means that data can be processed several times and their values do not diminish over time (Samuelson, 1954). In addition, their values arise from what they reveal in the aggregate, namely innovative things, which can be realized by combining data in new ways. The constant enhancement of digital applications for producing, storing and manipulating data has brought the focus onto data-driven and data-led science even in Humanities. Indeed, nowadays the management of large digital datasets is a normal routine for archaeologists, who deal with very different types of data coming from fieldwork or lab research. Data of various natures (statistics, topographic measurements, laboratory analysis, historical sources) can be collected, stored, processed and analysed through the use of digital tools. This trend is going to increase even more in the following years with the gradual integration of digital technologies to most of the archaeological topics.

This process has increased exponentially the amount of data that can be processed. However, most of the raw data are normally excluded by the common process of knowledge sharing, because usually are not disseminated along with the results of researches, preventing any possible further use and reanalysis by different subjects, for crosschecking purposes or for other aims. Digital technologies particularly facilitate the exchange and interoperability, offering new possible approaches to the management of data and to the entire discipline. Many archaeologists seem to be unaware that the value of research data increases if they are available as open access. The use of digital technologies is fostering the development of e-research that is the way scientific knowledge is produced and shared (Beaulieu & Wouters, 2009; Boulton et al., 2012). Sharing has become a new scientific paradigm, and, if properly sustained by economic and political choices, will lead to open access research data, making data openly available to public/private stakeholders, and citizens (Wessels et al., 2014). Moreover, web infrastructures can easily store and distribute datasets composed by digital images (raster files) or vector files related to archaeological fieldworks. Due to the constant diffusion of digital surveying techniques, such as digital photogrammetry, more and more data are "born digital" providing a large number of digital data which can reciprocally increase what previously described; 3D outputs had also a significative increase over the last years (Olson, Gordon, Runnels, & Chomyszak, 2014). The spread of digital technology caused a real revolution in archaeology, modifying methodological approaches and field practices (Campana, Scopigno, & Carpentiero, 2016).

In this context, the main aim of this paper is to present a workflow for archaeological data sharing as open data[1] with a large level of technical usability and interoperability. In order to achieve a good level of automation during on-site activities and increase usability in the following processing stages, acquisition of primary data is mainly carried out by using digital acquisition techniques (e.g. digital cameras and terrestrial laser scanning). The processing of the raw data is performed by means of commercial software for scan registration (Faro Scene) and image processing (Agisoft PhotoScan) that allows for simple and semi-automated workflow. Outputs obtained from this step are then processes in modelling and drawing environments to generate digital models, both 2D and 3D, of the archaeological area. Those crude geometrical data are then enriched with further information to generate an Information System which is finally published as open data by using Open Geospatial Consortium (OGC) standards to maximize interoperability.

The paper is structured as follows: Section 2 presents an overview of the open data approach applied to archaeology and a brief overview of the Italian legislative framework, as well as one of the major European initiatives regarding archaeological open data. The description of the presented method to manage fieldwork documentation in digital format is presented in Section 3, together with the possibility to share different kinds of data through Web Map Services (WMS), so as to ensure a better data circulation among different subjects. Datasets are mainly composed of raster and vector files, as well as by 3D digital objects (sub-Section 3.1). Further interoperability with 3D GIS is also presented (sub-Section 3.2). Sections 4 and 5 are focused on discussion and conclusions, respectively.

2. Open data and archaeology

2.1. Open data policy: property, sharing and reusing

The term "open data" means a set of generic raw or partially interpreted data made freely available for multiple uses. Obviously, this approach highly increased with the wide use of digital data, which can be easily and remotely shared (Kitchin 2014). As a first consequence of the increasing importance, a lack of clarity about key terms in literature and public debates related to open data has to be highlighted. In particular, the ambiguity of widely-used terms like "open" and "free" has caused misunderstanding, mixing-up concepts like "free usage" and "free of charge", and consequently nourishing the *gratis* (i.e. for zero price) vs *libre* (i.e. with little or no restriction) debate (Zuiderwijk, & Janssen, 2014, Dawes, Vidiasova, & Parkhimovich, 2016). Access and availability, reuse and redistribution as well as a wide participation are the key points of every open data initiative; as stated by the Open Knowledge Foundation, "Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike". Starting from this definition it is explicitly clarified that, in this definition, free matches the libre concept. The key features of openness can briefly be summarized as follows:

- Availability and access: data must be available as a whole and at no more than a reasonable reproduction cost, preferably by doing unloading over the internet. The data must also be available in a convenient and modifiable form. More details on this point are discussed in Section 2.2.
- Reuse and redistribution: data should be provided under terms that permit reuse and redistribution including the intermixing with other datasets.

[1] Data can be accessed as WMS at: http://geoserver.atlas.polimi.it/geoserver/ows?service=wms&version=1.3.0&request=GetCapabilities

- Participation: data can be used, reused and redistributed according to certain specific restriction (e.g. non-commercial).

Data can be published considering different grades of openness. Copyright law in most countries by default grants copyright holders monopolistic control over their creations. Distributing a work under public or open copyright licenses, copyright holders give permission for others to copy or change their work. Generally, the so-called 5Rs (Wiley, 2014) are put forward as a framework for assessing the extent to which content is open:

- Retain: the right to make, own, and control copies of the content (e.g. download, duplicate, store, and manage);
- Reuse: the right to use the content in a wide range of ways (e.g. in a class, in a study group, on a website, in a video);
- Revise: the right to adapt, adjust, modify, or alter the content itself (e.g. translate the content into another language);
- Remix: the right to combine the original or revised content with other open content to create something new (e.g. incorporate the content into a mashup);
- Redistribute: the right to share copies of the original content, your revisions, or your remixes with others (e.g. give a copy of the content to a friend).

At the international level, the most used public copyright license is Creative Commons (Creative Commons, 2018a; Creative Commons, 2018b). A CC BY-SA license is used when an author wants to give people the right to share, use, and build upon a work that they have created. The CC licenses all grant the “baseline rights”, such as the right to distribute the copyrighted work worldwide for non-commercial purposes, and without modification. The detail of each of these licenses depend on the version, and comprises a selection out of four conditions:

- Attribution (BY): Licensees may copy, distribute, display and perform the work and make derivative works based on it only if they give the author or licensor the credits in the manner specified by these;
- Share-alike (SA): Licensees may distribute derivative works only under a license identical to the license that governs the original work;
- Non-commercial (NC): Licensees may copy, distribute, display, and perform the work and make derivative works based on it only for non-commercial purposes; and
- No Derivative Works (ND): Licensees may copy, distribute, display and perform only verbatim copies of the work, not derivative works based on it.

2.2. Open data openness

From a user point of view, accessing open data is a process that goes from discovery, to access and finally to data utilization and exploitation. Data discovery involves the use of services such as metadata catalogues to find data of particular interest over a specific geographic region or according to some keywords or tags. On the other hand, access involves the order, packaging and delivery (according to the publication standards and protocols) of the data specified. Finally, exploitation is what the final user does with the data for their own purpose. In the last decade with the growth of the Web-based technologies, access has become a demand-driven operation. Consumers expect simple discover and access to cheap (or free) data in simple standard formats that can be easily used. Indeed, while in the past the focus of data access was on the supplier side (e.g. public administration, municipalities, etc.) with a strong emphasis on standards (e.g. OGC Standards, Comprehensive Knowledge Archive Network (CKAN), etc.), nowadays, increasingly non-traditional suppliers are offering services allowing broader participation in the open data world. Moreover, the further democratisation of access to geospatial data enables value-added suppliers to create new data products and services. For this reason, the trend is to worry about the interfaces to the data. This allows the data to be managed in the best manner possible while providing open, standards-based access. These aspects in the open data world introduce two interconnected issues: 1) data openness (how much data are really easily accessible?); and 2) data formats and standards (which formats and standards are used?). Data openness is strictly connected to data usability. Indeed, the possibility to re-use data is reduced in case a proprietary format is used for its publication. The need of assuring usability of data, was recognized since the beginning of the open data movement with the suggestion by Tim Berners-Lee (Bizer, Heath & Berners-Lee, 2009), inventor of the World Wide Web, of a five-star rating system (http://5stardata.info/en/) based on the openness characteristics of data (Table 1). According to that classification, the more stars an open dataset has, the more it is usable. For example, an Excel table is more usable than a scanned version of the table published as PDF. Similarly, the availability of an open format (like CSV) makes the data more accessible than the publishing in property formats (such as Excel). The availability of products on the web through the Uniform Resource Identifier (URI) and standard interfaces is a further step in the data openness.

Table 1: 5-star open data classification as proposed by Tim Berners-Lee.

Star #	*Description*
★	Make your stuff available on the Web (whatever format) under an open license
★★	Make it available as structured data (e.g. Excel instead of image scan of a table)
★★★	Use non-proprietary formats (e.g. CSV instead of Excel)
★★★★	Use URIs to denote things, so that people can point at your stuff
★★★★★	Link your data to other data to provide context

2.3. Archaeological data and open approach

The “open data wave” is gaining large importance in many fields such as research, engineering, science, government, public administration, etc. (Egger-Peitler & Polzer, 2014; Burkel, 2016). This wave is having its first impact, and more is expected in the next future, also in the archaeological field, changing in some aspects the on-field activities and the dissemination of excavation results. Indeed, the adoption of this paradigm is strictly connected with two important problems: 1) regulations about the property, the sharing and the use of data; and 2) the architecture used for data sharing. Here the term "data" obviously refers to native digital data, although paper-based (or other formats) data can be digitalized and partially considered as well. This section focuses on discussing those two main issues with specific attention to the archaeological domain.

The open data approach is considered by several scholars as an exceptional opportunity for archaeology, able to improve the overall research, stimulate new collaborations, widen the perspectives and allow better circulation of knowledge. Increased stress on data would also bring to a change in the consideration of archaeological data, as an effective part of the archaeological research, not only the basis where interpretations are grounded (Costa, Beck, Bevan & Ogden, 2012). Nevertheless, there are some points which have to be carefully considered. The copyright of data is usually assured by Creative Commons licenses; oppositions on this topic are far from being rare within the archaeological community, fearing possible plagiarism or misuse of data which are considered part of personal works and researches (Anichini & Gattiglia, 2015).

When dealing with archaeological data, there is a difference between publications, which communicate data but especially information, and raw data, which can be analysed or reanalysed, depending on needs and aims. The openness of the latter is of major importance for archaeology: this is due to the unrepeatability of the archaeological investigations (Anichini & Gattiglia, 2015). Scientific publications need to be based on proved sets of data, even if these sets could be sometimes not directly available to other researchers (Richards, & Winters, 2015). Moreover, the archaeological data can be considered as different from those ones belonging to other subjects, and characterized by a strong heterogeneity and personal influences. Among the different sources, SUs (stratigraphic units) lists, SU datasheets, stratigraphic charts, plans and sections, images, drawings, videos, written texts, geographic data, etc. have all been considered as archaeological data (Anichini & Gattiglia, 2015).

This concept of data openness is relevant also when publishing archaeological data. In particular, archaeology deals with large datasets which are mainly the result of field activities and researches. These datasets can be very various in nature, as well as a number of attributes and formats. For instance, researches on findings, such as potteries, coins or metal items, usually produce large sets of records with a relevant number of attributes for each record; this kind of data can be normally managed through relational databases, which organizes contents into tables. Other important data coming from archaeological investigations are spatial data, including maps, ancient cartography, plans, sections, and 3D data (Portalés et al., 2017). As the use of digital surveying techniques has been constantly increasing over the last decade, new datasets are available. Topographic measurements, such as points measured by total station or Global Navigation Satellite System (GNSS) devices, digital elevation models (DEMs), georeferenced orthophotos, vector drawings, GIS vector layers, three-dimensional meshes, point clouds are all data which can be used to support traditional archaeological research. These kinds of new datasets can be the result of previous processing phases, such as orthoimages, or come as raw data to be processed, such as sets of digital images or topographic points. This is a relevant distinction, since before the digital revolution most of the archaeological outputs were made public only in processed versions; the possibility to get also raw data and to use them in various ways increases the possible outcomes, boosting innovative and alternative uses of data which could be exploited for purposes other than those ones they were collected for (Scopigno et al., 2017). It also allows users to compare obtained results using different tools or methodologies.

2.4. Regulatory framework: Italian and European overview

The Italian approach in the past was quite conservative and non-digital formats did not help a wider diffusion different from usual publications (Gattiglia 2009; Richards, & Winters, 2015). The archaeological documentation has been considered for a long time as belonging to the public administration. The L. 633/1941, Art. 11, states that public administrations own the rights on works created on their behalf or with their funds. The real status of the archaeological documentation in Italy is still a matter of debate (Trabucco, 2009). However, nowadays the archaeological documentation is usually accessible by citizens only by explicit request (L. 241/1990, *capo* V and D. P. R. 184/2006); furthermore, the commercial use of these data is generally not taken in consideration (D. Lgs. 32/2010) and the free access has been explicitly guaranteed to archaeologists involved into projects of preventive archaeology (Circular 10/2012 by *Direzione Generale per le Antichità*). Despite the traditional approach followed in Italy, several changes and experiments happened in the last years to go towards a more open availability of archaeological data. Some Italian public administrations started to make available datasets concerning statistics, catalogue information and cartographic data related to their territories. At national level, some datasets related to archaeological museums or archaeological presences are retrievable online, as well as at regional level[2]. The Italian Ministry of Culture has its own open data page[3]. Several public regulations (CAD 82/2005; D. Lgs. 36/2006–D.L.179/2012 and D. Lgs. 33/2013) stress the openness to information; the 179/2012 states that public data, when shared without any particular prescription, are to be considered as open ("open by default"). Moreover, one of the latest regulations issued by the Italian Ministry of Culture (n. 1–20 gennaio 2016, attached n. 3, p. 2), explicitly refers to a possible open use of archaeological data. Currently, one of the active online repositories for archaeological data in Italy is one created for the MAPPAproject, and carried

[2] Web link: http://www.dati.gov.it; last access: March 2018. https://www.dati.lombardia.it; last access: March 2018. http://www.sistemonet.it/sistemonet/index.do; last access: March 2018.

[3] Web link: http://www.beniculturali.it/mibac/export/MiBAC/sito-MiBAC/MenuPrincipale/Trasparenza/Open-Data/index.html; last access: March 2018.

out by the University of Pisa[4]. Despite the lack of a common line of action in Italy, several movements supporting the open approach for archaeological data have recently raised, gathering archaeologists for the common aim of better research and development.

At the European level, an important initiative is the ARIADNE Project[5]. Its aim is primarily "to create a community of use fostering the creation, sharing, use and reuse of digital data" and "to bring together and integrate existing archaeological research data infrastructures so that researches can use the various distributed datasets". This Portal allows users to retrieve hundreds of thousands of datasets belonging to fieldworks and archaeological objects all over Europe and the project aims to be a landmark for European archaeologists. The service is intended for both public users and researchers; the interoperability is improved by using a vocabulary mapping which allows also multilingual queries. ARIADNE's mission is also to provide best practices for the creation of national archaeological databases.

In the UK, the Archaeological Data Service has been sharing for over twenty years documentation coming from archaeological investigations carried out all over the country and is a national reference for the management, storage and distribution of archaeological data[6]. The Journal of Open Archaeological Data already provides for some archaeological datasets, freely downloadable from the web[7].

3. Operative workflow

As previously anticipated, the aim of this paper is to present a methodology to share the results of archaeological fieldwork as open data with a large level of technical usability. In order to achieve a high level of usability and interoperability, the presented workflow highly relies on digital and automated techniques both in the acquisition and processing stage as well as in the publication phase. An overview of the proposed procedure is presented in Fig. 1. In particular, the main steps can be summarized as follows:

- On field survey: digital surveying techniques are used to document the different phases of the archaeological excavation in order to collect both geometric and semantic information. In this step, photogrammetry plays a major role in the acquisition of geometric information while electronic spreadsheets can be used to collect information about archaeological features (such as dimensions, characteristics, composition, etc.).
- 3D digital modelling: the raw data acquired on-site are than processed to obtain some intermediate products such as point clouds or meshes. Although many studies refer to those digital photogrammetric results as final outputs (Kersten & Lindstaedt, 2012; De Reu et al., 2013) for 3D registration of archaeological features, it is worth underlining that 3D products represent only a digital surface; they do not record and represent digital objects. For this reason, a 3D modelling stage is needed to create geometric objects with an archaeological meaning.
- 3D information system generation: the crude geometrical model derived from the previous step is coupled with the electronic spreadsheets generatedon-site giving specific information to the different elements (e.g. SU number, etc.) enriching this way the digital reconstructed objects creating a 3D GIS system.
- Data publication: the results of previous elaborations are published as open data by using open formats and released as OGC compliant services.

The importance of testing 3D representations of archaeological layers in a GIS environment is in adding information to 3D contents. The first goal reached through the processing of point clouds obtained by on-site surveying was to pass from 3D surfaces to 3D objects. In this way, archaeological layers ceased to be represented only by their upper surfaces but in their full three dimensions. The second step was to add information to these objects, improving the method from visualization of stratigraphy to analysis of stratigraphy through the GIS tools. In this way the digital objects gain double independence: the first one is reached through the 3D modelling; and the second one through the addition of informative contents able to enrich the digital representation.

3.1. Digital surveying and 3D modelling

Digital surveying techniques gained a lot of popularity over the last decade, being increasingly applied to archaeological fieldworks all over the world (Russo, Remondino, & Guidi, 2011). Despite some drawbacks, which should be carefully considered by the scientific community, digital surveying usually overtakes traditional methods for what concerns accuracy, time of execution and interoperability. Accuracy is assured by the use of topographic measurements realized through total stations or laser scanners, which allows keeping the overall error constant and controlled; time spent on documentation activities on-site is substantially reduced when it is compared to manual drawings, a key-point especially when dealing with limited resources; interoperability is a concept basically introduced by digital data, since paper-based documentation is not very suitable to be processed and exchanged within digital ecosystems.

Tests presented in this paper were carried out on the S. Calocero site in Albenga (Italy) during a two-years campaign occurred in 2014 and 2015. The S. Calocero archaeological area is an important Early Christian site, situated along the hillside of a small elevation close to the historical town of Albenga (Fig. 2a). The main remains belong to a religious building founded during the 6th century and to the monastery. The entire area preserves evidence from the Roman period to the 16th century, partially discovered and excavated during the 20th century (Spadea Noviero, Pergola, & Roascio, 2010). In 2014 and 2015 several trenches (Fig. 2b) were set to explore preserved stratigraphy within the archaeological area (Pergola et al., 2014); documentation was entirely carried out through the use of digital tools. The entire documentation of excavation progress was performed through the use of

[4] Web link: http://www.mappaproject.org/; last access: March 2018.
[5] Web link: http://www.ariadne-infrastructure.eu/; last access: March 2018.
[6] Web link: http://archaeologydataservice.ac.uk/; last access: March 2018.
[7] Web link: http://openarchaeologydata.metajnl.com/; last access: March 2018.

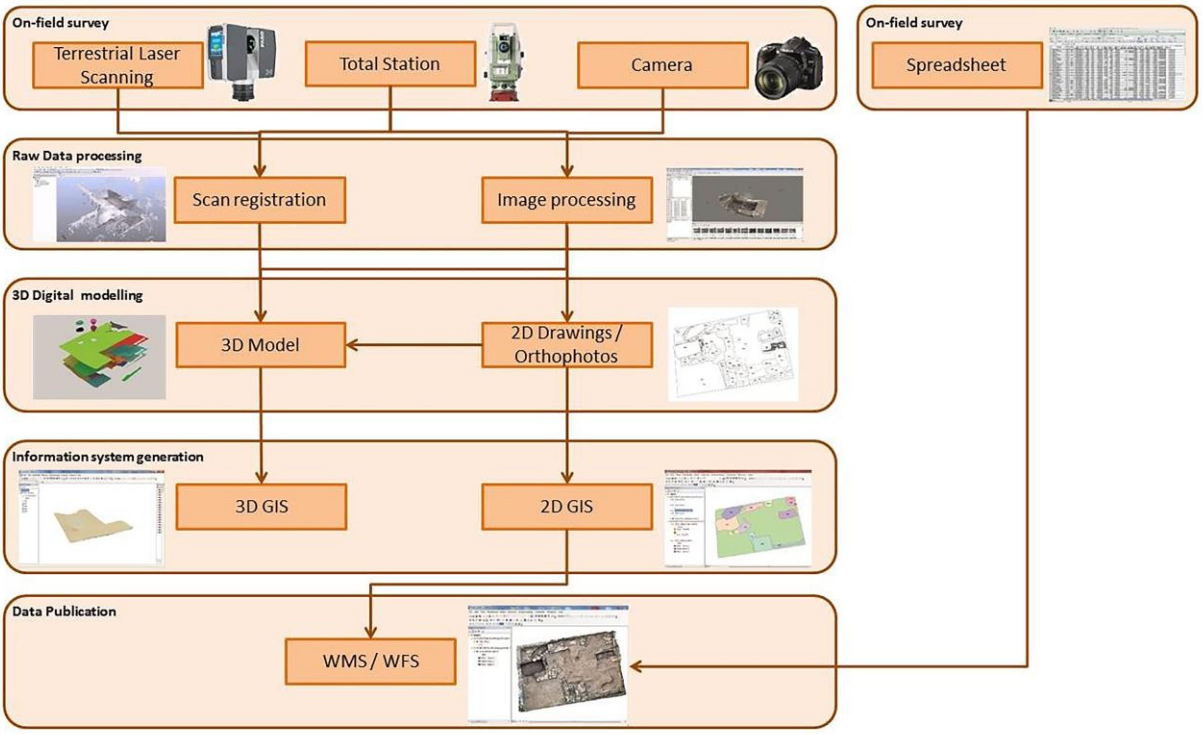


Figure 1: Workflow of the proposed approach.

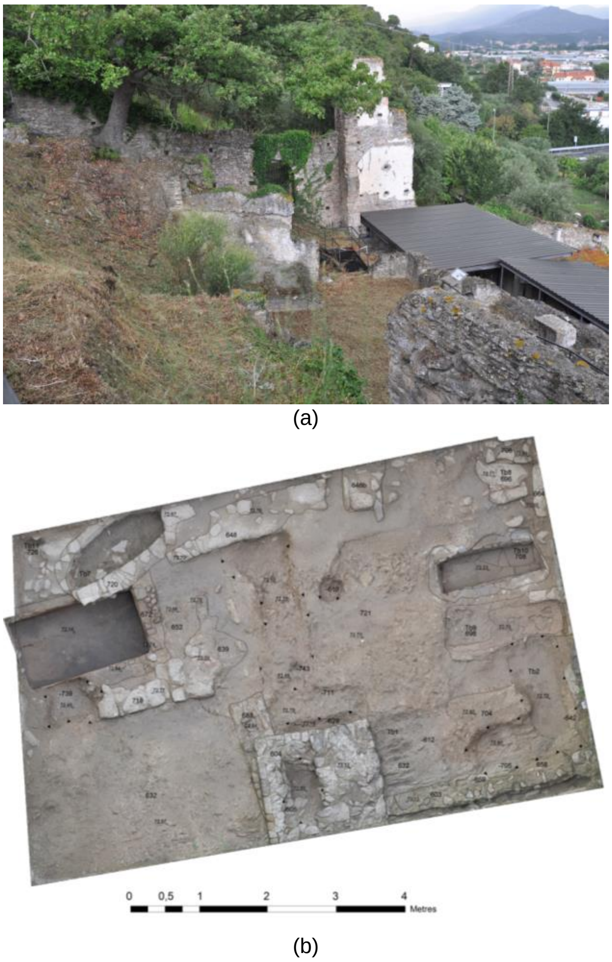


Figure 2: The S. Calocero site: a) An overview image of the archaeological area; b) One of the archaeological trenches of the presented case study.

digital photogrammetry, instrumental topographic measurements and digital modelling (Valente et al., 2017); stratigraphic sheets, sketches and field reports are the only paper-based documentation produced during the campaign; manual measurements and drawings have been skipped, relying on digital surveying. The recording activity produced a large set of different kinds of data (Table 2). Primary raw data were images and topographic points exported from the instrument in TXT or CSV formats; only the last ones was immediately usable in GIS and converted into shapefiles.

Images were normally processed with photogrammetric software: after this phase some secondary intermediate data, such as orthoimages, digital elevation models (DEMs), point clouds, meshes, and 3D PDFs were available.

Part of these outputs, i.e. orthoimages and DEMs, can be easily imported, together with the topographic points, into GIS to be georeferenced. A series of ground control points (GCPs) were set around the excavated areas, using 12 bit coded markers provided by Agisoft PhotoScan Pro v. 1.2, the selected photogrammetric software. GCPs were measured with the total station using WGS84 datum and Universe Transverse Mercator Cartographic System UTM 32N, to provide a correct georeferencing of the entire area: they were located outside the single trenches, in order to minimize the effect of excavation progresses and to avoid multiple measurements of the GCPs network. Single archaeological features (such as deposits, cuts, structures) were recorded not only with traditional photography but also with digital photogrammetry, in order to obtain a reliable image of every stratigraphic unit which could be used as a reference for drawing. Orthoimages reproducing the entire trench were also created, to better enhance excavation progresses.

Table 2: Outputs of the recording and survey activities.

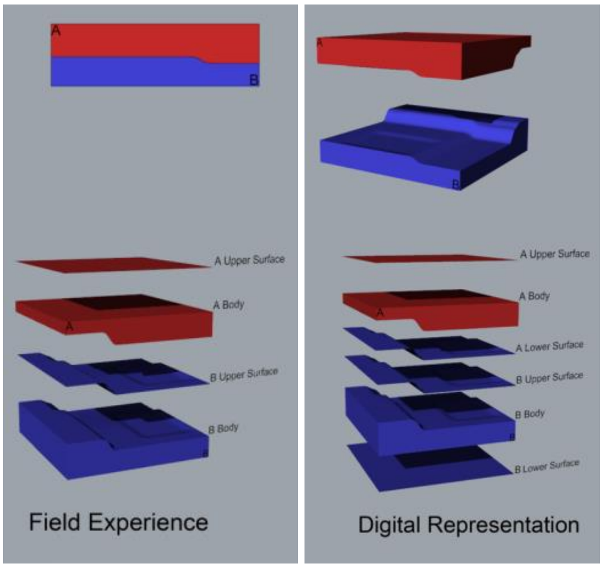
	Field Data	*Automatically Processed Data*	*Manually Processed Data*
Primary Raw Data	Topographic measurements Images		
Secondary Intermediate Data		Orthophotos Point Clouds Meshes DEMs	
Re-elaborated Data			Vector Drawings 3D Models

The third level of outputs is composed of vector drawings and digital models. Vector drawings of detected features (deposits, structures, burials, findings,…) can be obtained from georeferenced outputs (e.g. either point cloud or orthophotos) after the interpretation of archaeologists. Digital models (meshes) of archaeological features and human remains can be also obtained. Most successful tests have been performed with deposits and human remains, since only a few remains of structures were detected during the excavations.

This kind of data is different from the previous ones since they are the direct result of human interpretation, not the outcome of automatic processing. All of these digital data can be successfully imported, analysed and modified in other software and easily exchanged among different researchers. Their nature is different, but they can be comprehended into some wide categories such as TXT data (topographic measurements), raster data (images, orthoimages), vector data (vector drawings, shapefiles) and 3D geometry data (3D surfaces and models in different formats).

Some data obtained after image processing, i.e. point clouds and meshes, can be manually processed by modelling software to obtain 3D objects of archaeological features. In order to obtain a digital stratigraphic sequence from the previous datasets, a certain amount of manual operations need to be done. The only available data coming from the field are the photogrammetric recording of the upper surfaces of the discovered features. This approach is much more challenging due to the complex nature of archaeological features, with very irregular morphologies, but it goes towards an effective representation in three dimensions of the archaeological record. This aspect can be very important for improved documentation of stratigraphic sequences. Archaeological evidence discovered during fieldworks is mainly perceivable only from above (Fig. 3a).

This means that upper surfaces are the only parts which can be directly experienced by human operators: lower surfaces of deposits could be even argued not to exist, since the lower boundaries of stratigraphic layers are normally composed by the upper surfaces of more ancient and deep deposits. These peculiarities compel to face 3D representation of archaeological stratigraphy with a different approach (Valente et al., 2017).



(a) (b)

Figure 3: Differences between a) real and b) digital experience of archaeological features.

The thickness of a single layer has to be obtained by the union of upper surfaces belonging to different deposits; since every feature needs clear boundaries, surfaces could be doubled in order to make every deposit as an independent unit (Fig. 3b). A similar approach has to be extended also to feature interfaces, i.e. cuts or removals. These ones are immaterial elements of the stratigraphic sequence, which can only be detected in their material modification of the soil, i.e. in the changes of deposit morphology. However, they can be digitally modelled as pure surfaces, not belonging to any material object as stated by stratigraphic principles (Harris, 1989). These surfaces are obtained once again from the upper surfaces of deposits. If repeated for every single feature discovered during the excavation, this method would allow to obtain a 3D representation of the stratigraphic sequence only using field data obtained from digital photogrammetry; point clouds of layer and feature interfaces are converted into surfaces using a modelling software and successively joint together (Fig. 4).

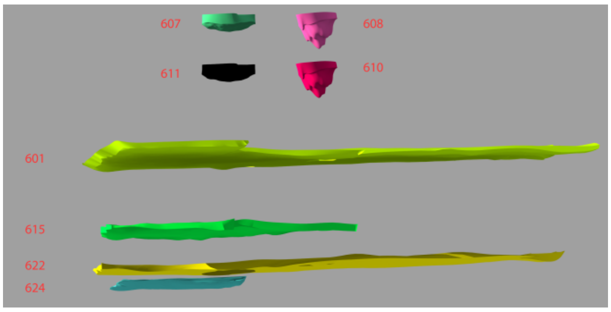


Figure 4: 3D modelling of archaeological features detected after excavation.

To perform this modelling point clouds exported as TXT files from photogrammetric projects are imported into Rhinoceros® v. 5 modelling software. This approach can be extended to other archaeological features, such as human remains. Buried individuals are perceivable in their unity only after excavation, since they are made out of single bones which are usually collected as findings. A 3D approach for archaeo-anthropological purposes allows users to obtain a 3D model of human remains as they have been discovered on-site, preserving the position of single bones and the unity of each individual. Starting again from photogrammetric outputs, i.e. textured meshes, a digital model of burials can be obtained positioning on the 3D mesh a set of digital bones, which is shared on the web under Creative Commons license (http://lifesciencedb.jp/bp3d/; last access: November 2017).

3.2. 3D GIS and WMS export

The management of 2D (plans, orthophoto) and 3D (modelled objects) layers obtained from the processing of photogrammetric outputs has been experimentally tested in GIS environments and obtained results were published as Web Map Service (WMS) and Web Feature Service (WFS).

Since all of the outputs obtained during on-site excavation are georeferenced, they can be managed also in a 2D GIS environment. For instance, being orthophotos, within the same Cartographic Coordinate System, they can easily show the diachronic sequence of the excavation progresses. A shapefile (.SHP) can be created in order to draw the single elements which compose the excavation (Fig. 5).

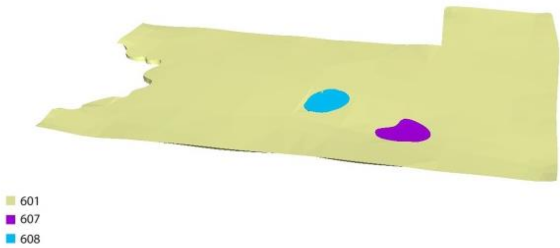
Figure 5: Vector drawing of excavation results managed in GIS.

The common ways to represent data into GIS are points, polylines and polygons. For GIS applied to excavated areas, the most used elements are points, for findings, and polygons, for SU. A relevant series of data can be connected to each element through external databases, storing, for instance, the information kept by SU sheets. Although extremely useful to store and display information, as for performing spatial analysis within the excavated area or among different excavated areas, the drawing tools of GIS are not specifically designed for detailed characterization.

Every 3D SU has been exported from Rhino in DAE format as polysurfaces. The size of the obtained polysurfaces is very small if compared to meshes: 816 kb was the greater size among the 3D layers exported. Single SUs have been successively imported in ArcScene v. 10.4 from ArcGIS suite, GIS software able to manage 3D elements. ArcGIS is one of the most used suites for spatial analysis and ArcScene has been selected because it also supports 3D contents. SUs have been imported in separated DAE files, but they kept the same positioning, so they automatically occupied the right position.

3D deposits have always been correctly uploaded in ArcScene, with no particular issues; both representations of deposits and feature interfaces have been uploaded. Digital SUs keep the same morphology, so no unwanted overlaps have been detected. Since every SU was uploaded as a different shapefile (SHP), they have been successively merged into a single file. In ArcScene it was possible to correlate different data to three-dimensional layers thanks to related charts. For what concerns deposits, the data normally stored in the SU datasheet were directly related to the relative 3D content (Fig. 6). The entries of traditional SU datasheet were adapted to the structure of the 'attribute table', becoming immediately accessible. The SU id number, measurements, descriptions of the main characteristics (such as colour, composition and compactness), the stratigraphic relationships and so on have been transferred in digital charts.

(a)



(b)

Figure 6: 3D models of archaeological features exported into 3D GIS: a) Side view; b) Top view.

Some tests have been performed in the 3D GIS environment also with the modelled skeletons. Two import methods have been tested. With the first one, shapefiles are constituted by the entire undifferentiated skeleton; with the second one, every single bone has been previously exported as a separate shapefile and then merged together with the other ones. With this second method it is possible to label single bones throughout the table of contents, i.e. adding informative contents to single bones; this aspect can be interesting for archaeo-anthropological analysis, where much information come from bone analysis.

Integration with the modelled layers has been reached within the GIS environment, making possible to display at the same time some 3D layers and burials. Some 2D outputs, mainly archaeological drawings and orthophotos, were released as open data by using as a communication interface OGC standard services published by the GIcarus Lab Server.

In particular, the server-side software used is GeoServer v. 2.6.1, a powerful open source platform for publishing spatial data and interactive mapping applications on the web. GeoServer allows the visualization of produced data through the use of OGC standards like Web Map Service (WMS, WMTS, WFS and WCS). The architecture of the adopted solution is presented in Figure 7.

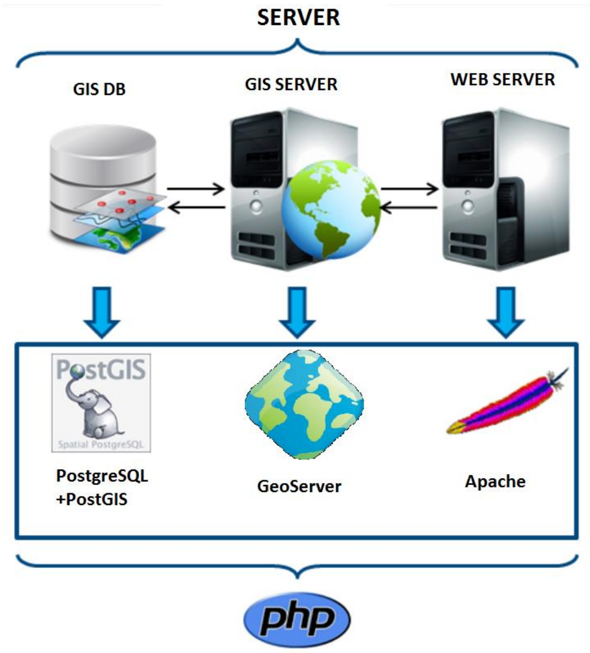


Figure 7: Architecture of the geographic open data server. Adapted from Carrion et al. (2015).

The adopted solution relays on a typical server structure:

- GIS database is implemented by using PostgreSQL with the PostGIS extension to manage geographical data;
- The GIS server used to render the data set, as previously defined, is the GeoServer open source solution;
- Tomcat Apache technology is used to deploy the system on a web server.

Data were organized into two main categories: (i) orthophotos and (ii) vector drawing of excavation results. More precisely, orthophotos were served, due to their raster nature as WMS and WMTS for fast display and loading as simple background maps. To allow raster algebra in GIS software, datasets are also provided as Web Coverage Service (WCS). Vector drawings are organized to present excavation results. In particular, data as the SU number, the nature of the SU (deposit or feature interface), and SU chronology were added as attributes of vector drawing. In this second case, data are provided with both WMS and WFS. WMS layers are mainly provided for visualization purposes even if some basic queries can be performed by Common Query Language (CQL). Instead, WFS main advantage is the possibility to perform advanced queries and analysis into a GIS environment.

4. Discussion

For what concerns archaeological data in Italy, it is important to highlight that at the moment their distribution in open format is not easily integrated within the normal documentation workflow. An important part of archaeological outputs is still produced in analogue formats, which need to be successively digitalized. Moreover, although recent regulations are more oriented towards an open data approach, national standards to archive and publish open archaeological data are still missing. Functional integration with currently applicable legislation is correspondingly missing, because there is a lack of shared coordination which could bring to a national strategy, in agreement with other European countries.

The method presented for documenting archaeological stratigraphy in three dimensions does not require any other field operation to collect data and perform elaborations, making easier its integration within the documentation workflow during fieldworks. Nevertheless, the steps necessary to obtain the final 3D stratigraphic objects can be quite time-consuming and need to be carried out manually. Some topological inconsistencies (Fig. 8) have been also observed among different archaeological layers, especially when the original thickness was very limited to a few centimetres or less, while when deposits have a solid body error seem to be acceptable. This is a constraint of the method because in case of very complex stratigraphy similar cases are expected to increase. Further research on this aspect could avoid the unwanted intersections between surfaces belonging to different stratigraphic units.

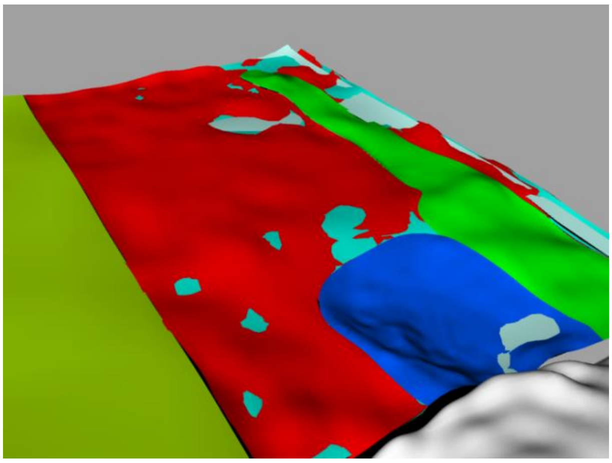


Figure 8: Example of possible topological inconsistencies of modelled features.

Some further problems may arise when applying the presented methodology to large archaeological sites. In particular, importing large 3D files in DAE format may result in errors when loading them into ArcScene software. Indeed, when loading large dataset it may be necessary to split it into several parts. For example, some human remains were also added to the GIS environment. In this case, each bone was imported separately to prevent such data overload. Some problems may also arise due to single-precision floating point accuracy of GIS software. This may jeopardize the presentation of meshes characterized by centimetre accuracy if data are generated in cartographic coordinates. In the presented case, due to the limited size of the investigated area, the generation of 3D models was performed by using false origins, truncating UTM coordinates. For example, for a point having coordinates (X = 437196.613, Y = 4877467.024) in WGS84/UTM 32N, their real-world coordinates were truncated to (X = 196.613, Y = 467.024). After importing in ArcScene the generated 3D models, UTM coordinates were restored applying a shift to the entire model (shift in X = 437000, and a shift in Y = 4877000). This trick allowed us to manage cm-resolution meshes while keeping the 3D model in real-world coordinates.

5. Conclusions

The paper presented a digital workflow principally based on topographic measurements and digital photogrammetry to document progress during archaeological fieldworks. Digital photogrammetry was proven to be very useful, modifying minimally the traditional recording workflow integrated with the excavation practice. It also allowed us to obtain various kinds of data that can be successively processed to obtain further results. For instance, 3D field data can be used to obtain 3D objects and represent in a more effective way the archaeological stratigraphy as experienced on-site. This is a step towards a fully 3D approach for archaeological documentation of stratigraphy, possibly applied on large scales that do not require complex or long procedures to collect field data. The method still has some drawbacks, such as the time needed to manually process the data or some topological inconsistencies, which could be solved in the future with more testing and research.

Most of the datasets collected during fieldworks or later processed can be shared through WMS. The constant and gradual diffusion of the open data approach in the archaeological field would rapidly increase the number of digital data available for reuse and sharing. This passage needs to be reasonably supported not only by technological improvements but also by new regulations which are necessary to control and supervise the production, sharing and use of open data, especially in those countries where data are more likely to be static. Widespread circulation of raw data could even open a new branch of archaeological research, focused on the validation and reuse of older datasets.

Acknowledgements

The authors want to thank Prof. Philippe Pergola, who had the scientific direction of the archaeological fieldworks in Albenga, and dott. Stefano Roascio and dott.ssa Elena Dellù, who managed the on-field operations; dott. Giovanni Svevo for the help with on-field survey. Thanks to Fondazione Lamboglia and Fondazione De Mari which funded fieldworks. A special thanks to Fondazione Fratelli Confalonieri which funded the research.

References

Anichini, F., & Gattiglia, G. (2015). Verso la rivoluzione. Dall'Open Access all'Open Data: la pubblicazione aperta in archeologia. *Post-classical archaeologies*, *5*, 298–326.

Beaulieu, A., & Wouters, P. (2009). E-research as intervention. In N. Jankowski (Ed.), *E-research: transformation in Scholarly Practice* (pp. 54–69). New York, NY: Routledge.

Bizer, C., Heath, T. & Berners-Lee, T. (2009). Linked data-the story so far. *International journal on semantic web and information systems*, *5*(3), 1–22. https://doi.org.10.4018/jswis.2009081901

Boulton, G., Campbell, P., Collins, B., Elias, P., Hall, W., Laurie, G., & Walport, M. (2012). *Science as an open enterprise*. London, England: Royal Society.

Burkel, V. (2016). Environmental Health Tracking Rides the Open Data Wave. *Journal of Environmental Health*, *78*(7), 36–38.

Campana, S., Scopigno, R., Carpentiero, G. & Cirillo, M. (Eds.) (2016). *CAA2015. Keep The Revolution Going: Proceedings of the 43rd Annual Conference on Computer Applications and Quantitative Methods in Archaeology*. Oxford: Archaeopress Archaeology.

Carrion, D., Migliaccio, F., Minini, G., & Zambrano, C. (2015). Open Web Services: new tools for Medievalist Historians to manage and share their research work. *Geomatics Workbooks*, *12*, 55–66. Retrieved from https://re.public.polimi.it/retrieve/handle/11311/980787/39679/FOSS4G-eu15_submission_119.pdf

Creative Commons, (2018a). *Homepage*. Retrieved April 10, 2018 from https://creativecommons.org

Creative Commons, (2018b). *About the licences*, Retrieved April 10, 2018 from https://creativecommons.org/licenses/?lang=en.

Costa, S., Beck, A., Bevan, A., & Ogden, J. (2012). Defining and advocating open data in archaeology. In G. Earl, T. Sly, A. Chrysanthi, P. Murrieta-Flores, C. Papadopoulos, I. Romanowska & D. Wheatley (Eds.), *Archaeology in the Digital Era: Proceedings of the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA) University of Southampton on 26-30 March 2012*, I (pp. 449–456). Amsterdam University Press: Amsterdam.

Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly*, *33*(1), 15–27. https://doi.org/10.1016/j.giq.2016.01.003

De Reu, J., Plets, G., Verhoeven, G., De Smedt, P., Bats, M., Cherretté, B., & Van Meirvenne, M. (2013). Towards a three-dimensional cost-effective registration of the archaeological heritage. *Journal of archaeological science*, *40*(2), 1108–1121. https://doi.org/10.1016/j.jas.2012.08.040

Gattiglia, G. (2009). Open Digital Archives in archeologia. Good practice. *Archeologia e Calcolatori supplemento, 2*, 49–63.

Egger-Peitler I. & Polzer T. (2014). Open Data: European ambitions and local efforts. Experiences from Austria. In M. Gascó-Hernández (Ed), *Open Government. Public Administration and Information Technology*, *4*. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-9563-5_9

Harris, E. (1989). *Principles of archaeological stratigraphy*. London: Academic Press.

Kersten, T. P., & Lindstaedt, M. (2012). Image-based low-cost systems for automatic 3D recording and modelling of archaeological finds and objects. In M. Ioannides, D. Fritsch, J. Leissner, R. Davies, F. Remondino, & R. Caffo (Eds), *Progress in Cultural Heritage Preservation. EuroMed 2012. Lecture Notes in Computer Science, 7616*, 1–10. Springer, Berlin, Heidelberg. https://www.doi.org/10.1007/978-3-642-34234-9_1

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. London: Sage.

Olson, B. R., Gordon, J. M., Runnels, C., & Chomyszak, S. (2014). Experimental three-dimensional printing of a lower Palaeolithic handaxe: an assessment of the technology and analytical value. *Lithic Technology*, *39*(3), 162–172. https://doi.org/ 10.1179/2051618514Y.0000000004

Pergola, P., Spadea, G., Roascio, S., Dellù, E., Castiglia, G., Svevo, G., & Valente, R. (2014). Il complesso di San Calocero ad Albenga alla luce dei nuovi dati (campagna di scavo in concessione al Pontificio Istituto di Archeologia Cristiana del 2014). *Rivista di Archeologia Cristiana*, *90*, 365–421.

Portalés, C., Alonso-Monasterio, P., & Viñals, M. J. (2017). 3D virtual reconstruction and visualisation of the archaeological site Castellet de Bernabé (Llíria, Spain). *Virtual Archaeology Review*, *8*(16), 72–85. https://doi.org/10.4995/var.2017.5890

Richards, J. D., & Winters, J. (2015). Digging into Data: Open Access and Open Data. *European Journal of Post-Classical Archaeologies*, *5*, 28–298.

Russo, M., Remondino, F., & Guidi, G. (2011). Principali tecniche e strumenti per il rilievo tridimensionale in ambito archeologico. *Archeologia e calcolatori*, *22*, 169–198.

Samuelson, P. A. (1954). The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*, *36*(4), 387–389.

Scopigno, R., Callieri, M., Dellepiane, M., Ponchio, F., & Potenziani, M. (2017). Delivering and using 3D models on the web: are we ready? *Virtual Archaeology Review*, *8*(17), 1–9. https://doi.org/10.4995/var.2017.6405

Spadea Noviero, G., Pergola, P., & Roascio, S. (2010). *Albenga. Un antico spazio cristiano. Chiesa e monastero di San Calocero al Monte*. Genova: Fratelli Frilli.

Trabucco, M. (2009). Pubblico ma non pubblico: prospettive normative sulla proprietà intellettuale dei dati archeologici. *Archeologia e Calcolatori supplemento, 2*, 65–70.

Valente, R., Brumana, R., Oreni, D., Banfi, F., Barazzetti, L., & Previtali, M. (2017). Object-oriented approach for 3D archaeological documentation. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-2/W5*, 707–712. https://doi.org/10.5194/isprs-archives-XLII-2-W5-707-2017

Wessels, B., Finn, R. L., Linde, P., Mazzetti, P., Nativi, S., Riley, S., Smallwood, R., Taylor, M. J., Tsoukala, V., Wadhwa, K. & Wyatt, S. (2014). Issues in the development of open access to research data. *Prometheus: Critical Studies in Innovation*, *32*(1), 49–66. https://doi.org/10.1080/08109028.2014.956505

Wiley, D. (2014). The Access Compromise and the 5th R. Iterating toward openness. Retrieved December 13, 2015 from http://opencontent.org/blog/archives/3221

Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, *31*(1), 17–29. https://doi.org/10.1016/j.giq.2013.04.003