SITAR, the Geographic Archaeological Information System of Rome: some challenging issues in opening archaeological data

Valeria Boi, Ilaria Jovine, Milena Stacca & Mirella Serlorenzi

Abstract – The Geographic Archaeological Information System of Rome (SITAR) is a project of the Special Superintendency for the archaeological heritage of Rome, a branch office of the Italian Ministry for Cultural Heritage and Activities and for Tourism. The project aims to provide open archaeological heritage data from urban and suburban areas of Rome to the general public (http://sitar.archeoroma.beniculturali.it/). The SITAR webGIS portal enables the online consultation of archaeological data. For every record in the database, the topographic location is provided, along with a descriptive sheet containing general information about the research and scientific information. The information provided corresponds to a “minimum level of knowledge”, which is adequate to allow for an aware re-use of data for academic purposes, conservation and urban planning. The design of such a project must take into account the inherent tension between the public interest of making archaeological data available to the public and the protection of personal data and intellectual property. In reaching its goal, SITAR also has to tackle a wide range of challenging issues, such as data quality and accessibility, long-term preservation and economic sustainability, which will only be possible by establishing a wide cooperation at European level, through a network of institutions concerned with cultural heritage data and services, as now suggested by European trends.

Key words – Archaeology, Open Data, PSI, GIS, archaeological database, authorship, privacy


Schlüsselwörter – Archäologie, Open Data, PSI-Richtlinie, GIS, Bodendenkmälerdatenbank, Urheberrecht, Privatsphäre

Introduction
To understand the complexity of cultural heritage management in Italy, and the differences when compared to other countries’ systems, it is useful to explain the main principles adopted in the Italian cultural context. The most recent law regarding the matter is the Code of Cultural Heritage and Landscape, approved by Legislative Decree 22 January 2004, n. 42. Our attention must be principally focused on the competences conferred to the State and the other regional institutions in managing cultural heritage in the best and most successful way. This decree is based on article 9 of the Italian Constitution, which is one of the most fundamental because it declares that “The Italian Republic promotes the development of culture. It protects the landscape and the historical and artistic heritage of the Nation”; in this way the cultural, social, legal nature of national cultural heritage becomes the unifying principle of the sector. In article 1 of the Code, it is written that “the Republic protects and enhances the cultural heritage (...) The safeguard and enhancement contribute to preserve the memory of the nation and its territory and to promote the development of culture”. So, it is clear that cultural heritage is considered a means of self-identification, able to represent the national identity and, for this reason, is the most important rule within this law. In this way, the importance of cultural heritage as a unifying concept, within a double element made up of cultural and landscape heritage, is highlighted. It is a complex, but uniform combination, established over time, which we inherited from the past and that we have to transmit to future generations.

At the same time, the safeguard action is considered a dynamic attitude, because cultural assets do not need to be only passively conserved, but also used as instruments to promote cultural growth. Heritage, as in the heritage of memory, language and history, plays a civic role and is an integral part of national identity. Preserving the national memory through cultural heritage and referencing the memory of the national community and its territory is crucial, especially in view of the current process of European integration.
Article 1 goes on: “The State, regions, metropolitan cities, provinces and municipalities provide and support the preservation of cultural heritage and its public use and exploitation... The other public bodies ensure the conservation and the public use of their cultural heritage” finally, “the private owners of properties belonging to the cultural heritage are required to preserve them”. These simple sentences mean that there are two kinds of people involved in cultural management – the public and the private - and that there are three main functions to ensure: preservation and public use for the first ones, and conservation for the second. So, the State carries out its duties through the Ministry of Culture, and the Ministry itself operates within the national territory through more than 20 Superintendencies that guarantee the direct connection with local situations (MiBACT 2014). Moreover, in the Italian law, every object found underground belongs to the State and nobody can sell, buy, damage or modify it. So, even if you find an archaeological object in your private garden, you do not own it and it belongs to the State. That is why all excavations in Italy are carried out under the Ministry permit and under the scientific direction of the Superintendency.

In the Italian Law for the Digital Administration “public data” are “data that everyone can know” (Legislative Decree No. 36/2006). More recently, a new meaning has been attached to it and public data are also seen as “data produced by public administrations within their institutional goals”, they are “paid” by the community and, as a consequence, belong to the community itself (http://www.funzionepubblica.gov.it/media/982175/vademecumopendata.pdf). SITAR (Geographic Archaeological Information System of Rome), a project of the Superintendency for Archaeological Heritage of Rome, has the purpose of producing a complete and updated web database of the archaeological heritage of Rome, processing all data coming from archaeological findings. SITAR is an essential tool to monitor the protective actions exercised by the Superintendency in the area, allowing identification of the ongoing investigations, and to see the results of those already completed. Since the genesis of the project, the instrument chosen for data publication has been identified as webGIS, which is considered the best solution to enable editing and consultation with the Superintendence georeferenced database in different locations, but also to share the content. Within the web platform, the topographic data is associated with a summary record that collects context and “history” of the investigation data and the scientific data related to the information we can infer.

SITAR represents information that must be shared with the local authorities, businesses and professionals involved - in various capacities - in the process of decision-making related to land management. The archaeological data fall within the category of spatial data, which is of general interest and, because of their impact on management and planning policies, must be shared with organisations in charge of the management of the territory. The scientific data must also be made available to researchers, because of the contribution they can offer to the progress of knowledge. Sharing data with the citizens, finally, can allow SITAR to become a mediator between information of a specialist nature and civil society, with the hope that the latter becomes more involved in the protection of cultural heritage. In this way, use of the SITAR Project can allow for the improvement of skills, competences and professional proficiency re-qualification by users, thanks to the inclusion in new workflows, procedures and the use of the web platform. This enhancement of human resources seems likely to be the most important consequence of the entire process of implementation, with regard to the added values also in terms of job opportunities, as well as of a deeper awareness of the spread of knowledge.

The presentation of data is one of the key points of reflection about the system; because it was created within the Ministry, the need to identify the datasets that can be public has become a priority. In this paper, we will examine the logic levels provided by the system, with the aim to determine which ones may be considered public data and may be diffused freely on the network (for a full description of SITAR, see Serlorenzi 2013).

Origin of information and privacy protection policies

To locate an archaeological find and to represent it using its geographical coordinates is equivalent to declaring its birth; in this way, it can exist and be unequivocally be recognised. Therefore, the first logical level of SITAR indexing, called “Origine dell’Informazione”, “origin of information”, describes the area of intervention from a cartographic point of view, and it collects all data about the “history” of the archaeological investigation, allowing people to identify the context of data production. In this sense, the information
provided at this level can be considered a sort of “metadata” of the research itself. This section includes:

- START AND END OF RESEARCH
- DISTRICT
- MUNICIPALITY
- AREA
- ADDRESS
- RESEARCH METHOD
- WORK DESCRIPTION
- APPLICANT
- SCIENTIFIC TEAM
- OWNER AUTHORITY
- MANAGING OFFICIAL
- PROPERTY REAL ESTATE
- GEO-REFERENCING CODE
- GEO-REFERENCING METHOD
- KIND OF REPRESENTATION
- ARCHIVE REFERENCES

These data represent the general information about the archaeological research, and may contain personal data about people or companies involved in the conduct of the work: archaeologist, scientific director, etc., as well as the natural or legal person who has requested the authorisation to carry out the work, generally the owner of the property.

Personal data belonging to the person/organisation who requested the authorisation to begin archaeological digging are often present. If it is an individual, the presence of the two fields “applicant” and “address” could make the owner of the land or house in which the investigations are carried out identifiable, in contravention to the principle of the “anonymisation” policy. For this reason, such information, registered for office use only, will not be publicly displayed (Bassi 2011). Despite the need for a maximum flow of information the protection of personal data is vital, and this is a basic factor to increase confidence in the institutions and the main object of attention by the Superintendence: in order to reconcile the right to information with respect for privacy, the SITAR asked the Italian Guarantor Authority for Personal Data about the dissemination of such private information. The answer has ruled out the possibility of spreading personal data in the absence of a law or regulation. The display of information about the identity of those who produced the archaeological data does not represent a violation of their rights but a duty towards those people who have worked in the field and on the preparation of documentation, and this is regarded as a recognition of their work. At the same time, it represents a guarantee for the users, who will be able to quote the source of information whenever they use it.

Archaeological data and intellectual property rights

The second logical level of SITAR database stores the archaeological data; the so called “partizione archeologica”, “archaeological part”, stores the elements of scientific knowledge by distinguishing them on the basis of chronological and functional criteria. This logical level is particularly suitable for data coming from preventive archaeology, which represent the great majority of the records in the database. These data constitute only partial data and fragments of the past city; therefore, scientific data first need to be described, then to be interpreted and connected to each other, in order to reconstruct the historical landscapes and the whole context (Campàna 2011). The archaeological part record sheet includes the minimum set of data which characterise it in a chronological, functional and cultural way; each archaeological part corresponds to something which constitutes a clearly identifiable human activity: the building of a structure, its restoration, its destruction. Basically this logical level leads to the answers to the fundamental questions: What is this? When was it built? In other words, to describe the period of foundation or construction, its functional life, its reoccupations, neglect and decay, discovery and exploitation (Serlorenzi 2013). This section includes:

- RESEARCH METHOD
- OBJECTIVE DEFINITION
- SPECIFIC DEFINITION
- DEFINITION OF INTERPRETATION
- DESCRIPTION
- BUILDING TECHNIQUE
- GEO-REFERENCING CODE
- GEO-REFERENCING METHOD
- KIND OF REPRESENTATION
- MAXIMUM ALTITUDE
- MINIMUM ALTITUDE
- CHRONOLOGY

This is the “essential information” necessary to ensure a proper understanding of the archaeological and topographical context, which allow people to re-use data with a clear awareness. The information represents the results of a shared process of interpretation, and both scientific directors and professional archaeologist/s must be quoted as authors of the information, as mentioned above. In particular, there are two fields which need a de-
talled discussion, as they do not represent a simple description, but an interpretation: the “specific definition” field and the “interpretation definition” field. It is therefore essential that whoever uses this information cite the authors (as for the authorship of archaeological data; see Anichini, Gatticella, Gualandri & Not 2013, and, on the same project, the lawyer’s legal opinion in Ciurcina 2013. See also, in general, Trabucco 2009 and, more recently, Serlorenzi, Jovine, Boi & Stacca 2013).

One aspect of primary importance is related to the quality of published data, because the Superintendence, as an organ of the State, needs to validate and verify the data before its publication. At the same time, this necessary process of data validation must be carried out in a brief time after the completion of the field work, especially in relation to the preventive archaeology, which is directly connected to urban planning. In order to give an efficient contribution to the town planning, archaeological data must be rapidly shared with all the actors involved. Recently, in 2012, dealing with the subject of preventive archaeology, the Ministry suggested that the time to carry out this control must not exceed one year after the end of the work in the archaeological site.

Archaeological unit

So far we analysed the data storage mode of “objective” archaeological information. There is, however, a further logical level, the “archaeological unit” aimed at an interpretational reading of the territory. The unit represents a synthesis in which individual archaeological parts, coming from field surveys carried out at different times and for different purposes, are combined to reconstruct the elements that constitute the settlement system. The archaeological units thus contribute to the identification of historical and topographic sets that have formed the different historical landscapes of the city and its region, and, as we said before, are particularly suitable for preventive archaeology data, which are partial and fragmentary ones.

The information stored in this logical level comes from an interpretation of data and so the author of the record owns the intellectual property, and he/she must always be quoted when the information is reused.

SITAR procedures and security policy

Overcoming the initial weak integration of ICT, technologies and updated methodologies in office workflows, SITAR’s standardised and constantly updated procedures are strongly improving productivity of users in data/documents archiving, preparing, analysing, sharing and dissemination. This procedural approach also enables to progressively reduce the gap between the rapid increase of material archives and the formation of new digital archives and libraries, just now started within the Superintendence, for an effective ubiquitous availability of data and information. Moreover, to supervise the typology of SITAR’s WebGIS accesses, a simple role-based access control system has been created. This method enables the assigning of a specific role to each user associated with different levels to edit the data. The system permission scheme consists of various user roles, each one associated with different permission levels as far as data access and modification is concerned. Thus the minimal level is read-only, with the exclusion of some personal data. Other roles allow data insertion and modification; in this way, the risk of an uncontrolled alteration of data is avoided.

Licenses

The possibility to reuse and re-aggregate the information made available in SITAR must be guaranteed by the choice of an open license, given the eminently public character of this database. Italian law, with a provision of 2012, has made substantial changes to the policies of access and re-use of public data. In May 2014, the „Agenzia per l’Italia digitale”, the public agency charged with the process of digitisation of public sector information in Italy, published the “National Guidelines for the enhancement of public information”, which clarifies some key points of the new law. In particular, they argue for the adoption of licenses by Creative Commons, by virtue of their automatic supranational inter-operability and therefore the possibility of maximum re-use of the data. With the exception of information protected by the “Law regarding the protection of personal data”, we can say that all the other data in the record, i.e. general data, can be disseminated with a CC-0 license, and so fall back into the public domain. This has already been established in the Memorandum of Understanding signed within the European project ARIADNE (http://www.ariadne-infrastructure.eu/). To support the project purposes and to publicise the valuable informative heritage of its repositories, the Superintendence provides the shareable and already available SITAR dataset related to the
first level of information, intended as the georeferencing, methodological, temporal information and metadata, and short scientific references of each archaeological study carried out in the territory of Rome and Fiumicino.

The scientific data, on the other hand, can be distributed under a CC-by license, which only requires attribution, i.e. to cite the source at the time of re-use. The adoption of this type of license would allow for derivative works, both for commercial and non-commercial purposes. In this regard, we believe it is important to reflect on whether or not to properly cite all those who contributed to the production of archaeological data. The “personalisation” of the license in this direction would recognise the work done in various ways, by many professionals, while providing a boost to the advanced sharing of knowledge.

Users can not only consult archaeological records in the database from the web platform, but also carry out the download of information; these are the main functionalities:

**For all users**
- Download of metadata describing the origin of information, archaeological part, archaeological unit.

**For registered users**
- advanced search by location, by address, by unique code of origin of information or archaeological part, site monument, which will be downloadable in different open-formats.
- download of digitised documents about the origin of information, archaeological part, archaeological unit.

**Data validation and update**

SITAR is a database created and managed by the Superintendence for the architectural heritage of Rome and so, as an institutional platform, it has a duty to guarantee the reliability and quality of the published data. The quality of information plays a key role in enabling the re-use of information by professional archaeologists and private companies working in preventive archaeology and in archaeological impact assessment, which are at present the main SITAR users. For them, the most important information needed is the accuracy of positioning, when the data set was last updated, and the accuracy of the chronological and functional identification. In order to ensure an appropriate control of the archive data before publication, at first it is necessary to establish who is responsible for the data accuracy, and so who has the right and duty to “certify” the data, validating it before exposing it to the public. In Italy, at the moment, the scientific direction of the excavation is attributed to the officials of the Superintendencies. In fact, they are in charge of the coordination of work and are responsible for deciding on the size of the excavation, the method of work, and any additions and insights, so it is also their task to verify the quality of the documentation produced. Therefore, according to this principle, it is possible to affirm that the data published by SITAR is available to the public, and that the Ministry has the right to guarantee their public dissemination and re-use. In fact, the data published by the Superintendence also functions to ensure efficient cooperation with organisations charged with planning, in order to avoid the risk that the design of strategic works for the city is affected by inaccurate data. At the same time, both the scientific director and the archaeologists charged with executing the field research have the right to be correctly cited when the information is reused by third parties. Another important aspect with regard to the quality of information provided to the public, is the constant updating of the database, which could be guaranteed by publishing updates to the data at regular intervals.

**Costs, long-term preservation and sustainability**

SITAR was realised with the contribution of internal personnel and contractors, and thanks to funding from the Ministry of Culture. In the first seven years of the project, public funding guaranteed system design and development, as well as its implementation. In particular, up to now the activities carried out focused on:

- the design of the logic levels of the system;
- the realisation of the data entry platform;
- the realisation of the web database and webgis interface;
- the digitisation of a corpus of documents, previously dispersed in local archives and now available in digital format to all accredited users;
- the implementation of the web database, freely accessible online;
- the constant technological upgrading;

The Superintendence, as the official holder of the archives on behalf of the community, is obliged for its institutional role to ensure the conservation and the public access of this huge repository of
knowledge. For this reason, over the years a series of agreements with other public institutions has been signed, including the CNR (National Research Council) and the Consortium GARR (Group for Research Networks Harmonization), which provide their cutting-edge technology in the management of the Research network, and guarantee both the preservation of digital resources and the online accessibility to the database.

Those themes inevitably bring with them the issue of the costs which must be sustained to ensure the efficiency of the platform, its technological upgrade and the constant update of new data. As said, up to now the costs have been fully supported by public funding, but in recent years public spending has been consistently cut, particularly in the field of cultural heritage, and so it becomes necessary to think - in a concrete way - how the project can be sustainable in the future. In fact, the expected rapid growth of SITAR’s user community will lead to an increasing interest in exploitation of the SITAR data bank, applications and web services which have to be faced with a strong economic strategy. It is necessary to programme not only funding for the development of the system, but also the cost of long-term data preservation.

In the event of substantial reductions in public funding, we could imagine that the costs of new data implementation could be reduced by entrusting the data entry to the private companies or professionals who carry out the archaeological investigation; on the other hand, costs could partially be incurred by the commissioning bodies. The fact that also the study and publication costs have to be previewed since the project is sustained by the commissioning bodies is envisaged by the Circular 10/2012 of the Directorate General for the Antiquities; private participation in the financing of archaeological valorisation is also foreseen by Art Bonus (L. 106/2014).

The reflection on the private sector leads us to consider the role of the users: no doubt, in fact, private companies and professionals also represent one of the major users of the system. While it is clear that the use for purposes of promotion, study or research is always free of charge, in some cases in Italy it was thought that data released for commercial use, e.g. preventive archaeology, could be subject to fees, to consumption or one-time charges.

This is, for example, the solution adopted by the Italian Cadastre, with a law of 2012, after a period of free access to the services of data consultation and downloading, they now request the signing of an agreement, which is free of charge for public bodies, while individuals are subject to the payment of an annual fee, in order to contribute funding towards the running costs of the platform (see also http://www.linkedopendata.it/implicazioni-economiche-open-data).

An opposite approach is the one adopted by the Spanish Cadastre: it is the biggest data provider of the Spanish Spatial Data Infrastructure. Citizens have open, free of charge access to all data except those subject to Data Protection law. Since 2003, when the services offered on the digital platform started their activity, there has been an extraordinary increase in use of data, that has been stimulated by a progressive increase in the services offered: WFS, WMS, KML. The internet services have been a phenomenal success, with more than 8 million accesses a day. Many reasons influenced the cadastre to offer the data for free; in fact, they believe it is a contradiction to seek economic compensation for access to cadastral information when their public task is to collect the data. Furthermore, they estimated that the cost of bureaucracy in determining the pricing, management and invoicing of the system would be greater than the economic benefit derived from exploiting the data. Moreover, the increase in transparency of the system has resulted in many users having reported errors or potential improvements of data, thus improving the quality of the data with the users’ feedback.

For all these reasons, the case of the Spanish cadastre seems to be preferable, also with regard to profit margins offered by this policy and, in contrast, to the positive effects on the economy of a wider openness of these data, both in terms of economic sustainability and of the potential of correcting and improving data quality thanks to the users’ feedback.

Anyway, before making use of solutions resulting in costs for users, it should be far preferable to explore the possibilities of the realisation of a network of institutions concerned with cultural heritage data and services, which could contribute to the costs, which is one of the proposed solutions in the European debate. This issue, in fact, has been recently addressed within the European project Advanced Research Infrastructure for Archaeological Dataset Networking in Europe, ARIADNE (http://www.ariadne-infrastructure.eu/) and in other initiatives in Europe, such as the IANUS project (http://www.ianus-fdz.de/), the Digital Preservation Coalition (http://www.dpconline.org/) and others.

We truly believe that the sharing of issues and experiences is the best way to improve data qua-
lity and service reliability, in order to foster free sharing of knowledge and to raise awareness and involve all the stakeholders.

References


About the authors

VALERIA BOI: archaeologist (2008), PhD candidate in Methodology of Archeological Research at the University of Sassari, with a research thesis focused on the evaluation of archaeological potential in urban areas. Freelance archaeologist, expert in GIS and Open Data. Since 2009 she has collaborated with the Heritage Office of Rome on the project SITAR (Geographic Archaeological Information System of Rome), for which she is addressing the issue of open data.

ILARIA JOVINE: graduated in Classical Literatures (2003) at the University of Rome “Sapienza” with a thesis in Greek epigraphy, postgraduated with a thesis in Latin Palaeography (2011). Since 2000 she has been working for the Ministry of Culture at the Heritage Office of Rome and since 2010 she has been collaborating on the protection and enhancement of the first and eighteenth Municipalities of Rome. She is member of the project SITAR (Geographic Archaeological Information System of Rome), cooperating in the dissemination and implementation of the system and addressing the issue of open data.

MIRELLA SERLORENNZI: is an archaeologist officer of the Ministry of Culture. She has been involved in the main archaeological urban excavations in Rome, like the ones on the Palatine Hill and in the Crypta Balbi. From 2000 to 2006 she worked as director archaeologist at the Heritage Office of Ostia; since 2006 she has been working as director and coordinator archaeologist at the Heritage Office of Rome; since 2013 she has been the scientific director of the National Museum of the Early Middle Ages. From 2007 onwards she has been the managing director of the project SITAR (Geographic Archaeological Information System of Rome), a high technological innovation project.

MILENA STACCA: is an expert in GIS and Open Data. Since 2009 she has collaborated with the Heritage Office of Rome on the project SITAR (Geographic Archaeological Information System of Rome), for which she is addressing the issue of open data.

Corresponding author
Milena Stacca
Servizio SITAR della SSBAR
Sede di Palazzo Massimo
milena.stacca@gmail.com