

Toscano, M. (2018). Where the researcher cannot get: open platforms to collaborate with citizens on cultural heritage research data. En E. Romero Frías y L. Bocanegra Barbecho (Eds.), *Ciencias Sociales y Humanidades Digitales Aplicadas* (pp. 538-561). Granada, España: Universidad de Granada [ISBN: 978-84-338-6318-8]. New York, USA: Downhill Publishing [ISBN: 978-0-9897361-7-6].

Where the researcher cannot get: open platforms to collaborate with citizens on cultural heritage research data

MAURIZIO TOSCANO

Universidad de Granada

maurizio.toscano@gmail.com

ABSTRACT

Over the years, many research projects and research groups have accumulated large collections of raw data, nowadays they are normally compiled and stored in local archives, more often on personal computers and sometimes on Intranet servers. Local data represents a missed opportunity for research projects, in particular in the cultural heritage field. The vast majority of these datasets, even those that are already available in a digital format, remain hidden, underutilized and incomplete because they are linked to the effort and limited resources of a single researcher or a specific research group. This paper proposes a shift from Local Data to Open Platforms, from data stored locally and accessible only to a restricted number of researchers to remotely accessible, editable and sharable datasets. Online platforms perfectly fit the needs of

a research group that is experiencing the limits of first-hand data collection and verification or that intends to extend the potential area of source collection. To demonstrate the effectiveness of this proposal, three different case studies will be explored. They have all been recently developed as the result of collaborations between the author and university research groups in the humanities. Despite sharing a common background in cultural heritage, they are very different with respect to the object of study. The core of this paper is the discussion, where we analyse the lessons learnt from applying this approach, highlighting its benefits and the intrinsic issues of the process.

Keywords: Web Information Systems, collaborative platforms, co-creation, digital humanities, Science with and for Society

RESUMEN

Muchos proyectos y grupos de investigación han llegado a acumular a lo largo de los años grandes cantidades de datos en bruto, por regla general acumulados y almacenados localmente, a veces en servidores locales o, más a menudo, en ordenadores personales. Los datos locales representan una oportunidad perdida para los proyectos de investigación, en particular en el campo del patrimonio cultural. La gran mayoría de estos conjuntos de datos, incluso aquellos que ya están en formato digital, permanecen ocultos, subutilizados e incompletos porque están vinculados al esfuerzo y los recursos limitados de un solo investigador o grupo de investigación específico. Esta contribución propone de pasar desde Datos Locales a Plataformas Abiertas, desde datos accesibles solo a un número restringido de investigadores a conjuntos de datos accesibles, editables y compartibles a distancia. Las plataformas en línea se ajustan perfectamente a las necesidades de un grupo de investigación que experimenta los límites de la recopilación y verificación directa de los datos o tiene la intención de ampliar el área potencial de la recopilación de fuentes. Para concretar esta

propuesta, se describirán tres diferentes casos de estudio. Todos han sido desarrollados recientemente y son el resultado de la colaboración del autor con diferentes grupos de investigación en humanidades. Incluso si todos están relacionados con el patrimonio cultural, son muy diferentes con respecto al objeto de estudio. El núcleo principal de este documento reside en la discusión, donde se analizan las lecciones aprendidas al aplicar este enfoque, destacando sus beneficios y los problemas intrínsecos en el proceso.

Palabras Clave: Sistemas de Información Web, plataformas colaborativas, co-creación, humanidades digitales, Ciencia con y para la Sociedad

1. BACKGROUND

This paper is the result of the author's professional experience, who over the past few years worked with several researchers and research groups and, as a consequence of these collaborations, built several online collaborative platforms, created to collect unpublished research data or collaborate with other research groups or the civil society over large collections of research data. The author has a background in humanities and cultural heritage, so the case studies described in this paper have been derived from these fields, as well as the proposal and the discussion apply better to projects in human sciences, even if not exclusively. This paper is part of the work that the author is developing within the H2020 European Project REACH - RE-designing Access to Cultural Heritage for a wider participation in preservation, (re-)use and management of European culture¹.

¹ REACH project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement No. 769827, https://cordis.europa.eu/project/rcn/212215_es.html

2. WHY DOES LOCAL DATA REPRESENT A MISSED OPPORTUNITY?

Many research projects and research groups have compiled large collections of raw data over the years, nowadays predominantly in the form of electronic files or databases. Those archives are normally accumulated and stored in local archives, more often on personal computers and sometimes on Intranet servers. This generates several problems, unknown to analogue records, in the form data integrity, backups, and long-term consistency amongst other things. The aim of this paper is not to discuss these issues, but instead to focus on the missed opportunities that local data represents for research projects, in particular in the field of cultural heritage.

The vast majority of these datasets, even those that are already available in a digital format, remain hidden, underused and incomplete because they are linked to the effort and limited resources of a single researcher or a specific research group.

The publication of raw research data online –Open Data– is becoming increasingly popular², however somewhat slowly within the humanities, and is already a requirement of many European and national calls. This is a very positive initiative, focused on the use, re-use and verification of research data, but it only serves as a partial solution to the local data issue described above, as it does not contribute directly to the improvement, verification and enlargement of those corpora. Data is shared raw, without the necessary tools to allow direct interaction with the general public or other research groups.

² The state of Open Research Data in Europe is described by the Open Science Monitor, at <https://ec.europa.eu/research/openscience/index.cfm?pg=home§ion=monitor>

3. WHY DO OPEN PLATFORMS REPRESENT A VALID INSTRUMENT OF RESEARCH?

Over the last few decades, Digital Humanities have taken advantage of the widespread increase in the use of the Internet and have developed tools that make it possible to collaborate electronically with citizens (Fresa and Justrell, 2015) and other research groups. In order to deal with the above-mentioned issue, a possible approach is to move from Local Data to Open Platforms, from data stored locally and accessible only to a restricted number of researchers into content available online, where the datasets are accessible, editable and shareable.

Online platforms perfectly fit the needs of a research group that is experiencing the limits of first-hand data collection and verification or that intends to extend the potential area of source collection. They provide several advantages over other instruments of data management. They are collaborative for their own nature, allowing concurrent users working simultaneously, tracking individual activity and allowing version control and restore. When required, they can be managed and customized remotely. They are compatible with widely accepted and well-maintained Web standards, to create attractive and easy-to-use graphic interfaces. In this sense, they belong to an environment, the Web, that the vast majority of potential contributors use on a daily basis for work and during their spare time. The interfaces have a familiar appearance and therefore a complicated manual is not required when interacting with them also to non-specialists. More significantly, they offer built-in instruments to work dynamically on a dataset, either inputting new data, commenting or modifying existing data and uploading any type of archive.

According to the needs of the research project, the platform can be configured to accept real-time changes or through required moderation and approval. We will return to this subject in the discussion, but for now, we can say that allowing real-time changes

produces better engagement since by seeing their input published straightaway, contributors are encouraged to keep contributing; on the other hand, incorporating changes only after administrators' approval improves data quality control.

To effectively involve citizens in scientific research, online live collaboration is a key feature. Theoretically, a standalone PDF form can be sent individually or shared in large social network groups to collect new data, but this practice makes it harder to keep track of new contributions, requires extra management work to consolidate new and existing records and can affect data integrity and version control; additionally, online collaboration offers other benefits, such as creating a sense of community among contributors and foster data sharing. Finally, and most significantly, these collaborative tools can be custom made: not all platforms are made equal, in the same way that not all research projects are made equal. They can be tailored to fit the dataset that is going to be published and to create a separate, simple custom webform for data collection.

When structuring the back-end database, it is crucial to keep the aim of the research project in mind. A clear distinction should be made and different steps should be taken if producing a collection of data is the ultimate goal of the project as opposed to just a necessary step towards analysis and interpretation. In this last case, which is quite common, data collection standards³ can be taken in limited consideration, and the primary focus should be the use of the information within the research context.

³ For humanities and social science, main metadata standards are: Text Encoding Initiative (<http://www.tei-c.org/index.xml>), the Visual Resources Association Core (<http://core.vraweb.org>), Dublin Core (<http://dublincore.org/documents/dc-mi-terms>) and the Data Documentation Initiative (DDI) (<http://www.ddialliance.org>).

4. CASE STUDIES

To consolidate this proposal to shift from local data to open platforms and briefly highlight the benefits of this approach and the intrinsic issues of the process, we are going to present three different case studies. They are all recently developed and are the result of the collaboration of the author with different research groups within the humanities. Even if they share a common background in cultural heritage, they are very different with respect to the object of study, one being focused on the shared artistic heritage among Andalucía and Latin America⁴, another on traditional medieval irrigation systems in the provinces of Granada and Almería⁵ and the last on eighteenth century houses in rural villages in the Lecrín Valley (Granada)⁶. For the purpose of this paper, what is important is that they are similar from a technical and methodological point of view, sharing the same issue and choosing a similar path in order to overcome it.

4.1. IBERO-AMERICAN ARTISTIC HERITAGE

For the platform *Colabora con Andalucía y América*⁷ (Toscano, 2018), the local data was represented by a large catalogue of artworks created by Andalusian artists throughout twelve Latin American countries, dating from the Renaissance to the present. Approximately 2,600 records have been collected since

⁴ Research project *MUTIS. Patrimonio artístico y relaciones culturales entre Andalucía y América del Sur*, Ref. HAR2014-57354-P, principal investigator professor Rafael López-Guzmán.

⁵ Research project *Acequias. Proyecto de ciencia participativa basado en el estudio de los valores culturales y ambientales de los sistemas históricos de regadío de Granada y Almería*, Ref. FCT-15-9814, principal investigator professor José María Martín Civantos.

⁶ Research project *Familias, desigualdad social y cambio generacional en la España centro-meridional, 1700-1900*, Ref. HAR2013-48901-C6-6-R, principal investigator professor Francisco García González.

⁷ <http://colabora.andaluciayamerica.com>.

2008, with detailed information about the artist, the style, the iconography, the chronology and the localization, with the latter organised from country down to the building level. The awareness of the breadth of the subject and of the effort required to compile the files filled to date, has worked as an incentive to launch this initiative. The dataset has been published as individual records, presented in a catalogue view, with an advanced faceted search engine. Registered users have two options to collaborate: improving existing records or report a new artwork. Many catalogue records do not have illustrations or they are not high quality, because it has not always been possible to directly access the artworks to take pictures; in those cases, the online platform is an opportunity to get much better illustrations taken directly onsite. On the other hand, the form is very simple and focuses on the most valuable data for the researcher, i.e. a photo of the artwork and its location, even if the possibility of providing more detailed information remains, such as the author, the typology and a short description. In order to facilitate filling in the form directly in front of the artwork, kept in a museum, a gallery, a church, a private collection or in a public space, full compatibility with mobile platforms was ensured, as well as the ability to directly upload a picture from a smartphone or tablet. Finally, to speed up possible contributions, registration has been simplified through the availability of login via social networks (Facebook and Google), basically omitting the user verification step via email. The aim was to create a means to broadening and improving the catalogue of artworks at a global level. Taking advantage of the immense potential of the internet, it is now possible to gather contributions directly from the study areas or from other research groups that are working on related topics.

Catalogue

Search artwork

Initial date


Use just the year (i.e., 1500)

Final date

Search found 39 items


Virgen de la Bella

Author: Anónimo
Date: 1701 to 1800
Typology: Esculturas de bulto redondo
Style: Barroco
Location: Iglesia de San Bartolomé Arani Cochabamba Bolivia
Elaboración: P07-HUM-03089




San Juan de Dios en contemplación

Author: Pérez Holguín, Melchor
Date: 1701 to 1800
Typology: Pinturas de caballete
Style: Barroco
Location: Universidad Mayor, Real y Pontificia de San Francisco Xavier Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089




San Juan de Dios con el Niño Jesús al hombro

Author: Anónimo
Date: 1701 to 1800
Typology: Pinturas de caballete
Style: Barroco
Location: Museo Universitario de Charcas Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089




San Juan de Dios penitente

Author: Anónimo
Date: 1701 to 1800
Typology: Pinturas de caballete
Style: Barroco
Location: Universidad Mayor, Real y Pontificia de San Francisco Xavier Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089




San Juan de Dios con el Niño Jesús al hombro

Author: Anónimo
Date: 1701 to 1800
Typology: Pinturas de caballete
Style: Barroco
Location: Museo Universitario de Charcas Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089




San Juan Bautista

Author: Cueva, Gaspar de la
Date: 1633 to 1638
Typology: Esculturas de bulto redondo
Style: Barroco
Location: Iglesia de San Miguel Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089




San Francisco Solano

Author: Anónimo
Date: 1601 to 1700
Typology: Esculturas de bulto redondo
Style: Barroco
Location: Ermita de San Francisco Solano Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



San Juan de Dios con el Niño Jesús

Author: Anónimo
Date: 1667 to 1701
Typology: Pinturas de caballete
Style: Barroco
Location: Museo Convento de la Recoleta Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



Search artwork

Initial date

Use just the year (i.e., 1500)

Final date

Apply

Reset

Location

- (-) Bolivia
- Potosí (19)
- Chuquisaca (10)
- La Paz (8)
- Cochabamba (1)
- Oruro (1)

Style

- (-) Barroco

Typology

- Pinturas de caballete (54)
- Esculturas de bulto redondo (12)
- Pinturas (2)
- Pinturas sobre tabla (1)

Author

- Anónimo (17)
- Cueva, Gaspar de la (8)
- Anónimo potosino (4)
- Pérez Holguín, Melchor (2)
- Anónimo popular (1)
- Anónimo sevillano (1)
- Berrio, Gaspar Miquel de (1)
- Cordoba, Ignacio (1)
- Herrera y Velarde, Francisco de (1)
- Martínez Montañés, Juan (1)
- Pérez Holguín, Melchor (1)
- Zamudio, Bernabé de (1)

Iconography

- San Juan de Dios (16)
- San Francisco Solano (4)
- Divina Pastora (3)
- Jesús atado a la columna (3)
- San Fernando (3)
- Crucifixión de Jesús (2)
- Virgen de la Asístiga (2)
- Crucificado (1)
- Dios padre (1)
- Immaculada Concepción (1)
- San Juan Bautista (1)
- Virgen de la Bella (1)

Figure 1: Colabora con Andalucía y América online platform. Catalogue of 2,600 artworks created by Andalusian artists in Latin America, with faceted search filters in the sidebar. Source: <https://colabora.andaluciayamerica.com/>

546

4.2. HISTORICAL IRRIGATION COMMUNITIES IN SOUTHERN SPAIN

In the project *Regadíos históricos*⁸, core local data has been obtained from a combination of fieldwork surveys and interviews with local irrigation communities. In the field, researchers recorded irrigation channels and cultivated plots with handheld GPS and a portable geodatabase, while also carrying out interviews with local farmers about traditional water management practices. Additionally, to enlarge and enrich the dataset, two other strategies have been carried out: a telephone campaign contacting all irrigation communities in the provinces of Granada and Almería and a desk-based cartographic survey, based on the *Inventory and Characterization of the Irrigation of Andalusia. Mediterranean and Atlantic Districts* (Inventario, 2008), modified according to well-known historical transformations and more recent modernizations that affected those systems. The resulting dataset that is now available online covers more than 200,000 hectares of irrigated land and counts more than 24,000 km of irrigation ditches, managed by 550 historical irrigation communities, for a total of 1,424 individual records in the database. Considering the cartographic nature of these data, the most appropriate way to present them to the public was using a spatial data infrastructure, more commonly known as a map.

This product represents an important result by itself, but in order to create a true collaborative map, registered users now have the option to include new records and to comment on existing ones. To facilitate collaboration, a dynamic online form has been setup, divided into two sections: a series of text fields to describe the irrigation community, with an additional field to collect files (documents, images, spreadsheets etc.), and an editable map, where the user is invited to draw on the map the cultivated plots,

⁸ <http://regadiohistorico.es>

as polygons, the path of the irrigation channels, as polylines, along with springs and water sources, as points. A short introductory video illustrates the whole process. The concept is similar to many other online collaborative projects that collect user generated geo-data, of which the most widely known is OpenStreetMap⁹, a worldwide collaborative effort to create a free editable map of the world. Registration has been made very easy by combining all the stages into a one-step online web form process. The system requests the user email, automatically generates an account, associates the submitted information with it, logs the user in, and sends the username and password via email. In this way, the platform dramatically decreases the initial barrier of user registration and verification, simplifying the whole process of contributing new data.

⁹ <https://www.openstreetmap.org>

Mapa colaborativo de regadíos históricos de Granada y Almería

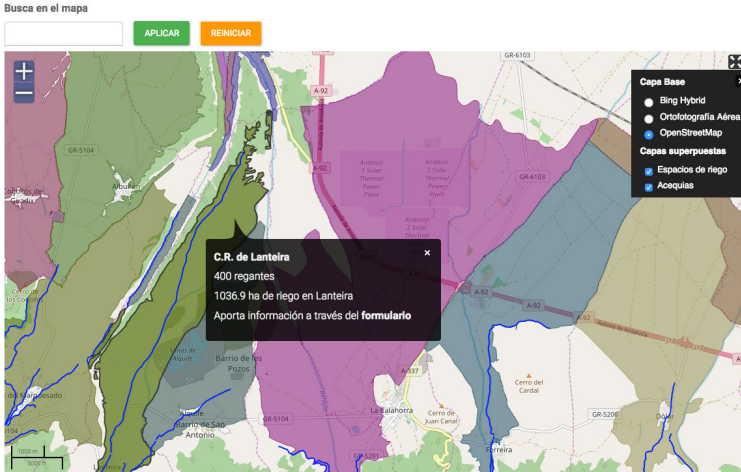


Figure 2: Regadíos históricos online platform. The collaborative map showing the example of the irrigation community of Lanteira (Granada), its cultivated area and main irrigation channel (acequia madre). Source: <http://regadiohistorico.es/>

Crear Ficha de riego

Nombre de la comunidad de Regantes *


Provincia * Nombre de la acequia principal

Granada

Municipio * Número de regantes

Lanteira

LOCALIZA EL MUNICIPIO PARA EMPEZAR A DIBUJAR



Utiliza los iconos arriba a la derecha según lo que quieras dibujar: área de riego o acequia. Usa el ratón para desplazarte y la barra de zoom para acercarte.

Hectáreas de riego

Riego eventual

- Ninguno -

Archivos adjuntos

Añadir archivo nuevo

No file selected.

Los archivos deben ser menores que 25 MB.
Tipos de archivo permitidos: jpg png gif tiff pdf txt doc docx xls xlsx

Contacto comunidad

Persona de referencia (presidente o secretario)

Teléfono

Información adicional

LEER TÉRMINOS Y CONDICIONES

Figure 3: Regadíos históricos online platform. Webform to send new data, with fields to collect structured information, send archives and to draw cultivated plots and irrigation channels on a dynamic map. Source: <http://regadiohistorico.es/>

4.3. EIGHTEENTH CENTURY RURAL HOUSES

In the case of *Las casas del Valle del Lecrín (S. XVIII)*¹⁰, local data derived from a long project carried out independently by the historian Margarita M. Birriel Salcedo, on the sections of the Cadastre of the Ensenada Marquis kept in the historical provincial archive of Granada. The research is focused on rural houses, the information is transcribed directly from the XVIII century manuscripts and registered, until then, using handwritten notes on notebooks organised with colour coded stickers. In this third case study, original data was not only local, but also stored in analogue format, presenting an additional exciting challenge that unfortunately we cannot expand on here, as it is outside the scope of this paper. The database structure has been based on two content types, houses and homeowners, connected by a many-to-many relationship. Recorded data basically concerns the houses, their dependencies, dimensions, locations, neighbouring, owner(s) and census; for the homeowners, their neighbourhood and properties. Primarily, the platform has been designed as an individual research instrument, to improve the recording process and above all the historical analysis, offering search capabilities, structured information and automatic quantifications. In addition, having used an online database from the beginning, instead of a local IT solution, has drastically simplified the process of opening the dataset to the public. Now, anonymous users can browse all published data, which is divided into two main views with several combinable search filters, and four galleries, with sketch plans from the cadastre, hypothetical reconstructions of the houses, photos of the villages, houses, annexes and architectural elements. Registration is required to comment on existing records and upload files to the platform: the two means of collaborating available at the moment. Furthermore, to maintain the original objective of the platform - a working tool

¹⁰ <http://casaslecrin.es>

for the researcher to store and retrieve research data - we have introduced a “Save as draft” option for ongoing records that should not yet be visible to standard users.

Las casas del Valle del Lecrín (S. XVIII)
 Catálogo del Marqués de la Ensenada

Búsqueda propietarios | Búsqueda casas | Glosario propietarios | Glosario casas | Galerías | Contacta | Acceder

Búsqueda avanzada casas

Tu búsqueda ha encontrado 604 resultados

Tipología: 1ª dependencia, 2ª dependencia, 3ª dependencia, 4ª dependencia
 Superficie (en metros): Lugar, Arrendada, Censo, Fecha documento

Inserir valores mín y máx

Casa	Lugar	Barrio	Propietarios	Dependencias	Superficie	Arrendada	Alquiler	Catastro	Total	Censo	Fecha documento
Casa en Las Abuñuelas Barro Alto	Las Abuñuelas	Barro alto	1. PALLA NAWA/RO Antonio de	1. - Planta baja 2. - Planta alta	44.02 m.	No			150 metros y 0 mrs y 33.00 mrs		1752 población, otros
Casa de cambio en Las Abuñuelas Pago del Jaral	Las Abuñuelas	Pago del Jaral	1. JIMÉNEZ Manuel	1. - Planta baja	41.92 m.	No			730 reales y 8,00 mrs		1752 habicón, población, otros
Casa en Las Abuñuelas Barro Alto	Las Abuñuelas	Barro alto	1. JIMÉNEZ Manuel		44.02 m.	No			730 reales y 8,00 mrs		1752 habicón, población, otros

Figure 4: Las casas del Valle del Lecrín (S. XVIII) online platform. Table of recorded houses, with combinable filters at the top. Source: <http://casaslecrin.es/>

5. DISCUSSION

Once the issue has been presented, a possible solution proposed and illustrated using several ongoing initiatives, we are now able to discuss some of the lessons learnt from applying this approach.

Considering that all collaboration will be voluntary, it is wise to offer something in exchange. Potential contributors are likely to be people interested in the subject, so the availability of good quality, well-structured open research data will definitely be a valuable asset for the platform visitors and will encourage their engagement. The existing dataset will generate ideas on where and how it can be improved as well as display what kind of new data is useful to search for: users are more likely to contribute if they can preview where their contributions will end up. Additionally, once

published, the data will collaborate to increase the visibility of the platform in organic search engine results.

Users behaviour on the platform is unpredictable, so carefully plan the permissions given to registered users, simplify the workflow process to contribute and offer key instructions and help texts, in short chunks written in a font that is easily readable.

When registration is required, as we recommend, sign-up and log-in steps should be as straightforward as possible, to reduce the initial barrier and increase participation¹¹. One possible strategy is the use of social login (i.e. Facebook, Twitter, Google, LinkedIn), basically skipping the user verification step via email. This approach has several advantages: it reduces spam users, as it requires a valid social media profile, and it can automatically provide additional user information, including a profile photo, place of residence, age, and so on. The downsides are that, since the user verification process is outsourced, additional work may be required to ensure it remains compatible with future changes in the backend API, and if the users later stop using the social network account used to register, they will no longer have access to the website. Additionally, if the project plans to work with schools, directly involving young students in data collection, the use of social login can cause concerns regarding privacy and should therefore be used with caution. Social login has clear value for marketing websites (Goings and Abel, 2013), but for the purposes of an open platform of research data, a better solution comes by integrating the registration process within the form for submitting new data. This way, the web form can be presented directly to the end user, who can begin completing it straightaway, with all the registration and verification steps completed

¹¹ According to research compiled by WebHostingBuzz in 2013, 86% of users report being bothered by having to create new accounts on websites. Retrieved from <http://www.webhostingbuzz.com/blog/2013/03/21/whos-sharing-what/> [12/02/2018].

upon submission. In this scenario, in order to keep a trustworthy community of users and receive good quality contributions, a mechanism should be put in place to control and prevent spam, as spamming is inevitable for all public platforms that acquire a certain amount of visibility.

Ownership of contributions and Intellectual Property Rights (IPR) issues should be addressed outlining the terms and conditions of participation in the project. Basic principles are that authorship of individual contributions should be acknowledged and recorded, and the distribution and reuse of the research data collected should always be guaranteed. Some citizen science projects are more likely to provoke IPR and copyright issues, and those should not be underestimated.

As demonstrated for major collaborative websites, including Wikipedia, Yahoo! Groups, YouTube, etc., also online open platforms for research are affected by the 90-9-1 principle, or 1% rule (Nielsen, 2006). It basically suggests that 90% of users are just browsers who never contribute, they just read, 9% of users contribute a little, and 1% of users account for almost all the activity. The ratios might not be accurate, but the general structure still applies and we were able to confirm this behaviour first-hand within the Historic Graves project¹² (Bocanegra Barbecho, Toscano and Delgado Anés, 2017: 329), a collaborative effort to geotag and transcribe all Irish memorials. Keeping this benchmark in mind is useful in order to manage expectations, to estimate the number of passive users needed in order to start receiving a certain amount of contributions, and to implement special features that facilitates multiple contributions by the most engaged users (the top 1%).

Although the platform can function independently, being a readily available resource on the internet, a good practice is to plan

¹² <http://historicgraves.com>, developed by John Tierney and the author of this paper at Eachtra Archaeological Projects Ltd.

and carry out *ad-hoc* campaigns, local events and training days to advertise the initiative and encourage contributions through engagement with the public. Instead of trying to involve individuals on a personal basis, we suggest identifying and engaging with existing communities of citizens or reference collaborators, linked in some way to the research theme, who can adopt the role of local mediators. The use of online and traditional media to publish press releases about the initiative is also very effective, especially to create peaks of audience, from which long-term users can be acquired. These outreach activities are extremely important, especially at the beginning, as then search engine ranking and word-of-mouth can compensate for the lack of targeted distribution.

In this sense, human sciences research subjects are very well suited to attract the interest of civil society. Local communities sometimes are privileged potential contributors of the initiative, and the research topic often relates to a common cultural heritage, that is sometimes in danger of disappearing. Such is the case, for example, of the project about traditional irrigation channels in Andalucía, where the local communities featured in the dataset are a living research subject, strongly linked with the contemporary process of transformation and modernization in agriculture; or the above-mentioned Irish project where rural communities take care of their own local graveyard, geolocating and transcribing individual headstones (Toscano and Tierney, 2016); or, again, the case of Spanish associations in Argentina, that have had a strong role in the transnational online collaborative project *e-xiliados*¹³, focused on retrieving unpublished historical sources about the Spanish republican exile (Bocanegra Barbecho and Toscano, 2015). In cases like these, creating an open collaborative platform is also a way to raise awareness among citizens about the historical and cultural values related to that heritage, and share those values

¹³ <http://exiliadosrepublicanos.info>.

with communities and individuals living further away. On the other hand, establishing collaborations with others research groups can be more challenging, as that is not as common in the humanities as among researchers in natural science; the use of such platforms can stimulate building cooperation and data sharing, encouraging also common research activities.

People who will register and contribute to these initiatives can be greatly varied, from researchers specialised in the object of study, to passionate people related in some way to those places, objects or events, down to ordinary citizens, that come across the initiative and are willing to contribute by a spirit of networking, co-creation and participation. Wherever users come from, the platform should protect data integrity and security, being as open as possible yet as secure as necessary. Anonymous users should be able to browse all content; registered contributors should be allowed to publish content in real-time, comment on existing records or submit new data through moderation and approval. The latter approach is more common in research, with the intention of keeping a tighter control over data published online, but a warning should be made that it can affect user engagement and reduce the number of contributions. In every case, it is recommended to implement a versioning system, to keep track of all changes, their authors and timestamps. Being able to restore previous versions of a record and compare the changes between different versions is a much more powerful feature than a simple backup mechanism, which should be in place regardless as a means of protection in order to cover full restores in the event of server outages, data corruption or security breaches. Finally, an automatic notification system should be put in place as well, to generate communications among the three subjects involved: the user, the platform and the research team. New users, new data, new comments, and new versions of existing records, the availability of security patches: every important event should be notified so that, if needed, appropriate

actions can be taken. In our opinion, real-time publishing and editing, combined with a strong versioning mechanism and a solid notification system appears to be the best combination for a collaborative platform, to find a balance between a seamless and engaging experience to access and contribute to the dataset and the control over the platform security and data integrity in the medium-long term.

If the participatory project succeeds, a part-time platform administrator, ideally originating from the research team itself, should also be taken into account, to keep the system running, verify data sent, check for spam and notifications and support users when needed. The commitment of the research group in terms of use and maintenance of the platform in the medium term is important to build trust amongst users, but is also legitimate, at some stage, to consider closing or significantly modifying the platform according to the needs of the research group. In this case, it is highly recommendable to move the whole dataset to an online repository¹⁴, as open data that will support the research findings and can be reused, with appropriate attribution, in other research projects.

¹⁴ In social science and humanities, some of the most common data repositories are Zenodo (zenodo.org), Harvard Dataverse (dataverse.harvard.edu), Archaeology Data Service (archaeologydataservice.ac.uk), openICPSR (openicpsr.org), figshare (figshare.com) and Mendeley Data (data.mendeley.com). The Registry of Research Data Repositories (re3data.org) is an Open Science tool that offers an overview of existing international repositories for research data.

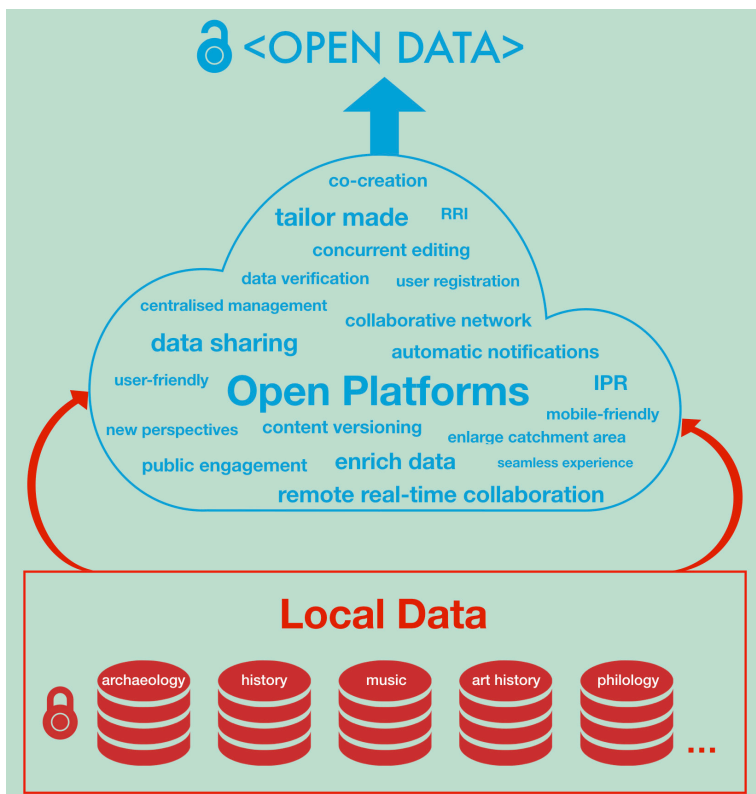


Figure 5: Diagram showing the ideal path from Local Data to Open Data, with the intermediate step of collaborative Open Platforms, and their main benefits. Source: compiled by the author

6. CONCLUSION

In conclusion, we can say that online collaborative platforms should be included in the toolkit of today's cultural heritage research, as a valuable option to establish new collaborations, to gather new data and to enrich the quantity and quality of existing datasets. There are several keys to success, that we can briefly summarize: develop a user-friendly and feature-rich platform to publish online the existing dataset; choose a modern and clean

mobile-friendly layout; implement data integrity mechanisms, automatic notifications and content versioning to keep track of contributions; carry out engaging campaigns, targeted directly at potential contributors, interest groups and local communities; commitment from the research group to disseminate, use and maintain the platform, as well as revise and use, if relevant, the contributions received.

Embracing the Open research methodology has some clear benefits. Moving from local data to open platforms allows researchers to deepen their investigations and expand the collaboration network to other groups, as far as it would have otherwise been impossible, due to logistical, time and resource limits. In addition, this approach enriches the corpus of data with new perspectives and forces the fine-tuning of techniques for verifying information. The ability to actively contribute to the research process brings science closer to civil society and opens new methods of collaboration, helping to align research and development results with the needs and expectations of citizens.

REFERENCIAS BIBLIOGRÁFICAS

- Bocanegra Barbecho, L., Toscano, M., Delgado Anés, L. (2017). *Co-creación, participación y redes sociales para hacer historia. Ciencia con y para la sociedad*, in *Historia y Comunicación Social*, Vol. 22-2, Universidad Complutense de Madrid, 2017, ISSN: 1137-0734 [<http://hdl.handle.net/10481/48545>]
- Bocanegra Barbecho, L., Toscano, M. (2015). *El exilio republicano español: Estudio y recuperación de la memoria a través de la web 2.0. Nuevo enfoque metodológico con el proyecto e-xiliad@s*, en *Migraciones & Exilios*, 15: 113-136. [<http://hdl.handle.net/10481/45829>]

- Dombrowski, Q. (2016). *Drupal for humanists*. College Station: Texas A&M University Press, ISBN: 9781623494728. <https://lccn.loc.gov/2016009099>
- Fresa, A., Justrel, B. (2015). *Roadmap for Citizen Researchers in the Age of Digital Culture*, CIVIC EPISTEMOLOGIES FP7 project. Retrieved from http://www.civic-epistemologies.eu/wp-content/uploads/2014/08/CE_Roadmap-Handbook.pdf
- Goings, K., Abel, P. (2013). *The Value of Social Login. Solving the Engagement Gap. Insights from Consumer Research*. Retrieved from <http://www1.janrain.com/rs/janrain/images/Industry-Research-Value-of-Social-Login-2013.pdf>
- Inventario y Caracterización de los Regadíos de Andalucía. Distritos Mediterráneos y Atlánticos, Junta de Andalucía, Consejería de Agricultura y Pesca. 2008, [Dataset]. Retrieved from <https://goo.gl/7XSzHq>
- Nielsen, J. (2006). *The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities*. Retrieved from <https://www.nngroup.com/articles/participation-inequality> [12/02/2018]
- Toscano, M. (2018). *Andalucía y América. Plataforma colaborativa sobre patrimonio artístico iberoamericano*, Quiroga. Revista de Patrimonio Iberoamericano, Num. 14 (Julio-Diciembre 2018). Universidad de Granada, ISSN: 2254-7037.
- Toscano, M., Tierney, J. (2016). *HistoricGraves.com. Explore, Protect, Promote*. Retrieved from <http://historicgraves.com/sites/default/files/presentation/index.html> [12/02/2018]

SHORT BIO OF THE AUTHOR

The author is an archaeologist from the Università della Tuscia, faculty of Conservazione dei Beni Culturali (Viterbo, Italy), and in 2007 he obtained a Master's degree in computer applications and IT in cultural heritage from the Università di Siena (Italy). In the last decade, he has specialized in the development of Web platforms for research, online databases and geographical information systems, in topics related to cultural, historical and artistic heritage. As a specialist in digital humanities, he collaborates with various institutions and private companies in the heritage and research sectors in Ireland, Italy and Spain, providing customized ICT solutions for projects in Cultural Heritage, Archaeology, History, History of Art, Cultural Tourism, Rural Heritage and Historic Cities. His research focus on the use and application of ICT and the Internet as scientific research tools that enhance the development of Digital Humanities.

El autor es arqueólogo por la Università della Tuscia, facultad de Conservazione dei Beni Culturali (Viterbo, Italia), y obtiene en 2007 un Master en gestión informática del patrimonio cultural por la Università di Siena (Italia). En la última década se ha especializado en el desarrollo de plataformas web para la investigación, bases de datos en línea y sistemas de información geográfica, en temáticas relacionadas con el patrimonio cultural, histórico y artístico. Como especialista en humanidades digitales, colabora con varias instituciones y empresas privadas del sector patrimonial y de la investigación en Irlanda, Italia y España, brindando soluciones TIC realizadas a medida para proyectos en Patrimonio Cultural, Arqueología, Historia, Historia del Arte, Turismo Cultural, Patrimonio Rural y Ciudades Históricas. Su investigación se centra en el uso y aplicación de las TIC y el Internet como herramientas de investigación científica que potencian el desarrollo de las Humanidades Digitales.