Interreg
Italy – Croatia

Co-funded by the European Union

CATCHING THE WAVES OF COOPERATION

DEMO PLATFORM ARCHITECTURE

Final version of November 27th, 2024

Deliverable Number D.1.2.2.

SUSTAINABLE AND SMART DIGITALIZATION IN PUBLIC SERVICE SECTOR

Project acronym: DIGITsmart

Project ID: ITHR0400304

Document History			
Project: DIGITsmart– Interreg Italy-Croatia		**Distribution:** Public	
Version	***Status***	***Date***	***Author***
0.1	Final	27/11/2024	PP2 – Romagna Tech
Notes:			

Summary

1. Introduction 3
2. Guidelines for digitising archive documentation 4
2.1 The Working Methodology 4
2.2 Initial Assessment - City of Kaštela 4
2.3 Initial Evaluation - Di Vagno Foundation 5
2.4 Objectives 6
2.4.1 Expected Functionality - Back-end 7
2.4.2 Expected Functionality - Front-end 7
2.5 Technological Reference Architecture for the Objectives 8
2.5.1 The characteristics of client-server systems 9
2.6 Technology Scouting for Partners 11
2.6.1 Specific needs of individual partners 11
2.6.2 Document Scanning 11
2.7 Guidelines for the Implementation of the Technological Architecture 12
2.7.1 Guidelines - Architecture City of Kaštela 13
2.7.2 Guidelines - Di Vagno Foundation Architecture 14
Bibliography and sitography 16

1. Introduction

The "Digitsmart" project was established with the objective of promoting the digitisation of the historical and documentary archives of two distinct partners: the Di Vagno Foundation and the City of Kaštela. This digitisation process is based on two fundamental aspects: the general objectives of the project, which provide a common frame of reference, and the specific objectives of each partner, which define the practical implementation of the technological systems.

The general objectives of the project include the digitisation of archival materials of different nature, both cultural and cadastral, and the creation of platforms that make these materials easily accessible to the public. Another important aspect is the exchange of good practices between partners and stakeholders, both public and private. Furthermore, it is crucial to comply with the accessibility requirements of European and national legislation, thus ensuring inclusive access to digitised content.

With regard to the partners' specific objectives, the Di Vagno Foundation focuses on the digitisation of photographic material and the implementation of a freely accessible archival platform. This platform is to use the archival description standard F v3.0 and will be required to manage a large amount of data (Big Data). Furthermore, the Foundation plans to digitise further relevant archives over time, in addition to those included in the project.

The City of Kaštela, on the other hand, aims to optimise its business processes through an interactive platform, making cadastral documentation easily accessible through GIS technology. Another objective is to simplify the process of managing and uploading archival data/documents and to migrate data from the current site to the new web platform. Again, the management of a large amount of data is planned, but not the integration of other archives, in addition to those included in the project.

These guidelines are intended to be a useful tool to promote the process of digital transformation in the different sectors of the cultural ecosystem, addressing museums, archives, libraries, superintendencies, institutes and public cultural sites that preserve, protect, manage and enhance cultural heritage.

The document is the result of a study and analysis of the current panorama with the aim of creating clear and comprehensible guidelines for all users. It should be noted that the document analyses everything that concerns digitisation and the technological tools that are useful to facilitate this process. The subject of archiving and archiving techniques are specific to organisations and are not dealt with in this document.

Starting from this outline, all users are called upon to define their own digitisation path: this is therefore not a prescriptive document, given that digital transformation is not a univocal and standardised process. Each cultural institute, in relation to its specific objectives, the resources available and its level of digital maturity, will define the actions to be implemented in line with the technologies and tools shared here.

2. Guidelines for digitising archive documentation

Starting with the project objectives, a detailed survey was conducted with the partners to gather information on the state of the art and to identify the technology needs necessary to achieve the objectives. The survey helped to understand which technologies needed to be scouted out.

The common need to digitise a historical archive facilitated the exchange of good practices consistent with the overall project objectives. However, as the content to be archived differs substantially among the partners, these practices refer to methodologies to deal with digitisation and operational steps to implement a digital system. The definition of the methodology and implementation steps made it possible to focus on the technological needs and the systems needed to meet them.

2.1 The Working Methodology

Digitising a paper archive involves several steps and it is essential to outline a clear and defined process and working methodology.

First, it is essential to conduct a thorough initial assessment of the paper archive. For both partners, it was decided to provide an excel document summarising the documents in the partners' possession, identifying some essential characteristics including: file type, number of files to be digitised, file size, digitisation goals, state of the art and special notes to be made.

Next, the documents must be properly prepared to facilitate the scanning process. The scanning phase transforms paper documents into digital images using dedicated scanners or professional digitisation services. It is important to ensure that the digital images are of high quality and readable.

After scanning, digital images must be indexed and classified so that they are easily searchable and retrievable, adding metadata such as title, date, author, and organising them in a logical filing structure. The digitised documents must then be stored and managed in an information system that allows easy and secure access by users.

Finally, it is crucial to provide training to staff involved in managing and accessing digitised documents, including training on the use of the information system, document search and retrieval procedures.

2.2 Initial Assessment- City of Kaštela

Type of document that needs to be digitised: Construction/building documentation

- Type of content (cadastral map, permission to build, archival photos, correspondance, etc): Construction/building documentation
- Detailed content description: Act for building permits; Act for the use of the building; Location permit; Registration of the start of construction

Analog archive size

- Number of documents: 4366

- Average size of documents: Approximately 1-3 pages

Need of digitization

- Describe to objective of digitisation: Each digitised (scanned) document will be incorporated into a GIS system that will be easily searchable according to certain attributes.
- % of archive already digitised: 82%.
- Digitised file format: *.pdf
- Format expected for files that needs to be digitized: *.pdf
- Occupation in Hard Drive by the documents already digitised: 58 GB
- Hypothesis of Hard Drive occupation for the digitisation of the remaining part of archive that needs to be digitised: 6 GB

State of art

- Indicate any technology system already in use in your organisation for digitization and archiving material: eArhiva
- Indicate any reference of IT systems at local, national, European level for archiving the type of content and significance within the project: eArhiva

Notes

Indicate any other consideration for this type of content: anything.

From the initial assessment conducted, several key aspects emerge for the city of Kaštela. Firstly, it is noted that the document format is standard and the amount of documents to be uploaded is medium.

The size of the documents is not large and this assumes a low occupation of disk space. It is also noted that a system is already in use for digitising and archiving that lacks integration with mapping systems.

2.3 Initial Evaluation - Di Vagno Foundation

Document type being digitised

- Type of content (cadastral plan, building permit, archive photos, correspondence, etc.): photographs kept in archival fonds
- Size of the analogue archive: it is intended to digitise a large part of the photographic heritage currently owned by the Foundation; photographs belonging to the following archive fonds will be filed and digitised: Nicola Baldi, Nicola Damiani, Rocco Errico, Oberdan Laforgia, Gianvito Mastroleo, Nicola Occhiofino, Giulio Orlando, Tommaso Sicolo
- Number of documents: 10,000
- Average document size: 10,000

Need for digitisation

- Describe the objective of digitisation: to safeguard the physical integrity of documents by avoiding their handling for consultation; to have web-enabled formats for the dissemination and dissemination of archival content
- % archive already digitised: 0.05%.
- format of digitised files: tiff, jpg
- expected format for files to be digitised: tiff, jpg
- Disk space occupied by already digitised documents: 40 GB
- Assumption of disk space needed for digital archiving of the remaining part of the archive to be digitised: 1 TB

State of the Art

- Indicate any computer systems already in use in your organisation for archiving content: sw Archiui
- Indicate any local, national, European reference information systems for storing the type of content and significance within the project: europeana.eu
- https://www.censimento.fotografia.italia.it/

Notes

Please indicate any other considerations regarding this type of content: no.

Several considerations emerge from the initial analysis conducted with the di Vagno Foundation. Firstly, the document format is standard, but the amount of documents to be uploaded is high, but given the high volume of documents, a high disk space occupancy is assumed.

The Foundation also has in use a digitisation and archiving system that lacks a consultation system to meet users' needs.

2.4 Objectives

The project aims to ensure the digitisation of material in the possession of the City of Kastela and the Di Vagno Foundation. Based on the initial assessment, several key steps were identified in order to achieve the set objective.

1. Data collection and digitisation.
2. Definition of the technologies used in the architecture of an information system for the management of digitised data.
3. Implementation of a digital system for uploading digitised archive data (back-end application).
4. Development of an online platform for consulting digitised documents (front-end application).

2.4.1 Expected Functionality - Back-end

The project provides a back-end system with advanced features to ensure a secure, efficient and user-friendly document management experience. One of the key features is the authentication and authorisation system, which allows for robust verification of user identity and ensures that only authorised persons can access document uploads. This system ensures that every action performed within the platform is traceable and controlled, reducing the risk of unauthorised access.

In terms of file management, the back-end supports the upload of a variety of file types, ensuring that these are stored securely and that data integrity is maintained. The user interface is designed to be intuitive and user-friendly, allowing quick and uncomplicated uploading of documents or files. Users can interact with the system easily, thanks to a design that facilitates the upload process.

Error handling is another key component, with the system providing clear and informative messages to users in the event of problems during file uploads. This approach minimises confusion and allows users to quickly resolve any problems.

The security measures implemented in the system are extremely robust, designed to protect the uploaded data from unauthorised access and possible security breaches. Finally, the system is designed to be scalable, capable of handling an increase in workload over time without compromising performance. This scalability is essential to ensure that the system remains efficient and reliable even with an increasing volume of documents and users.

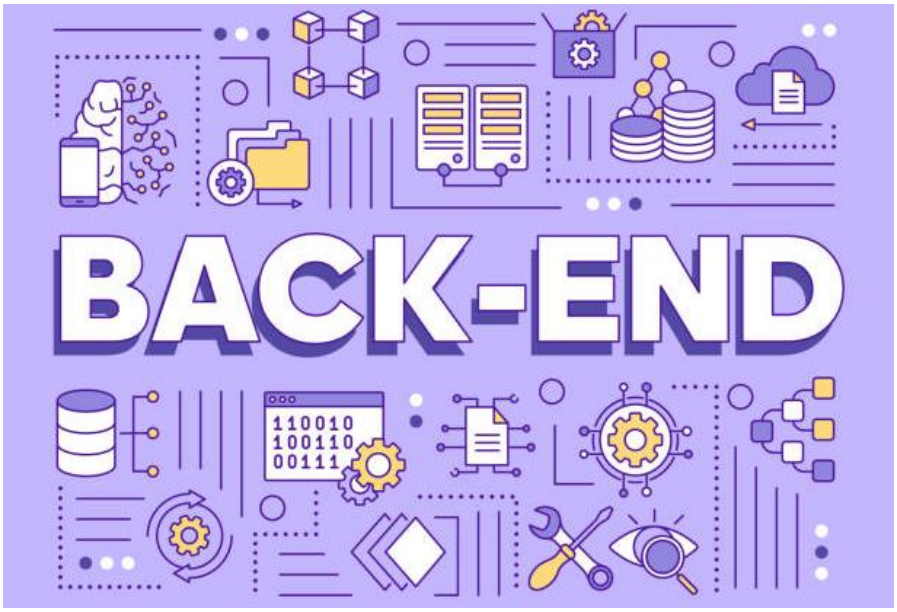

Figure 7 - Back-end.

2.4.2 Expected Functionality - Front-end

The project involves the realisation of a front-end characterised by an intuitive user interface, designed with a user-friendly design that allows easy and immediate navigation. Advanced search functionality is a key element, allowing users to perform effective searches through advanced options and specific filters. Search results will be presented in a clear and readable manner, ensuring quick access to the desired information.

The interface will also be fully responsive, ensuring a consistent and optimal user experience on desktops, tablets and mobile devices. This approach ensures that all users can access and use the platform regardless of the device they are using.

Particular attention will be paid to accessibility, ensuring that the front-end is also usable by users using assistive technologies. This commitment reflects the will to make the platform inclusive and usable by all, in accordance with current accessibility regulations.

Front-end performance will be optimised to ensure fast loading times and smooth navigation, improving the overall user experience. Security remains a priority, with robust measures in place to protect user data and sensitive information from any kind of threat.

Finally, the front-end will be closely integrated with the back-end, ensuring efficient data flow and synergetic operations. This integration maintains operational consistency between the different components of the system, ensuring that all functionalities work harmoniously to offer a high quality service to users.

Figure 8 - Frontend.

2.5 Technological Reference Architecture for the Objectives

From a technological point of view, the development of an online content storage and consultation system is managed through client-server architectures that guarantee high standards of reliability, security and customisation.

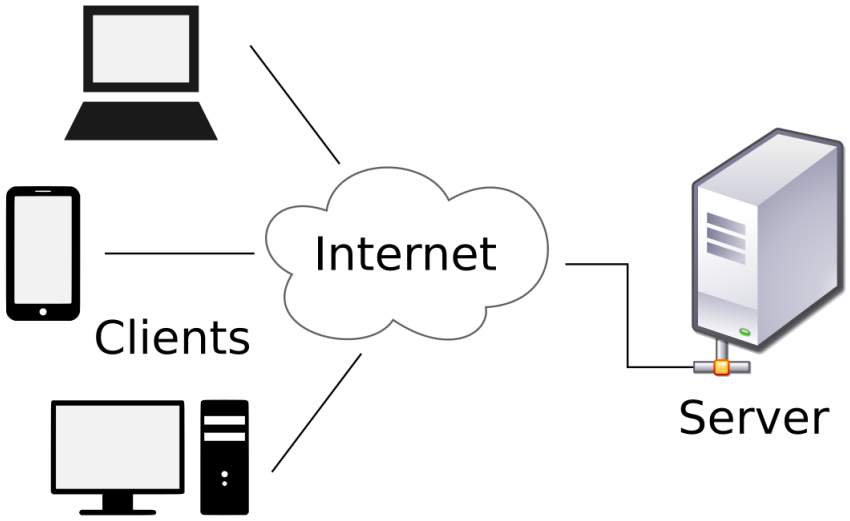

Figure 9 - Different types of clients connecting to the server.

2.5.1 The characteristics of client-server systems

Client-server systems are a network architecture that divides functions between service providers (servers) and service requesters (clients). These systems have certain special characteristics:

- Distribution of roles: In a client-server architecture, the roles of applications are clearly defined. Clients are responsible for interacting with end-users and sending requests to servers to obtain responses or services, while servers are responsible for providing the requested services and managing the associated data.
- Communication based on requests and replies: Communication between client and server takes place via a request-response scheme. Clients send requests to servers, specifying the desired action or data required, and servers respond by providing the requested results or services.
- Centralisation of data and resources: In client-server systems, data and resources are generally centralised on servers. This allows for centralised management of data access and sharing, improving system consistency and security.
- Scalability and deployment: The client-server architecture supports the scalability and deployment of the system by adding or removing servers to manage workloads as required. This allows the capacity of the system to be increased to handle more users or workloads without compromising performance.
- Separate user interface: In client-server systems, the user interface is separate from the business logic and data of the system. This allows the user interface to be updated or replaced without having to change the underlying business logic or data, facilitating maintenance and evolution of the system over time.
- Reliability and availability: Client-server architecture can improve system reliability and availability by distributing servers across multiple physical or virtual machines and providing

backup and redundancy mechanisms to ensure continuous operation even in the event of hardware failures or software errors.

- Security: The client-server architecture allows centralised security measures to be implemented on the servers to protect access to data and system resources. This includes role-based access controls, data encryption and authentication and authorisation mechanisms.
- Centralised resource management: In client-server systems, resource management is centralised on the servers, enabling efficient and scalable monitoring and management of resource utilisation, such as memory, network bandwidth and storage capacity.

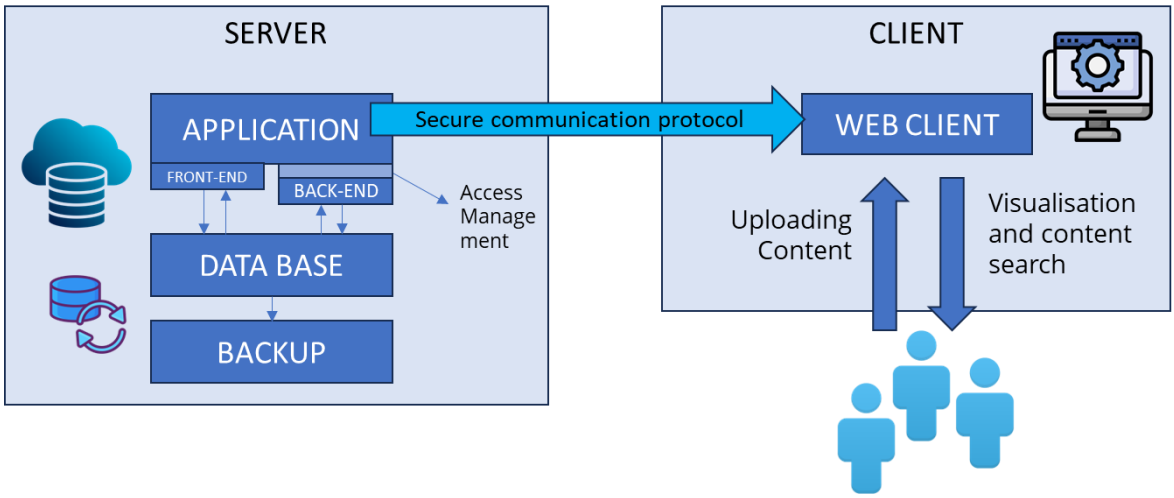

Figure 10 - User interaction in the client-server system.

2.6 Technology Scouting for Partners

During technology scouting, we will focus on two fundamental aspects: the technologies used to scan documents and the architecture of the web platform. With regard to scanning technologies, based on the initial analysis we conducted, we will identify the best solutions considering the type of documents to be digitised and the specific scanning objectives.

In parallel, we will examine the architecture of the web platform, also taking into account existing structures offering cultural content. The objective will be to identify scalable and flexible architectural approaches that can support a gradual and adaptable growth of the platform over time. This approach will enable the development of a solid foundation that can evolve and respond effectively to future needs, while ensuring an optimal user experience and high operational efficiency.

2.6.1 Specific needs of individual partners

On the other hand, there are different needs between the two partners: both from the point of view of the type of data collected, the size of the files and the development of the platform.

With regard to the type of archive document, the city of Kastela needs to scan cadastral documents, while the di Vagno Foundation needs to scan historical and cultural material that is particularly old and therefore delicate.

The Foundation of Vagno experiences a need to search for historical archive documents according to search keys typical of users of such archives; the City of Kaštela, on the other hand, intends to use the classification representation of the archive document with the GIS system for geolocalisation.

Finally, the two partners show a very different amount of data to be stored: 1Tb for the Di Vagno Foundation versus 70Gb for the City of Kaštela.

2.6.2 Document Scanning

While the City of Kaštela can use an ordinary office scanner for document scanning as the documents do not have any special features to consider, the same cannot be said for the di Vagno Foundation.

In this case, in fact, the documents to be scanned are old and particularly delicate: mere contact could alter them. In this sense, the Foundation has defined special criteria according to which the scanning should take place:

- Use of specialised non-invasive equipment
- Use of cold light lamps and adequate filtering of ultraviolet radiation.

Based on the analysis of available technologies for document scanning carried out in the previous paragraphs, we can state that there are several solutions available to meet the Foundation's needs:

- Planetary scanners: These scanners have a dome or sphere-shaped structure under which the document to be digitised is placed. A camera or set of cameras are mounted above the dome, so that a complete image of the document can be captured at once.
- Auto-feed scanners: These scanners use an auto-feed mechanism to transport documents through the device without touching them directly.

- Laser reading or 3D imaging scanners: These scanners use laser reading or 3D imaging technologies to digitise documents without touching them directly.
- Ambient light or infrared scanners: Some scanners use ambient light or infrared light to capture images of documents without the need for direct contact. These scanners can be used to digitise documents that are sensitive to light or vulnerable to physical contact.
- Ultrasonic scanners: Some scanners use ultrasonic waves to detect the shape and structure of documents without touching them directly. These scanners can be used to digitise sensitive or fragile documents without damaging them.
- Non-contact camera scanners: These scanners use high-resolution cameras to capture images of documents without touching them directly. They can be used to digitise documents in various formats and sizes without damaging them.

For needs of this kind, planetary scanners are frequently used in the cultural, academic and archival fields:

- Flexibility and adaptability especially for books, maps, manuscripts, works of art and more.
- High and detailed image quality, allowing every detail of scanned documents to be captured with precision and clarity.
- High digitising speed even at high image quality.

2.7 Guidelines for the Implementation of the Technological Architecture

Based on the scouting evidence, we are able to outline a reference technology architecture for the digitisation of the project partners' filing system.

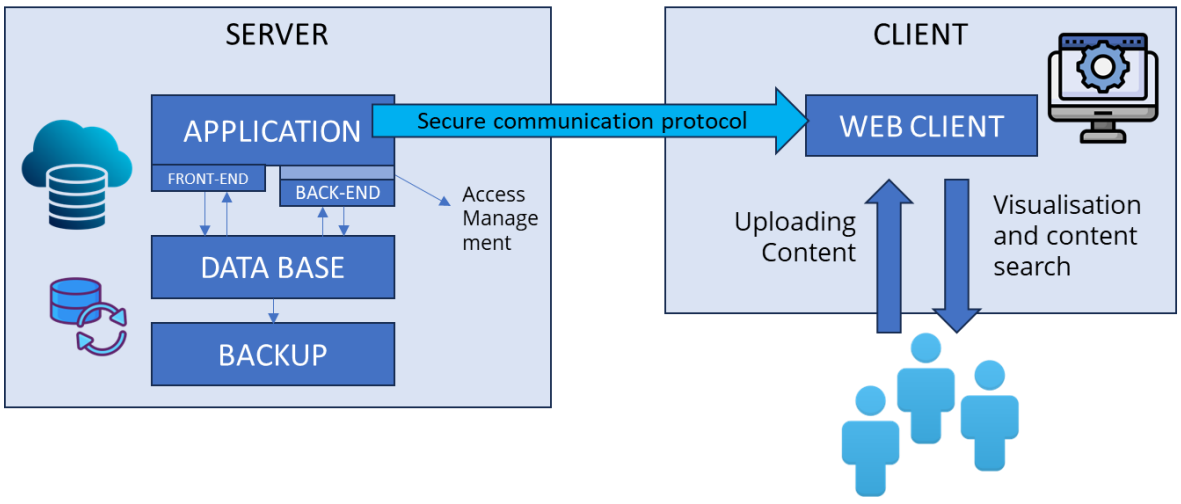

Figure 11 - Client-Server Architecture in detail.

The guidelines were developed by defining a common approach to the digitisation goals of the project partners in the methodological and project phases.

However, during the analysis for the development of the web platform, the different requirements between the two partners involved in terms of:

- Type of archive documents
- Searching for historical archive documents
- Geolocation
- Mole of data to be archived

These elements of diversity led to the need to use a different approach to the implementation of the architecture.

The common approach based on sharing methodological practices and technological architectures thus leads to two customised implementation solutions optimised to the partners' objectives.

2.7.1 Guidelines - Architecture City of Kaštela

Figure 12 - Wordpress

For the City of Kaštela, the management of a significant amount of data is accompanied by the need to represent documents through GIS technology. With this in mind, it seems most appropriate to consider the creation of a new platform based on a CMS such as WordPress. Open source technology allows development costs to be considerably limited and enables the development of a customised system. However, there is no direct connection between GIS and WordPress; GIS data are imported once into their own dedicated application and have no direct connection to WordPress.

GIS data are saved in the database used by the GIS application itself. With regard to storage and backup of the platform, data backups are made once a day on the executor's servers, which also take care of space and traffic management. There are no plans to migrate GIS data, and it is not clear which type of migration the initial question refers to. If it is WordPress, there is no GIS data migration process linked to it.

The decision to maintain two separate platforms is motivated by several strategic and operational considerations. First of all, separating the platforms allows optimising the performance and reliability of each system, avoiding potential conflicts and compatibility problems that could arise from direct integration. In addition, maintaining GIS on a dedicated platform ensures that specialised spatial data management and visualisation functionalities are handled by a system optimised for such tasks, improving the effectiveness and efficiency of operations.

Using WordPress for non-GIS content management allows you to take advantage of its flexibility and wide availability of plugins and themes to create a feature-rich website without the added complexity of directly integrating GIS. This approach allows you to manage and upload files via your own online storage space, and maintain compatibility with other cloud services for regular backups, without overloading the GIS system.

Finally, the separation of platforms offers greater security, as it isolates sensitive GIS data within their own dedicated environment, reducing the risk of unauthorised access and security breaches that could result from vulnerabilities in the CMS

Examples of relevant sites developed with Wordpress:

1. Historic Jamestowne: https://historicjamestowne.org/education/for-educators/lesson-plans/
2. Frederiksborg Castle: https://dnm.dk/en/
3. DESIGNMUSEUM DANMARK: https://mus.reindex.net/DMD/main/Landing.php?Lang=dan
4. Nasa: https://www.nasa.gov/
5. Digital Public Library of America: https://dp.la/
6. Open Crops: https://www.openculture.com/
7. Biblioteca Nacional De Espana: https://www.bne.es/
8. The Walters Art Museum: https://art.thewalters.org/

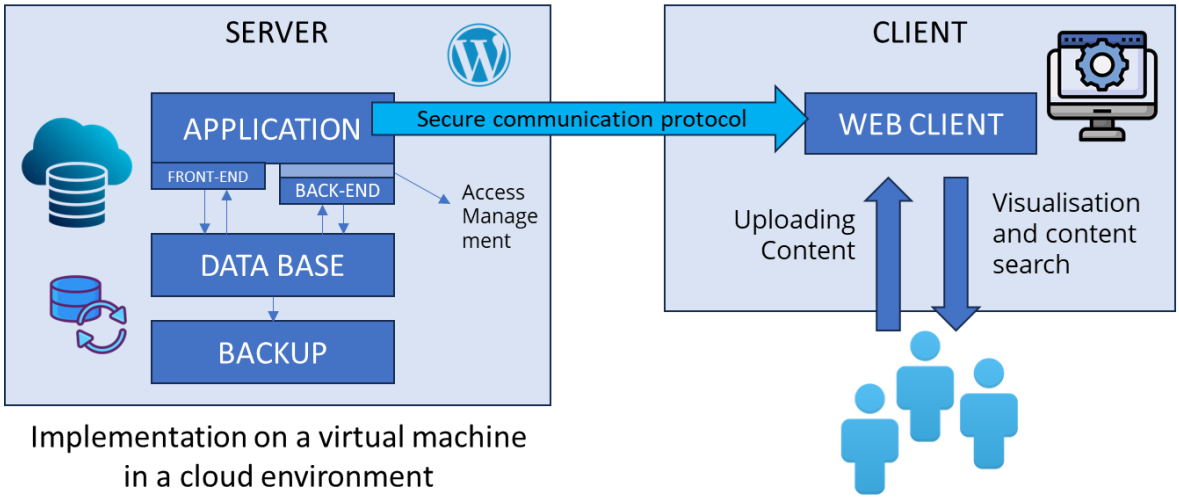

Figure 13 - Communications in the Client-Server architecture of Wordpress.

2.7.2 Guidelines - Di Vagno Foundation Architecture

For the Di Vagno Foundation, the need for large dedicated storage space emerges, making the use of consolidated cloud-based systems appropriate. In this context, it could be extremely advantageous to consider a re-engineering of the current platform in use, Archiui. This cloud platform, which is already a reference system for document archiving in Italy, could be further optimised to improve search criteria and make them more functional for the public interested in historical archive documents.

Implementing a re-engineering of Archiui would not only significantly expand the archiving capabilities, but also improve data management in an effective and consistent way as the Foundation's information assets continue to grow. With the passage of time and the increase of digitised historical documents, this solution would offer scalable and sustainable management, ensuring accessibility and usability of content for researchers, historians and the general public. The optimisation of search criteria would also

make it easier and more intuitive to consult the documents, thus increasing user interest and engagement. In summary, this strategy would allow the Di Vagno Foundation to proactively address the challenges of managing an expanding archive, while ensuring the integrity and enhancement of its documentary heritage.

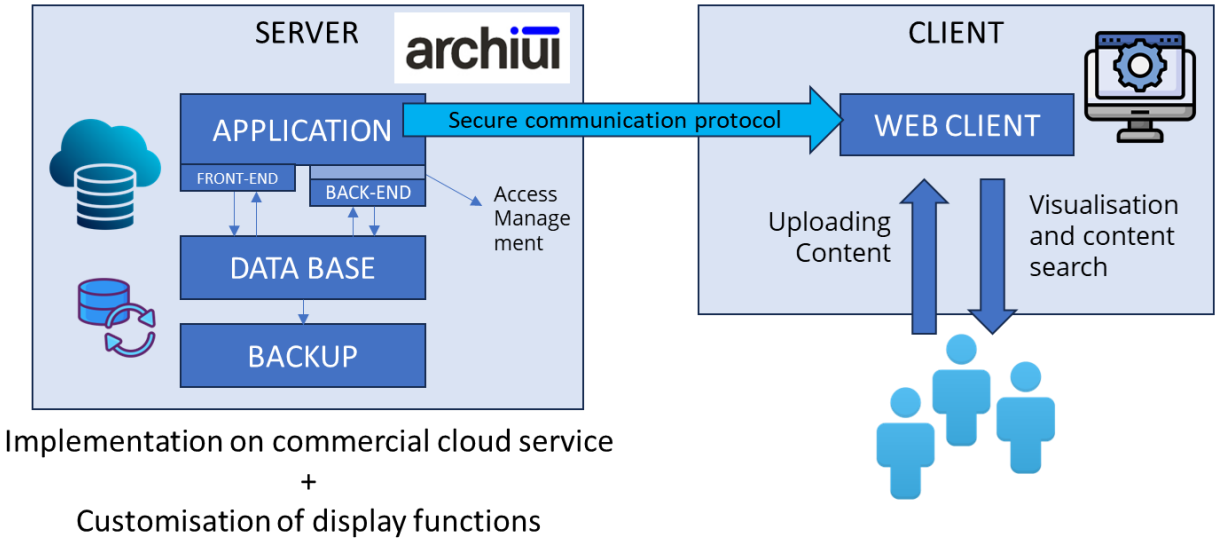

Figure 14 - Communications in Archiui's Client-Server architecture.

Bibliography and sitography

Allegrezza, Stefano. "Artificial intelligence applied to the digitisation of cultural heritage." The fiftieth anniversary of the UNESCO Convention for the Protection of Cultural Heritage (1972-2022): multidisciplinary paths: 9.

Cartelli, A., Daltri, A., Errani, P., Palma, M., & Zanfini, P. (2006). The open catalogue of Malatesta manuscripts. AIB Bulletin (1992-2011), 46(3), 207-220.

Compagnone, Denise. "Acea, memory recovered. The historical archive and the immersive museum as tools for the protection and enhancement of a century of corporate cultural heritage." DigItalia 17.2 (2022): 78-86.

Donadio, Elisabetta. Laser scanning and GIS technology for the stratigraphic analysis of the built heritage: the church of Santa Maria dell'Abbazia di Staffarda. Diss. Polytechnic University of Turin, 2013.

OF DIGITALISATION, I. PROFESSIONALS, AND DOCUMENTAL AND. PRIVACY. "The new Guidelines on the formation, management and storage of computerised documents."

Ferrara, Luigi. "Artificial intelligence tools in the national plan for the digitisation of cultural heritage and the evolution of the European regulatory framework. Initial observations." (2024).

Garlatti, A., and S. Iacuzzi. "Digitisation: challenges, opportunities and risks of smart technologies for local authorities." ACCOUNTING AND BUSINESS ECONOMICS SERIES (2020): 75-88.

Moses, R. P. (2005) A Glossary of Archival and Records Management Terminology. Chicago: The Society of American Archivists.

Masciariello, Antonio. "MEMORY, IDENTITY AND PROJECT: THE BUSINESS ARCHIVE IN FASHION DESIGN AS A CULTURAL PROJECT AND AS A PRODUCTION TOOL. THE CASE OF THE VERSACE HISTORICAL ARCHIVE." (2020).

Marras, Anna Maria. "Digital Glossary and Technologies." (2020): 1-21.

Palamoni, Gloria Mancini. "The sustainable development of cultural heritage between emergencies and digital technologies." Rivista italiana di informatica e diritto 4.1 (2022): 261-272.

ZORZI, CHIARA. 'Digitisation as a tool for the protection and enhancement of Cultural Heritage.'

https://docs.italia.it/italia/icdp/icdp-pnd-digitalizzazione-docs/it/v1.0-giugno-2022/index.html

https://www.forumpachallenge.it/premio_pa_colori/dalla-carta-alla-nuvola/

https://www.html.it/articoli/cms-cosa-e-migliori/

https://www3.diism.unisi.it/~benelli/scienze_della_comunicazione/dispense/2005_06/ClientServer.pdf

https://boolean.careers/blog/front-end-e-back-end-web-developer-ecco-la-differenza