

The ARQUIGRAFIA project: A Web Collaborative Environment for Architecture and Urban Heritage Image

Vânia Mara Alves Lima^{1†}, Cibele Araújo Camargo Marques dos Santos¹,
Artur Simões Rozestraten²

¹Department of Information and Culture, University of São Paulo, São Paulo, Brazil

²Department of Architecture Technology, University of São Paulo, São Paulo, Brazil

Abstract

Purpose: This paper presents the ARQUIGRAFIA project, an open, public and nonprofit, continuous growth web collaborative environment dedicated to Brazilian architectural photographic images.

Design/methodology/approach: The ARQUIGRAFIA project promotes the active and collaborative participation among its institutional users (GLAMs, NGOs, laboratories and research groups) and private users (students, professionals, professors, researchers), both can create an account and share their digitized iconographic collections in the same Web environment by uploading their files, indexing, georeferencing and assigning a Creative Commons license.

Findings: The development of users interactions by means of semantic differentials impressions recording on visible plastic-spatial aspects of the architectures in synthetic infographics, as well as by the retrieval of images through an advanced system search based on those impressions parameters. By gamification means, the system often invites users to review images' in order to improve images' data accuracy. The pilot project named Open Air Museum that allows users to add audio descriptions to images in situ. An interface for users' digital curatorship will be soon available.

Research limitations: The ARQUIGRAFIA's multidisciplinary team gathering professors-researchers, graduate and undergraduate students from the Architecture and Urbanism, Design, Information Science, Computer Science faculties of the University of São Paulo, demands continuous financial resources for grants, for contracting third party services, for the participation in scientific events in Brazil and abroad, and for equipment. Since 2016, significant budget cuts in the University of São Paulo own research funds and in Brazilian federal scientific agencies can compromise the continuity of this project.

Practical implications: The open source template called +GRAFIA that can freely help other areas of knowledge to build their own visual Web collaborative environments.

Citation: Lima, Vânia Mara Alves, Cibele Araújo Camargo Marques dos Santos, and Artur Simões Rozestraten. "The ARQUIGRAFIA project: A web collaborative environment for architecture and urban heritage image." *Journal of Data and Information Science*, vol.5, no.1, 2020, pp. 51–67.

DOI: 10.2478/jdis-2020-0005

Received: Jan. 17, 2020

Revised: Mar. 2, 2020

Accepted: Mar. 10, 2020



[†] Corresponding author: Vânia Mara Alves Lima (E-mail: vamal@usp.br).

Research Paper

Originality/value: The collaborative nature of the ARQUIGRAFIA distinguishes it from institutional image databases on the internet, precisely because it involves a heterogeneous network of collaborators.

Keywords Metadata; Architectural images; Collaborative web environment; Digital repository

1 Introduction

Knowledge Organization (KO) in the context of the new epistemological challenge called Digital Humanities—defined as the articulation of knowledge and methods used in the human sciences with the digital world (Guerreiro & Borbinha, 2014)—needs to establish new methods and procedures, as well as develop increasingly efficient tools, to represent and retrieve the iconographic information, which has grown exponentially as a result of increasing digitization of physical collections of images to be made available on the web.

For some years now we have had initiatives to preserve and guarantee access to cultural heritage, such as Europeana (<http://www.europeana.eu/portal/pt>) and the Getty Research Portal (<http://portal.getty.edu/>). Europeana is a free curatorial image platform that gathers several European institutions aiming to transform the world through culture, sharing high-quality iconographic collections. The Getty Research Portal (<http://portal.getty.edu/>) is a free online search platform providing worldwide access to an extensive collection of digitized art history texts and images from a worldwide range of institutions. It is an inter-institutional collaborative project initiated by the Getty Research Institute responsible for the most well-known tool for describing architecture and urban cultural heritage, the Art & Architecture Thesaurus.

Besides these large scale GLAMs (Galleries, Libraries, Archives, and Museums) many smaller institutions also run digitization centers and a steadily growing number of platforms to render their digital contents globally accessible for research and the interested public as the ETH Library (Gasser, 2017). ETH Zurich's main library with its shift in strategic focus towards the “digital library” has implemented crowdsourcing, a form of user participation used very successfully by a wide variety of institutions all over the world to enhance digital collections and raise their profile. In the ETH Library initiative, the volunteers have free access to the digitized images and they have their names included in the comments field on the Image Database, which is an incentive to continue collaborating with the growing amount of improved metadata. Similar to the ETH Library initiative, but expanding the collaborative approach, we present the ARQUIGRAFIA project.



ARQUIGRAFIA is a web collaborative environment for the preservation, research and dissemination of images of Brazilian architecture and urban spaces, which enables interactions between people and institutions. This digital environment contributes to research on architectural and urban heritage, as well as allowing the organization of Brazilian architectural images on the web (Lima et al., 2016). The collaborative nature of the ARQUIGRAFIA project distinguishes it from institutional image databases on the internet, precisely because it involves a heterogeneous network of collaborators: institutional users such as GLAMs, NGOs, Universities and Research Groups, together with private users such as students, teachers, photographers and people in general.

Since 2010, the ARQUIGRAFIA project has been facing scientific and technological challenges for creating apps and system features that promote active and collaborative participation among its users. Institutional and private users can create an account and share their digitized iconographic collections in the same Web environment by uploading their files, indexing, georeferencing and assigning a Creative Commons license. However, collaboration goes beyond uploading images and deals with user interactions—exchanging assessments, impressions, judgements on the architectural qualities represented in the photographs.

Most of the ARQUIGRAFIA users are architecture students (38.6%); but there are also architects (23.4%); undergraduate students in other areas (10.3%); teachers of architecture (4.3%); photographers (4.1%); and interested lay people (19,3%). Most of these users are between 20 and 30 years old. Together, they collaborate to tag, georeference and assign specific licenses (Creative Commons) to each image of their collections. Being an environment of teaching and researching, we can also include it in the field of Digital Humanities, since this field is defined by research in collaboration with teaching activities, combining computing and information technologies with academic practices in the field of humanities (Lima, Rozestraten, & Orth, 2016).

ARQUIGRAFIA has several technological and scientific integrated fronts: from the cleaning and conservation of original photographic images to its digitization, from the training of researchers to the development of a perpetual beta open source software. From the point of view of knowledge organization systems (KOS), ARQUIGRAFIA comprises both the broader and the narrower meanings of Knowledge Organization (KO) as defined by Hjørland (2008). The more general and broader meaning, because it establishes an inevitable relationship between knowledge and its organization in the society, since its images also represent the social organization of knowledge and its tags represent the conceptual structure in the field of the architecture and urbanism. The more specific and narrower meaning,



because the construction of its controlled vocabulary seeks to organize knowledge and information intellectually and cognitively, facilitating its management and retrieval, including, in addition to institutional indexing, social indexing as well. Therefore, ARQUIGRAFIA provides an opportunity to look accurately at contemporary issues, specifically regarding terminologies and controlled vocabularies for the representation of images. Other issues include the sharing of metadata between systems; as well as the consolidation of standards that respond both to international interoperability requirements and to local needs for information access and organization.

From a technological perspective, ARQUIGRAFIA plays the role of a pilot program for a template called +GRAFIA that was based on PHP Laravel and can offer free help for other areas of knowledge to build their own visual collaborative environments, such as, for example, a hypothetical BOTANYGRAFIA dedicated to the flora, or an ARTGRAFIA, dedicated to visual arts.

Conceptual and technological challenges concerning the design and the operation of ARQUIGRAFIA allow us to characterize it as an online experimental laboratory and a case study on the opportunities and the risks of digital projects on the humanities based on image collections. ARQUIGRAFIA proposes a new interface for the information environment of the institutional collection and, in so doing, it confronts controlled procedures of knowledge representation with the semantic capability of organizing the information for the web and networks.

2 First steps

The main source of the images for ARQUIGRAFIA is the photographic collection of the Iconographic Material Sector of the Library of the School of Architecture and Urbanism of the University of São Paulo (FAUUSP). A set of 42,000 images (34,000 slides and 8,000 black and white photographs) were scanned. Of this total of digitized images, more than 8,000 images were uploaded to the ARQUIGRAFIA system, precisely those that already have the authorizations of the copyright holders as a Creative Commons license. The remaining images are in the curatorial stage of requesting authorization and analysis for upload. In addition, 3,000 images were uploaded and cataloged by private users as well, and other 1,782 belonging to other institutional collections such as the Republican Museum of Itu (<http://www.mp.usp.br/museu-republicano-de-Itu>) and the QUAPÁ, the Panorama of Brazilian Landscaping Project (<http://quapa.fau.usp.br/wordpress/>). Figure 1 shows ARQUIGRAFIA's home and login page.

Between 2012 and 2017, a team of scholarship-funded undergraduate students carried a wide effort towards cleaning, identifying and organizing the original images, as well as digital files and backups. Having to deal with FAUUSP Library's



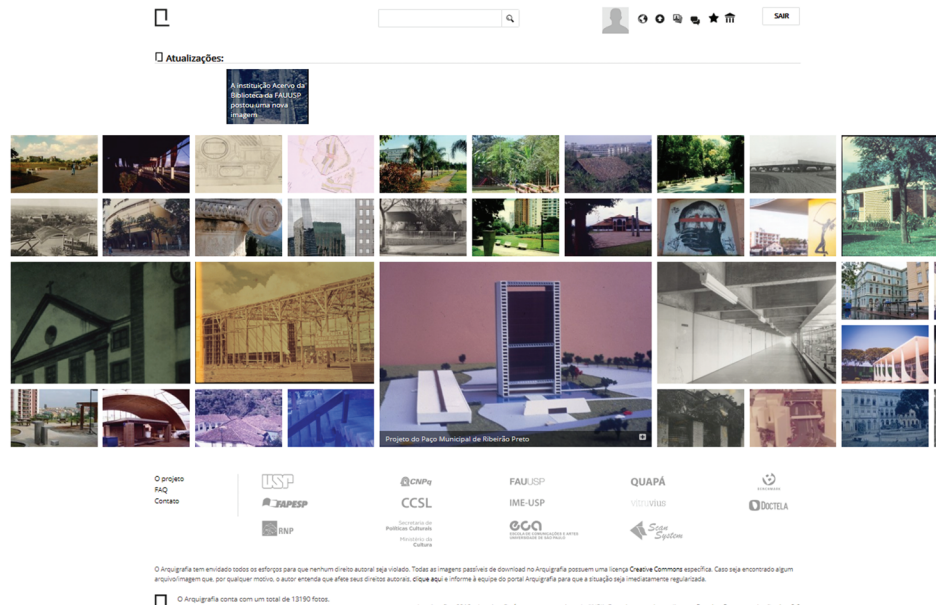


Figure 1. ARQUIGRAFIA homepage.

Source: <http://www.arquigrafia.org.br/home>

sizeable collection of photographs, ARQUIGRAFIA chose to focus on Brazilian Architecture and its urban spaces, letting aside foreign architecture images for now. This set was then cleaned and stored with proper materials, such as folders and mounting corners, which helps the conservation of the originals.

At the same time, cataloguing results were revised based on a survey of procedures and standards for the descriptive and thematic representation of images in order to define the set of metadata which best suits the organization and retrieval of information.

For the digitization of the institutional collection, a third-party company was hired using Plustek Optic 120 film and Silver Fast Ai Studio 8 (64-bit) scanning software that helps in the removal of dust and scratches. Each image was scanned without color correction in order to preserve the original appearance of the photographs and slides, keeping the time stamps (color changes, smudges, saturation, etc.) and its historical aspect. Each generated file is 5 MB with 4,000 dpi resolution and is saved in TIFF, JPEG, and PDF formats and recorded on DVDs and external hard drives. After the scan, backups of the set of images were created and the images were uploaded to ARQUIGRAFIA.

Each image received a registration number, which allows the association with the metadata of its cataloguing and description. A program transformed the metadata



into content objects used by the system. In order to make this transformation, the Apache ODF Toolkit software (<http://incubator.apache.org/odf toolkit>) was used to create a communication interface between the metadata and the ARQUIGRAFIA system for information mining and transformation.

Then, the information storage activity allows the creation of associations between the content objects and their representation in the database. Once the object association is made (an author is associated to an image and this image to an address), the system uses the Hibernate persistence library (<http://www.hibernate.org>) to store the database in Mysql (<http://www.mysql.com>).

3 Metadata

To define the metadata, some cataloging standards were analyzed, such as the Anglo American Cataloging Code—AACR2; the International Standard for Bibliographic Descriptions for Non-Book Material—ISBD (NBM) and content standards such as Cataloging Cultural Objects—CCO. From the analysis of these standards and the identification of the information required in ARQUIGRAFIA, a spreadsheet was developed to integrate the metadata necessary for the representation of the images and the data administration in the collaborative web environment. In this way, a set of metadata was established according to Table 1.

Table 1. ARQUIGRAFIA Metadata.

Image Metadata Level	Type of Information
Descriptive metadata	Title, Number of the classification, Name, Country, State, City, District, Street, Image author, Tags, Image date, Project author, Construction date, Notes, Date of registration number, Date of cataloging
Structural metadata	Dimensions, width, height, resolution, color depth, color model
Administrative metadata	License (<i>Creative Commons</i>), Harvesting, Donors, Authorization form for web distribution

Source: elaborated by the authors.

Regardless of whether they have personal or institutional access users must fill in at least the title, the name of the author of the image, the country and some tags that represent the subject to upload an image to ARQUIGRAFIA. This procedure can be performed on the website through any device, from notebooks to smartphones (Rozestraten, Lima, & Santos, 2017).

In order to encourage users to closely observe images and formulate judgments about buildings and urban spaces represented in the photographs, from the point of view of Architecture, ARQUIGRAFIA proposes to its users the recording of impressions based on pairs of opposing plastic-spatial qualities, called binomials (Figure 2).



Residência
Manuel Antônio
Mendes André

Suas impressões da Residência Manuel Antônio Mendes André

- Eu conheço pessoalmente esta arquitetura.
- Estou no local.

Para cada um dos pares abaixo, quais são as qualidades predominantes na arquitetura que são visíveis nesta imagem?

Aberta (10 %)	Fechada (90 %)
Interna (20 %)	Externa (80 %)
Complexa (75 %)	Simplex (25 %)
Simétrica (89 %)	Assimétrica (11 %)
Translúcida (33 %)	Opaca (67 %)
Horizontal (85 %)	Vertical (15 %)

Tags: fachada, residência, concreto, muro

Imagens interpretadas com média similar (2) Imagens

Figure 2. ARQUIGRAFIA binomials.

Source: <http://www.arquigrafia.org.br/>

The binomials are organized as semantic differentials such as: open/closed; internal/external; translucent/opaque; complex/simple; symmetrical/asymmetrical; horizontal/vertical. The conceptual underpinnings for these pairs of opposite qualities come from Henrich Wölfflin (1864–1945) “Principles of Art History” (1950), adapted to Architecture, organized as Charles E. Osgood’s (1916–1991) semantic differentials (1990). ARQUIGRAFIA invites users to record their impressions of the architecture represented in an image, based on six pairs of opposing qualities. Gathering multiple individual impressions as collective interpretations, the system can guide cross-system navigations by means of interactions between images with similar and/or mirrored profiles.

To do so, the average of interpretations is calculated and shown in a chart (Figure 3) and compared with the average of other images that had their binomials



Research Paper

analyzed in the system, allowing the identification and retrieval of images with similar patterns.

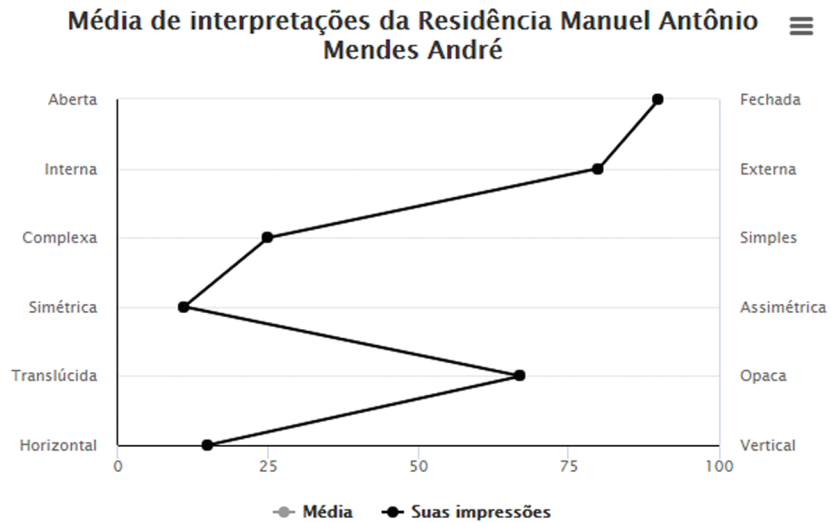


Figure 3. An example of an image interpretation average with suggestions of possibly similar images.
Source: <http://www.arquiografia.org.br/>

In addition, it allows the system to establish, for every image, a comparative perspective between its original interpretation, the later interpretations, and the average of all interpretations already made (Rozestraten et al. 2010). The various interpretations are recorded and can change the classifications of similarities. Additionally, the interpretations of each single user can also establish its profile and preferences over the years.

4 Tags vs controlled vocabulary for a Knowledge Organization System

The inclusion of tags or markers by users with natural language is called folksonomy, social tagging or social indexing. Tagging can greatly contribute to the



creation and management of digital collections, as it is carried out collaboratively, distributing resources, activities, and reducing costs. Its importance for the organization, retrieval and access to digital information is demonstrated by the studies of Angus et al. (2010) and Bradley (2011) on Flickr, a photo sharing and social networking site.

Specifically, the study by Angus et al. (2010) explored the potential use of the Flickr image site as an academic image resource by identifying tagged images belonging to academic subject categories. Image content analysis and term frequency analysis provided information from the context of the image. The results of this study showed the possibility of using the tool as a resource for specific images in some areas of knowledge and for individual academic studies, which reinforces the relevance of using social indexing in an image repository developed for teaching, research and extension purposes such as ARQUIGRAFIA.

The tagging performed at ARQUIGRAFIA is related to Vander Wal's specific folksonomy model, where one or a few people insert the tags (Moreiro Gonzalez, 2011; Rafferty, 2018). Thus, we find in this collaborative environment a mixed knowledge organization system where any user can upload an image and tag it using the suggestion list (controlled vocabulary) or adding terms, but only this user can edit the information and the tags that he inserts. However, other users may contribute by reviewing the information and indicating additions and corrections via the contributor's own editing system. Despite the ease of use of social indexing and the approximation with the active vocabulary of users, the lack of language control can present difficulties for information retrieval.

Rafferty (2018), in an entry on the ISKO Encyclopedia of Knowledge Organization, refers to tagging as the practice in which web users use keywords to describe, categorize or comment on digital content.

This marking allows an individual response to information objects by configuring a triad formation: the user, the information object and the keyword, keeping them connected as observed in Figure 4. The tags are such as: "*representar2015*" and "*#representar2015*" for the Architecture event held in 2015, "*library*" the type of building; "*ufu*" the abbreviation of the institution where the event happened; "*uberlândia*" the city; "*minas gerais*" the province; "*paulo zimbres*" the architect; "*cobogó*" the type of the wall with hollow elements and "*tijolo de barro*", mud brick.

For Font, Serra & Serra (2013), collaborative tagging has emerged as a solution for labeling and organizing digital content on the web. However, these collaborative tagging and social indexing systems present problems regarding tag ambiguity, synonyms and the amount of content words used, and it can be inferred that the organization and navigation of content marked in this way may present difficulties for the efficient retrieval of information.





Figure 4. Central Library of Universidade Federal de Uberlândia.

Source: <https://www.arquiografia.org.br/photos/6298>

The ANSI/NISO Z39.19 “Guidelines for the construction, format, and management of monolingual controlled vocabularies” defines controlled vocabulary as a list of explicit and controlled terms, and these terms may not be ambiguous and must contain definitions which are not redundant (National Information Standards Organization, 2010).

Bearing in mind that in ARQUIGRAFIA both personal and institutional users can upload and index the images of collections of photos, then it is possible to understand the inherent tension between the use of free-form user terminology and the control of indexed information by the institutional users. It happens because this information system needs to fulfill the demands of organizing the institutional images, for the purpose of academic retrieval and preservation, and at the same time count with social indexing and personal user participation for system feedback with images and markers.

Therefore, it was necessary to do the terminology standardization between the lists of subjects used by the library for the indexing of the photographs and slides; the terms of architecture and urbanism of the Controlled Vocabulary of the Integrated Library System of the University of Sao Paulo; the list of default tags of ARQUIGRAFIA, based on the expertise of its team and the tags employed by the users which were harvested in the database.

The 1,145 terms of the list of tags assigned by personal users (Santos & Santos, 2017) were analyzed between 2017–2019 with the purpose of allowing their inclusion in the controlled vocabulary under construction, considering their definitions and the equivalence relationships between them and the terms of the lists of the library, the VOCAUSP and the default list of ARQUIGRAFIA.

The tags that belong to the users' semantic universe surely enrich the vocabulary, showing how they think and retrieve the information. The result constitutes the first version of a collaborative controlled vocabulary that acts as a suggestion of terms for all users, maintaining the possibility of inserting new tags later.

Besides the user guarantee, an indexing tool needs to obtain the literary guarantee too. The literary guarantee of the ARQUIGRAFIA controlled vocabulary is based on research of the terms in dictionaries, glossaries, encyclopedias and specific terminologies and in the terminological method.

The terminological method consists in defining the term from the characteristics of the concept and its definition in the domain based in the works of Dahlberg (1978, 2009, 2011, and 2014), Cabré (1995), the ISO 25964-1 (2011) and their application in the construction of the controlled vocabulary of arts by Lima, Costa, and Guimarães (2017). The characteristics indicate the extension and the intension of the term, the extension being understood as the class of all things that the term applies to, and the intension as the properties that an object must have to be in the scope of the term definition. Each true statement about a certain property or characteristic of an object delivers a knowledge element about it. The sum of the statements about such an object forms the whole of characteristics of its concept. These statements also form its definition, such as in Dahlberg's example: a museum is a public building; it serves for the exhibition of objects; it possesses collections of certain fields of study; it presents collections thematically; it has certain times for visitors and controls visitors (in general) by means of tickets (Dahlberg, 2009).

So far, it has been possible to standardize and define 1,300 terms which have been included in five categories: form (building type), function (use of the building, past or present), materials (materials used in the construction), technique (construction technique used) and history. These categories make up the first level of the controlled vocabulary structure and after that we have been working to establish its logical and ontological relations based on the relationships between their characteristics.

This list was shared in a Google Drive spreadsheet (Figure 5) with students and teachers from the research group containing: definition; source (dictionary or thesaurus from which the definition originated); proposed definition (based on characteristics and according to terminological method); references and synonyms; if the term is in the USP Controlled Vocabulary; hierarchy in USP Vocabulary; possible relationships in the ARQUIGRAFIA vocabulary; suggested hierarchy and consistency with the indexed images.

Finally, these relations are established by using the terminological procedures and the procedures for the construction of controlled vocabularies indicated in the ISO standards (ISO, 2000; 2011). At the same time, the categories and terms were included in a mind map software (Figure 6) to visualize the hierarchical relations



Research Paper

1	LISTA DE TAGS	Definição	Fonte	Definição Proposta	REMISSIVAS / SINÓNIMOS	VOCAUSP	Hierarquia VOCAUSP	Relações possíveis no Vocabulário Arquigrafia	Hierarquia st
3	Acesso ao metrô	[ACESSO] - O ato de chegar ou entrar [METRÔ] - 1. Sistema de estrada de ferro, geralmente subterrâneo, destinado ao transporte rápido de passageiros em meios urbanos.	Priberam	Espaço interno ou externo por onde é possível entrar no Metrô (VL)		não			
4	aço corten	É um tipo de aço patinável, ou seja, com pequenas adições de elementos, como cobre, fósforo, níquel e cromo, que em condições ambientais contribuem para a formação de uma película que protege esses aços de ação corrosiva na atmosfera oxidante de muitos ambientes urbanos. Apresenta-se com uma coloração avermelhada.	E-civil	Tipo de aço patinável (OP)		não		Não tem no Arquigrafia	
5	aço inoxidável	Aço resistente a oxidação e corrosão pela inclusão na sua composição de alto teor de cromo. É adequada sua utilização em locais próximos ao mar, pois resiste bem à maresia. É empregado em diversos elementos da construção, como ESQUADRIAS, CORRIMÃOES e LUMINÁRIAS. É também chamado aço-cromo.	Albermaz	Tipo de aço resistente a oxidação, também chamado de aço-cromo (OP)	Aço cromo	sim	Engenharia metálica > Metalurgia extrativa > Aço > Aço inoxidável	Não tem no Arquigrafia	

Figure 5. Spreadsheet of the controlled vocabulary under construction.

Source: ARQUIGRAFIA research team.

and make decisions about the position of the term in the controlled vocabulary structure.

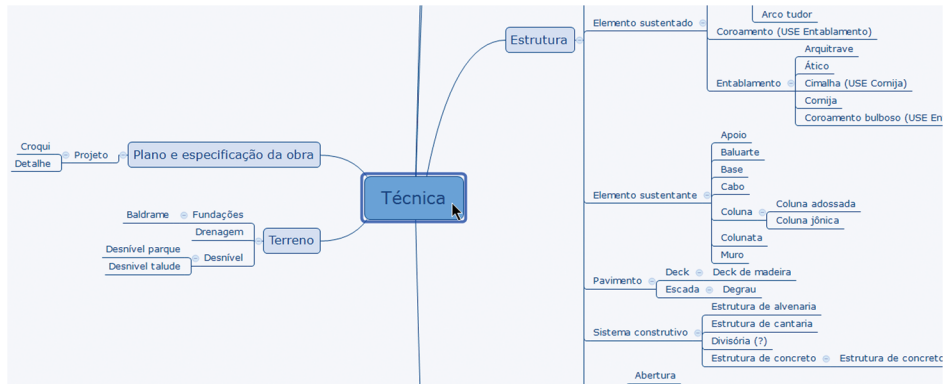


Figure 6. Mind map by categories.

Source: ARQUIGRAFIA research team.

Periodically, it will be necessary to harvest new tags uploaded by the users and submit them to this process of standardization for further inclusion in the controlled vocabulary, avoiding synonymy (two words with the same meaning) and polysemy (a word with several meanings), contributing to the improvement of information retrieval as shown in Figure 7.

For the consistency of the indexing, it is necessary to have knowledge of the subject area, to analyze the characteristics of the support and its contents, and to develop clear rules for the use of the controlled vocabulary.

ARQUIGRAFIA’s indexing policy indicates that it is advisable to tag the materials used in the construction of the work that are visible in the foreground of the image



Research Paper

albums, and insertion of contributions (Figure 8), which encourage the users to review the cataloguing of the images by means of gamification processes.

The screenshot displays the ARQUIGRAFIA social interaction resources interface. It features a main post for "Residência Manuel Antônio Mendes André" with a photograph of a building facade. The post includes various metadata fields: "Inserido em: zsmzsm", "Autor(es) da Imagem: Index LDs", "Data da Imagem: 18 de junho de 2018", "Data de conclusão da obra: Ano de 1966", and "Endereço: Rua Coronel Arthur Godói, 185, Vila Mariana - São Paulo SP - Brasil". A map shows the location. The interface also features a comment section and a "social mídias" callout. Other callouts highlight "notifications albums messages contributions" and "georeferencing".

Figure 8. ARQUIGRAFIA social interaction resources.

Source: <http://www.arquigrafia.org.br/>

The User-Centered Design (UCD) procedures were included based on gamification elements aimed at a greater user engagement especially with interface elements related to collaboration, such as: notifications; posts; complementing information about images and comments; the possibility to follow and be followed by other users.

6 Summing up

Due to the collaborative nature and the characteristic of enriching its metadata with user-generated content, ARQUIGRAFIA fits in the definition of the Digital Humanities (DH), understood as a new epistemological challenge where we have the articulation of knowledge and methods used in the human sciences with the digital world. At the same time, it faces challenges related to the sustainability of a system in continuous growth, which includes development and programming;



storage and preservation; management and insertion of data and images. The ending of the digitization of the set of 42,000 images in addition to the digital preservation of the information uploaded into the system still brings us to a challenge regarding its digital curatorship, as well as issues related to copyright and the obtaining of licenses for insertion in ARQUIGRAFIA.

Currently, the ARQUIGRAFIA research team deals with new short and medium term objectives such as:

- the implementation of a first version of a moderation system integrated with gamification;
- the development and dissemination of the +GRAFIA template;
- the studies related to the plastic-spatial qualities of the binomials arranged as semantic differentials, seeking to define an image evaluation model based on visual similarities that may be useful for information retrieval;
- the engagement of an interactive community around the images and their information, seeking long-term sustainability for the system.
- the evaluation of the descriptive, technical and administrative metadata to make them more interoperable in a web collaborative environment.
- the improvement of the controlled vocabulary as a visual knowledge organization system

Finally, we must conduct usability studies to evaluate the current beta version and redesign the system from the critical observations made by the users, to expand the iconographic base of ARQUIGRAFIA, including digital video, drawings and other audiovisual resources, as well as to deepen the research into new relevant topics and future developments.

Author contributions

Vânia Mara Alves Lima (vamal@usp.br) conducted the knowledge organization literature review, analyzed the available materials generated by the stakeholder-driven research project and wrote the paper. Cibele Araújo Camargo Marques dos Santos (cibeleac@usp.br) analyzed the available materials generated by the stakeholder-driven research project and wrote the paper. Artur Simões Rozestraten (artur.rozestraten@usp.br) analyzed the available materials generated by the stakeholder-driven research project and reviewed the paper.

References

- Angus, E., Stuart, D., & Thelwall, M. (2010). Flickr's potential as an academic image resource: An exploratory study. *Journal of Librarianship and Information Science*, 42(4), 268–278.
- Bradley, P. (2011). From Flickr to Playbills: How to find the right images. *CILIP UPDATE with gazette*, (Jul), 23–23.



- Cabré, M.T. (1995). La terminologia hoy: Concepciones, tendências y aplicaciones. *Ciência da Informação*, 24(3), 289–298.
- Dahlberg, I. (2009). Brief communication: Concepts and terms—ISKO’s major challenge. *Knowledge Organization*, 36(2/3), 169–177.
- Dahlberg, I. (2011). Brief Communication: How to improve ISKO’s standing: Ten desiderata for knowledge organization. *Knowledge Organization*, 38(1), 69–74.
- Dahlberg, I. (2014). Brief communication: What is knowledge organization? *Knowledge Organization*, 41(1), 85–91.
- Dahlberg, I. (1978). Teoria do conceito. *Ciência da. Informação*, 7(2), 101–107.
- Font, F., Serra, J., & Serra, X. (2013). Folksonomy-based tag recommendation for collaborative tagging systems. *International Journal on Semantic Web and Information Systems*, 9(2), 1–30.
- Gasser, M. (2017). Research partnerships, user participation, extended outreach—some of ETH Library’s steps beyond digitisation. DH. Opportunities and Risks. *Connecting Libraries and Research*, Berlin, Germany: fihal-01660814f.
- Guerreiro, D., & Borbinha, J.L. (2014). Humanidades digitais novos desafios e oportunidades. *Revista Internacional del Libro, Digitalización y Bibliotecas*, 2(2), 63–75.
- Gil-Leiva, I. (2002). Consistence of document indexation involving novel indexers. *Anales de Documentación*, 5, 99–111.
- Hodge, G. (2000). *Systems of knowledge organization for digital libraries: Beyond traditional authority files*. Washington, DC: The digital Library Federation.
- Hojrland, B. (2008). What is knowledge organization (KO)? *Knowledge Organization*, 35(2/3), 86–101.
- Hughes, A.V., & Rafferty, P. (2011). Inter-indexer consistency in graphic materials indexing at the National Library of Wales. *Journal of Documentation*, 67(1), 9–32.
- International Standard Organization (2011). *ISO 25964-1: Information and documentation—thesauri and interoperability with other vocabularies—part 1—thesauri for information retrieval*. Geneva: ISO.
- Lancaster, F.W. (2004). *Indexação e resumos: Teoria e prática*. (2nd ed.). Brasília: Briquet de Lemos Livros.
- Lima, V.M.A., Costa, I.G., & Guimarães, M.O. (2017). A Organização do Conhecimento no domínio das Artes: o fazer terminológico na gestão do vocabulário controlado. Recife: Ed. UFPE.
- Lima, V.M.A., Rozestraten, A.S., Santos, C.A.C.M., Marques, E.A., & Sampaio, L.A. (2016). Arquigrafia: um repositório digital de imagens em ambiente colaborativo web. *RBBB. Revista Brasileira de Biblioteconomia e Documentação (Online)*, 12, 103–107.
- Moreiro González, J.A. (2011). *Linguagens documentárias e vocabulários semânticos para a web: elementos conceituais*. Salvador: EDUFBA.
- National Information Standards Organization. (2010). *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. Baltimore: NISO.
- Osgood, C.E. (1990). The nature and measurement of meaning. In Osgood, C.E., & Tzeng, O.C.S. (Eds.). *Language, meaning and culture: The selected papers of C. E Osgood*. New York: Praeger Publishers.
- Rafferty, P. (2018). “Tagging”. *Knowledge Organization*, 45(6), 500–516.
- Rozestraten, A.S., Andrade, B.M., & Figueiredo, F.G. (2018). *Manual de Procedimentos Técnicos do Projeto ARQUIGRAFIA*. 2nd. Ed. São Paulo: FAU/USP.



- Rozestraten, A.S., Chou, A., Valentini, S., Gerosa, M.A., Reganati, G., Valente, C., & Claro, R. (2013). Smart audio city guide: Um sistema colaborativo para apoio ao deslocamento urbano de pessoas com deficiência visual. Proceedings of the SBSC Conference (pp. 175–178). Porto Alegre: Brazilian Computer Society.
- Rozestraten, A.S., Lima, V.M.A., & Santos, C.A.C.M. (2017). ARQUIGRAFIA: Digital images in the collaborative environment on the Web. IFLA Satellite Meeting: Digital Humanities. Berlin. Hague: IFLA.
- Rozestraten, A.S., Martinez, M.L., Gerosa, A.M.A., Kon, F., & Santos, A.P.O. (2010). Rede social ARQUIGRAFIA-Brasil: Estudos iconográficos da Arquitetura Brasileira na web 2.0. Seminário Nacional de Documentação do Patrimônio Arquitetônico com o uso de Tecnologias Digitais. Salvador: UFBA.
- Santos, L., & Santos, C.A.C.M. (2017). Arquigrafia: Um projeto de indexação de fotografias no meio digital, São Paulo. In: 25. Simpósio Internacional de Iniciação Científica e Tecnológica da USP. São Paulo: USP.
- Wölfflin, H. (1950). Principles of art history: The problem of the development of style in later art. New York: Dover Publications.



This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

