

Machado de Assis in the digital collections of Brazilian heritage institutions

Carlos H. Marcondes

Graduate Programa of Knowledge Management and Organization, Minas Gerais Federal University

Belo Horizonte, MG, Brazil

ch_marcondes@id.uff.br

Abstract— Digital heritage collections in Brazil hold great potential for culture, education, creative industries, and citizenship. They are complementary and offer many synergies and thematic overlaps. However, they are not integrated; they cannot be accessed from a single access point; it is not possible to capitalize on their synergies and complementarities; the reuse of their objects is problematic. The aim of this study is to create a so called a Thematic Narrative Path, about Machado de Assis, a Brazilian writer. A Thematic Narrative Path is cross-cutting theme across digital collections. It is formed by digital objects from collections of different Brazilian heritage institutions. It is represented using RDF, TriG, maned graphs, and linked data technologies.

Keywords— *Digital collections. Heritage data, Digital Humanities, Reuse. Interoperability, Open Data*

I. INTRODUCTION

Digital heritage collections published on the Web have additional potentialities that physical collections do not have. They have - an increased range, i.e., can be accessed from everywhere, by everyone, at any time; and - plasticity, i.e., their digital objects, once published on the web and having an IRI assigned, became independent from any hosting institution and can be combined and mixed, forming unsuspected, new and richer resources.

The latest SURVEY ON THE USE OF INFORMATION AND COMMUNICATION TECHNOLOGIES IN BRAZILIAN CULTURAL INSTITUTIONS [1] reports that a significant number of Brazilian archives, libraries and museums have digitalized their collections, many of them are also available online.

These digital collections hold great potential for culture, education, creative industries, and citizenship. They are complementary and offer many synergies and thematic overlaps. Digital collections are also one of the greatest inputs of the so-called Digital Humanities [2], [3], [4], viewed as ‘Collections as Data’ [5]. Digital heritage objects (DHO) are also considered as ‘epistemic objects’ [6], they are key resources to education. Recently, the concept of digital humanities has been evolving broadly into cyber humanities [7].

However, such collections are not integrated; they cannot be accessed from a single access point; it is not possible to capitalize on their synergies and complementarities; the reuse of their objects is problematic; there are not common infrastructure

or services to support their use and integration. They are ‘information silos’ [8]. The National Culture Plan, 2010-2020 does not mention any public policy concerning interoperability or a common metadata format for Brazilian digital collections [9].

The aim of this study is to create a so called a Thematic Narrative Path (TNP), about Machado de Assis, a Brazilian writer. A TNP is a cross-cutting theme across different digital collections. It is formed by digital objects from collections of different Brazilian heritage institutions. A TNP is a story told by a digital curator, in which the characters are mostly DHOs. A proof of concept is created to show - that the digital collections of these institutions are thematically complementary, have many overlaps; - and how difficult is to identify and select DHO from these collections to integrate the TNP due to the lack of standards and interoperability mechanisms among them.

Once selected, integrated and interconnected within a TNP, DHOs constitute new digital resources with added value, increasing their potential for reuse for Culture, Science, Education and Citizenship. With such aim, DHO from these collections were identified, selected, and collected. Brazilian digital heritage collections don’t have aggregator nor meta-search services where users can submit a query to different collections; so, Google was used to identify collections and candidate DHOs within them. Once selected, such DHO were integrated, and interconnected into a TNP, represented with RDF technologies, linked data, named graphs, and knowledge graphs. The results show that the Brazilian digital collections of Brazilian heritage institutions have great potential that is not used because interoperability technologies and standards are not adopted and, mainly, due to the lack of public policies.

This paper is organized as follows: After the introduction, Section II discusses the methodological procedures used. Section III outlines the development of the TNP. Section IV describes the MIC model. Section V describes the implementation of TNP modeled according to MIC, a conceptual model previously developed [10], in RDF and named graphs. Section VI presents remarks on this research and its future developments.

II. METHODOOGY

This is an ongoing research project. It builds on previous research results related to the management of digital heritage collections. The research aims to evaluate the interoperability

and extraction capabilities offered by the collection management systems of heritage institutions in Brazil. This paper presents a proof of concept that addresses this issue. After collecting information about Machado de Assis, primarily from his Wikipedia page, Google was used to identify digital records and collections relevant to the research. These records were then manually extracted. The MIC (Integrated Model of Curationship), developed in previous research, was then used to create a TNP focusing on Machado de Assis, Rio de Janeiro and slavery in Brazil. The conceptual model was developed using Visual Paradigm online tools [11]. Such TNP is serialized in RDF [12] and named graphs [13]. Online tools provided by isSemantic [14] were used to validate the RDF serialization and convert it into triples. GraphDB was then used to load the set of triples and provide a visualization of the knowledge graph. Data collected, the RDF serialization are available in a github repository [15].

III. THE DEVELOPMENT OF THE THEMATIC NARRATIVE PATH

Today digital curation has a broad meaning encompassing activities from digital preservation as promoted by the Digital Curation Centre (DCC) [16] to promoting access and reuse. During the 20th century, curatorship in culture shifted its focus from collections to exhibitions [17]. Currently such exhibitions are gaining a global dimension. Examples are “Da Vinci: Mechanics of Generous” at the Science Museum in London, curated by engineer Jim Bennett, “Da Vinci” at the Louvre Museum, Paris, which was principally curated by Vincent Delieuvin and Louis Frank, and “Human Bodies”, curated and catalogued by Prof. Gunther Von Hagens. With LOD technologies these exhibitions can now become new digital, authorial, and permanent resources.

The instrument for disclosing and registering an exhibition is a catalog, a kind of special work. A catalog is a typical publication that permanently materializes the curatorship of a collection or exhibition. The contents of an exhibition catalog are descriptions, analysis, or comments about sets of heritage objects and why they are put together and related. An example of a printed catalog is the one of the “Da Vinci” exhibition at the Louvre Museum, Paris [18]

In such a catalog, a curator proposes a new look at that set of objects, contextualizes them, reveals new relationships, often unusual, between them, and proposes a narrative about those objects. Curating an exhibition is an authorial work.

Today it is possible to develop curatorial work on collections of digital objects. Such curatorial work is expressed by a special kind of publication, the digital exhibition catalog. A curatorial work is thus a kind of digital document which somehow mimics the structure and purposes of an exhibition catalog. MIC was developed based on the structure of the previously mentioned catalogs. The Thematic Narrative Path Machado de Assis in the digital collections of Brazilian Heritage Institutions is modeled according to the MIC as digital Curatorial Work.

The creation of a TNP requires a screenplay, developed by a digital curator with extensive knowledge of the digital collections and their digital objects that may be of interest to the exhibition. In this case, as a proof of concept, the Wikipedia

page about Machado de Assis [19] is used as a screenplay, to guide the selection of DHOs. A total of 08 DHOs were selected to integrate the TNP; in the following Table, the DHOs selected are shown.

TABLE I. DHOs SELECTED

DHO, Description	Institution's Collection	IRI attributed
Machado de Assis, Brazilian writer	Wikidata	<http://exemplos.org/000001>
Candido Portinari, Brazilian painter	Wikidata	<http://exemplos.org/000002>
Rio de Janeiro, the city where Machado de Assis was born in 1839	Wikidata	<http://exemplos.org/000003>
Slavery in Brazil	Wikidata	<http://exemplos.org/000004>
Posthumous Memories of Braz Cubas	Portinari Project	<http://exemplos.org/000005>
Caterpillar, drawing by Cândido Portinari to illustrate Posthumous Memoirs of Braz Cubas	Portinari Project	<http://exemplos.org/000006>
Posthumous Memoirs of Braz Cubas, a movie inspired by Machado de Assis's novel	YouTube	<http://exemplos.org/000007>
Letter by Machado de Assis commenting a law concerning the abolition of slavery in Brazil	Brazilian National Digital Library	<http://exemplos.org/000008>

There are several related workd with focus on modeling narrative as a sequence of events describing a specific theme or domain, as in [20], [21], [22], [23].

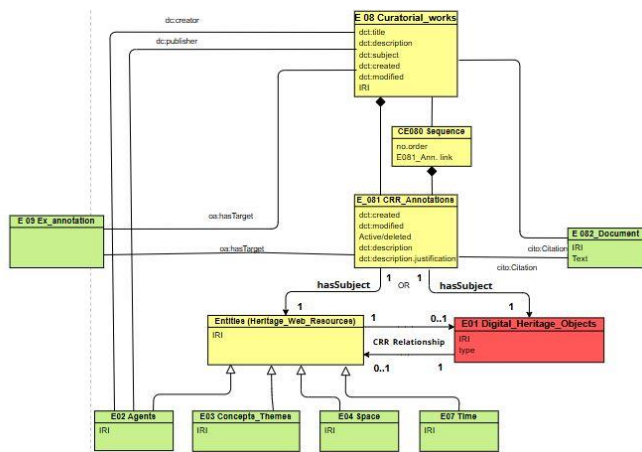
IV. THE MIC MODEL

There are different conceptual models in the heritage domain as IFLA LRM [24], CIDOC CRM [25], BIBFRAME [26], EDM [27]. MIC is not an alternative model to these. The aim of MIC is not to represent heritage objects and the processes they are involved in.

The MIC project aims to represent a collection of digital heritage objects (DHOs) published by different heritage institutions and represented according to different conceptual models. These DHOs are selected and compiled by a digital curator to create a digital exhibition catalogue, regardless of the institution hosting them. Such DHOs are interrelated with each other and with entities such as creators, publishers, artistic movement and styles, periods and events, and places, represented as digital resources. MIC uses linked data technologies to achieve these objectives.

The core of MIC are the digital DHOs selected by a digital curator to integrate a TNP. According to the MIC, a Curatorial Work, in this case, the TNP *Machado de Assis in the digital collections of Brazilian heritage institutions*, has as constituent parts Annotations and DHOs. As a rule, for each selected object, the digital curator makes an Annotation. An Annotation is a curator's commentary or analysis of a DHO and its relationships to other DHOs or to its creator, artistic movement, periods, concepts, places, etc.

An Annotation is the reification of a relationship between two DHOs, or between a DHO and an external resource. MIC also enables Annotations without referring to DHOs, such as those about external entities, i.e., any web resources. In MIC Annotations are modeled using the W3C Web Annotation Data Model [28]. According to the Web Annotation Data Model an Annotation is a relationship expressed by a Target predicate, specifying the URL of the annotated resource, and a Body predicate, specifying the annotation made about the Target.



The MIC has the following Entities: Curatorial Works (E 08), CRR Annotations (E 081), Digital Heritage Objects (E 01), and Entities, Heritage Web Resources, identified in the class diagram by the yellow color; an external Entity has as specific subclasses Agents (E02), Concepts/Themes (E03), Space (E04) and Time (E05), identified by the green color.

cited by the digital curator to justify or support his/her argumentation

The constituents of a TNP are E081_CRR_Annotations. Each one is implemented as a specific named graph. Each one is identified by a specific IRI and connected to the whole Curatorial Work by the property `dc:hasPart`. Each E081_CRR_Annotation is a reification of a **relationship between resources**. Within each specific E081_CRR_Annotation's named graph there is an **Open Annotation Data Mode**, `oa:target`, specifying the resource being annotated and a `oa:body`, connecting the E081_CRR_Annotation's named graph to a specific named (sub) graph in which is specified a **relationship of the previous specified target resource with another one**.

TABLE II. ANNOTATIONS

	Cubas related to the movie inspired by it
<http://exemplos.org/crr/Machado_Letter_Abolition_slavery>	Machado de Assis related to a letter in which he comments about the abolition of slavery in Brazil

An Annotation is a named graph. Each has an `oa:target` specifying the resource being annotated and an `oa:body` specifying the IRI of another named graph where specific relationships with the `oa:target` resource (from CRR Vocabulary and from other vocabularies) are specified. Thus, Annotations are reifications of relationships. An example of a 081_CRR_Annotation coded in RDF, named graphs is shown below.

```
{
<http://exemplos.org/crr/Machado_Letter_Abolition_slavery>
  rdf:type mic:E081_Annotation;
  oa:target <http://exemplos.org/000008>; -> The letter
  oa:body
    <http://exemplos.org/crr/Machado_Letter_Abolition_slavery_b
    ody> .}
{
<http://exemplos.org/crr/Machado_Letter_Abolition_slavery_b
ody>
  rdf:type oa:body;
  dct:creator <http://exemplos.org/000001> -> Its creator
  skos:related <http://exemplos.org/000004> .} -> Slavery in
Brazil
```

DHOs from different Brazilian heritage institutions were selected to develop the TNP Machado de Assis in the digital collections of Brazilian heritage institutions. Such DHOs are shown in the table. As coded in RDF, for each DHO is created and record, consisting in an IRI, a RDF type within the MIC, i.e., `mic:E01_Digital_Heritage_Object`, a TOP class, a `crr:181_Incollection`, specifying the heritage institution page from where the DHO was selected, and a `dct:identifier.URLinInstitution`, specifying the original DHO URL within the heritage institution collection. The creation of a specific record of each DHO is justified as few of the DHOs collected have an URIs and thus enable that additional links can be made to the DHO and even to the Machado de Assis in the digital collections of Brazilian heritage institutions TNP, enhancing it; however the original DHO URL in the heritage institution collection is preserved, thus guaranteeing a path to the original digital heritage record. An example of the named (sub) graph encoding a DHO is shown below.

```
1- {
2- <http://exemplos.org/000005>
3-  rdf:type mic:E01_Digital_Heritage_Object;
4-  skos:label "Posthumous Memoirs of Brás Cubas, novel by
5-  Machado de Assis";
6-  mic:crr_181_Incollection <https://www.portinari.org.br/>;
7-  dct:identifier.URLinInstitution
8-  <https://www.portinari.org.br/acervo/bibliografico/25019/
9-  memorias-posthumas-de-braz-cubas/>.
10- }
```

The TNP *Machado de Assis in the digital collections of Brazilian Heritage Institutions* comprises three content blocks: the curatorial work itself, the annotations, and the objects referred to in the annotations. Serializing these in RDF TriG named graphs proved satisfactory for separating each content block.

Methodologically, a curatorial work can be developed bottom-up according to the MIC framework, by creating the three building blocks that comprise it as named graphs:

1. A named graph for each DHO or external entity of interest to the specific curatorial work.
2. A named graph for each annotation, i.e. for each relationship between each DHO or entity and another heritage object or entity. In each annotation, the property `mic:hasSubject` relates a specific DHO or entity to another.
3. A named graph containing a set of `dct:hasPart` relationships relating the curatorial work to each of the annotations that comprise it.

The TNP Machado de Assis in the digital collections of Brazilian heritage institutions generates a set of RDF triples which are loaded GraphDB triplestore. Within GraphDB the resulting knowledge graph can be visualized, transversed, and navigated. The following Figure shows a partial view of the TNP Machado de Assis in the digital collections of Brazilian heritage institutions knowledge graph.

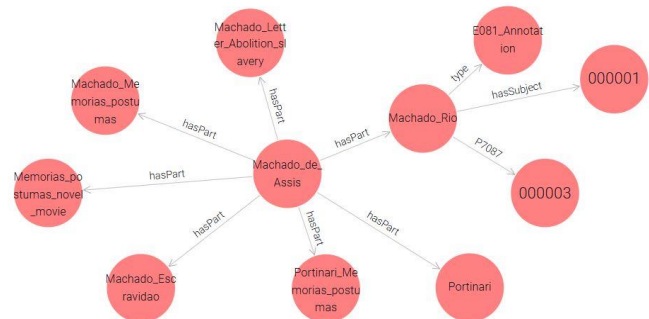


Fig. 2. The TNP Machado de Assis knowledge graph

In turn, the following Figure shows a DHO, the novel *Posthumous Memoirs of Brás Cubas* and its relationships.

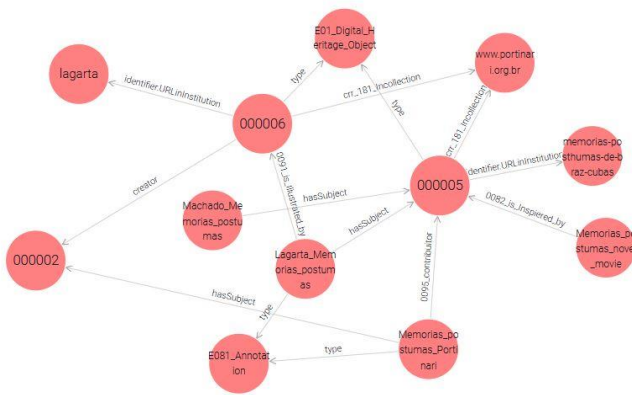


Fig. 3. A DHO and its relationships

Representing relationships between DHOs as annotations (which reify these relationships) makes more difficult to identify the DHOs that are related to one another. The following SPARQL query returns which DHOs as mic:hasSubject of annotations are related to other DHOs.

```
'PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX mic: <http://exemplos.org/mic/>
PREFIX crr: <http://exemplos.org/crr/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core/>
SELECT ?DHO1 ?Relation ?DHO2 WHERE
{?s mic:hasSubject ?DHO1.
?s ?Relation ?DHO2. } ,
```

However, this modelling decision emphasizes the structure of curatorial work as an interpretative narrative about the relationships between DHOs.

VI. FINAL REMARKS

MIC and its serialization in RDF, TriG, named graphs and linked data is conceived to serve as input to web exhibition authoring tools [32]. GraphDB provides a graphic interface through where the knowledge graph can be explored.

TNP implementations could form part of infrastructures such as the future digital repository of the INBCM [33] - National Inventory of Museum Cultural Assets - Such TNP would interrelate the digital objects within the repository with each other and with those outside it, thereby promoting the objects and the institutions that hold them. This would be an added-value service for both the future INBCM and Brazilian heritage institutions.

However, there is a need for more intuitive content management systems to be used by digital curators and teachers in developing authorial thematic narrative paths to be used in classes, lectures, and conferences. Such users need a content management system for hosting the thematic narrative paths they developed. And so do the students, the end users of the TNPs. Also, as a proof-of-concept future TNPs will need a

development methodology, and infrastructure as an authoring tool environment where cultural heritage curators can create their TNPs. Such development methodology should be customized to their use, simple, intuitive, and assisted, as cultural heritage curators may not be familiarized with technological issues.

This is an initial proposal, ongoing research. Future developments will include the evaluation and validation of the model by its preferred users, the digital curators, and identifying potential content management platforms where TNPs can be stored by their authors and accessed by end users. Some initial experiments concerning such questions can be seen in [34] using the Tainacan platform [35].

ACKNOWLEDGMENT

This work was supported by the Brazilian agencies CAPES (Financing Code 001) and CNPq under Grant n. 303379/2022-7.

REFERENCES

- [1] Comitê Gestor da Internet no Brasil, Pesquisa sobre o uso das tecnologias de informação e comunicação nos equipamentos culturais brasileiros: TIC Cultura 2022. São Paulo, Brazil: CGI.br, 2023
- [2] T. G. Padilla and D. Higgins, "Library collections as humanities data: The facet effect," *Public Services Quarterly*, vol. 10, no. 4, pp. 324–335, 2014.
- [3] K. Fenelon, "Modeling digital humanities collections as research objects," in *Proc. ACM/IEEE Joint Conf. Digital Libraries (JCDL)*, 2019, pp. 138–147. doi: 10.1109/JCDL.2019.00028
- [4] M. L. Zeng, "Semantic enrichment for enhancing LAM data and supporting digital humanities," *El Profesional de la Información*, vol. 28, no. 1, e280103, 2019. doi: 10.3145/epi.2019.ene.03
- [5] B. C. G. Lee, "The 'Collections as ML Data' checklist for machine learning and cultural heritage," *J. Assoc. Inf. Sci. Technol.*, 2023.
- [6] K. Tybjerg, "Exhibiting epistemic objects," *Museum & Society*, vol. 15, no. 3, pp. 269–286, 2017. [Online]. Available: <https://pdfs.semanticscholar.org/a152/1b29555cb7982794a6c80a1e3504ba2d8782.pdf>
- [7] Giovanni Adorni, and Emanuele Bellini. "Towards a Manifesto for Cyber Humanities: Paradigms, Ethics, and Prospects." 2025 IEEE International Conference on Cyber Humanities (IEEE-CH). IEEE, 2025.
- [8] A. Shahrokni and J. Söderberg, "Beyond information silos," in *Proc. BigMDE Workshop*, Ottawa, Canada, 2015, 63.
- [9] Brasil, Ministério da Cultura, Secretaria da Diversidade Cultural, Análise e avaliação qualitativa das metas e o monitoramento do Plano Nacional de Cultura (PNC). Salvador, Brazil: Universidade Federal da Bahia (UFBA), 2018, 256 p., ISBN: 978-85-8292-185-2.
- [10] C. H. Marcondes, "A curatorial model for digital heritage collections as linked open data and named graphs," *Scire*, vol. 29, no. 2, pp. 85–95, Jul.–Dec. 2023.
- [11] Visual Paradigm, "Visual Paradigm Online Tools." [Online]. Available: <https://online.visual-paradigm.com/>
- [12] W3C, "RDF 1.1 TriG: RDF Dataset Language," W3C Recommendation, Feb. 25, 2014. [Online]. Available: <https://www.w3.org/TR/rdf12-trig/>
- [13] J. J. Carroll et al., "Named graphs," *J. Web Semantics*, vol. 3, no. 4, pp. 247–267, Dec. 2005. doi: 10.1016/j.websem.2005.09.001
- [14] isSemantic, "isSemantic." [Online]. Available: <https://issemantic.net/>
- [15] GitHub, "GitHub repository." [Online]. Available: <https://github.com/carlos-marcondes/Thematic-Narrative-Paths---MIC---Integrated-Model-of-Curationship>
- [16] Digital Curator Centre.. "DCC". [Online]. Available: <https://www.dcc.ac.uk/>
- [17] P. O'Neil, "The curatorial turn," in *Issues in Curating Contemporary Art and Performance*, J. Rugg and M. Sedgwick, Eds. Bristol, U.K.: Intellect Books, 2007, pp. 13–28.

- [18] Musée du Louvre, “Leonardo da Vinci Exhibition,” Paris, France, 2019. [Online]. Available: <https://www.louvre.fr/en/what-s-on/exhibitions/leonardo-da-vinci>
- [19] Wikipedia, “Machado de Assis.” [Online]. Available: https://pt.wikipedia.org/wiki/Machado_de_Assis
- [20] P. Mulholland, A. Wolff, and T. Collins, “Curate and storyspace: An ontology and web-based environment for describing curatorial narratives,” in Proc. 9th Extended Semantic Web Conf. (ESWC 2012), Heraklion, Greece: Springer, 2012, pp. 748–762. doi: 10.1007/978-3-642-30284-8_54
- [21] M. Daquino and F. Tomasi, “Historical Context Ontology (HiCO): A conceptual model for describing context information of cultural heritage objects,” in Proc. Metadata and Semantics Research Conf., Cham, Switzerland: Springer, 2015, pp. 424–436. doi: 10.1007/978-3-319-24129-6_37
- [22] G. Renda and M. Daquino, “Storytelling with linked open data,” in La Memoria Digitale, AIUCD, 2023, pp. 164–167.
- [23] C. Meghini, V. Bartalesi, and D. Metilli, “Representing narratives in digital libraries: The narrative ontology,” Semantic Web, vol. 12, no. 2, pp. 241–264, 2021. doi: 10.3233/SW-200418
- [24] P. Riva, P. Le Boeuf, and M. Žumer, IFLA Library Reference Model: A Conceptual Model for Bibliographic Information. The Hague, Netherlands: IFLA, 2017. [Online]. Available: <https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017.pdf>
- [25] ICOM/CIDOC, “CIDOC CRM, Version 7.1.3,” 2024. [Online]. Available: https://www.cidoc-crm.org/html/cidoc_crm_v7.1.3.html
- [26] Library of Congress, “BIBFRAME Model Documentation,” Washington, DC, USA, 2025. [Online]. Available: <https://www.loc.gov/bibframe/docs/index.html>
- [27] Europeana Foundation, “Europeana Data Model (EDM).” The Hague, Netherlands. [Online]. Available: <https://pro.europeana.eu/page/edm-documentation>
- [28] W3C Web Annotation Data Model. [Online]. Available : <https://www.w3.org/TR/annotation-model/>
- [29] C. H. Marcondes, “Towards a vocabulary to implement culturally relevant relationships between digital collections in heritage institutions,” Knowl. Org., vol. 47, no. 2, pp. 122–137, 2020. doi: 10.5771/0943-7444-2020-2-122
- [30] Wikidata, “Slavery in Brazil.” [Online]. Available: <https://www.wikidata.org/wiki/Q136038554>
- [31] Wikidata, “Machado de Assis.” [Online]. Available: <https://www.wikidata.org/wiki/Q311145>
- [32] U. Klentien. “Development of virtual museum model for youth.” International Journal of Information and Education Technology 12.4 (2022): 313-317.
- [33] Instituto Brasileiro de Museus – Ibram, “Resolução Normativa Ibram nº 6, de 31 de agosto de 2021”, 2024. [Online]. Available: <https://www.gov.br/museus/pt-br/assuntos/legislacao-e-normas/outros-instrumentos-normativos/resolucao-normativa-ibram-no-6-de-31-de-agosto-de-2021>
- [34] Collection Machado de Assis nos acervos digitais das instituições brasileiras de Cultura, Memória e Patrimônio, Available <https://profimarcondes.org.br/teste/>
- [35] Um software livre, flexível e poderoso para criação de repositórios de acervos digitais em WordPress. Tainacan, 2026. [Online]. Available: <https://tainacan.org/>