Workflow with Emphasis in Technologies for the Construction of an Aggregator and a Dashboard of Brazilian Museums Digital Collections

Joyce Siqueira[1](✉), Dalton Lopes Martins[1], and Vinícius Nunes Medeiros[2]

[1] Universidade de Brasília, Brasília, DF, Brazil
joycitta@gmail.com, daltonmartins@unb.br
[2] Tainacan.Org, Brasília, DF, Brazil
vnicius.nm.ba@gmail.com

Abstract. The Brazilian Institute of Museums - Ibram, began the process of disseminating its museum collections on the web, through the Tainacan software, a result of the Collection Platform Project, carried out in partnership with the Federal University of Goiás and the University of Brasília. With the project's developments, a new information service is under study, the creation of Brasiliana Museums, an aggregator of cultural digital objects from museums managed by Ibram, which aims to offer a single search and data retrieval interface. Beyond that aggregator it is possible developed another new service, a dashboard for Ibram's managers, very important for an analytical thinking through graphical interfaces. This research this is configured as a case study, because is aimed at solving Ibram's data aggregation problem, in addition to being a qualitative, exploratory and descriptive research. This paper aims to present the development process of Brasiliana Museums and a dashboard, highlighting the technological tools used: 1, mapping origin museums; 2, automated collection of items the collections; 3, aggregation of collections and data transformation; 4, submission of items to Tainacan; 5, data integrate storage; 6, dashboard development and; 7, publication of the aggregated collections for the search and recovery of Brazilian museum items. The main technologies used in the prototype was the Elastic Stack, the new plugin, developed in Ruby programing language, and the Tainacan repository software. This result is the aggregation than more 17 thousand objects from 20 Ibram's museums, that allow in single interface to search and retrieve digital Brazilian museum items. The prototyping of the service reveals itself, therefore, as an efficient solution, both from the point of view of socialization and dissemination of digital objects and in the monitoring and monitoring of information from the collections through analytical panels. In addition to presenting a cheap and viable technological solution to be implemented.

Keywords: Aggregator · Elastic stack · Tainacan · Museums · Technologies

A. L. Pinto and R. Arencibia-Jorge (Eds.): DIONE 2022, LNICST 452, pp. 84–97, 2022.
https://doi.org/10.1007/978-3-031-22324-2_8

1 Introduction

The Brazilian Institute of Museums - Ibram is an autarchy of the Brazilian Federal Government, responsible for the National Policy on Museums and for the improvement of services in the sector, for example, increase in museums visitation, promotion of policies for the acquisition and preservation of collections, enable access to culture to a part of the population that would not be able to enjoy it, in addition, of administration of 30 museums (Gov.br 2022b).

In 2014, Ibram, in partnership with the Federal University of Goiás, through the Participatory Public Policy Laboratory, started the development of the project "Platform Collections: inventory, management and dissemination of museum heritage", to promote the fundamentals of a policy for digital collections, which resulted, in 2016, in the first version of Tainacan, a software free and open-source for the management and publication of digital collections. In 2018, the project, now entitled Tainacan Project, is being developed in partnership with the University of Brasília - UnB, through the Networks Intelligence Laboratory.

In 2022, 20 Ibram's museums they had already migrated the documentation of their collection to Tainacan and are available for access via the web (Gov.br 2022a). With this panorama and the growing number of migrated museums, the Laboratory has been carrying out several studies with the aim of creating a new information service, Brasiliana Museums, an aggregator of cultural digital objects from the collections of different museums, which aims to offer an interface unique search and retrieval of information using Tainacan both in the origin museums, that is, in the data providers, and in the own aggregator, as the final repository. Beyond that, with digital items from all museums integrated went too possible a dashboard was built for Ibram's managers.

In this paper, we present the workflow for data aggregation and dashboard construction, however, with emphasis on the technologies used to implement a pilot and experimental service for Ibram.

The steps taken to develop the prototypes of the aggregation service are presented. The main technological solutions used are the Elastic Stack (Beats, Logstash, Elasticsearch and Kibana) and the Tainacan digital repository, as well as the aggregation workflow, which has seven steps, the know: step 1, mapping origin museums, in Tainacan; step 2, automated collection of items the collections, with Beats - Filebeat; step 3, aggregation of collections and data transformation, with Logstash; step 4, submission of items to Tainacan, with Logstash plugin development in Ruby; step 5, data integrate storage, in Elasticsearch; step 6, dashboard development, with Kibana and; step 7, publication of the aggregated collections for the search and recovery of Brazilian museum items, with Tainacan.

The paper is organized as follows: Sect. 2 describes the methodological procedures; Sect. 3 presents the results of the study, Brasiliana Museum's aggregation process; Sect. 4 discusses the results; and finally, Sect. 5 identifies important aspects for further research.

2 Methodology

Qualitative research, with an exploratory and descriptive nature, which addresses the process of aggregating Ibram's digital museum collections and dashboard creation, configuring the research as a case study. As a technical procedure, bibliographic, exploratory (Siqueira and Martins 2020) and systematic review literature - SRL (Siqueira and Martins 2021) research was used in order to provide theoretical sustainability to the study.

In the studies mentioned above, 20 different international projects for the aggregation of cultural data were found, many of which deal with the aggregation for the Europeana Foundation, that is, different projects that use similar processes and technologies. In both studies were found some data aggregators, Table 1.

Table 1. Data aggregators

Data aggregator	Link	Current
Ariadne	https://ariadne-infrastructure.eu/	Yes
CHO repository	https://github.com/psu-libraries/cho	No
D-NET	https://github.com/dnet-team/dnet-basic-aggregator	No
META-SHARE	https://github.com/metashare/META-SHARE	No
MINT	https://github.com/mint-ntua/Mint	No
MoRe	http://more.dcu.gr/	No
Repox	https://doi.org/10.1007/978-3-642-04346-8_65	No
Supplejack	https://digitalnz.org/developers/supplejack	Yes

Siqueira and Martins (2021, 2020)

Not all articles Siqueira and Martins (2021, 2020) indicated which technologies are used. From the articles that did it, we found that there were several different technologies in use and not a more used one. In order to complement the answer, we verified whether it is current or not. For the purpose of this survey, we considered the updates that have occurred in the last two years. In general, it was not clear how these technologies are used or how they are integrated. Concomitantly, the documentation proved to be rather lacking in detail and with lengthy discussions on the subject. We identified the D-NET technology as more frequently; however, in the GitHub we observed that the last update repository was conducted four years ago, which makes it more difficult to adopt in new projects. We can observe that the tools developed specifically for aggregation are mostly outdated.

This is the international scenario. In Brazil, although there are initiatives that aim to integrate cultural collections, these are limited, isolated and require technical support and scalable governance models for the Brazilian territory (Martins et al. 2022). We could see that the simple replication of international solutions would not be possible, because, they are diversified and deal with the specificities of each location; there is not enough documentation available to try to implement something and there is no consensus on better solutions and/or technologies. That is, even if it were possible to replicate, it

would probably not be effective in the Brazilian context, since they are different realities, considering infrastructure, investment, specialized labor, among others (Siqueira and Martins 2021; Siqueira et al. 2021).

Looking for modern, free tools with an active community that would scalability to the process, but that were also relatively simple to deploy, we selected to carry out the aggregation of the data with the Elastic Stack[1], in addition to a plugin for Logstash developed by the authors, using the Ruby[2] programming language and the Tainacan[3].

We emphasize that this tool was chosen because it meets the minimum criteria we are looking for and there is no intention here to define it as the best, but rather as the most appropriate for the context of Brazilian museums.

With the development of the prototype, it was possible to prove that the Elastic Stack with Tainacan achieves excellent results, as it allows extremely fast queries, text search, horizontal scalability, is capable of handling petabytes of data and the model with the Elastic Stack proved to be of medium complexity and customization for a viable possibility of use and analysis of data or models of learning configurations.

Tainacan is a technological solution for creating digital collections on the Internet. It was designed to meet the reality of cultural institutions, allowing the management and publication of collections in an easy and intuitive way. It is a plugin and a theme from WordPress[4], which offers an administrative panel for, for example, the creation of user profiles, with different levels of access to the collections, as well as web pages for the communication of the collections (Tainacan.org 2021b).

The Elastic Stack is composed by four technologies tools: Beats, Logstash[5], Elasticsearch[6] and Kibana[7] for "get data from any source, in any format, to search, analyze and visualize everything in real time" (Elastic 2021a). The Filebeat[8] is an Elastic Beat, is a lightweight shipper for forwarding and centralizing log data. Installed as an agent on your servers, monitors locations that you specify, collects log events, and forwards them either to Elasticsearch or Logstash for indexing (Elastic 2021b). The Logstash is data collection engine with real-time can dynamically unify data from disparate sources and normalize the data into destinations of your choice. Cleanse and democratize all your data for diverse advanced downstream analytics and visualization use cases (Elastic 2021i). Elasticsearch is a distributed search and data analysis engine that stores data centrally for fast search. Finally, Kibana is a user interface for viewing Elasticsearch data and navigating the Elastic Stack (Elastic 2021d).

It is possible to develop new plugins for Logstash: add your own input, codec, filter, or output plugins to Logstash following the guidelines and best practices can help you build a better plugin (Elastic 2021c). The filter developed for Logstash, tainacan_submission,

[1] https://www.elastic.co/elastic-stack/.
[2] https://www.ruby-lang.org/about/.
[3] https://tainacan.org/.
[4] https://wordpress.org/.
[5] https://www.elastic.co/logstash/.
[6] https://www.elastic.co/elasticsearch/.
[7] https://www.elastic.co/kibana/.
[8] https://www.elastic.co/beats/Filebeat.

aims to retrieve the data processed by Logstash and send them to a new Tainacan installation, through its submission API.

The aggregation service and the dashboard, in principle, consists of digital collections from 20 museums, with 22 collections, resulting in 17,315 digital objects, Table 2.

Table 2. Digital museums and their respective numbers of digital objects online

N.	Online digital collections	Digital objects number
1	Museu Casa da Hera	1220
	Museu Casa da Hera (Indumentaria)	67
2	Museu Casa da Princesa	799
3	Museu Casa de Benjamin Constant	983
4	Museu Casa Histórica de Alcântara	631
5	Museu da Abolição	301
6	Museu da Inconfidência	4625
7	Museu das Bandeiras	401
8	Museu das Missões	90
9	Museu de Arqueologia de Itaipu	1040
10	Museu de Arte Religiosa e Tradicional	132
11	Museu de Arte Sacra da Boa Morte	783
12	Museu do Diamante	895
13	Museu do Ouro	126
14	Museu Histórico Nacional	847
	Museu Histórico Nacional (Moedas de Ouro)	1215
15	Museu Regional Casa dos Ottoni	463
16	Museu Regional do Caeté	243
17	Museu Regional São João Del Rey	328
18	Museu Solar Monjardim	77
19	Museu Victor Meirelles	237
20	Museu Villa Lobos	1812
Total items		**17315**

For this research, were used the Tainacan version 0.15 or higher. Aggregation Tainacan uses the latest version. Elastic Stack use version 7.9. For the entire process to work correctly, Tainacan must have a minimum version 0.15 and Elastic Stack, version 7.9.

3 Technologies Workflow or the Construction of an Aggregator and a Dashboard

Although it uses the same technologies, there are two different workflow paths, each for its own purpose. The path 1 refers to the creation of the dashboard and the path 2 to the creation of Brasiliana Museums. The Fig. 1 presents the architecture.

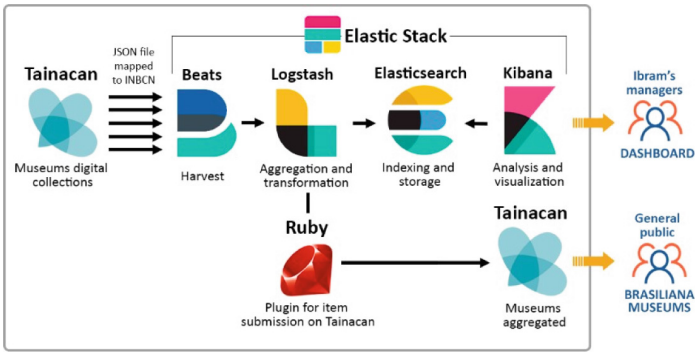

Fig. 1. Brasiliana museums workflow. Prepared by the authors (2022)

The paths are described below.

3.1 Tainacan: Mapping Origin Museums

By default, Tainacan offers alternative URLs for viewing data in Application Programming Interface: JavaScript Object Notation - JSON API formats; HyperText Markup Language - HTML and CSV files (Tainacan.org 2021a), enabling new applications to benefit from the data.

In Tainacan, from version 0.15 onwards, allows, through a plugin developed for Ibram, to map the metadata to the Ibram's standard entitled 'National Inventory of Museum Cultural Assets" – INBCM (BRASIL 2021).

The plugin titled "Tainacan Mappers IBRAM INBCM" was developed using the Hypertext Preprocessor - PHP[9] programming language, and is available on GitHub[10] under the GPLv3 license[11].

It is also available for installation on WordPress and offers a simplified interface for mapping a collection's metadata, in a one-to-one degree of equivalence, that is, a metadata from the originating museum can be mapped to an INBCM metadata. In elements where there is no exact equivalence, the metadata is disregarded. Regarding semantic mapping, it is performed by specialized personnel.

Beyond the INBCM's metadata are includes in JSON file: the pagination information (the items are presents in pages on Tainacan), the collection's identification number (id),

[9] https://www.php.net/.
[10] https://github.com/tainacan/tainacan-mappers-ibram-inbcm.
[11] https://www.gnu.org/licenses/gpl-3.0.pt-br.html.

Item URL in the museum of origin and the thumbnail URL, Fig. 2. All are important metadata for museums identification after aggregation.

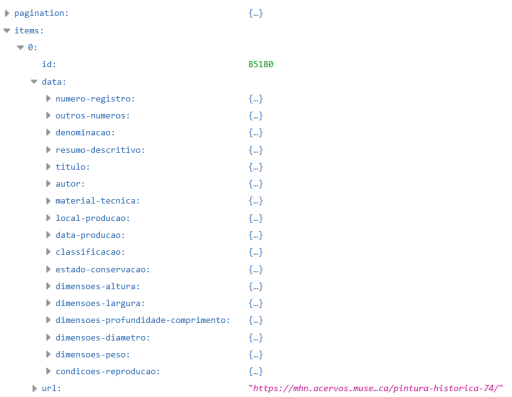

Fig. 2. Mapped JSON file. National Historic Museum (2022)

After the mapping, Tainacan makes available the JSON API, "Simple JSON: inbcm-ibram mapper"[12], which will be collected by Filebeat.

3.2 Beats - Filebeat: Automated Collection of Items from Museums

In the case of aggregation, only the metadata is collected, that is, documents or attachments are not considered. In the moment, ways are being studied to collect thumbnails of documents to provide users with a preview of them.

We emphasize that the interest is for the user to visit the museum of origin, through the access link to the item, and not just use the aggregator, so collection efforts are focused on metadata. The idea is to allow the search and retrieval of information, to lead the user to the museum of origin of that item.

To harvest the JSON file of each museum, we use the Filebeat, from two configuration files. The first is the configuration necessary for each collection, for each museum, which will have its data collected. This file basically contains the URL of the JSON API, information about the pagination and the inclusion of four new metadata: museum: the name of the museum, the name of the collection and its location, with city and state. All Ibram's files are available in GitHub[13]. The second is a general configuration file, which orchestrates access to individual files (.yml) and sends their results to Logstash. This file too is available in GitHub[14].

[12] National Historic Museum JSON file example: https://mhn.acervos.museus.gov.br/wp-json/tainacan/v2/collection/24/items/?perpage=96&order=ASC&orderby=date&exposer=json-flat&mapper=inbcm-ibram&paged=1.

[13] https://github.com/tainacan/tainacan-elk/tree/master/Filebeat/inputs.d.

[14] https://github.com/tainacan/tainacan-elk/blob/master/Filebeat/Filebeat.yml.

3.3 Logstash: Aggregation of Collections and Data Transformation

After the collection of data sources made by Filebeat, Logstash receives hem, aggregates them and performs transformations on the data, when necessary.

> A Logstash pipeline has two required elements, input and output, and one optional element, filter. The input plugins consume data from a source, the filter plugins modify the data as you specify, and the output plugins write the data to a destination (Elastic 2021h).

In the input plugin, data is received from Filebeat. Afterwards, the filters are executed. "A filter plugin performs intermediary processing on an event. Filters are often applied conditionally depending on the characteristics of the event" (Elastic 2021d). There are three standard Logstash filters: prune, mutate and fingerprint, and one, tainacan_submission, developed by the authors using the Ruby programming language, in order to submit items to a Tainacan installation. In the output plugin the data is sent to Elasticsearch storage. Figure 3 shows the pipeline with the main steps.

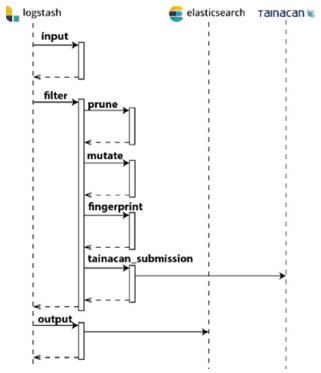

Fig. 3. Logstash pipeline. Prepared by the authors (2021).

The prune filter removes fields from events based on whitelists or blacklist. This can be useful for a json filter that creates fields with names that you don't know the names of beforehand, and you only want to keep a subset of them (Elastic 2021g). The prune filter performs the cleaning of received data, delivering the relevant data, configured in Logstash, removing any others sent by Filebeat.

The mutate filter allows you to perform general mutations on fields. You can rename, remove, replace, and modify fields in your events (Elastic 2021f). In Ibram, the following

mutate filters were used: capitalize[15], lowercase[16], remove_field[17], rename[18], strip[19] and split[20].

Subsequently, the Fingerprint filter:

> Create consistent hashes (fingerprints) of one or more fields and store the result in a new field. You can use this plugin to create consistent document ids when events are inserted into Elasticsearch. This approach means that existing documents can be updated instead of creating new documents (Elastic 2021e).

To create the unique code, a hash was created using three fields: the installation, that is, the museum's name, the item's id and registration number, collected from the museums of origin.

The last filter, tainacan_submission, was developed by the authors to perform the submission of items collected in Tainacan, for Brasiliana Museums.

In output, the data is sent to Elasticsearch and used by Kibana to build the dashboard.

3.4 Ruby Plugin in Logstash: Item Submission in Tainacan

For the best performance of submissions, a queue of processes, Fig. 4, that organize and forward them to the Tainacan API was implemented, since Logstash sends a high number of items simultaneously. All processed items are included at the end of the queue, while a set of threads (four, by default, in Fig. 4) are responsible for processing the output of items from that queue, effecting the items submission in Tainacan. The Ruby code is available in GitHub[21].

After all items in the queue are processed, they will be available for access in the Tainacan installation. It is important to note that after this step, the items will also be sent to Elasticsearch.

3.5 Elasticsearch: Data Storage

All items are stored in Elasticsearch, where you can perform various data search and retrieval operations. The Fig. 5 show the Elasticsearch interface.

---

[15] https://www.elastic.co/guide/en/logstash/current/plugins-filters-mutate.html#plugins-filters-mutate-capitalize.

[16] https://www.elastic.co/guide/en/logstash/current/plugins-filters-mutate.html#plugins-filters-mutate-lowercase.

[17] https://www.elastic.co/guide/en/logstash/current/plugins-filters-mutate.html#plugins-filters-mutate-remove_field.

[18] https://www.elastic.co/guide/en/logstash/current/plugins-filters-mutate.html#plugins-filters-mutate-rename.

[19] https://www.elastic.co/guide/en/logstash/current/plugins-filters-mutate.html#plugins-filters-mutate-strip.

[20] https://www.elastic.co/guide/en/logstash/current/plugins-filters-mutate.html#plugins-filters-mutate-split.

[21] https://github.com/tainacan/tainacan-elk/tree/master/logstash/plugins/src/filters/logstash-filter-tainacan_submission.

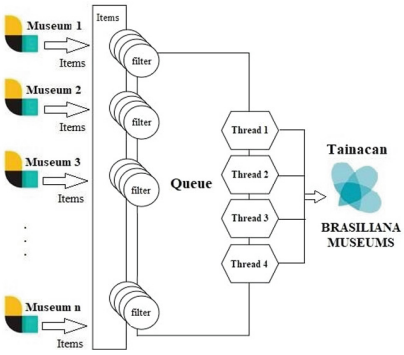

Fig. 4. Submission queue scheme. Prepared by the authors (2022)

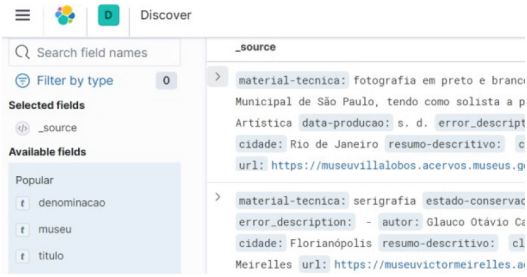

Fig. 5. Elasticsearch interface. Prepared by authors (2022)

In the case of the aggregator, the data will be used to feed the dashboard developed in Kibana.

3.6 Kibana: Dashboard

The Kibana tool accesses the integrated data stored in Elasticsearch, enabling the creation of interactive panels, composed of visualization elements, such as graphs, titles, metrics and maps.

The panels, which can be easily created and/or edited, offer managers different views of the collections, in a single, fully interactive interface that reacts to each new search. As it is directly related to the data aggregation process, the updates carried out in the museums of origin are also carried out in the panels.

We emphasize that the panels are designed to bring information considered important by managers, however, Kibana offers numerous and agile forms of research. We cite an example, if you want to find all items typed "chair" (cadeira, in Portuguese). As shown in Fig. 6 from a simple search, all your panel shapes the result, showing, for example, which museums have chairs, in addition to listing the links that direct to the objects in the museums of origin.

The prototype is available for access at: http://200.137.217.144:6601/goto/613663b72fc306a908e86a6233e5219c.

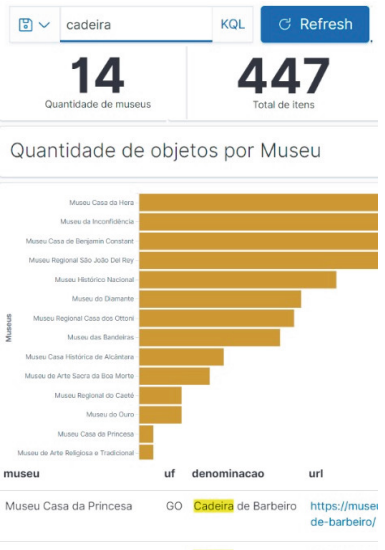

Fig. 6. Search example in Ibram's dashboard. Prepared by authors (2022)

3.7 Tainacan: Final Repository

The aggregated data is sent to a new Tainacan installation via the plugin, and this installation is set up to provide efficient searching and browsing. Figure 7, shows part of the interface.

Fig. 7. Brasiliana Museums. Prepared by authors (2022)

As it is a prototype, the Brasiliana Museums can be seen at http://integracaoibram.tainacan.org/, in which 20 museums are integrated, resulting in 17,315 digital objects.

4 Discussions

Different countries offer the cultural data aggregation service, providing search and retrieval portals, such as the Europeana Portal, of Europe, and the Portal of the Digital Public Library of America - DPLA, of the United States of America, which bring together millions of digital objects from countless different data sources.

In Brazil, although there are initiatives of this type, they are still limited, isolated practices that need technical support and scalable governance models for the Brazilian territory. Thus, the aggregation process exposed in this article was designed to meet the reality of Brazilian institutions managed by Ibram, both budgetary and specialized technical support, in a scalable way, and it is a simple and efficient solution to integrate the collections in an automated way, using only free software.

The research demonstrates the viability of the technologies and, therefore, validates the technological model and the flow of information circulation. It is noteworthy that the combination of Tainacan and Elastic Stack proved to be of medium complexity, becoming a viable learning and customization option.

We emphasize that important goals are achieved, such as allowing users to search for museum objects in a simplified way without the need for more advanced knowledge about museums.

The aggregator can bring other relevant benefits, such as offering resources to improve the management of museum documentation, as it presents, in a unified and interactive way, a macro view of museums; demonstrate syntactic or semantic inconsistencies, enabling museums to improve their information at the base; compare different institutions, making it possible to find patterns and connections of interest, finding indicators that can ensure better strategic planning, greater productivity and agility in accessing data, generating effects in administrative, cultural and educational dimensions and enriching the data source for research.

Another important topic is the financial return that aggregators can bring. Two leading institutions, the British Library and Europeana, carried out studies, both published in 2013, indicating that the British Library's online services are valued at £19.5 million a year, generating benefits for UK students, researchers and citizens. Europeana concluded that the economic impact for the European continent would be approximately 78.8 million euros in a baseline scenario, 59.6 million euros in a pessimistic scenario and 97.7 million euros in an optimistic scenario (Martins et al. 2018).

As it is something new at the national level, in addition to the aggregator of cultural data, it is necessary to provide institutions with best practices to maintain their collections on the web, in order to provide better results related to integrated search and, consequently, to improve the national collections as a whole.

It is important to highlight that to retrieve content in the aggregator it is important to establish quality parameters to be applied to the source data, which requires the creation of guidelines for better data quality, such as Europeana and DPLA, enabling museums to improve their data continuously. These guidelines are also under study by the Network Intelligence Laboratory.

For Ibram, as an institution that manages museums, aggregation also brings advantages related to the management and monitoring of collections, which makes it possible

to carry out more accurate diagnoses of the status of collections and information about them, in addition to the generation of reports broader and more reliable analytics.

5 Conclusions

Technological evolution has enabled the reinvention of the museum space, opening up potential paths for the dissemination and socialization of memory and culture. Thus, international and national institutions have invested in making their digital collections available, resulting in countless museum data distributed on the web, generating different possibilities for searching and retrieving information.

This scenario, although very rich, brought challenges to the discovery of cultural heritage resources. To this end, in order to optimize the location of digital objects present in different museum sites across the internet, several institutions around the world have become data providers for services specialized in the aggregation of cultural collections of cultural heritage, as is the case of Europeana and of the DPLA. In this sense, aggregation makes the search and retrieval of information about collections a more efficient information retrieval operation for the users of these collections.

For example, to find a specific digital object, it is necessary to know the collection in which it belongs and some of its characteristics, which for a lay user would be practically impossible, greatly limiting the actual dissemination of the collection. In this perspective, providing museums and users with aggregation services for efficient search and retrieval becomes an essential theme.

The study and prototyping of the service for the aggregation of digital museum collections, Brasiliana Museums, reveals itself, therefore, as an efficient solution, both from the point of view of socialization and dissemination of digital objects and in the monitoring and monitoring of information from the collections through analytical panels, which can change the way managers, researchers and users interact with cultural data, bringing countless benefits to society. In addition to presenting a cheap and viable technological solution to be implemented.

As future works, research continues and improvements will be carried out during the studies, with the next step being the collection of thumbnails of images of digital objects from the museums of origin for the aggregator.

References

BRASIL. Instituto Brasileiro de Museus, Resolução Normativa nº 2, 2014. Diário Oficial. https://www.in.gov.br/web/dou/-/resolucao-normativa-ibram-n-6-de-31-de-agosto-de-2021-342359740. Accessed 03 Mar 2022

Elastic: Elastic stack (2021a), https://www.elastic.co/elastic-stack/. Accessed 10 July 2021a

Elastic: Filebeat reference [7.13]. Filebeat overview (2021b). https://www.elastic.co/guide/en/beats/Filebeat/current/Filebeat-overview.html. Accessed 10 July 2021

Elastic: Logstash reference [7.13]. Contributing to Logstash (2021c). https://www.elastic.co/guide/en/logstash/current/contributing-to-logstash.html. Accessed 10 July 2021

Elastic: Logstash reference [7.13]. Filter plugins (2021d). https://www.elastic.co/guide/en/logstash/current/filter-plugins.html. Accessed 09 July 2021

Elastic: Logstash reference [7.13]. Filter plugins. Fingerprint filter plugin (2021e). https://www.elastic.co/guide/en/logstash/current/plugins-filters-fingerprint.html. Accessed 09 July 2021

Elastic: Logstash reference [7.13]. Filter plugins. Mutate filter plugin (2021f). https://www.elastic.co/guide/en/logstash/current/plugins-filters-mutate.html. Accessed 09 July 2021f

Elastic: Logstash reference [7.13]. Filter plugins. Prune filter plugin (2021g). https://www.elastic.co/guide/en/logstash/current/plugins-filters-prune.html4. Accessed 09 July 2021

Elastic: Logstash reference [7.13]. Getting started with Logstash. Stashing your first event (2021h). https://www.elastic.co/guide/en/logstash/current/first-event.html. Accessed 09 July 2021h

Elastic: Logstash reference [7.13]. Logstash introduction (2021i). https://www.elastic.co/guide/en/logstash/current/introduction.htm. Accessed 10 July 2021

Gov.br: Museus Ibram (2022a). https://www.gov.br/museus/pt-br/museus-ibram. Accessed 02 Dec 2022

Gov.br: Sobre o órgão (2022b). https://www.gov.br/museus/pt-br/acesso-a-informacao/institucional/sobre-o-orgao. Accessed 02 Dec 2022

Martins, D.L., Silva, M.F., do Carmo, D.: Acervos em rede: perspectivas para as instituições culturais em tempos de cultura digital. Em Questão **24**(1), 194–216 (2018). https://doi.org/10.19132/1808-5245241.194-216

Martins, D.L., Carmo, D.do., Silva, M.F.S.: Modelos de governança em serviços de acervos digitais em rede: elementos para a produção de uma política pública nacional para objetos culturais digitais. Perspectivas Em Ciência Da Informação, **27**(1) (2022). https://periodicos.ufmg.br/index.php/pci/article/view/39044

Siqueira, J., Martins, D.L.: Recuperação de informação: descoberta e análise de workflows para agregação de dados do patrimônio cultural. Ciência Inf. **49**(3) (2020). http://revista.ibict.br/ciinf/article/view/539910

Siqueira, J., Martins, D.L.: Workflow models for aggregating cultural heritage data on the web: a systematic literature review. J. Assoc. Inf. Sci. Technol. 1–21 (2021). https://doi.org/10.1002/asi.24498

Tainacan.org: Páginas do Tainacan (2021a). https://tainacan.github.io/tainacan-wiki//tainacan-pages?id=as-páginas-especiais-do-tainacan. Accessed 09 July 2021

Tainacan.org: Tainacan. Home (2021b). https://tainacan.org/en/. Accessed 10 July 2021