Received: 2 February 2023 | Revised: 16 July 2023 | Accepted: 7 September 2023

DOI: 10.1002/asi.24835

RESEARCH ARTICLE

JASIST WILEY

An approach to assess the quality of Jupyter projects published by GLAM institutions

Gustavo Candela[1] | **Sally Chambers**[2,3,4] | **Tim Sherratt**[5]

[1]Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante, Alicante, Spain

[2]Ghent Centre for Digital Humanities, Ghent University, Ghent, Belgium

[3]KBR, Royal Library of Belgium, Brussels, Belgium

[4]DARIAH, Digital Research Infrastructure for the Arts and Humanities, Paris, France

[5]Centre for Creative and Cultural Research, University of Canberra, Bruce, ACT, Australia

Correspondence
Gustavo Candela, Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante, carretera Sant Vicent s/n, 03690 Sant Vicent del Raspeig, Alicante, Spain.
Email: gcandela@ua.es

Abstract

GLAM organizations have been digitizing their collections and making them available for the public for several decades. Recent methods for publishing digital collections such as "GLAM Labs" and "Collections as Data" provide guidelines for the application of computational methods to reuse the contents of cultural heritage institutions in innovative and creative ways. Jupyter Notebooks have become a powerful tool to foster use of these collections by digital humanities researchers. Based on previous approaches for quality assessment, which have been adapted for cultural heritage collections, this paper proposes a methodology for assessing the quality of projects based on Jupyter Notebooks published by relevant GLAM institutions. A list of projects based on Jupyter Notebooks using cultural heritage data has been evaluated. Common features and best practices have been identified. A detailed analysis, that can be useful for organizations interested in creating their own Jupyter Notebooks projects, has been provided. Open issues requiring further work and additional avenues for exploration are outlined.

1 | INTRODUCTION

GLAM (Galleries, Libraries, Archives and Museums) have been digitizing their collections and making them available for the public since the mid-1990s (Ayris, 2010; Hughes, 2004; Nielsen, 2008). Reusing digital collections in innovative and inspiring ways has become an active challenge for cultural heritage institutions (Mahey et al., 2019). Collaborations with digital humanities researchers are increasing (Wilms, 2021). Recent approaches provide guidelines and best practices for publishing digital collections for computational use (Padilla et al., 2019).

Jupyter[1] has emerged as a popular interactive computing environment for Open Science, enabling the exploration and reproduction of results, simulations and documentation of workflows (Beg et al., 2021). A Jupyter Notebook combines code, text, images and charts in a single document that is executable in a local or cloud environment. The growing popularity of Jupyter Notebooks for data analysis is reflected in the fact that GitHub currently hosts more than 700,000 repositories using Jupyter,[2] while the collaborative Machine Learning platform Hugging Face hosts over 7000 notebooks.[3]

Jupyter is also becoming a key tool in GLAM institutions for public engagement as it can be used to introduce users to accessing and reusing datasets (Candela et al., 2022; Sherratt, 2021). GLAM organizations are making large quantities of open data available to the public, but the scale, scope, and research possibilities of this data is not always obvious. Jupyter Notebooks provide a means of documenting pathways for access and reuse, encouraging researchers to explore these rich data

offerings. The GLAM Workbench was a pioneering project, and the use of Jupyter Notebooks has since been taken up by a range of GLAM institutions.[4] More institutions are likely to follow, for example, the 77 organizations that are members of the International GLAM Labs Community.[5] By measuring the quality of the existing Jupyter Notebook projects made available by GLAM institutions, best practices and guidelines can be identified in order to help support new adopters.

Previous work has proposed best practices and guidelines for publishing Jupyter Notebooks (Australian Research Data Commons, 2023; Rule et al., 2019). "Quality" in terms of Jupyter Notebooks has been also addressed in a general context (Pimentel et al., 2019, 2021). However, these approaches are applied to projects relating to Jupyter Notebooks in general. Our approach focuses on GLAM institutions for the following reasons: (i) they have played a leading role in the publication of openly available digital collections and datasets of different types of materials, such as newspapers, metadata, images, text and maps; (ii) the potential of notebooks for stimulating the reuse of the digital collections by researchers, and (iii) to foster uptake and encourage other institutions to adopt computational access to digital collections in a coherent and coordinated way.

The objective of the present study is to introduce a methodology to assess the quality of Jupyter Notebooks projects made available by GLAM institutions. This methodology was applied to a list of projects made available by relevant institutions. The results of this study are publicly available and can be applied to other domains such as digital humanities and data science.

The main contributions of this paper are as follows: (a) a compilation of Jupyter Notebooks published by relevant institutions; (b) a methodology to assess the quality of Jupyter Notebook projects published by GLAM organizations; and (c) the results of a quality assessment obtained by applying these criteria to a selected set of Jupyter Notebooks. In addition, we propose improvements for discoverability and access of such Jupyter Notebook projects by providing machine-readable metadata through a collaborative platform such as Wikidata. Furthermore, these contributions are intended to encourage GLAM institutions to adopt Jupyter as a key method for introducing researchers to (re)using datasets.

We also hope that the paper will contribute to discussions around best practice for the development, documentation, and maintenance of Jupyter Notebook repositories in the GLAM sector and beyond. While the quality measures described focus on the *use* of notebooks, attention to accepted standards and best practice guidelines will facilitate the automation of management processes, foster collaboration, and increase prospects for long-term sustainability. For example, the GLAM Workbench embeds basic metadata, licensing, and management tools within a reusable repository template.[6] In this way, the paper seeks to engage with recent initiatives around software sustainability, and the application of FAIR principles to research software (Australian Research Data Commons, 2022; Barker et al., 2022).

The paper is organized as follows: after a brief description of the state of the art in Section 2. Section 3 describes the methodology employed for assessing the Jupyter Notebook projects. The application of the methodology and results are shown in Section 4. In addition, an approach for improving the discovery of Jupyter Notebooks with Wikidata is introduced. The paper concludes with an outline of the adopted methodology, general guidelines on how to use the results, and future work.

2 | RELATED WORK

During the last decade, GLAM institutions have been exploring new ways to make digital collections available for the public. There is a wide diversity of materials (e.g., historical text, maps, images and metadata), standards, formats, and delivery methods (e.g., compressed files and APIs). Some examples include the Data Foundry at the National Library of Scotland (NLS), the historic American newspapers provided by Chronicling America, Smithsonian Open Access[7] and the Metropolitan Museum of Art Collection API.[8] Such approaches are facilitating collaborations with digital humanities researchers (Vandegrift & Varner, 2013).

Recently, there has been a growing interest in the application of advanced techniques and methods such as Machine Learning (ML), Computer Vision and Artificial Intelligence to the collections of cultural heritage institutions. Several initiatives are focused on the identification of challenges, common issues and best practices in developing machine learning projects that utilize data provided by cultural heritage institutions such as libraries and museums (Berlin State Library, 2022; Lee, 2022; Lorang et al., 2020; Murphy & Villaespesa, 2020; Padilla, 2019).

The publication of Jupyter Notebooks by GLAM organizations has recently increased in order to demonstrate to researchers how to access and reuse their digital collections and materials. For instance, the US Library of Congress (LoC) has made available a Jupyter Notebook collection based on their catalogs including the US Library of Congress Digital Collections[9] and Chronicling America historic American newspapers.[10] The GLAM Workbench is a collection of tools and examples using Jupyter Notebooks to explore digital collections provided by GLAM institutions, particularly in Australia and New Zealand (Sherratt, 2021). The NLS has created a Jupyter Notebook project based on the datasets published

in their Data Foundry (Ames & Havens, 2022). The Biblioteca Virtual Miguel de Cervantes (BVMC) has published a collection of Jupyter Notebooks that applies a wide range of research methods to different datasets provided by several relevant GLAM institutions (Candela et al., 2022). Recently, a collection of Jupyter Notebooks for processing of historical text resources from Europeana Newspapers with CLARIN natural language processing (NLP) tools has been made available (CLARIN ERIC, 2022). Outputs from research collaborations based on GLAM and Cultural Heritage institutions such as AI4LAM (Artificial Intelligence for Libraries, Archives and Museums) and NewsEye[11] have made collections of Jupyter Notebooks available to present their results. Other approaches include Jupyter being embedded into teaching and learning platforms such as Constellate.[12] Additional examples have been published in support of climate studies and to inform policy decisions (NASA, 2022). Table 1 shows an overview of projects published by GLAM institutions.

Data quality and standards are crucial in order to enable reuse of digital collections. Several methodologies have been proposed to assess the quality of data, defining criteria classified by a number of dimensions (ISO 25000, 2014; World Wide Web Consortium, 2017; Zaveri et al., 2016). In addition, new ontologies such as the Data Quality Vocabulary has been created to enable data providers to describe and share information about the quality of their datasets (World Wide Web Consortium, 2016).

The quality of notebooks has been addressed in previous works that identify common issues (Oli et al., 2021; Pimentel et al., 2019, 2021). Examples of these include: the execution order of cells, and problems such as unnamed notebooks (e.g., Untitled1.ipynb) or the reproducibility of results. Other approaches focus on the publication of best practices to facilitate reproducibility, to better understand the code and documentation, and to enhance discoverability (Rule et al., 2019).

Climate change and environmental degradation due to electricity consumption is becoming an important issue for cultural heritage institutions and digital humanities researchers (European Commission, 2019). Due to the expansion of digitization and emerging technologies such as artificial intelligence and machine learning, the

TABLE 1 Overview of Jupyter Notebook projects published by GLAM institutions, individual researchers or created as a result of research projects using cultural heritage data.

Institution	Title	URL
AI4LAM	AI4LAM Metadata Working Group	https://github.com/AI4LAM/metadata-working-group/tree/main/notebooks
Archives Unleashed Project	Archives Unleashed Notebooks	https://archivesunleashed.org/notebooks/
Austrian National Library	Scripts and Jupyter Notebooks	https://labs.onb.ac.at/gitlab
Biblioteca Virtual Miguel de Cervantes	GLAM Jupyter Notebooks	https://data.cervantesvirtual.com/glam-jupyter-notebooks
British Library	Jupyter Notebooks using the British Library's Digital Collections & Data	https://github.com/BL-Labs/Jupyter-notebooks-projects-using-BL-Sources
Europeana	Europeana Newspapers Notebooks	https://marketplace.sshopencloud.eu/training-material/duVII1
German National Library	German National Library's Jupyter Notebooks	https://github.com/deutsche-nationalbibliothek/dnblab
US Library of Congress	Data Exploration	https://github.com/LibraryOfCongress/data-exploration
National Library of Estonia	Access to the National Library of Estonia newspaper and periodical collections	https://data.digar.ee/samples/access_eng.html
National Library of Scotland	Jupyter Notebooks	https://data.nls.uk/tools/jupyter-notebooks/
NewsEye	NLP Notebooks for Newspaper Collections	https://github.com/NewsEye/NLP-Notebooks-Newspaper-Collections
Penn Libraries	Collections As Data Notebooks at Penn Libraries	https://github.com/upenndigitalscholarship/collections-as-data-notebooks
Tim Sherratt	GLAM Workbench	https://glam-workbench.net/
Victoria and Albert Museum	Data Explorations	http://developers.vam.ac.uk/notebooks/data-explorations/intro.html

European Commission is exploring measures to improve the energy efficiency in cloud computing and data centers (European Commission, 2020; European Commission, 2019). For instance, cultural heritage initiatives focused on climate change have been recently organized.[13] The carbon impact of training machine learning and artificial intelligence models is becoming increasingly important in the research community (Bannour et al., 2021; Lacoste et al., 2019; Schwartz et al., 2020). Recent software libraries aim at facilitating the distribution of pretrained models to help reduce the carbon footprint (Wolf et al., 2019).

Collaborative initiatives such as Wikidata have become very popular among the GLAM sector to enrich their datasets by describing and linking their resources. Wikidata enables the creation of new properties to link resources (e.g., authors, works or locations) and allowing a global community to edit and maintain the data as well as increasing the visibility of the resources. Previous works are focused on the quality of Wikidata concerning several aspects such as constraints violations in data and the definition of data quality criteria (Färber et al., 2018; Shenoy et al., 2022).

This overview demonstrates how Jupyter Notebooks can be made available for the public and assessed. Nevertheless, to our best knowledge, none of the work to date provides a methodology to assess the quality of the Jupyter Notebook projects published by GLAM institutions. This analysis will be useful for the GLAM community to identify and agree on best practices.

3 | METHODOLOGY

This section introduces the criteria for assessing the quality of Jupyter Notebook projects published by GLAM institutions that reuse digital collections by means of computational access methods (e.g., APIs, HTTP downloads). For this study, a Jupyter Notebook project is understood as a collection of notebooks as well as the configuration files and datasets used by the notebooks, either as downloadable dump files or by means of an API. The methodology employed is based on previous works, in particular guidelines and best practices (ISO 25000, 2014; Pimentel et al., 2019; Rule et al., 2019; Zaveri et al., 2016). Note that while the previous definitions of the quality criteria refer to data, the criteria in this study have been adapted in order to assess the content of a Jupyter Notebook project. A number of dimensions are used to classify the criteria. Each criterion includes a function with values ranging from 0 to 1. The following section presents the proposed quality criteria: *understandability*, *availability*, *efficiency*, *traceability*, *portability*, *recoverability* and *credibility*.

Previous work has identified issues regarding the use and publication of Jupyter Notebooks (Beg et al., 2021; Grus, 2018; Pimentel et al., 2019, 2021; Xie, 2018). In what follows, we discuss these issues adapted to the GLAM sector, and propose a set of criteria for analyzing the quality of Jupyter Notebooks made available by GLAM organizations. Note that in some cases projects include a large collection of Jupyter Notebooks or consist of several GitHub repositories. In such cases, the criteria can be applied to a random sample of notebooks included in the project. Table 2 shows the criteria grouped by dimensions and including examples of how the projects can be analyzed.

3.1 | Understandability

Understandability refers to the ease with which data can be comprehended without ambiguity using appropriate languages and symbols, and be used and interpreted by humans (ISO 25000, 2014; Zaveri et al., 2016). For instance, information to aid understandability can be provided as text, metadata description and documentation in several formats (e.g., pdf files, websites, research articles, etc.). Understandability also includes intended use, enabling potential users of a collection to be identified. Based on other approaches, the use of a descriptive name for the notebooks, as well as storing the output, can help users to better understand a collection.

3.1.1 | Using literate programming features

While a traditional computer program consists of code that can include comments describing its functionality, in literate programming the programmer creates documentation addressed to humans that includes code.[14] One key benefit of Jupyter is that it combines code with text descriptions to create a computational narrative (Rule et al., 2019). These computational narratives are crucial to help users to better understand how data are processed, to encourage individual exploration and to share the results with the community (Granger & Pérez, 2021). Textual descriptions can be included using the plain-text formatting syntax, markdown (Gruber, 2004). Depending on the audience and the context, markdown cells can include, for example, an introduction to the task performed, the dataset or API used, textual descriptions of the code, and comments on the results or references. Figure 1 shows an example of markdown cell.

TABLE 2 Description and structure of the criteria used to assess the quality of Jupyter Notebooks published by GLAM institutions.

Dimension	Criterion	Example of analysis
Understandability	Using literate programming features	To what extent the notebooks include markdown cells?
	Including additional documentation and guidelines	The collection include additional documentation about the project?
	Naming of the notebooks	The names used for the notebooks are self-describing?
	Storing cell output	Does the notebooks contain the outputs?
	Audience/intended use	Who is the audience of the collection?
	Provisioning of metadata	Information as metadata (e.g., title, author, etc.) to describe the collection is included?
Availability	License	Is a license available? Which license has been used to publish the collection?
Efficiency	Size	What is the size of a Jupyter Notebook project including the datasets used?
Traceability	Versioning	Is the collection published using a versioning system such as GitHub?
Portability	Providing dependencies	Does the project provide a dependencies file (e.g., requirements.txt)?
Recoverability	Providing citation information	Is a persistent URI provided? Is a citation format such as BibTeX available?
	Last run date	When was the Jupyter Notebook project last used or run successfully?
Credibility	Trustworthiness on project level	Does the project provide reliable provenance information such as acknowledgements, sources or awards received?

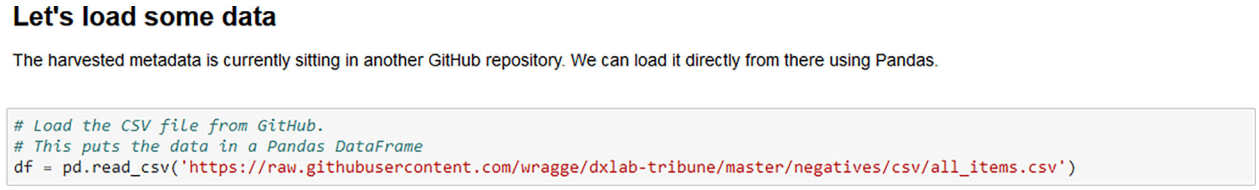

FIGURE 1 Example of markdown cell describing the process of harvesting a dataset from GitHub. The markdown cell is followed by the Python code.

For example, if c is a cell in a notebook nb, included in a project including a collection of Jupyter Notebooks p, this criterion measures the average difference between markdown and code cells per notebook:

$$avgCells(p) = \frac{1}{n}\sum_{i=1}^{n} \mid m_{nbMarkdownCells}(nb_i) - m_{nbCodeCells}(nb_i) \mid \quad (1)$$

Then we can define the metrics $m_{nbMarkdownCells}(p)$ and $m_{nbCodeCells}(p)$ as follows:

$$m_{\text{nbMarkdownCells}}(nb) = \frac{\mid \{nb \epsilon p \mid c \epsilon nb \wedge cellType(c) = md\} \mid}{\mid c \epsilon nb \mid}, \quad (2)$$

$$m_{\text{nbCodeCells}}(nb) = \frac{\mid \{nb \epsilon p \mid c \epsilon nb \wedge cellType(c) = code\} \mid}{\mid c \epsilon nb \mid} \quad (3)$$

After which, the values are inverse normalized to map the highest size value to 0 and the lowest to 1. Let c be a selection of Jupyter Notebook projects, $minCells(c)$ and $maxCells(c)$ the minimum and maximum size value of c, then we can define the criterion $m_{cells}(p)$ as follows:

$$m_{cells}(p) = 1 - \frac{(avgCells(p) - minCells(c))}{(maxCells(c) - minCells(c))} \quad (4)$$

Note that the quality and semantics of the text provided in the notebooks is out of the scope of this

approach. In addition, it is worth noting that the inclusion of in-code comments, apart from the markdown cells, is also important in helping users to better understand the code.

3.1.2 | Including additional documentation and guidelines

Publishing documentation about how to find, use, and understand the projects in multiple places including blogs, README files, and user-friendly tutorials in peer-reviewed journals of digital humanities such as Programming Historian, is a key element to improve understandability and to foster community engagement and research (Padilla et al., 2019).

$$m_{\text{documentation}}(p)=\begin{cases}1 & \text{dedicated website and turorials}\\ 0.5 & \text{README file}\\ 0 & \text{otherwise}\end{cases} \tag{5}$$

3.1.3 | Naming of the notebooks

Meaningful names are useful for users and for search and retrieval systems (e.g., Europeana Newspapers Notebooks). They aid findability. By default, Jupyter creates notebooks titled "Untitled." Jupyter allows users to create a copy of a notebook adding the text -Copy to the name. However, if a user does this, it is important that they rename the notebook appropriately. In addition, filename conventions such as POSIX define a set of rules for filename portability (Pimentel et al., 2021). An additional aspect to consider is consistency in file naming practices, as creators may be inconsistent with the naming of files within a notebook project.

Then we can define the metrics $m_{naming}(p)$ as follows:

$$m_{\text{naming}}(p)=\frac{1}{n}\sum_{i=1}^{n}|\{nb_i \epsilon p \wedge isDescriptiveAndConsistent(nb_i)\}| \tag{6}$$

3.1.4 | Storing cell output

Displaying execution results is part of the narrative aspect of Jupyter Notebooks. A user is able to store and clean cell outputs (e.g., plots, prints, etc.) using the editor (Pimentel et al., 2021). Previous work suggests cleaning execution results before committing to avoid noise when comparing the content of notebooks (Xie, 2018). However, for the cultural heritage domain, providing users with examples in the cells can help improve the understandability of the notebooks, especially for less-experienced users. Note that the formula includes notebooks that fail to execute fully or return errors in their results.

Then we can define the metrics $m_{output}(p)$ as follows:

$$m_{\text{output}}(p)=\frac{1}{n}\sum_{i=1}^{n}|\{nb_i \epsilon p \wedge storeOutput(nb_i) \wedge noError(nb_i)\}| \tag{7}$$

3.1.5 | Audience/intended use

How Jupyter Notebooks are made available to the public depends on the audience and intended use. Potential users include the general public, students, less digitally-literate researchers and highly skilled researchers such as data scientists. According to the target audience, the narrative and code may be different in terms of details, complexity and length (Rule et al., 2019).

A Jupyter Notebook project may be addressed to different types of users. Notebooks addressed to advanced users can be identified by using packages aiming at applying advanced computational methods and concepts such as computer vision, the semantic web and Named Entity Recognition (NER). On the contrary, notebooks intended for less experienced users are based on specific and simple tasks such as retrieving and analyzing a dataset.

$$m_{\text{audience}}(p)=\begin{cases}1 & \text{advanced and less experienced users}\\ 0.5 & \text{advanced or less experienced users}\\ 0 & \text{otherwise}\end{cases} \tag{8}$$

Note that this criterion has been manually measured by reviewing for instance whether the collection provides information regarding the audience, and the inclusion of introductory and advanced notebooks.

3.1.6 | Provision of metadata

Metadata provides additional information that helps users to better understand the item that has been published. Key elements are its structure (e.g., file format), the organization that owns the project, the date of publication, datasets and access methods used, and the version of the project (World Wide Web Consortium, 2017). Publishers are encouraged to provide human and machine-readable information in multiple languages, focusing on the intended audience.

Collaborative-editing platforms such as Wikidata have emerged as multilingual, free and open metadata

repositories that can be updated by humans and machines. Wikidata stores metadata about entities that are identified by a QID (or Q number). Each entity is described by means of properties that define statements in the form of subject-predicate-object (e.g., Shakespeare is_author_of Hamlet). The information is accessible by means of a public API.[15]

3.2 | Availability

Availability is defined as *the degree to which data has attributes that enable it to be retrieved by authorized users and/or applications in a specific context of use* (ISO 25000, 2014).

3.2.1 | License

A Jupyter Notebook project can be made available by means of open licenses allowing researchers to reuse the content and reproduce the results without restrictions. Creative Commons provide a set of licenses that can be adapted to the requirements of the authors.[16] For example, CC-BY[17] licenses require attribution, which is important for notebooks used within a research context, including a link to the license and an indication if changes were made. Best practices and guidelines on the application of open licenses to data have been provided by organizations such as the Open Knowledge Foundation (Open Knowledge Foundation, 2015). In some cases, the license may be based on national regulations. However, in other cases the license may not be clear or no licensing information has been provided. Additionally, some datasets used in a particular Jupyter Notebook project may be restricted, for example, access is only available inside the library's reading room. In this case, metadata describing the restricted dataset should be included, including details of how to access the particular dataset.

$$m_{\text{license}}(p) = \begin{cases} 1 & \text{using licenses enabling reuse} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

3.3 | Efficiency

Efficiency is defined as the extent to which data has attributes that provide expected outcomes while using the appropriate amounts of resources in a specific context of use (World Wide Web Consortium, 2016). Several indicators can be used to measure the carbon footprint such as the size of the resources, the size of the datasets reused or the distance that the data and resources have to travel.[18]

$$m_{\text{metadata}}(p) = \begin{cases} 1 & \text{multilingual, human and machine-readable metadata available} \\ 0.75 & \text{machine-readable metadata available} \\ 0.5 & \text{human-readable metadata available} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

3.3.1 | Size

Jupyter Notebooks are in general hosted on GitHub. The larger the size of the data and notebooks, the more computing capacity and energy is required to transfer and store the data. According to a selection of projects, this criterion measures the size of a Jupyter Notebook collection. This criterion includes the size of the datasets that are downloaded as dump files or by means of an API. After running the code of each notebook, the size of the folder of the project is computed.

Given a Jupyter Notebook project p, and *size* the function that measures the size of the each notebook *nb* including the datasets used, the average of the size of the notebooks is computed as follows:

$$avgSize(p) = \frac{1}{n}\sum_{i=1}^{n} |\{nb_i \epsilon p \wedge size(nb_i)\}| \quad (11)$$

After which, the values are inverse normalized to map the highest size value to 0 and the lowest to 1. Let c be a selection of Jupyter Notebook projects, $minSize(c)$ and $maxSize(c)$ the minimum and maximum size value of c, then we can define the criterion $m_{size}(p)$ as follows:

$$m_{\text{size}}(p) = 1 - \frac{(avgSize(p) - minSize(c))}{(maxSize(c) - minSize(c))} \quad (12)$$

3.4 | Traceability

The traceability dimension is defined as *the degree to which data has attributes that provide an audit trail of access to the data and of any changes made to the data in a specific context of use* (World Wide Web Consortium, 2016).

3.4.1 | Versioning

Jupyter Notebook projects may change over time. For instance, particular notebooks may be updated (e.g., repository access or dataset used) and improved (e.g., new versions of software libraries). In order to manage these changes, new versions of a project may be created and made available to the public (World Wide Web Consortium, 2017). For example, releases in GitHub are versioned software packages including code and notes providing a full project history. In addition, software versions used can be described as metadata in platforms such as Wikidata See, for example, the property[19] as well as using vocabularies such as Schema.org.[20]

$$m_{\text{version}}(p) = \begin{cases} 1 & \text{project versions and metadata available} \\ 0.5 & \text{project versions or metadata available} \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

3.5 | Portability

In the context of data, portability is defined as *the degree to which data has attributes that enable it to be installed, replaced or moved from one system to another preserving the existing quality in a specific context of use* (World Wide Web Consortium, 2016).

3.5.1 | Providing dependencies

Jupyter Notebooks can be executed locally and in cloud environments such as Binder (Jupyter et al., 2018) and Google Colab.[21] However, notebooks themselves do not include the versions of the Python libraries used in projects (Grus, 2018). This can lead to issues in terms of incompatibilities when running the notebooks in other environments. Notebooks have limited control over the environment in which they are run. A Jupyter Notebook can include code to install specific versions of packages via the pip command, though this is limited by the environment. Python provides different methods to include the dependencies based on standard configuration files including setup.py, requirements.txt and Pipfile (Pimentel et al., 2021). These files can be used within cloud environments to install the correct version of the software libraries to enable users to run the code.

$$m_{\text{dependencies}}(p) = \begin{cases} 1 & \text{using dependencise file} \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

More advanced initiatives use Docker images built from the Jupyter Notebooks and requirements files to package the software for reuse. This approach has several benefits in terms of portability since repositories can be run locally or in a wide range of cloud services. In addition, it offers efficiency benefits as Docker images are built once and deployed as required rather than being built multiple times. Tools based on code repositories such as GitHub are available to build, run, and push Docker images.[22]

3.6 | Recoverability

Recoverability in a data context is defined as *the degree to which data has attributes that enable it to maintain and preserve a specified level of operations and quality, even in the event of failure, in a specific context of use* (World Wide Web Consortium, 2016).

3.6.1 | Providing citation information

In general, citation information for software and datasets (e.g., title, authors or date of publication) is less common than in other research outputs such as articles and books.

The use of persistent URIs allows users to find and cite projects through time (World Wide Web Consortium, 2017). They are standardized unique and permanent strings assigned to online works such as books and articles that provide a persistent link to their location on the internet. Examples of persistent identifiers (Madden et al., 2020) include: Archival Resource Keys (ARKs)[23]; Digital Object Identifiers (DOI)[24] and Handles.[25]

Recent approaches provide innovative methods to include citation information describing how to correctly cite the software. For example, `CITATION.cff` files are plain text files with citation information that are supported by major code and data publication repositories (Druskat et al., 2021).

Platforms such as Zenodo[26] and DataCite[27] provide services to make code available and citable by means of persistent interoperable identifiers such as DOI and

standard citation styles as BibTeX.[28] In addition, notebooks may cite datasets that are used or analyzed for a particular purpose that can be included as a list of references.

$$m_{\text{citation}}(p) = \begin{cases} 1 & \text{using persistent URI and full citation} \\ 0.5 & \text{using persistent URI or full citation} \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

3.6.2 | Last run date

Jupyter Notebook projects can depend on multiple resources to run successfully. For example, a Jupyter Notebook might use an API that has been decommissioned or changed. Providing the last run date can help users identify how up-to-date the notebooks are, and when they were last used.

$$m_{\text{lastrun}}(p) = \begin{cases} 1 & \text{providing last run date} \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

3.7 | Credibility

In terms of information, credibility is introduced as *the degree to which data has attributes that are regarded as true and believable by users in a specific context of use* (ISO 25000, 2014).

3.7.1 | Trustworthiness on project level

This criterion measures whether a Jupyter Notebook project includes reliable and clear information. The concept of source criticism has been previously used as an alternative to other practices of information verification in several fields such as digital humanities and journalism, providing guidelines for the practical assessment of sources and source material (Koch & Kinder-Kurlanda, 2020; Koolen et al., 2019; Steensen et al., 2022). In this context, relations in source material to other sources are particularly important. Another aspect to consider is analyzing what information is missing in the source material (e.g., information about the author missing). Some examples include acknowledgements, awards received, provenance, funding sources, attributions, references in research articles, and so forth. For instance, initiatives such as Core Trust Seal promote sustainable and trustworthy data infrastructures.[29] Awards may be an indication of trustworthiness in terms of the impact and the institutions involved.[30] Advanced approaches include the provision of provenance information based on interoperable and standard vocabularies such as the PROV Ontology (PROV-O) (World Wide Web Consortium, 2013a).

Another relevant aspect to measure trustworthiness is evidence that a repository is being actively maintained. Developer platforms such as GitHub provides information (e.g., date) about changes of the code as commits. A recent date is an indicator that the code is actively updated and maintained.

$$m_{\text{trustworthiness}}(p) = \begin{cases} 1 & \text{reliable information available and actively maintained} \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

Note that this approach has considered updates within the last 12 months to measure that the code is actively updated and maintained.

4 | RESULTS

This section presents the application of the method proposed in Section 3 to a list of Jupyter Notebook projects published by relevant GLAM organizations or making use of cultural heritage collections. The projects have been selected according to the following criteria: (i) they are included in Section 2; (ii) they are available using open licenses; and (iii) they are available without requiring the user to login or create a user account. Each of the projects selected were assessed against the criteria proposed in Section 3. Table 3 shows the results obtained per project.

In order to analyze the content of the notebooks, the Python library *nbformat* has been used to identify whether the notebooks contained output, and the use of markdown cells in combination with code.[31]

4.1 | Discussion and findings

In general, the notebook projects included in this article are based on the datasets provided by the creating organization, such as the NLS and the LoC. However, in some cases such as the GLAM Workbench and the BVMC, the notebooks reuse datasets from several institutions.

Overall, the dimension **understandability** achieves good results for the notebook projects. In addition, documentation is in general provided as dedicated websites or README files. In some cases, such as the German National Library, the information is provided using a

TABLE 3 Results of the assessment of the quality of Jupyter Notebook projects.

Project	m_{cells}	$m_{documentation}$	m_{naming}	m_{output}	$m_{audience}$	$m_{metadata}$	$m_{license}$	m_{size}	$m_{version}$	$m_{dependencies}$	$m_{citation}$	$m_{trustworthiness}$	$m_{lastrun}$	Total
Archives Unleashed Notebooks	0.58	1	0.82	0.65	1	0.5	1	0	0.5	0	0	1	0	7.05
BL's Jupyter Notebooks	0.63	1	0.92	0.67	1	0.5	1	0.81	0.5	1	0	1	0	9.03
BVMC—GLAM Jupyter Notebooks	0.94	1	1	0.5	0.5	0	1	0.94	1	1	1	1	0	9.88
Europeana Newspapers Notebooks	0.98	1	1	0.9	1	0.5	1	0.99	1	1	1	1	0	11.37
German National Library's Jupyter Notebooks	0.87	1	1	1	1	0.5	1	0.99	0.5	1	0	1	0	9.86
GLAM Workbench	0.99	1	1	1	1	0.5	1	0.98	1	1	1	1	0	11.47
LoC—Jupyter Notebooks	0.96	1	0.83	0.83	1	0.5	1	0.99	0.5	0	0	1	0	8.61
NewsEye—NLP Notebooks for Newspaper Collections	0.86	1	1	1	0.5	0.5	0	1	0.5	0	0	1	0	7.36
NLS—Jupyter Notebooks	0.92	1	1	1	1	0	1	0.95	0.5	1	1	1	0	10.37
Penn Libraries—Jupyter Notebooks	0.96	1	1	0.6	1	0.5	1	0.89	0.5	1	0	1	0	9.45
VAM—Data Explorations	1	1	1	0.92	1	0.5	1	0.99	0.5	1	0	1	0	9.91

PDF file (German National Library, 2021). The predominant language used for documentation in the list of projects analyzed is English. Documentation provided in other languages may facilitate the reuse of the materials, in particular, projects published by institutions located in countries in which English is not the official language. In general, the *naming* used for the notebooks are descriptive. Note that some filenames contained embedded spaces and they can cause issues (e.g., scripts considering there are no space characters in filenames). In general, the notebooks include the *output* of the code cells. If required, the interface provided by Jupyter enables users to clean the output of the cells. The *intended audience* is not explicitly stated in the projects. However, it can often be identified by reviewing the content of the notebooks, the datasets and software libraries used. Most of the projects are dedicated to a wide audience including beginners and advanced users, providing examples arranged according to the level of difficulty and the complexity of the issues. Only the BVMC and Newseye are dedicated to advanced users since these projects are based on the application of more sophisticated techniques based on NLP and the use of knowledge graphs. In terms of *metadata*, in general the projects provide human-readable metadata to describe their notebook collections.

The **availability** of most of the projects is high as they are available using *open licenses* (e.g., CC-BY). The projects provide the licensing information in several ways such as a README file, LICENSE file and dedicated websites. Using a consistent and standard method of publishing this information may help users to clearly identify the license upon the contents are available. In some cases, the projects refer to notebooks provided by other institutions. For instance, the projects provided by the BL and BVMC provide references to external notebooks.

In terms of **efficiency**, larger sizes are found in some examples such as Archives Unleashed. Another relevant aspect to consider is whether the organizations providing the datasets are in compliance with green policies (Dodd et al., 2020).

Regarding **traceability**, and according to the *versioning* criterion, all the projects analyzed use GitHub as a code repository. Note that only a few of them use the versioning and release features when publishing their projects, such as the GLAM Workbench. Platforms such as Zenodo provide integrated environments to facilitate the administration and publication of releases based on GitHub repositories.[32] Zenodo enables the automatic creation and preservation of citable archived repositories of code that are available in public platforms such as GitHub.

Portability achieves a lower score since some of the projects do not include a standard file to describe the *dependencies* used. Binder is a cloud environment that can be used to run a Jupyter Notebook. However, Binder requires a dependencies file to create the environment in which the notebooks will be run. Only a few projects provide the dependencies file (e.g., requirements.txt) and the link to Binder. Only some of them link to Google Colaboratory (Colab). It is important to notice that when using Binder there are limitations: (i) memory used (a maximum of 2GB); (ii) a time-out will occur after 10 min of inactivity; and (iii) the maximum concurrent users per repository is 100.[33]

With regard to the **recoverability** dimension, only four of the projects provide a *persistent URI* and full citation information. This issue can be solved by adopting a citation solution such as `CITATION.cff` and using online platforms such as Zenodo that provides a DOI for the publications.

With regard to the **credibility** dimension, all the projects provide provenance information about the owner, datasets used, references, awards, and so forth demonstrating their *trustworthiness*.

The methodology provided addresses a selection of quality measures. Some ways in which the methodology could be extended or improved based on best practices include: (i) the assessment of the quality of the code provided by the Jupyter Notebooks by using Python packages such as Pylint[34]; (ii) the use of validation tools to test all the Jupyter Notebooks included in a repository when updating Python packages[35]; and (iii) the integration of a Black[36] extension for Jupyter to apply standard formatting to the notebooks.

5 | IMPROVING THE DISCOVERABILITY OF JUPYTER NOTEBOOKS: A WIKIDATA APPROACH

Search engines have started to harvest and process machine-readable cross-domain metadata provided by knowledge bases such as YAGO, DBpedia and Wikidata

TABLE 4 Wikidata identifiers for the Jupyter Notebook projects.

Project	URL
BL's Jupyter Notebooks	https://www.wikidata.org/wiki/Q111421205
BVMC—GLAM Jupyter Notebooks	https://www.wikidata.org/wiki/Q111396450
GLAM Workbench	https://www.wikidata.org/wiki/Q111396660
LoC Jupyter Notebooks	https://www.wikidata.org/wiki/Q111450546
NLS—Jupyter Notebooks	https://www.wikidata.org/wiki/Q111411199

TABLE 5 Overview of Wikidata properties used to describe the Jupyter Notebook projects.

Property	Description
Instance of (`P31`)	The class of which the project is an example. The values used for this property are the items software (`Q7397`) and collection (`Q2668072`)
Owned by (`P127`)	Organization owner of the subject. For instance, the British Library (`Q23308`)
Source code repository (`P1324`)	The GitHub repository of the project
Official website (`P856`)	Official website of the project
Uses (`P2283`)	Each dataset used in the project will be described by means of this property
Country (`P17`)	Sovereign state of the project. For instance, Spain (`Q29`) or Scotland (`Q22`)
Logo image (`P154`)	Graphic mark used to identify the project
Title (`P1476`)	Title of the project
Inspired by (`P941`)	Work, human, place or event which inspired the project
Award received (`P166`)	Used to describe that the project received an award
Depends on software (`P1547`)	The project depends on the software Jupyter Notebook (`Q105099901`)
Copyright license (`P275`)	License under which the project is released. For instance, CC-BY (`Q20007257`)
Software version identifier (`P348`)	Numeric or nominal identifier of a version of a software program

in order to increase the visibility of resources (Färber et al., 2018). In particular, Wikidata enables the crowdsourced curation of items by the community providing rich and updated information about resources.

While most of the projects analyzed in this paper provide rich metadata in different ways such as dedicated websites, README files and more advanced methods like CITATION.cff plain-text files, none of them are included in a collaborative editing platform such as Wikidata.

Collaborative editing platforms such as Wikidata have become popular in the GLAM sector providing an environment to describe resources using structured information by means of properties (e.g., official website, creator and title) (Jean et al., 2019; Padilla et al., 2019). Wikidata stores the information in a knowledge graph as structured machine-readable data using the Linked Data principles (Tim Berners-Lee, 2006). Wikidata defines entities to identify resources and properties to add information about them. The information is publicly available by means of an open SPARQL API (World Wide Web Consortium, 2013b).

A selection of the projects analyzed in Section 4 have been published in Wikidata to provide machine-readable metadata. Table 4 shows the Wikidata identifiers for the projects edited and Table 5 shows the properties used to describe them. Listing 1 shows a SPARQL sentence to retrieve all the information stored in Wikidata about the Jupyter Notebooks projects provided by GLAM institutions described in Table 4.

LISTING 1 Query used to retrieve all the information stored in Wikidata about the Jupyter Notebooks projects provided by GLAM institutions. The instruction *values* is used to provide the identifiers of the entities used in the sentence.

```
SELECT DISTINCT ?nbs ?nbsLabel ?linknb
                ?ownerLabel
                ?maintainerLabel
                ?datasetLabel
WHERE
{
  values ?nbs {wd:Q111421153
               wd:Q111421205
               wd:Q111450546}
  ?nbs wdt:P2283 ?dataset .
  ?nbs wdt:P856 ?linknb .
  OPTIONAL {?dataset wdt:P127 ?owner.}
  OPTIONAL {?dataset wdt:P126 ?maintainer.}
  SERVICE wikibase:label {bd:serviceParam
    wikibase:language "en" }
}
```

As an example of application and reuse of Wikidata, the International GLAM Labs website reuses the information provided by Wikidata about the Jupyter Notebooks projects provided by its members for a new section regarding the computational access to digital collections.[37] The section includes a chart to explore the relationships between the datasets used in the projects, using the data visualization capabilities of the Wikidata SPARQL endpoint.

6 | CONCLUSIONS AND FUTURE WORK

Over the past few years, there has been a growing interest in publishing and reusing the digital collections made available by GLAM institutions. Jupyter Notebooks are becoming increasingly popular in the GLAM community as a key mode of fostering reuse of materials made available by such organizations.

Based on previous work, we defined a methodology for assessing the quality of projects based on Jupyter Notebooks and published by relevant GLAM institutions. This includes a proposed set of quality criteria: *understandability*, *availability*, *efficiency*, *traceability*, *portability*, *recoverability*, and *credibility*. The methodology was applied to assess the quality of 11 Notebook projects. Our evaluation showed that the methodology can be useful to identify common mistakes and best practices. In addition, a machine-readable metadata description model has been proposed to enhance discoverability.

Future work to be explored includes the evaluation of additional collections of Jupyter Notebooks, the extension of the quality criteria and the analysis of the usability and understandability of the notebooks for example, using qualitative methods, together with the intended audiences. In addition, the use of an ontology to make available the results, the automation of the assessment process and the inclusion of the carbon impact as a criterion will be explored. Finally, the use of Jupyter Notebooks for improving the digital literacy of both cultural heritage professionals and digital humanities researchers is another further avenue of potential exploration.

ACKNOWLEDGMENTS

The authors would like to thank Sarah Ames from the National Library of Scotland for her feedback on an initial draft of this article. The authors would also like to note that two of the authors of this article; Gustavo Candela and Tim Sherratt, created notebook projects (the GLAM Jupyter Notebooks at the Biblioteca

Virtual Miguel de Cervantes [BVMC] and the GLAM Workbench) which were evaluated as part of this article.

ORCID

Gustavo Candela https://orcid.org/0000-0001-6122-0777
Sally Chambers https://orcid.org/0000-0002-2430-475X
Tim Sherratt https://orcid.org/0000-0001-7956-4498

ENDNOTES

[1] https://jupyter.org/.

[2] Based on a search by programming language available at https://github.com/search?q=language:"Jupyter+Notebook"&type=Repositories.

[3] See more details at https://huggingface.co/blog/notebooks-hub.

[4] See, for example, the six Jupyter Notebook projects at https://glamlabs.io/computational-access-to-digital-collections/.

[5] https://glamlabs.io/list-members/.

[6] https://github.com/GLAM-Workbench/glam-workbench-template.

[7] https://www.si.edu/openaccess.

[8] https://metmuseum.github.io/.

[9] https://www.loc.gov/collections.

[10] https://chroniclingamerica.loc.gov/.

[11] https://github.com/NewsEye/NLP-Notebooks-Newspaper-Collections.

[12] https://constellate.org/.

[13] See, for example, https://blogs.bl.uk/digital-scholarship/2022/10/open-and-engaged-2022.html and https://sas-dhrh.github.io/dhcc-toolkit/.

[14] http://literateprogramming.com/.

[15] https://query.wikidata.org/.

[16] https://creativecommons.org/.

[17] https://creativecommons.org/licenses/by/4.0/.

[18] https://cloudblogs.microsoft.com/industry-blog/en-gb/technetuk/2021/10/12/how-to-measure-and-reduce-the-carbon-footprint-of-your-application/.

[19] See, for example, the property https://www.wikidata.org/wiki/Property:P348.

[20] https://schema.org/softwareVersion.

[21] https://research.google.com/colaboratory.

[22] https://repo2docker.readthedocs.io.

[23] https://arks.org.

[24] https://www.doi.org.

[25] https://www.handle.net/index.html.

[26] https://zenodo.org/.

[27] https://datacite.org/.

[28] See, for example, https://zenodo.org/record/5495619.

[29] https://www.coretrustseal.org/.

[30] See, for example, https://data.bl.uk/bllabsawards/.

[31] https://nbformat.readthedocs.io/en/latest/.

[32] See, for example, https://zenodo.org/record/5584195.

[33] https://mybinder.readthedocs.io/en/latest/about/user-guidelines.html#resources-available.

[34] https://pypi.org/project/pylint/.

[35] https://nbval.readthedocs.io/.

[36] https://black.readthedocs.io/.

[37] See the new computational access section at https://glamlabs.io/computational-access-to-digital-collections/.

REFERENCES

Ames, S., & Havens, L. (2022). Exploring National Library of Scotland datasets with Jupyter Notebooks. *IFLA Journal*, *48*(1), 50–56. https://doi.org/10.1177/03400352211065484

Australian Research Data Commons. (2022). *A national agenda for research software*. https://doi.org/10.5281/zenodo.6378082

Australian Research Data Commons. (2023). *FAIR for Jupyter Notebooks: A practical guide*. https://ardc.edu.au/resource/fair-for-jupyter-notebooks-a-practical-guide/

Ayris, P. (2010). The status of digitisation in Europe: Extensive summary of the second LIBER-EBLIDA workshop on the digitisation of library materials in Europe. *LIBER Quarterly*, *19*(3/4), 193–226. https://doi.org/10.18352/lq.7961

Bannour, N., Ghannay, S., Névéol, A., & Ligozat, A. (2021). Evaluating the carbon footprint of NLP methods: A survey and analysis of existing tools. *Proceedings of the Second Workshop on Simple and Efficient Natural Language Processing, SustaiNLP@EMNLP 2021*, Virtual, November 10, 2021, 11–21. https://aclanthology.org/2021.sustainlp-1.2

Barker, M., Chue Hong, N. P., Katz, D. S., Lamprecht, A.-L., Martinez-Ortiz, C., Psomopoulos, F., Harrow, J., Castro, L. J., Gruenpeter, M., Martinez, P. A., & Honeyman, T. (2022). Introducing the FAIR Principles for research software. *Scientific Data*, *9*(1), 622. https://doi.org/10.1038/s41597-022-01710-x

Beg, M., Taka, J., Kluyver, T., Konovalov, A., Ragan-Kelly, M., Thiéry, N. M., & Fangohr, H. (2021). Using Jupyter for reproducible scientific workflows. *Computing in Science & Engineering*, *23*(2), 36–46. https://doi.org/10.1109/MCSE.2021.3052101

Berlin State Library. (2022). *Qurator. automated curation technologies for the digitised cultural heritage*. https://ravius.sbb.berlin/

Berners-Lee, T. (2006). *Linked data*. https://www.w3.org/DesignIssues/LinkedData.html

Candela, G., Sáez, M. D., Esteban, M. E., & Marco-Such, M. (2022). Reusing digital collections from GLAM institutions. *Journal of Information Science*, *48*(2), 251–267. https://doi.org/10.1177/0165551520950246

CLARIN ERIC. (2022). *Jupyter Notebooks for Europeana newspaper text resource processing with CLARIN NLP tools*. https://marketplace.sshopencloud.eu/training-material/duVII1

Dodd, N., Felice, A., Oliveira, M. N. D., Maya-Drysdale, L., Viegand, J., Flucker, S., Tozer, R., Whitehead, B., Wu, A., & Brocklehurst, F. (2020). *Development of the EU green public procurement (GPP) criteria for data centres, server rooms and cloud services* (Policy assessment). Publications Office of the European Union. https://doi.org/10.2760/964841

Druskat, S., Spaaks, J. H., Chue Hong, N., Haines, R., Baker, J., Bliven, S., Willighagen, E., Pérez-Suárez, D., & Konovalov, A.

(2021). *Citation file format*. https://doi.org/10.5281/zenodo.5171937

European Commission. (2020). *Energy-efficient cloud computing technologies and policies for an ecofriendly cloud market*. https://digital-strategy.ec.europa.eu/en/library/energy-efficient-cloud-computing-technologies-and-policies-eco-friendly-cloud-market

European Commission. (2019). *The European green deal*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0640&from=EN

Färber, M., Bartscherer, F., Menne, C., & Rettinger, A. (2018). Linked data quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. *Semantic Web*, *9*(1), 77–129. https://doi.org/10.3233/SW-170275

German National Library. (2021). *Jupyter Notebook tutorials in the browser*. https://www.dnb.de/SharedDocs/Downloads/EN/Professionell/Services/dnblabJupyterNotebookKurzanleitung.pdf

Granger, B. E., & Pérez, F. (2021). Jupyter: Thinking and storytelling with code and data. *Computing in Science & Engineering*, *23*(2), 7–14. https://doi.org/10.1109/MCSE.2021.3059263

Gruber, J. (2004). *Markdown*. https://daringfireball.net/projects/markdown/

Grus, J. (2018). *I don't like notebooks*. https://conferences.oreilly.com/jupyter/jup-ny/public/schedule/detail/68282.html

Hughes, L. M. (2004). *Digitizing collections: Strategic issues for the information manager*. Facet Publishing.

ISO 25000. (2014). *ISO/IEC 25012*. https://iso25000.com/index.php/en/iso-25000-standards/iso-25012

Jean, G., Smith-Yoshimura, K., Washburn, B., Davis, K., Detling, K., Eslao, C. F., Folsom, S., Li, X., McGee, M., Miller, K., Moody, H., Tomren, H., & Thomas, C. (2019). *Creating library linked data with Wikibase: Lessons learned from project passage*. https://doi.org/10.25333/faq3-ax08

Jupyter, P., Bussonnier, M., Forde, J., Freeman, J., Granger, B., Head, T., Holdgraf, C., Kelley, K., Nalvarte, G., Osheroff, A., Pacer, M., Panda, Y., Perez, F., Ragan-Kelley, B., & Willing, C. (2018). Binder 2.0—Reproducible, interactive, sharable environments for science at scale. In F. Akici, D. Lippa, D. Niederhut, & M. Pacer (Eds.), *Proceedings of the 17th Python in Science Conference* (pp. 113–120). https://doi.org/10.25080/Majora-4af1f417-011

Koch, G., & Kinder-Kurlanda, K. (2020). Source criticism of data platform logics on the internet. *Historical Social Research*, *45*(3), 270–287. https://doi.org/10.12759/hsr.45.2020.3.270-287

Koolen, M., van Gorp, J., & van Ossenbruggen, J. (2019). Toward a model for digital tool criticism: Reflection as integrative practice. *Digital Scholarship in the Humanities*, *34*(2), 368–385. https://doi.org/10.1093/llc/fqy048

Lacoste, A., Luccioni, A., Schmidt, V., & Dandres, T. (2019). *Quantifying the carbon emissions of machine learning*. CoRR, abs/1910.09700. http://arxiv.org/abs/1910.09700

Lee, B. C. G. (2022). *The "Collections as ML Data" checklist for machine learning & cultural heritage*. https://doi.org/10.48550/ARXIV.2207.02960

Lorang, E., Soh, L.-K., Liu, Y., & Pack, C. (2020). *Digital libraries, intelligent data analytics, and augmented description: A demonstration project*. https://digitalcommons.unl.edu/libraryscience/396/

Madden, F., van Horik, R., van de Sandt, S., Lavasa, A., & Cousijn, H. (2020). *Guides to choosing persistent identifiers—Version 3*. https://doi.org/10.5281/zenodo.4192174

Mahey, M., Al-Abdulla, A., Ames, S., Bray, P., Candela, G., Derven, C., Dobreva-McPherson, M., Gasser, K., Chambers, S., Karner, S., Kokegei, K., Laursen, D., Potter, A., Straube, A., Wagner, S.-C., & Wilms, L. (2019). *Open a GLAM lab*. International GLAM Labs Community, Book Sprint. https://doi.org/10.21428/16ac48ec.f54af6ae

Murphy, O., & Villaespesa, E. (2020). *AI: A museum planning toolkit*. https://themuseumsai.network/toolkit/

NASA. (2022). *ARSET—Measuring atmospheric carbon dioxide from space in support of climate related studies*. https://appliedsciences.nasa.gov/join-mission/training/english/arset-measuring-atmospheric-carbon-dioxide-space-support-climate

Nielsen, E. K. (2008). Digitisation of library material in Europe: Problems, obstacles and perspectives anno 2007. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, *18*(1), 20–27. https://doi.org/10.18352/lq.7901

Oli, P., Banjade, R., Tamang, L. J., & Rus, V. (2021). Automated assessment of quality of Jupyter Notebooks using artificial intelligence and big code. In E. Bell & F. Keshtkar (Eds.), *Proceedings of the Thirty-Fourth International Florida Artificial Intelligence Research Society Conference*, North Miami Beach, Florida, USA, May 17–19, 2021. https://doi.org/10.32473/flairs.v34i1.128560

Open Knowledge Foundation. (2015). *Guide to open licensing*. https://opendefinition.org/guide/

Padilla, T. (2019). *Responsible operations: Data science, machine learning, and AI in libraries*. https://doi.org/10.25333/xk7z-9g97

Padilla, T., Allen, L., Frost, H., Potvin, S., Russey Roke, E., & Varner, S. (2019). *Final report—Always already computational: Collections as data*. https://doi.org/10.5281/zenodo.3152935

Pimentel, J. F., Murta, L., Braganholo, V., & Freire, J. (2019). A large-scale study about quality and reproducibility of Jupyter Notebooks. In *Proceedings of the 16th International Conference on Mining Software Repositories, MSR 2019*, 26–27 May 2019, Montreal, Canada (pp. 507–517). IEEE/ACM. https://doi.org/10.1109/MSR.2019.00077

Pimentel, J. F., Murta, L., Braganholo, V., & Freire, J. (2021). Understanding and improving the quality and reproducibility of Jupyter Notebooks. *Empirical Software Engineering*, *26*(4), 65. https://doi.org/10.1007/s10664-021-09961-9

Rule, A., Birmingham, A., Zuniga, C., Altintas, I., Huang, S., Knight, R., Moshiri, N., Nguyen, M. H., Rosenthal, S. B., Pérez, F., & Rose, P. W. (2019). Ten simple rules for writing and sharing computational analyses in Jupyter Notebooks. *PLoS Computational Biology*, *15*(7), e1007007. https://doi.org/10.1371/journal.pcbi.1007007

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, *63*(12), 54–63. https://doi.org/10.1145/3381831

Shenoy, K., Ilievski, F., Garijo, D., Schwabe, D., & Szekely, P. A. (2022). A study of the quality of Wikidata. *Journal of Web Semantics*, *72*, 100679. https://doi.org/10.1016/j.websem.2021.100679

Sherratt, T. (2021). *GLAM workbench*. https://doi.org/10.5281/zenodo.5603060

Steensen, S., Belair-Gagnon, V., Graves, L., Kalsnes, B., & Westlund, O. (2022). Journalism and source criticism. Revised approaches to assessing truth-claims. *Journalism Studies*, *23*(16), 2119–2137. https://doi.org/10.1080/1461670X.2022.2140446

Vandegrift, M., & Varner, S. (2013). Evolving in common: Creating mutually supportive relationships between libraries and the digital humanities. *Journal of Library Administration*, *53*(1), 67–78. https://doi.org/10.1080/01930826.2013.756699

Wilms, L. (2021). Digital humanities in European research libraries: Beyond offering digital collections. *Liber Quarterly: The Journal of European Research Libraries*, *31*(1), 1–23. https://doi.org/10.18352/lq.10351

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., & Brew, J. (2019). *Huggingface's transformers: State-of-the-art natural language processing*. CoRR, abs/1910.03771. http://arxiv.org/abs/1910.03771

World Wide Web Consortium. (2013a). *PROV-O: The PROV Ontology*. https://www.w3.org/TR/prov-o/

World Wide Web Consortium. (2013b). *SPARQL 1.1 query language*. https://www.w3.org/TR/sparql11-query/

World Wide Web Consortium. (2016). *Data on the web best practices: Data quality vocabulary*. https://www.w3.org/TR/vocab-dqv/

World Wide Web Consortium. (2017). *Data on the web best practices*. https://www.w3.org/TR/dwbp/

Xie, Y. (2018). *The first notebook war*. https://yihui.org/en/2018/09/notebook-war

Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., & Auer, S. (2016). Quality assessment for linked data: A survey. *Semantic Web*, *7*(1), 63–93. https://doi.org/10.3233/SW-150175

How to cite this article: Candela, G., Chambers, S., & Sherratt, T. (2023). An approach to assess the quality of Jupyter projects published by GLAM institutions. *Journal of the Association for Information Science and Technology*, 1–15. https://doi.org/10.1002/asi.24835