

A Checklist to Publish Collections as Data in GLAM Institutions

<input type="checkbox"/>	01	Provide a clear license allowing reuse of the dataset without restrictions	<ul style="list-style-type: none">• CC0, CC BY, Public Domain Mark• National initiatives• No known copyright
<input type="checkbox"/>	02	Provide a suggestion of how to cite the dataset	<ul style="list-style-type: none">• BibTeX, APA• DOI• Versions
<input type="checkbox"/>	03	Include documentation about the dataset	<ul style="list-style-type: none">• README file• Tutorials and websites• Programming Historian & research articles
<input type="checkbox"/>	04	Use a public platform to publish the dataset	<ul style="list-style-type: none">• GitHub, Zenodo, DataCite• Hosting
<input type="checkbox"/>	05	Share examples of use as additional documentation	<ul style="list-style-type: none">• Prototypes & tools• Jupyter Notebooks (reproducible)• GLAM Labs
<input type="checkbox"/>	06	Give structure to the dataset	<ul style="list-style-type: none">• Folder structure• Using self-describing folder and file names• BagIt File Packaging Format & Data Package
<input type="checkbox"/>	07	Provide machine-readable metadata (about the dataset itself)	<ul style="list-style-type: none">• Dublin Core• Vocabulary of Interlinked Datasets (VoID)• Data Catalog Vocabulary (DCAT)
<input type="checkbox"/>	08	Include your dataset in collaborative edition platforms	<ul style="list-style-type: none">• Increase visibility• Title, author, location, license, main subject, etc.
<input type="checkbox"/>	09	Offer an API access to your repository	<ul style="list-style-type: none">• OAI-PMH, JSON, XML• IIIF• SPARQL
<input type="checkbox"/>	10	Develop a portal page	<ul style="list-style-type: none">• GitHub Pages• New section in the Lab• e.g. Chronicling America & Data Foundry
<input type="checkbox"/>	11	Add a terms of use	<ul style="list-style-type: none">• e.g. section detailing copyright, liability and access statements

<https://doi.org/10.48550/arXiv.2304.02603>