

# Citation builder: specification document

Target release	Paperpal release: Q2 FY24
Issue	<div><div></div> AH-3869: Citation generator / Reference formatter <div>WORK IN PROGRESS</div></div>
Document status	<div>DRAFT</div>
Task owner	@Govind
Tech lead	@Zaineb Penwala
Tech scoping	Citation generator

## Overview

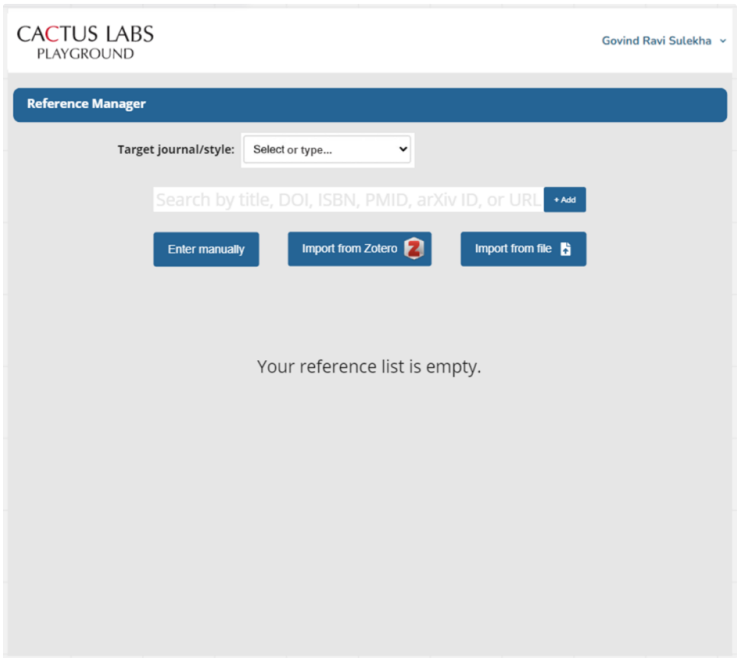
The citation builder is designed to simplify the process of creating references and citations by providing a platform for users to easily import, search, upload, or manually enter references. The feature uses the Citation Style Language (CSL) and supports 9000+ styles, covering nearly all of academic publication. Under the hood, the citation builder is based on a popular JavaScript-based implementation of CSL called [citeproc-js](#).

Development will be in a phased manner.

## User input methods

The citation builder will support four user input methods:

- Import from third-party apps:** Users will be able to import references from third party apps, starting with Zotero in phase 1. Users must login to their third-party account and select the desired references/collections to import. The idea is to extend this functionality to R Discovery and strengthen the ecosystem when Paperpal launches a citation builder.
- Search by Title/DOI:** Users can search for references by entering the title or DOI of the publication. The citation builder will then retrieve the corresponding reference information from Dyson.
- Upload BibTeX/JSON Files:** Users can upload BibTeX or JSON files containing references.
- Manual Entry:** Users can manually enter reference information by filling out a form that includes fields such as reference type, author names, title, and year of publication.



Home screen of the citation builder. All four input methods are illustrated. The home screen also allows the user to select the desired style.

## Phase 1: Zotero import

In phase 1, only the Zotero import method will be supported. Users will be able to create reference lists and reorder references manually. They will also be able to copy inline citations to clipboard. In phase 1, the API will support only inline citations to a single reference.

zotero

[Groups](#) [Documentation](#) [Forums](#) [Get Involved](#) [Log In](#) [Upgrade Storage](#)

[Register for a free account](#) [Forgot your password?](#)

## Login

Username or Email

Password

☐ Remember Me

Keep me signed in

[Login to Zotero](#)

[Log in with OpenID](#)

zotero

Welcome, govinds · [Settings](#) · [Inbox](#) · [Download](#) · [Log Out](#)

[Upgrade Storage](#)

[Home](#) [Web Library](#) [Groups](#) [Documentation](#) [Forums](#) [Get Involved](#)

[Home](#) > [Settings](#) > [Feeds/API](#) > New Key

## New Private Key

An application would like to connect to your account

The application test\_1 would like to access your account.  
Create a new private key to share with a third party so they can access your data.

Key Description

test\_1

[Accept Defaults](#)

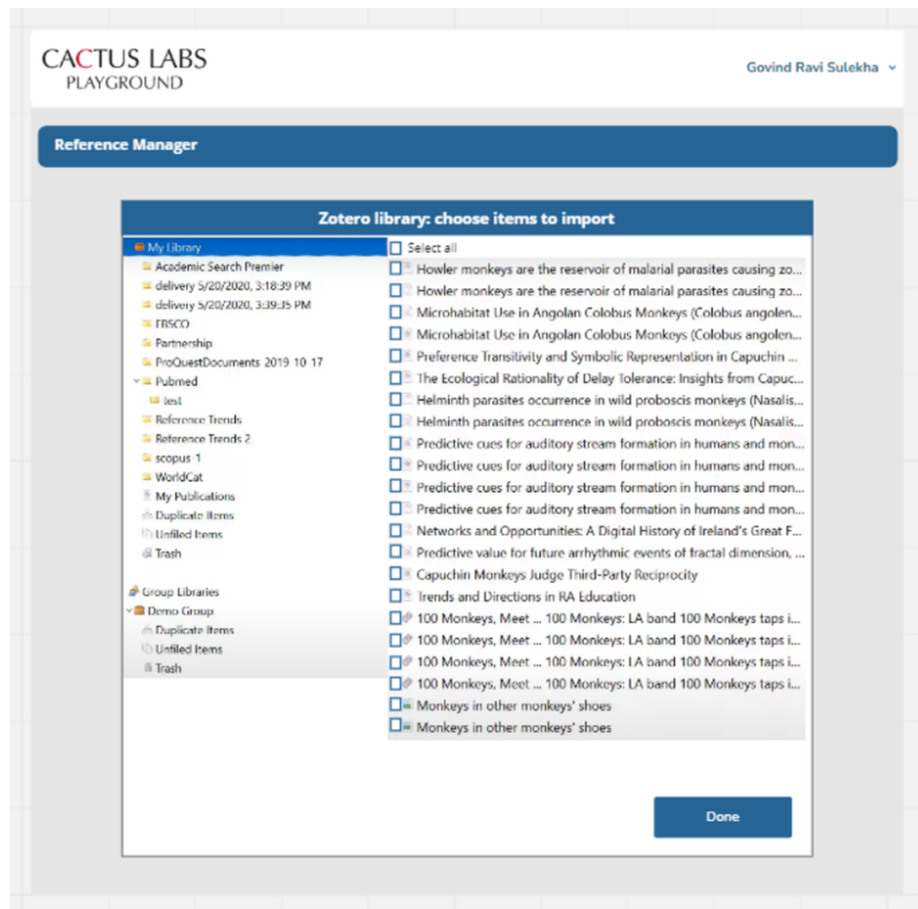
[Change Permissions](#)

**Personal Library**

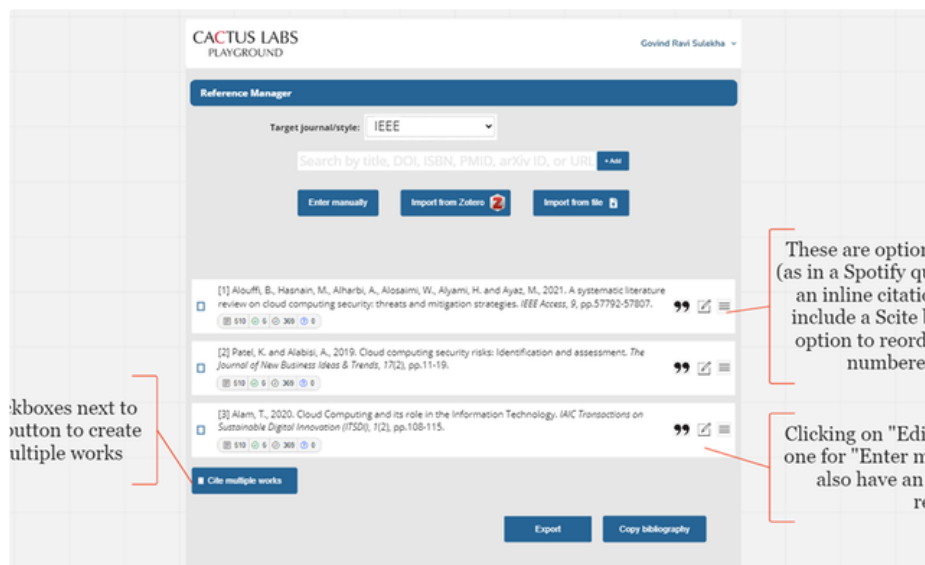
☒ Allow library access  
Allow third party to access your library.

☐ Allow notes access  
Allow third party to access your notes.

Upon clicking the "Import from Zotero" button, users are taken to the Zotero login page and then the account access permissions page.



Subsequently, users can select references in their Zotero library through a file-explorer-like screen. Collections/folders are shown on the left; references with checkboxes are shown on the right.



Home screen after the addition of 3 references in the IEEE style. Note that the "Cite multiple works" functionality will be added in phase 3. Each reference is accompanied by a Scite badge, showing citation statistics. Note that the "Edit" button will be added only in Phase 2, along with support for Manual Entry. Phase 1 includes the generate inline citation option and the reorder option (in the case of numeric styles).

## Scite badge

Each citation will be accompanied by a Scite badge, if available. The Scite badge indicates the health of a reference with detailed citation statistics (number of citations, supporting citation statements, contradicting citation statements, mentioning citation statements, and retractions, if any). The Scite badge is returned along with references in the JSON response of the citation builder API.

[Scite report](#)

## Differences between numeric and alphabetical citation styles

In numeric styles (e.g., IEEE and AMA), references are ordered as per the user's wish. In alphabetical styles (e.g., APA and MLA), citations are ordered alphabetically (alphabetical order of the first author's surname). There is a complex logic for alphabetical ordering when two references have the same first author.

## Master list of numeric/alphabetical styles

For **numeric styles**, the UI should support the reordering of references. Any change to the order will require changes to the citation numbers, which can be handled in the frontend without additional API calls. Further, **when new reference(s) are added to an existing list**, the API output will have citation number starting from "1." These numbers will need to be revised in the frontend to reflect the position(s) of the new reference(s) in the reference list.

For alphabetical styles, the UI will not support reordering. When new reference(s) are added to an existing (alphabetically ordered) list, a new API call should be performed with the *complete list of reference IDs* to place the new reference(s) at the correct positions in alphabetical order.

## API response structure

For each item, the API response includes the following components: `inline_citation`, `full_citation`, and `scite_badge_data`. Note that the text in `inline_citation` and `full_citation` contain HTML tags for bold, italics, superscript, subscript, underline, small caps, etc. These will need to be parsed appropriately.

```

"status": true,
"key": "IIR-B",
"inline_citation": "Zhongjie <span style='font-variant:small-caps;'>Wu</span>, Renata M. <span style='font-variant:small-caps;'>Wentzcovitch</span>, Koichiro <span style='font-variant:small-caps;'>Umemoto</span>, Baosheng <span style='font-variant:small-caps;'>Li</span>, Kei <span style='font-variant:small-caps;'>Hirose</span> & Jin-Cheng <span style='font-variant:small-caps;'>Zheng</span>, « Pressure-volume-temperature relations in MgO: An ultrahigh pressure-temperature scale for planetary sciences applications », </div><div>Journal of Geophysical Research</div>, 113-B6, 2008, p. B06204.",
"full_citation": [
  "data": "Zhongjie <span style='font-variant:small-caps;'>Wu</span>, Renata M. <span style='font-variant:small-caps;'>Wentzcovitch</span>, Koichiro <span style='font-variant:small-caps;'>Umemoto</span>, Baosheng <span style='font-variant:small-caps;'>Li</span>, Kei <span style='font-variant:small-caps;'>Hirose</span> & Jin-Cheng <span style='font-variant:small-caps;'>Zheng</span>, « Pressure-volume-temperature relations in MgO: An ultrahigh pressure-temperature scale for planetary sciences applications », </div><div>Journal of Geophysical Research</div>, 113-B6, 2008, p. B06204."
],
"scite_badge_data": {
  "status": true,
  "message": "only class='scite-badge' data-doi='10.1029/2007JB006276' data-layout='horizontal' data-show-zero='true' data-small='true' data-show-labels='false' data-tally-show='true'></div><script async type='application/javascript' src='https://cdn.scite.ai/badge/scite-badge-latest.min.js'></script>"
}
}

```

Sample API response structure for an alphabetical style.

**For alphabetical styles,** full citation contains only one part: data.

```
{
  "status": true,
  "key": "ITEM-8",
  "inline_citation": "[8]",
  "full_citation": [
    {
      "data_1": "[8]",
      "data_2": "Z. Wu, R. M. Wentzcovitch, K. Umemoto, B. Li, K. Hirose, and J.-C. Zheng,
        "Pressure-volume-temperature relations in MgO: An ultrahigh pressure-temperature scale
        for planetary sciences applications," <i>J. Geophys. Res.</i>, vol. 113, no. B6, p.
        B06204, Jun. 2008, doi: 10.1029/2007JB005275."
    }
  ]
}
```

### Sample API response for a numeric style

**For numeric styles,** full\_citation contains two parts: data\_1 (citation number) and data\_2 (full citation text). If the user re-orders references, data\_1 values need to be revised.

## Phase 2 and beyond

**Phase 2:** add support for the other three input methods: searching by title/DOI (Dyson input), BibTeX/JSON file upload, and manual entry; also add support for an Edit button for each reference.

**Phase 3:** add support for multi-reference inline citations such as "[1-3]" and "(Smith et al., 2019; Maxwell, 2013a; Maxwell, 2013b)".

CACTUS LABS PLAYGROUND Govind Ravi Sulekha

Reference Manager

Add reference manually

Item Type: Book

Title: \_\_\_\_\_

Author: last name first name

Series: \_\_\_\_\_

Series Number: \_\_\_\_\_

Volume: \_\_\_\_\_

# of Volumes: \_\_\_\_\_

Edition: \_\_\_\_\_

Place: \_\_\_\_\_

Publisher: \_\_\_\_\_

Date: \_\_\_\_\_

Original Date: \_\_\_\_\_

# of Pages: \_\_\_\_\_

Language: \_\_\_\_\_

ISBN: \_\_\_\_\_

Short Title: \_\_\_\_\_

URL: \_\_\_\_\_

Accessed: \_\_\_\_\_

Archive: \_\_\_\_\_

Loc. in Archive: \_\_\_\_\_

Library Catalog: \_\_\_\_\_

Call Number: \_\_\_\_\_

Done

The fields below depend on the item type selected.

Form for manual entry. Note that the selection in the first field determines which subsequent fields will be displayed.

How it'll look in a text editor

CACTUS LABS PLAYGROUND Govind Ravi Sulekha

Reference Manager

Target journal/style: Select

Search by title, DOI Add

Your reference list is empty.

Import Zotero library

Import from file

Clean Slate Add/Edit Citation Load Sample Text

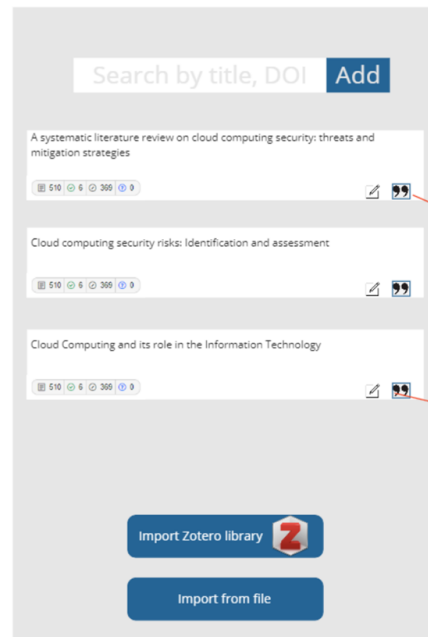
Text formatting tools may be shown here.

The user can add items to the reference list through a Dyson search.

Opens the Zotero sign-in page for importing the user's library

Optional; not required for MVP. File types like BibTeX and CSL JSON can be supported.

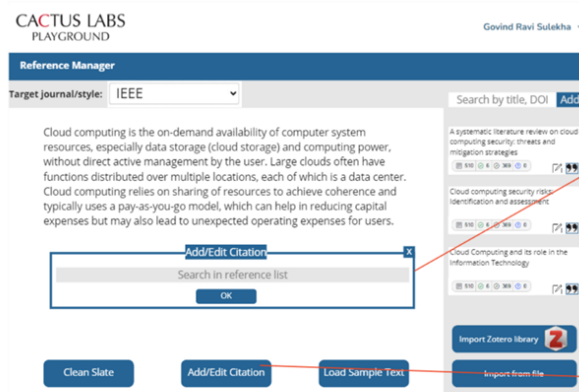
The citation builder shows up on the sidebar, similar to synonyms, consistency checker, etc.



Zoomed-in view of the sidebar.

Button 1: edit  
Button 3: Copy  
full/inline citation to  
clipboard.

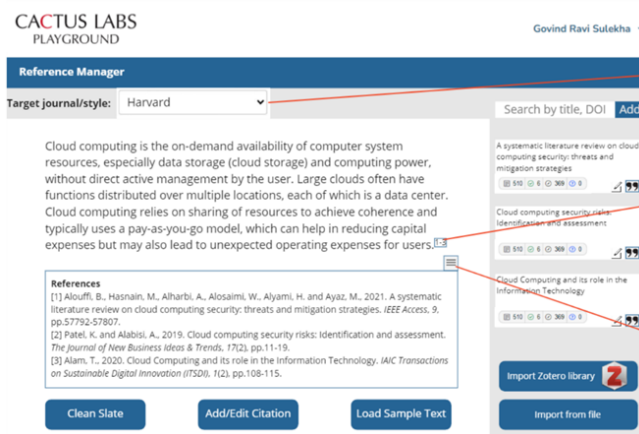
This list must be stored  
in our backend and not  
the user's Zotero library.  
Citations and  
bibliography are  
dynamic objects that are  
linked to items in this  
list.



The user can place the cursor  
anywhere in the text and click  
on "Add/Edit Citation" to  
open this search box. From  
this search box, the user can  
select one or more items from  
the reference list. Clicking on  
OK inserts an inline citation  
to the selected items at the  
cursor location. The search  
box behaves in the same  
manner as that of Zotero's  
Word add-in

Can also be used to edit  
an existing inline  
citation

After clicking on Add/Edit citation. The search box will be similar to the [search box in Zotero's Word add-in](#).



The user can change the  
target style at any time.  
The citations and  
bibliography are  
automatically updated.

Citations are selectable  
dynamic objects and  
not plain text.

The bibliography can be  
moved by clicking and  
dragging. Double-  
clicking hides the  
bibliography.

After adding 3 references.

## ⚠ Out of scope

The citation builder will not edit references in a free-text format. LLMs are currently not able to do this accurately. Editage uses a tool called [Merops](#) for this purpose.



Enter a URL, ISBN, DOI, PMID, arXiv ID, or title

Cite

Manual Entry

## Bibliography

IEEE

[1] E. Troja *et al.*, "A nearby long gamma-ray burst from a merger of compact objects," *Nature*, vol. 612, no. 7939, pp. 228-231, Dec. 2022, doi: 10.1038/s41586-022-05327-3.

[2] Google Quantum AI *et al.*, "Suppressing quantum errors by scaling a surface code logical qubit," *Nature*, vol. 614, no. 7949, pp. 676-681, Feb. 2023, doi: 10.1038/s41586-022-05434-1.

Home screen of zbib.