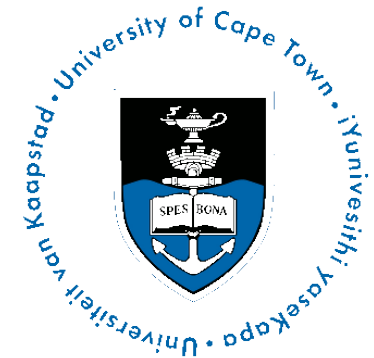


# Simple Archives

## *a different scale for data archives*



Hussein Suleman <[hussein@cs.uct.ac.za](mailto:hussein@cs.uct.ac.za)>

*Digital Libraries Laboratory  
Department of Computer Science  
School of IT  
University of Cape Town  
14 September 2021*

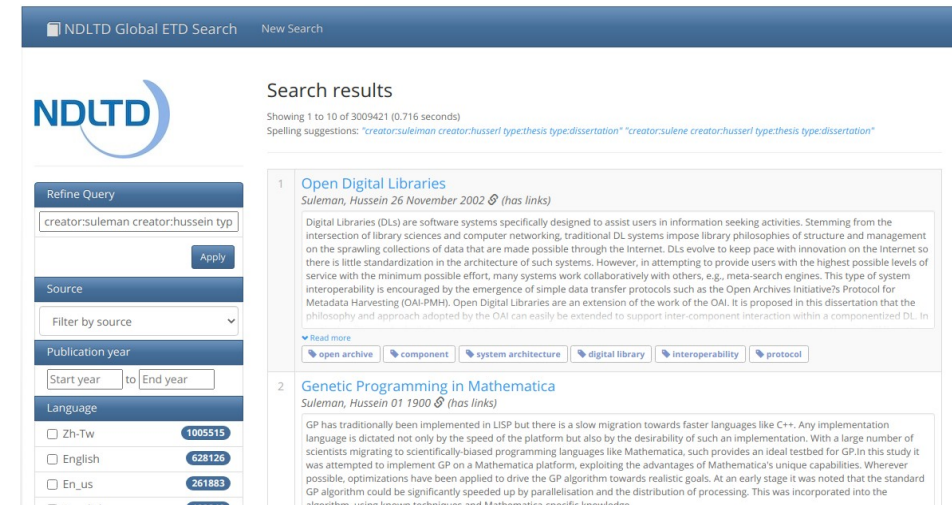


*Last edited: 13 September 2021*

*Slide count: 8*

# Scale in Digital Libraries

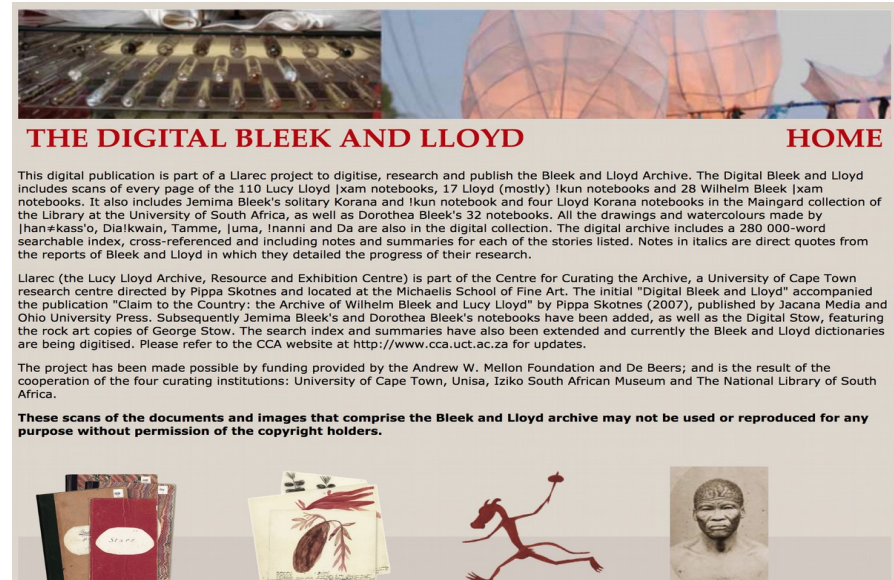
- ❑ Nothing starts off big.
- ❑ NDLTD Global ETD Search (<http://search.ndltd.org>) has millions of records - but based off many small archives.
- ❑ Scaling a single central server is easy.
- ❑ How do we support thousands of small archives?



<http://search.ndltd.org/>

# Scale in the Digital Humanities

- ❑ There isn't one metadata format and one object type.
- ❑ Every collection:
  - Is relatively small.
  - Is not particularly fixed.
  - Has unique metadata needs.
  - Has unique digital objects.
  - Has special interaction paradigms.
  - Encodes context/scholarship into archiving.
- ❑ How do we scale in this world?



<http://lloydbleekcollection.cs.uct.ac.za/>

# Scaling in Low-Resource Environments

- ❑ Low-resource environments:
  - Poor countries/regions.
  - No skilled technical staff.
  - Little sustainability for archives.
- ❑ Are archives designed for:
  - Rescue/Migration?
  - Archive decommissioning?
  - Minimal cost management?
  - Minimal resource utilisation?



[Login / Register](#)

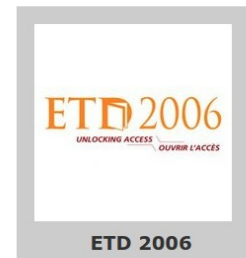
[Home](#) | [About](#) | [Authors](#) | [Search](#) | [Contact Us](#)

## Search Document Archive

Enter your search terms:

GO

## Browse Collections



## About Docs

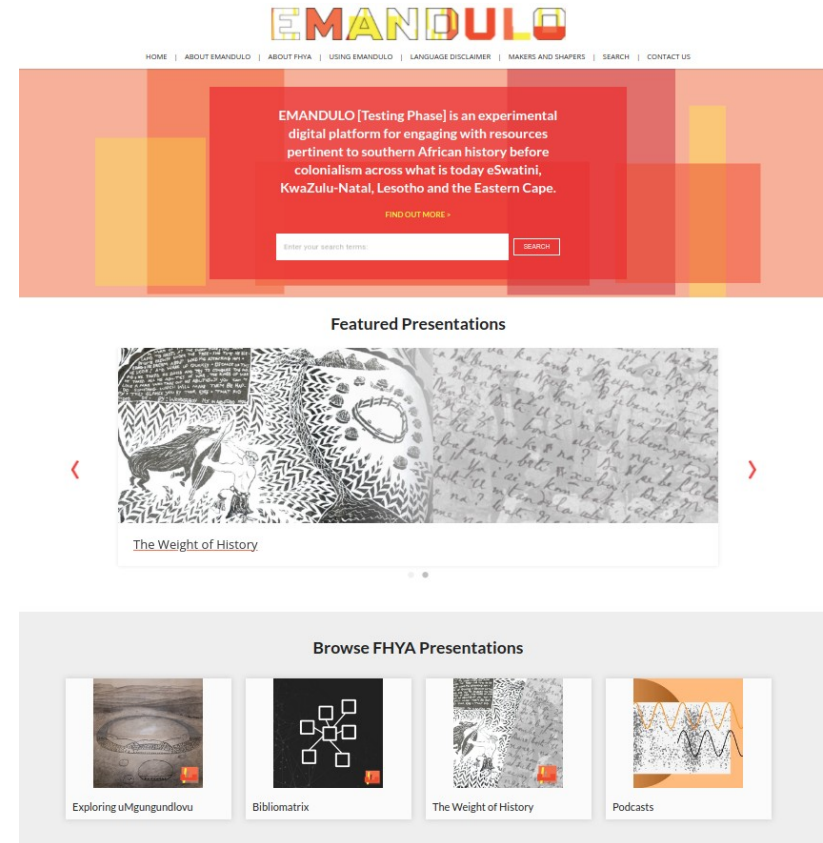
This is a collection of research papers, presentations and related files from the annual symposia of the [Networked Digital Library of Theses and Dissertations](#).

It is hosted on an instance of Simple DL.



# One Solution: Keep Things Simple

- ❑ Manage metadata using spreadsheets.
- ❑ Create metadata in simple static files.
- ❑ Create a website to organise and provide access.
- ❑ Use client-side scripts where needed.
- ❑ No Web application needed for access.
- ❑ No OS dependency.
- ❑ Simple design maximises flexibility.



<http://emandulo.apc.uct.ac.za/>

# Simple DL

- ❑ Toolkit to create simple archives.
- ❑ Based on many successful ideas in DL architecture, but built for low-resource environments and for adaptation.
- ❑ Experimental but:
  - Can change how we develop systems.
  - Provides a platform to experiment.

slumou / **simpliedl** Public

<> Code Issues Pull requests Actions Projects Wiki Security Insights

main 1 branch 0 tags

Go to file Code

About  
Simple Digital Library Toolkit  
Readme  
MIT License

Releases  
No releases published

Packages  
No packages published

Languages

Language	Percentage
JavaScript	47.2%
Perl	45.9%
Less	2.2%
Raku	0.4%
SCSS	2.4%
CSS	1.8%
Ruby	0.1%

Simple DL

This is an experimental toolkit to create simple digital libraries, where the goal is ease of preservation, ease of rescue and ease of development.

The toolkit manages and generates a self-contained Web experience around a collection of digital objects and their associated metadata.

Generation a site is a 3-step process: import data/configuration from sources; generate

<https://github.com/slumou/simpliedl>

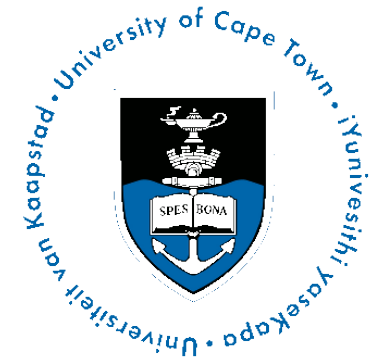


# Some Reflections

---

- ❑ One size does not fit all.
- ❑ One design/architecture does not meet all needs.
- ❑ How people engage with data is a rapidly evolving and unsettled idea.
- ❑ Context matters and bias is real.
- ❑ An archive is not the final location for an object.
- ❑ An archive is one temporary location for an object.

# Thank you for listening!



*Questions, comments and suggestions are welcome...*

Contact me at: [hussein@cs.uct.ac.za](mailto:hussein@cs.uct.ac.za)

*Digital Libraries Laboratory  
Department of Computer Science  
School of IT  
University of Cape Town*