# Digitally Preserving African Heritage

Hussein Suleman hussein@cs.uct.ac.za

University of Cape Town
Department of Computer Science
Centre for ICT for Development
Digital Libraries Laboratory





*April* 2016

# Why am I here?

- To talk about Digital Libraries/Preservation.
- To share some research findings.
- To collaborate and develop research links.
- To inspire you to think differently.
- To convince you to preserve our heritage!





## What is the Digital Libraries Lab?

- Research in technologies for research and education, specifically digital libraries:
  - African language search engines, machine translation
  - cultural heritage preservation
  - technology for education
- Teaching
  - 2 staff supervising about 20 MSc and PhD students
- Advocacy
  - a little wherever we can
- Collaboration
  - industry, academic, (govt?)







Mapungubwe Collection University of Pretoria







Timbuktu Manuscripts http://www.timbuktufoundation.org/Manuscripts/index.htm







Kirby Collection http://web.uct.ac.za/depts/sacm/kirby.html





#### SPECIAL ISSUE

15th November, 1966.

#### WE SHALL NEVER FORGET.

WE SHALL NEVER FORGET our comrades-in-arms who have paid the supreme price for the lofty cause of freedom for our country.

WE SHALL NEVER FORGET the callous murder of MINI, KHAYINGA, MKABA, BONGCO, NGUDIE, SALOOJEE and many others; their death remains a permanent indictment against the South African Apartheid Rulers.

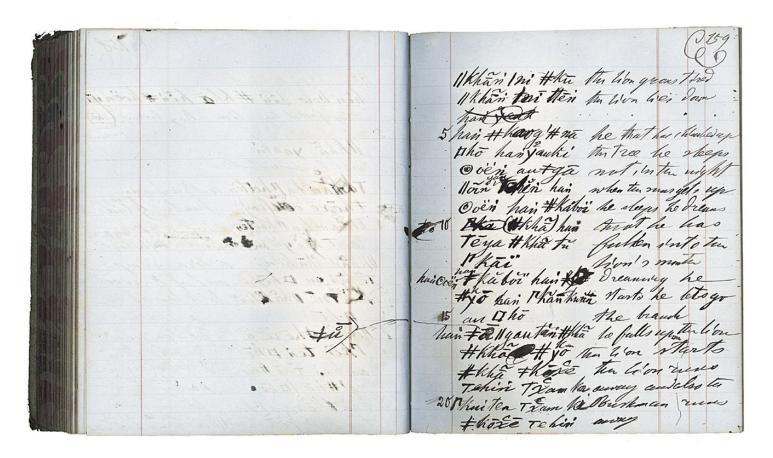
WE SHALL NEVER FORGET these comrades, who have braved Vorster's torture chambers and faced fascist brutality boldly in the face. Their sacrifices SHALL NOT BE IN VAIN.

WE SHALL NEVER FORGET our thousands of freedom-fighters who are languishing in Vorster's dungeons. These gallant sons and daughters of South Africa, whose only crime is their militant and uncompromising struggle against White Domination, symbolise the highest in morality. Some have been

Digital Imaging South Africa http://www.disa.ukzn.ac.za/



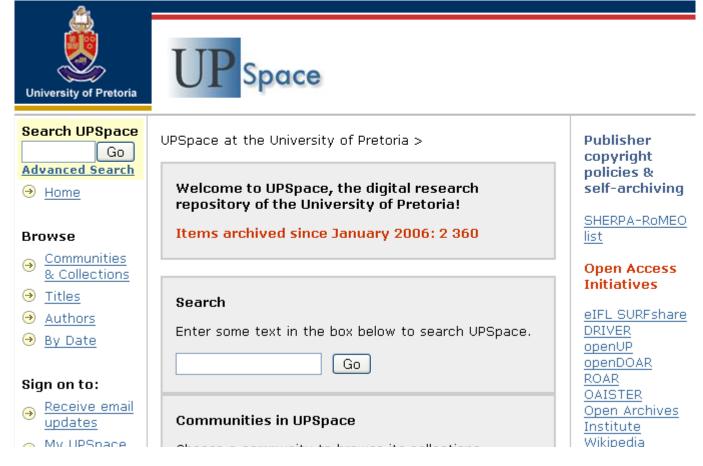




Bleek and Lloyd Collection http://www.lloydbleekcollection.uct.ac.za/



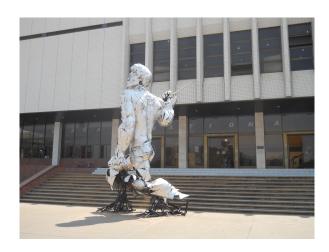




















### **Urgency**

- Some documents and storage media are rapidly deteriorating.
- Some storytellers are the last in their generations.







### **Rewriting History**



- We now know there were powerful ancient civilizations all over Africa.
  - Colonial governments suppressed this information for centuries!
- History must be preserved what little evidence we have left.





### **Skills and Education**



- Typical archivists are not as highly skilled as counterparts elsewhere.
- Digital media is still not the norm.
- Education levels of general population hinders preservation – end-user data curation is very difficult.



### **Funding**

- Typically, there is little.
- Many preservation projects are funded by external agencies, but with restrictions on data accessibility.
- There is a desperate need to do more with less.





# Internet Bandwidth (Digital Divide)



- Non-existent in some places and poor everywhere else.
- Preservation projects designed for high bandwidth are not suitable.
- All online solutions must be bandwidth-friendly.



# Is Africa Special?

- Definitely NOT!
- The same problems are faced by some other communities.
- Many communities face some of the problems.
- Most communities can benefit from solutions to these problems.



### Solutions: Lightweight and Reusable Systems

- Simplicity
- XML
- Minimalist Archives.
- Metadata management using office suite.
- Multi-purpose software tools (repositories).
- Shared skills in common tools, e.g., DSpace



### Solutions: Bandwidth

- Collections accessible over CD/DVD-ROM, local drives, network, etc.
  - Preservation by copying.
- Static collections rather than dynamic.
  - Preserve files instead of services.

Minimal bandwidth use e.g., using AJAX





# Solutions: Experience Recreation

Storytelling in virtual environments.

Low-cost virtual environments.

Virtual recreation of historical districts.



### Solutions: Basic Digitization

- Scan documents.
- Take photographs.
- Take 3D laser scans.
- Record audio.
- Build digital libraries / archives to preserve.
- Share and reuse information.





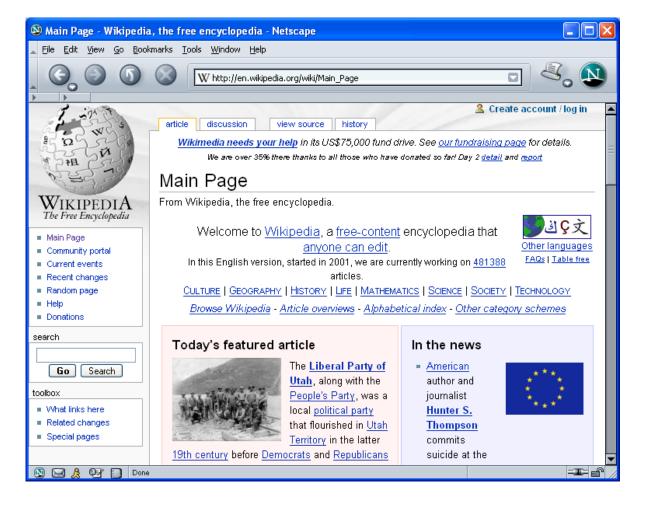
### What is a Digital Library: Example 1/3







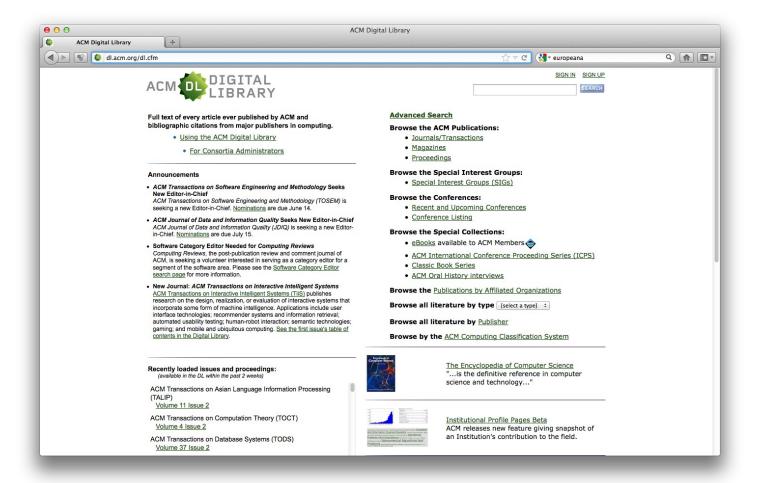
### What is a Digital Library: Example 2/3







### What is a Digital Library: Example 3/3







# Typical DL Services

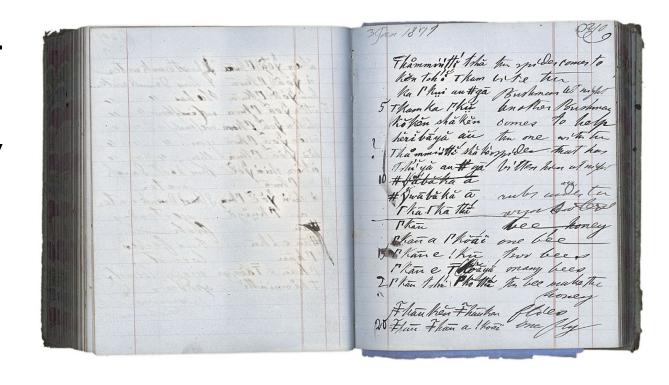
- User Management: accounts, auth, profile
- Searching: info retrieval, Google, indexing
- Browsing: categories, classification, subsets
- Submission: explicit/harvested/crawled
- Review: quality, workflow
- Annotation: reviews, ratings, discussions
- Recommendation: suggestions, collab filtering





### Case Study: Bleek and Lloyd Collection 1/2

- Books and drawings documenting nowextinct culture of [xam and !kun (Khoi-San?) groups.
- Documented by Wilhelm Bleek, Lucy Lloyd and others in late 1800s in Cape Town.
- ~20000 page images

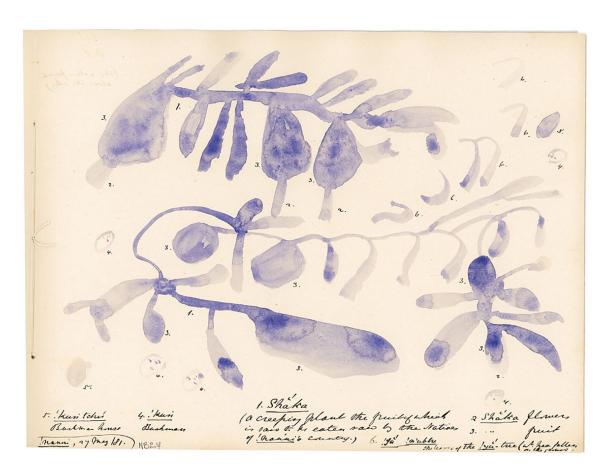






### Case Study: Bleek and Lloyd Collection 2/2

- □ ~800 drawings
- On UNESCO Memory of the World register.
- Curated by UCT/NLSA/Iziko-SAM/UNISA/...
- Digital preservation funded by Mellon, led by Michaelis School of Fine Arts, UCT







# Bleek and Lloyd Core Requirements

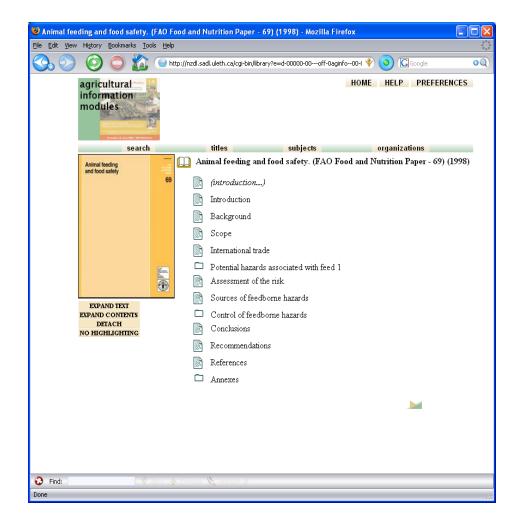
- Make the collection accessible as widely as possible:
  - Over the Web,
  - Off a CDROM,
  - Off a network-shared drive,
  - Etc.
- Platform independence (Mac/XP/Linux/etc.).
- Low barrier to use.
- Standards-compliance.





# Option 1: Greenstone

- Greenstone, a digital library tool, creates standalone CDROM collections.
- It still requires software installation.
- It does not work on ALL platforms.

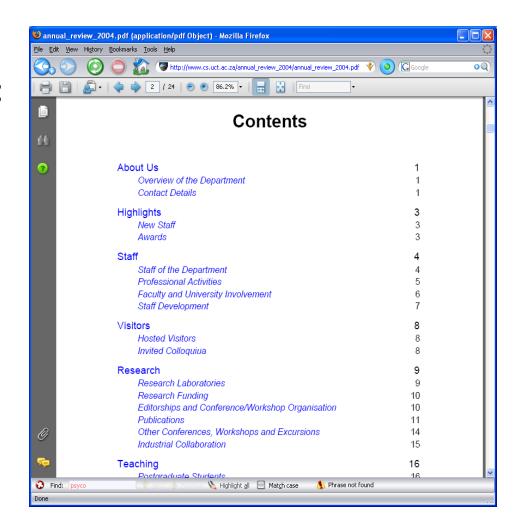






# Option 2: XSL-FO

- XSL-FO can be used to create hyper-linked static PDFs, like books.
- Does not work for large books.
  - PDF file sizes increase dramatically...







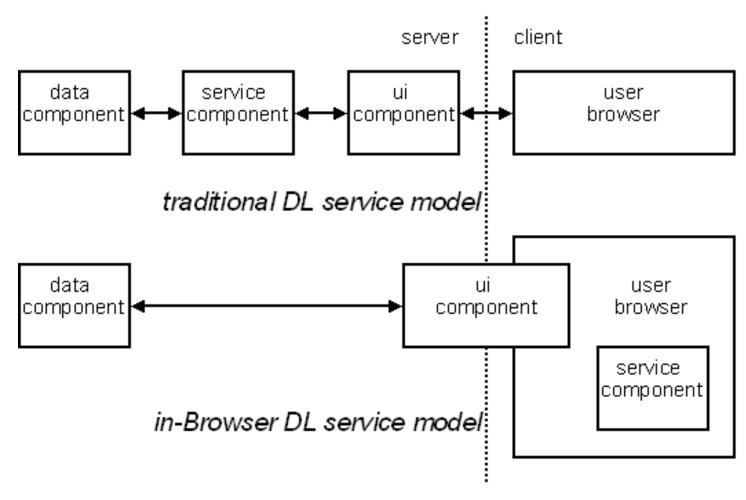
### Solution 1: XML + XSLT → XHTML

- Encode all descriptive information using XML.
- Write XSL transformations to convert the XML into multiple formats, each corresponding to an HTML page view.
- Needs advanced XSLT techniques to deal with size of data.





### Solution 2: in-Browser Services







This digital publication is part of a Llarec project to digitise, research and publish the Bleek and Lloyd Archive. Llarec (the Lucy Lloyd Archive, Resource and Exhibition Centre) is a University of Cape Town research centre located at the Michaelis School of Fine Art.

The project has been made possible by funding provided by the Andrew W. Mellon Foundation and De Beers; and is the result of the cooperation of the three curating institutions: University of Cape Town, Iziko South African Museum and The National Library of South Africa.

Scanning and retouching was done by Scan Shop, Cape Town.

DVD developed by Hussein Suleman,

Department of Computer Science, University of Cape Town.

Publication: Llarec Michaelis School of Fine Art University of Cape Town 31-37 Orange Street 8001 Cape Town http://michaelis.uct.ac.za/

#### © Pippa Skotnes 2007

The copyright for this publication belongs to Pippa Skotnes. Copyright for individual images belongs to the home institution. Scan numbers reflect the home institutions.

All rights reserved.

No part of this DVD may be reproduced or utilised in any form and by any means, without permission in writing from the publisher.



#### CLAIM TO THE COUNTRY



#### THE DIGITAL BLEEK AND LLOYD







#### THE DIGITAL BLEEK AND LLOYD

The Digital Bleek and Lloyd includes scans of every page of the 114 Lucy Lloyd [xam notebooks, 13 Lloyd (mostly)] kun notebooks and 28 Wilhelm Bleek [xam notebooks. It also includes Jemima Bleek's solitary Korana and !kun notebook (but excludes three Korana notebooks in the collection of the Maingard Library at the University of South Africa) and all the drawings and watercolours made by [han=kass'o, Dia!kwain, Tamme, [uma, !nanni and Da. This represents the coming together of an archive dispersed across three institutions: the National Library, Iziko South African Museum and the University of Cape Town. With the scans is a 280 000-word searchable index, cross-referenced and including notes and summaries for each of the stories listed. Notes in italics are direct quotes from the reports of Bleek and Lloyd in which they detail the progress of their research.

KEYWORD SEARCH



Books

Cover to cover | Contributor | Story | Category | Keyword



Drawings/Watercolours

Contributor | Category | Keyword





### THE DIGITAL BLEEK & LLOYD

#### HOME

#### Notebooks

Cover to cover Stories Contributors

> Categories Keywords

#### Drawings/Watercolours

Contributors Categories Keywords

SEARCH

#### NOTEBOOKS

Cover to cover

#### Wilhelm Bleek notebooks

- 1. BC 151 A1 4 001
- 2. BC 151 A1 4 002
- 3. BC 151 A1 4 003
- 4. BC 151 A1 4 004
- 5. BC 151 A1 4 005
- 6. BC 151 A1 4 006
- 7. BC 151 A1 4 007
- 8. BC 151 A1 4 008
- 9. BC 151 A1 4 009
- 10. BC 151 A1 4 010
- 11. BC 151 A1 4 011
- 12. BC 151 A1 4 012
- 13. BC 151 A1 4 013
- 14. BC 151 A1 4 014
- 15. BC 151 A1 4 015
- 16. BC 151 A1 4 016
- 17. BC 151 A1 4 017
- 18. BC 151 A1 4 018
- 19. BC 151 A1 4 019





#### THE DIGITAL BLEEK & LLOYD

#### HOME

#### Notebooks

Cover to cover
Stories
Contributors
Categories
Keywords

#### Drawings/Watercolours

Contributors Categories Keywords

SEARCH

#### **NOTEBOOKS**

Book: BC 151 Al 4 002

Collection: Wilhelm Bleek notebooks

Page Range: Al 4 2 00334.JPG - Al 4 2 00428.JPG



Image File: A1\_4\_2\_FRCOV.JPG

Story: Covers and first pages of Bleek's Book II or



Image File:
A1\_4\_2\_IFCOV.JPG

Story: Covers and first pages of Bleek's Book II or BC151\_A1\_4\_002



Image File: A1\_4\_2\_00334.JPG

Story: A man falls upon the Lion



Image File: A1\_4\_2\_00335.JPG

Story: A man falls upon the Lion



Image File: A1\_4\_2\_00336.JPG

Story: Jacob Nein and the leopard



Image File: A1\_4\_2\_00337.JPG

Story: Jacob Nein and the leopard



Image File: A1\_4\_2\_00338.JPG

Story: Jacob Nein and the leopard



Image File: A1 4 2 00339.JPG

Story: Jacob Nein and the leopard





#### THE DIGITAL BLEEK & LLOYD

#### HOME

#### Notebooks

Cover to cover
Stories
Contributors
Categories
Keywords

#### Drawings/Watercolours

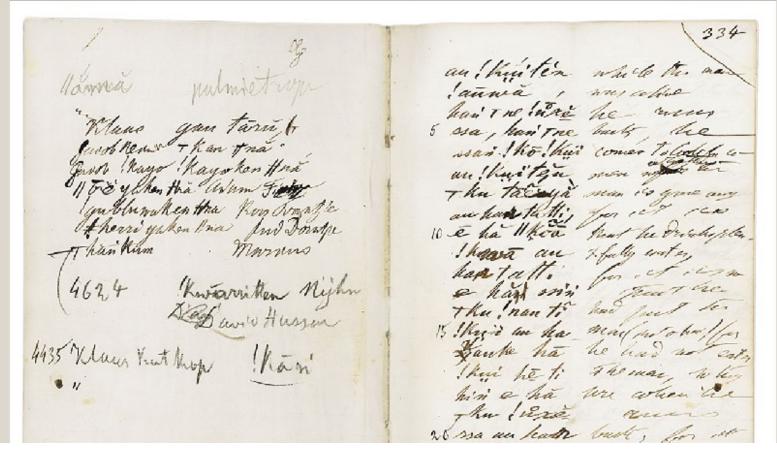
Contributors Categories Keywords

SEARCH

#### **NOTEBOOKS**

Page Image: A1\_4\_2\_00334.JP0

Previous Book Page	Book: BC 151 A1 4 002	Next Book Page
Previous Story Page	Story: A man falls upon the Lion	Next Story Page







HOME

#### Notebooks

Cover to cover Stories

Contributors

Categories

Keywords

#### Drawings/Watercolours

Contributors

Categories

Keywords

SEARCH

#### **NOTEBOOKS**

Stories

#### Wilhelm Bleek notebooks

- 1. Covers and first pages of Bleek's Book I or BC151 A1 4 001
- 2. Words and sentences (given by Adam Kleinhardt)
- 3. Word lists, including 'Bushman/Hottentot' vocabulary
- 4. Words and sentences (given by |a!kunta)
- 5. A man falls upon the Lion
- 6. Woman transformed into a lion
- 7. Baboons take a child
- 8. Veldschoen
- 9. The spider
- 10. Words and sentences (dictionary and genealogies)
- 11. The Moon pierced by the Sun
- 12. Words and sentences: names of animals
- 13. Leopards, lions and phrases
- 14. Moon and stars: an incantation
- 15. The man carries away the ostrich to his home
- 16. The eating of jackal
- 17. The Lion becomes a star
- 18. The piercing of ears
- 19. The xam and the Dutch
- 20. About 'seacows'
- 21. The |xam and the hyena





#### HOME

#### Notebooks

Cover to cover
Stories
Contributors
Categories
Keywords

#### Drawings/Watercolours

Contributors Categories Keywords

SEARCH

#### [a!kunta (Stoffel) (I)

#### Date

20 May 1871 (it is unclear whether this story was told on this date as there are others on the same page)

#### Categories

History (Early Race)

#### Keywords

Lion (a man falls on), Lion (chases a man up a tree), Lion (and a man), man (falls on a Lion), man (hides in a springbok skin sack), hides him), man (falls asleep in a tree), man (dreams of falling), springbok skin (a man hides from a Lion under), skin (of springbok hides from a Lion under), tree (man falls asleep in and falls on a Lion), Early Race (a story of the), Early Race (a man falls upon the

#### Story Pages

158-159, 334-335

#### Page Images



Image File: A1\_4\_1\_00158.JPG

Book: BC\_151\_A1\_4\_001



Image File: A1\_4\_1\_00159.JPG

Book: BC\_151\_A1\_4\_001



Image File: A1\_4\_2\_00334.JPG

Book: BC\_151\_A1\_4\_002



Image File: A1\_4\_2\_00335.JPG

Book: BC\_151\_A1\_4\_002





#### HOME

#### Notebooks

Cover to cover

Stories

Contributors

Categories

Keywords

#### Drawings/Watercolours

Contributors Categories

Keywords

SEARCH

#### **NOTEBOOKS**

#### Contributors

#### !khannumup (Petros Willems)

- 1. !khannumup (or Petros Willems): his personal history
- 2. !nauxa (or Willem) at the Museum, 24 September 1880
- 3. Words and sentences: at the Museum, 24 September 1880 (!nauxa and !khannumup)

#### !kweiten ta |ken (Rachel) (VI)

- 1. A lion's story, or, The child who saved her sleeping parents from the lion
- 2. About maidens and how they adorn young men with ||ka or 'rooi klip'
- 3. Names of !kweiten ta |ken's relations
- 4. Story of kua ka khumm
- 5. The Anteater's story, or, The Anteater, Springbok, Lynx and Partridge
- 6. The Crow's story: the Crows are sent out to search for husbands, or, The !kagen ka Kkomm's story and the |hunn ta kkomm's story (including What happened when the !kagen found the missing men, p.3995)
- 7. The Lion's story
- 8. The Quagga's story
- 9. The Rain's story, and |kannu the waterhole
- 10. The Water's story: more about how maidens adom young men
- 11. The maiden's story; the frog's story
- 12. The man who went to sleep when out hunting alone
- 13. The story of the Leopard Tortoise
- 14. The story of the Springbok's kid, who was carried off by the Elephants (an incomplete account)
- 15. What is done with a 'new' maiden
- 16. What the man did to his wife when she was pregnant
- 17. Words and sentences: by !kweiten ta |ken
- 18. Words and sentences: by !kweiten ta |ken
- 19. Words and sentences; parts of the body
- 20. Xumi ko killed by a lion
- 21. a khumm called by a lion

Inanni (VT)

http://banzai.cs.uct.ac.za/~hussein/bleek\_lloyd/data/stories/5/index.html





#### HOME

#### Notebooks

Cover to cover

Stories

Contributors

Categories

Keywords

#### Drawings/Watercolours

Contributors

Categories

Keywords

SEARCH

#### **NOTEBOOKS**

Categories

#### Artefact and dress

!kuppen (imitation of the sound of horses)

!nanni's father's dress

A drum

A necklace of reed used to cure a little child's cold

Arrow-making

Arrowheads

Cutting with stone knives

Digging sticks used by men

Elephants (eating their hearts)

Feathered arrows and poison

Grass in Bushmanland that resembles brushes

Heron's song (The song of the Blue Crane)

Horns burnt for rain

How an Elephant steals a young Springbok from [kaggen (the Mantis) or Pet Springbok carried off by an Elephant

How the prepared springbok's ears are tied on

Huts

Jewellery and ornaments worn by women

Karosses must not be beaten upon the ground, for fear of bad consequences

Making fire with two pieces of sticks; and tinder-making

Mat sieves

Names and descriptions of bags, nets and things worn

Notes on chippings, or, Chipping no. 4

Notes on rock painting copy no. 2 (!nu'sa and other groups)

Notes on rock painting copy no. 5, no. 6, and no. 7

Notes on rock painting copy no. 8: the rain and the rain-bull

Pieces of wood for divining, or |xu





HOME

#### Notebooks

Cover to cover
Stories
Contributors
Categories
Keywords

#### Drawings/Watercolours

Contributors Categories Keywords

SEARCH

#### NOTEBOOKS

Keywords

```
!au
  is a utensil for eating 584
  shaped rib-bone 584
  the making of 584
  the working or shaping of 584
  used for eating gambro and kuiten 584
!gabbaken-!gabbaken
  his wife found fault with his appearance 676
  how he punished his wife for making personal remarks 676
  or Mason Wasp 676
  possessed weapons 676
  shot his wife and killed her with an arrow intended for the Hare 676
  was formerly a man 676
!gaunu
  a song sung by, especially by women 518
  and 'Aquilae's water' 690
  and flowers 517
  and stars 517
  and the Dawn's stars 505
  and the names of stars 517
  asks the ||garraken flower to open 517
  great Star 505
  his song 505
  his song to the Star 517
  is !guonni 517, 519
```





#### HOME

#### Notebooks

Cover to cover Stories

Contributors

Categories

Keywords

#### Drawings/Watercolours

Contributors

Categories

Keywords

SEARCH

#### **DRAWINGS**

#### Contributors

- 1. !nanni
- 2. !nanni (and Tamme)
- 3. <u>Da</u>
- 4. Dia!kwain
- 5. Tamme
- 6. han≠kass'o
- 7. uma
- 8. unattributed

drawing book covers





#### HOME

#### Notebooks

Cover to cover
Stories
Contributors
Categories
Keywords

#### Drawings/Watercolours

Contributors Categories Keywords

SEARCH

#### DRAWINGS

Contributor: Inann



Image File: Iziko\_SAMLB\_!nBook1\_001 Category: plants and animals



Image File: Iziko\_SAMLB\_!nBook1\_002 Category: plants and animals, custom and daily life



Image File: Iziko\_SAMLB\_!nBook1\_003 Category: plants and animals



Image File: Iziko\_SAML Category: pl:



Image File: Iziko\_SAMLB\_!nBook1\_005 Category: plants and animals



Image File: Iziko\_SAMLB\_!nBook1\_006 Category: plants and animals



Image File: Iziko\_SAMLB\_!nBook1\_007 Category: plants and animals



Image File: Iziko\_SAML Category: pl: relations with



#### HOME

#### Notebooks

Cover to cover Stories Contributors

Categories

Keywords

#### Drawings/Watercolours

Contributors Categories Keywords

SEARCH

#### **DRAWINGS**

Image: Iziko\_SAMLB\_!nBook1\_001

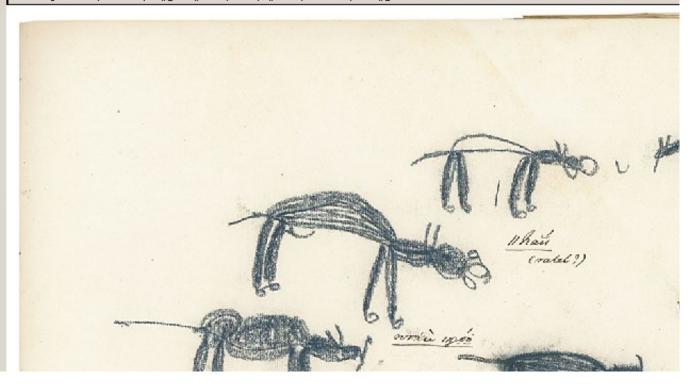
Contributor : !nanni

Medium: crayon/pastel/chalk (blue & some black)

Date: 16 May 1881 Size: 350x250

Category : plants and animals Keywords : animal (ratel)

Description : n|xu de, n|xu ||goo, ||hau (ratel?), ||hau, |kere de, |kere ||goo







HOME

#### Notebooks

Cover to cover Stories Contributors Categories

Keywords

#### Drawings/Watercolours

Contributors Categories Keywords

SEARCH

#### DRAWINGS

- animal
- 2. artefact and dress
- 3. celestial bodies and aeroscopy
- 4. custom and daily life
- 5. healing and ailing
- history (personal)
- personal history
- plants and animals
- 9. relations with others
- 10. relations with plants and animals
- 11. the Rain and Rain's liquid
- 12. transformation



HOME

#### Notebooks

Cover to cover Stories Contributors Categories Keywords

#### Drawings/Watercolours

Contributors Categories Keywords

SEARCH

#### **DRAWINGS**

Keywords

#### !kun

'another kind of Bushman' NLLB D Ta058 called the 'Sun Bushmen' NLLB D !n006 fruit eaten by Iziko SAMLB TaBook3 010 the |na NLLB D Ta058

#### !nanni

and the country of the Biriko !kun <u>Iziko SAMLB D !n003</u> his grandfather Karu's place and things <u>Iziko SAMLB D !n005</u> plant drawn out of his head <u>Iziko SAMLB !nBook4 006</u> tree in his father's country NLLB D !n039

#### !nanni's country

food eaten in NLLB D !n031

#### adomment

Makoba's NLLB D !n064

necklace NLLB D !n064

thong, roots worn by men around neck NLLB D Ta054

animal Iziko SAMLB InBook4 002, NLLB D Ta069, NLLB W Ta005, Iziko SAMLI

!khau lizard or Kochelman <u>UCTLB\_D\_Di003</u>, <u>UCTLB\_D\_Di004</u>

"Nettson" Iziko SAMLB D!h001

'mere thing' Iziko SAMLB !nBook2 005

?NLLB D A006, NLLB D A027

a bird eaten by the !kun, and one eaten by the Makoba Iziko SAMLB TaBook3 00

a bird, eaten by the !kun, which utters its own name NLLB W u021

a large bird in uma's country Iziko SAMLB ubookl 003

and its young Iziko SAMLB TaBook1 003

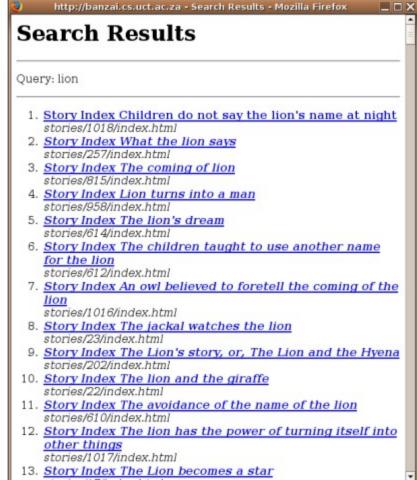
and Ovaherero NLLB D Ta023

and Ovambo NLLB D. In111











# Principles of DL for African Heritage

- Efficient bandwidth use
- Advanced technology
- Appropriate technology
- Local relevance
- Modernization instead of Africanization
- Global applicability of solutions
- Minimalism of staff/money
- Multicultural/multilingual inclusivity





## More Case Studies in Heritage DLs

recent research at UCT

- alternative archive infrastructures
- cloud computing
- multilingual IR
- heritage preservation
- visual dictionaries
- rock-art exploration using mobile device
- mobile | Xam input





## Cloud-based Archives



"It was much nicer before people started storing all their personal information in the cloud."

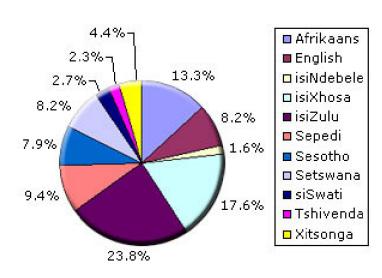
Lebeko Poulo, Lesotho Mushashu Lumpa, Zambia

- individual services and whole archives in private clouds
- install locally
- reduces need for skilled staff
- instant archives
- shared resources
- automatic scalability
- acceptable performance, after cache priming
- user studies in progress





# Multilingual Information Retrieval



(source: SA Info)

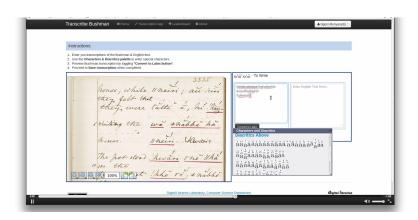
Mohammed Mustafa Ali, Sudan

- search queries with multiple languages
- current systems biased to one language
- rerank documents by understanding query and reweighting languages/results
- better quality results found, higher up in results





# Document Transcription: Bleek and Lloyd Stories



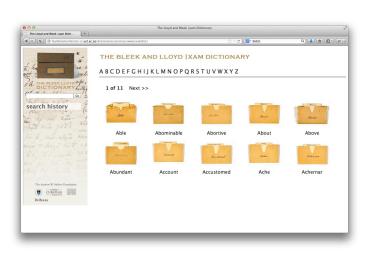
Ngoni Munyaradzi, Zimbabwe

- crowdsourced transcription application
- volunteers to convert images to text
- automated algorithms to check and assess quality
- interactive Web interface for users to enter text
- 10% better than AI approaches!





# Language Preservation: Online | Xam dictionary



visual dictionary of |Xam language

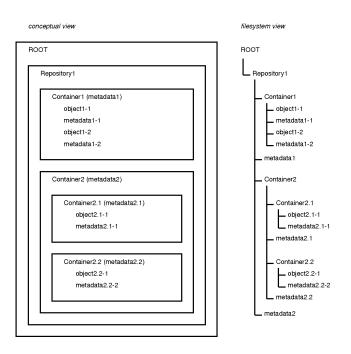
- simple archive foundation
- client-side processing as far as possible
- linked into Bleek and Lloyd

Kyle Williams, South Africa





# simplyCT



Phiri Lighton, Zambia

simple archive architecture

- performance
- understandability
- flexibility
- applicability
- good performance for small to medium collections
- easy to use and expand





## What We Have Learnt

Digital preservation in Africa has special problems.

But all problems can be addressed adequately with appropriate and innovative use of current technology.





# Future Challenges

- Scalability of preservation efforts.
  - How to create similar collections easily?
  - A national heritage archive?

- Tools for management and dissemination?
  - Extend Greenstone?





## Future Challenges

- Standard tools to manage metadata/data?
  - Usability
  - Scalability
  - Extensibility
  - System independence
- Many current repository tools better suited to papers and not heritage collections.



## The Future – Past Preserved







# That's all Folks!

direct all questions and comments to:
 hussein@cs.uct.ac.za
 Facebook/slumou
 Twitter@slumou