

Institutional Repositories: *Why they are important for African Research*

*Hussein Suleman
hussein@cs.uct.ac.za*

*University of Cape Town
Department of Computer Science*

November 2007



sivulile



open access
south africa


WHAT



What Exactly is an Institutional Repository?

UCT CS Research Document Archive [Home](#) || [About](#) || [Browse](#) || [Search](#) || [Register](#) || [User Area](#) || [Help](#)

Welcome to UCT CS Research Document Archive



Welcome to the [UCT Computer Science](#) Research Document Archive, which archives and makes accessible documents that are products and by-products of research in the department.

Search the Titles, Authors, Abstracts and Keywords :

[Browse](#)

Browse the archive by [Subject](#), [Year](#), [Lab](#) or [Type](#).

[Latest Additions](#)

View items added to the archive in the past week.

[Simple Search](#)

Search the archive using the most common fields.

- ☑ Collection of research-related documents
- ☑ Author self-submission
- ☑ Checking of submissions
- ☑ Archive-everything!
- ☑ Hierarchical browsing
- ☑ Simple and fielded searching
- ☑ OAI-PMH compliance



WHY



because of Direct access from GOOGLE 1/3



Web

Results 1 - 10 of about 73 for **suleman fox madalli**. (0.59 seconds)

[UCT CS Research Document Archive - Digital Libraries](#)

Digital Libraries. **Fox**, Edward A., Hussein **Suleman**, Devika **Madalli** and Lillian Cassel (2003) Digital Libraries, in Practical Handbook of Internet Computing. ...
pubs.cs.uct.ac.za/archive/00000016/ - 6k - [Cached](#) - [Similar pages](#)

[UCT CS Research Document Archive - Design and Implementation of ...](#)

Suleman, Hussein, Edward A. **Fox** and Devika **Madalli** (2003) Design and Implementation of Networked Digital Libraries: Best Practices. ...
pubs.cs.uct.ac.za/archive/00000017/ - 6k - [Cached](#) - [Similar pages](#)
[[More results from pubs.cs.uct.ac.za](#)]

[Documentation Research and Training Centre: Items for Author ...](#)

... 2003, Design and Implementation of Networked Digital Libraries: Best Practices, Hussein, **Suleman**; **Fox**, Edward A.; **Madalli**, Devika P. ...
<https://drtc.isibang.ac.in/items-by-author?author=Madalli%2C+Devika+P>. - 10k - [Cached](#) - [Similar pages](#)

[Documentation Research and Training Centre: Items for Author Fox ...](#)

... Date of Issue, Title, Authors. 2003, Design and Implementation of Networked Digital Libraries: Best Practices, Hussein, **Suleman**; **Fox**, Edward A.; **Madalli**, Devika



because of Direct access from GOOGLE 2/3

UCT CS Research Document Archive

[Home](#) || [About](#) || [Browse](#) || [Search](#) || [Register](#) || [User Area](#) || [Help](#)

Digital Libraries

Fox, Edward A., Hussein Suleman, Devika Madalli and Lillian Cassel (2003) *Digital Libraries*, in *Practical Handbook of Internet Computing*. CRC Press.

Full text available as:

[PDF](#) - Requires [Adobe Acrobat Reader](#) or other PDF viewer.

EPrint Type: Book Chapter

[H Information Systems: H.0 GENERAL](#)

[H Information Systems: H.1 MODELS AND PRINCIPLES](#)

Subjects: [H Information Systems: H.3 INFORMATION STORAGE AND RETRIEVAL](#)

[A General Literature: A.1 INTRODUCTORY AND SURVEY](#)

[H Information Systems: H.4 INFORMATION SYSTEMS APPLICATIONS](#)

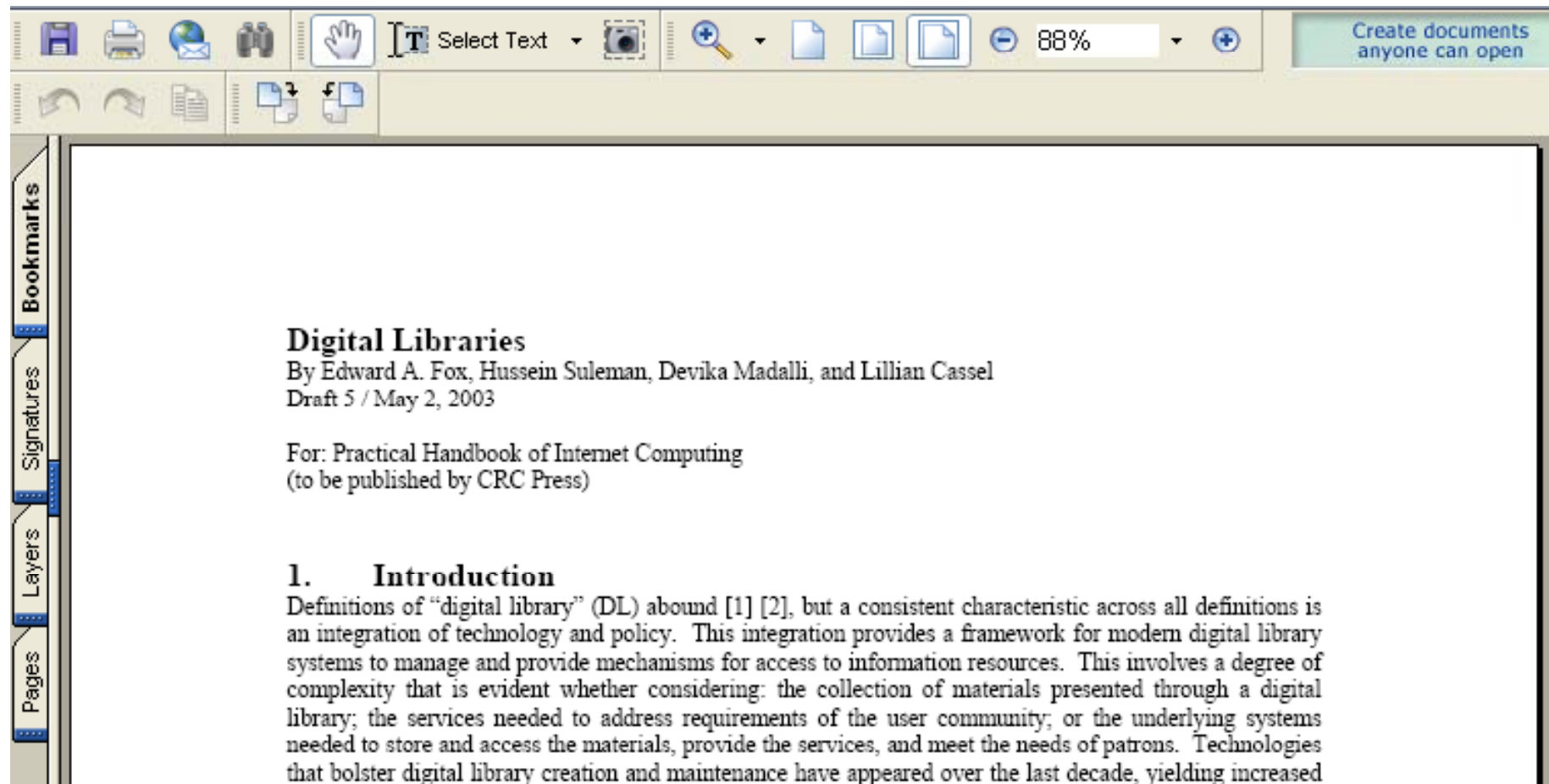
ID Code: 16

Deposited By: [Suleman, Hussein](#)

Deposited On: 11 July 2003



because of Direct access from GOOGLE 3/3



The screenshot shows a Google Docs interface. The top toolbar includes icons for saving, printing, sharing, and zooming. The document content is as follows:

Digital Libraries
By Edward A. Fox, Hussein Suleman, Devika Madalli, and Lillian Cassel
Draft 5 / May 2, 2003

For: Practical Handbook of Internet Computing
(to be published by CRC Press)

1. Introduction
Definitions of “digital library” (DL) abound [1] [2], but a consistent characteristic across all definitions is an integration of technology and policy. This integration provides a framework for modern digital library systems to manage and provide mechanisms for access to information resources. This involves a degree of complexity that is evident whether considering: the collection of materials presented through a digital library; the services needed to address requirements of the user community; or the underlying systems needed to store and access the materials, provide the services, and meet the needs of patrons. Technologies that bolster digital library creation and maintenance have appeared over the last decade, yielding increased



because of the No of ACCESSES from Google

- 1734 hits directly from Google in March 2007.
- Example:
 - <http://www.google.com/search?q=questionnaire+system+UML>
 - Kritzinger, Pieter, Marshini Chetty, Jesse Landman, Michael Marconi and Oksana Ryndina (2003) ChattaBox: A Case Study in Using UML and SDL for Engineering Concurrent Communicating Software Systems. In Proceedings Southern African Telecommunications Networks and Applications Conference, George, South Africa.



because academics can manage a research OVERVIEW

2004

Conference Paper

1. [Arnab, Alapan and Andrew CM Hutchison \(2004\) Digital Rights Management - An Overview of current challenges and solutions. In Venter, HS, JHP Eloff, L Labuschagne and MM Eloff, Eds. *Proceedings Information Security South Africa \(ISSA\)*, Gallagher Estate, Midrand, South Africa.](#)

Conference Poster

1. [Landman, Jesse and Pieter Kritzinger \(2004\) DS-CDMA Fading Channels with Bursty IP Traffic Arrivals. In *Proceedings WiOpt 2004*, Cambridge, England.](#)

Departmental Technical Report

1. [Arnab, Alapan and Andrew Hutchison \(2004\) Digital Rights Management - A current review. Technical Report CS04-04-00, Department of Computer Science, University of Cape Town.](#)
2. [de Wet, Nico Dirk and Pieter S. Kritzinger \(2004\) Model-Based EIS Performability Analysis. Technical Report CS04-01-00, Department of Computer Science, University of Cape Town.](#)
3. [Perumal, Sameshan and Pieter Kritzinger \(2004\) A Tutorial on RAID Storage Systems. Technical Report CS04-05-00, Department of Computer Science, University of Cape Town.](#)



because WEB pages are not sufficient

□ Problems:

- Persistence – will the documents always be there?
- Authority – can we trust the authenticity of the website?
- Standards – are the formats and metadata the same across all websites?
- Discovery – how do we find these documents?
Google?



because of REPORTING

- ❑ You can create a departmental or university listing of all research output.
- ❑ Where copyright does not allow, include just a citation – no files – for completeness.

NDLTD Union Catalog Project

Suleman, Hussein (2004) *NDLTD Union Catalog Project*, in Fox, Edward A., Shahrooz Feizabadi, Joseph M. Moxley and Christian R. Weisser, Eds. *Electronic Theses and Dissertations: A Sourcebook for Educators, Students, and Librarians*, chapter 6, pages 73-77. Marcel Dekker, Inc..

Full text available as:

EPrint Type: Book Chapter

Subjects: [H Information Systems: H.4 INFORMATION SYSTEMS APPLICATIONS](#)

ID Code: 14

Deposited By: [Suleman, Hussein](#)

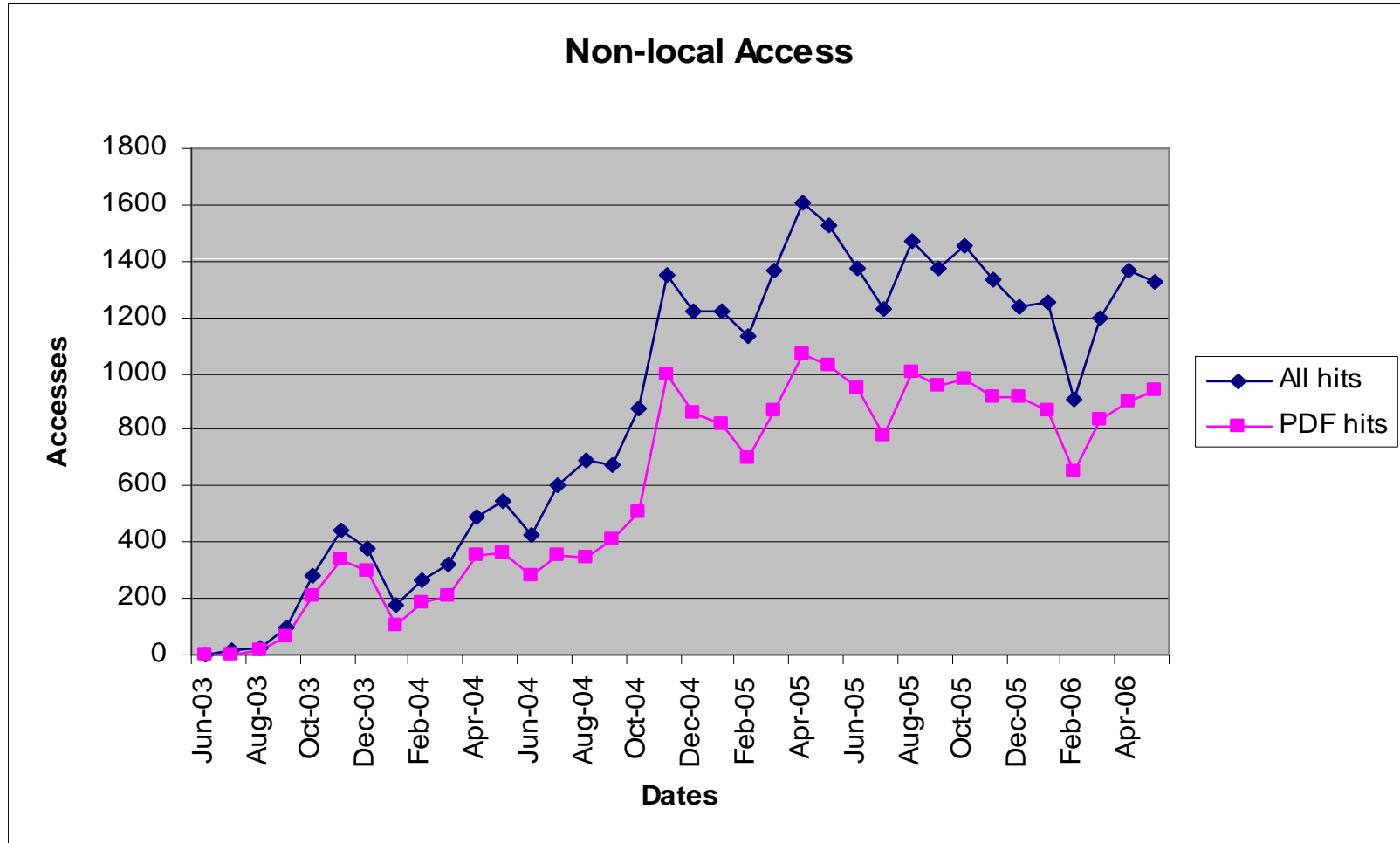


because of SHARING

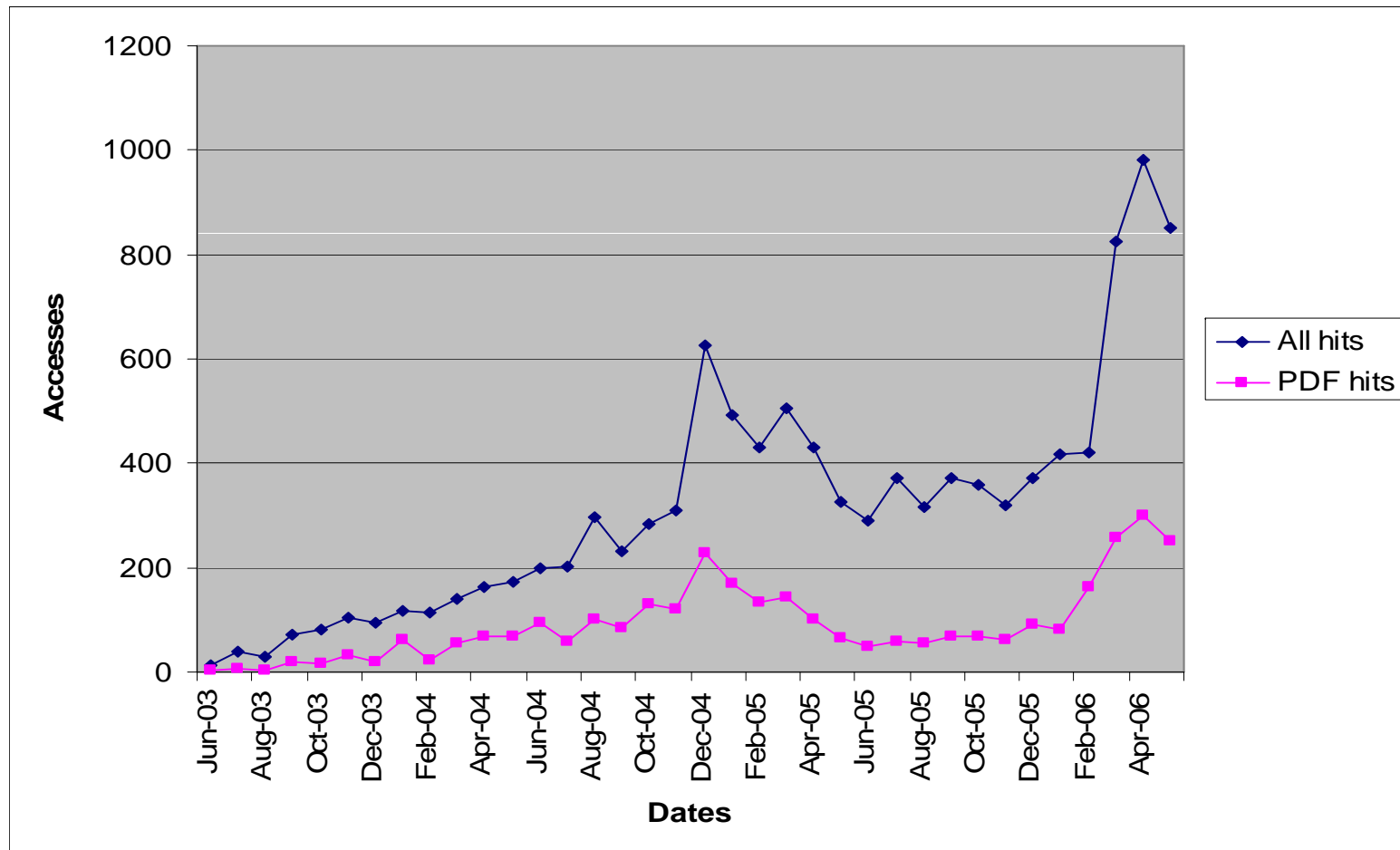
- ❑ Most Institutional Repositories are compliant with Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) v2.0.
- ❑ Metadata can be freely harvested by any service provider, e.g., search engines.
- ❑ Thus more people know what you have and can find it easily.



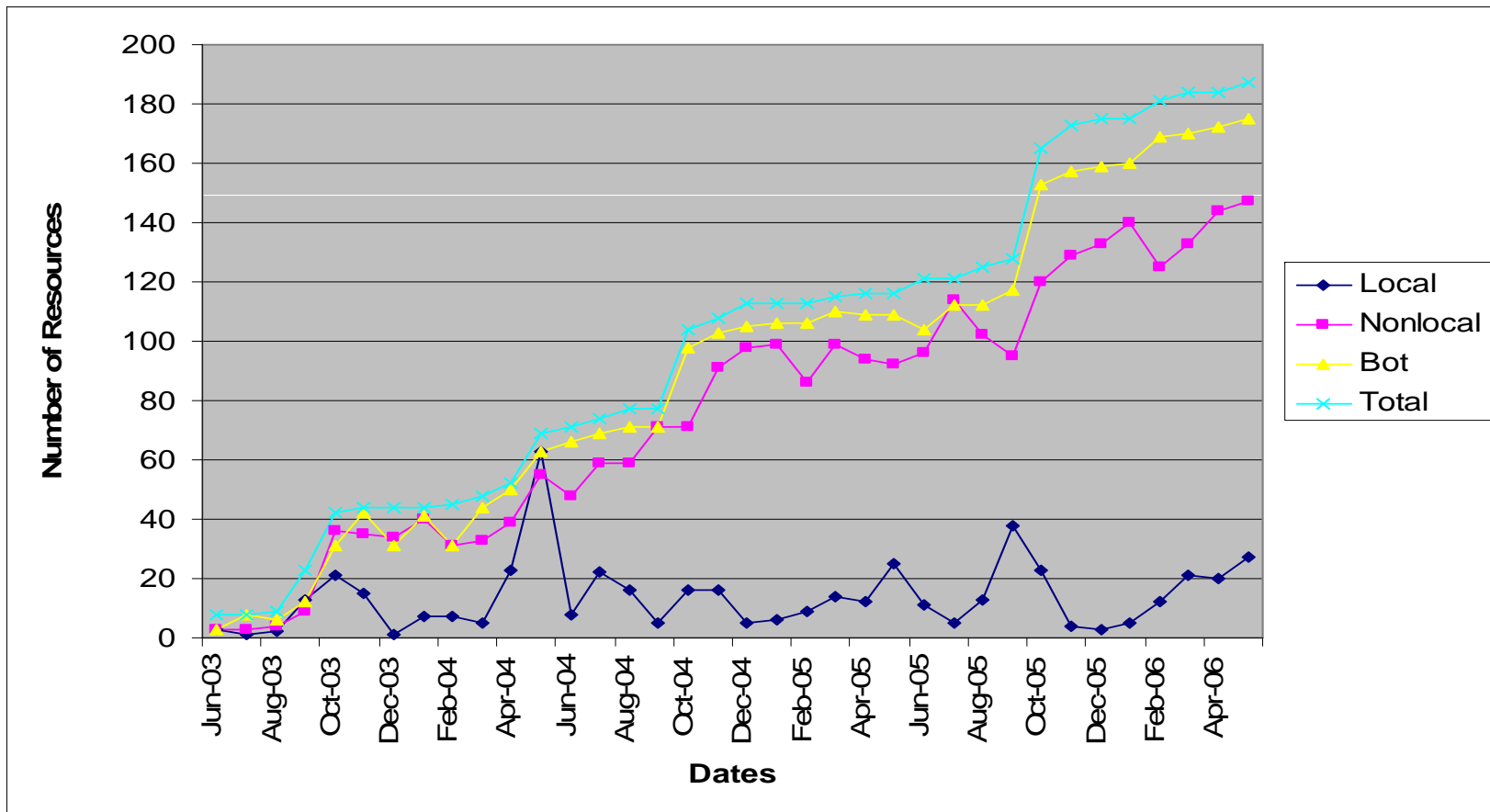
because of how many **USERS** we get



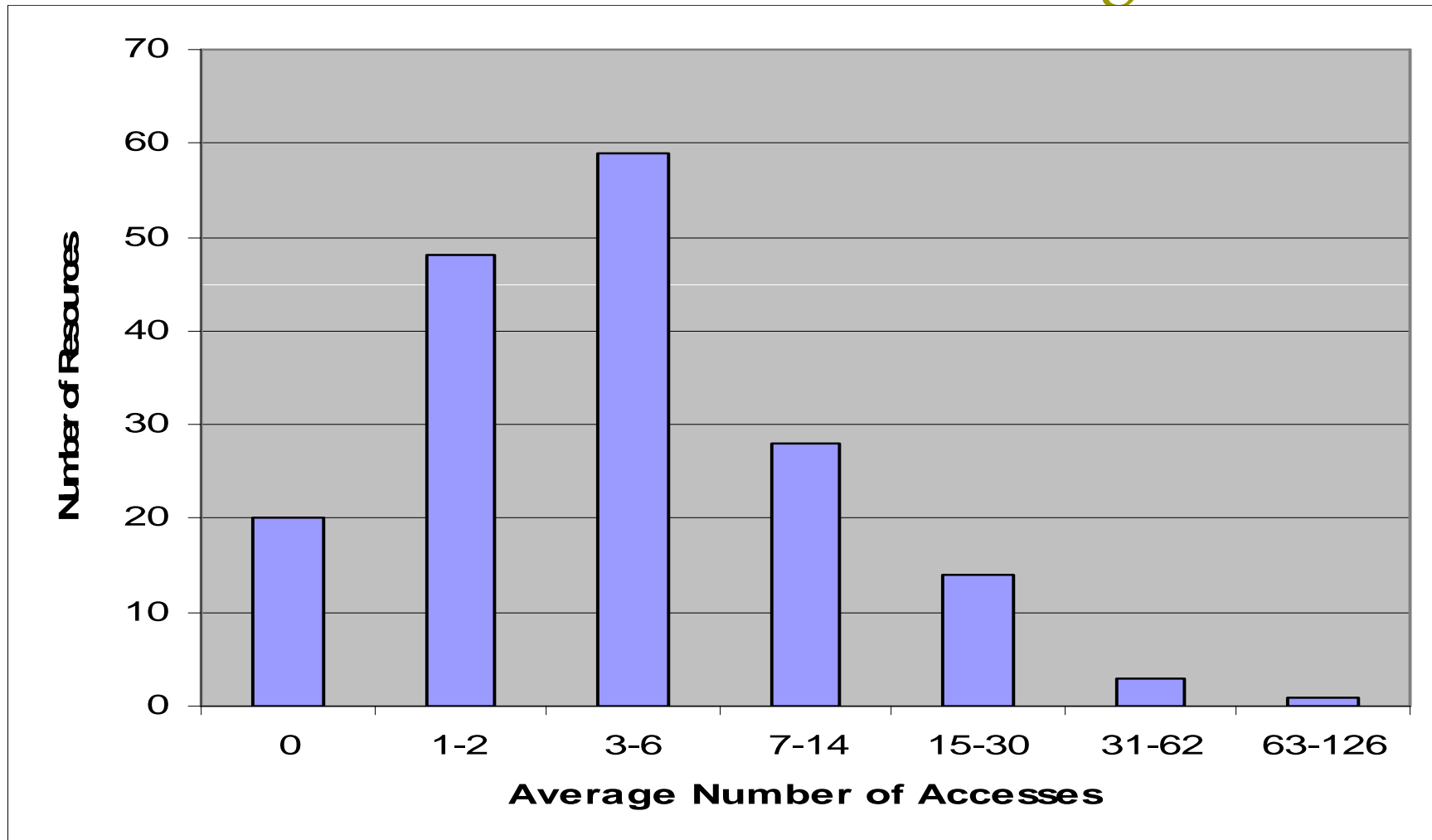
because of no. of SEARCH ENGINES we get



because of how many RESOURCES get viewed



because of how OFTEN resources get viewed



because SUBJECT repositories failed

- ❑ In the 90's subject archives were popular e.g., arXiv, RePEc, NCSTRL.
- ❑ Problems:
 - Sustainability – repositories run by organisations with limited funding.
 - Development skill – staff needed to develop and maintain all software needed to run the archive.
 - Should the archive be centralised or distributed?
 - Owner of data was not best provider of services.



because COPYRIGHT is changing in our favour

- ❑ Most society publishers will allow archiving on a website or IR e.g., ACM
- ❑ Most commercial publishers allow archiving on a website or IR after some time (typically 12-24 months), if not immediately.
- ❑ Newer commercial publisher agreements make greater allowance for IRs.
- ❑ You can always negotiate with a publisher!



because 400000 ETDs are already out there

OCLC ONLINE COMPUTER LIBRARY CENTER

A Project of OCLC Research

XTCat NDLTD
NDLTD Union Catalog



[Identify](#) | [GetRecord](#) | [ListIdentifiers \(Resumption\)](#) | [ListMetadataFormats](#) | [ListRecords \(Resumption\)](#) | [ListSets](#)

```
responseDate    2007-11-04T13:20:01Z
request         http://alcmec.oclc.org/ndltd/servlet/OAIHandler?verb=ListSets
```

ListSets

setSpec	setName	DC record count
ADTP	Australasian Digital Theses Program	30853
AUCKLAND	University of Auckland	375
BGMU	Brigham Young University Theses	910
BICBF	Bibliothèque interuniversitaire de la Communauté française de Belgique	880
CALTECH	California Institute of Technology	3776
CCSD	CCSD theses-EN-ligne, France	7936
CRANFIELD	Cranfield University	109



HOW



Set up Infrastructure

- ❑ Task for: ICT / Libraries.
- ❑ Acquire software (EPrints, DSpace, etc.).
- ❑ Test installation.
- ❑ Acquire hardware.
- ❑ Install and configure IR.
- ❑ Integrate into campus websites.
- ❑ Perform ongoing maintenance.



Convince Academics

- ❑ Leadership from senior management.
- ❑ Arm-twisting + Ego.
- ❑ Rewards.
- ❑ Payments for submission.
- ❑ Generate reports from IR.
- ❑ Link IR prominently.
- ❑ Work with younger, more dynamic, staff.



Discussion...



*to find me, search on Google or
Facebook for "hussein suleman"*