

Access to Individual Harvested Sites in a Web Archive



Tracy Meehl
DLF Fall Forum, Providence, RI
November 13th, 2008

Library of Congress Web Archives

EVENT-DRIVEN

- September 11th, 2001
- Winter Olympic Games 2002
- U.S. Congresses 107th, 108th, 109th, etc.
- U.S. Elections 2000, 2002, 2004, 2006, 2008, etc.
- Iraq War 2003-
- Papal Transition 2005
- Supreme Court Nominations 2005-2006
- Crisis in Darfur, Sudan 2006
- Egypt 2008

FORMAT/COLLECTION-DRIVEN

- Organizational Sites corresponding to Papers/Archives collected by LC's Manuscript Division
- Sites corresponding to creators whose works are collected by/represented in LC's P&P Division
- Legal Blawgs identified by the Law Division

Iraq War, 2003 Web Archive

The New York Times of the bloggers.
-- Pravda

The Grand Central Station of
Blogerville! -- American Journalism
Review

"Glen also has a rather dry sense of
humour, which I'm guessing not
enough people pick up on." -- Andreas
See

[Make a Donation](#)

Navigation
[Instapundit for PDA](#)
[Instapundit Backup](#)
[Instapundit Extra!](#)
[Instapundit Print](#)
[PsyWar Update](#)
[Other Writings](#)
[Print Format](#)
[XHTML](#) | [PDF](#)
[About Me](#)
[Archives](#)

iraqpolicy.com

THE IRAQ POLICY INFORMATION PROJECT

[The issues](#) [About us](#) [Expert sources](#) [Links](#) [Contact us](#)

Get Iraq updates: [your email address]

The Potential Costs of War With Iraq ... Is the Administration playing "Russian Roulette" with the economy?

The White House suggests that war with Iraq will begin before the end of the month, but the Bush administration has stubbornly refused to provide a substantive estimate of what it might cost. In the words of Secretary of Defense Donald Rumsfeld, it is not "a very useful exercise" to prepare cost estimates in light of the many variables.

[Read Full Feature>>](#)

Consequences of War Could Be Disastrous for Fragile Economy

A comprehensive analysis by a group of experts convened by the Center for Strategic and International Studies (CSIS) predicts that a war in Iraq could send prices careening way up as much as

100 Celebrities



The Command Post



Now Open: TCP Forum

We've had readers requesting a discussion board for a while now, and decided to try it out and see how it goes. Now open: [The Command Post Forum](#) Enjoy!



Most Recent Pub. Desk Posts - [Go To The Publisher's Desk](#)

- > Blog Spotting - No Passard - 1
- > TriviaQuizze - All The President's Hair - 0
- > Site Matters - The New Tree 650 - 0
- > MSM - What CNN Can Learn From Fep Boys - 0
- > Blog Spotting - DeepBlog - 1
- > JETB - Campaigns & Elections Blogging Panel - 0
- > Go See - GO IGLESIA - 4



You are not reading this!

Command Post Poll

John Negroponte:

- Great choice!
- Good choice
- Bad choice
- None of the above



Updated 02 May 2003

WAR REPORT

Select articles, documents, and analyses compiled and frequently updated by the [Project on Defense Alternatives](#)

Iraq War

Iraq's Other Disarmament Challenge: Small Arms
Warren Richey, *Christian Science Monitor*, 02 May 2003.

Iraq Situation Critical, Warn Aid Agencies
The Guardian, 02 May 2003.

Quiet Power Struggle in Southern Iraq
Peter Slemon, *Washington Post*, 02 May 2003.

Text of President Bush's Speech
New York Times, 01 May 2003.

Critic Accuses Media of Aiding U.S. War Propaganda
David Morgan, *Reuters*, 01 May 2003.

The War is Over (Except for Iraq)
Phil Reeves, *The Independent*, 01 May 2003.

Villified Weapons Inspectors May Have Got It Right
Sydney Morning Herald, 01 May 2003.

Afghan Aftermath

Declared No Longer a Combat Zone, There's No Peace in Afghanistan

Todd Pitman, *Associated Press*, 02 May 2003. Posted on the Boston Globe website.

Rumsfeld Announces End of Afghan Combat

Vernon Loeb, *Washington Post*, 02 May 2003.

Afghanistan Offers Lesson on Rebuilding

Indra A.R. Lakshmanan, *Boston Globe*, 02 May 2003.

Security Remains Major Concern for Afghanistan

Mike Patterson, *Agence France-Presse*, 02 May 2003. Posted on the ReliefWeb website.

Afghanistan, Once More the Melting Pot

Syed Saleem Shahzad, *Asia Times*, 01 May 2003.

Rumsfeld Arrives in Afghanistan

BBC News, 01 May 2003.

INSTAPUNDIT.COM

"Nothing is more powerful than an individual acting out of his conscience, thus helping to bring the collective conscience to life." - Norman Cousins

NOTICE: This web site reflects the understanding that humanities chosen course of ecological destruction is the essence of ignorance, wholly unsustainable and akin to collective suicide. Worse yet, humanity is now seriously flirting with World War III and nuclear annihilation. To deny this is the height of stupidity and criminal negligence with regard to those who have produced children.

Humanity today has inherited a world where our very survival is at stake. As such the status quo is a virtual death sentence for all life as we know it. Our greatest responsibility is to fully acknowledge the natural rights of all life being interconnected, our universal kinship. In other words, "do unto others as we would have them do unto us." Because indeed our fate is inseparably linked to that of our victims, and/or our benefactors.

Those rejecting Universal Kinship will undoubtedly find my writing "extreme" and/or "offensive." I make no apologies whatsoever. Those in touch with reality will understand that my words are a sincere and passionate call for massive Direct Action & Civil Disobedience based on Truth Justice & Peace, and genuine love for all life.

Ken Nichols O'Keefe - Founder

Universal Kinship Society - Truth Justice Peace Human Shield Action



The Few. The Proud. [Recruiting](#) | [HOMC](#) | [Units](#) | [Career](#) | [Marine Online](#) | [Marine 4 Life](#) | [News](#) | [Family](#) | [Publications](#)

News From the Front

[Story List](#)

[Previous Page](#) | [Next Page](#)

- [Photo Gallery - Recent Deployed Imagery](#) Mar. 27, 2003
- [Storm doesn't slow Marines' advancement](#) March 26, 2003
- [Deployed Marines Remember Fallen Comrades](#) March 23, 2003
- [II MEF grants protect I MEF command](#) March 25, 2003
- [I MEF establishes forward command post in Iraq](#) March 22, 2003
- [Interacting with the locals \(Photo Essay\)](#) March 24, 2003
- [Rumaylah oil fields secured, preserved for Iraqi people](#) March 25, 2003
- [Wings of Freedom: 3rd MAW aircraft support Operation Iraqi Freedom](#) March 20, 2003
- [Photo Gallery - March 24th - Operation Iraqi Freedom](#) Mar. 24, 2003
- [Rodeo kicks up sand in Kuwait Desert](#) March 23, 2003

Crisis in Darfur, Sudan 2006 Web Archive



On May 3, 2006, a commission of Sudan activists and leaders in the movement to stop genocide and slavery in Sudan embarked on a week-long fact-finding and relief mission in Southern Sudan and Darfur. [Click here to see photos and find out what it was like for all involved...](#)

Sudan Freedom Walk a Success

Thanks to the efforts of myriad participants and volunteers, the 300-mile Sudan Freedom Walk from NYC to Washington, D.C. was a rousing success. [Click here to see photos and video and read what it was like for everyone involved...](#)

It's a Divestment Bonanza!

We've recently received word that the state of Maine has divested from Sudan. That's already fantastic news, but to top that off, one of California's major retirement funds – CalSTRS – has voted to divest as well. **About \$6 billion has been kept out of the hands of the Sudanese regime** thanks to these victories. [Can you help keep this momentum going by starting your own divestment campaign?](#)

Dear Friend:

You have a key role to play in helping stop the Sudanese genocide. International leaders need to see that the American people are outraged about the killing fields and slave raids in Sudan. Join individuals and communities across the country in demanding immediate action to stop the genocide perpetrated by the Sudan regime.



The movement to stop the genocide in Sudan is growing stronger every day. Through rallies, candlelight vigils, petitions and fundraisers, activists like you are n activism highlights, [click here](#).

If you have limited time but don't want genocide, this center is for you. It is di you to make an impact. Here are a few



[HOME](#) [ABOUT US](#) [LEARN](#) [TAKE ACTION](#) [CAMPAIGNS](#) [BLOG](#) [NEWSROOM](#) [FOR ACTIVISTS](#) [STORE](#)

Co-sponsored by HCUconnect.com

[SUBSCRIBE](#) to transAfrica Forum eNews!

[Home](#) | [Our Mission](#) | [Staff](#) | [Board](#) | [Contact Us](#) | [Internships](#) | [Link](#)

Global Justice Campaigns

ONE STANDARD
For a Fair & Just US Policy towards Haitian Refugees
Haiti, Free & Sovereign

Arthur R. Ashe Foreign Policy Library

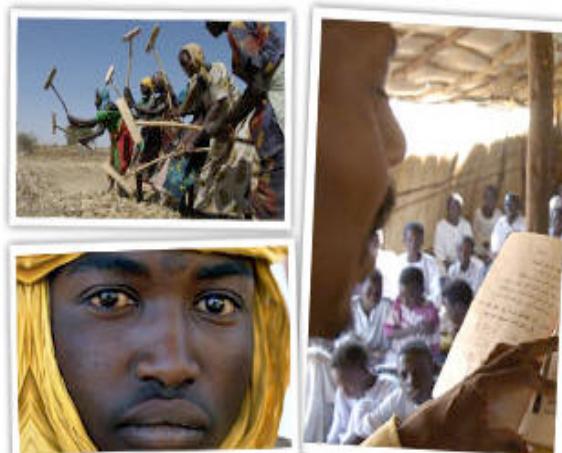
Resources
African World Visions
Film Series
Viewpoint Lecture Series
Readers' Corner
Writers' Corner

transAfrica FORUM NEWS ARTICLES, COMMENTARIES

TransAfrica Forum is seeking to fill the position of [President](#)

TransAfrica Forum News, November 2006
An October Surprise: Grassroots Movement's 20th Anniversary Foreign Policy Victory
By Dr. Sylvia Hill,
Vice Chair, TransAfrica Forum

On October 2, 1986 the USA anti-apartheid



EMBASSY OF THE REPUBLIC OF SUDAN

THIS IS SUDAN VISA / PASSPORT TRAVEL / TOURISM

Monday, November 27, 2006 | WASHINGTON, D.C.

View the Online Newsletter!



To download this brochure for printing or distribution, [click here](#) to download the original PDF file.

EMBASSY OF SUDAN

HEADLINES

Darfur Rebels'

Washington, DC, Jia total disregard for non-signatories to words Justice and of the Sudan Liber attack outside the 260 miles from Khartoum. The outlaws attack heavy weapons. And almost the entire homes...

Consular Office

The Consular Office Embassy Building I Washington, D.C. 202-336-0565. Our Monday to Friday.

Sudan Easter

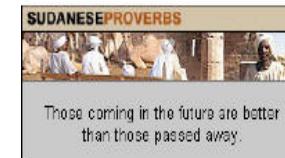
Rebels in eastern Sudan have agreed to end hostilities, the Entente

An Argument. Violence Gen

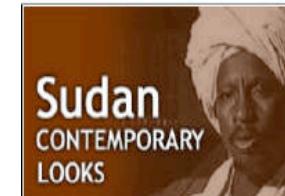
To call something what's happening in Darfur a civil war is inaccurate. The Africa Program Sunday's Currents wrote that calling it incoherent, but cou

Strategic Victi

THOUSANDS OF African military in government from p in Darfur must be p



Those coming in the future are better than those passed away.



LC Manuscript Division Archive of Organizational Web Sites

The Civil Rights Coalition for the 21st Century

Founded by LCCR and LCCREF

Saturday October 11

Issues select an issue Site Search advanced search Login

Ground Broken on Memorial to Honor the Rev. Dr. Martin Luther King, Jr.

The Rev. Dr. Martin Luther King, Jr., the most inspirational leader of the Civil Rights Movement, will soon be memorialized in the nation's capital. On Monday, November 13, civil rights leaders, two Presidents, three of Dr. King's children, and thousands of others gathered for the groundbreaking ceremony of the first monument to a black American

King, Jr. full story | more stories

2006 Supreme Court Term Could Be Critical for Civil Rights

The Supreme Court will hear a number of important cases during this session that... full story | more stories

Lawsuit Threatens Minority Journalism Program

In yet another attack on affirmative action



News & Events

Finding Aid Freud Chronology

LEAGUE OF WOMEN VOTERS Making Democracy Work® Grassroots leadership since 1920

ABOUT US SUPPORT US JOIN US CONTACT US FIND A LOCAL LEAGUE LOGIN

Stay Informed

Sign up for the League's e-newsletter and get all the latest information delivered to your inbox. your email: »

Civil Liberties

Election Reform

Lobby Reform and Ethics

Redistricting Reform

Voting Rights Act

Arctic National Wildlife Refuge

Clean Air

DC Voting Rights

Find Your Representative

Opinion Column by LWVUS President Featured in Roll Call

The League is taking on today's toughest issues.

» Campaign Finance Reform

» Civil Liberties

» Election Reform

» Lobby Reform and Ethics

League 411.org

411.org, an easy-to-use website that provides essential information for the upcoming election.

VOTE411.org

ELECTION INFORMATION YOU NEED

Act Now!

Stand up for Democracy Today!

Get Action Alerts

Donate Now

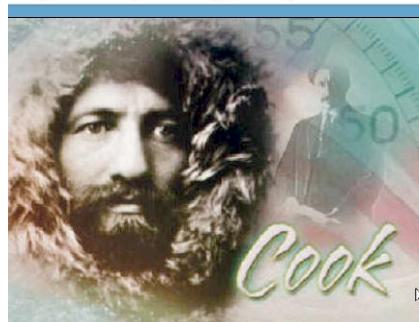
Join the League



"Frederick Cook remains as one of the most perceptive as well as the most vilified of polar explorers...Cook had made a discovery that no amount of humiliation could take from him and one that only the most fortunate and observant of arctic explorers ever learned."

Rob McGhee
"The Last Imaginary Place: A Human History of the Arctic World"
Oxford University Press, 2005

Dr. McGhee is curator of Arctic Archaeology at the Canadian Museum of Civilization and Massey Medalist of the Royal Canadian Geographical Society.



NCJW®

National Council of Jewish Women

ABOUT NCJW NEWS ISSUES & ACTION PROGRAMS JOIN NCJW SUPPORT NCJW

A FAITH IN THE FUTURE. A BELIEF IN ACTION.

SPEAK OUT TODAY READ THIS

welcome to the National Consumers League

about NCL publications NCL news Join NCL search contact us

October is NATIONAL CYBER SECURITY MONTH

Learn to Keep Safe from ID Thieves and Other Ghouls on social networking sites

STAYSAFEONLINE.org National Cyber Security Alliance

Fraud.erg • LifeSmarts • StepChildLabor.erg • SOSRx • PhishingInfo.org

NCL 15+ years of experience fighting for the health, safety, and welfare of consumers

let the National Consumers League be your guide to these issues >>>

your family's health

telemarketing & Internet fraud

money & credit finances

child labor and fair labor standards workers' rights

safety & nutrition food

products, services & security technology

phones, electricity & more utilities

protecting your personal information privacy

product safety, regulatory issues, etc. other interests

Visual Image Web Archive



SOCIAL DOCUMENTARY PHOTOGRAPHER



Greg Lynn FORM

1817 Lincoln Boulevard
Venice, CA 90291, USA
+01 (310) 821-2629 tel
+01 (310) 821-9729 fax

[email](#)



Prokudin-Gorskii Images

Color photographs from the Russian Empire taken a century ago (1909-1915).



BOSTON CONVENTION & EXHIBITION CENTER
Boston, Massachusetts, 2004

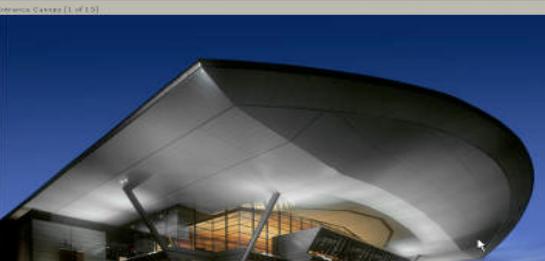
CIVIC CONVENTION CENTER
Completed
200,000 sq meters / 2,100,000 sq feet

IMAGES **DRAWINGS**

DESIGNERS

Described by Boston Globe architecture critic Robert Campbell as "the most dramatic new building in the city in years," the Boston Convention Center (BCC), located in the developing Historic Seaport District, is a 200,000-square-foot convention center and hotel developed by Rafael Viñoly Architects, P.C. (in joint venture with HNTB Architects). The 200,000-square-foot building extends from the new business and commercial district along the waterfront. The BCC is the first major project completed by the firm's Boston office.

ENTRANCE GALLERY (1 of 15)



Legal Blawgs Web Archive



The Harvard Law School
Corporate Governance Blog

CONTRIBUTORS:

HLS FACULTY & FELLOWS

Lucian Bebchuk
Robert Charles Clark

John Coates

Allen Farrell

Oliver Hart

Howell Jackson

Robert Jackson

Ronald Kraakman

Mark Ramseyer

Mark Roe

Hal Scott

Robert Sitkoff

Holger Spämann

Leo Strine, Jr.

Guhan Subramanian

Beth Young

PROGRAM ADVISORY BOARD

Peter Atkins

Joseph Bachelder

John Finley

Byron Georgiou

Robert Mendelsohn

Theodore Mirvis

Countrywide's Corporate Governance: Definitely Subprime

Posted by J. Richard Finlay, Centre for Corporate & Public Governance, thecentreforgovernance.org, on Wednesday October 31, 2007 at 9:48 pm

Countrywide Financial
subprime meltdown
scrambling to re-jig
corporate governance
important clues to what
stakeholders.

The lesson of Countrywide to retreat from Enron is necessary. On the contrary, when it comes to corporate governance, it is still the Typhoid Mary calamity strikes, as it did more. It also shows the consequences for other interests, which is why that of shareholders...

... continue reading
[Subprime](#)



Microsoft On China Software Counterfeiting Feels Good, Can I Have Some More?

Posted by Dan on August 2, 2007 at 08:21 am

The [Cndn Hst](#) media empire recently came out with a new bus called Portfire (which I have yet to see in the flesh) along with some other interesting websites. One of them is interesting a business, business, and business "lifestyle" type information number of blogs and those are mostly well done also.

Among the blog pieces was one by [Edith Salomon](#), entitled, "[Bodge Statistics, Software Edition](#)" that makes two very interesting points about Microsoft's software counterfeiting.

First, the blog talks about how the [recent news](#) regarding the jo Chinese operation nabbing "500 million" in Counterfeit Software misleadingly flat out wrong.

Salomon gives the reported facts as follows:

...and, she notes that the Chinese police and local authorities had led to the arrest of 25 people and the seizure of \$500 million worth of counterfeit Microsoft and Symantec software that was being made in China and distributed world wide. In the last couple of weeks, the operation led to the seizing Chinese government of 290,000 counterfeit discs and certificates of authenticity.

HealthLawBlog

Monday, June 04, 2007

TB quarantines are usually a local matter, rarely federal



That's what [last night's AP story](#) on Drew Speare reported. That local Maryland officials to detain many state health workers being kept in order -- the first in 44 years -- but they were replaced by a local Denver order

Lenz Blog

Copyright, Patents, WTO and whatever else interests me

Working Group Opinion on Search Engines

April 26th, 2008

As noted at [EU Law Blog](#), the EU Working Group on Data Protection has published an opinion on data protection law regarding search engines.

They say that the [Surveillance Directive](#) (data retention Directive 2006/24) does not give search engines any rights to store search histories, because these are not traffic data, but content of the communication, which may not be retained under Art. 5(2).

I agree with this result, but not with the reason given.

A search engine is retaining contents of a communication of which it is a party. If someone sends a query to Yahoo!, they are not using Yahoo!

ABOUT TLB

Philip Jessup proposed the idea of a transnational law course. His vision of the subject was broad, including public and private international law, state and non-state actors, business, administration, politics, culture, as well as negotiations and litigation. Inspired by his idea, TLB is established by its pursuit to address occurring national frontiers.

Authors

ody

igkins

xson

vinia

All For

ers
a committed interest in
al issues and you're a rising

August 01, 2007

Think that Clickwrap Contract is Iron-Clad? You May Want to Take a Closer Look

As an article posted on [Wired](#) today illustrates, clickwrap contracts are no cure-all for the companies that employ them -- especially large, consumer-oriented versus B2B ones. The title of this particular article -- "[Courts Turn Against Abusive Clickwrap Contracts](#)" -- though, is somewhat ambiguous. While the article does highlight a few recent cases in which courts have delved beyond the prima facie validity of a clickwrap contract -- turning a more critical eye toward its substantive terms, the reach of these cases is likely more limited than it first appears.

In the first case mentioned, [Citation v. T-Mobile](#), a California State Court of Appeal had recently found an arbitration clause in a T-Mobile clickwrap contract to be unenforceable; the plaintiff(s) had argued that the clause was, essentially, an end-run around California's [more liberal class action laws](#). The Court of Appeal agreed, finding a 'high degree' of substantive unconscionability in the arbitration clause. That the contract was a clickwrap one, then, provided the minimum degree of procedural unconscionability necessary to warrant the dismissal of T-Mobile's motion to compel arbitration. (By way of background, California law requires both an element of substantive unconscionability and of procedural unconscionability in order to support an overall finding of 'unconscionability').

Search

Pages

> About

Archives

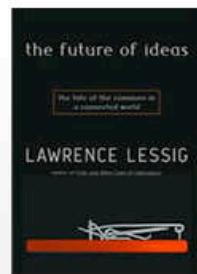
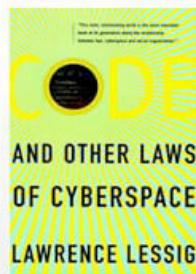
- > April 2008
- > March 2008
- > February 2008
- > December 2007
- > November 2007
- > October 2007
- > September 2007
- > August 2007
- > July 2007
- > June 2007
- > ...



Lessig 2.0



Blog News Wiki Mixer Whine



Egypt, 2008 Web Archive



[Random image](#)



tags used in blog
posts

political prisoners
technology **Multimedia**
bloggers in jail a7a
department مقالات
بـالعـربـي kefaya torture
breaking news sharkawy
guest writers manal
waragi tales freedom of
speech **english posts**
announecments and
statements politics
Web sexual harassment
citizen journalism protests
italy prison FOSS mubarak
feminism **Egypt** Culture
alaa journalism

Library of Congress Web Archives

<u>Election 2000</u>	800
<u>Election 2002</u>	4000
<u>Election 2004</u>	1945
<u>Election 2006</u>	2098
<u>Election 2008</u>	2000*
<u>107th Congress</u>	579
<u>108th Congress</u>	583
<u>109th Congress</u>	580*
<u>110th Congress</u>	580*
<u>September 11, 2001</u>	2300
<u>Winter Olympics 2002</u>	62
<u>Iraq War 2003-</u>	231
<u>Papal Transition 2005</u>	192
<u>Crisis In Darfur, Sudan 2006</u>	218
<u>Visual Images</u>	17
<u>Organizational Sites, Manuscript Division</u>	30
<u>U.S. Supreme Court Nominations 2005-2006</u>	281
<u>Legal Blawgs 2007</u>	90
<u>Egypt 2008</u>	30

Web Archives Processing Workflow

- **Identify and select sites**
- **Create a “seed list” of sites to be crawled, determine how frequently they will be crawled and submit to IA**
- **IA captures selected web sites as W/ARC files**
- **Create catalogers’ list and a MODS template for metadata extraction and submit them to IA**
- **IA extracts metadata from archived web sites (W/ARC files) into the MODS template**

Web Archives Processing Workflow

- **Metadata extraction results in a preliminary MODS record for each archived site**
- **Enhance record, reviewing & revising some values if needed (title, language, abstract, keywords) and adding some values (LCSH headings—subjects and sometimes names)**
- **Register item-level handles**
- **Load MODS records onto server, index, generate item-level search/browse**
- **Create collection-level record in ILS and register collection-level handle**

W/ARC Files

- **ARC file format used by Internet Archive to store web archives since 1996**
- **Access to archived web sites in ARC files depends on large-scale indexing of ARC files**
- **ARC file indexing can only support access by URL and date**
- **WARC has since been developed as an extension to ARC and is now an ISO standard, it carries a little more metadata, but access to web sites in WARC files is still very limited**
- **As tools are developed to support WARC files, WARC files will be preferred to ARC files for storing web archives**

NutchWAX Keyword Indexing & Search

- **NutchWAX is a web archive search technology based on Nutch an open source web search software—designed to improve access to W/ARC files**
- **NutchWAX ("Nutch + Web Archive eXtensions") can search keyword indexes of ARC files—so it extends more basic access (by URL and Date) to include keyword access**
- **However, building/rebuilding indexes for each archive is still cumbersome and expensive**
- **Building/rebuilding comprehensive indexes that include more than one web archive is even more cumbersome and expensive**
- **And even with NutchWAX's keyword access, archived sites are not searchable/browseable/integratable with other web archives or library resources**

Why Provide Site Level Access to these Sites?

- **Access limitations of W/ARC files and NutchWAX ("Nutch + Web Archive eXtensions")**
- **Use of controlled vocabularies**
- **Leverage subject cataloging & language expertise to quickly and substantially enhance subject access**
- **Resources become highly integratable with other library resources at the item level**
- **Better precision and recall**
- **Persistent IDs/handles allow for stable citations and digital scholarship at site-level**
- **Leverage use of existing search/browse systems**

How Do We Provide Site-Level Access to these Sites?

- **Boilerplate as much relevant archive-level and site-level metadata as is possible into the MODS template**
- **Extract as much useful metadata as is possible from archived web sites W/ARC files (using a perl script or other method that grabs the metadata from meta tags in the W/ARC files)—titles, dates, file types, abstracts, subject keywords, etc.**
- **Leverage LC subject cataloging & language expertise and controlled vocabularies to add subject access**

Overview of MODS Record Data Elements

Title	- Extracted from W/ARC file/HTML title tag - Cataloger uses if viable, otherwise supplies
Alternative Title	- Cataloger supplies if another useful and different title displays on piece
Name Personal	- Included for some archives, when relevant, cataloger supplies
Name Corporate	- Included for some archives, when relevant, cataloger supplies
Type of Resource	- Boilerplate "text"
Genre	- Boilerplate "Web site"
Origin Info	- Extracted from W/ARC file – first/last dates captured YYYYMMDD(iso8601)
Language	- Boilerplate in if known (iso639-2b code) - Cataloger can supply additional languages
Physical Description	- Extracted from W/ARC file/MIME type, e.g., text/css, image/jpeg
Abstract	- Extracted from W/ARC file/META name=description content - Cataloger can edit/enhance
Subject/Keywords	- Extracted from W/ARC file/META name=keywords content - Cataloger can edit/enhance
Subject/LCSH	- Cataloger supplies
Collection Title/PID	- Boilerplate, collection title & collection PID/handle
Identifier	- Boilerplate, variant of handle, e.g, hdl:loc.natlib/mrva0000.0000
Note	- Extracted from W/ARC file, resolves to URL for active site
Location/Usage	- Boilerplate item-level PID/handle - PID is registered to resolve to archived Web site URL
Access Condition	- Boilerplate rights info/permissions info – imported from OSI records
Record Info	- Boilerplate record creation date - Boilerplate record identifier, handle suffix mrva0000.0000

Crisis in Darfur, Sudan 2006 Web Archive

Archive size: 218 sites

Harvest info: 1 phase, multiple captures

Frequency: Varies--weekly to monthly crawls for each site

Metadata: 1 collection-level MARC record, with collection level PID
218 item-level MODS records, with item-level PIDs

LCSH: 1 boilerplate LCSH heading
Unlimited specific LCSH headings at site level—these are selected by cataloger from a list of about 20 LCSH terms that relate to the content in the archive

Catalogers' List for Darfur, 2006 Web Archive

Wayback Machine links to seeds:

-- **Live/Proxy** links to the live Web. To enable proxy viewing
manually enter the proxy host and port informat:
Proxy host: loc-wm.archive.org
Port number: 9006
-- **Wayback 1.4** links to Wayback Machine 1.4.

-- **WARC** links to WARC deployment.

- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://againstoti.blogspot.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://arabist.net/arabawy/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://egyworkers.blogspot.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://ghosty1313.blogspot.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://masgon.maktoobblog.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://masradam.blogspot.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://tadamonmasr.wordpress.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://wa7damastrya.blogspot.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.6-april.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.al-araby.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.alwafid.org/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.alwasatparty.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.dakahliaikhwan.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.egymasr.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.egytiangreens.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.el3amal.net/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.elbehira.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.gharbiaonline.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.harakamasria.org/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.ikhwanonline.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.ikhwanpress.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.imbh.net/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.keepourgas.blogspot.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.manalaanet.net/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.mdedalghalaa.org/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.nowabikhwan.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.saveegyptfront.org/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.sharkiaonline.com/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://www.tortureinegypt.net/>
- [[Live/Proxy](#)] [[Wayback 1.4](#)] [[WARC](#)] <http://ymeenshemal.blogspot.com/>

Resource Page for an Archived Web Site, Darfur, 2006 Web Archive

Crisis in Darfur, Sudan, Web Archive, 2006 Resource Page

Each date listed below represents one capture of the site. Clicking any date will display the web site record as it appeared at that time. Following other links from that page will return the closest matches available from this Web Archive, however, some linked pages may not be available in the Archive. Please use the browser back button to return to this page if you wish to visit other archived dates.

7 archived captures for <http://www.afrika.no/>

* denotes when site was updated.

Search Results for Jan 1, 2006 - Dec 31, 2006					
Jan 2006 - Feb 2006	Mar 2006 - Apr 2006	May 2006 - Jun 2006	Jul 2006 - Aug 2006	Sep 2006 - Oct 2006	Nov 2006 - Dec 2006
0 pages	0 pages	0 pages	4 pages	2 pages	1 page
			Jul 17, 2006 * Jul 20, 2006 Aug 21, 2006 * Aug 26, 2006 *	Sep 18, 2006 * Oct 16, 2006 *	Nov 20, 2006 *

[home](#) >> [previous page](#) >> Crisis in Darfur, Sudan, Web Archive, 2006 Resource Page

[The Library of Congress](#) [More Online Collections](#)

[Contact Us](#)

Bilingual (eng/nor) Archived Web Site - Darfur, 2006 Web Archive

afrika.no
The Norwegian Council for Africa

Fellesrådet | The Index on Africa | Africa News Update | English

Updated: Friday 14, July

Africa at large: Rich nations are offside in Geneva
Source: [Business Day \(South Africa\)](#), by [Brendan Vickers](#), 07.06.2006
The past week's Soccer World Cup matches in Germany spawned far more superlatives than did the sombre drama of Geneva, where a meeting of 60 of the world's trade ministers failed to revive the stalled Doha Round of the World Trade Organisation (WTO). Indeed, the virtuosic quarterfinal victory of the French delegation over their Brazilian counterparts is perhaps representative of the history of multilateral trade regulation: northern protectionist interests have always played an adroit game, skilfully outflanking the legitimate developmental ambitions of the south. And while the Africans put up a brave fight in this tournament, most teams — as in past trade negotiations, notably the Uruguay Round — were soon knocked out and sent packing.
[read more >>>](#)

Africa at large: Aid For the Poor, Not For the Consultants
Source: [Inter Press Service \(Africa\)](#), by [Moyiga Nduru](#), 07.06.2006
Johannesburg - No less than a quarter of annual development aid -- about 20 billion dollars -- is being used by donor countries to fund technical assistance of sometimes dubious worth, says ActionAid International in a new report.
[read more >>>](#)

Uganda: Museveni amnesty to Kony illegal- ICC
Source: [Monitor \(Uganda\)](#), by [Frank Nyakairu](#), 07.06.2006
THE International Criminal Court has reminded Uganda of her obligation to arrest Lord's Resistance Army rebel leader Joseph Kony.
[read more >>>](#)

Uganda: Donors warn of more aid cuts
Source: [Monitor \(Uganda\)](#), by [Elias Biryabarema and Hussein Bogere](#), 07.06.2006
The government should prepare for more announcements of cuts in budget support, sources within the diplomatic community in Kampala have warned.
[read more >>>](#)

Kenya: President's office perceived most corrupt
Source: [Standard \(Kenya\)](#), by [Alex Ndegwa](#), 07.06.2006
It's now official: The Office of the President is perceived to be the most corrupt in the land. A new survey by a Government watchdog, whose chief executive is the President's own appointee, has arrived at this damning verdict.
[read more >>>](#)

Entire afrika.no | Index on Africa | News Update | Search

The Norwegian Council for Africa
Osterhaugst. 27
N-0183 Oslo
Norway
Phone: +47 22989312
Fax: +47 22989301
Email: afrika@afrika.no

[about the nca](#) | [contact us](#)

The Index on Africa
Index on Africa is a gateway to information on Africa on the Internet, with over 3.800 links sorted by country, subject and news.

The Index has been created by The Norwegian Council for Africa (NCA), as a part of NCA's efforts to raise awareness about Africa and African affairs.

Africa News Update
Join our free newsletter, [Africa News Update](#), and keep yourself updated on what's happening on the continent!

60

Join | Remove

Preliminary MODS Record – Darfur, 2006 Web Archive

The screenshot shows the Oxygen XML Editor interface. On the left is the Project Explorer pane, which lists several XML files under a project named "darfurProjectDone.xpr". One file, "mrva0011.0001.xml", is selected and highlighted with a blue background. The main workspace on the right displays the XML code for this selected file. The code is a MODS record for a web archive, structured with various elements like titleInfo, typeOfResource, genre, originInfo, language, physicalDescription, subject, relatedItem, identifier, note, location, and recordInfo.

```
<mods xmlns="http://www.loc.gov/mods/v3" version="3.2">
  <titleInfo>
    <title>Darfur Crisis</title>
  </titleInfo>
  <typeOfResource>text</typeOfResource>
  <genre>Web site</genre>
  <originInfo>
    <dateCaptured encoding="iso8601" point="start">20060517</dateCaptured>
    <dateCaptured encoding="iso8601" point="end">20061120</dateCaptured>
  </originInfo>
  <language>
    <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
  </language>
  <physicalDescription>
    <internetMediaType>application/pdf</internetMediaType>
    <internetMediaType>text/html</internetMediaType>
  </physicalDescription>
  <subject authority="lcsh">
    <geographic>Sudan</geographic>
    <topic>History </topic>
    <temporal>Darfur Conflict, 2003-</temporal>
  </subject>
  <relatedItem type="host">
    <titleInfo>
      <title>Crisis in Darfur, Sudan Web Archive, 2006</title>
    </titleInfo>
    <locations>
      <url>http://hdl.loc.gov/loc.natlib/collnatlib.00000011</url>
    </locations>
  </relatedItem>
  <identifier>hdl:loc.natlib/mrva0011.0001</identifier>
  <note type="system details">65.109.167.118/pipa/pdf/jun05/Darfur_Jun05_rpt.pdf</note>
  <location>
    <url displayLabel="Archived site">http://loc.archive.org/darfur/2006*/65.109.167.118/pipa/pdf/jun05/Darfur_Jun05_rpt.pdf</url>
  </location>
  <location>
    <url usage="primary display">http://hdl.loc.gov/loc.natlib/mrva0011.0001</url>
  </location>
  <accessCondition>None</accessCondition>
  <recordInfo>
    <recordCreationDate encoding="iso8601">20070516</recordCreationDate>
    <recordIdentifier source="dlc">mrva0011.0001</recordIdentifier>
  </recordInfo>
</mods>
```

MODS Subject Heading List - Darfur, 2006 Web Archive

The screenshot shows a software application window with a toolbar at the top and two main panes below. The left pane, titled 'Project', contains a tree view of files under 'darfurProjectDone.xpr'. The right pane, titled 'darfurLCSH.xml*', displays the XML code for the subject headings.

```
1   Additional subject headings (catalogers may choose additional headings from among those listed below):
2
3   151 $a Sudan $x History $y Civil War, 1983-2005
4   <subject authority="lcsh">
5     <geographic>Sudan</geographic>
6     <topic>History</topic>
7     <temporal>Civil War, 1983-2005</temporal>
8   </subject>
9
10  151 $a Sudan $x Politics and government $y 1985-
11  <subject authority="lcsh">
12    <geographic>Sudan</geographic>
13    <topic>History</topic>
14    <temporal>Civil War, 1985-</temporal>
15  </subject>
16
17  151 $a Sudan $x Economic conditions $y 1983-
18  <subject authority="lcsh">
19    <geographic>Sudan</geographic>
20    <topic>History</topic>
21    <temporal>Civil War, 1983-</temporal>
22  </subject>
23
24  151 $a Sudan $x Civilization
25  <subject authority="lcsh">
26    <geographic>Sudan</geographic>
27    <topic>Civilization</topic>
28  </subject>
29
30  150 $a Public opinion
31  <subject authority="lcsh">
32    <topic>Public opinion</topic>
33  </subject>
```

Completed MODS Record – Darfur, 2006 Web Archive

```
<mods xmlns="http://www.loc.gov/mods/v3" version="3.2">
<title Info><title>afrika.no: The Norwegian Council for Africa</title></title Info>
<type Of Resource>text</type Of Resource>
<genre>Web site</genre>
<origin Info>
<date Captured encoding="iso8601" point="start">20060717</date Captured>
<date Captured encoding="iso8601" point="end">20061120</date Captured>
</origin Info>
<language>
<language Term authority="iso639-2b" type="code">eng</language Term>
<language Term authority="iso639-2b" type="code">nor</language Term>
</language>
<physical Description>
<internet Media Type>application/download</internet Media Type>
<internet Media Type>application/x-javascript</internet Media Type>
<internet Media Type>image/bmp</internet Media Type>
<internet Media Type>image/gif</internet Media Type>
<internet Media Type>image/jpeg</internet Media Type>
<internet Media Type>image/pjpeg</internet Media Type>
<internet Media Type>text/css</internet Media Type>
<internet Media Type>text/html</internet Media Type>
</physical Description>
<abstract>afrika.no - The Index on Africa and Africa News Update. Features news on and links to all countries in Africa. With sections on Culture, Development, Economy, Education, Environment, Health, Human Rights, News and Politics. By the Norwegian Council for Africa.</abstract>
<subject authority="keyword"><topic>afrika, africa, culture, development, economy, education, environment, health, politics, travel</topic>
</subject>
<subject authority="lcsh">
<geographic>Sudan</geographic>
<topic>History</topic>
<temporal>Darfur Conflict, 2003-</temporal>
</subject>
<subject authority="lcsh">
<topic>International relief</topic>
</subject>
<subject authority="lcsh">
<geographic>Sudan</geographic>
<topic>Economic conditions</topic>
<temporal>1983-</temporal>
</subject>
<related Item type="host">
<title Info><title>Crisis in Darfur, Sudan Web Archive, 2006</title></title Info>
<location><url>http://hdl.loc.gov/loc.natlib/collnatlib.00000011</url></location>
</related Item>
<identifier>hdl:loc.natlib/mrva0011.0037</identifier>
<note type="system details">www.afrika.no/</note>
<location><url display Label="Archived site">http://loc.archive.org/darfur/2006*/www.afrika.no/</url></location>
<location><url usage="primary display">http://hdl.loc.gov/loc.natlib/mrva0011.0037</url></location>
<access Condition>Access restricted to on-site users at the Library of Congress.</access Condition>
<record Info>
<record Creation Date encoding="iso8601">20070516</record Creation Date>
<record Identifier source="dlc">mrva0011.0037</record Identifier>
</record Info>
</mods>
```

Displayed MODS Record - Darfur, 2006 Web Archive

Title: afrika.no: The Norwegian Council for Africa

Abstract: afrika.no - The Index on Africa and Africa News Update. Features news on and links to all countries in Africa. With sections on Culture, Development, Economy, Education, Environment, Health, Human Rights, News and Politics. By the Norwegian Council for Africa.

Date Captured: July 17, 2006 - November 20, 2006

[Archived Site](#)

Subject(s): Sudan--History--Darfur Conflict, 2003-
International relief
Sudan--Economic conditions--1983-

Language(s): English
Norwegian

Genre: Web site

Access Condition: Access restricted to on-site users at the Library of Congress.

URL at time of capture: www.afrika.no/

Citation ID: <http://hdl.loc.gov/loc.natlib/mrva0011.0037>

Record ID: mrva0011.0037

Collection Title: [Crisis in Darfur, Sudan Web Archive, 2006](#)

Library of Congress Web Archives Homepage

The Library of Congress >> More Online Collections

Library of Congress Web Archives *Minerva*

BROWSE | SEARCH | TECHNICAL INFORMATION

[home](#)

Web Archives Available:

- [United States 107th Congress Web Archive](#)
- [United States 108th Congress Web Archive](#)
- [Crisis in Darfur, Sudan, Web Archive, 2006](#)
- [Library of Congress Manuscript Division Archive of Organizational Web Sites](#)
- [Papal Transition 2005 Web Archive](#)
- [September 11, 2001 Web Archive](#)
- [United States Election 2000 Web Archive](#)
- [United States Election 2002 Web Archive](#)
- [Visual Image Web Sites Archive](#)



The Library of Congress Web Archives (LCWA) is composed of collections of archived web sites selected by subject specialists to represent web-based information on a designated topic. It is part of a continuing effort by the Library to evaluate, select, collect, catalog, provide access to, and preserve digital materials for future generations of researchers. The early development project for Web archives was called MINERVA.

[home](#)

The Library of Congress >> More Online Collections
March 6, 2008

Contact Us

Collection Overview - Darfur, 2006 Web Archive

OVERVIEW

SEARCH

BROWSE

Collection Overview

Citing Resources - Copyright Information

Scope: Described by the United Nations as the world's worst humanitarian crisis, the current conflict in Sudan's western region of Darfur began in early 2003, when rebels attacked government targets and then expanded to the widespread killing of civilians and the destruction of villages. These attacks have led to more than two million people being displaced and thousands killed. More than 10,000 aid workers, both Sudanese and international, are actively involved in assisting victims of this man-made disaster.

In the spring of 2004, the conflict began to receive broad international press coverage; scholars and development specialists produced in-depth studies and reports. Governments issued position papers and sent representatives to Sudan to see the situation firsthand. New organizations, each with Web sites, have been formed because of the Darfur crisis. Reminiscent of those seen during the 1970's and 1980's which opposed the oppressive apartheid system in South Africa, campaigns have begun urging U.S. states and universities to divest from companies doing business in Sudan and asking the public to petition its government officials to take action. Web sites and blogs have replaced many of the pamphlets, flyers, and newsletters used to disseminate information in an earlier era. A Web Archive of the Crisis in Darfur, Sudan, will preserve the documentation of this humanitarian crisis for future historians, and will include sites of key organizations, a sampling of news reports, and the responses of government, international organizations and the general public in the U.S. and worldwide.

This collection is part of a continuing effort by the Library of Congress to evaluate, select, collect, catalog, provide access to, and preserve digital materials for future generations of researchers.

Collection Period: March 20, 2006 - November 20, 2006

Search Page - Darfur, 2006 Web Archive

The Library of Congress >> More Online Collections

Library of Congress Web Archives *Minerva*

BROWSE | SEARCH | TECHNICAL INFORMATION

[home](#) >> [overview](#)

Crisis in Darfur, Sudan, Web Archive, 2006

OVERVIEW

SEARCH

BROWSE

Search this Collection

Enter Search Word(s):

GO!

>> [search all collections](#)

Search On (select only one):

- All fields
-
- Name
- Title
- Subject
- Abstract
- Language
- Year captured (YYYY)
- Record ID (mrva1234.1234)

[Search Help](#)

[home](#) >> [overview](#)

The Library of Congress >> More Online Collections
March 6, 2008

Contact Us

Browse Page - Darfur, 2006 Web Archive

The Library of Congress >> More Online Collections

Library of Congress Web Archives *Minerva*

BROWSE | SEARCH | TECHNICAL INFORMATION

[home](#) >> [overview](#) >> [browse results](#)

Browse:

Crisis in Darfur, Sudan Web Archive, 2006
mrva0011

Browse by » [Subject](#) [Name](#) | [Title](#)

(16 subjects)



- ➊ [Disaster relief](#)
- ➋ [Disinvestment](#)
- ➌ [Genocide](#)
- ➍ [Human rights](#)
- ➎ [Humanitarian intervention](#)
- ➏ [International relief](#)
- ➐ [Peacekeeping forces](#)
- ➑ [Public opinion](#)
- ➒ [Refugees](#)
- ➓ [Sexual abuse victims](#)
- ➔ [Sudan--Civilization](#)
- ➕ [Sudan--Economic conditions--1983-](#)
- ➖ [Sudan--History--Civil War, 1983-2005](#)
- ➗ [Sudan--History--Civil War, 1985-](#)
- ➘ [Sudan--History--Darfur Conflict, 2003-](#)
- ➙ [Sudan--Politics and government--1985-](#)

[home](#) >> [overview](#) >> [browse results](#)

The Library of Congress >> More Online Collections
March 6, 2008

Contact Us

MARC Collection-Level Record - Darfur, 2006 Web Archive

[Brief Record](#)

[Subjects/Content](#)

[Full Record](#)

[MARC Tags](#)

Crisis in Darfur, Sudan, Web archive, 2006

Relevance: 

LC Control No.: 2007700187

LCCN Permalink: <http://lccn.loc.gov/2007700187>

Type of Material: Book (Print, Microform, Electronic, etc.)

Main Title: Crisis in Darfur, Sudan, Web archive, 2006 [electronic resource].

Published/Created: Washington, D.C. : Library of Congress, 2007-

Related Names: [Library of Congress.](#)

Access Advisory: Descriptions of the archived Web sites are searchable without restriction; access to some archived sites is restricted to on-site users of the Library of Congress.

Summary: Selective collection of 216 Web sites, archived from March 20, 2006 to Nov. 20, 2006, relating to the humanitarian crisis in Darfur, Sudan. The archived sites are international in scope, and include those related to organizations involved in human rights, refugees, disaster relief, etc., government sites, research and public opinion group sites, sites of religious organizations, press organization sites involved in reporting on the Darfur crisis, and other miscellaneous sites.

Notes: Title from overview page, as viewed on Feb. 16, 2008.

"Library of Congress Web Archives, Minerva."

Subjects: [Disaster relief --Sudan --Darfur.](#)

[Atrocities --Sudan --Darfur.](#)

[Disinvestment --Sudan --Darfur.](#)

[Genocide --Sudan --Darfur.](#)

[Human rights --Sudan --Darfur.](#)

[Humanitarian intervention --Sudan --Darfur.](#)

[International relief --Sudan --Darfur.](#)

[Peacekeeping forces --Sudan --Darfur.](#)

[Public opinion --Sudan --Darfur.](#)

[Refugees --Sudan --Darfur.](#)

[Sexual abuse victims --Sudan --Darfur.](#)

[Web archives.](#)

[Sudan --History --Darfur Conflict, 2003-](#)

[Sudan --History --Darfur Conflict, 2003- --Atrocities.](#)

[Sudan --History --Darfur Conflict, 2003- --Civil relief.](#)

Google Search – Item in Darfur, 2006 Web Archive

Web Images Maps News Shopping Gmail more ▾

Google darfur afrika.no Search Advanced Search Preferences

Web

[afrika.no - Sudan: Darfur's lifeline radio \(Comment\)](#)
afrika.no - The Index on Africa and Africa News Update. Features news on and links to all countries in Africa. With sections on Culture, Development, ...
www.afrika.no/Detailed/13699.html - 24k - [Cached](#) - [Similar pages](#)

[afrika.no - Darfur and Somalia: Two peace efforts, different ...](#)
In **Darfur**, peace efforts have taken a multilateral approach, and with Sudan caving Norway
Phone: +47 22989312 Fax: +47 22989301 Email: afrika@afrika.no ...
www.afrika.no/Detailed/14078.html - 29k - [Cached](#) - [Similar pages](#)
[More results from www.afrika.no »](#)

[afrika.no: The Norwegian Council for Africa\(Library of Congress ...](#)
URL at time of capture: www.afrika.no/. Record ID: mrv0011.0037. Collection Title: Crisis in **Darfur**, Sudan Web Archive, 2006 ...
lcweb2.loc.gov/diglib/lcwa/mrv0011.0037/ - 7k - [Cached](#) - [Similar pages](#)

[Search Results \(All Collections: Library of Congress Web Archive ...](#)
5. AfricaFocus home page. From: Crisis in **Darfur**, Sudan Web Archive, 2006. [bibliographic information]. 6. **afrika.no**: The Norwegian Council for Africa ...
lcweb2.loc.gov/diglib/lcwa/searchAll?query=++collection:mrv0011%20...History--
Darfur%20Conflict%2C%202003... - 10k - [Cached](#) - [Similar pages](#)
[More results from lcweb2.loc.gov »](#)

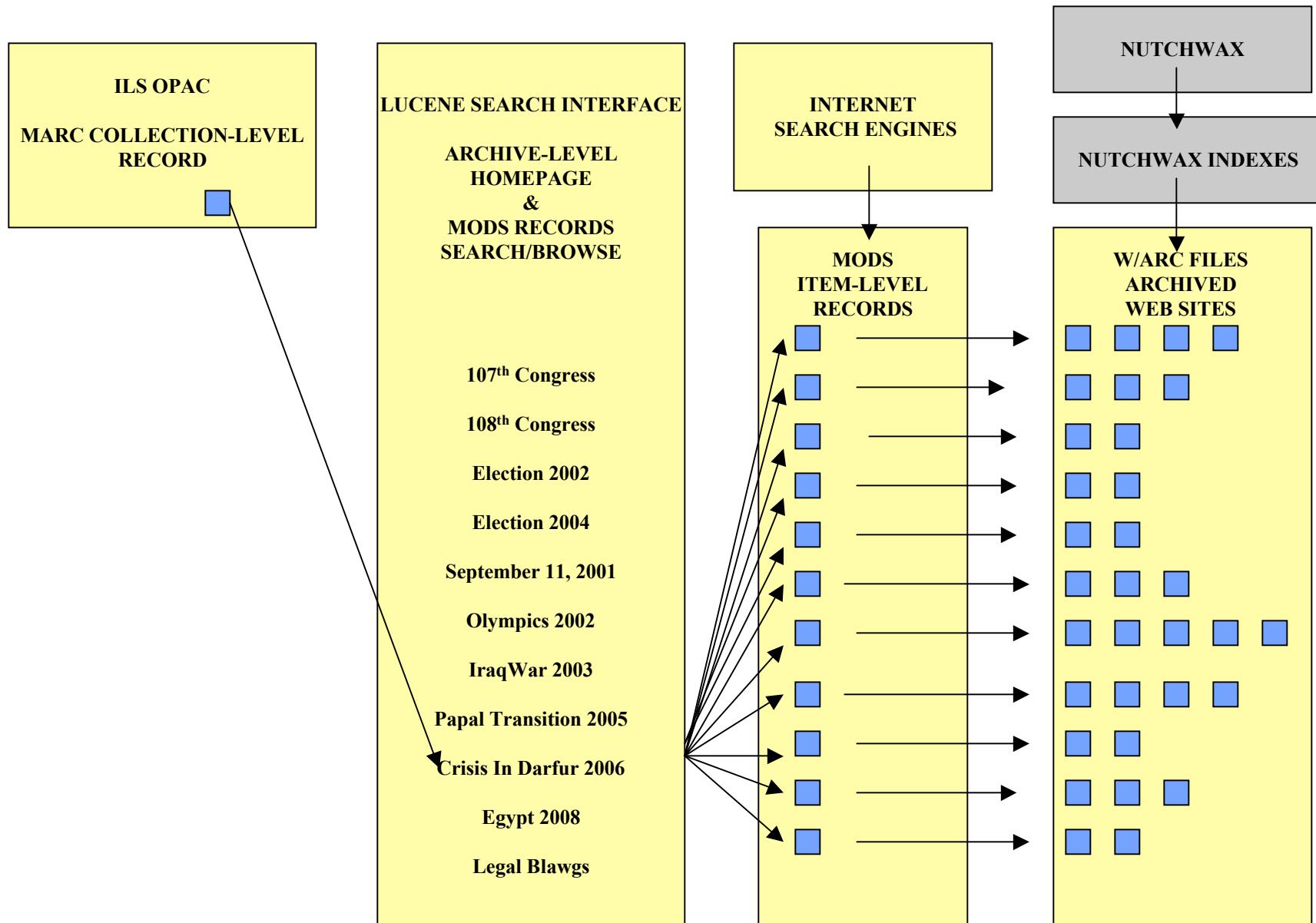
[Sud Afrika: No shoes](#)
Sud **Afrika**. "Pray that God may open to us a door for the word, to declare the mystery of ...
Disculpen que **no** se pueden ver las fotos en esta pagina. ...
gabysalazar.blogspot.com/2007/06/no-shoes.html - 69k - [Cached](#) - [Similar pages](#)

[Afrika News - Darfur: Is this ceasefire for real?](#)
Written by Ernest Waititu: Editor-in-chief, **Afrika** News ... they were monitoring a paper ceasefire that had **no** resonance on the sun-baked ground in **Darfur**. ...
afrikane.ws.org/index.php?option=com_content&task=view&id=28&Itemid=49 - 22k - [Cached](#) - [Similar pages](#)

[Afrika News - Legendary NBA Player Appeals for Help in Darfur](#)
Mar 2, 2007 ... Written by Ernest Waititu: Editor-in-chief, **Afrika** News ... Why not **Darfur**, he wondered? He urged Americans to write letters to their ...
afrikane.ws.org/index.php?option=com_content&task=view&id=61&Itemid=26 - 25k - [Cached](#) - [Similar pages](#)
[More results from afrikane.ws.org »](#)

[DARFUR: THE FIGHT FOR PEACE](#)
'Nigeria sends additional troops to **Darfur**', www.afrika.no/Detailed/6484.html. 38. ' EU envoy pledges aid for **Darfur**', http://news.bbc.co.uk/2/hi/africa/ ...
www.issi.org.pk/Journal/2004_files/no_4/article/5a.htm - 73k - [Cached](#) - [Similar pages](#)
by M DARFUR - [Related articles](#)

LC Web Archives – Levels of Access



Results - Pros

- **Archived resources are searchable and indexable along with other library collections and online resources**
- **Item-level and collection-level subject access and controlled vocabularies make these resources highly integratable at the item level and collection-level**
- **Site-level access facilitates searching and browsing within and across web archives—ability to find, refind & cite resources**
- **Good use and reuse of extracted and human-created metadata—friendly environment in which traditional catalogers learn XML and MODS—project benefits from specialized subject cataloger expertise**
- **Flexible and sustainable infrastructure for making web archives available for digital scholarship—stable/citable persistent IDS at the site level and the collection level**

Results - Cons

- **Scalability—approach works well with archives of up to 2,000 sites, but hasn't been tested w/much larger archives**
- **Project investment is basically the same for each archive—whether it's 100 sites or 2000 sites--project setup still requires template creation, metadata extraction, LCSH analysis at archive level, handle registration, etc.—so essentially the same amount of resources regardless of archive size**

Future Considerations

- **MODS tools—need for a flexible MODS input/editing form that would hide boilerplate and extracted metadata that the cataloger does not need to see—we have experimented w/XMLSPY's Authentic and XForms, but we lose flexibility w/regard to parsed subjects with both of these**
- **Future plans to integrate the NutchWAX component to provide more comprehensive keyword access to W/ARC files—this will complement existing collection and site-level access**
- **Experiment tag cloud generators to increase subject keyword access**

Tag Cloud Generated from Archived Web Site Darfur, 2006 Web Archive

#111 \$1 \$1 billion. ActionAid Africa African Aids Alex Archive Amy Belinda Beresford Biryabarema Bogere Brazilian Brendan Business Congo Congress Consultants Council Court Criminal Cup DRC Day Democratic Doha
Donors Ecowas Elias Emad Email English Entire External Fax Fellerdet Flow Frank French Friday Geneva Germany Guardian HIV Hussein ICC Immigration Index Inter International Internet It's Johannesburg Join Joseph Jul July
Kampala Kenya Kony Leaders Lesotho Liberia Liberian Library Light Lord's Mail Mekay Minerals Misery Monitor Moyiga Museveni N-0183 NCA's Ndegwa Nduru Needed NewZimbabwe.com Norad Norway Norwegian Nyakairu
Observer Organisation Oslo Osterhaugt Phone Poor President President's Press Remains Remove Republic Resistance Rich Service Sibanda Soccer Source Standard Testing Tichaona Tuesday UK Uganda Update
Uruguay Vickers WTO Washington Web Wednesday World Zimbabwe Zimbabwean account adroit affairs .africa news update afrika.no agents aid ambitions amnesty announcements annual appointee archived arrest arrived
assistance asylum asylm awareness billion boxes brave brutal budget case charged chief chilling civil cobalt collection communities community companies contact continent copper corrupt counterparts countries country country's created cuts damning
delegation destroyed developed development developmental diplomatic dollars donor drama dubious efforts elements from elites embarked energy epidemic evidence ex-CIO executive failed fight forms free function fund game gateway government group happening
hearing hears history illegal- impact information .inter press service africa. interests knocked land large leader legitimate links local matches meeting ministers .monitor uganda. multilateral nations nca negotiations network new .new
report news news update. newslist noop northern norwegian council notably obligation office official offside opposition outflanking packing page part perceived played pocketing poverty prepare previous
production protectionist put quarter quarterfinal raise read rebel rebuild reduce regulation reminded report representative revenues revive revolutionary search security seekers sharing site skillfully sombre sorted sources south south africa.
spawned spending spread stalled state strategy subject superlatives support supported survey teams technical test tournament trade tribunal undesirable updated verdict victory viewing virtuosic virus war wam warned watchdog webmaster week's
world's worth deal ♦ ♦

THAT'S ALL FOLKS



tmee@loc.gov