



MULTIPLE ACCESS POINTS

CREATING THE HUMAN RIGHTS WEB ARCHIVE PORTAL
AT COLUMBIA UNIVERSITY

ALEX THURMAN
Archive-It Partners Meeting
12/3/2012

- * In 2009 Columbia University Libraries got a grant from Andrew W. Mellon Foundation to implement a web resources collection program
- * Grant funded two 3-year Web Curators and one 2-year Programmer
- * Curators to build Human Rights Web Archive
- * Programmer to help build local access portal

* background

* www.archive-it.org collection page

Ability to search all web archive collections, browse by url with limited metadata (as of 2009)

* MARC records in CLIO catalog for individual archived websites and for web archive collections

* Local CUL portal for human rights collection

Branded, fuller integration of web archive into CUL website.
Enhanced browsing, searching features



multiple access points for web archives

Select an Institution ▾

...Or Search All Collections ▾

GO

Advanced Search

COLLECTION: Human Rights



Partner: [Columbia University Libraries](#)

Crawling Activity: 2008 - present

Topics: Society & Culture

Videos Archived: [list videos](#)

An initiative of CUL's Center for Human Rights Documentation and Research, the Human Rights Web Archive is an effort to preserve and ensure access to freely available human rights resources created mainly by non-governmental organizations, national human rights institutions, and individuals. Funded by the Andrew W. Mellon Foundation.

Enter Search Terms Here...

GO

Advanced Search

Prefer to search by URL? Enter URLs Here...

http://

Take Me Back

• 1. Non-governmental organizations (NGOs)-Amnesty International sites

- <http://amnesty.ca> - Amnesty International. Canadian Section-English-Speaking (Mar 11 2011 - Apr 22 2011)

Canadian English-speaking section of Amnesty International. For francophone section, see <http://www.amnistie.ca/>

- <http://amnesty.hu/> - Amnesty International Magyarország (Mar 17 2011 - Apr 05 2011)

Hungary section of Amnesty International. In Hungarian.

- <http://amnesty.ie/> - Amnesty International. Irish Section. (Mar 11 2011)

Ireland section of Amnesty International. In English.

- <http://amnesty.or.jp/> - Amnesty International. Nihon Shibu. (Mar 11 2011 - Apr 22 2011)



www.archive-it.org/collections/1068
(2009)

[Basic Search](#)[Advanced Search](#)[My Searches](#)[My List](#)[My Library Account](#)[» CLIO classic](#)[» CLIOmobile](#)

New Search :

Keyword

Go

[Search History](#)

Abuelas de Plaza de Mayo

Author: [Asociación de Abuelas de Plaza de Mayo](#)
Title: Abuelas de Plaza de Mayo [electronic resource].
Published: Argentina : Abuelas de Plaza de Mayo.
Online Link(s): [Current site](#)
[Archived site](#)
LC Subjects: [Asociación de Abuelas de Plaza de Mayo](#).
[Human rights--Argentina](#).
[Children--Argentina](#).
[Disappeared persons--Argentina](#).
[Human rights advocacy--Argentina](#).

Holdings Information:

Location (guide):	Online
Call Number:	ERESOURCES
Status:	No information available

Other Subject Terms: Non-governmental organizations[Web sites](#)

Summary: "Grandmothers of May Square is a civil non governmental organization which is not affiliated to any political party. Its aim is to locate the children who disappeared as a result of the political repressive methodology implemented by the military dictatorship and to restore them to their legitimate families."

Notes: Viewed on March 6, 2009.
Preserved by Columbia University Libraries' Web Resources Collection Program.

Languages: In Spanish; some text also in English and Italian.**Material Type:** Website

Place Request:

- [Place a Recall/Hold](#)
- [Borrow Direct](#)
- [ILL](#)
- [Document Delivery](#)
- [Inter-Campus Delivery](#)
- [In Process / On Order](#)
- [Precataloging](#)

Actions:

- [Print](#)
- [Export](#)
- [E-mail](#)
- [Send to phone](#)
- [Add to My List](#)

Help:

- [Ask a Librarian](#)
- [Item Not on Shelf?](#)
- [E-Resource Problem](#)
- [Recommend a Title](#)

This item

- Record View**
- [Staff View](#)

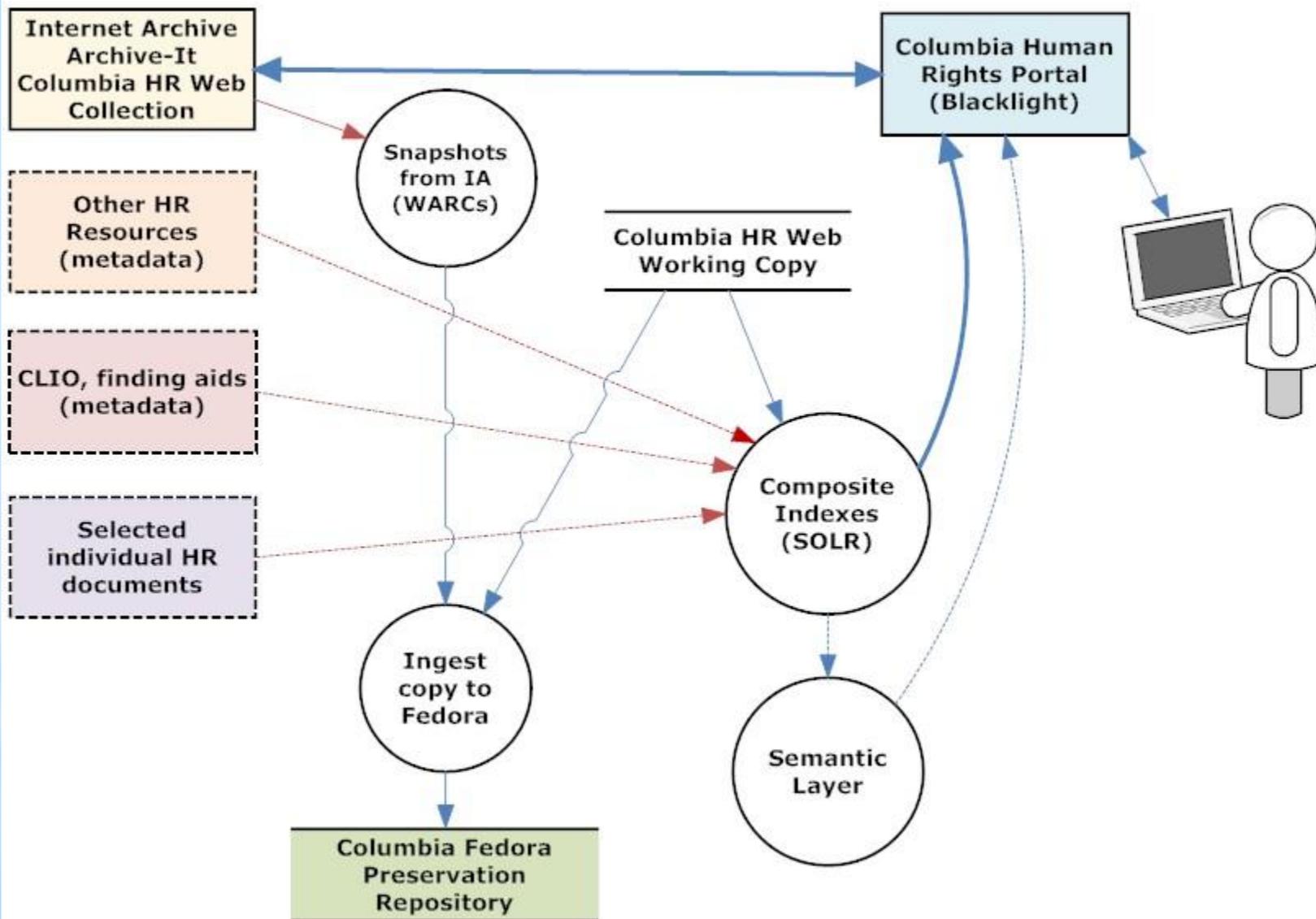
* CLIO

(Columbia University Libraries Online Catalog)

<i>Portal Field</i>	<i>MARC Tag</i>
Title	245
Creator	1XX/7XX
Organization Type	653 0
Organization Based In	"Country"
Geographic Focus	043
Subjects	650
Summary	520
Languages	"Language" and 041
Seed URLs	856 40
Archived URLs	856 41
Nonpublic portal identifier	"965hrportal"

* mapping MARC data for portal use

Columbia Human Rights Web Archive Portal (HRWA) Schematic Overview



Most challenging and innovative LDPD project to date.

- * Most data in single project (ca. 2 TB)
- * Largest indexes
- * Greatest number of servers for indexing / production
- * Most complex data (WARC / Web)
- * Most challenging end-user design requirements
- * Most uncharted in terms of users, possible uses, possible value added features, scoping, etc.
- * Most cutting edge, most unanswered tech questions

*** LDPD perspective on building local portal**

* Beta release for internal testing, 9/12-

* User testing, 1/13-

User groups to be tested include:

Human rights researchers (faculty, students)

Content providers (NGOs, activists)

Web archiving/library community

* **HRWA portal status**

- * Need to establish technical procedure for regular updating of portal CLIO metadata and WARC download from Archive-It
- * Determine if local indexing of data is sustainable
- * Improve quality of SOLR full-text search results for web archives
- * Determine if application of site-level metadata to individual documents is net benefit
- * Enhance full-text searching via Semantic web/Linked data layers?

* issues/
further development

HUMAN RIGHTS WEB ARCHIVE

BETA

CENTER FOR HUMAN RIGHTS DOCUMENTATION & RESEARCH AT COLUMBIA UNIVERSITY

[Home](#)
[LEARN MORE](#)
[NOMINATE A SITE](#)
[CONTACT](#)

[SEARCH HISTORY](#)

The Human Rights Web Archive @ Columbia University is a searchable collection of archived copies of human rights websites created by non-governmental organizations, national human rights institutions, tribunals and individuals. Collecting began in 2008 and has been ongoing for active websites. New websites are added to the collection regularly.

The HRWA is an initiative of the [Center for Human Rights Documentation & Research](#) and is a key focus of the Columbia University Libraries' [Web Resources Collection Program](#). The HRWA was made possible by generous support from the Andrew W. Mellon Foundation. [Learn More](#) »

FEATURED SITES

TITLES

URLS

SUBJECTS

PLACES

LANGUAGES

A selection of sites relating to the subject **Transitional justice**.



Featured Sites ▾



International Criminal Tribunal for the f...

Site Details ➤



Institute for Justice and Reconciliation

Site Details ➤



Hay'at al-Inṣāf wa-al-Muṣālaḥah

Site Details ➤

<http://hrwa.cul.columbia.edu/search/7005623>

* hrwa.cul.columbia.edu

FEATURED SITES

TITLES

URLS

SUBJECTS

PLACES

LANGUAGES

Browse websites by title.

A-Z

Z-A



"İnsan Hüquqları XXI əsr-Azərbaycan" Fondu

5-18 Kinyöm Chaedan

ABONG--Associação Brasileira de Organizações Não Governamentais

Abuelas de Plaza de Mayo

Access to Justice

Action Group for Health, Human Rights and HIV/AIDS (AGHA)

Ação Brasileira pela Nutrição e Direitos Humanos

Adalah

Human Rights Congress for Bangladesh Minorities

Human Rights Council of Australia

Human Rights Education Institute of Burma

Human Rights First

Human Rights First Society

Human Rights Law Network

Human Rights Press Point

Human Rights Trust of Southern Africa

Human Rights Watch

* browse by title

FEATURED SITES

TITLES

URLS

SUBJECTS

PLACES

LANGUAGES

Browse websites by URL.

A-Z

Z-A



<http://acda.co/>

<http://advocacyforum.org/>

<http://against-torture.net/>

<http://aldameer.org/>

<http://amigosdemujeres.org/>

<http://amnesty.ca/>

<http://amnesty.hu/>

<http://amnesty.ie/>

<http://amnesty.or.jp/>

<http://amnesty.org.uk/>

<http://www.dzeno.cz/>

<http://www.ecmi.de/>

<http://www.ecpat.net/>

<http://www.egjustice.org/>

<http://www.ehrco.org/>

<http://www.ehrea.org/>

<http://www.eisa.org.za/>

<http://www.elsam.or.id/>

<http://www.en.mosharka.org/>

<http://www.en.mosharka.org/>

* browse by URL

FEATURED SITES

TITLES

URLS

SUBJECTS

PLACES

LANGUAGES

Browse websites by subject.

#

A-Z

Z-A



Academic freedom 1

Adultery (Islamic law) 1

African Americans -- Civil rights 2

African diaspora 1

Africans -- Civil rights 1

Agricultural laborers 1

Agricultural laborers -- Social conditions 1

Agriculture and state 1

AIDS (Disease) 2

Islam and politics 1

Islamic law 1

Journalists -- Crimes against 1

Judges 1

Judicial power 1

Judicial process 1

Jumma (Bangladeshi people) -- Crimes against 1

Justice, Administration of 3

Kurdish diaspora 1

* browse by subject / alphabetical

FEATURED SITES TITLES URLS **SUBJECTS** PLACES LANGUAGES

Browse websites by subject. # A-Z Z-A 

Human rights (General) 425	Ethiopians -- Ethnic identity 1
Civil rights 67	Ethnic relations 1
Democracy 38	Exiles 1
Women's rights 29	Farmers -- Social conditions 1
Ombudspersons 22	Food sovereignty 1
Civil society 15	Forced migration 1
Transitional justice 15	Foreign workers 1
Indigenous peoples 14	Free speech 1
Torture 14	Freedom of association 1
Truth commissions 12	Freedom of movement 1

* browse by subject / number of sites

FEATURED SITES

TITLES

URLS

SUBJECTS

PLACES

LANGUAGES

Browse websites by geographic focus.

#

A-Z

Z-A



[Global focus] 49

Egypt 20

Israel 19

Gaza Strip 16

West Bank 16

Africa 14

Argentina 11

Mexico 11

Russia (Federation) 10

Latin America 9

France 2

Georgia (Republic) 2

Ghana 2

Greece 2

Hungary 2

Ireland 2

Italy 2

Jordan 2

Korea (South) 2

Kosovo 2

* browse by geographic focus -
number of sites

Showing item 1 of 1.

Centro de Derechos Humanos y Ambiente

Alternate Title	Center for Human Rights and Environment
Creator	Centro de Derechos Humanos y Ambiente
Organization Type	Non-governmental organizations
Organization Based In	Argentina
Geographic Focus	Latin America
Subjects	Human rights -- Environmental aspects ; Sustainable development ; Environmental protection ; Conservation of natural resources
Summary	The work of CEDHA focuses "on promoting greater access to justice and guarantee human rights for victims of environmental degradation, or due to non-sustainable management of natural resources, and to prevent future violations."
Languages	Spanish ; English



[View site captures for http://wp.cedha.net/](http://wp.cedha.net/)

[View site captures for http://www.cedha.net/](http://www.cedha.net/)

[View site captures for http://www.cedha.org.ar/](http://www.cedha.org.ar/)

[Email This](#)

* site details page

HRWA Impunity Sites

HUMAN RIGHTS WEB ARCHIVE

DOCUMENTATION & RESEARCH AT COLUMBIA UNIVERSITY

ACT SEARCH HISTORY

ED SITES TITLES URLS SUBJECTS PLACES LANGUAGES

of sites relating to the subject **Truth commissions.** Featured Sites

Search:

- Website descriptions
- Page full text

* search options

Displaying all 4 sites (0.004 seconds) for: **Impunity** 



Collectif des associations contre l'impu...

Network of organizations fighting **impunity** for human rights violations in Togo....

[Site Details](#)



Global Witness

"Global Witness exposes the corrupt exploitation of natural resources and international trade systems, to drive campaigns

[Site Details](#)



Comité de Familiares de Detenidos-De...

The Committee of Relatives of Missing Prisoners in Honduras in a non-governmental human rights organization founded in 1982. Its goal is

[Site Details](#)



Uganda Coalition on the International ...

"The Uganda Coalition for the International Criminal Court, initiated in 2004 by Human Rights Network – Uganda, is a local Ugandan

Narrow These Results

GEOGRAPHIC FOCUS

- [Honduras](#) (1)
- [Togo](#) (1)
- [Uganda](#) (1)
- [\[Global focus\]](#) (1)

LANGUAGE

- [English](#) (2)
- [French](#) (1)
- [Spanish](#) (1)

ORGANIZATION BASED IN

- [Honduras](#) (1)
- [Togo](#) (1)
- [Uganda](#) (1)
- [United Kingdom](#) (1)

ORGANIZATION TYPE

- [Non-governmental organizations](#) (3)
- [Other organization types](#) (1)

SUBJECT

- [Human rights \(General\)](#) (4)
- [Civil rights](#) (2)
- [Impunity](#) (2)
- [Conservation of natural resources](#) (1)
- [Corruption](#) (1)
- [more »](#)

* website descriptions (metadata)
search results

Displaying page 1 for about **1,500,000** archived web page results (0.682 seconds) for: **impunity** 

« Previous 1 2 3 4 ... Next »

Results per page ▾

Impunity

Jun 3, 2012 at 3:08AM UTC

http://www.safhr.org/index.php?option=com_content&view=section&id=9&Itemid=535

Impunity... **Impunity** Username Password Remember Me Forgot your password? Forgot your username? Login Back...

Size: 53.5 KB - HTML

Other Relevant Captures

[Mar 3, 2012 at 3:20AM UTC \(53.5 KB\)](#)

[Domain-](#) [Domain+](#)

Impunity

Jun 3, 2012 at 3:07AM UTC

http://www.safhr.org/index.php?option=com_content&view=section&id=9&Itemid=87

Impunity... **Impunity** Username Password Remember Me Forgot your password? Forgot your username? Login Back...

Size: 53.4 KB - HTML

Other Relevant Captures

[Mar 3, 2012 at 3:18AM UTC \(53.4 KB\)](#)

[Domain-](#) [Domain+](#)

Impunity

Jun 3, 2012 at 3:08AM UTC

http://www.safhr.org/index.php?option=com_content&view=section&id=9&Itemid=710

Impunity... **Impunity** Username Password Remember Me Forgot your password? Forgot your username? Login Back...

Size: 53.6 KB - HTML

Other Relevant Captures

[Mar 3, 2012 at 3:19AM UTC \(53.6 KB\)](#)

[Domain-](#) [Domain+](#)

Narrow These Results

DOMAIN

[cpj.org](#) (409,457)
[knchr.org](#) (263,867)
[humanrights.asia](#) (143,426)
[peacebrigades.org](#) (127,108)
[wluml.org](#) (35,471)
[more »](#)

DATE OF CAPTURE

[2011](#) (590,495)
[2010](#) (476,505)
[2012](#) (413,278)
[2008](#) (17,541)
[2009](#) (7,845)
[custom range »](#)

FILE TYPE

[HTML](#) (1,448,681)
[PDF](#) (50,021)
[Document](#) (4,129)
[XML](#) (2,756)
[Presentation](#) (54)
[Spreadsheet](#) (23)

GEOGRAPHIC FOCUS

[\[Global focus\]](#) (755,229)
[Kenya](#) (264,783)
[Asia](#) (173,496)
[Africa](#) (35,695)
[Mexico](#) (20,510)

* page full-text search results

Nominate Your Site for the Human Rights Web Archive

This form is for site owners who wish to nominate their organization or site for inclusion in the **Human Rights Web Archive**. If you are not the owner of the site you are nominating, please use the [public nomination form](#).

Please note that we cannot guarantee the inclusion of a nominated website in this program. We review the nominations carefully and assess the subject, technical, and copyright issues that may affect a site's suitability for archiving. Send questions and comments to See also our [FAQ](#).

Site Information

Name of
website/organization

URL of website

Description of
website/organization

Additional Information

Permissions

If we select your site for inclusion in the Human Rights Web Archive, will you give Columbia University Libraries/Information Services permission to crawl your site and provide access to the archived content in the Human Rights Web Archive?

* site owner nomination form

Send any portal feedback or questions to me at
at2186@columbia.edu

Or use site Contact form

Thanks!