

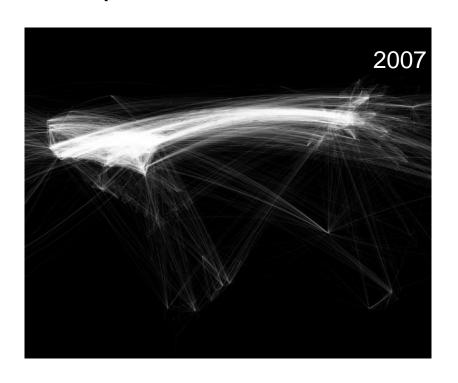
May 18, 2018

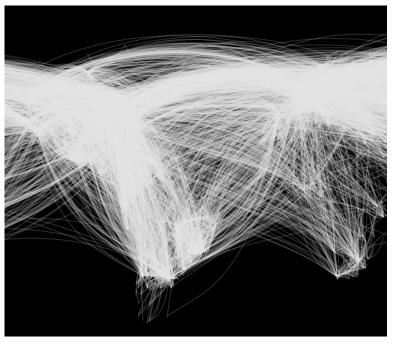
Harvesting Web Content and Data from Emerging Regions: a Status Report

Jeffrey Garrett Independent Consultant



Expansion of the Web in Latin America during 2007–2011





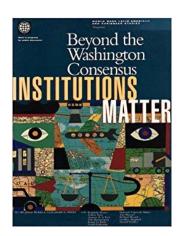
2011

Source: Worldwide City-to-City Internet Connections in 2007 (I.) and 2011 (r.). Courtesy of Chris Harrison, Carnegie-Mellon University.

Why the Web Matters—especially in International and Area Studies, even more especially in Latin American and Caribbean Studies

- Harrison's mapping "only reflects density of connections, and not usage hundreds of people may utilize a single connection in an internet cafe, often the only form of connectivity people have access to in developing nations."
- Of 3.9 billion Internet users in users in the world (as of June 30, 2017), 10.4% reside in Latin America and the Caribbean, while 8.2% are in North America. Almost half—49.7%—are in Asia.
- "... in Latin American Studies, ... much digital publishing is not channeled or distributed through commercial publishers but is instead only taking place on the freely accessible web ..."

Sources: Chris Harrison (2017); Internet World Stats: Usage and Population Statistics (2017); Graham & Norsworthy (2018)



Beyond the
Washington
Consensus: Institutions
Matter. Burki, Shahid
Javed, Perry, Guillermo
Washington, D.C.:
World Bank
Publications, 1998

Research & Citation Practice 20 Years Ago

Academy for Educational Development (AED) (1996). Educational Reform Projects Supported by the World Bank, the Inter-American Development Bank, and USAID in Latin America and the Caribbean. Washington, D.C.: Academy for Educational Development.

Akerlof, George (1970). "The Market for Lemons, Qualitative Uncertainty and the Market Mechanism." Quarterly Journal of Economics, 59:473-494.

Akerlof, G., and P. Romer (1993). "Looting: The Economic Underworld of Bankruptcy for Profit." Brookings Papers on Economic Activity, 2:1-73.

Alesina, Alberto, and Allan Drazen (1991). "Why Are Stabilizations Delayed?" The American Economic Review, 81(5):1170-1188.

Alexander, William E.; Jeffrey M. Davis; Liam P. Ebrill; and Carl-Johan Lindgren (1997). Systemic Bank Restructuring and Macroeconomic Policy. Washington, D.C.: International Monetary Fund.

Alvarez, Benjamin (1998). "Life Cycle and Legacy of the Educational Reforms in Latin America and the Caribbean." *International Journal of Educational Reform*, 7(1):34-45 (January).

Alvarez, Benjamin, and Monica Ruiz-Casares (eds.) (1997). Paths of Change: Education Reforms Under Way in Latin America and the Caribbean. Washington, D.C.: U.S. Agency for International Development/Agency for Educational Development.

American Bar Association (1989 and 1993). Multinational Commercial Insolvency. Wisconsin: MG Publishing.

Ames, Barry (1998). "Institutions and Democracy in Brazil." Mimeograph. Pittsburgh, Penn.: Department of Political Science, University of Pittsburgh.

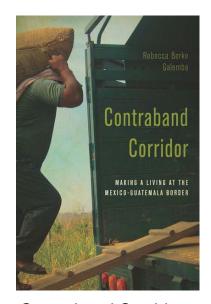
Angell, Alan (1996). "Improving the Quality and Equity of Education in Chile: The Program 900 Escuelas and the MECE-Básica." In *Implementing Policy Innovations in Latin America: Politics, Economics and Techniques*, edited by Antonia Silva. Washington, D.C.: Inter-American Development Bank Social Agenda Policy Group.

Arrow, Kenneth J. (1985). "The Economics of Agency." In J. Pratt and R. Zeckhauser, eds., Principals and Agents: The Structure of Business. Cambridge, Mass.: Harvard Business School Press.

Arruñada, Benito (1996). "The Economics of Notaries." European Journal of Law and Economics, 3(1):5-37.

Print and Online Monographs & Reports

Print & Online Journals



Contraband Corridor:
Making a Living at the
Mexico-Guatemala
Border. Galemba,
Rebecca B. Stanford,
Stanford University
Press, 2018

Research & Citation Practice Today

Asencio, Karen Mercado. April 12, 2012. "The Under-Registration of Births in Mexico:

Consequences for Children, Adults, and Migrants." *Migration Policy Institute*. Retrieved on August 12, 2016, from www.migrationpolicy.org/article/under-registration-births-mexico-consequences-children-adults-and-migrants.

Asturias, Miguel Ángel. 1993. Translated by Gerald Martin. *Men of Maize*. Critical Edition. Pittsburgh and London: University of Pittsburgh Press.

Azurdia, Mercedes. September 2, 2013. "New Free Trade Agreement with Mexico Takes Effect in Guatemala." *PanAm Post*. Retrieved on July 28, 2016, from https://panampost.com/mercedes-azurdia/2013/09/02/new-free-trade-agreement-with-mexico-takes-effect-in-guatemala/.

Bacon, David. 2008. Illegal People: How Globalization Creates Migration and Criminalizes

Immigrants. Boston: Beacon Press.

Bargent, James. February 4, 2016. "Mexico Impunity Levels Reach 99%: Study." *InSight Crime*.

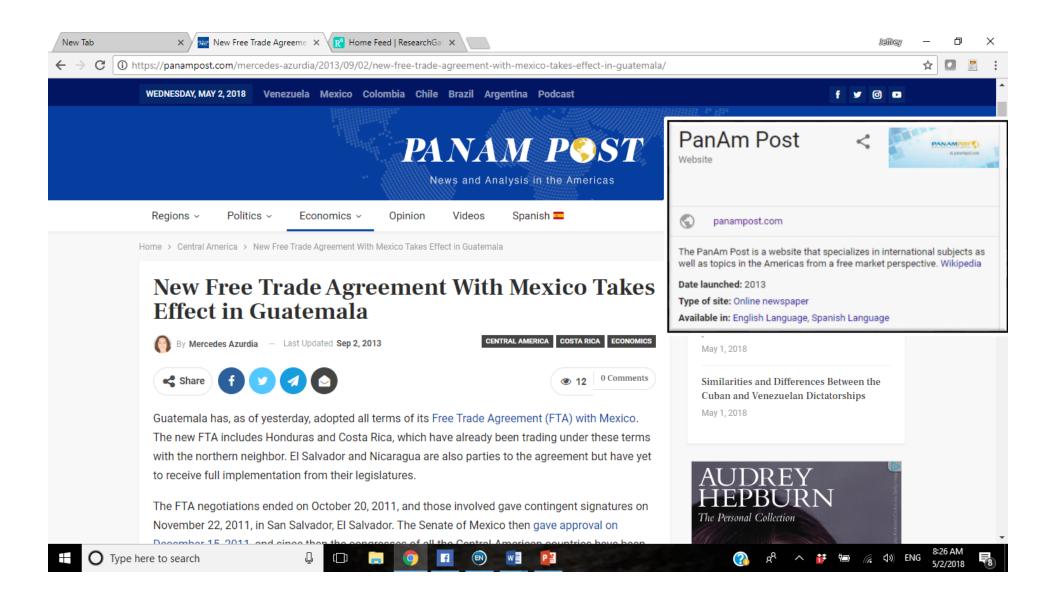
Retrieved on August 24, 2016, from www.insightcrime.com/component/content/article?

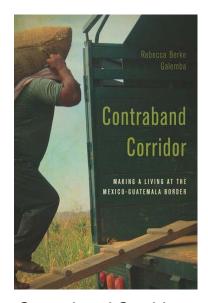
id=7447:mexico-impunity-levels-reach-99-study.

Barrientos, Ricardo. October 30, 2012. "Contrabando: expresión de poder real e impunidad." *Plaza Pública*. Retrieved on October 22, 2015, from www.plazapublica.com.gt/content/contrabando-

Print and Online Monographs & Reports

Open Access Web Content





Contraband Corridor:
Making a Living at the
Mexico-Guatemala
Border. Galemba,
Rebecca B. Stanford,
Stanford University
Press, 2018

Research & Citation Practice Today

Asencio, Karen Mercado. April 12, 2012. "The Under-Registration of Births in Mexico:

Consequences for Children, Adults, and Migrants." *Migration Policy Institute*. Retrieved on August 12, 2016, from www.migrationpolicy.org/article/under-registration-births-mexico-consequences-children-adults-and-migrants.

Asturias, Miguel Ángel. 1993. Translated by Gerald Martin. *Men of Maize*. Critical Edition. Pittsburgh and London: University of Pittsburgh Press.

Azurdia, Mercedes. September 2, 2013. "New Free Trade Agreement with Mexico Takes Effect in Guatemala." *PanAm Post*. Retrieved on July 28, 2016, from https://panampost.com/mercedes-azurdia/2013/09/02/new-free-trade-agreement-with-mexico-takes-effect-in-guatemala/.

Bacon, David. 2008. Illegal People: How Globalization Creates Migration and Criminalizes Immigrants. Boston: Beacon Press.

Bargent, James. February 4, 2016. "Mexico Impunity Levels Reach 99%: Study." *InSight Crime*.

Retrieved on August 24, 2016, from www.insightcrime.com/component/content/article?

id=7447:mexico-impunity-levels-reach-99-study.

Barrientos, Ricardo. October 30, 2012. "Contrabando: expresión de poder real e impunidad." *Plaza Pública*. Retrieved on October 22, 2015, from www.plazapublica.com.gt/content/contrabando-

Print and Online Monographs & Reports

Open Access Web Content

17 of the 120 Links to Web Content in *Contraband Corridor* (Stanford UP, 2018)

http://foreignpolicy.com/2016/02/23/obama-pena-nieto-mexico-corruption/

http://nofrackingmexico.org/nueva-ley-para-criminalizar-la-protesta-social-y-limitar-el-libre-flujo-de-informacion-en-el-marco-de-las-reformas-estructurales/

www.animalpolitico.com/2013/11/hacienda-cierra-12-garitas-aduanales-en-4-estados-fronterizos/

www.elfinanciero.com.mx/archivo/economia-ilegal-genera-perdidas-por-950-000-mdp.html

http://rightsaction.org/sites/default/files/Rpt_130220_Aguan_Final.pdf

https://nacla.org/news/mexico-abuses-against-us-bound-migrant-workers

www.prensalibre.com/economia/Conflicto-ingreso-maiz-blanco-Mexico 0 230976910.html

www.huffingtonpost.com/2011/05/23/10-countries-with-worst-income-inequality_n_865869.html

www.cipamericas.org/es/archives/15407

http://fpif.org/mexicos-oil-privatization-risky-business/

www.cipamericas.org/archives/1834

www.migrationpolicy.org/article/mexico-caught-between-united-states-and-central-america

<u>www.prensalibre.com/pl/2006/septiembre/07/lectura_dept.html#151072</u> (link no longer available at time of publication)

https://loschapincitos.wordpress.com/2013/07/21/en-guatemala-cae-hermano-de-jefe-narco-que-ordeno-la-matanza-de-9-policias/

<u>www.worldcrunch.com/business-finance/the-threat-of-mexico-0-s-massive-undergound-economy/c2s16698/www.conasami.gob.mx/formatestimonios.aspx</u> (link no longer available at time of publication) www.telesurtv.net/english/news/Informal-Economy-Makes-Up-26-of-Mexicos-GDP-20140808-0044.html

Advice from Style Manuals I: MLA Handbook

- "While URLs define where online material is located, they have several disadvantages: they can't be clicked on in print, they clutter the works-cited list, and they tend to become rapidly obsolete."
- "Even an outdated URL can be useful, however, since it provides readers with information about where the work was once found."

MLA Handbook, 8th edition (2016), p. 48

MLA Handbook, 8th ed.: Criticism

- This 8th ed. of the MLA Handbook, published in 2016, shows no awareness of the existence of archival Web content.
- Its advice is to depart from an author's responsibility to the reader: to research and document one's sources.
- Jill Lepore: "The footnote, a landmark in the history of civilization, took centuries to invent and to spread. It has taken mere years nearly to destroy." (2015)

Advice from Style Manuals II: Chicago 17th ed.

- "If a site ceases to exist before publication, or if the information cited is modified or deleted, this information should be included in the text or note. . . . Such dates, together with the URL, give interested readers a chance to find the information through the Internet Archive or other means." (2017, 14.207)
- I.e, no recommendation to cite the archive as the actual source or as a backup and more persistent source.

17 of the 120 Links to Web Content in *Contraband Corridor* (Stanford UP, 2018)

http://foreignpolicy.com/2016/02/23/obama-pena-nieto-mexico-corruption/

http://nofrackingmexico.org/nueva-ley-para-criminalizar-la-protesta-social-y-limitar-el-libre-flujo-de-informacion-en-el-marco-de-las-reformas-estructurales/

www.animalpolitico.com/2013/11/hacienda-cierra-12-garitas-aduanales-en-4-estados-fronterizos/

www.elfinanciero.com.mx/archivo/economia-ilegal-genera-perdidas-por-950-000-mdp.html

http://rightsaction.org/sites/default/files/Rpt_130220_Aguan_Final.pdf

https://nacla.org/news/mexico-abuses-against-us-bound-migrant-workers

www.prensalibre.com/economia/Conflicto-ingreso-maiz-blanco-Mexico 0 230976910.html

www.huffingtonpost.com/2011/05/23/10-countries-with-worst-income-inequality_n_865869.html

www.cipamericas.org/es/archives/15407

http://fpif.org/mexicos-oil-privatization-risky-business/

www.cipamericas.org/archives/1834

www.migrationpolicy.org/article/mexico-caught-between-united-states-and-central-america

www.prensalibre.com/pl/2006/septiembre/07/lectura_dept.html#151072 (link no longer available at time of publication)

https://loschapincitos.wordpress.com/2013/07/21/en-guatemala-cae-hermano-de-jefe-narco-que-ordeno-la-matanza-de-9-policias/

<u>www.worldcrunch.com/business-finance/the-threat-of-mexico-0-s-massive-undergound-economy/c2s16698/www.conasami.gob.mx/formatestimonios.aspx</u> (link no longer available at time of publication)

www.telesurtv.net/english/news/Informal-Economy-Makes-Up-26-of-Mexicos-GDP-20140808-0044.html

Not secure https://www.cipamericas.org/es/archives/15407



Your connection is not private

Attackers might be trying to steal your information from **www.cipamericas.org** (for example, passwords, messages, or credit cards). <u>Learn more</u>

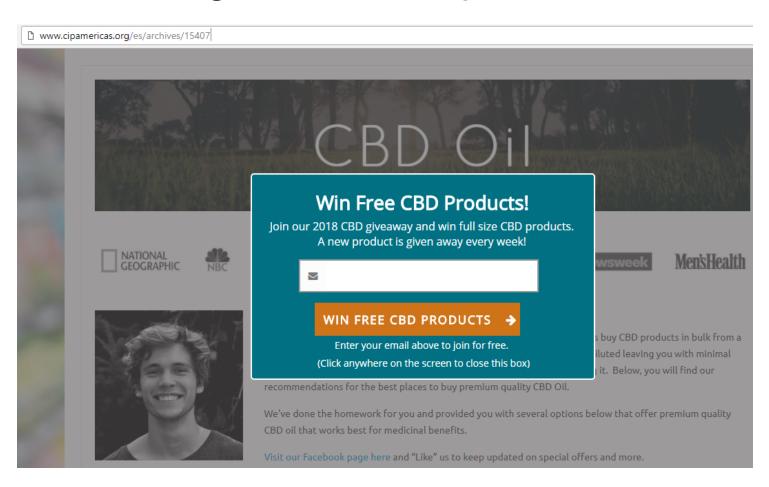
NET::ERR_CERT_COMMON_NAME_INVALID

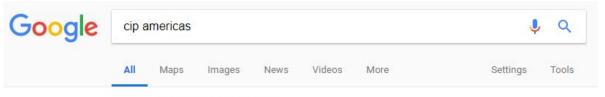
Automatically send some <u>system information and page content</u> to Google to help detect dangerous apps and sites. <u>Privacy policy</u>

ADVANCED

Back to safety

Surprise! "cipamericas.org/es/archives/" is now a site advertising cannabis oil products . . .





About 659,000 results (0.43 seconds)

The Americas Program | Programs | Center for International Policy

https://www.ciponline.org/programs/americas-program •

With more than 35 years of experience in Latin American news and analysis, the **Americas** Program of the Center for International Policy is a leading source of information for activists, academics and citizens concerned about US foreign policy in Latin **America**, human rights and movements for social justice within the ...

You've visited this page 2 times. Last visit: 2/8/18

| The Americas Program

www.americas.org/ -

Radiography of the VIII Summit of the Americas in the Trump Era. The era of Trump marks a change that replaces the construction of hemispheric agendas with mere demands that other nations back up the U.S. government in its war policies, and an attitude of intolerance and aggression toward political situations that ...

CIP Americas Program - Home | Facebook

https://www.facebook.com > ... > Broadcasting & Media Production Company ▼

CIP Americas Program, Mexico City. 2.4K likes. The Americas Program seeks to advance a new world of citizen action, analysis, and policy options.

CBD Oil » PremOffers

www.cipamericas.org/archives/1009

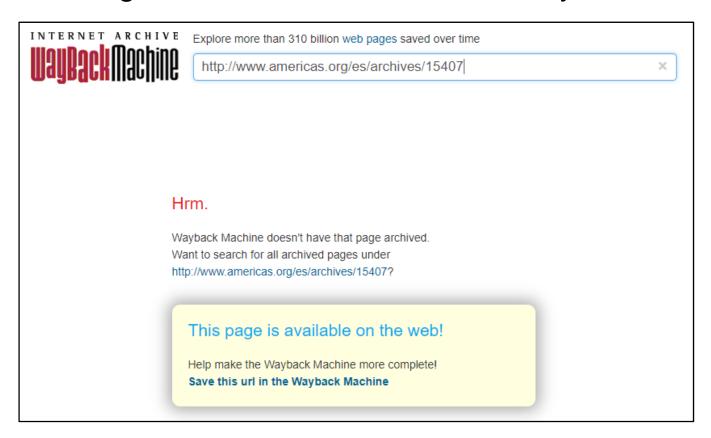
Think twice before buying CBD products from local stores. Many local stores buy CBD products in bulk from a variety of wholesalers in foreign countries that are poor quality and highly diluted leaving you

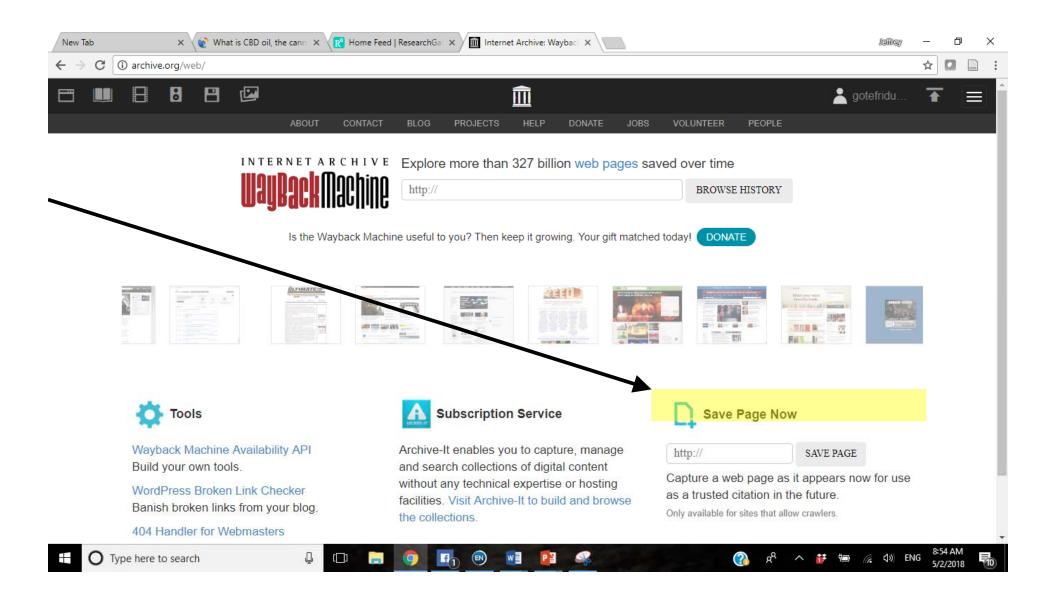
So I called the CIP (Center for International Policy) in Washington, D.C., and asked where their content had gone and what they were doing to get it (and their "archival" website) back . . .

Content Wrangling in the Wild West of the Web

- CIP said that after their site had been hijacked/corrupted, they had abandoned "cipamericas.org" and moved their archive along with other compromised subdomains to "americas.org".
- As documented by a review of crawls of this domain performed by the Internet Archive, "americas.org" had been given up in 2007 by the Resource Center of the Americas in Minneapolis, then used briefly by La Conexión de las Américas before taken over by CIP.
- I found the referenced article at the new location.
- I notified the author at the University of Denver. She was not happy—but what could she do? The book had already been published.
- I decided to archive the content at its new location . . .

Self-Archiving Your Sources Is Fun and Easy!









Programa de las Américas

Un nuevo mundo de acción y comunicación para el cambio social

INICIO TEMAS REGIONES BLOG NOSOTROS COLABORADORES CONTACTO



añol

Successfully self-archived!

Deportación, detención ilegal y abusos en la frontera de EEUU con Guatemala

By Laura Carlsen | 7/julio/2015

Tweet



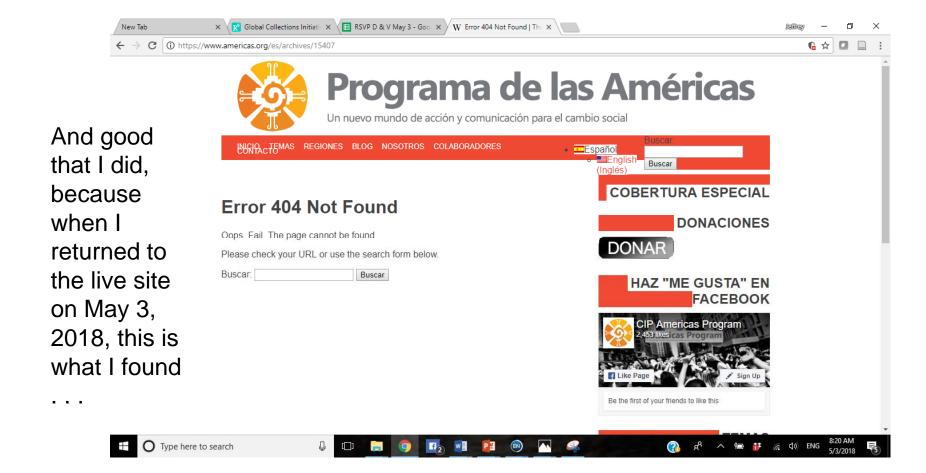
La frontera sur de México se ha vuelto la linea de contención del país más poderoso del mundo. Las víctimas de esta política extraterritorial son las personas migrantes de Centroamérica que cruzan cada día, buscando salvar su vida de la violencia y el hambre que azotan sus países.

Desde hace años, y de manera intensificada en el último año, el gobierno de los Estados Unidos ha reclamado la frontera sur de COBERTURA ESPECIAL

DONACIONES

HAZ "ME GUSTA" EN FACEBOOK

TEMAS



So . . .

- ... as an author or a student ... why not just cite the stable, archived URL in the first place??
- And if there isn't one, why not make one yourself?? (Go ahead, be your own archivist!)
- It could save your readers inconvenience . . . and yourself embarrassment!
- Why aren't style manuals, why aren't we giving this advice to our communities??

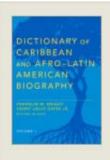
The Way It Could Be Done (I)

Ellinic una /vigration Studies 31 (2003): //4-/94.

- 12 "Tunnels beneath U.S. border proliferate," SECRECY NEWS_from the FAS Project on Government Secrecy, http://www.fas.org/sgp/news/secrecy/2009/02/022309.html, accessed January 20, 2011.
- 13 Vicente Fox, "Mensaje de Vicente Fox, durante la Sesión Solemne del H. Congreso del Unión," December 1, 2000, http://wayback.archive-it.org/176/20060818191623/http://www.presidencia.gob.mx/actividades/? contenido=6&pagina=377, accessed January 20, 2011.
- 14 Ana Arana, "How the Street Gangs took Central America," Foreign Affairs 84 (May/June 2005): 98 Academic OneFile. Web. 20 January 2011.
- 15 James Brooke, "Fighting in Panama: Latin America; U.S. Denounced by Nations Touchy About Intervention," New York Times, December 21, 1989.
- 16 Douglas W. Payne. "Why Drug Traffickers Love Free Trade." Dissent 44 (Summer 1997), 60.

An explicit link to LAGDA content (University of Texas at Austin) in a recent monograph (Williams 2012, 140), retrieved by searching for instances of http://wayback.archive-it.org/176/ in Google Books.

The Way It Could Be Done (II)



Dictionary of Caribbean and Afro-Latin American Biography

Edited by Franklin W. Knight and Henry Louis Gates, Jr.

Publisher: Oxford University Press Print ISBN-13: 9780199935796 Current Online Version: 2016 Print Publication Date: 2016 Published online: 2016

Bibliography

Chadbrand, Ian, "Shelly-Ann Fraser's Rise from Poverty to One of the World's Best Sprinters is Remarkable." *The Telegraph*, 29 October 2009.

Find this resource:

Dania, Bogle. "'It Was Her Smile,' Says Jason Pryce." *Jamaica Observer*, 10 September 2012. http://www.jamaicaobserver.com/sport/-lt-was-her-smile----says-Jason-Pryce_12486130.

Find this resource:

Ince, Basil A. Black Meteors: The Caribbean in International Track and Field. Kingston, Jamaica: Ian Randle, 2012.

Find this resource:

Rayon, Kayor. "Fraser Expects Great Results in 100 Metres." *Jamaican Observer*, 13 August 2008. https://web.archive.org/web/20090506015347/http://www.jamaicaobserver.com/sports/html/20080813T000000-0500_139009_obs_fraser_expects_great_results_in_____metres.asp.

Find this resource:

"Shelly-Ann Fraser-Pryce." IAAF Athletics. http://www.iaaf.org/athletes/jamaica/shelly-ann-fraser-pryce-198930.

"Shelly-Ann Fraser-Pryce, Track and Field Athlete (1986—)." Biography.com. http://www.biography.com/people/shelly-ann-fraser-pryce-20830469.

Fraser-Pryce, Shelly-Ann (1986–), track a

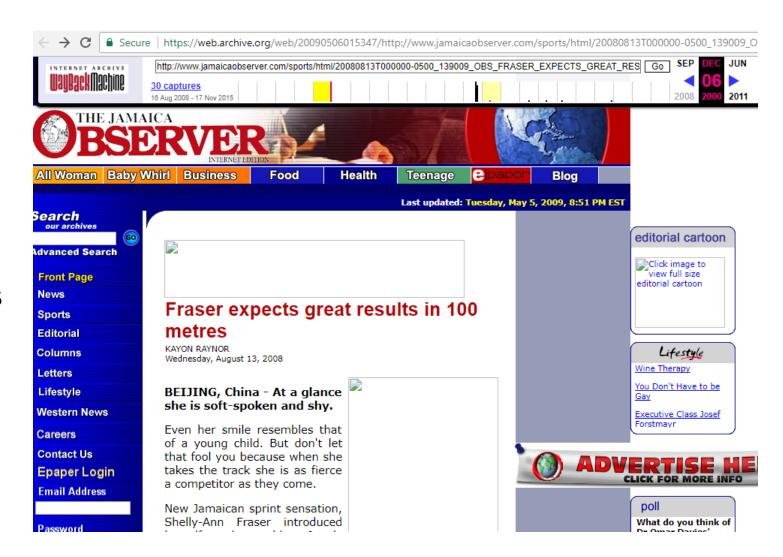
world champion, a

was born in Kingston, Jamaica, on 27 December 1986. Shelly-Ann room tenement house in Waterhouse, an impoverished, overcrowd mother, Maxine Simpson, worked as a street vendor, while her fath family, had very little to do with them.

First participating in track and field at George Headley Primary Sch
the Primary School Championships at the age of 10. After completing primary school, sine qualified schools school school schools at the age of 10.

Reliably archived, but note missing content!

New tools allow the capture of more content types.



The Way It Could Be Done (III)

Senior Honors Thesis, University of Maryland (2013), citing content from the Latin American Government Documents Archive (LAGDA) at the University of Texas at Austin

scholar Yaakov Ariel says that Latin American Christian Zionism is part of a global movement that is growing in new places like Korea and joining places traditionally supportive of Christian Zionism like the United States.³⁶

- "UMD librarian Pat Herron was very helpful in this process. Her demonstrations in our LASC library session introduced us to all the relevant databases, which would have materials from research to newspaper articles to how to locate an archive full of government documents from all around Latin America."
- "At the end of this process, I very much feel like a historian who is able to carry out independent research, and feeling like I really have those skills is a great end to my undergraduate career at UMD."

Sandra A. Shaker, Student at UMD (2013)

³² Cánovas, Literatura de inmigrantes árabes y judíos, 164.

^{33 &}quot;Solid Latam support for Palestine"

³⁴ "Declaration of the Government of Chile regarding the recognition of the State of Palestine," Government of Chile, January 7, 2011, accessed September 24, 2013, http://www.gob.cl/english/government-information/2011/01/07/declaration-by-the-government-of-chile-recognition-of-the-state-of-palestine.htm

³⁵ Chris Arsenault, "Evangelicals rise in Latin America," Al Jazeera English, March 26, 2013, accessed November 19, 2013, http://www.aljazeera.com/indepth/features/2012/03/201232593459332334.html.

³⁶ Yaakov Ariel, "An Unexpected Alliance: Christian Zionism and its Historical Significance," *Modern Judaism* 26.1 (2006): 81, accessed November 19, 2013, http://muse.jhu.edu/journals/modern_judaism/v026/26,1ariel.html.

What To Look Forward To . .

AN EVALUATION OF WEB ARCHIVING PROGRAMS IN THE US RELEVANT TO INTERNATIONAL AND AREA STUDIES

The Example of Latin America and the Caribbean

Contents

		luation of Web Archiving Programs in the US Relevant to International and Area Studies	
ne		ample of Latin America and the Caribbean	
		troduction: Latin America Takes to the Web	
I.	Со	enducting Research in the Live and the Past Web of Latin America: A Hypothetical Example	!
II.		Archiving the Latin American & Caribbean Web	1
Α		Web Archiving at the Library of Congress	1
	1.	A Brief History	1
	2.	Governance and Selection	1
	3.	Support and Collaboration	1
	4.	Metadata and Search	1
	5.	Inhouse Use Analysis and User Feedback	
	6.	Challenges, Future Plans, and Hopes	
В		The Human Rights Web Archive at Columbia University	
	1.	A Brief History	
	2.	Governance and Selection	
	3.	Support and Collaboration	
	4.	Metadata and Search	
	5.	Inhouse Use Analysis and User Feedback	
	6.	Challenges, Future Plans, and Hopes	
С		Web Archiving at the University of Texas at Austin: LAGDA and HRDI	
	1.	A Brief History	
	2.	Governance and Selection	
	3.	Support and Collaboration	
	4.	Metadata and Search	
	5.	Inhouse Use Analysis and User Feedback	
		Challenges Suture Plans and Hones	

Parts I, II, and III:

An introduction to web archiving as it relates to IAS—especially to Latin America & the Caribbean . . .

. . . and as practiced at 3 leading US programs:

Library of Congress Columbia University University of Texas

Part IV: An Independent Use Analysis

IV. An Independent Use Analysis

Use analysis that goes beyond the anecdotal—in other words, that is based on harvested occurrences of unique and explicit URL stems in large databases—is problematic for several reasons. For one, researchers who cite websites may choose to cite original, live-web URLs even if they used the archived content. Live web URLs are shorter, meaning less "clutter [in] the works-cited list" (MLA Handbook, 2016, 48), and authors may assume (or just hope) that readers corroborating research findings will go to the Internet Archive or to specific institutional archives if they can't locate/relocate the original content. Contributing to this practice may be that the standard style guides (Chicago, MLA, APA, others) are either silent or equivocating on how or even whether to reference archived web content (Davis 2016), reflecting either a lack of awareness of the existence of these archives or lacking community consensus on how to treat them in scholarly work. A further complication stems from the fact that most Archive-It content is also loaded into the *public* Internet Archive (archive-org) and these URLs are not explicitly traceable to the collecting institutional Archive-It participant. For example,

Part V: Accelerating the Integration &

Mainstreaming of IAS-Relevant Web Archives

- 1. Greater Interinstitutional Collaboration
- 2. "Postcustodial" Collaboration with Partners Abroad
- 3. "Desiloization" within Libraries, Archives, & Communities Served
- 4. Metadata Standardization
- 5. Improved Citation Standards & Tools
- 6. Development and Promotion of Self-Archiving Tools
- 7. Credibility Enhancement through Clearer Scoping & Certification
- Outreach to Publishers, Associations, Funding Bodies
- 9. Ramp Up Outreach on Campus through Public Services
- 10. Promoting Opportunities for Whole-Collection Analyses & Datamining

In Closing: The Problems Are Not So Much About Technology Anymore

"We can assert that the primary obstacles to expanding [web archiving] activities in libraries are less on the "technology" side and more on the 'cultural' side."

(Graham & Norsworthy, 2018)

• "I . . . hope that we'll develop highly effective and functional ways of discovering and using the web archives—that some of the high hurdles that currently exist will be diminished or eliminated . . ."

(Pamela Graham, personal communication, 2018)