Latin American and Iberian Studies
Collection Development in the Age of Blogging:

Identifying, Collecting and Preserving Literary Blogs and Other Cyberliterature

Melissa Gasparotto
Librarian for African and Latin American Studies
Rutgers University Libraries
Who is blogging in the Hispanic world?

- The average Spanish-language blogger is a male between the ages of 25 and 34, and he is most likely to live in Spain, Argentina or Mexico
  - 52% of Spanish-language bloggers are from Spain
- In the United States, 20% of all bloggers are English-speaking Hispanics.
  - This 20% participation in blogosphere is particularly notable given the fact that Hispanics represent only 15% of the US population
What are they blogging about?

- Informe sobre el estado de la blogosfera hispana Bitacoras.com 2010
And why should this matter to librarians?

• Bloggers are citing resources we have in the library:

Medios más enlazados desde los blogs

• Our users are reading blogs, and they turn to blogs for a wide variety of information:

Términos de búsqueda -- Bitacoras.com

• Blogs, no matter what their thematic content, are the new ephemera and grey literature, and will prove indispensable as primary and secondary sources for future scholars.
Literary Blogs and Other Cyberliterature

- Serialized fiction: La blogonovela
- Microfiction/Flashfiction
- Collaborative literary projects utilizing comment threads
- Hypernovel
- Wikinovel
- E-poetics

Literatura Electrónica Hispánica
Portal from Biblioteca Virtual Miguel de Cervantes
The need for preservation

- According to the III Encuesta a Bloggers Hispanos, a full 72% of bloggers use free web hosting.
- Which is exactly as frightening as it sounds...
But What About Internet Archive?

- In my own informal survey of the top blogs on the cultural list on Bitacoras.com, the bulk had never been captured by Internet Archive, or hadn't been captured since 2008!!!
How does one go about selecting blogs to preserve?

- Blog ranking sites such as Bitacoras.com
- Visible collections of blogs, such as Boomeran(g)
- References and linkrolls on sites that are already known
- Must be a continually updated project, since literature appears and disappears so quickly
Existing Blog Harvesting and Preservation Projects

• Human Rights Web Archive, from the Center for Human Rights Documentation and Research at Columbia University
  • http://www.columbia.edu/cu/lweb/indiv/humanrights/hrwa/index.html
  • Mellon-funded project to preserve freely available human rights material on the web, primarily produced by NGOs, individuals and governmental initiatives
  • Uses Archive-It
  • Snapshots taken periodically
  • Does not capture outgoing links

- Web Archiving Section, Library of Congress
  http://www.digitalpreservation.gov/
  - Thematic pilot projects
  - Harvested by Internet Archive
  - Collection-level MARC record added to LC Catalog and WorldCat
  - Technical Background and Guidelines:
    http://www.loc.gov/webarchiving/technical.html
  - Part of International Internet Preservation Consortium
    http://netpreserve.org/about/index.php
LC Web Preservation, cont.

- Pilot project for blog preservation

- Legal Blawgs Project

"A selective collection of authoritative blogs associated with American Bar Association approved law schools, research institutes, think tanks, and other expertise-based organizations, containing journal-style entries, articles and essays, discussions, and comments on emerging legal issues, national and international."

Collection dates: March 2007 - 2009

- Legal Blawgs Web Archive
Existing Services for Archiving Blog Content

• Archive-IT
  – http://www.archive-it.org

• California Digital Library Web Archiving Service
  – http://webarchives.cdlib.org/was

• Web Curator Tool
  – http://webcurator.sourceforge.net/
  – Open source project
Sources

- "Informe sobre el estado de la blogosfera hispana Bitacoras.com" (2010) http://bitacoras.com/informe/10