

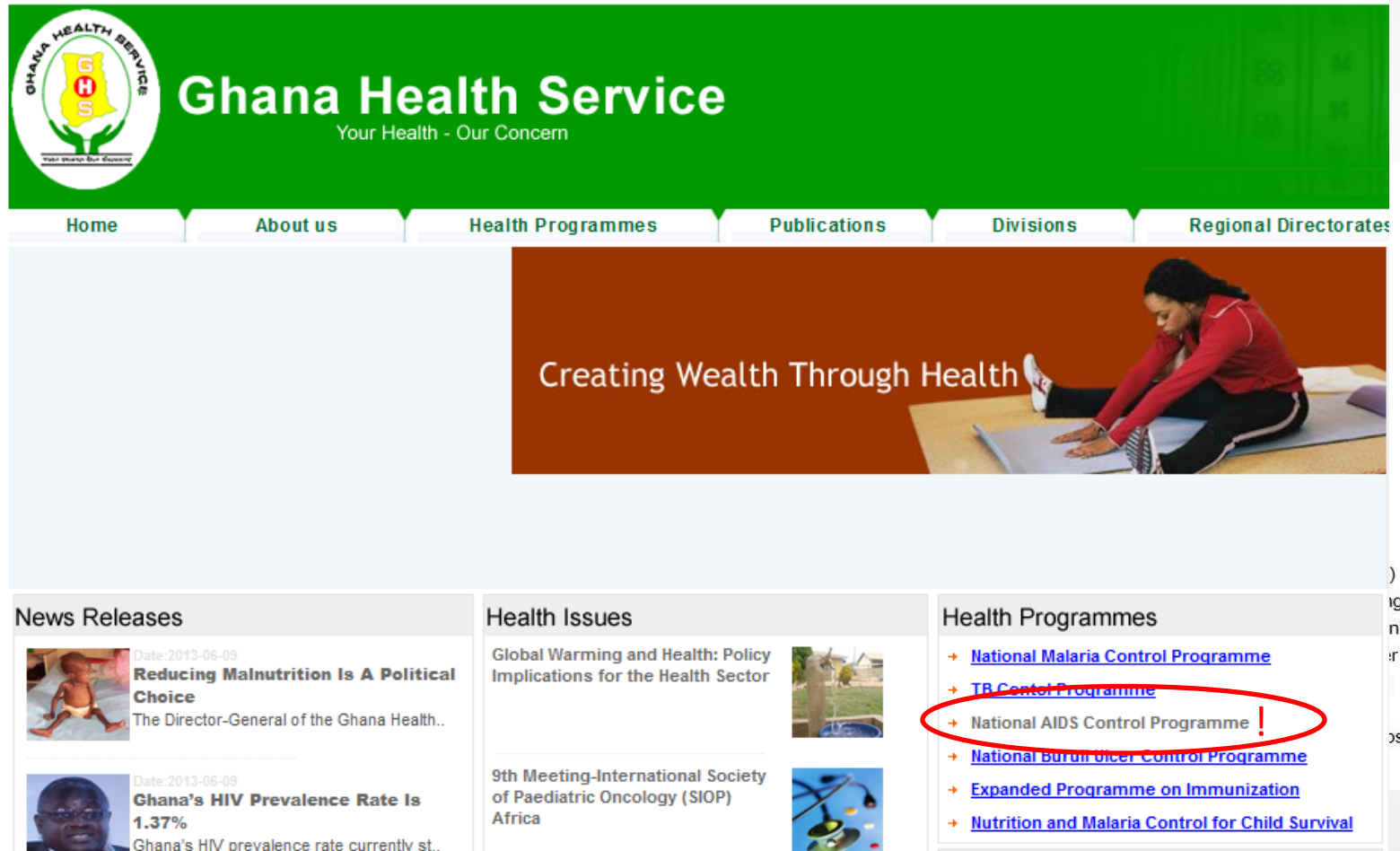
# Strategies for unearthing foreign government documents on the web

Jeremy Darrington, Politics Librarian

Princeton University Library

Documents Association of New Jersey (DANJ) 2013 annual conference

# The Problem



The screenshot shows the Ghana Health Service website. The header features the Ghana Health Service logo and the text "Ghana Health Service" with the tagline "Your Health - Our Concern". Below the header is a navigation menu with links for Home, About us, Health Programmes, Publications, Divisions, and Regional Directorates. A large banner image shows a woman in a red shirt stretching on a mat, with the text "Creating Wealth Through Health". Below the banner are three columns of content: News Releases, Health Issues, and Health Programmes.

**Ghana Health Service**  
Your Health - Our Concern

Home About us Health Programmes Publications Divisions Regional Directorates

Creating Wealth Through Health

**News Releases**

- Date: 2013-06-09  
**Reducing Malnutrition Is A Political Choice**  
The Director-General of the Ghana Health..
- Date: 2013-06-09  
**Ghana's HIV Prevalence Rate Is 1.37%**  
Ghana's HIV prevalence rate currently st..

**Health Issues**

- Global Warming and Health: Policy Implications for the Health Sector
- 9th Meeting-International Society of Paediatric Oncology (SIOP) Africa

**Health Programmes**

- [National Malaria Control Programme](#)
- [TB Control Programme](#)
- [National AIDS Control Programme !](#)
- [National Buruli Ulcer Control Programme](#)
- [Expanded Programme on Immunization](#)
- [Nutrition and Malaria Control for Child Survival](#)

# Caveats

- Look other places first
- These strategies have a high failure rate

# Try translation

- Use Google Translate to navigate through the native interface to potentially find docs, stats, etc.
- Ex:  
[http://verkiezingen2007.belgium.be/nl/cha/results/results\\_talb\\_etop.html](http://verkiezingen2007.belgium.be/nl/cha/results/results_talb_etop.html)

# Beware your browser

- Many foreign government sites only work properly in old IE (particularly Chinese sites)
- While some will only work in standards-compliant browsers (i.e., not IE)
- Ex: <http://www.stats.gov.cn/english/>

# Gain from Google

- Use **site:** to search Google's index of that particular domain or subdomain (e.g., stats.gov.cn)
- Use **filetype:** to search for particular document formats (**pdf** and **xls** are your best bets, no preceding dot)
- Use them in FF's search box or Chrome's omnibox.
  - Note: there is no space between the operator and the value, e.g., site:stats.gov.cn filetype:pdf
- Ex:  
<http://www.stats.gov.cn/tjsj/ndsj/2012/indexeh.htm>

# Access the Archive

- Wayback Machine at: <https://archive.org/>
- You need a website to start with
- Underlying interactive databases aren't typically captured, but documents (esp. PDFs) are.
- Website directory structure frequently changes in subtle ways from year to year, so when in doubt, go up to the top level domain and then navigate back down
- Ex: <http://www.electionethiopia.org/en/home.html>

# Persist patiently

- Think about which agency/ministry/ commission might produce the needed documents/info
  - Consider statistical agencies
  - Look at IGOs or NGOs for clues as to which agency may produce the document
- Go beyond the first page (or 2 or 3) of results
- Ex: <http://www.ghana.gov.gh/>

# Links to tools I used

- For Firefox:
  - Fire IE: <https://addons.mozilla.org/en-US/firefox/addon/fire-ie/>
  - Quick Translator: <https://addons.mozilla.org/en-US/firefox/addon/quick-translator/>
- For Chrome:
  - IE Tab: <https://chrome.google.com/webstore/detail/ie-tab/hehjbfgiekmjfkfjpbkbammjbdenadd?hl=en-US>