Indexing Issues for the Blogosphere:
The Case of Nofollow

Authors

Anne Helmond (Netherlands)

Rosa Menkman (Netherlands)

Michael Stevenson (USA)

Clifford Tatum (USA)

Esther Weltevrede (Netherlands)

Research Question

- What's a blog, what is the blogosphere and where is it?

- To what extent do search engines not only map the blogosphere but also construct it?

rel="nofollow"

- What is nofollow?

- Why is it important?

  - It's visible (unlike other indexing decisions)

  - It's an intervention by the software makers and the search engines and not the blogger.

  - It has an impact on the meaning and value of links

  - It influences ranking and indexing (different per search engine)

Case Study: Masters of Media

• Nofollow is one of many decisions indexers make, most of which we cannot see. But we can measure their sum effects.

• Google vs. Google Blogsearch vs. Technorati for our own blog

Method

- Indexing the indexers

  - Query link:mastersofmedia.hum.uva.nl in each index

  - Scrape search results using different DMI tools

- Compare and visualize results

Cross-device overlap backlinks MoM

Sources: www.google.com; blogsearch.google.com; technorati.com
Query: "link:http://mastersofmedia.hum.uva.nl"
Tool: Analyse/Compare lists
Results: All (unique hosts)
Date: 30/07/07

Technorati

Google

Google Blog Search

sargasso.nl   cityofsound.com
oearta.ie   technorati.com   aboutblank.nl   cocomment.com
astro.temple.edu   metaverse.jp   kenspeckle.net   blog.imessengr.com
en.wikipedia.org/wiki/Short_message_service   particlestream.motime.com
entregeeks.net   browserden.co.uk   swissestates.com
joepvandelft.net   feeds.technorati.com   surveillance-studies.org
greatestjournal.com   edugame.eu   omniture.com
webopedia.com   egaminglot.com   365questions.org
omdeziel.web-log.nl   crowdsourcing.typepad.com
unexpectedmedia.com   direct.shots.snap.com
friesian.com   creativecommons.org
romantoi.wordpress.com   wordpress.org
geocities.com/phersea   d-avide.it
ro-m-an.hyves.nl   blog.outer-court.com
socioomc.blogspot.com
soph72.wordpress.com
tomasgonsorcik.com
simplyappalling.blogspot.com
wildblueskies.com
textingvstalking.blogspot.com
placeblogger.com
gamespace.nl
ilyagram.org
gmpg.org
jeffhung.net
w3.org

eatmyfuckingstilettos.wordpress.com
polyglotconspiracy.net

kariahintikka.fi
komunikacii.net
c21org.typepad.com
deuxlits.wordpress.com
ideant.typepad.com
netzmedium.de   somelab.fi
higher-yearning.org   sethf.com
criticaldesign.wordpress.com
questiontechnology.org
academic-wasteland.blogspot.com
honorsinamsterdam2007.blogspot.com
newmedia.mediaonestudios.com
xer-files.blogspot.com

networkcultures.org

peppie.wordpress.com

mastersofmedia.hum.uva.nl

newmw.wordpress.com

annehelmond.nl

blog.waag.org
wordpress.com
dialogic.blogspot.com
jeffvansky.wordpress.com
internetundpolitik.wordpress.com
whatsnewmedia.wordpress.com
romantoi.wordpress.com
astore.amazon.com
craigbellamy.net
mediastudies.nl
edugame.eu

nyu.edu
flickr.com   rosa-menkman.blogspot.com
onahill.blogspot.com   technical-itch.co.uk
infocult.typepad.com