Mapping the Australian Political Blogosphere

Dr Axel Bruns, Creative Industries Faculty, Queensland University of Technology

Lars Kirchhoff, Thomas Nicolai, Sociomantic Labs
Berlin, Germany – lars.kirchhoff, thomas.nicolai@sociomantic.com – http://sociomantic.com/

1. Introduction
Crawling, scraping, visualising Australian political blogs:
- Based on growing list of blog feeds
- Custom-built content scraper
- Distinction between post content and other material

Timeframe:
- November 2007 to January 2008 – ~250 feeds
  (for Australian federal election on 24 November 2007)
- Second stage ongoing since January 2008

2. Method
- Identify blogs, subscribe to RSS feeds
- Follow RSS feeds, identify new posts
- Scrape entire blog page
- Isolate salient content – blog post only (plus comments?)
- Extract post text and links
- Map topics and link networks

Advantages over previous approached:
- Better data quality: excluding blog headers / footers / sidebars / ads / blogrolls
- More than links: capturing content for automatic and manual analysis
- Not retrospective: tracking and capturing activity as it happens
- Diachronic, diatopic: comparing activity across timeframes and blogs / blog clusters

3. Results
- Clear distinctions in topical focus of specific blogs
- Moderate clustering tendencies
- Links from mainstream online media very rare
- Mainstream media blogger-columnists in a separate category
- Left-leaning blogs more active than right-leaning blogs
- Weekly and seasonal activity variations
- Activity triggered in part by external events

4. Conclusions
- Successful proof of concept for our tools
- Further tweaking of data outputs required
- Able to provide weekly/monthly pointer to ‘hot’ topics/sites
- Need to operate for 6-12 months to establish baseline data
- Can be expanded to tracking activity in other user/led content creation spaces (Twitter, Facebook, YouTube, …)

5. References