

# Intro to Microformats

Chapter 2 of Russell, Mining the Social Web

## Outline

- Sematic Web & Microformats
- XFN to motivate BFS

## Semantic Web

- Make the Web (a collection of pages) into data useful to machines
- Goal, use markup to make the relationships, meanings (semantics) explicit
- Competing standards, formats
- Low incentive to change
  - What's in it for the person rewriting/reencoding their pages (data)?

## What is a Microformat?

- Snippet of well structured information
- Evolutionary, make small pieces of the web useful as data
- Markup
  - XHTML
  - Semantic HTML
  - RDF in XHTML

## Some standard microformats

- XFN – social networks in the web
- geo – specification of geocoordinates (lat, long) (draft)
- hCard – contact information
- hRecipe – recipe information (draft)
- XOXO – extensible outlines
- XMDP – meta data profiles

## XFN – Collecting Social Data

- Motivate Breadth First Search
- This has implications for all kinds of social data collection
- Look at this a bit ...

## Consider Twitter

The image shows a Twitter profile for @dwmcphd. The profile is centered, with a green smiley face icon and the handle @dwmcphd. To the left is a list of followers, and to the right is a list of people followed. Green arrows point from the followers list to the profile and from the profile to the following list.

**You have 143 followers**  
Here's more about them.

**People**

- BookinBookOut** Book in Book Out  
Helping students navigate textbooks with each other: <http://www.bookinbookout.com> <http://psu.bookinbookout.com>
- reclamemais** Rafael Fernandes
- Twenebowa** David T. Larkson  
Assistant professor at Ashland University College, design researcher, UX consultant, design and development
- brunlindeath** Brian Lindeman  
technology studies, ethnography of design, sociology of science, & motorcycles & eglife & bicycles &...
- infoqit52** infoqit52
- lena\_m** Lena
- BethPatt** Beth Patt  
PhD student from NCQA at the UW Information School. I am currently studying ethics, informatics and the value of libraries.

**You follow 79 people**  
Here's more about them.

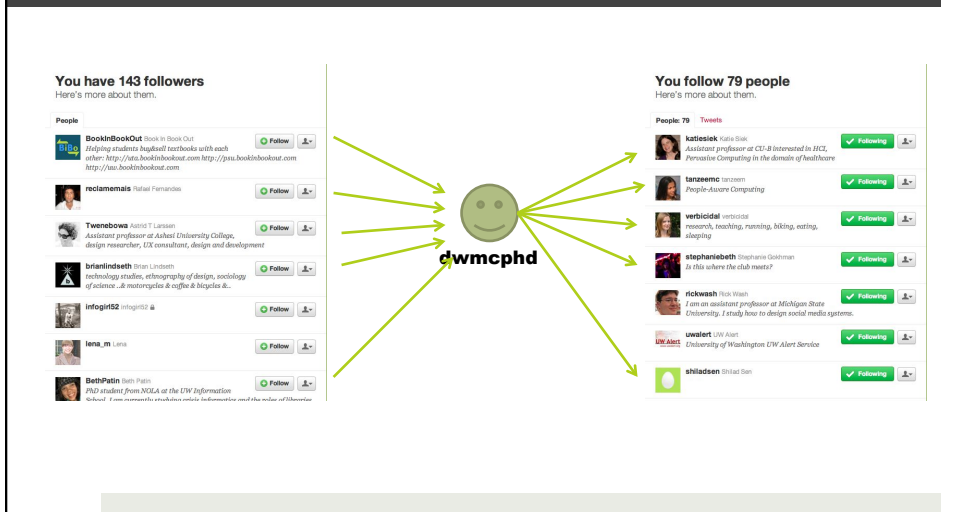
**People 79 Tweets**

- katesiak** Kate Sisk  
Assistant professor at CU-B. Interested in HCI, Pervasive Computing in the domain of healthcare
- tanzeemc** tanzeem  
People-Aware Computing
- verbiidid** verbiidid  
research, teaching, running, biking, eating, sleeping
- stephaniebeth** Stephanie Goldman  
Is this where the club meets?
- rickwah** Rick Wah  
I am an assistant professor at Michigan State University. I study how to design social media systems.
- UWAlert** UW Alert  
University of Washington UW Alert Service
- shildesen** Shilad Sen

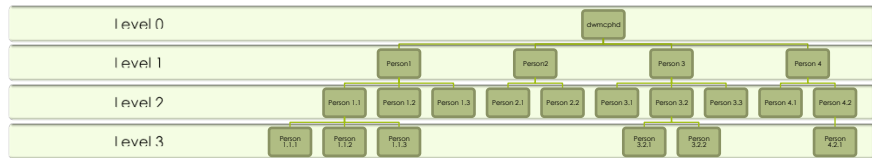
## Followers Graph

- On an individual basis somewhat clear
- This is really a directed graph (digraph)
  - Each node (person) has some number of incoming and outgoing edges.

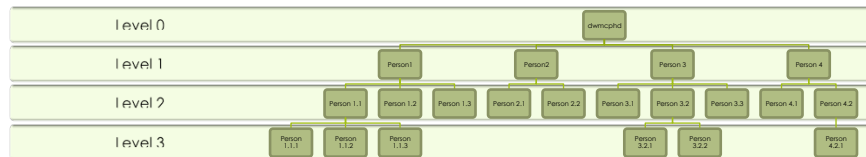
# Follower digraph



# Simplify as a tree



## Breadth First Search (BFS)



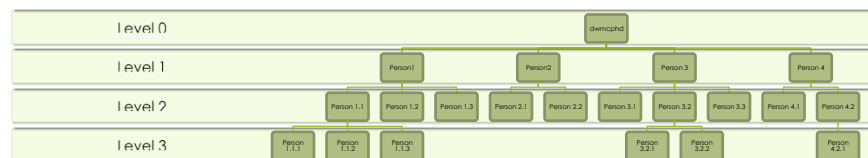
- Process each Level of the tree before going to the next level
  - Keep a queue, an ordered list
  - Process the current node
  - Add children to the queue
  - Take a node from the queue
  - Repeat

What are some other examples?

## What are some other examples?

- Web pages
  - Hyper links relate one page to another
  - Wikipedia pages
- Blogroll
  - Blogs that reference on another
- Twitter
  - Followers
  - Retweet, or hashtag semantics

## Depth First Search (DFS)



- How would that work?

## Thoughts

- Breadth First
  - Generally good for this kind of data
  - Still need to watch for "cycles" in the tree (digraph)
  
- Need to think about relational data in a different way
  - Not all social data is explicit articulation of "friend" or "follower"