Searching for Myself: Motivations and Strategies for Self-Search

Cathy Marshall and Siân Lindley
Microsoft Research

CHI 2014, Toronto
1 May 2014
back in 2006 I started a blog...
As long as we stick to the far distant past, I can dredge up some comparable experiences. But no-one wants to hear another joke about Keith Richards being the work of a particularly talented and ambitious embalmer. I’m tempted to talk about the time Rock Howard and I saw the Monkees at Disneyland: a show that’s so bad it’s good.
I am lucky in that Bevy is much more on top of musical happenings than I am. Also her taste in music has been influenced over the years by my own and so that works out well. Case in point: I introduced her to the Meat Puppets circa "Up on the Sun" and we went and saw them a few times back in the day. Nowadays the Meat Puppets have reformed and have relocated to Austin while Beverly has become the biggest Meat Puppets fan in the known universe. In fact, since Bevy does genealogy, I know more about the Kirkwood family tree than my own mother's ancestors.
how often do people look for themselves?

- 2009 Pew Internet poll: more than ½ of adults (57%) had searched for their own names
- 2008 privacy research: 30% of users had performed a vanity search over the study’s 70-day period (Jones, Kuman, Pang, & Tomkins, CIKM’08)
- The popular business press advocates frequent vanity searches as part of managing one’s ‘personal brand’ (but see Allison Woodruff’s CHI’14 paper about reputation management)
- And Lady Gaga? She looks for herself ‘daily’.
this work began via an interview study

re: online stuff
• what do people have?
• what do they value?
• how do they keep it now?

• if we (Microsoft) built an integrated archive for them, what might they do with it?

Lindley, Marshall, Banks, Sellen, Regan Rethinking the Web as a Personal Archive Proc. WWW’13

To remind participants of what they had online, we asked them to search for themselves. Sometimes they did this anyway.
why do people search for themselves?

• to manage their digital footprint
• to discover new content about themselves
• to re-find stuff they know is out there
• for entertainment / out of interest
• to unearth lost or forgotten content*

One thing we noticed was that the process itself was highly engaging for most participants.

In this talk, I’ll focus on surprises and implications, rather than on comprehensive results.
5 surprises...

- the search results page itself
- participants’ digital doppelgangers
- the way participants used image search
- the role of social media sites
- vanity search and personalization
Before the social media revolution, there was some thought that the SERP could be a digital calling card.
Managing one's digital footprint: SERP order

It wasn't what was on the SERP, it was the order that results appeared that surprised participants. They wanted to control the relative order.

"It's always been the same stuff... [but there's] more stuff that I would rather see higher up, like that story from China was a story that I enjoyed doing...." [Todd]

By contrast: "...when I see this weird Quora user page so prominent, or that MyLife thing, it just seems like something is not working quite correctly there. Because it's not relevant to me. ..." [Todd]
doppelgangers: the other yous

news anchor who paid for ‘reputation management’

me

author of religious novels for ladies

interior decorator (replaces South African winemaker)
managing one’s digital footprint: doppelgangers

participants DID NOT always want to weed out their doppelgangers and claim their own stuff.

• participants had strategies for seeing around their doppelgangers... (e.g., Kim used her middle name)

• doppelgangers offer playful competition and reassurance:

  “[I search for my name] Mainly to just make sure I’m above my brother.” [Oliver]

• but also doppelgangers offer a digital smokescreen:

  “I am thankful for [a woman whose name is spelled similarly], who is a jewelry designer, so when you search my name, she [comes up]. There's like three pages of [links to her pages]. I'm going, “thank god this woman exists.”” [Lynn]
managing one’s digital footprint: image search as strategy

some participants manage their digital identity through image search

– helps the searcher find unwanted images

– helps the searcher quickly zero in on desired content (it’s easier to scan a page full of photos than a page of text)
managing one’s digital footprint: social media

social media sites offered some participants newfound control over the SERP (e.g. by making some accounts public and others private)

• “If you use social media in the right way then if someone searches for you, you can make sure that the results that they find are about you and not about someone else.” [Ava]

• at the same time, social media could be the source of leakage between distinct personas, particularly FB ‘likes’.

  e.g. Ann used “BitBucket for the nerd persona” and SoundCloud for her musician persona, but by liking something with the wrong FB account, one persona leaked into the other.
managing one’s digital footprint: personalization

participants found personalization annoying (and difficult to turn off) when they were trying to manage their own digital footprints.

• “I don't know whether [personalization is] set right now. I think it might be the default…” [Todd]

• “…Google's strange, because it will pull up things you've already pulled before. ... I don't know what normal people would see.” [Kim]
3 design implications

- revisiting the idea of “SERP as profile”
- using the SERP as a privacy diagnostic
- supporting vanity search as a process
dimensions of SERP as dynamic profile

Controllability

Controllability: does self-searcher have the ability to directly change whether the item appears on the SERP?

[didn’t expect it]

Expectedness

Expectedness: Is self-searcher aware of the material and how it got there (its source and provenance)?

[can’t control it]
dimensions of SERP as dynamic profile

Controllability

e.g. candid Facebook photos posted by a friend

Expectedness

e.g. public social network feed

e.g. real estate data for your house

e.g. article in online newspaper archive; old newsgroup post
lightweight SERP control
(for controllable/expected content)

• social media results can be ordered or visible according to audience without compromising the integrity of the SERP
  – to find type of audience, compare relative path lengths (distance on different networks)
  – define categories of social media sites (pinterest and tumblr v. LinkedIn and Google Scholar)
  – e.g. Facebook friends can see Pinterest profile; LinkedIn contacts can’t
• doppelgangers can be suppressed, highlighted, or left alone, again according to audience
SERP can be presented diagnostically

• help self-searchers diagnose social media ‘leaks’ and unintended disclosures
  – e.g. using Stutzman & Kramer-Duffield’s analysis of intended v. expected audiences

• allow self-searchers to work backward from the SERP to identify desired privacy settings
  – allow them to see the tie between settings and SERP visibility
aid in self-search process

- assist in query formulation for the different functions of self-search
  - e.g. finding forgotten content v. showing digital footprint
- show changes in SERP over time
- organize results around specific content
  - help find and manage reuse, comments, attention
- reveal the effects of personalization
  - search without personalization
  - search as someone else:
    how would a prospective employer see me?
conclusion

- self-search should be treated as a distinct type of information seeking
- SERP-as-dynamic-profile should be reconsidered and supported in a social media era
contact info:

cathymar@microsoft.com
sianl@microsoft.com

http://research.microsoft.com/~cathymar
http://research.microsoft.com/en-us/people/sianl/

blog—http://ccmarshall.blogspot.com
twitter—http://twitter.com/ccmarshall
appendix
Aggregators pull records for a named entity from public and for-pay big data information services
• e.g. tax/voter rolls
• e.g. arrest records

May draw from social media sites like Facebook

May be domain-specific, pulling together real estate related data

Examples:
• Spokeo
• PeekYou
• Redfin...
managing one’s digital footprint: aggregators

participants found various aggregators in turn surprising, disturbing, and inaccurate

• they draw together identities kept distinct: “This is what I found unnerving because it linked together profiles that weren’t necessarily linked together” [Ava]

• they mix one’s own identity with that of doppelgangers

• they violate ‘privacy through obscurity’ and other social norms

• they contain out-of-date information: ‘I haven't lived there in almost 10 years’ [Todd]

Yet... several admitted they’d used aggregator sites’ data to look up other people.