

Self-Representation of Online Identity in Collected Hyperlinks

Terrell Russell

SILS, UNC-CH, CB #3360, 100 Manning Hall Chapel Hill, NC 27599 Email: unc@terrellrussell.com

Frederic Stutzman

SILS, UNC-CH, CB #3456, 213 Manning Hall Chapel Hill, NC 27599 Email: fred@metalab.unc.edu

Online identity is the representation of one's persona in a digital context. A primary factor in this representation is the collection of links that represent an individual in search. Designed by the authors, claimID (<http://claimID.com>) is a web service that enables individuals to use the hyperlinks about them to create a rich presentation of their online identity. In this paper, we analyze patterns of self-representation of online identity as observed in claimID.

Introduction

A primary factor in an individual's online identity is their representation in web search results. Simply, when an individual's name is entered as a search query, a set of matching results is returned. The quality of this result set of links may vary wildly, from topically appropriate to dramatically incorrect and/or misleading. In this context, the search engine presents an agentic representation of identity, over which the individual generally has little control. This representation of identity is particularly troublesome considering the value placed on this reputation – a recent survey found that 77% of employers “Google” potential employees (Donlin, 2006).

With the creation of claimID (Stutzman & Russell, 2006), users are enabled with the ability to create a self-representation of their identity by engaging the hyperlinks about them. In doing so, these individuals are provided a voice in the conversation happening about them online (between searchers and the search engines). This simple, effective approach to the management of online identity has proven popular with thousands of users.

Approach

Each link that a user collects onto their claimID page (<http://claimid.com/terrell> for example) may be categorized as either “By” them or “Not By” them, and “About” them or “Not About” them. There is also a third possibility of “Other” for each category, as a number of users wished to create representation of self outside the admittedly simple taxonomy.

As more URLs are collected in a user's account, the user begins to draw a more comprehensive picture of who he is online. As he connects the different facets of his online identity – news articles, social network profiles or blog posts, for example, he empowers the searcher to better find these URLs and disambiguate himself as compared to

others who might share the same identity characteristics. In addition, by creating links between related documents online, the search engines will likely better cluster those documents when presenting search results web-wide.

Observations

By/About

User categorization of online identity links in claimID is presented in Table 1. The vast majority of links that users are claiming are both “By and About” themselves (54%). The two adjacent categories (“Not By and About”, “By and Not About”) are used 11% and 12% respectively. The “Not By and Not About” category is not directly related to the person's online identity, but surprisingly, is used on nearly 8% of the links, mostly for disambiguation purposes (they point to sites by/about other people with the same name). This is higher than the authors expected and suggests that disambiguation is a larger issue than first imagined. The rest of the categories (using the “Other” categorization) are used less (1-3%), with only “By and Other” seeing a marked difference at over 6% of the links in the system.

Table 1. User-categorized identity links

	About	Not About	Other	Totals
By	23,647 (53.81%)	5,397 (12.28%)	2,751 (6.26%)	31,795 (72.35%)
Not By	4,844 (11.02%)	3,413 (7.77%)	1,411 (3.21%)	9,668 (22.00%)
Other	783 (1.78%)	629 (1.43%)	1,071 (2.44%)	2,483 (5.65%)
Totals	29,274 (66.61%)	9,439 (21.48%)	5,233 (11.91%)	43,946 (100.00%)

Based on 6418 accounts with 1 or more claimed links.

Users with at least one link in their account have claimed an average of 6.8 links, while the median number of links per account is 4. The largest number of links in a particular account is 161, and the IQR is 8 (25% percentile: 1, 75% percentile: 9). The distribution's standard deviation is 10.68, and it displays a positive skew of 9.8.

Verified Identity

Users at claimID have the option of verifying their claims of link ownership with the open identity protocol MicroID (<http://microid.org>). When a user verifies ownership of a

link at claimID, a small verification badge appears next to that link, creating the perception of verified identity. While research has explored the enriching effects of verified identity in a virtual community (Millen & Patterson, 2003), we wished to explore the effects of verified identity on self-disclosing behavior. In Table 2, we present the results of two analyses of the effects of verification of identity on self-disclosing behavior.

Table 2. Effect of identity verification on self-disclosure.

	Coefficient	F	R ²
RQ1* (n=6418)	5.18	271.63	.0405
RQ2* (n=1419)	2.06	109.47	.0711

* $p < .05$

In RQ1 we examine the effect of verification on number of links presented in a claimID profile. Treating verification as a dummy variable, we find a significant effect of verification on self-disclosing behavior. In RQ2 we examine the effects of level of verification on levels of self-disclosure, treating the number of verified links as a continuous variable. Again, we find that higher levels of verification display a significant positive effect on self-disclosing behavior. The effect of this is presented in Figure 1.

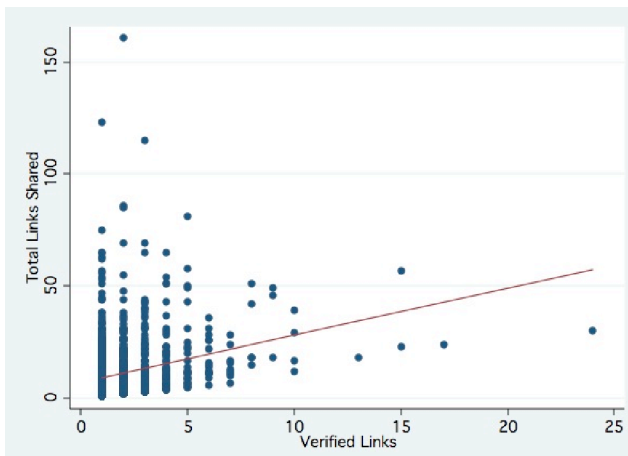


Figure 1. Effect of link verification on self-disclosure behavior (RQ2)

Discussion

Online identity is a complex issue. We are just beginning to struggle with our new reality of global, archived, persistent identity. Search engines have reduced the complexity of finding information about another human, but at what cost? The agentic representation of identity in

search engines is clearly a departure from our more traditional norms for identity presentation and dissemination.

As our preliminary findings show, individuals feel the need to play a self-agentic role in the presentation of their online identity. While the majority of representation is confirmatory (Table 1), a significant minority of this representation is non-confirmatory, in an attempt at disambiguation. At the same time, we found a surprising number of identity links per user – nearly 7 – illustrating that our digital identity future is made up of many different representations, spread throughout the net. Finally, we explored the effects of verification on levels of self-disclosure. We found that individuals who verify their online identities are more likely to engage in self-disclosing behavior, pointing to an interesting link between the ability to verify one’s self and motivations for engaging in self-disclosing behavior.

We must note the limitations in our methodology, however. Users of claimID are self-selected; while we have no reason to doubt they are a fairly representative sample of the internet, certain tendencies that draw an individual to claimID may be present in our overall analysis. With our verification analysis, we must note that individuals may not be able to verify all of their claimed sites. Our verifier requires the insertion of a small token into a webpage for this verification to take place. This may render some sites unverifiable, but does not lead us to believe that the analysis as presented is ecologically invalid.

Conclusions

Individuals who used claimID are actively taking a role in shaping their online identities by presenting their individual representation of self in hyperlinks. By linking to these resources from a central location, they are bringing these disparate resources closer together, and presenting a more cogent self. While this analysis is preliminary in nature, we feel that it provides a number of interesting and useful insights into how we may collectively manage our identities in this new identity future.

References

- Donlin, K. (2006, August 11) Job Hunting? Go Google Yourself. Minneapolis Star-Tribune, p. C1.
- Millen, D. and Patterson, J. (2003) Online Identity, community networks, social capital. CHI 2003. New York: ACM.
- Stutzman, F. & Russell, T. (2006). ClaimID: A System for Personal Identity Management. Proceedings of the 6th ACM/IEEE-CS JCDL. (pp. 367-367) New York: ACM.