Introspective Users and Introspective Text: Some Recent Results

Shomir Wilson – Carnegie Mellon University

Timeline

2011: PhD, Computer Science, University of Maryland
Metacognition in AI, dialogue systems, detection of mentioned language

2011-2013: Postdoctoral Fellow, Carnegie Mellon University
Usable privacy, mobile privacy, regret in online social networks

2013-2014: NSF International Research Fellow, University of Edinburgh
2014-2015: NSF International Research Fellow, Carnegie Mellon University
Characterization and detection of metalanguage
Also: collaboration with the Usable Privacy Policy Project

Faculty and Professional Collaborators

University of Maryland: Don Perlis

UMBC: Tim Oates

Franklin & Marshall College: Mike Anderson

Macquarie University: Robert Dale

National University of Singapore: Min-Yen Kan

Carnegie Mellon University: Norman Sadeh, Lorrie Cranor, Alessandro Acquisti, Noah Smith, Alan Black

University of Edinburgh: Jon Oberlander

University of Cambridge: Simone Teufel

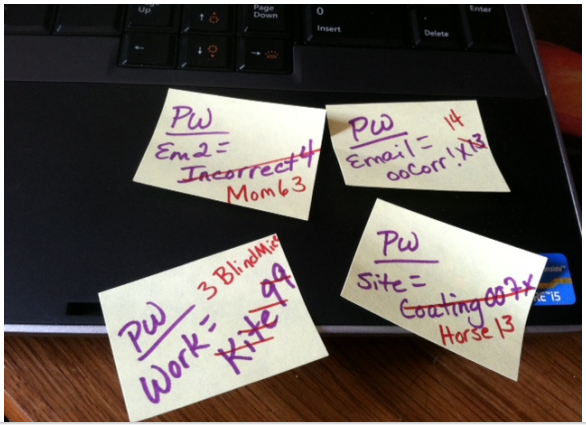
Student Collaborators

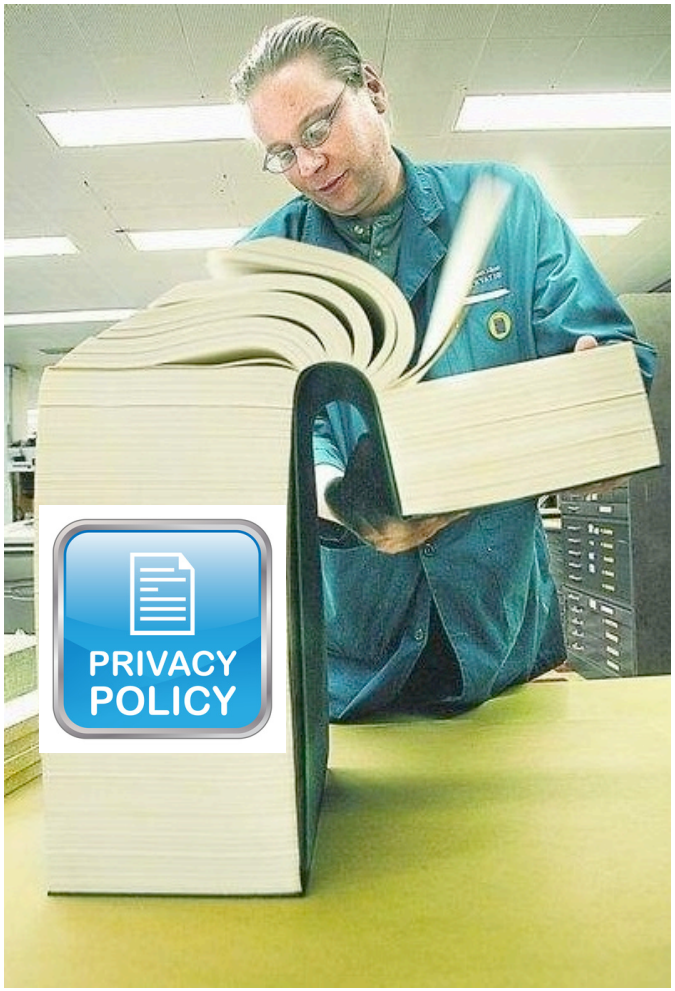
Carnegie Mellon University: Hazim Almuhimedi, Bin Liu, Salem Hilal, Jon Breiger, Rob Murcek, Tommy Doyle

University of Cambridge: Kevin Heffernan


language
privacy
policies
metalanguage
users
information
phenomenon
systems
distinction
technologies
approach
natural
identifying
instances
corpus
paper
sharing
use-mention
notice
website
humans
explicit
deixis
understand
artifacts
describes
use
corpora
text
work
websites
automatic
results
present
extract
computational
communicative
reason
three
cas
deleted
learning
expectations
communicate
frequently
analysis
research
location
e.g
potential
performance
words
key
wikipedia
future
detection
figure
make
features
word
materials
crowdsourcing
presents
must
essential
wordnet
scalable
sentence
texts
specific
practical
tweets
often
mcl
automated
wikibooks
lists
examine
illustrations
understanding
entities
made
set
ai
readers
senses
loop
better
existing
occurs
metacognitive
basis
settings
cause
system
complex
ability
using
mention
showing
clients
combines
artificial
content
first
references
questions
allows
tools
impact
perform
aspects
problem
preferences
behavior
without
designers
machine
decisions
go
policy
one
role
failures
well
attempt
study
human
pedagogical
suggests
vocabulary
aspect
phrases
recognize
common
framework
feasibility
unexpected
properties
patterns
discuss
recognition
become
relatively
read
problems
building
gathered
intelligent
different

Usable Privacy: Motivations





http://www.paintsquare.com/blog/images/PSN_1002_Blog_StickyNotes.JPG
http://stylettomag.co.uk/wp-content/uploads/2014/05/Diary.jpg

http://img3.wikia.nocookie.net/__cb20140304030658/degrassi/images/9/9a/Big-book.jpg
http://thebriberyact.com/wp-content/uploads/2010/09/privacy-policy.jpg

Oversharing, Regret, and Nudging

NEWSFEED BIZARRE

A Texas Teenager Got Fired for a Tweet Before Starting Her Job

Rishi Iyengar @iyengarrishi | Feb. 11, 2015

Employers use social media too, kids

A Texas teenager got fired from her new job less than 24 hours before she started after she used a couple of choice expletives to describe it on Twitter.

RECOMMENDED FOR YOU

Reveal too much	117	25%
Direct criticism	96	20%
Expressive	64	14%
Direct attack	62	13%
Blunder	51	11%
Implied criticism	34	7%
Group reference	13	3%
Agreement changed	3	1%
Behavior edict	2	0%
Lie	1	0%
Other	31	7%

Oversharing in an online social network (OSN) can lead to regret.

Can we identify OSN content that individuals are likely to regret?

Can we help people maintain their professed sharing preferences?

http://time.com/3706434/cella-tweet-fired-texas-jets-pizza/

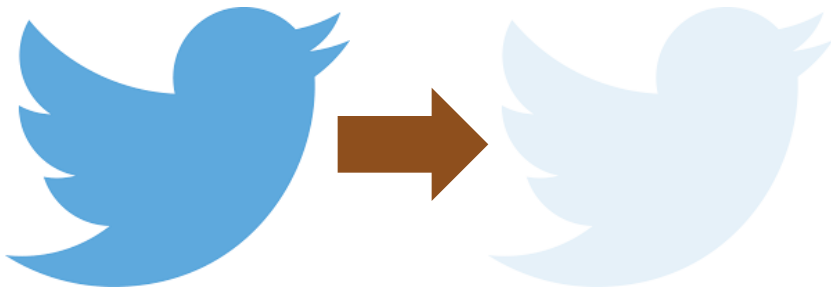
"I read my Twitter the next morning and was astonished": A conversational perspective on Twitter regrets. Manya Sleeper, Justin Cranshaw, Patrick Gage Kelley, Blase Ur, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh. CHI 2013.

Twitter Deletion Study

OSN post deletion is potentially an indication of regret. Can we study regret via deletion?

We tracked 292K active Twitter users for one week and collected their public tweets.

We used deletion notices from the Twitter API to track when tweets were deleted.



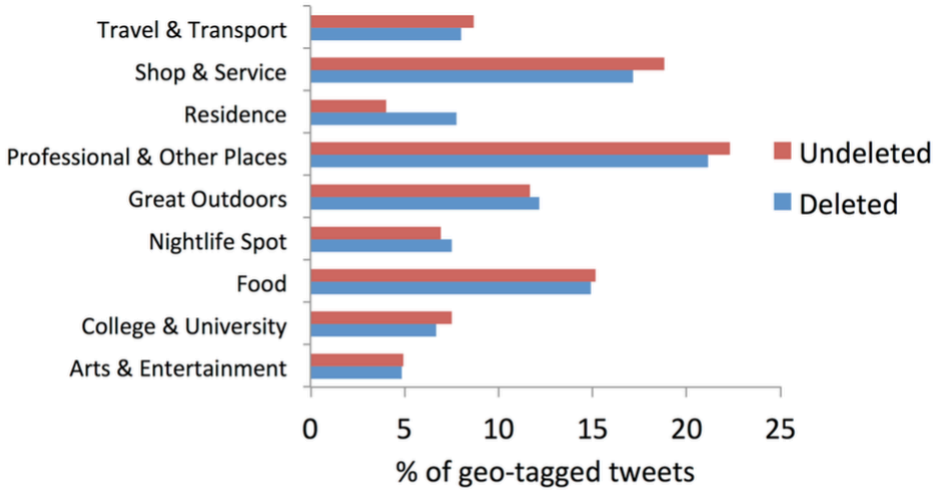
Hazim Almuhimedi, Shomir Wilson, Bin Liu, Norman Sadeh, and Alessandro Acquisti. Tweets are forever: A large-scale quantitative analysis of deleted tweets. In Proc. CSCW 2013.

How Are Deleted Tweets Different?

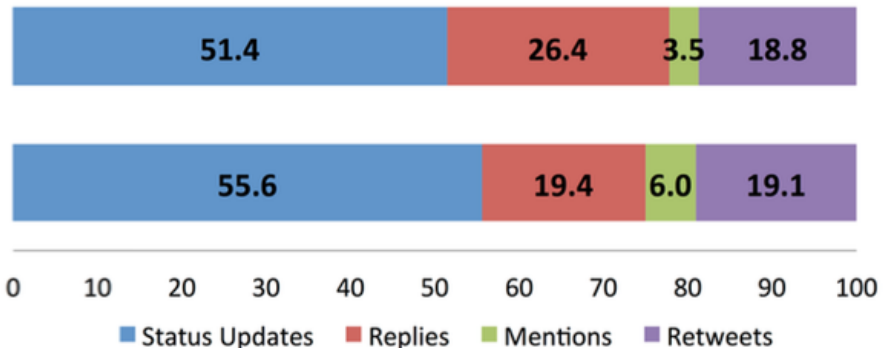
We collected a total of 6.7M tweets. 2.4% were deleted during the observation period.

In aggregate, there were some significant differences between deleted and undeleted tweets.

Tweet Location: Non-Deletion vs. Deletion



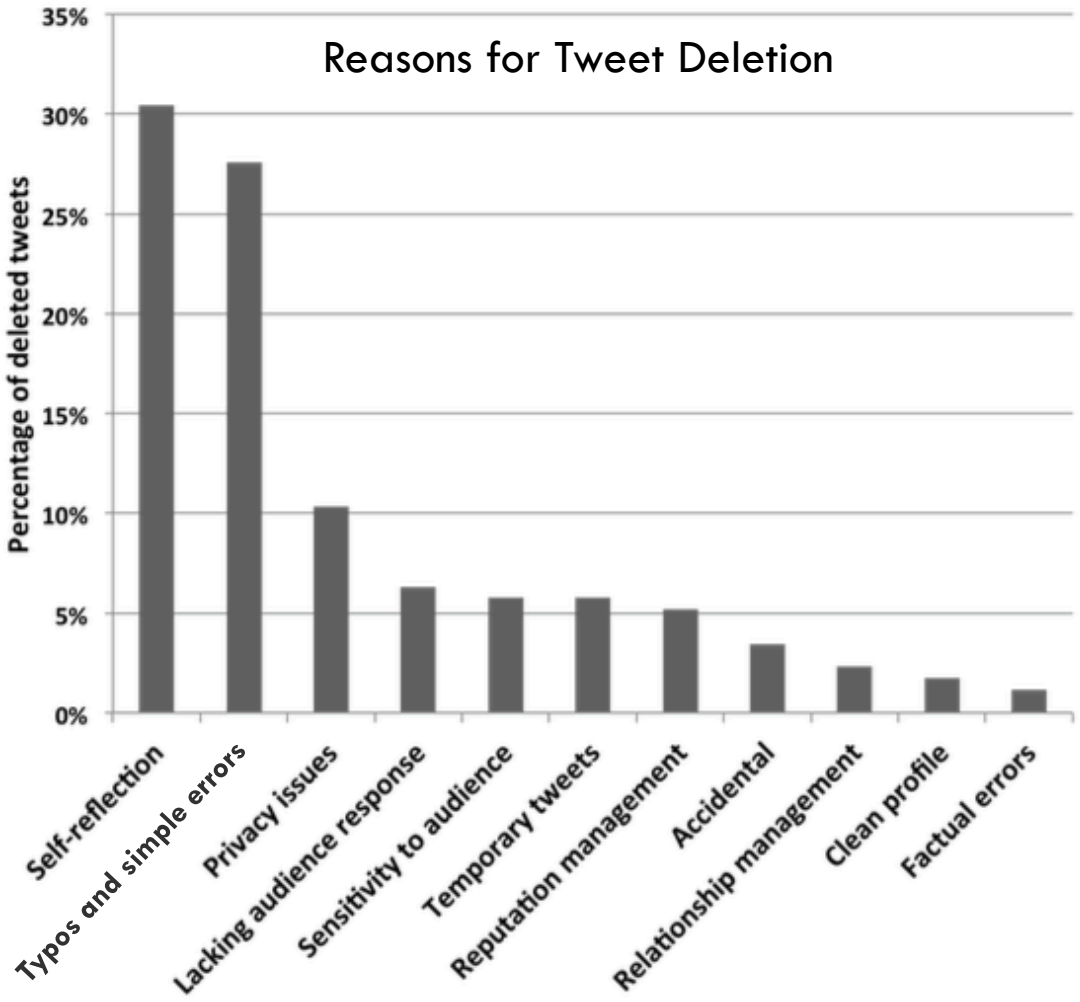
Tweet Origin: Non-Deletion (Top) vs. Deletion (Bottom)



Discussion

- Deleting a tweet doesn't mean it's completely gone
- In aggregate, deleted tweets show some intuitive traits
- Still, in aggregate, deleted tweets are just barely distinctive

In the Pipeline: A User Study



Ideal Scenario: Non-Retweets

Action	%
Make changes	38
Post nothing	34
No change	23
Other	5

Ideal Scenario: Retweets

Action	%
Do not retweet	47
No change	37
Add comments	13
Other	3

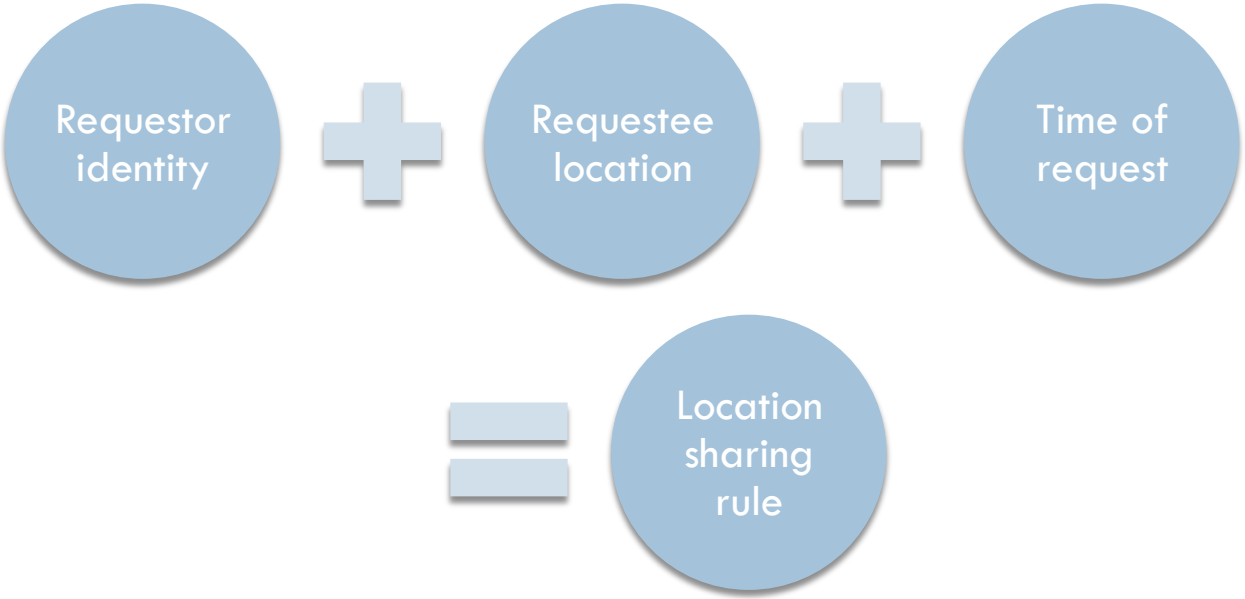
Location Sharing

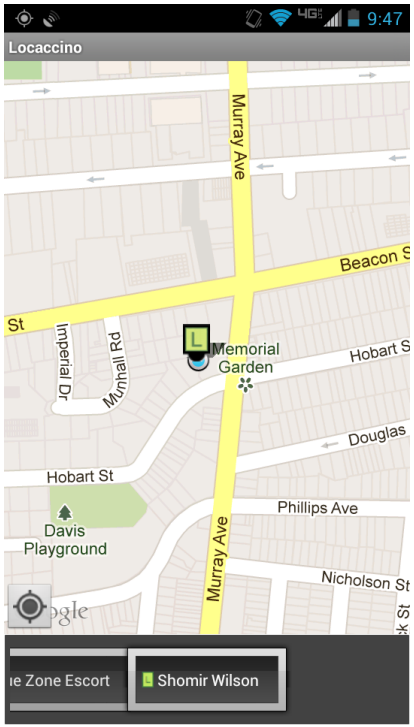
Locaccino (2010-2013)

Location sharing and CMU shuttle tracking

Available for iPhone and Android

~35,000 downloads





Shomir Wilson, Justin Cranshaw, Norman Sadeh, Alessandro Acquisti, Lorrie Cranor, Jay Springfield, Sae Young Jeong, and Arun Balasubramanian. Privacy manipulation and acclimation in a location sharing application. In Proc. Ubicomp 2013.

Study Motivation

Finely-configurable OSN privacy settings are

- good: they can reflect users' nuanced preferences
- bad: they require attention to configure and maintain

Privacy profiles can represent users preferences.

- Mugan et al. clustered OSN users' location sharing preferences.

How does presenting privacy profiles to users influence their comfort with location sharing?

Mugan, J., Sharman, T., and Sadeh, N. Understandable Learning of Privacy Preferences Through Default Personas and Suggestions. Technical report CMU-ISR-11-112: Carnegie Mellon University, 2011. Available at http://reports-archive.adm.cs.cmu.edu/anon/isr2011/CMU-ISR-11-112.pdf.

Conditions and Protocol

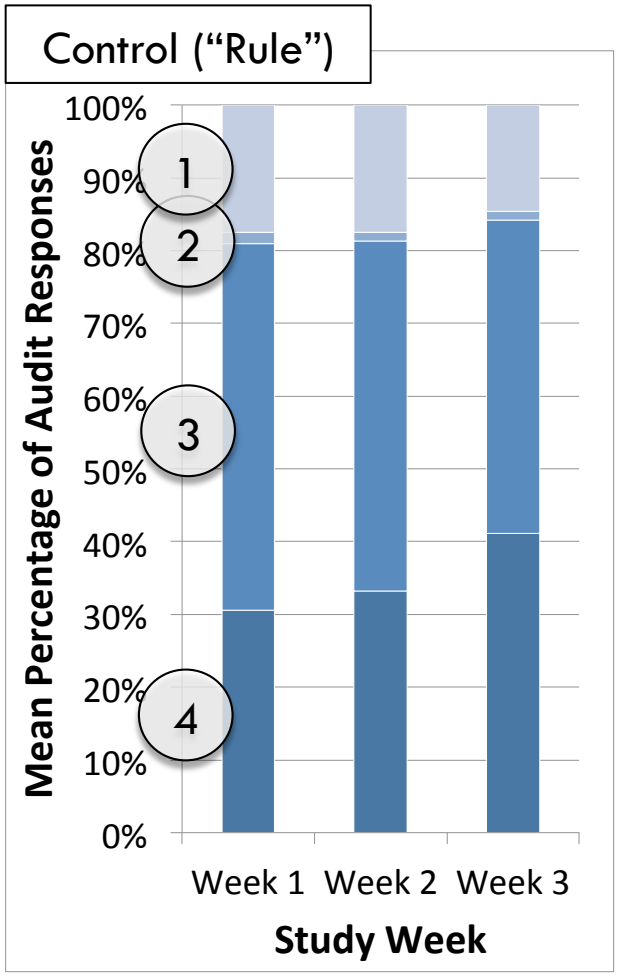
Subjects were randomly assigned to two conditions:

- Treatment ("profile"): 16 subjects
- Control ("rule") condition: 18 subjects

After initializing their settings, subjects used Locaccino for three weeks. Every night they audited real and hypothetical location sharing requests.

	Who and When	Type	Where was I?	Outcome	Feedback
	**Justin Cranshaw** Today, 22 minutes ago-22 minutes ago	What-if	Carnegie Mellon University, 4110 Wean Hall **(See map)**	Allow	
	**Justin Cranshaw** 57 minutes ago	What-if	Carnegie Mellon University, Skibo University Center **(See map)**	Deny	
	**Arun Balasubramanian** Yesterday, 11:59 PM-12:00 AM	Real	5737 Hobart St, Pittsburgh, PA 15217 **(See map)**	Deny	

Auditing: Composition of Results

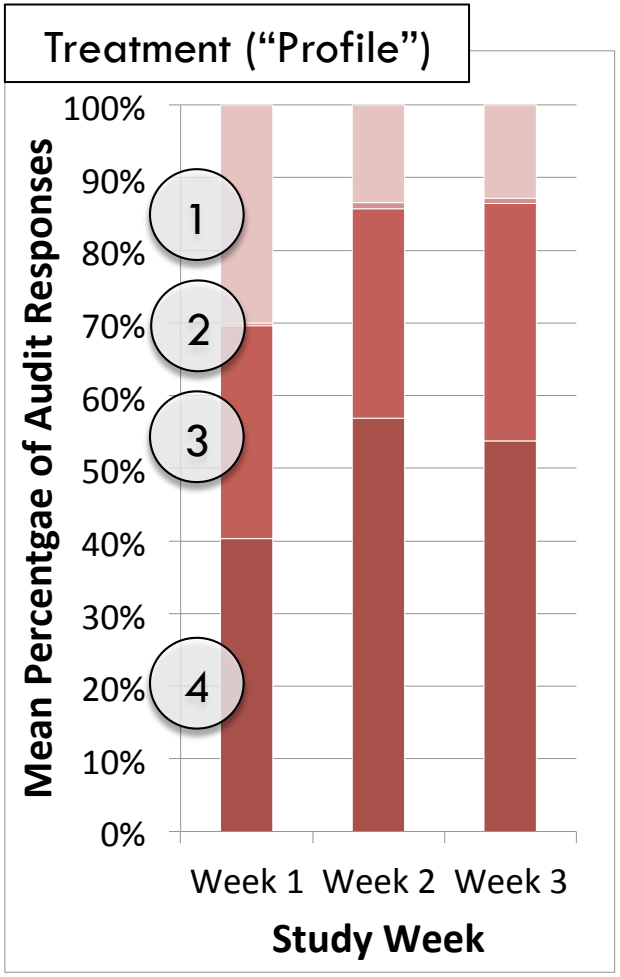


1 Request denied, unsatisfied

2 Request allowed, unsatisfied

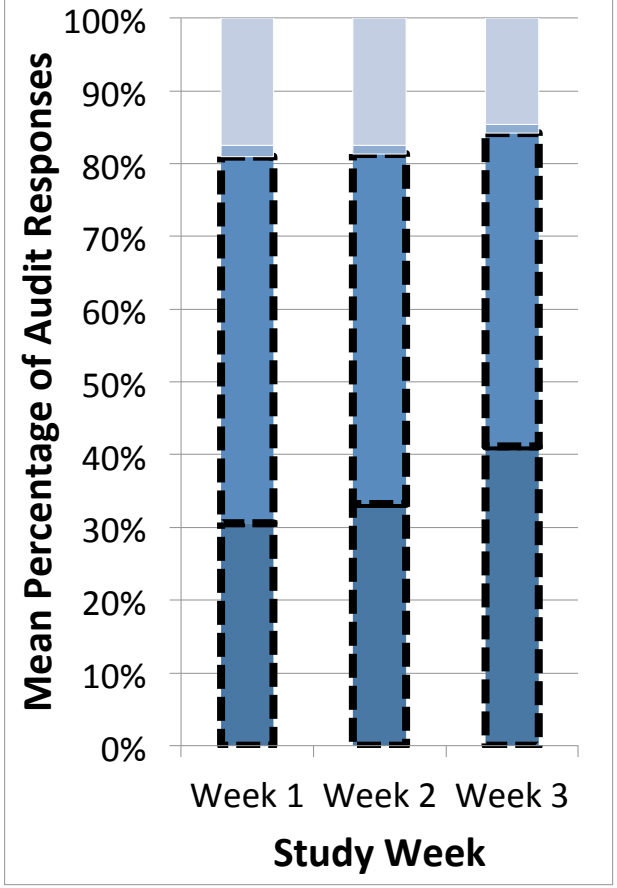
3 Request denied, satisfied

4 Request allowed, satisfied



Auditing: Satisfaction Rate

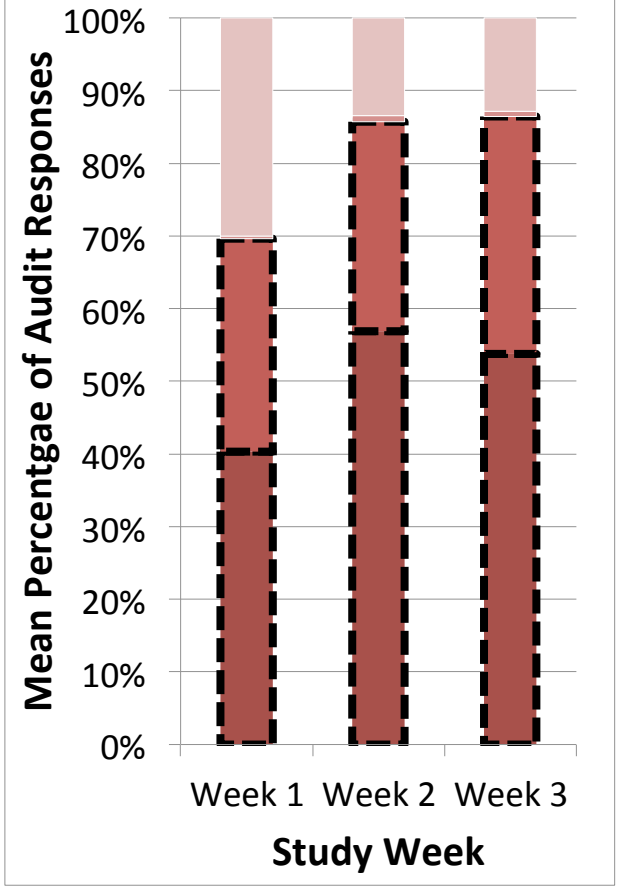
Control ("Rule")



The treatment group experienced a significant ($p=0.05$) increase in satisfaction from Week 1 to Week 3, but the rule condition did not ($p=0.23$).

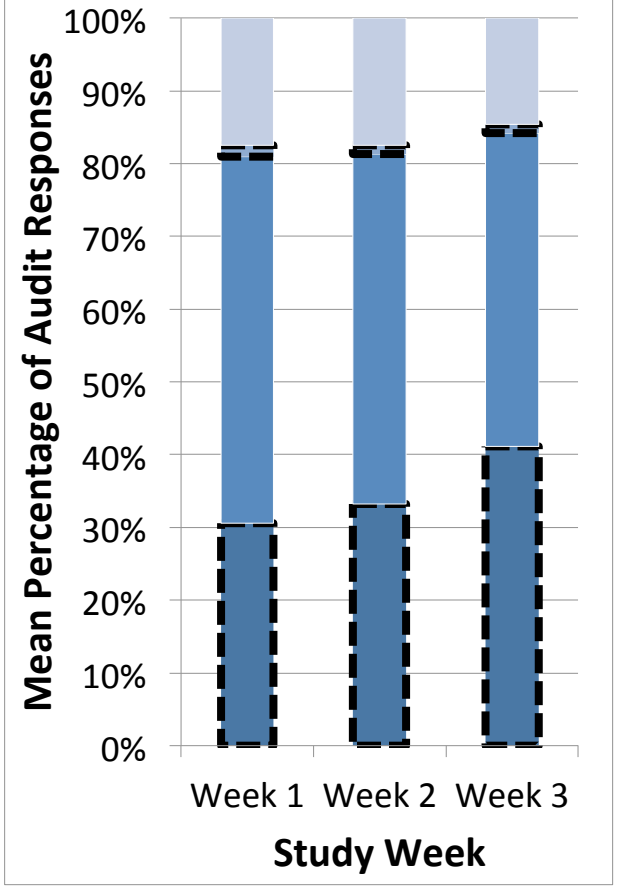
By-week differences between the groups were not statistically significant.

Treatment ("Profile")



Auditing: Sharing Rate

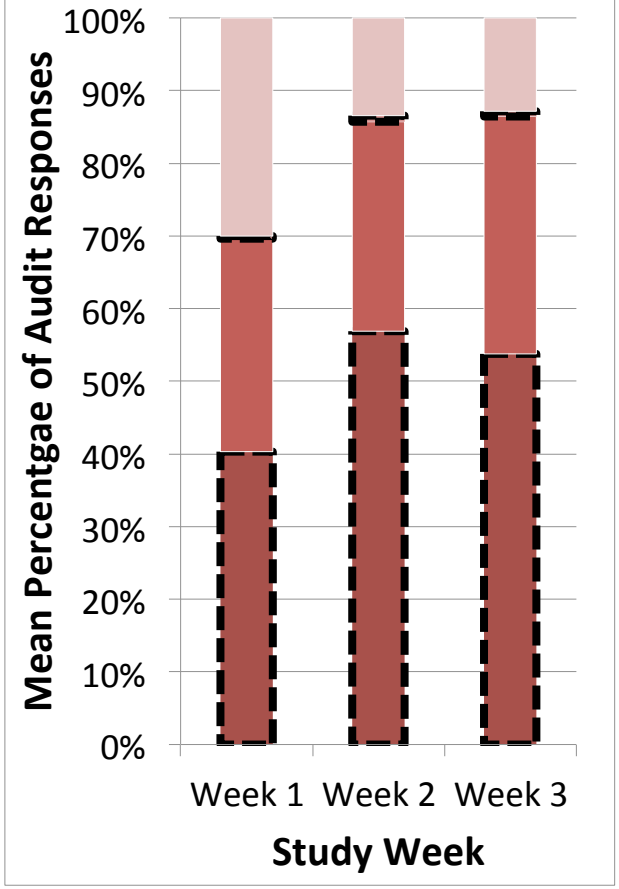
Control ("Rule")



Both groups showed trends towards greater sharing.

The treatment group shared significantly more during Week 2 ($p=0.01$) with mild indications of the same for Week 1 ($p=0.13$) and Week 3 ($p=0.093$).

Treatment ("Profile")



Discussion

Satisfaction in the conditions was roughly equal by the end of the study, but they never converged on an equal quantity of sharing.

Privacy profiles, as well as other efforts to simplify privacy choices, can have a significant impact on the levels of privacy that users select.

Privacy Policies: Status Quo

Last Revised March 11, 2014.

Kids and parents click here!

The following Privacy Policy summarizes the various ways that Condé Nast Digital ("Service Provider," "we" or "our") treats the information you provide while using www.wired.com ("Website"). It is our goal to bring you information that is tailored to your individual needs and, at the same time, protect your privacy.

Please read this Privacy Policy carefully. You can access the Privacy Policy any time at http://www.condenast.com/privacy-policy#privacypolicy. Your use of and/or registration on any aspect of the Website will constitute your agreement to this Privacy Policy. If you cannot agree with the terms and conditions of this Privacy Policy, please do not use the Website. This Privacy Policy does not cover information collected elsewhere, including without limitation offline and on sites linked to from the Website.

In addition to reviewing this Privacy Policy, please read our User Agreement. Your use of the Website constitutes agreement to its terms and conditions as well.

This Privacy Policy may be modified from time to time; the date of the most recent revisions will appear on this page, so check back often. Continued access of the Website by you will constitute your acceptance of any changes or revisions to the Privacy Policy.

I. THE TYPE OF INFORMATION THE WEBSITE COLLECTS

The Website generally collects personally identifying information with your specific knowledge and consent. For instance, when you enter a sweepstakes or contest, complete a survey, make a purchase, subscribe to our publication(s), or register for any portion of our services, you are asked to provide information such as your e-mail address, name or phone number. Optional information such as your age or gender may also be requested.

Our servers may also automatically collect information about your computer when you visit the Website, including without limitation the type of browser software you use, the operating system you are running, the website that referred you, and your Internet Protocol ("IP") address. Your IP address is usually associated with the place from which you enter the Internet, like your Internet Service Provider, your company or your university.

II. HOW THE WEBSITE USES INFORMATION PROVIDED BY YOU

Service Provider uses personally identifying information you supply through the Website to provide you with the service you have requested. For example, if you subscribe to any of our publications, we may use your e-mail address to send you a confirmation notice and your mailing address to send you the publication. Similarly, if you enter an online sweepstakes, we will use this information to notify you if you are a winner. We may also use the information to communicate with you about new features, products or services, and/or to improve the services that we offer by tailoring them to your needs.

Unless otherwise specified on the Website, Service Provider may sell or share personally identifying information with our affiliates and with carefully selected companies who we think can offer you services and products of interest to you. If you do not wish to have your personally identifying information shared, write to us at the street address set forth at the end of this document. If you do not wish to receive future commercial communications from us by e-mail, simply follow the unsubscribe instructions contained within the e-mail. If you've registered on any part of the Website, please use the mechanism on the Website that allows you to change or update your member preferences or information, if available, to keep all such date accurate and up-to-date. Otherwise, contact your Privacy Policy Coordinator as described below with your changes or click here to be removed from all Condé Nast Digital mailing lists.

We also allow access to our database by third parties that provide us with services, such as technical maintenance or forums and job search software, but only for the purpose of and to the extent necessary to provide those services. And if you choose to purchase items through features on the Website, we may forward your information to third parties for services such as credit card processing and order fulfillment. There are also times when you provide information about yourself to us in areas of the site that may be managed or participated in by third parties, such as auction services or shopping areas. In such cases, the information may be used by us and by such third party(ies), each pursuant to its own policies. We may also provide your information to our advertisers, so that they can serve ads to you that meet your needs or match your interests. While Service Provider will seek to require such third parties to follow appropriate privacy policies and will not authorize them to use this information except for the express purpose for which it is provided, Service Provider does not bear any responsibility for any actions or policies of third parties. We may also provide access to our database in order to cooperate with official investigations or legal proceedings, including, for example, in response to subpoenas, search warrants, court orders, or other legal process.

In addition, we reserve the right to use the information we collect about your computer, which may at times be able to identify you, for any lawful business purpose, including without limitation to help diagnose problems with our servers, to gather broad demographic information, and to otherwise administer our Website.

While your personally identifying information is protected as outlined above, we reserve the right to use, transfer, sell, and share aggregated, anonymous data about our users as a group for any business purpose, such as analyzing usage trends and seeking compatible advertisers and partners.

In addition, as our business changes, we may buy or sell various assets. In the event all or a portion of the assets owned or controlled by Service Provider, its parent or any subsidiary or affiliated entity are sold, assigned, transferred or acquired by another company, the information from and/or about our Website users may be among the transferred assets.

http://www.wired.com/about/privacy-policy/

Privacy Policies: Status Quo

III. COOKIES

You may have read about "cookies," nuggets of information that are placed by a Website in a storage place on your own computer. We may use cookies to control the display of ads, to track usage patterns on the site, to deliver editorial content, and to record registration and personalization information. For example, if you register on any part of the Website and are given the option to save your display name and password on your computer, we can provide this convenience to you by placing a cookie on your computer. Our cookies may contain personally identifiable information and such cookies may be shared with our affiliates and other companies.

Some of our advertisers occasionally serve you cookies as well. We do not have control over cookies placed by advertisers.

We may also use advertising service vendors to help present advertisements on the Website. These vendors may use cookies, web beacons, or similar technologies to serve you advertisements tailored to interests you have shown by browsing on this and other sites you have visited, to determine whether you have seen a particular advertisement before and to avoid sending you duplicate advertisements. In doing so, these vendors may collect non-personal data such as your browser type, your operating system, Web pages visited, time of visits, content viewed, ads viewed, and other clickstream data. The use of cookies, web beacons, or similar technologies by these advertising service vendors is subject to their own privacy policies, not ours, and Service Provider disclaims all liability in connection therewith. If you do not want the benefits of these advertising cookies, you may be able to opt-out by visiting http://www.networkadvertising.org/optout_nonppii.asp.

If you don't want any cookies, your Web browser likely includes an option that allows you to not accept them. However, if you set your browser to refuse cookies, some portions of the Website may not function properly.

IV. INFORMATION SECURITY AND NOTIFICATION

Because no data transmission over the Internet is completely secure, and no system of physical or electronic security is impenetrable, Service Provider can not guarantee the security of the information you send to us or the security of our servers or databases, and by using the Website you agree to assume all risk in connection with the information sent to us or collected by us when you use the Website. In the unlikely event that we believe that the security of your information in our possession or control may have been compromised, we may seek to notify you. If notification is appropriate, we may notify you by e-mail (provided we have your e-mail address).

V. KIDS AND PARENTS

This Website is not intended for use by children, especially those under age 13. No one under age 13 is allowed to provide any personal information or use our public discussion areas, forums and chats. Minors between the ages of 13 and 17 must get the permission of their parent(s) or legal guardian(s) before making purchases, including subscriptions, on this site.

If your children disclose information about themselves in publicly accessible areas of the Website, they may get unsolicited messages from other parties. Accordingly, you should tell them not to do so.

If you're worried about your children's activities or their privacy on our site, we encourage you to contact our Privacy Policy Coordinator as described below.

VI. PRIVACY POLICY COORDINATOR

To be removed from all Condé Nast Digital email lists, click here.

If you have any concerns or questions about any aspect of this policy, please feel free to contact our Privacy Policy Coordinator as follows:

Privacy Policy Coordinator

The Condé Nast Publications

1313 Market Street

Wilmington, DE 19801

Privacy_administration@condenast.com

Privacy policies are essentially read by no one

http://www.wired.com/about/privacy-policy/

But Ideally?

Nutrition Facts

Serving Size 3 oz. (85g)

Amount Per Serving	As Served
Calories 38	**Calories from Fat** 0
	% Daily Value
Total Fat 0g	0%
Saturated Fat 0g	0%
Cholesterol 0g	0%
Sodium 0g	2%
Total Carbohydrate 0g	3%
Dietary Fiber 0g	8%
Sugars 0g	
Protein 0g	

Vitamin A 270%	•	**Vitamin C** 10%	
Calcium 2%	•	**Iron** 0%	

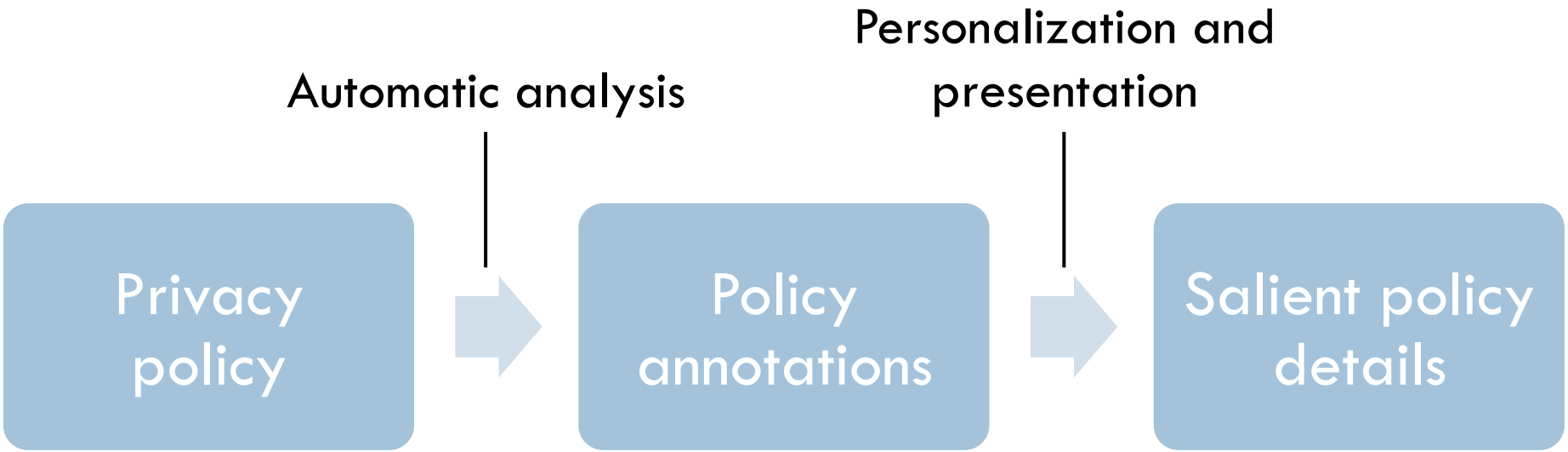
Percent Daily Values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs:

	Calories	2,000	2,500
Total Fat	Less than	65g	80g
Sat Fat	Less than	20g	80g
Cholesterol	Less than	300mg	300mg
Sodium	Less than	2,400mg	2,400mg
Total Carbohydrate		300g	375g
Dietary Fiber		25g	30g

Patrick Gage Kelley, Joanna Bresee, Robert W. Reeder, and Lorrie Faith Cranor. Design of A Privacy Label. In Proc. SOUPS 2009.

The Usable Privacy Policy Project

Goal: use crowdsourcing, machine learning, and NLP techniques to automatically (or semi-automatically) extract salient details from privacy policies.



For details, visit www.usableprivacy.org

Topic Change

Entity Linking and Artifact Reference

Communication in a document is not chiefly linear.

The entities that we refer to are not always external to the medium. Sometimes the referents are communicative artifacts.

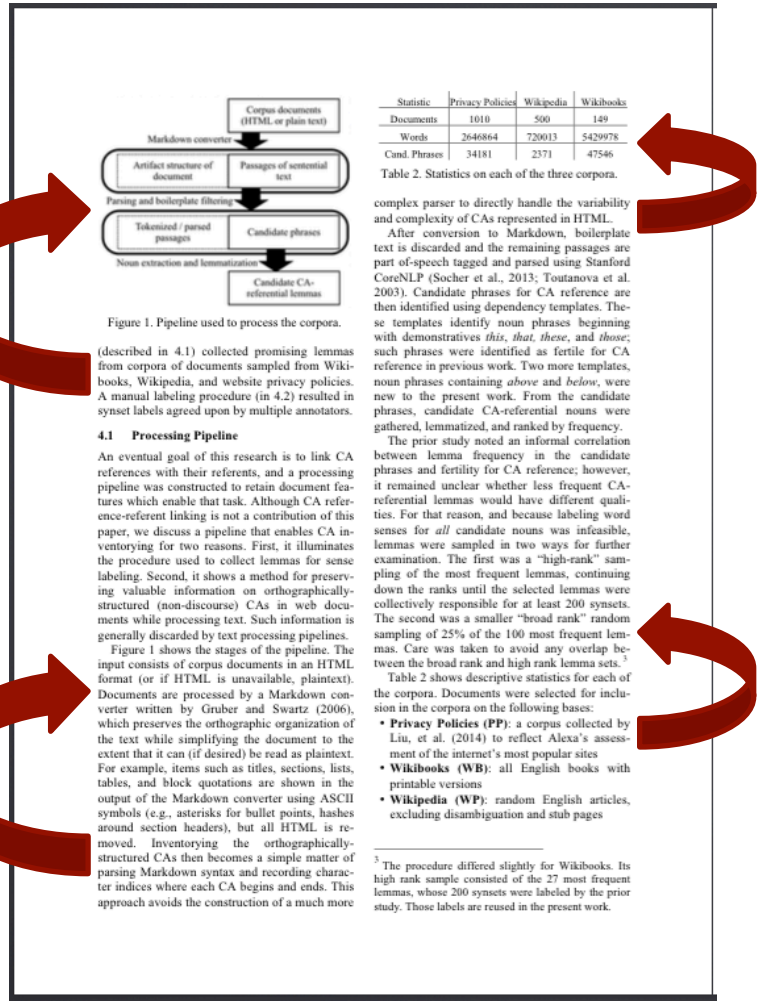


Figure 1. Pipeline used to process the corpora.

(described in 4.1) collected promising lemmas from corpora of documents sampled from Wikibooks, Wikipedia, and website privacy policies. A manual labeling procedure (in 4.2) resulted in synset labels agreed upon by multiple annotators.

4.1 Processing Pipeline

An eventual goal of this research is to link CA references with their referents, and a processing pipeline was constructed to retain document features which enable that task. Although CA reference-referent linking is not a contribution of this paper, we discuss a pipeline that enables CA inventorying for two reasons. First, it illuminates the procedure used to collect lemmas for sense labeling. Second, it shows a method for preserving valuable information on orthographically-structured (non-discourse) CAs in web documents while processing text. Such information is generally discarded by text processing pipelines.

Figure 1 shows the stages of the pipeline. The input consists of corpus documents in an HTML format (or if HTML is unavailable, plaintext). Documents are processed by a Markdown converter written by Gruber and Swartz (2006), which preserves the orthographic organization of the text while simplifying the document to the extent that it can (if desired) be read as plaintext. For example, items such as titles, sections, lists, tables, and block quotations are shown in the output of the Markdown converter using ASCII symbols (e.g., asterisks for bullet points, hashes around section headers), but all HTML is removed. Inventorying the orthographically-structured CAs then becomes a simple matter of parsing Markdown syntax and recording character indices where each CA begins and ends. This approach avoids the construction of a much more

Statistic	Privacy Policies	Wikipedia	Wikibooks
Documents	1010	500	149
Words	2646864	720013	5429978
Cand. Phrases	34181	2371	47546

Table 2. Statistics on each of the three corpora.

complex parser to directly handle the variability and complexity of CAs represented in HTML.

After conversion to Markdown, boilerplate text is discarded and the remaining passages are part of-speech tagged and parsed using Stanford CoreNLP (Socher et al., 2013; Toutanova et al. 2003). Candidate phrases for CA reference are then identified using dependency templates. These templates identify noun phrases beginning with demonstratives *this*, *that*, *these*, and *those*; such phrases were identified as fertile for CA reference in previous work. Two more templates, noun phrases containing *above* and *below*, were new to the present work. From the candidate phrases, candidate CA-referential nouns were gathered, lemmatized, and ranked by frequency.

The prior study noted an informal correlation between lemma frequency in the candidate phrases and fertility for CA reference; however, it remained unclear whether less frequent CA-referential lemmas would have different qualities. For that reason, and because labeling word senses for *all* candidate nouns was infeasible, lemmas were sampled in two ways for further examination. The first was a "high-rank" sampling of the most frequent lemmas, continuing down the ranks until the selected lemmas were collectively responsible for at least 200 synsets. The second was a smaller "broad rank" random sampling of 25% of the 100 most frequent lemmas. Care was taken to avoid any overlap between the broad rank and high rank lemma sets.[3]

Table 2 shows descriptive statistics for each of the corpora. Documents were selected for inclusion in the corpora on the following bases:

- **Privacy Policies (PP)**: a corpus collected by Liu, et al. (2014) to reflect Alexa's assessment of the internet's most popular sites
- **Wikibooks (WB)**: all English books with printable versions
- **Wikipedia (WP)**: random English articles, excluding disambiguation and stub pages

[3] The procedure differed slightly for Wikibooks. Its high rank sample consisted of the 27 most frequent lemmas, whose 200 synsets were labeled by the prior study. Those labels are reused in the present work.

Collecting Artifact References

Phrase templates can be used to retrieve many references to communicative artifacts (CAs).

Category	Examples
Structural	Many of the resources listed elsewhere in **this section** have…
	In **this chapter**, we will show you how to draw…
Illustrative	Consider **these sentences**: [followed by example sentences]
	[following a source code fragment] …the first time the computer sees **this statement**, 'a' is zero, so it is less than 10.
Discourse	Utilizing **this idea**, subunit analogies were invented…
	In **this case**, you've narrowed the topic down to "Badges."
Non-Artifact Reference	Devices similar to resistors turn **this energy** into light, motion…
	What type of things does a person in **that career field** know?

Shomir Wilson and Jon Oberlander. Determiner-established deixis to communicative artifacts in pedagogical text. In Proc. ACL 2014.

A Word Sense / Ontology Problem

Word senses separate artifact references from other kinds of references.

There are many words for artifacts(!).

Goal: discriminate between synsets (word senses in WordNet) that refer to CAs from those that do not.

Accomplishments and Work in Progress

Done: a supervised learning approach to discriminating CA and non-CA senses

- Senses gathered using vocabulary in candidate phrases from Wikibooks, Wikipedia, and privacy policies
- High recall, low precision

In progress / future work: student projects

- Linking artifact references to their referents
- Applications to dialog systems
- Applications to educational materials?

Potential WWBP Tie-Ins

Some general thoughts:

- Online privacy is about much more than secrets
  - How is control of personal information on OSNs related to happiness?
  - What are the effects of undersharing on OSNs?
- How do people discuss discussion on social media?
  - Is it a common thing?
  - Can it tell us anything about what drives a discussion or how people feel about it?

Shomir Wilson

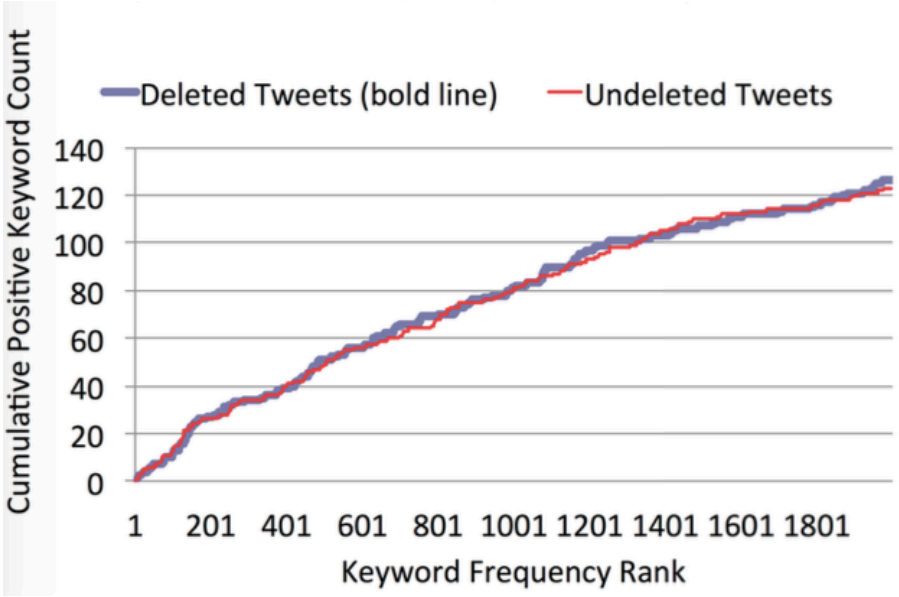
shomir@cs.cmu.edu

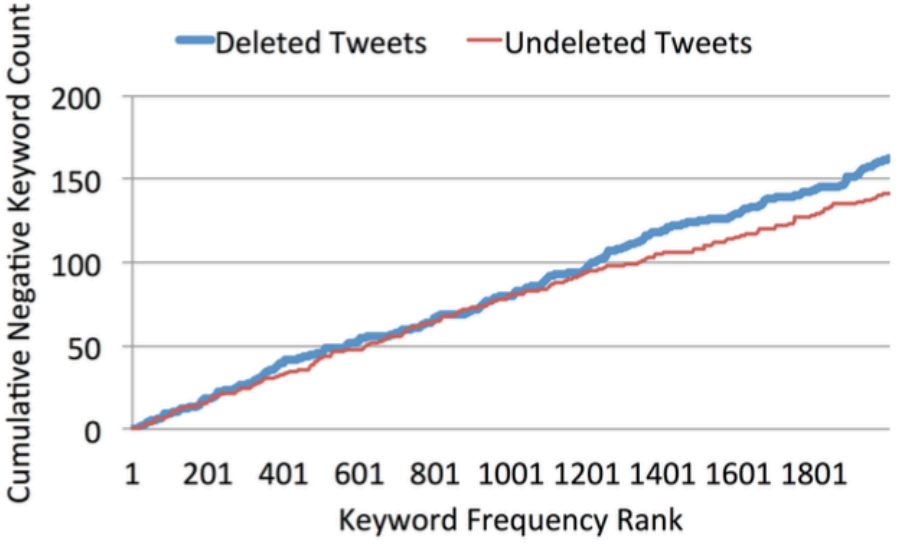
http://www.cs.cmu.edu/~shomir

Appendix

Sentiment Differences

Keywords: AFINN-111 word valence list





Location Sharing Study: Wizards