# Obscurity by Design: An Approach to Building Privacy into Social Media

**Fred Stutzman**
H. John Heinz III College
Carnegie Mellon University
Pittsburgh, PA
fred@fredstutzman.com

**Woodrow Hartzog**
Cumberland School of Law
Samford University
Birmingham, AL
whartzog@samford.edu

## Abstract
Responding to the privacy challenges of pervasive technologies such as social network sites, regulators propose that privacy protections be *built into* all phases of the software development lifecycle. The proposed solution, privacy-by-design, has been difficult to implement in practice. Therefore, we propose obscurity-by-design, an evidence-based, flexible set of principles designers can consider when building privacy in to interactional social media spaces.

## Author Keywords
Social media, privacy, privacy-by-design, obscurity

## ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Copyright is held by the author/owner(s).
*CSCW'12*, February 11–15, 2012, Seattle, Washington, USA.
ACM 978-1-4503-0556-3/12/02.

## General Terms
Human Factors, Design

## Introduction
As the beneficial outcomes of social media use are frequently linked to information seeking and sharing [5, 16], the effective management of privacy in social media is essential. Social media designers are challenged to construct systems and interfaces that promote interaction and disclosure while respecting individual privacy. Unfortunately, there is scant principled guidance for designers of social media privacy systems, and organizational reliance on heuristic-based approaches to privacy often backfire [e.g., 8]. The proposed solution for this problem, privacy-by-design, has been difficult to implement in practice [12]. Drawing on empirical and legal research, we propose a highly flexible, principled approach to social media privacy: obscurity-by-design.

## Privacy-by-Design
In recent years, privacy regulators have turned their attention to the accumulation of consumer data by companies, including social network sites (SNS). To enhance the privacy of these sites, regulators recommend that privacy protections be *built into* all phases of the development lifecycle [6, 7]. This approach, known as privacy-by-design (PbD), encourages companies to proactively address privacy

concerns so as to produce positive privacy outcomes for users [3]. Although well intentioned, PbD faces a number of challenges in implementation, including a lack of specificity in-principle, and weak market forces motivating adoption [12]. To date, applied PbD work has focused on back-end implementation principles, such as data minimization and security [9, 11, 13]. Very little work has focused on integrating PbD into the design of interfaces or interaction, which is critical for social media.

There are a number of practical reasons why PbD has avoided the interface. First, the translation of regulation to implementation is a complex process [9] and may be more efficient when applied to formal technologies (e.g., databases). Second, there is little guidance regarding how designers should approach the implementation of PbD in a contextually variant, interactional space. To address these challenges, we propose an approach to PbD in social media that is flexible and feasible to implement.

**Four Factors of Online Obscurity**

**Search Visibility**: Ease of discovery in search systems.

**Unprotected Access:** Degree of access restriction.

**Identification:** Degree to which individual is identified by direct or indirect disclosure.

**Clarity:** Ability for observer to comprehend discovered information.

## A Framework for Obscurity

Our proposed solution draws on the authors' research exploring privacy and identification in social media. In particular, we have explored how and why individuals shield aspects of their identity in online social interaction. Our research finds that much online interaction occurs in a state of *practical obscurity*, where information is not necessarily private, but is obscure enough to provide reasonable protection to the individual [14]. Motivated by this finding, we developed a four-factor definition of online obscurity (margin) that can be applied in policy and regulatory environments [10]. In this paper, we extend this definition into the design space, arguing that obscurity's four factors are a sensible, flexible way for designers to approach PbD in social media.

In our research, we find that obscurity is a common and effective approach to privacy and disclosure regulation in social network sites. We define obscurity as follows: Information is obscure online if it exists in a context missing one or more key factors that are essential to discovery or comprehension. We have identified four of these factors: 1) search visibility, 2) unprotected access, 3) identification, and 4) clarity. In hypothetical scenario 1 in figure 1, online information is strongly obscured by a blogger who only uses her or his first name (identification is limited) on the blog, she or he hides the blog behind a password or obscure URL (access protection) which prevents Google from finding an indexing the blog (search visibility).

## Toward Obscurity-by-Design

We propose that the four factors of online obscurity constitute a set of principles designers should consider when building privacy in to interactional social media

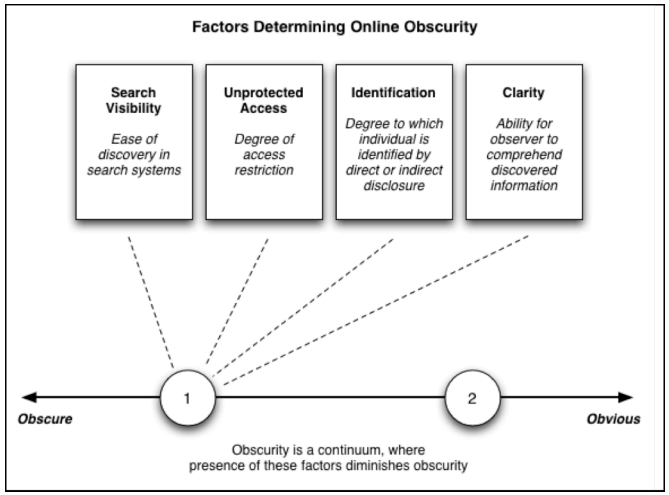


**Figure 1: Scenarios for online obscurity.**

spaces. This approach, which we term obscurity-by-design (ObD), is preferable to the design of PbD engineering requirements, as it provides designers a set of "starting points" which can be applied flexibly across culture and context. We now present suggestions for how each of the four factors of obscurity can be approached through design.

*Search Visibility*

Search visibility is the degree to which individuals, and the content they produce, are locatable and accessible through search. As search is a primary and common vector for discovery of individual content, designers should consider offering controls over inclusion in both internal and external search services. For example, some people may want their profile to appear in Google, while others would prefer to only be "searchable" one or two network degrees out (e.g., by friends-of-friends). Designers may also consider offering various levels of search engine obfuscation, where only certain aspects of the profile are placed into search, or search placement is manipulated to raise or lower placement of results.

*Unprotected Access*

Access protection covers the range of technologies and methods for controlling access to content. In SNS, access is generally governed by friendship status. In blogs or websites, access is often governed by credentials (such as passwords) or encryption. Consider the case where an individual would like to semi-privately share content without passwords or friendship connections. One could imagine a "paywall" that accepts links from certain sources restricting content access. While not perfect, this flexible approach could meaningfully enhance privacy while enabling selective disclosure.

*Identification*

Identification refers to the degree that individuals are identified through personal and interpersonal disclosures in online settings. SNS such as Facebook that have "real name" policies that require strong identification of site members through norms, and sometimes through enforcement action. These policies are controversial as the requirement of real names can disenfranchise a wide range of users (e.g., victims of abuse, political opposition) who face threats if they speak publicly with their "real names." We encourage SNS designers to support policies and technologies that allow for pseudonyms, name variants, and/or the use of multiple profiles to represent multiple facets of identity.

*Clarity*

Finally, clarity covers the degree to which an outside observer can *make sense* of content shared by an individual. If an individual is discussing insider information, we might assume that only a small proportion of the audience can actually translate the observed discourse into meaningful information. An example of this is social steganography [2], or the practice of hiding content in plain sight. This form of obfuscation is largely normative, where individuals disclose based on a "collective understanding of social contexts" [2, p. 17]. Designers can approach clarity by both recognizing and valuing individual strategies for managing clarity (i.e., respecting this normative practice in both policy and technology), and by considering the degree to which meta-data, data stores, and data recombination [1] allows outside individuals to programmatically *construct* clarity of observed information. Such considerations are especially important given the risks to persons (e.g., jobs security, safety) that can emerge from inadvertent disclosures.

## Implementation Considerations

ObD can be produced through technology, policy, and behavioral interventions ("nudges"). **Technologies** can include, but are not limited to, privacy-enhancing technologies (PETs) such as the option to hide individual content from internal and external search engines. **Policies** that allow for practices like the use of multiple profiles and pseudonyms, or those that prohibit obscurity-destroying practices such as scraping provide structural support for obscurity. Behavioral economics and social psychology provide instruction on how to "**nudge**" users toward obscurity-friendly practices [4, 15]. Gentle reminders like "showing" users the potential audience of their disclosure or reminding them their activity will be searchable by Google could promote a greater awareness of obscurity in the user.

In proposing ObD, we note the following important points. First, ObD is produced in site technologies and policies; the ObD approach necessarily brings together site policymakers with technical designers. Second, ObD's focus on interaction complements existing PbD scholarship and practice, which often is lower on the engineering stack [e.g., 9, 13]. Third, ObD's flexibility allows designers and policymakers to approach the four essential factors in context, so as to best fit ObD to the present culture or technology.

## References


[1] Acquisti, A., Gross, R. and Stutzman, F. Privacy in the Age of Augmented Reality. (Working paper).

[2] boyd d. and Marwick, A. E. Social Steganography: Privacy in Networked Publics. Paper presented at *ICA 2011*, Boston, MA, 2011.

[3] Cavoukian, A. *Privacy by Design.* Information and Privacy Commissioner of Ontario, Canada, 2009.

[4] Calo, M.R. Against Notice Skepticism in Privacy (and Elsewhere). *Notre Dame Law Review, 87* (2012).

[5] Ellison, N. B., Steinfield, C., and Lampe, C. Connection strategies: Social capital implications of Facebook-enabled communication practices. *New Media & Society 13*, 6 (2011), 873-892.

[6] European Commission. *A comprehensive approach on personal data protection in the European Union*, Brussels, 2010.

[7] Federal Trade Commission. *Protecting consumer privacy in an era of rapid change*. Washington, DC, 2010.

[8] Federal Trade Commission. *FTC Charges Deceptive Privacy Practices in Google's Rollout of Its Buzz Social Network*. Washington, DC, 2010.

[9] Gürses, S., Troncoso, C., and Diaz, C. Engineering privacy by design. Paper presented at *CPDP 2011*, Belgium, 2011.

[10] Hartzog, W. and Stutzman, F. The Case for Online Obscurity. Paper presented at *PLSC 2011*. Berkeley, CA, (2011).

[11] Kung, A., Freytag, J., and Kargl, F. Privacy-by-design in ITS applications. *Proc. IEEE WoWMoM* (2011), 1-6.

[12] Rubinstein, I. Regulating Privacy by Design. *Berkeley Technology Law Journal,* (2012).

[13] Spiekermann, S. and Cranor, L. F. Engineering Privacy. *IEEE Trans Software Engineering 35,* 1 (2009) 67-82.

[14] Stutzman, F. and Hartzog, W. Boundary Regulation in Social Media. *Proc. CSCW 2012,* Seattle, WA.

[15] Thaler, R. and Sunstien, C. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale, New Haven, CT, 2008.

[16] Yoder, C. and Stutzman, F. (2011). Identifying Social Capital in the Facebook Interface. *Proc. CHI 2011*, ACM Press (2011), 585-588.