

## **Amazon is my hangout! Self-disclosure and community building in Amazon's reviews**

**Federica Fornaciari**

University of Illinois at Chicago  
Department of Communication, ESP-IGERT Fellow  
fforna3@uic.edu

### **Abstract**

This study investigates patterns of self-disclosure in Amazon reviews measuring the level of sensitive information that users reveal. Informed by Goffman's work on the "presentation of self in everyday life", this study investigates sensitivity of information by using the software Linguistic Inquiry and Word Count (LIWC). Using a crawler we collected 3,485 reviews from Amazon, for six products. Sensitivity was measured through five main LIWC categories that include the following: pronouns, social processes, affective processes, biological processes, and personal concerns. As a result, Amazon reviewers in the sample collected disclosed significantly higher levels of sensitive information in these categories: family, humans, affect, positive emotions, negative emotions, sadness, cognitive mechanisms, concerns related to work, achievements, leisure and money. Results seem to suggest that users experience Amazon as a community built around people, whose participants often show their humanity, their offline social circles, their affective processes, their emotions, and their concerns.

### **Keywords**

Amazon reviews; online community; self-disclosure; sensitive information; textual analysis.

### **Introduction and relevant research**

With the introduction of online platforms and the gained popularity of user-generated content, individuals increasingly share personal information online (Jarvis, 2011; Li & Chen, 2010). Doing so, they expose different layers of the self, sharing them with "imagined communities" (Acquisti & Gross, 2006). The information posted online has various degrees of sensitivity and reveals different nuances of one's persona, contributing to refining one's identity (Fornaciari, 2012; Mehdizadeh, 2010; Mendelson & Papacharissi, 2011; Papacharissi, 2002). The current project explores whether and how individuals engage in self-disclosure when posting consumer reviews on Amazon. Doing so, it contributes to understanding the relationship between social practices and online platforms. It also furthers the comprehension of how presentation of self and participation are implemented in online retailing websites.

Sociality and presentation of self (Goffman, 1959) are performed across multiple audiences. Online platforms enhance such a networked sociability. For many, disclosing online has become key to enjoying the therapeutic benefits of social connection by fostering socialization, online community, and access to social capital (Andrejevic, 2011; boyd, 2011; Jarvis, 2011). As a consequence, the depth of disclosure is often positively correlated to the breadth of perceived opportunities for sociality, intimacy, community, social support, and customized information (boyd, 2006; Debatin et al. 2009; Ellison et al., 2011a; Ellison et al., 2011b; Johnson et al., 2011; Kramer & Haferkamp, 2011; Livingstone, 2008; Manen, 2010; Walther, 2011). In addition, research reveals that self-disclosure is positively correlated to sense of belonging in a community, desire for reciprocity, and trust; whereas it is inversely correlated to perceived privacy risks (Porter & Donthu, 2008; Posey et al., 2010). In sum, the process of disclosure entails an ongoing negotiation to attempt balancing privacy and publicity, thereby achieving desired levels of sociality (Altman, 1977).

The current research project investigates the patterns of self-disclosure on Amazon reviews addressing the following questions:

RQ1 – Is there an observed relationship between the disclosure of sensitive information and the use of a real name?

RQ2 – Is there an observed relationship between the disclosure of sensitive information and the disclosure of one's location?

RQ3 – Is there a relationship between type of reviewer and sensitivity of information disclosed?

## **Method**

The data collection that was operated through a crawler launched in amazon.com website in November 25<sup>th</sup>, 2012. The process generated 3,485 .txt files of reviews for six products (baby product, anti-aging, fragrance, electronic, sex-related, and weight loss). Each file included: text of review, length of review, location, and number of reviews posted. It also included the following Amazon's Badges: real name, top reviewer, hall of fame reviewer, vine voice, verified purchase.

The reviews were processed through the software LIWC (Linguistic Inquiry and Word Count) (Tausczik & Pennebaker, 2010) to measure multiple variables used as indicators of sensitive information. In particular the current study considered the "level of sensitivity" of information as a multidimensional variable. LIWC allowed measuring the percentage of words belonging in each of the following categories: use of pronouns (self or other-directed); social processes (social, family, friend, humans); affective processes (swear, affect, positive emotion, negative emotion, anxiety, anger, sadness, cognitive mechanisms); biological processes (biological, health, sexual); and personal concerns (work, achievements, leisure, home, money, religion, death). Data processed through LIWC were further analyzed using descriptive statistics, discriminant analysis, and ANOVA. Finally, the levels of sensitive information disclosed on Amazon were compared to the average levels of sensitivity of information disclosed (Pennebaker et al., 2007).

## **Discussion and conclusion**

Amazon reviewers in the sample collected, compared to Pennebaker and colleagues (2007), revealed higher levels of sensitive information in a number of categories. These included: family, humans, affect, positive emotions, negative emotions, sadness, cognitive mechanisms, concerns related to work, achievements, leisure and money. Amazon is understood as a venue built around people, whose participants often show their humanity, their offline social circles, their affective processes, their emotions, and their concerns. Previous research on online platforms for user-generated content suggests that users post information to foster their status and gain popularity within a community and to increase self-esteem through social comparison (Harper et al., 2007) Also, research reveals that people adapt their behavior to the perceived social norms of a community. Thus, for example, users who read sensitive content in existing reviews may be more willing to disclose personal information as well (Frey & Meier, 2004). These are possible explanations of why Amazon users engage in detailed self-disclosure. In addition, research on social capital shows that sharing information is often a necessary step to access social capital and foster participation (boyd & Ellison, 2008). Consistent with the findings of research addressing participation in SNSs (e.g. Acquisti & Gross, 2006), the current study reveals that Amazon reviewers often behave as active members of a community. It is possible that sense of belonging and participation positively influence users' trust and encourage them to perceive disclosure as a vehicle to develop their identity. In addition, increased disclosure may be motivated by increased familiarity with the amazon community. Research, in fact, suggests that familiarity with social media enhances trust, encourages disclosure, and decreases the perception of privacy risks (Fogel & Nehmad, 2009).

Combining the current findings with existing literature on self-disclosure, one may hypothesize that disclosure on amazon is also motivated by desire of building community, developed trust, and sense of belonging. In fact, individuals who disclose their identity on amazon (real name and/or location) tend to disclose significantly more information that reveals their social processes, shows their sadness, discusses biological and health processes and tackles concerns related to personal achievements. Also, data analyzed in the current study emphasized that "normal reviewers" (as opposed to professional reviewers) consistently tend to share higher levels of sensitive information thus increasing the personal participation in the amazon community. Another way to gauge the level of participation in the amazon community was through the number of reviews posted by a reviewer. Self-disclosure tended to increase for frequent (yet non-professional) reviewers who actively post reviews for the products they purchase, and have a sense of belonging in the amazon community. With some exceptions, the level of

sensitive information disclosed tends to progressively decrease very active reviewers and professional reviewers. This suggests that reviewers that post between 11 and 50 reviews are those who perceived themselves as elements of a community, who want to share their insights and contribute to the overall satisfactions of the other members and, perhaps, seek social support by disclosing concerns. Sense of belonging may be the discriminant element that encourages individuals to value participation and seek social capital as a benefit that outweighs the risks of privacy loss.

Admittedly, findings were limited by the fact that our sample included few reviewers belonging in the categories Top, Hall of Fame and Vine Voice. To address such a limitation, a future study could be conducted from a user's centered perspective (using the reviewer as unit of analysis) and collect reviews based on the use of badges. A comparison of equally sized groups of reviewers would allow a better assessment of these findings. Another limitation is that the motivations for sharing information in amazon were inferred from existing literature. Future research could implement a survey or a set of interviews to directly and empirically test the hypotheses that emerged.

### **Acknowledgments**

This material is based upon work supported by the National Science Foundation under Grant No. DGE-1069311.

## References

- Acquisti, A. & Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In P. Golle & G. Danezis (Eds.). *Proceedings of 6th Workshop on Privacy Enhancing Technologies*. Cambridge, U.K.: Robinson College.
- Altman, I. (1977). Privacy regulation: Culturally universal or culturally specific? *Journal of Social Issues*, 33(3), 66-84.
- Andrejevic, M. (2011). Social Network Exploitation. In Papacharissi, Z. (Ed.) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York : Routledge.
- boyd, d. (2006). Friends, Friendster, and Top 8: Writing community into being on social networking sites. *First Monday*, 11(12).
- boyd, d. (2011). Social Network Sites as Networked Publics Affordances, Dynamics, and Implications. In Papacharissi, Z. (Ed.) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York : Routledge.
- boyd, d. & Ellison, N. (2008). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13, 210-230.
- Debatin, B., Lovejoy, J. P., Horn, A-M., & Hughes, B. E. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication*, 15, 83-108.
- Ellison, N. B., Lampe, C., Steinfield, C., & Vitak, J. (2011a). With a Little Help From My Friends. How Social Network Sites Affect Social Capital Processes. In Papacharissi, Z. (Ed.) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York : Routledge.
- Ellison, N. B., Vitak, J., Steinfield, C., Gray, R., & Lampe, C. (2011b). Negotiating privacy concerns and social capital needs in a social media environment. In S. Trepte and L. Reinecke (Eds.), *Privacy Online*, Berlin: Heidelberg.
- Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, 25, 153-160.
- Fornaciari, F. (2012). *The language of technoself: storytelling, symbolic interactionism, and online identity*. In R. Luppiciini (Ed.), *Handbook of Research on Technoself: Identity in a Technological Society*. IGI Global.
- Frey, B., Meier, S.(2004). Social comparisons and pro-social behavior. Testing conditional cooperation in a field experiment. *American Economic Review* 94(5).
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday.
- Harper, F. M., Li, S. X., Chen, Y., & Konstan, J. A. (2007). Social comparisons to motivate contributions to an online community. *Persuasive Technology*, pp. 148–159.
- Jarvis, J. (2011). *Public Parts. How Sharing in the Digital Age Improves the Way We Work and Live*. New York: Simon & Schuster.
- Johnson, T. J., Zhang, W., Bichard, Z. L., & Seltzer, T. (2011). United We Stand? Online Social Network Sites and Civic Engagement. In Papacharissi, Z. (Ed.) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York : Routledge.
- Kramer, N. C., & Haferkamp, N. (2011). Online Self-Presentation: Balancing Privacy Concerns and Impression Construction on Social Networking Sites. In Reinecke, L. & Trepte, S. (Eds.). *Privacy online. Perspectives on Privacy and Self-Disclosure in the Social Web*. New York: Springer.
- Li, N., & Chen, G. (2010). Sharing location in online social networks. *IEEE Network*.
- Manen, M. (2010). The pedagogy of Momus technologies: Facebook, privacy, and online intimacy. *Qualitative Health Research*, 20(8), 1023-1032.
- Mehdizadeh, S. (2010). Self-presentation 2.0: Narcissism and self-esteem on Facebook. *Cyberpsychology. Behavior and Social Networking*, 13(4), 357–364.

- Mendelson, A. L. & Papacharissi, Z. (2011). Look At Us Collective Narcissism in College Student Facebook Photo Galleries. In Papacharissi, Z. (Ed.) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York : Routledge.
- Papacharissi, Z. (2002). The presentation of self in virtual life: Characteristics of personal home pages. *J&MC Quarterly*, 79(3), 643–660.
- Papacharissi, Z. (Ed.) (2011). *A Networked Self. Identity, Community, and Culture on Social Network Sites*. New York : Routledge.
- Pennebaker, J. W., Booth, R. J. & Francis, M. E. (2007) Operator's Manual. Linguistic Inquiry and Word Count: LIWC2007.
- Porter, C. E., & Donthu, N. (2008) Cultivating trust and harvesting value in virtual communities. *Management Science* 54(1), 113–128.
- Posey, C., Lowry, P. B., Roberts, T. L., & Ellis, T. S. (2010). Proposing the online community self-disclosure model: the case of working professionals in France and the UK who use online communities. *European Journal of Information Systems*, 19(2), 181–195.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54.
- Walther, J. B., Carr, C. T., Seung, S., Choi, W. DeAndrea, D. C., Jinsuk K., Tong, T., & Van Der Heide, B. (2011). Interaction of Interpersonal, Peer, and Media Influence Sources Online. In Papacharissi, Z. (Ed.) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York : Routledge.