# Impression formation in online peer production

Jennifer Marlow

CMU-HCII-14-110

September 2014

Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

## Thesis committee

Laura Dabbish, Carnegie Mellon University (chair)
Jodi Forlizzi, Carnegie Mellon University
Jim Herbsleb, Carnegie Mellon University
Sara Kiesler, Carnegie Mellon University
Scott Counts, Microsoft Research

Submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

# Abstract

Technology increasingly enables new forms of distributed and ad-hoc online collaboration, among them a form of large-scale, loosely coordinated, mostly voluntary work known as peer production. Some peer production environments now connect social media functionality directly to collaborative work artifacts, which provides participants with detailed information about other contributors' work history, interests and interactions.

Activity visibility provides informational signals individuals can use to make inferences about important characteristics of the people they interact with. However, the impact of the increased variety of information about collaborators and potential colleagues in a peer production setting is not well known.

Understanding how people form impressions of other contributors can inform the design of peer production environments. This thesis investigates the process and outcomes of using activity traces for interpersonal impression formation in online peer production.

I describe two interview studies with users of a social media enabled site supporting open source software development. The first study investigates what signals people use to form impressions about others' expertise and attitudes, and the second study identifies how they used this information to make decisions about work contribution acceptance. My results show that observers use cues that they see as reliable and easy to verify to draw conclusions about not only workers' abilities but also their underlying personal characteristics.

Finally, I present an experiment investigating how the visual presentation of activity history influences impressions of contributors and evaluation of work. In the study I vary the amount of detail and quality of work shown in an activity history and measure the influence on impressions and evaluation of work products. I find that greater detail enhances valence and persistence of initial impressions and bias towards an unknown worker as well as effort expended to correct the worker's output on the task.

My thesis advances our understanding of when and how social networking information and activity traces influence the process of making sense of unknown contributors' inherent qualities, and how this relates to work-related decision-making in peer production. The thesis also informs design principles for showcasing individuals' activity history in collaborative production sites.

# Acknowledgments

# Table of Contents

**Appendix: Visualizations of user activity history in GitHub and related sites** **153**

# List of figures

# List of tables

# 1 Impression formation in peer production

Peer production is growing as a form of online coordination and cooperation in which groups of loosely coordinated people interact to create a product (or service) of economic value, which they make available to contributors and non-contributors alike (Levine & Prietula, 2013). It is typically large-scale in nature, collaborative, and primarily voluntary (Algan, Benkler, Morell, & Hergueux, 2013).

Peer production, as defined and characterized by Benkler & Nissenbaum (2006), has some distinguishing characteristics as a form of creation and production:

- It involves the collective effort of individuals contributing towards a common goal in a more or less informal and loosely structured way
- No single entity "owns" the product or manages its direction
- No one is a formal manager who assigns projects to others

These features combine to create a setting in which medium-to-large scale collaboration and effort occur among a large and geographically dispersed group of otherwise unrelated participants.

Haythornthwaite (2009) distinguishes between two different patterns of participation: Lightweight peer production (e.g. crowdsourcing), in which people are relatively anonymous and independent in their contributions (for example, Smithsonian Volunteers, in which individual contributors transcribe text from historical images and documents), and heavyweight peer production (e.g. Wikipedia, open-source software development) in which involvement requires engagement with others and building internal norms along with coordinating contributions to a product as a whole.

Some benefits of the peer production setting, according to Levine & Prietula, (2013), are that people have the ability to build on others' work in a direct way and that they can interact directly with others, many of whom are strangers, to share and integrate knowledge and other resources.

Another feature of the peer production setting is that individuals are able to self-identify for tasks that attract them and for which they are suited (Benkler &

Nissenbaum, 2006).  While this increases the chances of receiving contributions from a wide pool of contributors, it also increases the uncertainty about any one individual's skills and qualifications.  In the case of open source software development, (Stewart, 2005) highlights that this creates a setting where "uncertainty about the quality of others is the norm rather than the exception".

Impression formation is important but a challenge in such a setting where participants come and go freely, contributing as they see fit, and mechanisms for peer review of contributions exist to help weed out mistakes.  Impression formation occurs whenever people come into contact with strangers for the first time, and in a large-scale peer production setting this may occur frequently.  Forming impressions of unknown others helps people make sense of each other and can significantly influence pre-interaction attitudes (Carr & Walther, 2014).

The many differences between members and the loose boundary around participation means that forming impressions of others' skills and attitudes given limited cues, time, and motivation, is not always easy.  When people do not fully evaluate and consider information about a person (for example, by only paying attention to some informational cues rather than others,) accuracy may suffer.  Inaccurate impressions can lead to conflict, mistrust, and problems with knowledge sharing (Johri, 2012). At the same time, cooperation, conflict resolution and knowledge sharing are critical in the more complex and knowledge based open collaboration settings like open source software or collaborative document creation (e.g. Wikipedia) where participants are more interdependent and need to cooperate and build on each other's work to create something meaningful (Zhu, Zhang, He, Kraut, & Kittur, 2013).

In heavyweight peer production settings, impressions of others feed into evaluations of contributors and their work, not only in terms of the quality of the work itself but also in terms of interpersonal behavior, helpfulness and support (Haythornthwaite, 2009). The implications of the peer production setting for the process of impression formation (in terms of goals and information used) may be distinct from other, previously studied settings due to the nature of the work.

The majority of previous work on online impression formation has focused on personality or expertise assessments in the context of Facebook or social networking sites and consisted of laboratory-based experiments where participants were exposed to and rated profiles of strangers in a simulated task (e.g. Gosling, Gaddis, Vazire, et al., 2007; Carr & Walther, 2014; Baruh, Chisik, Bisson, & Senova, 2012). Less is known about how impression formation unfolds in an online cooperative work setting.

Cooperative work settings have important differences from more recreational or purely social online environments that influence both the process and outcome of impression formation. For example, unlike impression formation in online dating, participants may or may not meet face to face (may or may not interact again). There may therefore be less of an emphasis in peer production on determining whether or not someone's self-description matches up with what they are actually going to look like in person. Unlike expertise location in the enterprise, participants in online peer production may come from a variety of backgrounds and lack the common ground of belonging to the same organization. Therefore, some level of standards or common practices cannot be assumed.

## 1.1 Social transparency and available information

While the details of the work involved in peer production differ across domains, their enabling websites are commonly focused around the creation and sharing of creative, digital artifacts (e.g. software code, digital text like Wikipedia articles or transcriptions of images and documents). These sites can provide insight into individuals' contributions and past behavior in the following ways:

- Work portfolio: Each person on the site has a personal profile that serves as a portfolio of all artifacts they have created or projects they have worked on.
- Activity traces: Personal profiles also contain a record of automatically-generated activity traces that make public actions people have taken around their own and others' artifacts, and reveals interpersonal interactions in the form of comments and discussions engaged in with others.
- Social network information: Connections between members in the form of followers and following are made explicit. It is possible to see both the number

of people a person is following and followers they have received on their profile. Further details of who they interact with are also visible in the activity traces (see above).

Each of the three sets of information listed above provide insight into what Stuart, Dabbish, Kiesler, Kinnaird, & Kang (2012) refer to as identity transparency (visibility of the identity of the site member), content transparency (visibility of the origin and history of actions taken on information), and interaction transparency (visibility of the information exchanges that occur between the member and others).

## 1.2 Implications for impression formation

The visibility of activity history and activity traces in peer production may have implications for the way in which work is done in peer production. When work is carried out online via a platform that records and renders individual interventions in the common project, it can facilitate discourse among members with regards to how their contributions do or do not advance the common enterprise (Benkler & Nissenbaum, 2006). Activity traces can also support what Halfaker, Geiger, & Terveen (2014) refer to as "social literacy", which can help people coordinate activities and support the development of practices and norms through observation. In settings like Wikipedia, being able to understand and interpret the traces left by others can help people not only see what events have occurred but also to observe and mentor newcomers.

Traditionally, the amount of detailed information provided about a person's past contribution histories in peer production has been limited. For example, Wikipedia profiles are intentionally very sparse, in an attempt to create a level playing field (Tausczik & Pennebaker, 2011). However, this lack of detail makes it difficult to learn from and leverage a person's past history to understand what they might be good at without directly interacting with them.

With regards to the role of activity traces in supporting interpersonal awareness, there is evidence from other peer production settings like Wikipedia and Newgrounds (a collaborative animation site) that personal attitudes are also important to assess (Derthick et al., 2011; Kurt Luther, Caine, Ziegler, & Bruckman, 2010). Thus, the ability

to see not only the work that people do but how they interact with others could both feed into an overall impression of a person's characteristics.

In open collaboration settings, members can access a great deal of information about other participants, including their contribution and communication history with other members. On one hand, this transparency of collaborations past may provide better information about how others behave before they interact directly. On the other hand, rich activity histories for every potential collaborator could lead to information overload and be difficult to wade through or make sense of efficiently and effectively.

We therefore might expect that additional social transparency and activity traces may be useful to individuals working in large-scale peer production environments, but little research as of yet has focused on how this information is used or what specific cues are most utilized to gather information and reduce uncertainty about unknown others.

This thesis contributes to social and psychological theory by advancing our understanding of the role of signaling theory (particularly signal evaluation effort) on interpersonal impressions in the relatively new context of heavyweight peer production. It also provides new insight into the relationship between impressions (attitudes) and behaviors in a peer review task, thus expanding our understanding of the positivity-negativity asymmetry in a lightweight peer production task.

This work also contributes to the field of HCI by systematically exploring and evaluating the role of online information presentation on psychological and behavioral outcomes. These insights can be used by designers of systems that display activity trace information to understand the potential tradeoffs and implications of presenting information in varying degrees of granularity and visual abstraction.

# 2 Overview of thesis work

To understand impression formation in online peer production environments equipped with personal profiles and activity traces, my thesis addresses the following research questions:

- How does impression formation happen in a peer production setting?
- How do impressions affect attitudes and behaviors towards unknown others?
- How does the design and display of individuals' peer production activity traces influence the impressions that are formed and their consequences?

In order to answer these questions, I have conducted a series of qualitative, interview-based studies in the context of a community to support open source software production and followed these up with an experiment to empirically test the effects of activity trace designs on impressions.

The qualitative studies are described in Chapters 5 and 6. I first examined when individuals proactively seek out additional information about the collaborative history and interaction patterns of other members of open collaboration communities. I found that they do so for three primary reasons: 1) to discover and explore the collaboration space relevant to their own work, 2) to assess and evaluate members' suitability for recruitment as a project contributor or a full-time company employee, and 3) to inform collaborative interactions around specific project artifacts and contributions.

My qualitative findings suggested that the visual design used to depict activity traces may influence impressions through various cognitive processes. Peer production participants described a variety of information-seeking processes that involved attending to signals that were easy to assess and interpret, as well as briefly or quickly attending to one or two recent pieces of detailed information. Understanding how design decisions have the potential to influence impressions' valence, accuracy and resistance to change is also of interest to those involved in the design of future socially transparent work sites.

This work suggested the important role of access to and interpretation of signals about others' work history on how people view and evaluate a specific piece of work. To further understand the interplay between impressions and work evaluations, I then conducted an experiment (see Chapter 9) manipulating display and quality of activity information to empirically test how the visual display of activity history information influences impressions of unknown workers' quality and editing behaviors of their work.

Findings from the experiment suggest that there are tradeoffs between presenting detailed textual displays of activity history and abstracted visual summaries of activity history, and that the quality of the work to be edited also plays a role. In our study, more detail in the traces induced positive first impressions of an unknown worker, but these positive impressions may have a detrimental effect on fixing bad work.

Overall, this thesis identifies the effects and implications of providing increased levels of access to information about individuals' past activity history and interactions in online collaborative work settings. My findings contribute to theory on impression formation by expanding our knowledge of interpersonal information seeking in an increasingly complex and detailed environment where perceivers seek information to achieve work-related goals. They give insight into what system-generated signals observers attend to in an information-rich environment under various collaborative and interactive conditions. The factors influencing signal choice and consequences of impressions may differ from other contexts because of the work setting in peer production (people are working together in a voluntary and ad-hoc fashion). I also provide design guidance to facilitate individuals' abilities to quickly and accurately learn about each other in an online collaborative setting.

# 3 Related Work

Because uncertainty about others is typically high in the large-scale, open, voluntary setting of peer production, people are constantly using whatever information is available or accessible to them in order to engage in the process of forming impressions of unknown others. I define impression formation here as the formation of a coherent cognitive representation of a target person, based on information acquired about that person and the inferences drawn from it (Hamilton, Katz, & Leirer, 1980).

Impression formation is an active process in which a perceiver organizes all the information available about a target person in order to form a coherent mental representation of him/her (Hamilton et al., 1980). When people encounter a stranger for the first time, they make judgments that rely on their impressions of this person and use this to inform their behavior towards the stranger, whether it is assigning them work, asking them for assistance or expertise, or deciding to interact with them socially (Johri, 2012).

Interpersonal impression formation shapes attitudes and behavior in a variety of settings ranging from online dating to friendship formation to e-commerce. Across these settings, people use available information to decide what others are like and how to behave towards them.

The information available for forming impressions has changed over time, first as computer-mediated interactions emerged as an alternative to in-person encounters, and then as the nature and quantity of information available about a person online expanded. In this section, I first outline impression formation as a form of uncertainty reduction. I then review models of impression formation in offline contexts and examples of work applying them in several relevant online contexts, including peer production.

## 3.1 Uncertainty reduction

People form initial impressions when they are exposed to a person for the first time. Impression formation is generally seen as a means of reducing uncertainty and ambiguity, and involves gathering information that allows the seeker to make sense of others' attitudes, intentions and behavior (Antheunis, Valkenburg, & Peter, 2010; Tanis & Postmes, 2003). The goal is to make others' behavior more predictable and to provide explanations for such behavior (Gibbs, Ellison, & Lai, 2011). Impression formation helps people predict and explain others' behavior, which can be useful in situations when people have little prior experience working with each other.

Reducing uncertainty is also closely linked to trust. By being able to predict how someone will behave, the evaluator can determine whether or not the target will do something beneficial or detrimental before the evaluator can really know for sure (Lockwood & Massey, 2012).

According to Fiske, Cuddy, & Glick, (2007), there are two primary dimensions of social cognition upon which people regularly assess each other: competence (perceived ability, skill, intelligence, etc.) and warmth (perceived intent, helpfulness, trustworthiness, etc.) Therefore, the process of impression formation involves assessing and reducing uncertainty both about what a person is good at and what they will be like to interact with. The relative importance of these dimensions may depend on the setting and the perceiver's motives for impression formation. For example, in situations where interpersonal interactions are of high importance (e.g. online dating, finding friends on social networking sites,) evaluating personality and attractiveness are paramount (Antheunis et al., 2010) whereas in online Q&A forums, perceived expertise is important in assessing the worth of responses (Gazan, 2010).

Research conducted in other peer production settings, such as Wikipedia, has examined impression formation in admin promotion decisions. This work suggests editors attend to history of interaction and work on the site in order to make decisions regarding promoting people to editor status on the site and allowing them greater control and responsibility over articles. The information that informs these decisions includes evidence of civil interactions with others, social networks and counts/types of edits

made (Derthick et al., 2011). Similarly, (Luther et al., 2010) found that people desired information about the quality of past work (through peer ratings) as well as individual characteristics (such as the personality of unknown collaborators) when deciding who to work with in an online animation creation community. However, the specific processes of using cues to evaluate people in these peer production settings (and the associated conclusions drawn from the cues) are not well known. Understanding how people currently use cues available to reduce uncertainty about others (and the consequences of the impressions they form) may reveal insights into when this process works well and why. It can also provide insights into how and when the impression formation process can be improved through displaying different types of signals in different ways.

## 3.2  Accuracy of impressions

An important issue when dealing with forming impressions of unknown others is to understand the degree to which the impressions are accurate (that is, do they correctly match the reality?) For example, in the presence of a very positive first impression, a person may disregard or downplay negative issues encountered later; potentially negative aspects may be generously overlooked. On the contrary, a negative first impression could lead to failure to revise the initial hypothesis, even in the presence of strong evidence to the contrary (Lindgaard, Fernandes, Dudek, & Brown, 2006).

Accuracy of impressions is important in the peer production setting because there may be consequences to incorrectly assessing people's skill level and attitude. For example, incorrectly thinking a novice contributor is skilled could damage the quality of a collaborative project if the novice's contributions are not adequately reviewed or scrutinized. Similarly, incorrectly thinking that a skilled contributor is a novice could cause the contributor to become offended at being treated as such (see, e.g. Zhu et al., 2013).

There are multiple ways of conceptualizing accuracy when it comes to impressions. According to Funder (1995), accuracy when judging the personalities of others stems from relevance, availability, detection and utilization of behavioral cues. That is, first, a

cue must be a good representation of the quality being evaluated, and second, it must be noticed by an observer in order to result in accurate impressions being formed.

Letzring, Wells, & Funder's (2006) four-stage accuracy model, similarly, states that for impressions to be accurate,

1. The target must display behavior that is RELEVANT in a

2. Context AVAILABLE to judge (that is, information is available and accessible.)

3. The judge must DETECT cues and

4. Correctly UTILIZE them to make a judgment

Failure at any of these four stages can lead to inaccurate impressions being formed.

Accuracy can also be defined in various ways. For example, it may be some sort of correspondence between a judgment and a real-world criterion, or it may be defined as consensus, or agreement among many observers (Kruglanski, 1989).

Research on accuracy of impressions suggests that in some cases, people are surprisingly good at forming accurate first impressions based on very little information. For example, literature pertaining to "thin slicing" suggests that people can form impressions within a very short time (around 10 seconds) that are as accurate as those formed with more information and over longer time (Ambady, 2010). Thin slicing behavior has been found in both offline settings (e.g. watching lecturers give a talk (Ambady, 2010); judging personality based on offices or screen names - (Gosling, Ko, Mannarelli, & Morris, 2002; Graham & Gosling, 2012) as well as in online settings where people are able to assess others' personalities via information on social networking profiles (Stecher & Counts, 2008).

At the same time, however, the perceiver's motivations to form an accurate impression are also instrumental in the ultimate accuracy of their conclusions. When motivation is low, information is scarce, or cognitive capacity is constrained, individuals may simply base their attitudes on the heuristic cue best suited to achieve their accuracy goals (Chen, Shechter, & Chaiken, 1996). For example, people may use category-based

impression formation or stereotyping as a potentially biased strategy to limit their cognitive load and increase cognitive efficiency (Neuberg & Fiske, 1987).

This is the basis for the heuristic-systematic view of impression formation, which suggests a dual-process model for impression formation (Chen et al., 1996). Heuristics are mental strategies or shortcuts used to process information. Particularly within information-abundant environments such as the web, heuristic (as opposed to systematic) cognitive processing is a common means of coping with information overload and uncertainty (Metzger, Flanagin, & Medders, 2010). They may be efficient but at the same time lead to incorrect conclusions or the presence of biases.

The confirmation bias is one such example of a situation in which individuals' first impressions color subsequent reactions. This bias refers to people's tendency to seek information that they consider supportive of existing beliefs and to interpret information in ways that are particular to these hypotheses or beliefs (Nickerson, 1998). A related concept is the primacy effect, where information utilized early in the process influences opinions, and subsequently acquired information is interpreted in a way that is partial to that opinion (Nickerson, 1998). Negative and positive first impressions may both be confirmed and resistant to change, but some evidence suggests that negative first impressions are particularly persistent. This phenomenon is known as the negativity bias (Kanouse, 1984).

When information about another person is being displayed online, choosing what information to display and how could potentially enhance or suppress some of these biases (such as the confirmation bias and negativity bias). Therefore, understanding how the design and display of information about people influences impressions can provide insights into whether or not biases can be enhanced or reduced by choosing what to display about others and how it is shown.

## 3.3 Signaling theory and accuracy

### 3.3.1 Principles of signaling theory

One factor that influences the accuracy of impressions is the nature of the information that perceivers seek out. Signaling theory gives some insight into why certain pieces of

information may be more reliable (or viewed as more reliable) by perceivers. According to this theory, we make assessments of others based on their visible characteristics and actions. These observable cues effectively act as "signals" of hidden qualities (such as experience or expertise) that are not directly observable (Donath, 2007).

Signaling theory provides a useful framework for understanding what makes people perceive something as accurate or reliable in the process of impression formation. It attempts to predict what makes certain cues viewed as more or less reliable indicators of an underlying quality. It does so by focusing on an analysis of the costs and benefits of producing or interpreting a given signal. For the signal producer, the costs of deceptively producing a signal must outweigh the benefits of trying to fake it in order for the signal to be reliable (DeAndrea, 2014).

This notion of costs and benefits feeds into two main types of signals in terms of how they are produced and interpreted: *Assessment signals* are thought to be more reliable indicators of the presence of a certain quality because they are costly to produce. They also typically require the individual producing the signal to actually possess that quality as well. *Conventional signals,* on the other hand, are less costly to produce and thus are more easily faked by someone not possessing the underlying quality they signal. With a conventional signal, the link between the signal and underlying quality is more arbitrary, and a matter of social convention (Donath, 2007). Therefore, in order to actually be signals (not just noise), conventional signals require a norm and/or enforcement mechanism.

For example, being able to lift a heavy weight is an assessment signal of someone's strength while simply wearing a Gold's Gym t-shirt would be a conventional signal of this underlying quality (something that can easily be acquired and worn even if the wearer is actually quite weak) (Lampe, Ellison, & Steinfield, 2007). If people tend to ridicule weak people wearing a Gold's Gym t-shirt, for example, wearing one would be a conventional signal, enforced by a community norm.

Costs and benefits of producing or relying on signals exist for both senders and receivers. On the sender side, costs are associated with the effort to produce a signal as

well as the potential consequences of being caught producing a misleading signal (for example, getting fired) (DeAndrea, 2014).

Receivers also face costs in terms of the amount of time and effort it takes for them to evaluate the veracity of a signal. They must undergo the process of translating signals into perceived meaning (Connelly, Certo, Ireland, & Reutzel, 2011). Two factors determine whether a receiver will find a signal useful for assessing a target. First, they look for evidence of signal fit (that is, whether the signal corresponds with the quality of the signaler they are trying to assess). Second, they must also determine the signal honesty (attempt to which a signaler might be trying to deceive) (Connelly et al., 2011).

This notion of costs and benefits that come with signal interpretation effort may also explain why an observer may utilize a less costly but also less reliable signal (Donath & boyd, 2004). In this case, the observer's motivations and the stakes of making an incorrect impression play a role. For example, if the observer is only casually interested in the target, they may be satisfied with using a signal that is easier to fake but requires less effort to seek out or verify. However, if the costs of being mistaken are high, then it is worthwhile to expend effort to invest in the cost of the assessment signal for the sake of reliability (Donath & boyd, 2004).

In the online peer production context, individuals typically did not have access to many signals about the qualities of other contributors. They may have seen the e-mail address of a person submitting code to a mailing list, which would be a conventional signal in terms of giving off information about the quality of the person's contribution. In a richer online environment, having access to more cues about a person could give deeper insight into what they are able to do. The signaling value and interpretation of different cues might influence what observers attend to and why.

### 3.3.2  Signaling theory online

Signaling theory was originally applied to online impression formation in the early days of the Internet. In this case, people were aware that individuals could easily fabricate their identities or claims about themselves in this new medium and as a result many self-presentations were viewed with suspicion (DeAndrea, 2014).

In the online realm, signaling theory has been applied to examine how individuals in online communities attempt to convey and interpret visible cues about others as signals of their underlying characteristics (e.g. user name as a low cost or conventional signal of interest in the topic of the community). This work suggests that the degree to which a certain type of cue is viewed as reliable may depend on the context of the site: For example, the interpretation of one's number of friends or connections on a site as a signal of popularity may be more or less reliable depending on how costly it is to make a connection (Donath, 2007). An additional important aspect of evaluating others' signals is the amount of effort involved, or how easy it is for the observer to verify the accuracy of these signals (Mavlanova, Benbunan-Fich, & Koufaris, 2012).

Online, there are also a variety of different signals that can be available about a person. These include self-generated cues such as what a person chooses to post or disclose about themselves, other-generated cues such as comments made on a profile by friends on social networking sites, and system-generated cues such as automatically generated activity traces. The impact of these different types of signals, and how they might differ from traditional cues used in impression formation, is discussed in the following sections.

## 3.4  Information available for impression formation

The impressions people form of one another are closely tied to the information they have access to about that person. As a result, impression formation processes in face-to-face settings differ from impression formation in a computer-mediated context.

For example, in face to face contexts, impressions about others are formed based on observable factors such as physical appearance, behavior, personality traits, language use, social markers in speech, etc. (Hancock & Dunham, 2001).

With the advent of computer-mediated communication (CMC), access to these types of physical cues was initially removed. New theories developed to explain how impression formation might be affected when direct access to people's physical presence was removed and replaced by a less rich, text-based medium.

The cues-filtered-out perspective (Culnan & Markus, 1987) suggested that CMC's diminished capacity to convey emotional and personal information would lead to decreased awareness of others and depersonalized communication, overall inhibiting interpersonal relationship formation.  This perspective relates to impression formation by implying that by restricting the available cues, impressions formed should be relatively vague and unspecific compared to face-to-face interactions.

The Social Identity model of Deindividuating Effects (SIDE) (Lea & Spears, 1992) incorporated the role of social identity as well, and predicted that when given minimal cues about another person, people are forced to rely on the remaining social cues that are available (such as cues to role, status, etc.)  The scarcity of social and interpersonal information available are assumed to lead to more intense and exaggerated impressions of others, based on their perceived group membership.  If the target is judged to be in a similar group to the observer, they will form strong positive feelings towards that person.  If the target is judged to be in an outgroup, the stereotypical impression will be more negative.

Studies of early computer-mediated settings gave insight into what people concluded about others based on the information available to them.  For example, in the early days of message boards, people would look to a person's email address to give clues about what type of person they were (Donath, 1999).  Participants in virtual communities often base their images of players on stereotypes. In many instances, the stereotypes are conjured up by screen names (Jacobson, 1999)  or usernames in games (Graham & Gosling, 2012).  Later work looked at the impact of adding a photo, and/or a biographical textual description of a person (Tanis & Postmes, 2003) on ambiguity and positivity of impressions.  It found that the presence of either or both of these cues led to more positive evaluations.

With advances in technology and the rise of social networking sites, computer-mediated communication has continued to expand to provide more cues about individuals beyond what was typically available on message boards or other text-only settings.  For example, video chat is more common, providing new forms of media richness.  Profiles on social media sites allow people to showcase not only textual

descriptions of themselves and their interests but also combine this with photos and messages provided by third parties.

The advent of social media began to increase the number of cues providing information about a person and liberated communicators to seek information in new and unique ways going beyond what was traditionally available face to face (Ramirez, Walther, Burgoon, & Sunnafrank, 2002). With people having access to more information and more types of information, the question then arose about what information they would seek out and why.

## 3.5 Impact of new technology

Social networking sites that allow for increasingly detailed transparency make one's virtual communication and information generated by their peers more accessible than before (Walther, 2011). Prior to the widespread use of social networking sites like Facebook, impression formation was primarily based on things a user could choose to share or reveal about themselves.

Now, more than ever before, people have a wider range of information available with which to form impressions about others and can choose between sources of information for passive information seeking. This can include self-generated information such as personally chosen photos or self-written texts, but it can also include other-generated information (like photos uploaded or comments written by friends) or even system-generated information (like automatically-generated activity traces of one's actions.)

The impact of these new and varied sources is still not well understood, but initial work gives some insight into when people might rely on or seek out different types of information (e.g. self-generated, other-generated, and system-generated cues). Work by Walther et al. (2009) compared the effects of self-generated versus other-generated cues on impressions in Facebook. This study found that friends' comments were more influential than self-made statements in assessing physical attractiveness of Facebook users, because these were seen to have greater warranting value and be less subject to manipulation. In another study, SNS profile owners were perceived as more attractive when their friends were attractive and when their profile featured positive (as opposed to negative) wall postings from others (Antheunis & Schouten, 2011).

System-generated cues are perhaps more novel to appear as options for assessing and forming impressions of others. It stands to reason that information generated by a computer automatically should be perceived as more reliable than self-generated information (DeAndrea, 2014) because technically the person generating the information should have less control over what is displayed (or at the very least, they actually have to undertake some action in order for evidence of that action to be recorded). However, signaling theory doesn't account for system-generated cues. Signaling theory assumes to some extent that individuals have some control over what signals they give off (DeAndrea, 2014). When most or all of the information available about a person is automatically generated and displayed, then the cue reliability may be less important. Instead, observers may be more influenced by signal fit and ease of interpretation than signal cost. The importance of signal fit (or what types of cues are sought out) might also relate to an individual's information-seeking goals.

## 3.6 Goals in online impression formation

In addition to understanding what kinds of information people seek out about unknown others, understanding the context of *why* they are seeking it (including their goals and motivations for learning about another) can give insight into what cues are most valued by observers. Ramirez, Walther, Burgoon, & Sunnafrank, (2002) proposed a typology of strategies that people can use to pursue more detailed information about a target. These include interactive strategies that involve interrogating another person or disclosing information about oneself, active strategies (e.g. acquiring information without direct interaction,) extractive strategies (searches of online written content), and passive strategies involving unobtrusive observation (e.g. reading messages someone has posted on a d-list, reviewing buddy profiles.)

The increased transparency of an individual's actions and interactions online could have the potential to change the impression formation process. For example, visibility of a person's work activity history expands the range of passively-obtained information sources.

According to Ramirez's model, one important precursor to information seeking relates to the observer's goals in reducing uncertainty about a person. People may have social,

instrumental, or emotional goals to achieve with respect to others, and they will seek out relevant information that helps achieve that goal.

Put another way, in situations where an individual is uncertain, that individual attempts to reduce ambiguity about the situation by seeking info pertaining to that specific environment (Westerman, Van Der Heide, Klein, & Walther, 2008). Previous work on impression formation in different settings has addressed how goals relate to the type of uncertainty being reduced and the sources of information sought to help reduce that uncertainty.

Impression formation is important in situations that require trust like collaborative work (Rusman, Van Bruggen, Sloep, Valcke, & Koper, 2010). Impressions of a person are important in situations like online dating, recruitment, working together in an enterprise and assessing the trustworthiness of someone as an information source.

Most work on online interpersonal impression formation has focused on forming impressions of personality or credibility (trustworthiness and expertise) in various settings. In each of these cases, observers' goals for forming impressions and what they attend to depend on the nature of the context. Table 1 highlights a few of these settings and related findings from studies.

| Setting | Uncertainty reducing goals | Information used | Sample studies |
|---|---|---|---|
| Online dating | Determine whether or not someone's lying about who they are<br>Determine if someone is a good match | Profile pictures<br>Self-descriptions | (Gibbs et al., 2011) |
| Facebook | Assess a person's personality based on profile<br><br>Decide whether or not to befriend someone | Profile information (e.g. hobbies, political views, etc.)<br><br>Pictures<br>Self descriptions<br>Friends<br>Comments left by friends | (Antheunis & Schouten, 2011; Stecher & Counts, 2008; Van Der Heide, D'Angelo, & Schumaker, 2012) |
| Enterprise Q&A / Expert location | Assess someone's expertise<br>Determine likelihood of responsiveness | Profiles in enterprise social networking system | (Kim, Oh, & Oh, 2007; Shami, Ehrlich, Gay, & Hancock, 2009; Yarosh, Matthews, & Zhou, 2012; Guy, Steier, Barnea, Ronen, & Daniel, 2013; Jacovi et al., 2014) |
| Twitter | Determine whether or not someone is a credible source of news | Picture<br>Alias | (Q. V. Liao, Wagner, Pirolli, & Fu, 2012; Morris, Counts, Roseway, Hoff, & Schwarz, 2012; Jiang Yang, Counts, Morris, & Hoff, 2013) |

Table 1.        Summary of uncertainty reducing goals and information provided in different impression formation contexts online.

The goals and information available vary greatly across these different contexts so that findings from one setting may not apply to another domain.  For example, profile pictures are important in a dating context to assess the perceived attractiveness of a

target, whereas physical attractiveness is a less salient or relevant cue in an expert finding scenario.

As discussed in Chapter 1, peer production is a unique context because it is not tied to a specific organizational context, individuals may come and go and choose how to contribute, and because uncertainty about others' credentials and skills is often high. However, for those involved in open source software development, the context of GitHub is a new setting in which the degree of uncertainty about others may be mitigated through the information the site provides. In the next chapter, I discuss the site in more detail as the setting for Studies 1 and 2.

# 4 The GitHub context

Studying impressions in the context of peer production (specifically using GitHub as an example of a socially transparent work environment) is interesting because of the fact that work in this setting involves coming into contact with and interacting with unknown volunteer contributors. The activities in peer production are often interdependent and require managing both code and people. At the same time, profiles on GitHub are publicly visible to other observers who may not be directly collaborating with an individual but are nonetheless using this information to evaluate them.

GitHub can be used for multiple purposes. First and foremost, it is a tool to actively accomplish collaborative work. However, anecdotal evidence indicates that it is also used as a place to observe and evaluate job candidates. By investigating in-depth two different information-seeking scenarios that both occur within GitHub, I could see how individuals' goals and associated level of uncertainty impacted the cues that they sought out and interpreted.

## 4.1 Description of GitHub

GitHub is a software hosting website with over 6.4 million users hosting over 14.1 million repositories as of July 2014 [https://github.com/about/press]. GitHub's site design integrates social media functionality directly with code management tools. Two unique aspects of the GitHub environment set it apart from other open source software development environments in terms of personal information available about project contributors.

The first unique aspect of GitHub is the integration of social networking site features with an online work environment. GitHub incorporates many traditional functionalities of social networking sites outlined by (boyd & Ellison, 2007) such as profiles, articulated social networks of users and publicly visible connections between people and projects. Figure 1 shows a typical user profile on the GitHub site at the time the research was conducted.

Profiles on GitHub display a variety of information about a developer. An individual's profile contains a) biographical data (such as the date they joined the site and optional details about location, employer, etc.), b) a list of their projects in public repositories (including whether they own the project or forked it from another user, the languages, and a histogram of project activity), as well as c) an "activity feed" that displays the most recent actions they have performed on the site (forking projects, watching other users or projects, submitting pull requests, commenting on code, discussing issues, etc.) Finally, the profile also highlights d) the number of people that follow the profile owner as well as the coders and projects that the profile owner has elected to watch.
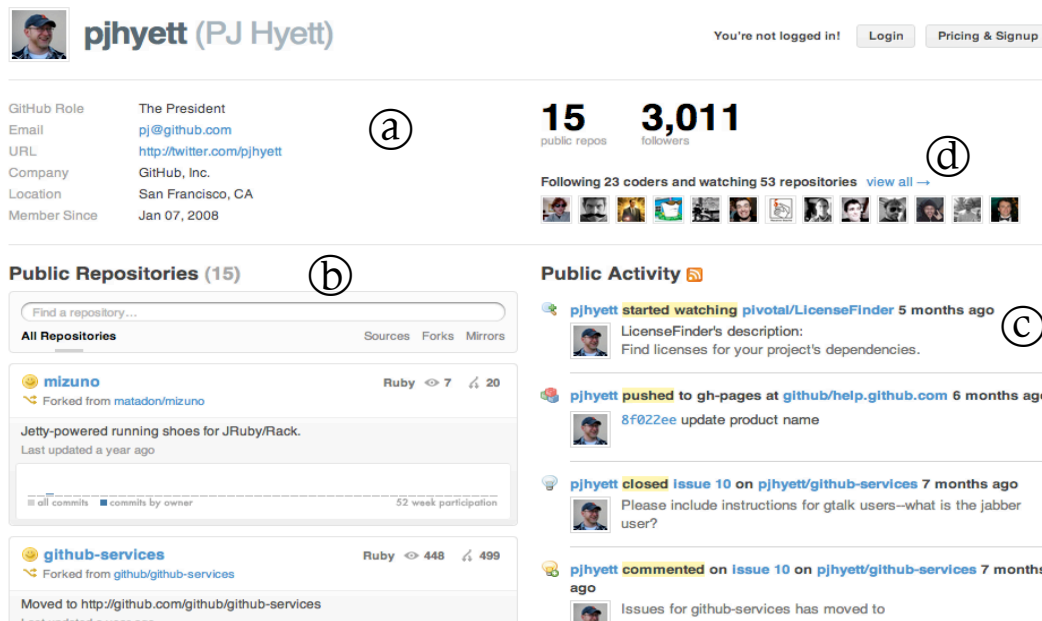


Figure 1.    Sample GitHub profile (profile features labeled a-d)

Similar to many other social networking sites, from the profile users can interact with other users (message them), view content they posted (their code repositories), or view an "activity stream" of their recent actions and behaviors. User profiles in other peer production environments (e.g. Wikipedia) do not include or link easily to user activity histories. GitHub is unique in terms of the visibility of activity and interaction history it provides compared to other software development and peer production environments more broadly.

The second unique aspect of the GitHub environment (after the integration of social media functionalities with work activities) is that coding is done much more publicly than in traditional software development organizations or other open source software platforms. Contributors can get involved in pre-existing projects by forking the project (i.e. creating a personal copy of the code), making changes to their personal copy, and then issuing a pull request to have their change merged back into the main branch of the project. The details of the work done by a contributor on the fork are visible to the project owner, making it easy to see what a contributor has done with the project. The low effort required to build on others' work is anecdotally associated with attracting more contributors to a project (Rao, 2012).

This accessibility can provide projects with many eager helpers, but comes with a complementary tradeoff that project owners have to evaluate a multitude of potential contributions before they are integrated into the master code base. At the same time, all code activity on the site is associated with a user's identity and public profile, meaning a code contribution can be vetted based on a developer's entire past history of contribution on the GitHub site.

## 4.2 Overview of GitHub studies

In our investigations of GitHub, I aimed to learn exactly when developers searched for more information about each other on the site, how they used this information to form impressions, and how it influenced their evaluation of other developers and the work they did. I conducted a series of interviews with people who used GitHub, focusing first on the question of when people had consulted the profiles of unknown others. This exploratory work (described in Marlow, Dabbish, & Herbsleb, 2013) identified the following three scenarios in which site members used profiles to form impressions of unknown others:

**1. Discovery**: Exploring new people and projects (for curiosity, keeping up to date)

**2. Hiring**: Recruiting and hiring software developers for full-time jobs

**3. Collaboration**: Assessing/learning more about unknown contributors to a project

Of these, the hiring and collaboration scenarios were of greatest interest because they were associated with decisions that had direct implications for future interactions among participants. To explore the hiring scenario further, I conducted a series of interviews with job seekers and employers using GitHub, described here as Study 1. The collaboration scenario, described in more detail in Study 2, examined in depth the role of profiles in evaluating unknown contributors in conjunction with their contributions.

# 5 Study 1: Using Profiles for Hiring

For the full paper describing this study, please see Marlow & Dabbish (2013).

Profiles that showcase an individual's public activities and interactions provide new, previously inaccessible insight into their work-relevant behaviors. The behavior visible on peer production site profiles can be used by external observers to evaluate and assess an individual's potential and suitability for a job. When activity traces are automatically generated and displayed by a system or work environment, individuals theoretically have less control over how they appear to others. Hiring and recruitment is an area where the activity made visible in social network enabled peer production environments could influence decision making for better or worse.

Anecdotally, the open source software community describes the traces provided by social media as affording more verifiable information about an individual's skills and abilities than a list of achievements on a resume. As John Resig, the creator of the jQuery interface library, recently tweeted: "When it comes to hiring, I'll take a Github commit log over a resume any day."
(https://twitter.com/jeresig/status/33968704983138304)

Social networking functionality, when tied with the work environment in a peer production site, provides moment-by-moment information about actions on artifacts and interactions around project decisions or activities. This means that information in these environments, compared to information on a resume, can provide a much more direct window into how someone works. In software development, for example, sites like GitHub allow employers to view the details of the code an individual writes in each commit, or contribution, to a project, and any interactions or discussions around the code. Potential employers can effectively reconstruct exactly what someone works on, how they work, what their code looks like, how they talk about their work or negotiate changes to collaborative projects, and their speed and style of work on public projects. This level of information about someone's working style is rarely available to a potential employer.

Additionally, as employers' interest in finding candidates via GitHub increases, developers may be attracted to participate in peer production environments such as open source software development as a means of building a publicly-visible portfolio that could help them get a job.

Research has not addressed how activity traces within online communities of practice play into the hiring process for more traditional jobs associated with the work of the community. Thus, we were interested in how employers in this environment made use of the displayed cues about developers' actions over time. This study focused on the following research questions:

> How are activity traces in an online peer production community used by potential employers to find and evaluate prospective software development hires?

> How do job seekers attempt to manage the impressions their activity traces give off to employers?

I conducted exploratory interviews with employers and job seekers in GitHub, an online open source software hosting repository with extensive social networking functionality integrated with the development environment. This means that potential employers can view an individual's profile of projects posted on the site, and see a history of their code related actions on these projects and other people's projects over time. As software developers commit changes to their software projects, these changes are broadcast to other developers watching the project. A history of commits (or contributions) to the code is recorded over time, along with conversation around changes in the form of comments. (See Dabbish, Stuart, Tsay, & Herbsleb, (2012) for a detailed description of the GitHub environment.)

I interpret the interview results through the lens of signaling theory to understand how and why certain cues were viewed as reliable signals of underlying characteristics of a potential hire. The results suggest that certain activity traces are used by employers more than others, in part because they serve as more reliable (assessment) signals of underlying characteristics that are often difficult to assess in traditional interviews, such as values and motivation, but also because they are easy to verify quickly. For example,

willingness to actively contribute to software and share projects openly on the site was viewed as a reliable signal of commitment to open source software ideology.

## 5.1 Impression formation and employment

The context of evaluating others for hiring and employment, while not directly a fundamental aspect of peer production work, is nonetheless interesting to explore because it is closely tied to a primary motivation for participation (Crowston, Howison, & Annabi, 2006) and it involves evaluating unknown others for future collaboration. Hiring is a high-stakes evaluation decision because in choosing an employee, one is in effect choosing a person who will be contributing to a company's revenue and work climate, potentially for a long amount of time (Weiss & Feldman, 2006). The cost of making a poor hiring decision can be higher than a bad decision about who can contribute to an open source project because a company can invest considerable time and effort in bringing a new person on board, and it is often hard to remove a poor employee.

When evaluating job candidates, either online or offline, a principal goal of the employer is to accurately assess applicants' job-relevant knowledge, skills, abilities, and other characteristics. These evaluations, in turn, affect selection decisions (Bohnert & Ross, 2010). For example, related work in the online peer production realm looking at admin permission granting in Wikipedia revealed that reviewers weighed evidence of interaction style, a candidate's social network, and the amount and type of past editing work when making these decisions (Derthick et al., 2011). Another factor influencing promotion decisions may be relative merit, or the comparison of how many edits a candidate has made compared to the person evaluating them (Leskovec, Huttenlocher, & Kleinberg, 2010). However, relative merit is based on very minimal cues such as number of edits or subjective criteria such as the relation of the voter to the candidate.

In addition to criteria relating to competence and expertise, employers also often place high importance on features that cannot be gleaned from a resume and need to be assessed in a job interview, such as a person's likeability and the potential person-organization fit (Weiss & Feldman, 2006). Other work suggests that other features that are hard to assess on paper such as whether a candidate is seen as "the right person for

a job" and perceived applicant-interviewer similarity can be strong determinants of hiring decisions (Kristof-Brown, Barrick, & Franke, 2002).

Increasingly, online presence on social networking sites (SNS) such as Facebook or LinkedIn are playing into the hiring process for full time jobs in the offline world. A 2008 survey (Grasz, 2009) found that information gleaned from SNS could damage an applicant's chances of being hired if it revealed that a person had lied about their qualifications. On the other hand, SNS information could be advantageous if it helped support their qualifications or portrayed a professional image. SNS members with personal profiles are aware that employers might look at their profiles (Bohnert & Ross, 2010) and occasionally engage in management techniques to present a professional image (DiMicco & Millen, 2007).

While it is known that employers supplement resumes with online information and use this to form impressions about candidates, less is known about what specific inferences hiring managers make from this information and how useful or accurate these impressions actually are. In the software development domain, recent surveys indicate that employers examining students' OSS experience look for compatible skills (while paying less attention to the popularity of the projects,) (Long, 2009) but it is unclear how exactly they go about this. Previous research has not examined how activity trace information would play into this process. Given the potential consequences of inaccurate hiring decisions being made on the basis of GitHub profiles and the possibility that job seekers may wish to manage the impression they give off, understanding how assessments currently are being made is important to inform the design and use of peer production sites for hiring purposes. In the next two sections we consider the impression formation and impression management process from the perspective of signaling theory.

### 5.1.1 Impression formation as signal assessment

As previously discussed in Section 3.4, the impression management and formation process in online peer production communities can largely be thought of as a signal production and evaluation process. Job seekers can attempt to convey or signal certain skills or abilities with the information they post on their profile or the activities they

engage in. Employers must determine which visible signals of developer expertise or personality to attend to on these sites (depending partly on how hard it is to verify them), and then interpret these signals to infer the developers' underlying or actual skill or expertise.

In the hiring domain, where deception about qualifications is a concern, being able to judge signal reliability is important. When the costs of forming an incorrect impression are high, for example, in hiring for a highly-paid job, perceivers may demand a more reliable signal that is costly to fake (Donath, 2007). Level of education attained is one example of a reliable signal of skills that is costly for a person to produce (Spence, 1973) but also potentially costly for an employer to verify. Closely related to the issue of signal evaluation is the issue of signal production, or impression management.

### 5.1.2 Impression management

Signaling theory also has important implications for impression management, because signalers may deliberately try to convey positive attributes to receivers. Given that signalers may have incentives to "cheat" (Connelly et al., 2011), understanding when and how they do this (and how receivers go about verifying the signals they produce) is an important topic that I investigated in this study.

Initial work by Goffman (1959) focused on ways in which individuals convey information about themselves to observers, which can be the "cues" they intentionally give, or their real behaviors, which may be "cues given off" through involuntary expressive behavior. Both types of information can be manipulated, either through overt deceit or through pretending. In the online realm, studies of impression formation and self-presentation in online settings have examined the cue management process in a variety of contexts and scenarios, ranging from honesty and lying about oneself in online dating profiles (Ellison, Heino, & Gibbs, 2006; Ellison, Hancock, & Toma, 2012) to friendship formation and other behavior on sites like Facebook (DiMicco & Millen, 2007;, Lampe et al., 2007; Walther et al., 2009) to blogging (Trammell & Keshelashvili, 2005). This work has primarily focused on understanding self-presentation in the context of interpersonal or non-work relationships, although Ellison et al. (2012) draw parallels between online dating profiles and resume submissions for

jobs, as using deception in either arena can be grounds for terminating a relationship. However, work on impression management in the social realm has focused largely on how people attempt to control the impressions conveyed by their profiles or pictures on these sites, and is not centered around impressions relating to work artifacts, skills, or behaviors.

Given that site design affects the reliability, visibility, and salience of signals (Donath, 2007), what signals do employers attend to in a peer production environment providing a plethora of trace information about work process and collaborative activity? How do prospective employees manage these signals? In order to understand how this new set of information plays into hiring, we examined how activity traces were used as signals in an online peer production environment instrumented with social media.

## 5.2  Method

I conducted a series of semi-structured interviews with thirteen GitHub users to identify how activity traces are used and assessed to infer a developer's abilities and personal qualities.  I began by sending a screening and recruitment questionnaire to 200 GitHub members with publicly available e-mail addresses on their profiles.  As there was no way to specifically filter for our two target groups (university students and employers,) I focused on targeting people located in North American and European cities that were likely to have large populations of both technical students and companies (e.g. San Francisco Bay Area, Boston/Cambridge, Pittsburgh, Seattle, Waterloo, Toronto, London, Berlin).

The questionnaire asked people if they had ever used GitHub as part of the job application or hiring process and if they would be willing to participate in a follow-up interview on the topic.  Overall, 128 people responded, 65 of whom volunteered for the follow-up interview.  Participants did not receive any compensation for taking part in the study.

I contacted respondents for interviews in the order in which they replied, sampling both employees and job seekers in order to understand the hiring process from both sides. Our participants for these interviews were seven employers who reported using GitHub to identify and evaluate job candidates and six job seekers who reported using

their GitHub profile to supplement their job applications. The interviews focused on how they had used GitHub during a recent hire or job application.

I asked employers to describe a recent past hire, focusing on how they used GitHub during that hiring process, what information they attended to on the site and what that information conveyed about the candidate. I also asked job seekers to describe how they used GitHub and how it had played into any recent job applications or interviews. To understand the role of impression management, I asked what they thought the information visible on their profiles said about them and whether and how they edited the information on their profiles or in other public places on the site.

In the analysis, I coded the interview transcripts to identify the different ways profiles were used in the hiring process, as well as the different types of inferences made about individuals being evaluated based on 'signals' in the GitHub environment. Using HyperResearch, a qualitative analysis software tool, I identified relevant sentences or broader segments in interview transcripts related to candidate evaluation, and then open-coded these segments for comments related to profile cues and inferences made from them. Next, specific instances of these themes were compared across interviewees and further refined as necessary, until a set of recurring themes about signals and the inferences drawn from them emerged. The interpretation of these themes from both employer and job-seeker perspectives is addressed in detail in the following sections.

## 5.3 Employers' use of GitHub

The first research question focused on how employers use GitHub profiles to evaluate new hires. All interviewees expressed the belief that a GitHub account gave more reliable insight into an individual's technical abilities and/or personal qualities than resumes or code samples taken out of context. GitHub profiles provided employers with a history of the individuals' contributions over time, and further guaranteed the candidate was indeed the author of any code submissions.

Table 2 summarizes the main GitHub signals and inferences mentioned by employers in the interviews. I also label signal-inference pairs based on their 'honesty' and evaluation cost: Signal reliability (or the degree to which an informational cue was seen as being easy to fake or an accurate indicator of an underlying quality), and the ease

with which a viewer can verify the signal (the amount of time it would take to view, interpret, and understand the significance of the signal).

Perhaps most notably, I found that employers were able to assess 'soft skills' (motivation, passion for the field) from the portfolio of activity visible on their GitHub profile. Relative comparisons across categories of activities (contributing to others' projects rather than just your own) or comparison across the types of work products created (side projects vs. purely functional or work relevant ones) signaled underlying personality and interest level of a developer.

In the rest of this section I describe in detail how employers used GitHub activity traces as signals of a job candidate's motivation, quality of code contributions, and soft skills or management abilities.

| Signal | Inference | Signal reliability, ease of verifiability |
|---|---|---|
| 1.  Active open source involvement | Shared open source values | Reliable Easy |
| 2.  Contributions accepted to high status project | Community acceptance of work, quality of contributions | Reliable Hard |
| 3.  Project ownership | Soft skills:  Initiative, project management | Reliable Easy |
| 4.  Side projects | Passion for coding | Reliable Hard |
| 5.  Number of watchers or forks of project | Project popularity | Unreliable Easy |

Table 2.      Summary of employer inferences from profile signals

## 5.4  Inferring motivation

Employers in the sample worked to assess how well job candidates would fit with their company or team culture (or person-organization fit). These factors are traditionally assessed during interpersonal interaction in face-to-face interviews through direct questioning (Rivera, 2012). The employers indicated that a job seeker's profile of activity

on GitHub signaled personal characteristics of the employee such as being a team player, showing commitment to their work, or demonstrating how he/she spent their free time.

### 5.4.1 Shared open source values and character

Employers care about value congruence with their employees. In the software development world, an important and hotly contested value is attitudes towards open source and whether software should be free. In fact, there is a well-documented ideology of open source software (Stewart & Gosain, 2006). Developers who differ in their software ideology may thus be said to come from different cultures. Thus, an important character property of a developer is their attitudes and commitment to the open source ideology.

Employers in the sample used presence on GitHub and activity levels on the site as signals of the level of commitment to the open source ideology. Simple presence on GitHub (having a profile and sharing even one repository) was viewed as an indicator of a potential employee's open source values by four of seven employers. The presence of code that was developed openly and shared with others signaled even more strongly that the developer valued openness, transparency, and participation in a community. This active involvement in the open source community was a signal of the candidate's selflessness and honesty. As one employer put it:

> *"If they've devoted time to this OS project, that's a good indicator that they're in [computer science] for the right reasons. Software engineering is becoming a pretty lucrative career…you could liken that to a doctor working with Doctors Without Borders. They're doing something because they want to give back to their community"* *(E7).*

Active participation in other people's projects was the most reliable signal of commitment to the open source mindset (mentioned by five out of seven employers). Cues such as recent and frequent commits to another person's project showed the candidate was indeed invested in the open source community (E2, E3, E4, E6, E7). It is rather trivial to create a profile on GitHub and fork other users' projects (meaning create a personal copy of the project in order to make changes to it). Having a copy of

someone else's project did not signal investment. The effort to fork a repository was negligible, while the effort and skill required to contribute meaningfully was much higher. This was widely understood, as one respondent described:

> *"a lot of people will just fork a lot of projects kind of to collect them but not actually do anything with them. So I look for a sign that these are things he's genuinely engaged in" (E4).*

Activities within forked projects were assessment signals of commitment, requiring much more effort to produce. Publicly building on another person's work served as a reliable signal that the candidate truly bought into the open source mindset (because active involvement was more costly than just setting up a profile on GitHub or simply forking projects). As one employer explained:

> *"[by looking for recent activity I was] sussing out whether they're a good sport about contributing to open source…if they're doing their job of keeping up to date and actually participating" (E2).*

### 5.4.2 Passion for programming

Organizations also differ in their working style or company culture around work life balance. Employers wanted to assess candidates' level of dedication to the work and their level of initiative. They were able to discern subtle motivational differences that suggested person-organization fit from the kinds of projects a developer worked on.

Interviewees described going through an individual's public repositories to figure out how they spent their time outside of work. They categorized projects on a user's profile as either work-related repositories that were part of an individual's "day job" (work or schoolwork,) and non work-related side projects, which could either be contributions to open-source projects not directly related to work, or personal projects done as a hobby or for fun but not necessarily intended for a wider audience.

For many employers, personal projects signaled a candidate's love for programming and willingness to do it in one's leisure time as well. One interviewee (E4) saw personal projects as a signal of interest in learning and developing one's career, while two employers (E3, E5) described using this signal to assess whether the candidate shared

the same enthusiasm for coding with the other members of their organization and were the type of people they liked to work with. This signal is valued because it ties into the aforementioned tendency for employers to like and seek out people who are similar to themselves and fit with their company's culture. As one employer explained:

> *"A lot of us spend our weekends working on [project name] so we want to work with people who are motivated to not just work on the code they've been assigned but to work on projects outside their job. It just shows a general excitement for the space and that's what we want to find – people that are really engaged" (E5).*

These side projects suggested a willingness to learn and revealed excitement about the software development domain. For employers this meant a potential employee who would spend their free time working, and show initiative and entrepreneurship in their work.

## 5.5  Inferring quality of contributions

Employers also care about a potential hire's competence and level of skill for the job. GitHub supported traditional methods of evaluating software development, allowing employers to look directly at the content of someone's code and the languages they had used. The cross-project visibility and the community on GitHub supported skill assessment beyond these traditional uses. Specifically, affiliation and accepted contributions to a popular project reliably signaled candidates' level of coding ability in the GitHub environment.

### 5.5.1  Accepted code as a seal of approval

If a candidate had contributions accepted to well-known open-source projects, it was seen as a community-level seal of approval. An accepted commit to a high-status project (a widely-used project with many contributors and watchers) signaled the candidate was someone who produced quality code. This acted as a reliable signal because it required approval of the code by others in the community, meaning it would be extremely difficult to falsify. For example, one employer described a candidate who seemed proficient because he had committed code to a high-status project:

*"Seeing that he had commits to jQuery, was filing tickets with jQuery, and I know that's a prestigious project to work on…Just by looking at his code, if nothing else seeing that it was being merged downstream into jQuery, I recognized that has demonstrated some level of proficiency" (E6).*

Another interviewee echoed this view for candidates who had contributed to open source projects, likening it to a reference:

*"someone else can vouch for your work because you were good enough to work on that project, be a part of that community" (E7).*

Since examining lines of a developer's code can be a time-consuming endeavor, it could be the case that using the reputation of previously-established projects that had accepted an individual's contribution as a proxy for quality (or lack thereof) was one way to reduce the perceiver's evaluation costs in forming impressions about a coder's abilities. An employer's opinion of a project's reputation in our sample was largely based on general knowledge of the wider community or past experience with its use versus visible cues at the project level such as watchers (people who have decided to "follow" the activity of a project) or forks (people who have saved a copy of the project to edit on their own.).

### 5.5.2 Popularity does not always equal quality

Employers also noted conventional signals of quality they did not trust. Primary among these were popularity signals: simple counts of watchers on a project or followers (people subscribed to a developer's activity feed). In some community settings, indicators of popularity (such as the number of votes given to an answer on a question-answering site) can serve as a proxy for the quality of the answer, while in other settings, popularity (e.g. having too many friends on a SNS) can be viewed negatively (Tong, Van Der Heide, Langwell, & Walther, 2008).

Project popularity on GitHub can be roughly assessed by the number of other people "watching" the repository along with the number of people who had forked that repository. Only one employer (E2) specifically mentioned looking at a candidate's main project to look for a large number of forks. Two employers (E4, E7) were more

skeptical about the utility of the watching/forking numbers as indicators of a developer's ability. As one interviewee explained:

*"I don't think I've hired or recruited someone specifically because they were working on a very popular project or something. There aren't enough popular projects and the popularity doesn't necessarily indicate quality for that to work." (E4).*

Popularity was thus to some degree viewed as a signal that developers could game. The interviewees noted that project popularity was an unreliable signal of code or developer quality because it had more to do with how much an individual promoted their work:

*"You can see if a lot of people have watched and forked and that's a good thing, but it kind of depends on how good a marketer that person was as well on GitHub." (E7).*

Rather than relying on numbers of forks and watchers, employers described looking at the project where the applicant had made the most commits (as presumably that was the work they were the most serious about or interested in) and then assessing the actual code that was written there to understand the individual's style and skill level. Although employers still were directing effort to inspect and understand the code being written, the higher-level cues about the individual's activity level across projects allowed them to better filter their attention.

## 5.6  Inferring developer "soft skills"

Finally, one employer mentioned inferring "soft skills," such as project management abilities, through a developer's activity traces. These cues were largely gleaned from their interactions in a project they had started rather than one they had forked (a distinction which is made obvious by the GitHub interface).

These projects allowed him to observe a candidate's project management skills and collaboration style. Owning a project involves tasks such as setting a design direction, managing incoming code contributions and patches, and interacting with potential collaborators. The employer described being able to infer these soft skills from projects that a person owned, stating that

*"[Projects he owned] would give me a better sense of his long-term design skills and ability to manage a community" (E4).*

## 5.7 Job seeker perspective

The second research question was how job seekers manage their presence on GitHub. In order to address this question, I also conducted interviews with six job seekers on GitHub (referred to here as J1-J6) from four different U.S.-based universities. These participants reported using GitHub to supplement their application materials as a job seeker in our recruitment survey. Of these, two were undergraduate students in their final year, one was a PhD student, and the other three recently graduated and had been working at their current position for one year or less. They had between 4 and 28 public repositories visible on their profiles (mean= 17.3) and an average of 19 followers.

All of the interviewees had originally joined the site to host projects they were currently working on and archive past work. The projects they shared included school and class projects, projects for hackathons, samples of code written for job interviews, side projects, and open source projects, both recent and older.

They varied in the ways they shared their work on GitHub with employers. Four of them had provided a link to their GitHub accounts on their resume and stated that this had been consulted during the hiring process. One person never provided this information to the employer and instead assumed the employer had found his profile through a Google search. The final interviewee did not actively promote his GitHub account in the hiring process but stated that he would provide a link if asked for it.

In the interviews, they described the role that their GitHub account had played in recent employment searches. We asked them to describe the impressions they thought their profiles conveyed to potential employers. Finally, we talked about how they managed their profile for an actual (or imagined) audience of potential employers.

## 5.8 Impressions given from profile

I asked interviewees what impressions they thought their GitHub profiles conveyed about them to a potential employer. Their responses highlighted two key qualities they

felt employers could see from their activity on the sight: passion for software development as an activity and field and technical expertise.

### 5.8.1 Activity signaling passion for the field

Interviewees mentioned that their visible level of activity in the site was a positive signal of passion for software development (J1, J2, J6). One person realized that it might be important from the employer's perspective to see evidence of extracurricular coding activity. Referring to some of the side projects, he mentioned,

> *"The fact that in my free time I went and coded something…I think that counts a lot"*
> *(J2).*

The interviewees also indicated their activity in terms of following also signaled an interest in the domain. They thought the fact that they followed a lot of projects conveyed that they possessed current awareness of the latest projects and technology (J2, J3). One interviewee followed several hundred coders and projects. He thought this would show employers his interest in coding and open source even though he himself had very few projects. Another explained,

> *"The impression I would like to give off is: This guy is current with what's going on in*
> *the open source world" (J2).*

### 5.8.2 Work signaling technical expertise

Developers were also keenly aware that the portfolio of projects on their profiles conveyed their areas of expertise to employers. One person's profile (J5) contained a lot of projects in "web-heavy" languages like Python and JavaScript. He thought these projects revealed that his skill set focused on web development. Several interviewees noted the languages they used in their projects would signal proficiency in those languages (J4, J5, J6).

Interviewees were also aware that details of their code would influence perceived skill level. They noted that good style such as descriptive commit messages, comments, and code indentation would convey competence. Two interviewees were aware that their code was not particularly readable or well-organized, partly because the projects were

originally just intended for personal use and trying new things. One person had many old projects on his site and his style had improved since. He explained,

> *"I have a lot of things in my GitHub account that I wouldn't feel comfortable putting on a resume because I didn't know how to code at the time…I'm definitely a better coder than some of the projects imply" (J5).*

## 5.9  Impression management

### 5.9.1  Attitudes towards profile manipulation

Although the job seekers in the sample were acutely aware their profiles gave off certain impressions about them, few reported having engaged in efforts (such as changing or editing any information visible on the profile) to control this impression. I learned in the interviews that this was partly because there were few ways of doing so and doing so was too costly and effortful. As one interviewee explained:

> *"A major component of your GitHub profile are your projects, and it's not really easy to refactor your code quickly" (J1).*

Job seekers also noted that even old or incomplete material could still be useful in conveying qualities about its creator. One interviewee intentionally left incomplete or unfinished projects on the site because they could signal thought processes such as *"how you are approaching and organizing something"* (J4).

### 5.9.2  Signal accuracy: Cleaning up the profile

The developers I interviewed realized there might be some need to edit information on their profile, either currently or in the future, to more accurately represent their work. This profile improvement consisted of either commenting code or removing old projects. For example, one developer expressed a desire to concentrate more on "cleaning up" his profile. In lieu of having the time to improve the accuracy of the signals on the site, he provided employers with a description of the current state of projects in the repository descriptions:

*"I try to say statuses of projects sometimes: I have a couple repositories where it says something like 'this is a super early proof of concept thing'…I make a slight effort to sort of downplay the things I'm not very proud of" (J5).*

He mentioned that in the future, he would consider stripping out everything that he was not proud of, including old and incomplete projects, in order to provide a more accurate picture of his current abilities.

Overall, I found that participants did not seem to engage in much "gaming" of the system for a variety of reasons. On one hand, they often felt that they were happy with the signals given off by their profiles regarding their active involvement and participation in projects. On the other hand, they did not try to fake information because it was effortful (i.e. difficult to give a false impression about actual code they wrote.) They could influence information given off by projects by downplaying them in the repository descriptions or by adding proper comments into the code. Keeping the profile as-is, with imperfections or old information, was seen as useful, both to other developers (who might be able to learn something by viewing old projects) and to employers.

## 5.10  Discussion

The interviews revealed that employers used cues on GitHub as signals of underlying characteristics of potential hires. These signals varied in their reliability as a function of production cost in the GitHub environment. In addition, employers' use of these signals was directly related to evaluation cost. Job seekers on the other hand, were aware that employers were using their behavior as signals of underlying properties. In addition, attempts to manipulate these signals were culturally or practically discouraged. The next section considers what made signals reliable and how evaluation cost influenced employer behavior.

### 5.10.1 Evaluation of signals and inferences:  Reliable signals

Employers expressed a belief that some cues were not easily manipulated and thus they were reliable indicators of a potential hire's competence, values and collaboration styles. My analysis revealed that two key properties of activity traces supported signal

reliability. These were the fact that traces were: (1) archived over time - signals generated by direction of time and effort expended (commitment to open source values, passion for programming), or (2) networked - signals generated by third party affiliation.

Activity traces on GitHub are archived and persistent, meaning employers could view a history of effort across projects over time. Employers used information about how much effort was put into different projects to draw conclusions about a person's values and motives. This signal (relative effort across projects) also appeared to be an accurate indicator of motivation. For example, job seekers who contributed to open source projects truly wanted to give back to the community. This kind of historical signal would be extremely costly to fake, making it a more reliable indicator of investment.

The network context of activity traces in GitHub meant that affiliation could also act as a signal. The acceptance of offerings to joint projects in the form of pull requests or code commits was archived and accessible. The fact that work had been accepted to well-known projects was not easy to fake because participants couldn't control whether a contribution was accepted or rejected. Connection to these high status projects increased employer perceptions of the candidate's competence or quality. These signals were deemed more trustworthy because they were generated by a third party. This is the concept of warranting introduced by Walther et al. (2009). Similar to signaling theory, the warranting value of information has to do with the perceptions of the extent to which information is immune to manipulation by the source it describes (DeAndrea, 2014).

 Research on warranting suggests that when forming an impression, information provided by third parties (for example, the fact that someone else had accepted the candidate's code into their work) and connection to a social network is more trustworthy than self reports. The offering and acceptance of work contributions provided a kind of third party verification of work quality.

### 5.10.2  Unreliable signals

Activity traces were deemed unreliable signals of a potential hire's characteristics if they were: (1) subject to manipulation, or (2) likely to be missing information because of

the way they were generated. An understanding of whether a signal was unreliable or not required employers to have intimate familiarity with the GitHub environment and the social context surrounding project activity.

Employers viewed some signals as conventional signals. The comments they made suggested that they were using social conventions surrounding the interpretation of signals like number of projects forked or the popularity of projects that people contributed to. For example, there was a general feeling or awareness that forking many projects was something people on GitHub did to "collect" them and did not actually mean they were actively contributing to them. As a result, while coders were not necessarily punished for forking many projects, this piece of information did not carry much weight in the evaluation process.

Project popularity was another example of an unreliable signal of a potential hire's coding ability. Although popularity was easy to assess through aggregate statistics on the number of people watching a project, it was seen as subject to manipulation and thus not closely tied to ability (the underlying characteristic employers wanted to assess). Project popularity could be inflated by an applicant, and was dependent on factors distinct from coding abilities such as how well an individual marketed his or her project or what type of project it was.

Signals were also deemed unreliable if they were known to be missing information. For example, employers acknowledged that projects on an individual's profile were not always representative of their full body of work. This meant they were only partial signals of properties like technical skills and coding abilities, since private projects were not visible on GitHub. If the majority of an individual's coding activity happened behind a firewall at work (meaning they could not share it), then their GitHub account would not completely reflect their true range of abilities.

Employers were aware that certain information about a potential hire was not visible on the profile and the possible reasons it might not be visible. Despite this knowledge of the broader context or perhaps because of it, employers valued an active presence on GitHub. There was a sense from the employers sampled that all things being equal, a

candidate with a GitHub profile would still be preferred because of the added signals. One person said,

> *"I don't think you can use it as the sole way to judge someone because a lot of it's going to depend on how much time the developer is donating to maintaining their GitHub, so if this is someone who has a wife and kids and not much time to be programming out of the workplace, that doesn't necessarily mean they are not as good of a developer. But if something is up there, it's definitely a huge plus and probably one of the first things we look for – are they sharing their source code?" (E7).*

### 5.10.3  Evaluation cost

Evaluation cost seemed to influence employer's use of activity traces in GitHub. The signals summarized in Table 2 are all observable on a user's profile on the site. It is easy to quickly verify active open source involvement just by seeing that a person is a member of GitHub and has evidence of recent activity in their profile's activity feed. It is also easy to tell whether a project was original or forked, via the presence or absence of an icon. It is perhaps not surprising that four out of the five most commonly used signals were effortful to produce and therefore seen as reliable. However, it should be noted that three of the five signals were also fairly easy to find and evaluate. This suggests that in addition to signal reliability, other characteristics of signals such as observability and effort to verify (Connelly et al., 2011) played an important role as well. This highlights the interplay between conflicting aspects of signals (e.g. reliability vs. effort) that previous work on signaling theory has suggested but not necessarily explored in depth.

Determining contributions to a high-status project required more effortful investigation if these commits were not in the top recent activities shown, while determining if something is a "side project," as one person mentioned, may require following up with a user to learn more about it (depending on how the repository is described in its textual summary.)

While one possible advantage of GitHub signals is the ability to view how people work with others, interviewees did not mention looking at evidence of past interactions with others to determine attitude and personality. Work on hiring has shown employers are

usually interested in candidates' likeability as a team member (Weiss & Feldman, 2006). In other contexts such as Wikipedia, traces of civil online interactions are used as a signal of likeability and collegiality (Derthick et al., 2011). However, in GitHub, it is non-trivial to view an individual's interactions. This information requires more effort to seek out and is time consuming to review.

Site design strongly influences the cost to access information about a person. This cost may have repercussions on what information feeds into the hiring decision. The employers I interviewed favored cues that took less effort to access and verify, such as relative effort across projects or presence of forked versus owned repositories (Figure 2).

**Public Repositories** (38)

Find a repository…

All Repositories | Sources Forks Mirrors

tilt | Ruby ⊙ 758 ⋔ 77

Generic interface to multiple Ruby template engines
Last updated 1 day ago

Graphs show relative participation across projects — all commits ■ commits by owner | 52 week participation

Text indicates if project is forked from another user —

backbone-fundamentals | JavaScript ⊙ 2 ⋔ 118
Forked from addyosmani/backbone-fundamentals

A creative-commons book on Backbone.js for beginners and advanced users alike
Last updated 4 days ago

Figure 2.    Sample cues used by employers. Relative effort across projects was visible in the bar graph below each repository (project) description – gray bars represent all commits to project and blue bars represent profile owner's contributions. Whether or not a project was owned versus forked is indicated in text below the project name.

They used these cues to identify people with extra passion in addition to skills, who were then set apart from general applicants or followed up with in person.

The second research question focused on impression management by job seekers. Participants acknowledged that they could not easily give a false impression about the actual code they wrote, but they could try to influence information given off by their projects through downplaying them in the repository descriptions or adding more proper comments into the code. Given that GitHub has relatively recently gained popularity as an employment tool, it is possible that future site members will be more conscious about managing their impressions for future employers.

## 5.11 Design implications

The interview findings have implications for the design of open transparent work environments. The effort required to access activity traces connected to a person may change what is considered during the hiring process. My results suggest it is important to strike a balance between providing enough information to be useful while avoiding overload. This could be useful in other online peer production settings such as Wikipedia: Visualizations of an editor's work history across pages could help editors understand what kinds of tasks they would be good at performing. This kind of history could also give context to debates and arguments.

In Table 3, I present additional design opportunities and suggestions for the type and visibility of different kinds of information that could be incorporated into personal profiles on peer production or other types of work sites:

| Interview finding | Design implication |
| --- | --- |
| Effort to access activity traces may change what is considered during the hiring process | Make it easy to access the information that the site wishes to emphasize or promote. Alternatively, provide a dashboard setting where viewers can choose to highlight the activities or qualities that are most important to them. |
| Employers looked for signals of individual qualities such as motivation | Additional ways of visualizing developer motivation could take into account other behaviors beyond number and frequency of commits or activity. These could include providing links, badges or symbols to group or categorize individuals as contributors (either system-generated or provided by collaborators) |
| Employers were interested in interpersonal qualities and cultural fit | Individual profiles could aggregate or automatically assess the nature of an individual's interactions with others (valence of language used) and number of others collaborated with to give a broader sense of working and collaboration style |
| Comparisons across projects/people | Since multiple candidates may be considered for a position, displaying a side-by-side profile comparison view would make it easier to compare candidates against each other on various dimensions |
| Employers used third-party evaluations of developers' quality | Provide alternative or easier ways to assess third party evaluations of developer quality (for example, could provide overall statistics of the percent of a candidate's contributions that had been accepted versus rejected) or allow project owners to give contributors quality ratings such as "kudos" or stars to exceptional contributors |

Table 3.    Design recommendations based on interview findings

My work also raises interesting broader questions about designing activity traces for multiple audiences. GitHub is a site that was not originally designed as a hiring tool. Its design also needs to support collaborative software development. Activity traces are used by developers to coordinate their work and transfer knowledge across projects. However, employers increasingly use these traces to evaluate new hires. There is a tension between using the site as a tool to carry out work and using it to manage the impression one gives to external observers.

The degree of overlap between signals that are important to manage for gaining acceptance within a given community and signals that are important to manage for external viewers such as employers or evaluators is another area that future work can tease apart and provide insight for transparent system design. As Begel, Khoo, & Zimmermann (2010) suggest, a social medium that reveals a user's knowledge, expertise, activities, or availability may be useful for finding knowledgeable others when consumed by the user's peers, but may feel like "corporate spyware" when consumed by his manager. Future systems could provide different views or privacy settings for different categories of users, while still providing the benefit of transparency of actions and reliable assessment cues.

## 5.12 Implications for other domains

The importance of passion for coding and active involvement in the community may be uniquely emphasized in GitHub due to the nature of the open source development community, in which people may be more motivated by self-development and reputation gaining compared to contributors to online content sites such as Wikipedia. In addition, contributing software to open-source projects also requires a certain level of expertise to pass the review process, compared to Wikipedia (Oreg & Nov, 2008).

However, some of the issues raised here may also apply to issues pertaining to the role of publicly visible behavior in other realms outside of GitHub. For example, there are other instances in which people in technical fields may wish to promote their skill or knowledge in publicly viewable online arenas.

While there has been much concern about employers accessing individuals' personal SNS profiles, there are other areas in which employers may wish to access public work-related information, and sometimes job-seekers may wish for employers to see this information. For example, there has been discussion in the developer community about when/if one should display their StackOverflow score on their resume, including varied reactions from employers about the meaning of this score (e.g. "Would you put your stackoverflow profile link on your CV / Resume?," programmers.stackexchange.com). Users of the Quora question-answering site have

also mentioned that they will strategically answer questions in order to attract the attention of influential people and as a form of networking (Paul, Hong, & Chi, 2012).

As people interact and collaborate online on sites geared towards "serious leisure," they may generate additional traces as they go about their work. It is possible to imagine a future of work in which evidence of one's abilities and what "type of person" they are is not solely gathered from their activity within a structured organization or educational program but from the aggregation of their behavior on the web. For example, are the details of a person's online Wikipedia editing activity considered to be useful by employers looking to hire journalists or writers, and do Wikipedia contributors looking for jobs actively promote the work they have done on the site?

These questions suggest interesting research directions to explore with respect to the role of visible online activity information in professional development and hiring in other creative digital fields. The kind of transparency pioneered by GitHub may have implications for the future of hiring. Employers may begin to expect applicants to provide a rich history of detailed work activity traces. Job seekers may in turn increasingly gravitate towards companies that will allow them to accrue a publicly-available (or shareable) portfolio of work. Businesses may find employees demand some open-source work sharing, and see policies of openness as a benefit. We are already seeing this kind of trend in fields like graphic design, as individuals often work for themselves to maintain the ability to visibly promote their work.

The implications of my results extend beyond software development as work becomes increasingly digital. Providing accessible, reliable traces of an individual's work history may support more accurate impressions of unknown contributors. These impressions will shape decisions about recruiting, hiring and promotion in tradition and new forms of organizations like Wikipedia or crowdsourcing. Such impressions are also likely to influence the collaborative dynamics of work. It is important for system designers and policy makers to consider what actions and activities can and should be recorded and made visible. Our results should help decision makers develop useful and efficient ways of providing various groups with the information they need while protecting individuals' rights.

## 5.13 Limitations and future work

The sample of interviewees in this study was limited to a small subset of people who volunteered to discuss the topic of GitHub and hiring. Another question that was not addressed in the current research was how impression formation occurs with other cues and information beyond the GitHub profile. It would also be valuable to follow up on employers' perceptions of new hires found or evaluated using GitHub after they have spent some time in their job role to understand how initial impressions correlate with actual performance, or how these change over time.

Future work can also include developers and employers who are not involved in the GitHub community or who participate in other sites such as BitBucket or Sourceforge to understand if they present or market themselves differently in these environments, as well as to understand why they choose not to use GitHub.

Furthermore, the role of impression formation and management in other sorts of online production communities should be examined, to uncover similarities and differences across domains outside of open source software development.

## 5.14  Discussion

In this work I examined how activity traces on GitHub are used as signals by potential employers. Past work on impression management online has suggested people actively manipulate various observable elements of their online presence to portray an idealized self. Interestingly, we saw very little evidence of such behavior in GitHub. This may be because the effort required to manipulate activity traces that actually matter to employers (such as active participation and code style) is too great.

At the same time, employers preferred easily verifiable signals of a potential hire's skills or abilities. This desire to minimize assessment effort meant they did not necessarily take advantage of the full range of transparency that GitHub can provide. They used candidates' presence and high-level cues of activity on the site to infer how well the person would fit within their organization. Employers favored developers who signaled similar passion and willingness to work on coding projects both during and outside of work.

Study 1 provided insights into how employers and job-seekers assessed and signaled coding ability and interest in programming through easily assessable and visible cues on GitHub profiles. Using GitHub for hiring provided an interesting scenario in which programmers were being evaluated based on behaviors out of the context in which they were originally generated. Returning to the peer production context, I was also interested in how GitHub profiles (and specific cues therein) were used by project owners to evaluate contributors and inform the decision making process. Therefore, Study 2 focused on the use of GitHub profiles in a more interactive and interdependent setting where project owners must engage with unknown contributors and coordinate contributions to the project as a whole.

# 6 Study 2: Impressions of contributors

For the full paper describing this study, please see Marlow, Dabbish, & Herbsleb, (2013).

In Study 1, I investigated the role of using activity trace information for hiring and recruiting developers. This scenario was a non-interactive evaluation in which employers were deciding what a candidate was like and whether or not they were worth following up with. Individual activity trace information was being viewed and interpreted outside of the original context in which it was created (e.g. in the process of working on open source software projects).

I was also interested in understanding how activity trace information is used in an interactive and collaborative setting to help people carry out work. In this case, the activity traces are related to and being interpreted in the same context in which they were created. An important aspect of collaboration in peer production is peer review of contributions. I was interested in how activity traces and profiles were being used during peer review of contributions to open source projects on GitHub.

In commons-based peer production, a fundamental problem is matching people to tasks (since there is no centralized management) (Benkler, 2002). This problem boils down to perceptions of expertise, credentials, etc., which impact concrete (decentralized) decisions about who can contribute what changes to what project artifacts, and whose voice is listened to. Such questions are potentially resolved by the impressions that people form of each other via social media. However, we don't yet fully understand when and how people form impressions of each other, and how these impressions influence work effectiveness in a peer production community where work environments are tightly coupled with social media.

The open source software development domain presents an example of commons-based peer production where "uncertainty about the quality of others is the rule rather than the exception" (Stewart, 2005). The issue of deciding how to deal with problematic code contributions can partially be attributed to this uncertain environment, where project

owners may have difficulty understanding the expertise, background, and credentials of unknown contributors who seek to become involved.

In typical work settings, we only have access to directed communication from our collaborators and revision information related to the projects we work on with them. However, in an open workspace instrumented with social media, there is the potential to know about the entire collaborative world of another user. This increased amount of information may change the way we understand what someone else knows, is good at, or what they are like as a person, and has the potential to inform how we assess their work and to shape our interactions.

Used in the context of collaboration, GitHub features give users unprecedented information about what other people are working on, who they have worked with and the history of their work across a community. Developers interact with unknown others through the site in the process of finding software to use, modifying code for their own purposes, discussing potential changes to software projects, or accepting contributions to a project they own. In the latter case, accepting a pull request from someone involves deciding whether code contributed by another person is worth being incorporated into the main code base.

## 6.1  Research Questions

Although social media functionality is increasingly prevalent in online peer production communities, we don't yet understand when or how users form impressions of each other in these online workspaces instrumented with social media functionality and how these impressions influence the work they do. Initial work by Dabbish et al. (2012) has revealed that users rely on behavior traces on the site to manage their projects, learn from others, and manage reputation and status. However, this work did not examine in detail when and why people attended to traces left by others, or how these perceptions influenced the work that was done. They noted that inferences were made about submitters' code and competence, but did not specifically explore how these inferences affected work outcomes.

We therefore still lack detailed understanding of how the impression formation process unfolds and how it directly impacts the project outcomes associated with assessing the

contributions of unknown others in a large-scale environment. To address this gap in the previous work, I focused on understanding how impressions of unknown collaborators are formed based on an aggregation of a developer's actions and interactions, and how these impressions affected work outcomes.

The research questions for Study 2 build on the distributed social cognition model (Smith & Collins, 2009), which posits that impression formation is an active process influenced by behavior in a network or group. When forming impressions of others, individuals often engage in an active process that involves several steps: First, choosing whether to obtain information about the target; next, choosing what information is elicited, then interpreting the elicited information and using this to form a person model (an integrated interpretation of what a person is like.) This theory gives us general guidelines about how information seeking occurs but does not provide details about what the process looks like in a specific setting. I was interested in investigating the following questions in the context of GitHub:

RQ1: Scenarios for information seeking
In what circumstances do people seek out information about unknown others in an online peer production community?

RQ2: Activity traces influencing impression formation
How do people use activity traces such as profile cues in an online peer production community to form an impression, or cognitive representation of others? What cues are more and less important in uncertainty reduction and impression formation?

RQ3: Impact of impressions on work outcomes
How do these impressions affect the work that is carried out?

In order to address these questions, I conducted an interview-based investigation of the role of profiles in influencing impression formation in GitHub. I found that users in this setting seek out additional information about each other to explore the project space, inform future interactions, and understand the potential future value of a new person. They form impressions around other users' expertise based on project and code-related cues, which combine with interaction traces to help influence judgments about how to

work with new contributors. The results from this chapter have implications for the design of social technology to support online peer production communities engaged in knowledge-based work.

## 6.2  Background

Open source software development is an example of a peer production community fueled by volunteer contributors interacting, via computer-mediated channels, from all over the world.

However, it is unclear to what extent the findings of impression formation in Wikipedia or artistic collaborations extend to evaluating non-managerial participants or to other online peer production communities, partly due to the nature of the domains. In open source development, participants have a wide range of technical abilities and skills, contribute to a project in different ways for different reasons, and may be motivated by career goals or a desire to build reputation and gain peer recognition (Lerner & Tirole, 2002). Attribution may be less of an issue in open source development than observed in artistic communities (Luther, Diakopoulos, & Bruckman, 2010; Monroy-Hernández, Hill, Gonzalez-Rivero, 2011) in part because projects are viewed as community property and also because systems like GitHub automatically provide a record of a project's origins and contribution history.

## 6.3  Challenges in software development

Supporting awareness of teammates in distributed software development to improve both task and social outcomes has been an important research area. For example, Trainer, Al-Ani, & Redmiles (2011) suggest that providing visual traces of work interdependencies between team members can influence trust in distributed team members and help people understand whom to ask for assistance. Other work (Treude & Storey, 2010) has also pointed to the ways in which tools such as dashboards and activity feeds can help teammates get a sense of the project and plan their tasks. Begel, DeLine, & Zimmermann (2010) point to the benefits of using social media at various stages in a software team's lifecycle, particularly with respect to coordination and communication. For example, they point to how social media can be used for effortless knowledge sharing or to help groups infer best practices by observing others' work.

However, most of these studies have focused on ongoing, organized teams of developers within an organization, unlike the more volunteer-based collaborations that can occur in open source development, where contributors can vary in the length of their involvement and may not share a specific organizational affiliation. Furthermore, in contrast to teams within organizations, developers in open source settings often lack guidance from a management hierarchy, and thus need to self-coordinate and make autonomous decisions based on whatever information is available to them.

Open source projects rely on contributions from a global community of developers to perform various tasks ranging from bug reporting to submitting feature requests and contributing patches and code. The success of a project depends on the proactive and constructive participation of contributors to the project (Sinha, Mani, & Sinha, 2011). However, these outside contributions may vary in quality depending on the skills and expertise of the people who contributed them.

Diversity in technical abilities can be helpful to a project because various types of contributions, from filing bugs to suggesting features, can be made by people with a range of expertise (Sinha et al., 2011). However, bug reports can waste time and divert developers' attention when they are misleading due to contributors' inexperience (Lotufo, Passos, & Czarnecki, 2011). In both of these cases, the potential benefits of receiving contributions from inexperienced, unknown others may be outweighed when editing their work becomes too time-consuming to deal with, and this often results in the suboptimal outcome of contributions not being accepted to the project (Rigby & Storey, 2011).

Although we know that individuals engaged in open source software development are continuously evaluating the contributions of others, few studies have addressed exactly when and why developers seek information about unknown others, how impressions of these people are formed, and what information is relied upon to infer an unknown developer's expertise, or other personal characteristics. In the next section we consider previous work on impression formation and uncertainty reduction to inform our study of this process in the peer production context.

## 6.4  Impression formation and uncertainty reduction

Early work by Goffman (1959) focused on how people created and managed the impressions they gave off to others in order to shape how others viewed them.  More recently, work has turned to understanding the cognitive processes of those formulating the impression (that is, the observers), and understanding how perceivers mentally organize discrete information about a person into a cognitive representation (e.g. Hamilton et al., (1980); Smith & Collins, (2009).

Impression formation with strangers can be thought of as an uncertainty reduction process motivated by the goal of understanding their behavior and predicting how they will behave (Gibbs et al., 2011). When people are faced with a previously unknown person, they can use direct social interaction or information gathering as means of reducing uncertainty about that person (Ramirez et al., 2002). During the process, the seeker fills out their mental models or mental representations of these new unknown people that help him or her to make sense of other people and their intentions, emotions and behaviors (Antheunis et al., 2010).  These can be models about how they will react to certain situations, but also what they know.

A great deal of research has examined how individuals form perceptions of what other people know or are good at.  For example, expertise finding is an important task in the corporate domain e.g. (Shami et al., 2009) and many internal tools have been developed, for example, to help people tag their own and others' expertise (Raban, Danan, Ronen, & Guy, 2012).  Consulting professional contacts' profiles on social media sites such as LinkedIn has been used as a strategy for learning more about someone before meeting them (Skeels & Grudin, 2009)..

However, in a large-scale online peer production community with hundreds or thousands of members, individuals cannot feasibly evaluate every new contributor that accesses a project or submits work. In this case, they may tend to rely on stereotypes as a means of more efficiently assessing people (Quinn, Mason, & Macrae, 2009).  Research in corporate settings has demonstrated that impression formation related to work skills does occur and can be enhanced through profile information (Raban et al., 2012; Shami et al., 2009).. It is unclear exactly how this impression process works and whether it can

occur outside of an organizational boundary. There are gaps in our knowledge about the mechanisms of impression formation in a peer production environment such as open source where participants span organizational boundaries and have extremely heterogeneous backgrounds. Open-source development has traditionally conducted much of the project-based interaction on message boards or via email lists, where it can be difficult to gain insight into who a person is or what they are good at without repeated interactions.

In this study, I add depth to the initial observations of the previous work by focusing in greater detail on when people seek information about each other, what information they use, and how they process it. I was specifically interested in the role that visible information about GitHub users plays in the collaborative software development process, with a particular focus on how this information (or lack thereof) influences the work process.

## 6.5 Method

I conducted interviews with 18 GitHub users focusing in detail on how they formed impressions of people they encountered on the site. Using information obtained through the GitHub API, I identified GitHub users who owned at least one open source project. Potential interviewees who had publicly-displayed e-mail addresses available on their personal profiles were contacted to see if they would like to participate in the study. Interviewees did not receive any incentives for participation.

I was interested in people who were in charge of managing projects that received contributions from a variety of different people, so the selection criteria for identifying potential participants were that they owned at least one project with six or more authorized "editors" (people who were authorized to make changes without approval) as some indicator of scope and size, although the projects also had many more watchers and forks and were open to submissions from non-editors as well. In this way, I identified people who held leadership positions by owning popular projects that others were involved in, and thus, presumably, were in charge of managing contributions from outside others to their projects.

Of the 18 interviewees (17 of whom were male), twelve were based in the U.S. and six were located in Europe. This sample was fairly representative of the GitHub community, which, according to a recent analysis, had around 80% of its users coming from North America and Europe (Takhteyev & Hilts, 2010). Except for two PhD students, all of them worked as professional software developers or consultants for various companies and organizations but also used GitHub to host personal projects that they worked on in their spare time. Most had been members of GitHub for three to four years (as revealed by the "joining dates" on their profiles), with the newest member having joined just under two years ago. Overall, they were active site users and the largest projects they owned had a median number of 59 project watchers, 21 project forks and 451 contributions.

Interviews lasted between 30 minutes and 1 hour. They were conducted over Skype, with screen sharing enabled so that both interviewer and interviewee could refer to various GitHub profiles and pages. The interviews followed a semi-structured format. To address RQ1, interviewees were first asked to identify scenarios in which they sought out more information about unknown others. Next, they were asked to go to these people's profiles and describe how good a sense they had of what the person was good at and what they were like as a person. To address RQ2, follow-up questions probed on what specific information they were basing these conclusions on. Each interviewee assessed the profiles of two or three others – some of whom they knew well and others whom they did not know well.

Finally, to assess the influence of profiles on work outcomes (RQ3), I also explicitly asked participants to show examples of recent pull requests (i.e. code contributions they had received from others) that they had recently dealt with and either accepted or rejected. They described the process of receiving the pull request and walked through the decision process of how they had handled the request, including whether or not they had consulted the profile of the requester.

## 6.6  Analysis

As the first research question concerned GitHub profiles, I began by analyzing interview transcripts for instances in which participants mentioned having consulted

others' profiles. The analysis process was structured around the three stages in the distributed social cognition model (Smith & Collins, 2009) and the corresponding research questions. Two researchers then made affinity diagrams grouping the instances around common themes, discussing until consensus was reached, and used these themes to develop categories for scenarios influencing profile use. We used qualitative analysis software (HyperResearch) to assist in coding interviews and in aggregating similar themes.

We then repeated this analysis process for the other research questions. To understand which cues individuals were using in impression formation, we identified all instances of new impressions in the transcripts and then open coded these instances to form themes around the cues being attended to and the impressions that resulted from them.

Finally, to understand the impact of impressions on the working process, we focused on one of the most commonly mentioned themes in the first round of coding: the context of receiving new pull requests from unknown others. We identified these kinds of pull request interactions described by the interviewees. Next, we generated a set of categories around each pull request in terms of the outcome of the request and factors leading to this decision. At some times, discussion was needed to determine the nature of some of the pull request conversations (whether there was a conflict or not) but this was resolved by matching what the interviewees said about the conversation with the correct segments of the visible discussion. Using focused coding (Charmaz, 2008), we compiled a set of pull requests that demonstrated these themes and then compared and contrasted specific pull requests to create a flow diagram of how decisions to accept pull requests were made, and how exposure to profile elements factored into this process.

## 6.7 Investigating a pull request sender

Because GitHub enables unknown people to fork a project without interacting with the owner, project owners often found themselves receiving pull requests from people they did not previously know. For many people, desire to reduce uncertainty about this person and what may have made them interested in the project propelled some users to examine the profiles of these unknown others in conjunction with examining their code.

In a way, learning more about a person was important to start building a working relationship: As one interviewee put it, *"I want to know you before I help you" (P6).*

### 6.7.1 Forming expectations about skills

For some interviewees (P3, P6, P14, P15), investigation into a contributor's skills and abilities occurred the first time a person submitted a contribution. Knowing what a person did helped owners to make sense of why the contributor might be interested in the project and also to understand *"who they are and what do they do"(P3).* They were interested in seeing if a person had contributed to other projects to understand in what capacity they might be able to help on their own project (or how much assistance or extra effort accepting their contribution would require, based on their technical abilities.)

For others (P10, P12), the information seeking did not occur until there had been multiple interactions with a given individual. One person explained that due to the sheer volume of people that he deals with, it is only useful to learn about those with whom he has some sort of ongoing relationship:

> *"for me there's so many different people I interact with on the Internet because my projects are pretty popular so that I just don't have the mental capacity to know each person who I interact with. But let's say if I get another bug from him and then maybe he makes other comments in the future, at some point he'll cross a threshold where I'm like okay, who is this person? What does he or she do?" (P12).*

## 6.8 Cues influencing impression formation

The second research question focused on which cues people attended to on GitHub profiles and on how they used these cues to form impressions about others. We found that profile cues were used to assess a user's general coding abilities, as well as determine specific areas of expertise and to learn more about an individual's personality. Table 4 presents an overview of general categories of impressions that were formed and some example cues that were used to derive these impressions:

| Impression category | Cues |
|---|---|
| C1. General coding ability | Amount of activity, frequency of commits, number of projects owned vs. forked, length of time on site, languages used |
| C2. Project-relevant skills | Types of visible activity (coding vs. discussing), specific languages used |
| C3. Personality and interaction style | Past discussion posts and threads |

Table 4.        Impression categories and associated cues

### 6.8.1  Assessing general coding ability

When describing the impressions formed of a person by viewing their profile, interviewees often made stereotypical judgments about a user's expertise level by quickly and superficially scanning the recent activity (or lack thereof) visible on the profile.  In general, impressions fell into one of three expertise-based personas: complete newcomers, novice or inexperienced users, and competent peers.

Complete newcomers were distinguished by a lack of projects or any activity on their profiles, as well as a recent joining date that corresponded with their contribution.  This led to conclusions that the account

> *"was created on the same day it forks my repository you can say. And then you can see that, basically, just works on that for about five days, submits the pull request, and then he hasn't done anything since."  (P3).*

This lack of information often made it *"kind of hard to tell how good he is, actually.  He hasn't contributed to anything else" (P14).*

One participant mentioned that seeing a person with an empty profile led him to characterize them as a certain "type of person" who signed up to report bugs but didn't actually contribute any code themselves.  These people predisposed him to expect a certain type of contribution:

> *"you can usually tell when you go to a profile and they don't have any projects, and they just sign up to communicate with developers, basically.  Yeah, it's a bit more frustrating*

*because… sometimes you get bad bug reports or duplicated bug reports, and people who report them are-- I wouldn't say they're-- not negative, but they're-- yeah-- complaining, I would say." (P13).*

Novice or inexperienced users, in contrast, could potentially have started their own projects in addition to forking those of others. However, their lack of expertise was assessed based on the kinds of projects that they chose to work on or the types of projects they had started themselves. In the former case, one interviewee explained:

*"Mongo DB is a database that's relatively well known for having a lot of fundamental architectural flaws and performance problems. The fact that he may choose to use this thing, again suggests to me that he's not a very experienced developer." (P7).*

Finally, competent peers were judged, again, based on the breadth and depth of the projects they owned (and the coding languages they used.) One person described that looking at the range and type of languages on a contributor's profile primed him to look favorably upon that person's work:

*"he seems like a quite talented coder. I mean, I see Pascal, I see Erlang, and yeah, he used VI, so, I mean, he validates his geek cred for all those things. So yeah, that would give me a pretty good first impression on the person, regarding the way he knows how to code." (P4).*

### 6.8.2 Assessing project-relevant skills

In addition to the high-level coding abilities deduced from the project information described above, such as languages used, time active on site and amount of active contributions made to projects (either as owner or contributor), these profiles also helped observers form impressions about specific areas of expertise – that is, the types of work a person was best suited to doing or preferred to do (e.g. did they spend more time writing code or editing and managing projects?)

One person highlighted how seeing the portfolio of languages a user had on their profile could give some insight into what they might be able to contribute to his projects (which were written in C code):

*"if someone forks and all their other projects are web based PHP stuff, well you can sort of guess that you won't be getting any code patches from them. But if someone writes something and you see that their profile's all really hardcore C libraries, and C stuff, then you can sort of expect them to actually help out and write good C code, for example."* *(P3).*

Thus, profile cues could not only lead to conclusions about a person's general competence as a coder, but they could lead the project owner to form some expectations about the ways a person would be able to assist their project.

### 6.8.3 Forming personality judgments

Beyond assessing technical ability level, or a contributor's area of expertise, interviewees also described how other profile cues (such as being able to see recent comments a user had made or discussions they had been involved in) gave them a sense of what a person was like to work with based on how they interacted with others. Sometimes, this information revealed negative personality traits, or led to the conclusion that someone was difficult to work with, if there was evidence of aggressive or hostile exchanges with others:

*"I think you can kind of look at how they respond in threads, if they're arrogant or if they're trying to be helpful, it does show through a little bit in threads" (P15).*

On the other hand, other times there was no evidence of a bad attitude and project owners concluded that contributors would be easy to work with, e.g. *"a nice guy who wants to help" (P8).* In either case, personality cues as assessed through visible interactions with others helped to provide reliable insights into how the person interacted with others, because these were generated in the context of actual previous discussions.

## 6.9  Pull request case studies

To help illustrate how social media related cues impacted the work process, interviewees provided several examples of recent pull requests they had received and interacted with.  We identified ten recent pull requests interviewees had received from a contributor they had not directly interacted with before, as we were interested in the

role of impression formation in deciding whether or not to accept the code from these new people. Examples of pull requests from people interviewees had collaborated with or worked with regularly were thus not part of this analysis.

### 6.9.1 Accepted versus rejected pull requests

A comparison of accepted versus rejected pull requests suggested that uncertainty was a critical factor in the code contribution review process. Owners were more certain about the value of simple changes that addressed features the owner had wanted to add, were small in scope, or fixed a known bug. Owners were less uncertain about the value of code that was suggesting a larger change, introducing a new feature, or conflicting with other existing functionality.

Code that was accepted "as is" was often either straightforward and easily verifiable, or accepted as a matter of principle (A1-A5). Few comments were involved. Declined pull requests, on the other hand, were not immediately accepted although they had the potential to be if certain fixes were made (D1-D5). These requests involved a great deal more uncertainty regarding the implications of accepting the change. These pull requests required some back-and-forth discussion between owner and contributor to explain the reasons why the code couldn't be automatically accepted or would cause problems and then to negotiate a final outcome.

When the value of submitted code was more uncertain, project owners often engaged in an assessment that involved weighing both code-based factors (e.g. perceived value of the code) and person-based factors (e.g. the perceived value of encouraging continued and sustained participation in the project by the new contributor.) For example, the cost of working with someone to fix their code so that it could be accepted (which could be high in cases where contributors were newcomers or novices) was weighed against the potential benefits of helping to mentor a new project member and potentially gain help in the future, or the risks of being annoyed by time-consuming arguments with a novice about why their contribution was not acceptable. Figure 3 summarizes the pull request acceptance decision process.
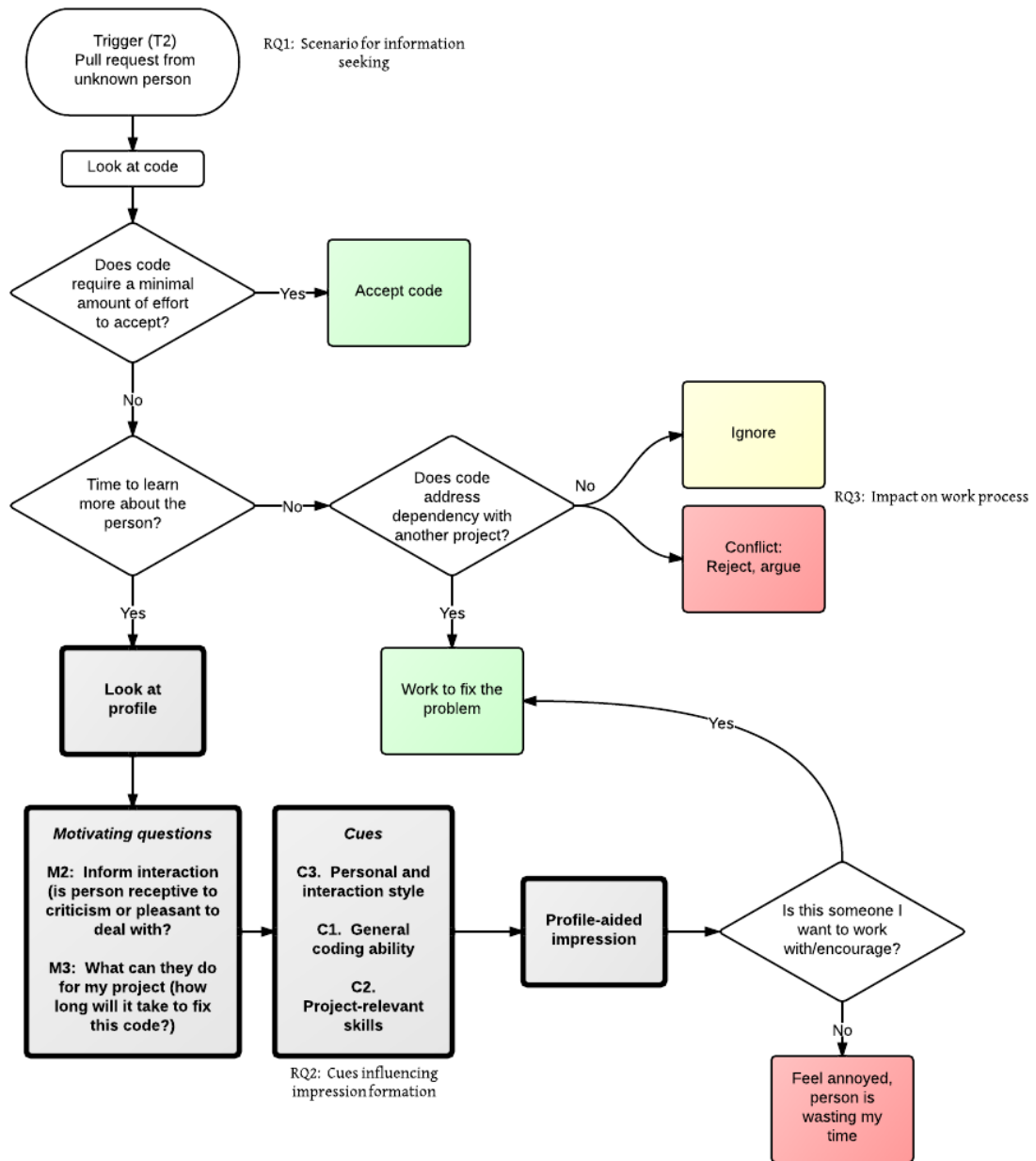
Figure 3.    Model of pull request acceptance process.

## 6.10 Impressions influencing mentorship: Examples

In this section I describe two of the five declined pull requests (D3, D4) where the owners reported having examined the profile of the contributor in the process of assessing the code. In both of these cases, the code being submitted was problematic in that it had stylistic issues, followed poor practices, and generally would require additional time and work by the project owner to fix to the point that it could be incorporated into the main branch of the project. However, the impressions they formed from glancing at the contributors' profiles influenced their willingness to accommodate the contributor's efforts.

When the owner of D4 consulted the contributor's profile, he discovered the contributor was a complete newcomer with no previous history. He described his subsequent reaction to this as:

> *"if I see this is their first pull request, I'm more like, 'Oh, thank you. Very nice.' And then try to be more gentle or…more friendly" (P14).*

The results of this impression were that as a result, he shaped his reaction to the code (which followed a poor practice of containing many different commits in one) to be more tolerant. In the end, the code that was submitted later proved to have some problems, so it was not accepted; however, the owner handled the issue in a way that illustrated his desire to encourage rather than sharply criticize someone whom he recognized to be a newcomer.

Similarly, in D3, the project owner received a pull request that was *"a bit problematic in the sense that it gives me a lot of work to accept a patch like this" (P3).* Since his assessment of the contributor, based on profile activity and interactions through the pull request was of "*an intelligent person who has not been coding too long, does not have too much experience in this*," he realized that this person was *"a good person to focus but needs hand-holding."*

For this reason, the owner was willing to look over his work, but rather than outright rejecting it he planned to look it over at a later time.

## 6.11 Impression accuracy: Examples

While there was evidence of project owners quickly forming impressions of new contributors and using these to frame their interactions with these people, an associated question was to what extent these impressions were accurate. In the six cases where project owners elected to consult the new contributors' profiles, two users had no evidence of activity on their profiles, one was assessed as a novice user, and three were deemed to be competent peers.

In four instances, participants made judgments of others' abilities without having consulted their profiles beforehand (or consulted them after the pull request interaction had occurred.) In three of these four instances, (D1, D2, D5), the default assumption made without looking at the users' profiles was that they were new to GitHub or to coding.

The owner of D2 described the contributor as:

> *"kind of a junior level experience guy, because anyone else would realize that these kind of changes are going to have a dramatic impact on this kind of project if they were more experienced." (P7).*

However, the sender of this pull request was also an interviewee for this study: He had been a professional programmer for over ten years and owned a popular open source project. In this case, had the project owner looked at his profile, he may have formed a different and more positive impression of the contributor's abilities.

There was also some evidence of people forming biased impressions, demonstrating some well-known cognitive biases and heuristics such as the primacy effect and confirmation bias.

Confirmation bias in the context of a negative first impression will lead to failure to revise the initial hypothesis, even in the presence of strong evidence to the contrary (Lindgaard et al., 2006). One participant's comment about an unknown contributor suggested a bias against a stranger potentially led to perceived lack of competence:

*"He's made some minor changes to Rails. Judging by his pull request to my project, I doubt they got accepted whatever they were."*

Since humans are "cognitive misers" (Fiske, 1980) who do not rely on more cues than necessary to form an impression, people may quickly judge others based on the first or limited information available to them rather than spending time to look through a variety of cues. For example:

*"I think if you read through his comments and actions that he does on GitHub here, you'll see that he's incredibly short-winded with how he says things to people…And this was the top interaction of this guy. I do not have to go look very far to figure this out."*

Thus, the impressions people were forming were occasionally inaccurate (e.g. the aforementioned example of P7 assuming a competent programmer was a novice in pull request D2.) This type of behavior has been shown to cause conflict in other settings such as Wikipedia (Zhu et al., 2013) and therefore, exploring ways of designing profiles to mitigate the effects of these biases was the focus of subsequent work.

## 6.12 Discussion and implications

Developers also sought out more information on others in response to uncertain code contributions. They sought out information about other developers' interaction style and interest to inform how they communicated with them. They also formed ability and expertise impressions based on profile information. These impressions influenced the way they handled project contributions, more so when the value of submitted contributions was uncertain, the contributor was unknown and future interactions were not expected.

Owners with a stronger tendency for helping would use this information to inform mentoring style interactions to improve a contribution. Bad contributor attitudes combined with low perceived skill and ability led to annoyance, inflexibility, long arguments, and delays. However, instances in which good attitudes combined low perceived ability suggested that attitude could trump poor skills by signaling that an initially contributor would be willing to learn and improve, and could be mentored to provide more useful contributions in the future.

Often, the project owners tended to assume that when unknown contributors sent something to a project for the first time, that they possessed inferior skills to the owner. This corresponds with the assertion made by Oreg & Nov (2008) that contributors need to have a reasonable level of expertise and to have this expertise made public in order for them to make a creditable contribution.

The decision of what information to display about users may depend on a site's goals and the behavior it desires to promote (for example, some sites like Wikipedia minimize visible information about contributors in an attempt to create a level playing field (Tausczik & Pennebaker, 2011). However, when direct cues are absent, it becomes difficult to identify authors and evaluate their expertise.

Findings from Study 2 suggest that detailed personal information can shape work outcomes in a peer production setting, particularly for complex contributions. The results from this study suggest that while details were used to infer personality and subjective characteristics of a contributor along with expectations of skill and contribution potential, these details were not always used in the most reliable way. Perhaps due to limited time and effort, project owners were nonetheless making cost-benefit tradeoffs by relying on stereotypes and heuristics, which, while efficient, may have led to misunderstandings and conflict that could have been avoided. For example, it was fairly easy to classify a contributor as a newcomer based on the first piece of information seen without taking the time to fully consider other pieces of information. The downside of this was the under-valuing of a person's skill set and failure to properly utilize this person's expertise for further contributions (the contributor gave up and left after some initial arguments about his contribution).

# 7 Summary of interview studies

Overall, my work with Studies 1 and 2 examined the impression formation process in two different scenarios within a common context (hiring vs. contribution evaluation in GitHub.) I was interested in the role of both interpersonal information-seeking goals to reduce uncertainty and the information people utilized to form impressions.

I also found evidence that within a common setting with access to the same range of information, the type of uncertainty observers had impacted what they looked for. In the hiring situation, one thing people were interested in determining was person-organization fit (the degree to which a candidate shared their values.) They were also more likely to be able to follow up with the candidate later via an interview. This led to a focus on general open source activity and involvement in side projects as assessment signals (that were also easy to evaluate).

In contrast, in the contribution evaluation situation, individuals had to balance two goals between accepting the work in the short term and encouraging participation in the long term. In this case, they utilized activity traces because these were useful to reduce uncertainty – by seeing how an unknown contributor had behaved in similar situations in the past, project owners could get some indication of how this potential new helper might be able to contribute and how they would react to criticism.

My qualitative studies suggest the cues provided about a person and the way in which these cues are presented can influence both what information appears to be most visually salient and the impressions formed about that person. Table 5 summarizes some of the design implications of the main site features as they were instantiated in GitHub.

| Study | Uncertainty | Information utilized |
|---|---|---|
| Study 1 (Hiring) | Person-organization fit | Work artifacts (side projects, popularity of projects, projects owned)<br><br>Involvement in open source (active participation on the site) |
| Study 2 (Contribution evaluation) | Effort required to fix a person's code contribution (and level of interaction with them required.)<br><br>Whether the person will be a useful contributor in the future | Work artifacts (for an overview of skills)<br><br>Activity traces (for details about interaction style and attitude) |

Table 5.       Role of site features across studies

In Study 1, there was not much use of the more detailed activity traces such as the chronological, detailed feed of recent activity. Despite the rich details that such a feed can provide, I found that in many of the hiring examples, people were relying on cues that were either easy to quickly scan (e.g. activity level) or easy to verify (e.g. number of forks or watchers).

This reliance on easy to find/easy to verify cues suggests that the plethora of information, particularly that which is provided by the activity traces, was not used to its full potential. This makes sense given that people form impressions quickly and are often limited by time. The downside of this selective information source seeking is that it could lead to biased impressions if the subset of information attended to is not representative of the quality being assessed.

In Study 2, biased impressions tended to surface most in terms of viewing activity traces (such as the activity history feed). In the pull request examples, project owners mentioned some cases where they looked at the activity traces of unknown contributors after they assessed the person's work in detail. Their comments indicated that their impressions of a code contribution often came first (before looking at the profile), with

no further information about the author. Impressions of the code often colored their impressions of the person that they subsequently formed by looking at their profiles (often interpreting activity trace information in a selective way that corresponded with what they first thought.)

Study 2 highlighted examples of mistaken (or possibly mistaken) impressions made by project owners in correctly assessing the abilities and personalities of unknown contributors. In many cases, this may have been due to a limited amount of time or motivation available for carefully considering all the available information about a new person. The three main types of problematic impressions that occurred, based on the examples provided by interviewees, fit into the following categories:

1) Not consulting the profiles of the contributors and instead drawing negative impressions based on code that was effortful to evaluate
2) Looking at just one (recent) piece of information at the top of the activity history
3) Looking at a few pieces of information and stereotyping a contributor as newcomer, novice, or expert (and then failing to revise this categorization even when faced with new information).

All three of these problematic impressions suggest the use of heuristic processing and shortcuts in the face of large amounts of information. These problems could potentially be ameliorated through interface design manipulations that make information about a user more visible, more accurate or easier to interpret. For example, a summary visualization that showcases an individual's history across many instances (rather than just one) could help to reduce the susceptibility to the primacy effect. On the other hand, a visualization that encourages observers to process the information more carefully might reduce the reliance on heuristic processing. This, contrary to the summary suggested above, might involve a more detailed display of information.

The interviews in Study 2 raised interesting questions about the impact of access to activity history, and the degree to which this information is summarized or provided in detail. Study 3 built on these observations in order to further empirically examine the impact of information presence and visualization on impressions. Specifically, it

focused on the potential of activity trace design to combat bias through summarization or abstraction.

# 8 Study 3: Activity visualization design experiment

This work will appear as: Marlow, J. & Dabbish, L. (in press). The effects of visualizing activity history on attitudes and behaviors in a peer production context. *To appear in Proceedings of CSCW 2015.*

In Study 1, I found that ease of assessment or verification of a cue was a factor driving which pieces of activity information were used (e.g. overall activity level as seen in graphs was used to infer enthusiasm as opposed to activity level assessed through the activity feed). In Study 2, project owners utilized both visual summaries of information (in the form of graphs) as well as the text-based public activity streams to assess unknown contributors. The activity streams provided some useful information that was not otherwise available (e.g. interaction style from prior discussions) but interviewee comments suggested that they attended to only one or two most recent pieces of information in the activity feed.

This suggests that there could be a tradeoff between impression accuracy and processing effort, and that the design and presentation of information about an individual's past history can influence what people look at and what they conclude based on this. In order to empirically test this tradeoff, I conducted an experiment in which I manipulated the amount and quality of information about another person to see its effects on impressions and subsequent behaviors.

This experiment was carried out in the context of a lightweight peer production setting (Amazon's Mechanical Turk). This was done for several reasons. One, it provided the ability to control the task setting and manipulate the conditions in a controlled way. Second, as Haythornthwaite (2009) mentions, in a lightweight peer production context, the system can provide reputation statistics about contributors but these must be designed into the system by organizers. Based on the results from Study 2 which suggested observers were subject to bias when they had limited time and effort to expend to assessing new contributors, I was interested in manipulating and observing the effects of system cues on perceptions of expertise to see if reducing the processing

effort of information helped by making it easier to get a quick overview, or if by removing details a visual summary reduced the helpfulness of activity history.

Prior work indicates that information visualization can influence individuals' opinions and attitudes. For example, this has been explored in areas such as making visual judgments based on social proof cues (Hullman, Adar, & Shah, 2011b) and judging the trustworthiness of Wikipedia articles based on visualizations of edit history (Chevalier, Huot, & Fekete, 2010; Suh, Chi, Kittur, & Pendleton, 2008). Of particular interest to this work are the effects of level of granularity versus abstraction (e.g. detailed textual display versus abstracted visual display). In particular, there is a tradeoff between providing additional detail (which can provide useful context) and providing this information in a succinct, summary format to increase processing fluency, minimize cognitive load and reduce effort of interpretation.

## 8.1  Overview of experiment

Exposure to activity history may affect how people evaluate work done by an unknown individual, particularly in a peer review setting where individuals review and correct each other's work.

I conducted an online experiment looking at the influence of activity history presentation on evaluation of an unknown individual's work in an online peer review/peer production context. My research questions were as follows:

> RQ1.  How does the design and presentation of work history information influence final impressions of that person and editing behaviors on their work?

> RQ2.  How does the valence or quality of an individual's work history influence initial attitudes and behaviors towards that person?

> RQ3.  How does subsequent experience with the individual's work artifacts influence final impressions of that person and effort expended to improve their work?

My study builds upon previous research in the areas of online impression formation and peer production by examining how interpersonal impressions can change over time and how they relate to work-related decision-making and evaluation. The experiment builds off of Study 2, where I found that impressions formed based on an unknown person's activity history influenced and may have possibly biased reactions to their work.

In order to create an experimental scenario that was as close to the GitHub scenarios as possible, these experiments employed a contribution-evaluation task designed to parallel the pull request submission process. I used a peer review task common in both lightweight and heavyweight peer production ( Benkler & Nissenbaum, 2006) and is often used in crowdsourcing contexts as a quality control mechanism (Hansen, Schone, Corey, Reid, & Gehring, 2013).

The goal of this study was to understand how the design of an activity history influences the first impressions formed and subsequent reactions to work. Initial findings from Studies 1 and 2 suggested a tradeoff between the amount of granularity being provided in activity histories and the cognitive capacity of an observer to attend to it all. For example, in both Study 1 and Study 2 there was evidence that observers would selectively focus on the first, most salient, or easiest to evaluate bit of information rather than trying to summarize and interpret it all.

Therefore, the goal of the experiment I conducted was to understand the attitudinal and behavioral effects of exposure to information about a person's past work history presented in different formats. In the next section, I present a review of alternative ways this information is currently presented and visualized on open source related sites, and then discuss the theoretical implications of different design parameters and choices.

## 8.2 Hypothesized effects of activity history information visualization and work quality

### 8.2.1 Design parameters: Detail versus visual abstraction

The design of activity histories may influence impressions of contributors because it can make certain information more accessible or change processing effort. There are several

key aspects of activity history design that may influence how it is processed and in turn what impressions the viewer forms of the person whose activity is represented including the amount of information shown, the simplicity or visual complexity of the way information is displayed, and the degree to which the information is presented in a granular and detailed (text-based) format vs. an abstract (visual) format (Thompson, Rantanen, Yurcik, & Bailey, 2007). In this experiment I focused on the latter dimension of granularity versus abstraction, since this is one of the primary factors that must be considered when displaying a large amount of information.

The activity history information provided in the GitHub interface is presented in a chronological, text-based list (see Figure 4), which contains detail but, as Studies 1 and 2 suggest, also is effortful to interpret. However, there are alternative ways of presenting this information. The GitHub interface also presents information about individuals' overall activity level in a visually abstract and aggregated format (see Figure 5).
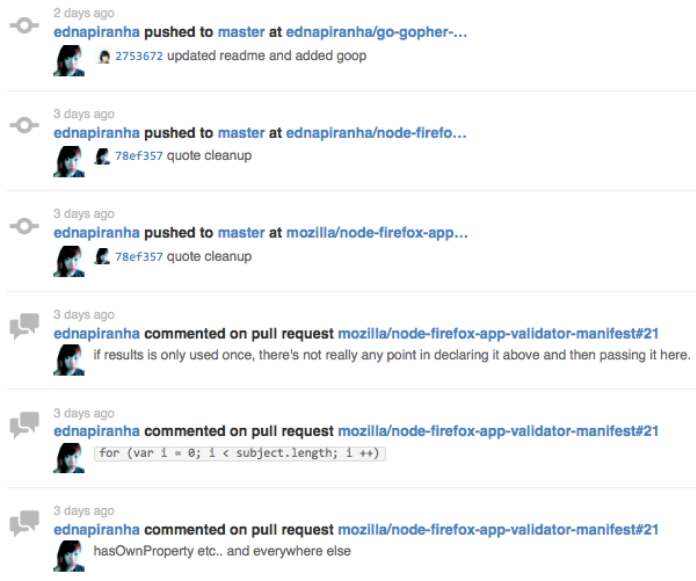
Figure 4.    Sample of text-based, chronological list of activity traces (detailed textual display)
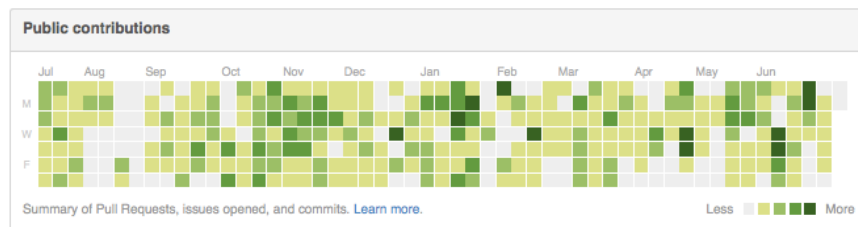


Figure 5.    Sample of visual, abstracted summary of activity traces

Other sites like Wikipedia and Stack Overflow make different design choices in terms of how they present information about users' activity history in varying degrees of abstraction or granularity (for example, Wikipedia activity traces are detailed text-based lists, whereas Stack Overflow utilizes a text-based feed, graphs, and visual icons to showcase various aspects of an individual's profile). There are tradeoffs between aggregating information about activity, such that focusing only on quantitative properties of user behavior (e.g. reputation score, number of answers) can cause people

to misclassify active users for knowledgeable ones, or be subject to manipulation (Yang, Tao, Bozzon, & Houben, 2014).

The implications of viewing information about this person in each of the aforementioned formats could be very different. In the following sections I discuss hypothesized effects of a detailed versus aggregate summary presentation of work. I also consider how the quality of the work history shown could influence both first and later impressions of a worker's quality.

### 8.2.2 Design and presentation of work history

RQ1 focused on how the design and presentation of work history information influence final impressions of that person and editing behaviors on their work.

Past work in a variety of online domains have explored the effects of using simple interface interventions to help people make sense of detailed and potentially overwhelming information about the past behavior of individuals and the artifacts that they create. For example, Liao & Fu have used various indicators of unknown online forum commenters' attitudes and expertise to help increase accuracy of perceptions of these commenters' positions and knowledge (Liao & Fu, 2014a; Liao & Fu, 2014b). Other research in the realm of peer production has focused on different ways of presenting and displaying information about work history. For example, work in Wikipedia explored ways of visualizing an article's edit history to indicate quality (Chevalier et al., 2010).

Figure 6 through Figure 9 showcase examples of previous activity visualization styles in previous work looking at visualizing Wikipedia and online forum history. As summarized in Figure 10, these visualizations vary on two major axes: First, they vary in the degree to which they use textual information to provide information about user behavior (low abstraction) or present it in a more visually summarized way (high abstraction). Second, they vary in the degree to which they provide the information at a very detailed, action-by-action level (high granularity) or provide an overview of aggregate behavior (low granularity).

| A: 2007/07/17 13:03:40 | /* Early life and education */ clarify the fish processing factory bit |
| T: 2007/07/17 13:00:57 | /* After Yale */ |
| T: 2007/07/16 18:09:42 | /* First Lady of the United States */ |
| A: 2007/07/16 18:08:27 | /* First Lady of the United States */ rm busted, duplicate image |
| A: 2007/07/16 04:37:50 | /* Early life and education */ common citation |
| A: 2007/07/16 04:37:03 | /* First Lady of the United States */ add a citation |
| A: 2007/07/16 04:25:27 | /* First Lady of the United States */ move in Eleanor Roosevelt material, mostly as fool controversies]] |
| A: 2007/07/15 07:01:36 | /* First Lady of the United States */ move Travelgate and Filegate here, from being-dis |
| A: 2007/07/15 06:34:05 | /* Marriage and family, law career and First Lady of Arkansas */ |

Figure 6.    An example of a low abstraction (text-based) and high granularity display.  This is an example of the "detail edit log" of Wikipedia editor behavior from the WikiDashboard tool (Suh et al., 2008).
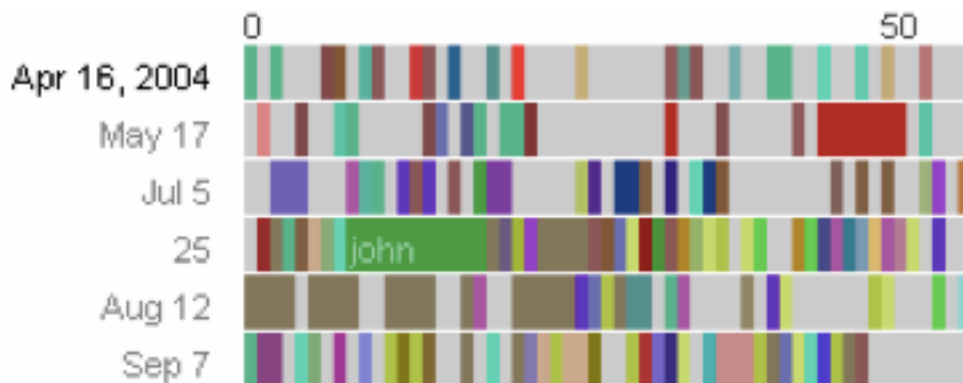


Figure 7.    An example of a high abstraction (visual summary-based) and high granularity (detailed) display in the form of a chromogram visualization of Wikipedia edits over time (Wattenberg, Viégas, & Hollenbach, 2007).

Figure 8.    A low granularity (text) and low-abstraction visualization of Wikipedia edit history in the form of a word cloud (McDonald, Gokhman, & Zachry, 2012)



Figure 9.    A low granularity and high-abstraction graph-based visualization summarizing an aggregation of a user's opinions across various interactions (Liao & Fu, 2014b).

High granularity (Detailed presentation)

Low abstraction (text-based)

High abstraction (visual summary-based)

Low granularity (Overview presentation)
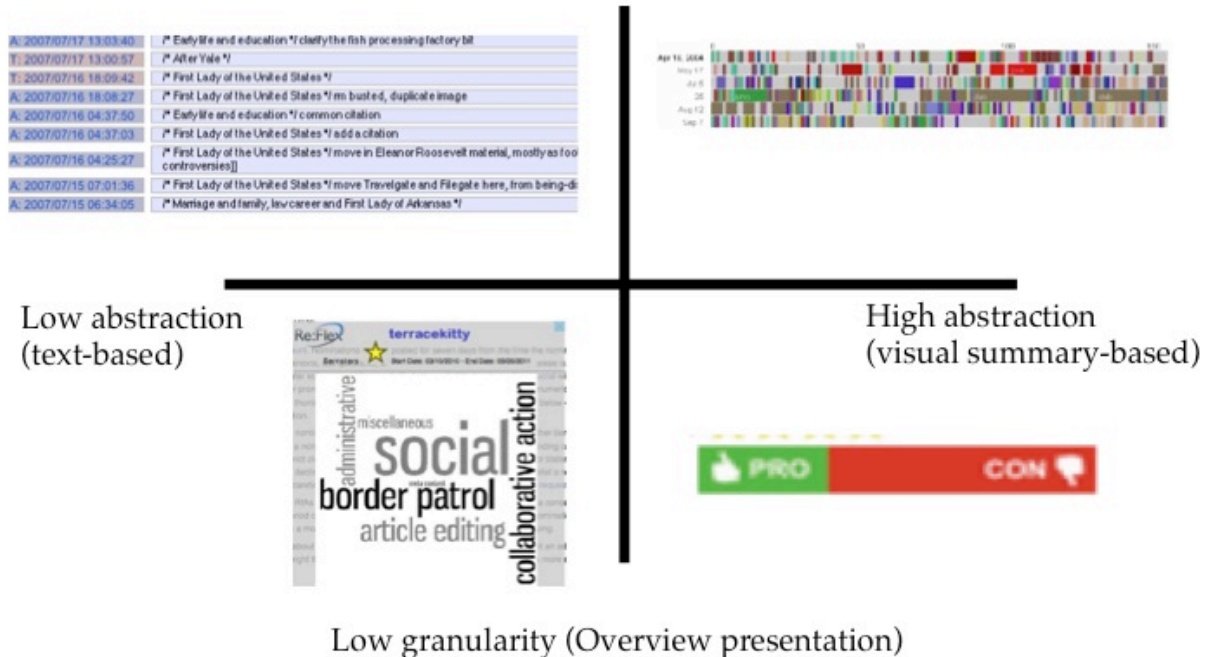
Figure 10.     Examples of activity history visualizations from prior work
categorized by abstraction and granularity

The design of these visualizations make tradeoffs in terms of how much information they show in what format, and as a result how they influence decision-making and performance. On one hand, detailed and granular information provided in a text-based format provides direct access to more fine-grained details and gives an overall more detailed picture of the information space (Thompson et al., 2007). The detail provided by text can be useful to reduce uncertainty (Berger & Calabrese, 1975) and give observers a more concrete and realistic picture of behavior. On the other hand, it is generally accepted that information overload can be a common problem in a text-based environments (Liao & Fu, 2014b). When there is a great deal of textual data, this can result in overload and make the interpretation of information more difficult and cognitively intensive (Thompson et al., 2007).

A common alternative to presenting information in a granular or text-based format is to use a visual summary or abstracted overview of the information. Visual summaries can help in a variety of ways, for example, by reducing the cognitive burden of interpreting detailed text through simplicity and ease of use. However, the lack of detail can make it more difficult to understand what is going on (Thompson et al., 2007).

Text-based lists of activity information are detailed but require effort to interpret (Thompson et al., 2007). Detail can be useful and an appropriate way to present information, such as in a leaderboard or a train schedule. The London Underground map previously was very detailed in terms of geographic and topological features and locations of stations until it was redesigned in the 1930s in a more abstract way. This visually abstracted representation of the same information is easier to process quickly, and provides some benefit in terms of helping people find the most direct routes. However, by abstracting out the notion of real distances, it can be less informative for finding the quickest or shortest routes (Ainsworth, 1999).

As the Underground example illustrates, there are important tradeoffs between detailed, granular, text-based information and visually summarized or abstracted information in the design of information visualization. The tradeoff between detail and abstraction informed the focus of this experiment as well.

Large volumes of information can be presented in great detail or aggregated. This has implications for the amount of processing effort it takes to make sense of the information. For example, work that is harder to perceive (such as text written in a font that is hard to read) requires more mental effort to process, and this has been found to lead to more negative impressions of the text's author (Oppenheimer, 2006). This finding suggests the easier a display is to process, e.g. the less detailed, the more positive feelings it may engender:

H1a: Reducing the amount of detail in the work history will increase impression positivity. Reducing detail by summarizing the information could be beneficial, by reducing information overload (which can impose a ceiling on accuracy (Tsai, Klayman, & Hastie, 2008) and make people more subject to bias).

At the same time, even uninformative details may increase the clarity of the judge's mental picture of the target (Tsai et al., 2008) and reduce uncertainty about the target (Berger & Calabrese, 1975), which have been associated with increased positive attitudes towards others (Tanis & Postmes, 2003). I therefore hypothesize that detail could provide a more complete image of the person as a concrete entity, and that visual abstractions could reduce the mental processing effort.

H1b: Increasing the amount of detailed information about another worker will increase impression positivity.

### 8.2.3 Effect of history quality on first impressions

RQ2 was about how the impact of exposure to information about an individual's work history might influence initial attitudes and behaviors towards that person. On one hand, making activity history visible can provide useful insight into details of behaviors that were inaccessible and encourage more careful review of difficult cases (Hansen et al., 2013). Being able to view the past work activity of collaborators can also help a new project member understand their collaborators' intentions and decisions through providing increased understanding of prior work history (D. Kim & Shipman, 2012).

Typically, only very limited information is available about others in peer production contexts like Wikipedia or social Q&A sites. As Tausczik & Pennebaker (2011) point out, minimizing information about a person's activity (e.g. what they have done in the past) can give individuals with little or no reputation the opportunity to produce high quality content.

The impact of providing information about a stranger's work history on attitudes and behaviors towards that person is not well known, although qualitative research suggests that this information would improve accuracy of perceptions. In creative collaborations, Luther et al. (2010) found that access to an individual's past work influenced how project leaders evaluated candidates. They equated a high "batting average" (reputation) and evidence of a long activity history with competence and expected them to be able to contribute to a project's success. Similar patterns have been observed in open source software development (Marlow et al., 2013), but these findings have not been experimentally tested.

One advantage of access to an individual's work process is the fact that an aggregated history of behavior should, overall, provide a good basis for reputation even if an individual might behave differently across specific interactions (Anderson & Shirako, 2008).

However, the quality (that is to say, perceived positivity or negativity) of the activity history's contents could make a difference. Work looking at the effects of exposure to different types of discussion histories for Wikipedia articles suggests that access to more information about how an article was written (specifically, how much conflict there was between editors) can shape readers' attitudes towards the quality of an article (Towne, Kittur, Kinnaird, & Herbsleb, 2013). In their study, Towne et al. found that visibility of negative conflicts among authors of an article decreased readers' perceptions of the article quality. However, we do not know how exposure to conflict affected perceptions of the articles' authors.

I therefore hypothesize that if work history is seen as indicative of an individual's past performance, then:

H2: Better work history will increase positivity of impressions of the person.

### 8.2.4 Enduring effects of exposure to activity history

RQ3 was about how subsequent experience with the individual's work artifacts after viewing their work history might influence final impressions of that person and effort expended to improve their work.

The potential disadvantage of access to work history is that people may be subject to bias based on whatever they see. When viewers are not motivated or unable to systematically evaluate information, they may be subject to using heuristic shortcuts or cues instead (Metzger et al., 2010). For example, the confirmation bias can cause individuals to form an initial impression and interpret subsequent information in light of this impression. This can be a problem when making social judgments in cases where the initial impression is inaccurate (Hullman et al., 2011b).

In general, the valence of information may also affect the nature of biases formed. While positive stimuli may elicit positive inferences (a positivity bias), and negative

stimuli may elicit negative inferences (a negativity bias) (Peeters & Czapinski, 1990), negative information is often stronger and more influential than positive information (Carr & Walther, 2014). Past work suggests that negative first impressions are often more resistant to change, and negative information is more likely than positive information to have an enduring effect (Kanouse, 1984).

Prior research on interactions around contributions in peer production provides some evidence of the negativity bias occurring, particularly towards contributors who are perceived to be novices (Marlow et al., 2013; Zhu et al., 2013). Based on this, we might expect that a negative first impression of a person will be less likely to change than a positive first impression, even in the face of disconfirming information. However, it could also be that observers will change their impressions after exposure to new evidence. Research on impression formation online suggests that while first impressions are often unexpectedly robust, they can be both confirmed and disconfirmed (Walther & Tong, 2014).

H3a: Initial impressions will persist even in the face of contrary evidence.

H3b: Initial impressions will change after new evidence is revealed.

## 8.3 Design of experimental visualizations

Based on the principles of visualization described in the previous section, I designed three visualizations to test in the experiment. The details of the final visualization choices appear in Section 8.4.2.1. The design process behind them involved a survey of related formats and a desire to match current practice as well as incorporate and utilize insights from past research.

The design of the experimental visualizations was a multi-step process. To understand how alternative visualizations of activity history currently prevail on the Internet, I conducted an informal survey of sites similar to GitHub to see how other web interfaces presented complex information about a developer's skills and activity history. I selected Stack Overflow, Masterbranch, and Open Hub as additional examples of sites that may make different design decisions in terms of how to display and aggregate user activity (see, e.g. Singer et al., 2013 for discussion of "profile aggregators" and users'

reactions to different design decisions in evaluating programmers.) Examples of different presentations can be found in the Appendix. I analyzed the different ways of presenting activity history and found that existing sites commonly presented information about a person in three ways: (1) in a summary format, providing aggregated statistics; (2) in a detailed, text-based list of recent activities and the details therein and/or (3) in a more visually abstract way that used visual representations of activity either alone or supplemented with text (e.g. graphical representations of commits to various projects or distribution of coding languages across projects).

These three formats corresponded to the text/abstraction and detail/overview design parameters that I was interested in exploring. After examining common ways of presenting the information, I focused on three types: on a text-heavy, detailed visualization, a hybrid text/visual abstraction overview, and a less detailed abstract overview. High-level characteristics of the visualizations (inspired by Hornbæk & Hertzum, 2011 and other examples from the Appendix) appear below in Table 6. Low-granularity and low-abstraction was not considered as an option because it did not match up well to current instantiations of history visualization on the Web in the sites surveyed, and therefore it may be more difficult to parse due to unfamiliarity.

|  | High abstraction (visual) | Low abstraction (text) |
|---|---|---|
| High granularity (detail) | Graph format with text and visual features such as bars, colors | Table format with text, lacking in visual cues |
| Low granularity (aggregate) | High-level aggregation with use of color coding and minimal text | N/A |

Table 6.        General principles for designs of visualizations

## 8.4  Method

### 8.4.1  Procedure

I conducted an online experiment on Mechanical Turk where participants first viewed a worker's past history and then went on to rate and edit new examples of their work.

Participants were recruited from Amazon's Mechanical Turk to complete what was advertised as a "transcription evaluation" task. The task was posted on January 6, 2014 and we provided a payment of 25 cents for successful completion. Participants were

told that the requesters were having people on Mechanical Turk type in the text from images of signs to help compile a database of transcriptions, and that we needed help from other Turkers to check them for quality.

Participants were then told they would first see some information about the past performance of the person whose work they were about to evaluate, including how other "editors" had evaluated their work. After this, participants were then told that they would be shown six new pieces of work that had not been evaluated yet. They were asked to rate the quality of the work and had the option to edit it if needed. After rating the six pieces of new work, participants completed a short questionnaire in which they evaluated the overall quality of the worker, gave a recommendation for using this person for future tasks, and provided demographic information such as age and gender.

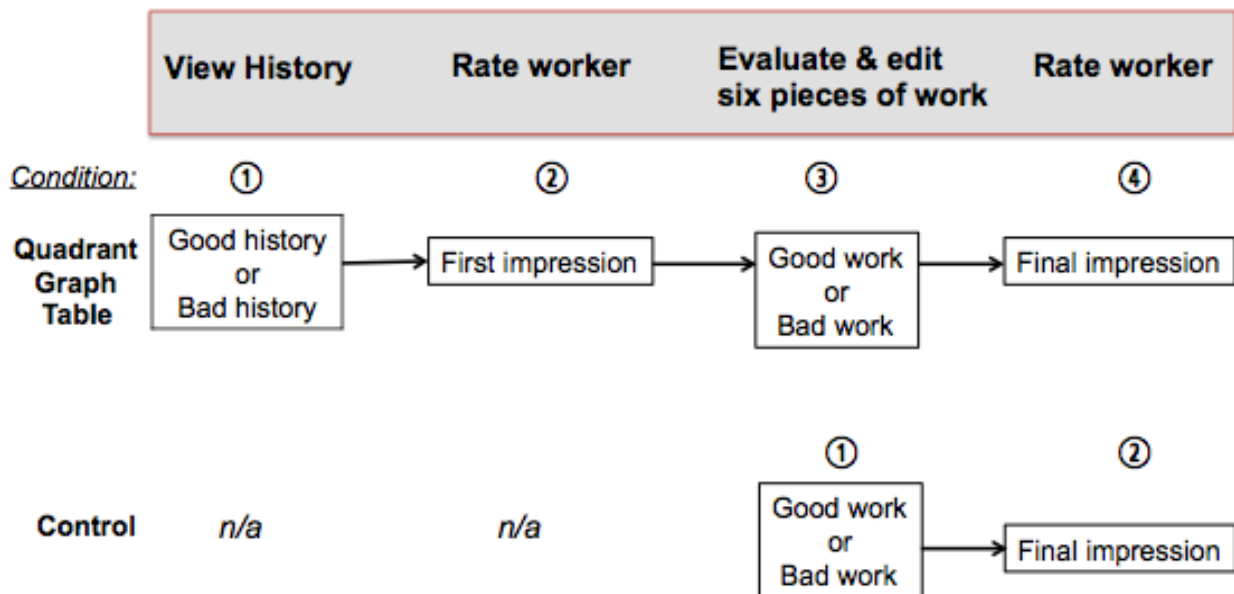Figure 11 summarizes the overall experimental flow:



Figure 11.    Overall experimental flow and manipulations

### 8.4.2 Independent variables

I manipulated three independent variables:

To test H1a and 1b, which were about the role of visualization of information on impressions, I designed three types of work history visualization format (*table*, *graph*, or *quadrant*). There were also two *control* conditions, where participants did not see any history information and evaluated either good or bad work.

In order to test H2, which was about the effect of initial work history quality on first impressions, there were 2 types of work history quality (good or bad).

Finally, to test H3a and 3b, which were about the persistence of initial impressions in the face of new evidence, I manipulated the quality of work to be evaluated (good or bad).

There were therefore a total of fourteen potential conditions and each participant was randomly assigned to one of these.

## 8.4.2.1 Work history visualization

Participants in the work history conditions saw information about how the worker they were evaluating had previously performed on transcribing six images, along with how their work was evaluated by another editor. The visualizations differed in terms of the amount of detail they presented and the degree to which this was listed in textual format or displayed in a more abstract, visual way. They were inspired by and based upon the features from GitHub profiles observed in Study 1 and Study 2 (refer to Figure 4 and Figure 5) and other sites as well (refer to Appendix).

The *table* condition (Figure 12) involved a grid filled with the text the participant had typed, the correct answer, the rating by the editor, and the percent of the work that had been completed.

## Work history for User 650

| Image | Correct text | Work from User 650 | Amount of text typed | Evaluation of work |
|---|---|---|---|---|
| 5 | SWEET BASIL & La FILIPINIANA | SWEET BASIL AND L. FILIPINIANA | 100% | MINOR EDIT: Typos or incorrect letters |
| 7 | GMITER | GMITER | 100% | ACCEPT |
| 8 | RESERVED PARKING THE RIGHT ONE UNAUTHORIZED VEHICLES TOWED AT OWNER'S EXPENSE | THE RIGHT ONE | 16% | REJECT: More than 50% of text is missing |
| 10 | BENKOVITZ SEAFOODS | BENKOVITZ SEA FOODS | 100% | ACCEPT |
| 13 | SUPER DOLLAR STORE WE ARE OPEN DURING DEMOLITION | SUPER DOLLAR STORE | 39% | REJECT: More than 50% of text is missing |
| 15 | DON'T BE A LITTER BUG HELP KEEP EDGEWOOD CLEAN | DON'T BE A LITTERBUG | 44% | REJECT: More than 50% of text is missing |

Figure 12.    Sample of table condition

The *graph* condition (Figure 13) contained the same information as the table condition, except the percent of the work completed and the rating by the editor were displayed in a visual format (through colored bar graphs.)



Figure 13.     Sample of graph condition

Finally, the *quadrant* condition (Figure 14) was the most abstract and plotted the participant on a grid showing the average completeness of their work and the average evaluation of their work.

Work history for **User 660**
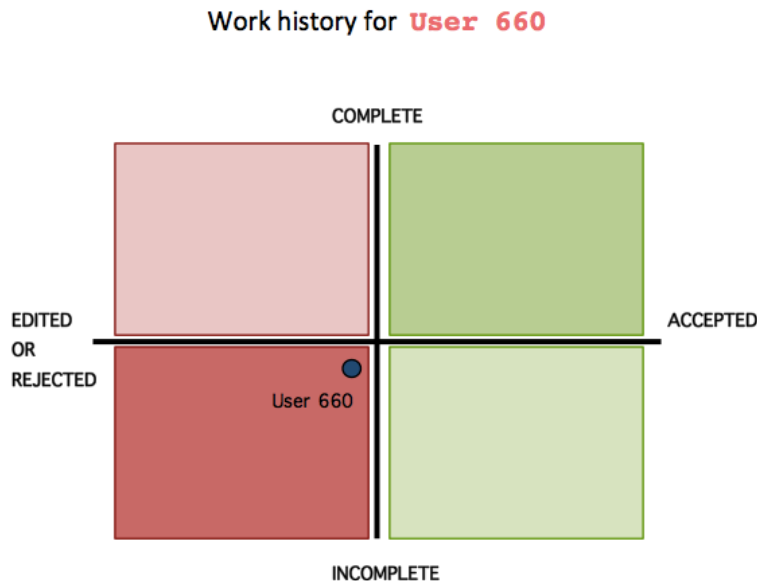


Figure 14.     Sample of quadrant condition

Because I was interested in varying the amount of detail and in varying the difficulty of processing the information, the table was designed to be the most detailed but also the hardest to process. The graph was designed to present activity information in a visual summary, using color and visual cues to help make it easier to process. The quadrant was designed as an alternative, easy to process visualization to make sure the particular design of the graph itself was not responsible for any observed effects. I used two different visual summaries to ensure that any differences between the table and the visual summaries was due to the difference in abstraction rather than the particular representation chosen.

## 8.4.2.2 Work history quality

I was also interested in how the quality of past work influenced attitudes towards subsequent work (H2). Therefore, I manipulated the quality of the work shown in the past history and in the six examples. The work shown was taken from actual activity history taken from Turkers who had performed the image transcription task previously.

For the "good" quality work, I used the responses given by a Turker who performed in the top 75% of the task. For the "bad" quality work, I used the work provided by a Turker who performed in the bottom 2.5% of the workers.

Figure 15 shows an example of the three visualization formats (in this case, showcasing "bad" work history). Participants in the control condition were not exposed to any of these work history visualizations. They viewed a neutral image (to ensure that they were still looking at something prior to beginning the work) but this was unrelated to the person whose work they were about to evaluate.
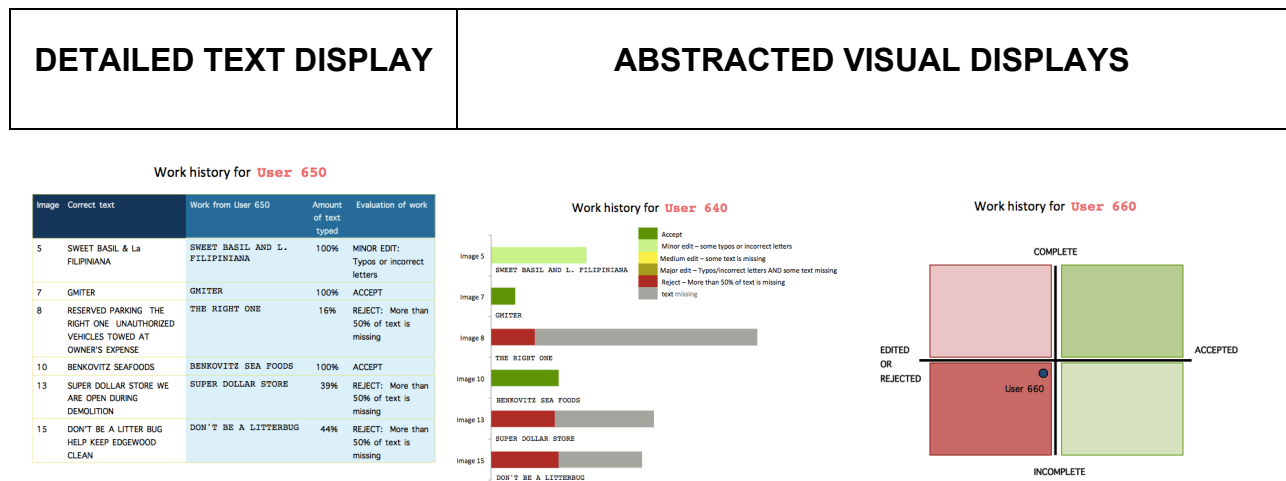


Figure 15.    Examples of bad work shown, from left to right: detailed text display (table) and abstracted visual display (graph, quadrant) formats

Figure 16 shows an example of all six visualization formats.



Figure 16.    Examples of work shown (from left to right) in table, graph, and quadrant format.  Good work history is in the top row, bad work history is in the bottom row.

### 8.4.2.3 Stimuli used for work quality evaluation

For the work task, participants were shown six images along with the worker's performance transcribing the image. The instructions for evaluating the work were as follows: *We would like the text transcribed to be as complete a representation of everything in the image as possible. The capitalization of letters does not matter.*

Figure 17 shows one of the sample images used along with examples of "good" and "bad" work.



Figure 17.      Sample image for transcription

Example of correctly transcribed text: `STAGNO'S BAKERY INC.  STAGNOS BAKERY`

Example of "good" work: `STAGNO'S BAKERY INC. STANGNOS` [missing 6 characters]

Example of "bad" work: `STAGNO'S BAKERY` [missing 18 characters]

### 8.4.3  Dependent variables

There were four dependent measures:  1) *First work impression:*  The initial impression of the quality of the Turker's past work, measured on a scale from 1-100, after viewing the history; 2) *Work evaluation:*  The mean of the six quality ratings given for the pieces of

work seen, on a five point Likert scale; 3) *Amount of work fixed:* A measure of relative amount of work edited, calculated as the average increase in number of characters from the original work to the participants' answer, and 4) *Final impression:* The aggregate sum of four five-point Likert scale items assessing overall impressions of the quality of the Turker's work.

## 8.5  Results

Overall, 284 participants completed the task (which had an overall completion rate of 82%). 55% were female, and 45% were male. 59% were located in the USA, 35% in India, and 6% in other countries.

### 8.5.1  Effects of visualization type on initial impressions

To compare among the three visualization types on first impressions of work quality (testing H1a and H1b), I conducted a one-way ANOVA comparing the three visualizations on first impressions of worker quality. There was no significant main effect of visualization type on first impressions of worker quality ($F(2,216)=1.53$, $p=.21$). However, planned contrasts between the table and the two visual summary types of visualizations (graph + quadrant) revealed that the table (aka detailed format) led to more positive first impressions of worker quality than the two summary conditions ($F(1,213)=10.93$, $p<.001$) (see Figure 18). This lent support to H1b, which predicted that more detail would lead to more positive attitudes
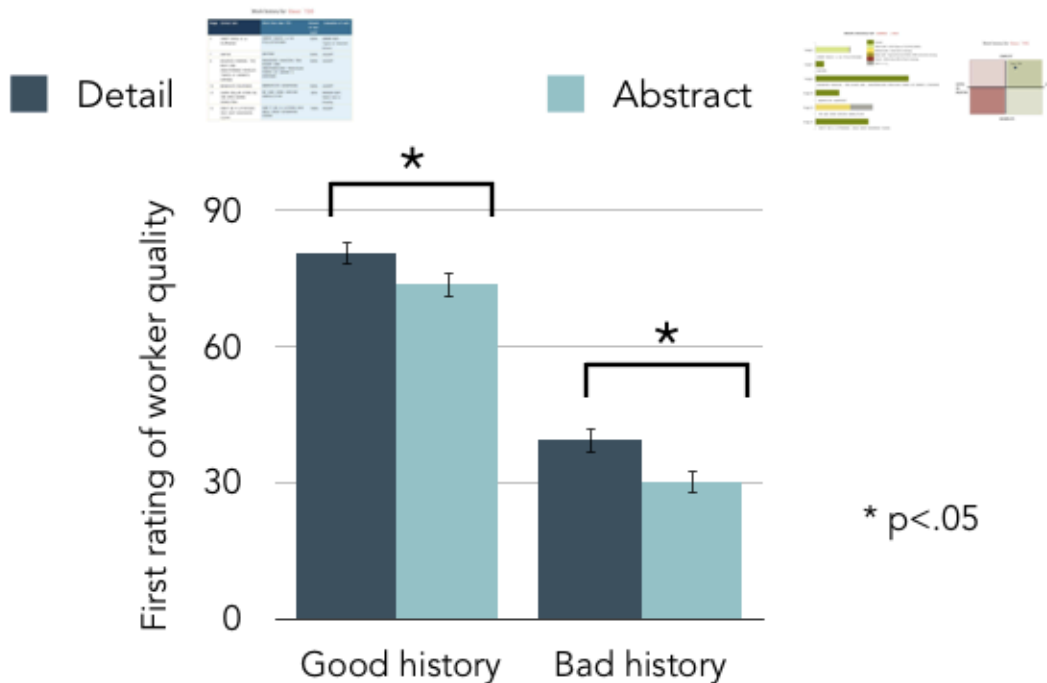
Figure 18.     First ratings of worker quality by visualization type and work history quality

### 8.5.2  Influence of history quality on first impression

RQ2 focused on the degree to which an initial impression was subject to influence based on the quality of work history.

I initially predicted that the valence of work history would influence initial impressions of the person.  I conducted a one-way ANOVA to assess the effects of work history on first impressions (on a 100-point scale, with 100 being the highest quality.) This was supported; participants exposed to the "good" work history gave a significantly higher rating of the quality of the Turker's past work (75.94 out of 100) than participants exposed to the "bad" history (33.39 out of 100) (F(1,217)=329.72, p<.001).

This served as a manipulation check and suggests that participants were able to fairly accurately assess work quality and correctly interpreted the history visualizations.

It also supported H2, which stated that a better work history should increase positive impressions.

### 8.5.3  Persistence of initial first impressions after exposure to work

Finally, Hypotheses 3a and 3b focused on the persistence of initial first impressions after direct exposure to work. In order to test these, I first subdivided the data based on whether the transcription work was good or bad. Then I ran a series of ANOVAs on good work and bad work separately in order to see if *what* people saw beforehand (no history, good history, or bad history) and *how* it was presented influenced their reactions when faced with equivalent work quality.
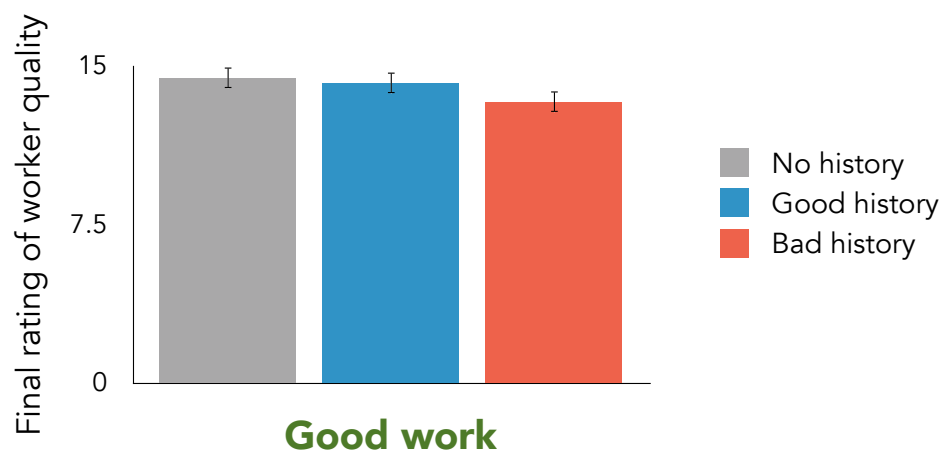


Figure 19.    Effects of history quality on final impression of worker (good work)

As indicated in Figure 19, when the work quality was good, the quality of history seen beforehand did not make a difference: There were no significant differences in final impressions of the worker across no history, good history and bad history conditions ($F(2,137)=1.90$, $p=.15$).
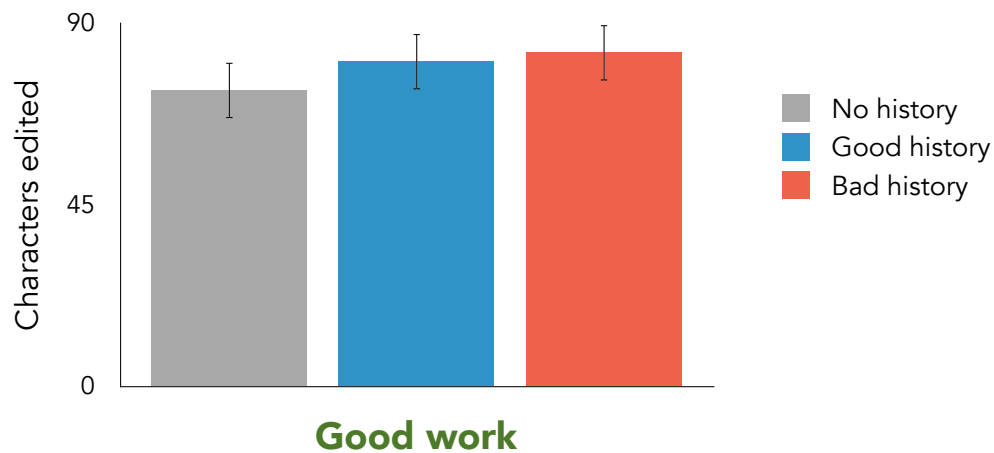
Figure 20.    Effects of history quality on work editing behaviors (good work)

As indicated in Figure 20, when the work quality was good, the quality of history seen beforehand did not influence the amount of work fixed:  There were no significant differences in editing behaviors of good work across no history, good history and bad history conditions ($F(2,137)=.44$, $p=.64$).

These findings suggest that when editors were exposed to evidence of good work, the work tended to speak for itself.  Evidence of good work seemed to override any initial effects induced by seeing no history, a good history, or a bad history, and led to similar patterns in final impressions and editing behavior across these conditions.

However, a different dynamic emerged when analyzing the influence of history on reactions to a bad worker.
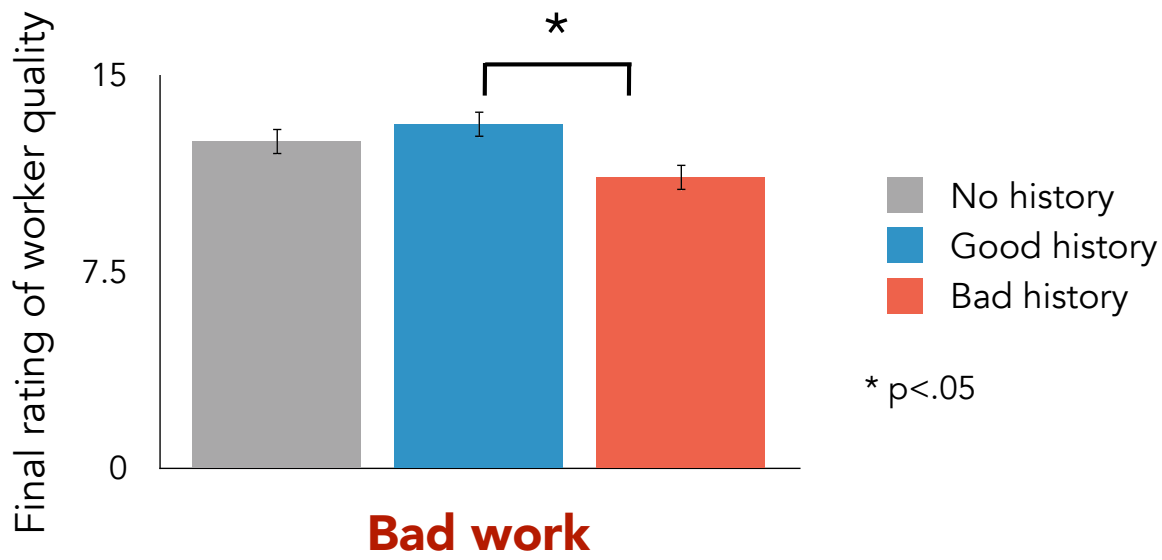
Figure 21.     Effects of history quality on final impression of worker (bad work)

As Figure 21 indicates, when the work quality was bad, the quality of the history mattered.  There was a significant main effect of history quality on final impressions of workers ($F(2,141)=3.77$, $p=.02$).  As post-hoc tests revealed, seeing a bad history initially made editors feel less positively about the worker than when they saw a good history beforehand.  Thus, a bad history followed by bad work was the most detrimental to final impressions.

This same pattern also manifested itself in terms of the effect of history quality on editing behavior of bad work.   As seen in Figure 22, there was a significant effect of history quality on editing behavior ($F(2,141)=4.74$, $p=.01$).  Specifically, editors who saw evidence of bad history edited significantly more of the bad work than either editors who saw no history information or those who saw good history information.
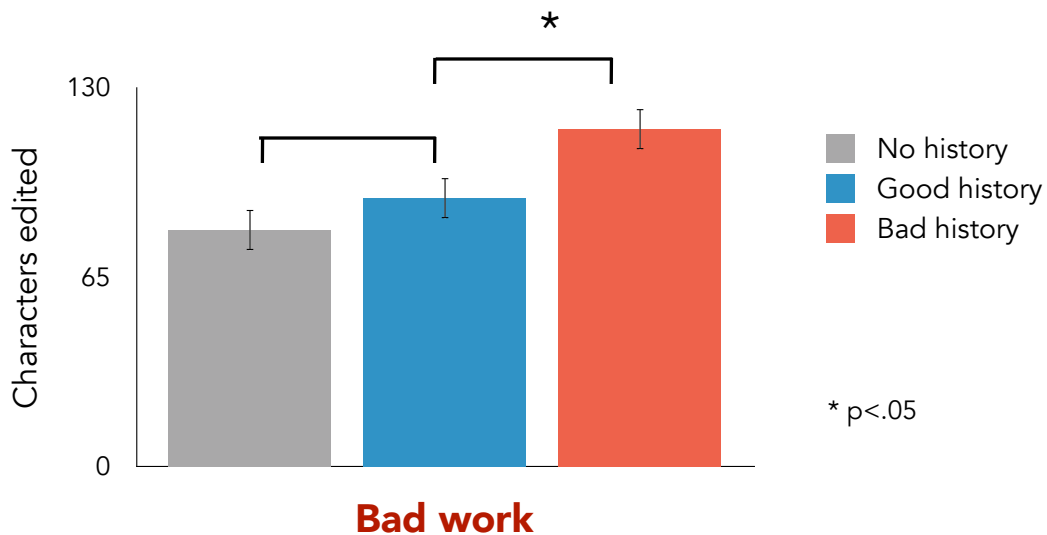
Figure 22.    Effects of history quality on editing of work (bad work)

Therefore, the influence of work history can be summarized as follows:  When the work that editors were viewing and rating themselves was good, history information did not have any effect on work evaluation or editing.  However, when the work was bad, editors who had seen bad history information rated the worker lower than those who saw good history, but edited the work more (thus improving its quality more).

I was also interested in seeing how (if at all) the visualization format of the history influenced later impressions and editing behaviors.  It could be that some elements of the visualization (such as level of detail) would lead to different downstream implications based on whether the work revealed or masked the details of good or bad performance.  Using a similar analysis approach to the role of history, I ran separate ANOVA tests on good and bad work, this time to see if the visualization type (detail vs. summary) influenced impressions and editing behaviors down the line.

Once again, when the work quality was good, the visualization type of the history people had seen previously did not matter.  There was no significant effect of visualization on final impressions or editing.

However, when the work was bad, visualization type influenced impressions and editing behaviors.  Although there was no significant main effect of visualization type

116

on final impression (F(2,141)=1.88, p=.15), there was a marginally significant difference between the detailed and abstract visualizations (F(1,141)=3.54, p=.06). Specifically, the abstract format lowered impressions compared to the detailed format.
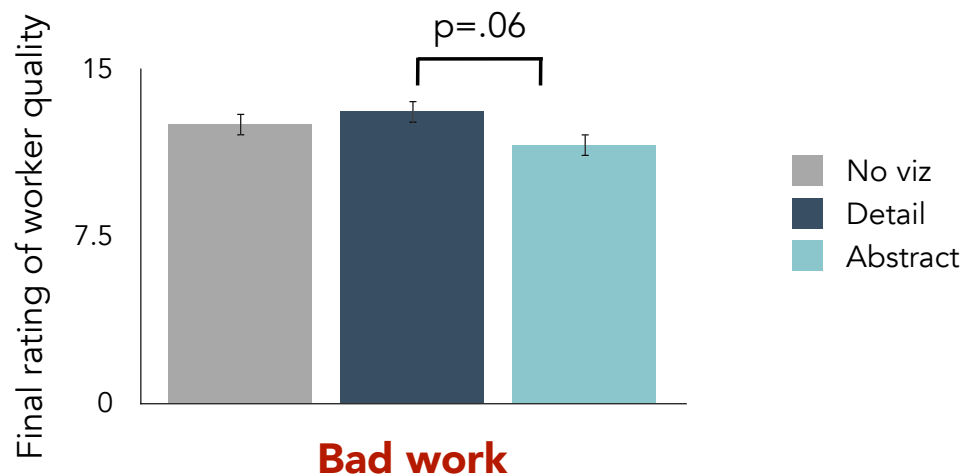


Figure 23.    Effect of history visualization type on final impressions (bad work)

Finally, history visualization affected editing behaviors for bad work only (F(2,141)=3.07, p=.04). Post-hoc tests showed this difference occurred between the abstract history visualization and the control (no history) condition (see Figure 24).
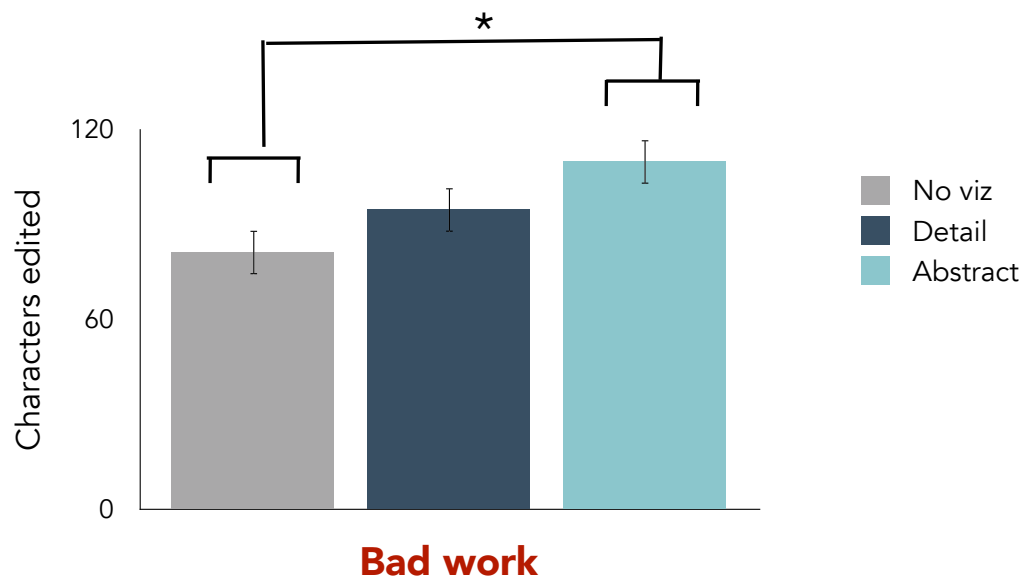
Figure 24.     Effect of history visualization type on editing behaviors (bad work)

To summarize the findings of history quality and visualization, I found that history quality and/or visualization did not influence impressions or editing behavior when the work that editors saw was good. However, when the work was bad, a bad history or an abstract visualization resulted in more negative final impressions of the worker. On the other hand, the bad history and abstract visualization also increased effort to edit the work when it was bad. In the next sections I consider theoretically justified explanations for the differences observed.

## 8.6  Discussion

### 8.6.1  Implications for theory

Figure 25 shows the hypothesized and observed results with regards to the role of history quality and detail on first impressions.
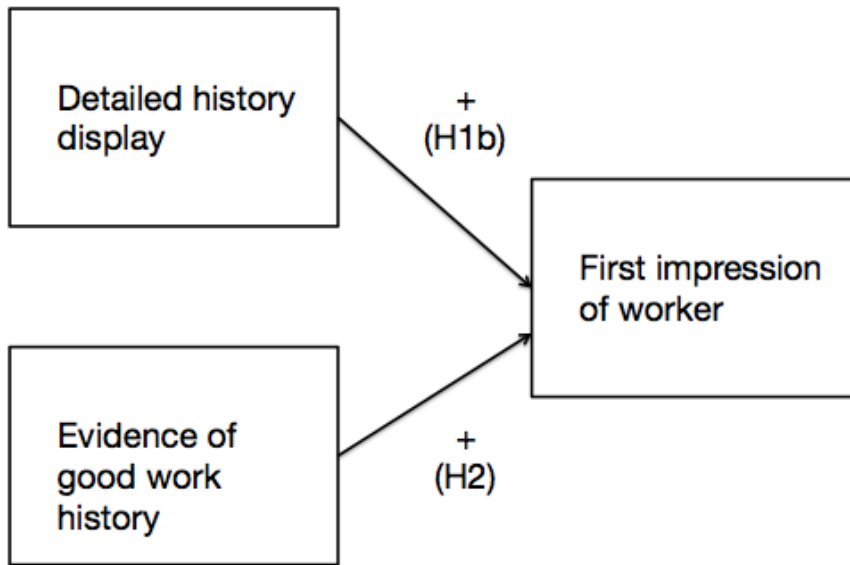
Figure 25.      Relationships between work history display and quality on first
        impressions

In conclusion, positive first impressions based on a "good" work history engendered a
positivity effect among participants in that condition.  This made them more positively
rate the quality of the worker, but at the same time made them neglect to fix more of the
work, especially when it was needed.  In contrast, a negative first impression seemed to
make participants more sensitive to fixing bad work (but did not make them unduly
critical towards good work.)
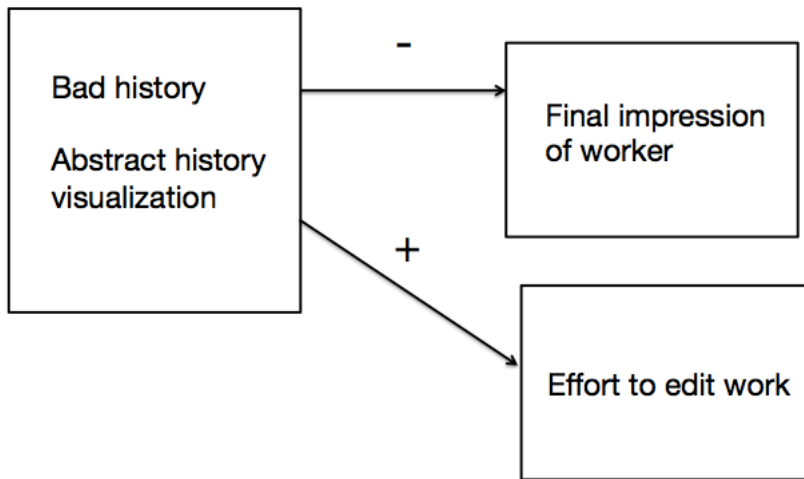
Figure 26.     Effect of history quality and visualization on final outcomes when
          the work quality viewed was **bad**

Here I examine possible explanations for the effects of various designs on attitudes and behaviors.

First, the more detailed activity history induced more positive initial first impressions than the abstracted visualizations. What may have caused the increased positivity of impressions after viewing the detailed display? As mentioned previously, participants expended more effort viewing the detailed display. This additional effort could have resulted in more positive impressions through greater perceived mental effort. Psychological research on the effort heuristic and the sunk-cost fallacy imply that perceived effort enhances evaluation (Tsai & McGill, 2011). Thus, the process of expending more effort to view the history visualization could have made observers feel more positively about the worker. Also, as originally predicted, it could be that the increased detail either reduced uncertainty about the worker and increased positivity, or that it obscured and obfuscated evidence of bad work along with good work.

I had expected that exposure to a negative activity history would cause editors in the experiment to carry this negative impression forward to the end, even if they were exposed to disconfirming evidence. Much previous research has shown the tendency for negative information to carry greater weight than positive information.

A negativity bias did not occur in the results and in fact there was the opposite effect, where people seemed to form positive first impressions that were resistant to change. While this enduring effect of positive information was initially surprising based on the rather negative tone that project owners took towards unknown contributors in Study 2, a possible explanation may lie in the nature of the experimental task that we used and may highlight different dynamics between impression formation in lightweight and heavyweight peer production.

For one, the task did not involve direct interaction between editor and worker. Lewicka, Czapinski, & Peeters (1992) highlight that the positivity bias (as opposed to the negativity bias) is most clearly observed in relatively neutral situations when evaluated targets are relatively novel or do not directly influence the individual during the process of evaluation. One reason why negative information gets more weight in interpersonal evaluative situations is because the cost of incorrectly assuming a bad person is good is higher than incorrectly assuming the reverse.

Another explanation for these findings may be that the editors/evaluators in the experiment were focused on forming impressions exclusively of targets' work output and not their attitude or personal qualities. As mentioned previously, most impressions of people carry a warmth (attitude) and a competence dimension. As Ybarra (2001) points out, in the warmth or morality domain, negativity effects are the norm just as described.

However, when assessing other people's competence, a positivity effect tends to occur. The dynamics behind this may tie back into the notion of signaling: Warmth is generally viewed as easy to fake (hence people are on the lookout for evidence of bad attitudes) while success or competent behaviors are not viewed as easy to fake. In other words, "once enthroned on the pedestal of competence, individuals tend to catch a break for periodic failures, but one instance of cold behavior may be enough to change warmth impressions for the worse" ( Cuddy, Glick, & Beninger, 2011).

Although the attitudinal effects of the positivity bias suggested by the results from this experiment confirm what has been observed in other settings, these findings

nonetheless provide additional insight into the consequences of attitudes and interpersonal competence evaluations on behaviors.

This result is somewhat surprising, since prior research on impression formation has suggested that negative information is typically more enduring than positive information (Peeters & Czapinski, 1990). If this was the case, we would not have seen a significant difference in editing behaviors between the two "bad history" conditions. However, the results from Study 3 are in line with related recent work by Carr & Walther, (2014), which also saw lack of support for the negativity effect in the context of social media information in a hiring context.

While most evidence of the negativity bias comes from direct interpersonal evaluations online or in the real world, it could be that negative information generated by others (in our case, by a previous editor's assessment) does not carry the same weight in informing impressions. For example, when users know little about the expertise or qualification of a previous rater, they may be less likely to trust past assessments (Metzger et al., 2010).

There may be various explanations for the lack of negativity effect that we observed that are related to the online peer review task and setting. One potential mechanism driving the persistence of negative information is that it is more novel and memorable and also as seen as more non-normative (Kanouse, 1984). The negativity bias may be reversed in situations where there is an expectation of negative behavior as the norm. This may have been the case in our task, where participants were primed to look out for work to correct. Additionally, the positivity bias is most clearly observed in neutral situations where the target of evaluation is relatively novel and does not directly influence the observer during the process of evaluation (Lewicka et al., 1992).

### 8.6.2 Implications for design

Differences in behaviors and attitudes between different visualizations of activity history suggest that design choices about how much information to provide and how difficult it may be to fully process have the potential to influence work outcomes.

One of Study 3's primary findings was that summary visualizations lowered impressions of the worker when the work was bad compared to the detailed visualization and that this effect was compounded by bad history followed by bad work. I also found that abstraction increased editing when compared to nothing, again when the work was bad. If we return to the ways outlined by (Hornbæk & Hertzum, 2011) in which overviews may affect information processing, we may find some insights in this regard.

With regards to the effects of summary (or overview) visualizations, one possibility is that they may reduce information processing effort. The other possibility is that they may encourage people to actively involve themselves in the information display to make sense of the visualization. My results suggest visual summaries reduced processing effort – participants who viewed a detailed history spent on average 33% more time (~13 more seconds) viewing the visualization than participants who viewed an abstracted bad history (this was a marginally significant difference). It could therefore be that by spending less time visualizing the bad history, participants did not take the time to make sense of and weigh all examples of work. For example, easy to process graphs may actually lead to more superficial interpretations (Hullman, Adar, & Shah, 2011a). As the flip side of this, it could be that spending more time on the detailed bad history allowed participants to give the worker the benefit of the doubt by seeing that their work did contain some correct transcriptions along with the bad (which may not have been uncovered by a casual glance).

Another hypothesized effect of overview visualizations, according to (Hornbæk & Hertzum, 2011) is that by using grouping, color coding or visual pop-out effects, they direct viewers' attention rapidly and without cognitive effort. One reason why the bad summary conditions may have led to differential editing behaviors based on the work could be that their designs featured color coding (using a green for good and red for bad) paradigm that made it easier to visually assess the overall work quality in a way that was not present in the table. Red in particular is associated with dangers and mistakes, and has the potential to make people more vigilant and risk averse (De Bock, Pandelaere, & Van Kenhove, 2013) so highlighting work with the color red may have

made editors more sensitive to seeking out and correcting mistakes. Adding color to the table in a similar way could potentially encourage more editing of the work.

While these effects are hypothesized, we did not directly measure participants' reactions to the visualizations in order to understand what features were most salient to them. Future work may involve observing how people interpret different visualizations using a think-aloud method to understand what they are focusing on and what they conclude based on various interface and design cues.

Table 7 presents a summary of design implications for the creation and visualization of activity trace information in response to the questions inspired by Study 2. To review, these dealt with (1) the impact of showing information about a worker's history compared to nothing; (2) the role of a summary or abstracted visualization, and (3) the role of a detailed visualization.

| Design parameter | Findings | Recommendations |
|---|---|---|
| Display of activity information versus nothing | Evidence of bad work influenced behaviors and attitudes compared to no information. | Consider whether it is detrimental to induce bad impressions and attitudes towards the person being observed. By exposing individual's activity information, showing bad work has the potential to change attitudes and behaviors compared to showing nothing at all (but showing good work does not have an effect). |
| Visual abstraction | Use of color may have reduced processing effort and made people more sensitive to fixing bad work by making mistakes "pop out"<br>Abstraction lowered impressions for bad work compared to detail | Visual abstraction may be a double-edged sword. While it did not affect impressions or behaviors when it highlighted good work, the use and salience of the color red to highlight bad work may have induced more negative impressions but also made people more vigilant to fix work. |
| Granularity (text vs. no text) | Text provided detail, which was associated with increased viewing time but also positivity of impressions | If the goal of the system is to encourage deep processing (and perhaps a more positive outlook on the worker in cases where interpersonal relationships are important), using text to obscure the quality of the work (or make it harder to assess) may be useful |

Table 7.    Initial design findings and implications

There could be other ways of enhancing a detailed work history to encourage more editing behaviors as well. For example, the system could insert artificial and fictitious

bad history in cases where a worker might have little to no prior work history in order to induce more sensitivity and attention to fixing their work. However, one tradeoff with doing this is that improved editing performance could come at the expense of more negative attitudes towards the worker. This could pose a problem in more collaborative situations with greater interactions between workers (e.g. in Wikipedia or software development).

In the experimental task, which was decoupled and anonymized, participants had little expectation of interacting with the person whose work they were fixing. As a result, in this type of situation it may be fine to highlight problems with their work in order to ensure more accurate editing. On the other hand, in a context where future interaction is expected, interpersonal attitudes and impressions of quality may play a bigger role in terms of outcomes to promote. Future research can address alternative design decisions and further explore their attitudinal and behavioral effects in a variety of collaborative task settings.

## 8.7 Conclusion

Overall, Study 3 provides some indication that observers are able to form fairly accurate impressions of an individual's past work history, but that both detail and valence of the history display play a role in the degree to which these impressions persist in the face of disconfirming evidence.

This work provides initial insights for designers of peer production systems about how various design parameters might influence both interpersonal attitudes and work evaluation behaviors (sometimes in different ways).

Given the increase in transparency of individuals' actions in online work contexts, there is great potential for leveraging this information to jump-start work relationships or to help recommend people for various tasks. However, my results suggest that not only *what* is presented but *how* it is shown can produce differing effects.

# 9 Contributions

## 9.1 Theoretical contributions

My thesis advances our understanding of impression formation in the specific context of online peer production. It builds upon prior research by delineating instances in which observers use activity traces to reduce uncertainty about unknown people in two different evaluative contexts: Hiring and deciding whether to accept unsolicited contributions to a project. GitHub provided a new setting in which many activity traces were generated by the system rather than explicitly under the control of the individual.

I applied the principles of signaling theory to the peer production work environment. My findings enhance understanding of this theory in a new online environment by identifying the factors and conditions that make perceptions of signals more or less salient in the peer production context. In this context, particularly in GitHub, many of the signals visible on individual profiles were not "intentional", that is, they were generated in the process of working and as Study 1 indicated, individuals were not always able or trying to manipulate or shape the signals that they gave off. As a result, all of the signals should have been equally reliable in terms of their "honesty". However, observers still chose certain cues over others in their impression formation process, which suggests that signal fit and effort required to interpret the signal were what drove them in their information-seeking process rather than just signal honesty.

In the peer production setting investigated, observers potentially had access to a variety of social media features providing cues they could use to reduce uncertainty about unknown others. These cues varied in terms of their ease of assessment, visual salience, and signaling value. The findings from Study 1 and 2 apply signaling theory to understand what is attended to and why. They emphasize not only the reliability of signals but also the ease of evaluation. Given the nature of peer production or recruitment and hiring, where time and mental resources are limited, my findings suggest that ease of evaluation may sometimes trump signal reliability, and that heuristics may be used either to help expedite the evaluation process or as a way to infer or extrapolate information (such as passion or interest) from a sea of details.

This work also begins to suggest linkages between impressions and behaviors. Much of the past work on impression formation has focused on rating individuals' expertise or personality in theoretical or lab-based scenarios. Through retrospective interviews based on actual encounters and collaborative decisions, this work provides the first information on when and how social networking information and activity traces influence contribution acceptance, rejection, and editing. My work also builds upon previous studies of self-reported impressions by including a behavioral component to see the trickle-down effects of impressions on action (particularly expending effort to edit another person's work).

The interplay between impressions of work and attitude suggests an interesting set of tensions that face peer production workers, particularly those in a leadership or editorial role. We see cases in which optimizing one outcome (e.g. personal relations) may come at the expense of positive work-related outcomes (such as working quickly and efficiently or improving the work product). The results from Study 3 provide further support from a new context to the notion that positivity biases seem to prevail when the criteria of evaluation deal with a person's skill or competence (as opposed to their attitude or character).

## 9.2 Contributions to HCI and CSCW

This thesis work also provides design contributions to the fields of HCI and CSCW. I examined a relatively new and emergent setting in the world of online collaborative work (an open, volunteer-based work environment where social media features are linked to and make it easy to see the work someone has done). The findings from Studies 1, 2 and 3 suggest ways in which this type of collaborative setting can improve impression formation through providing increased access to reliable signals about strangers. At the same time, lack of time and motivation to consider and evaluate the information can hinder impression formation in peer production (through the formation of biased impressions).

Broadly speaking, this thesis furthers our knowledge of how people use publicly visible activity traces to learn and draw conclusions about strangers contributing work to a collective endeavor. It also considers how seemingly small or simple design choices

(such as the amount of detail shown) may have differing psychological consequences for observers that influence perceptions and behaviors of others. These findings are useful not only for the context of online software development but may also be applied to peer review and online work settings in other fields such as online translation, citizen science or volunteer efforts, Wikipedia, and more.

Studies 1 and 2 provide insights into what profile and interface cues observers attend to in a variety of scenarios and why. A similar signal such as activity level was used in both settings but given different meaning/interpretation (for hiring, it was seen as a sign of passion for coding, while in a collaborative situation it was seen as an indicator of experience and expertise.) As a result, the same visible behavior may still be interpreted differently based on the observer's motivations. It may be hard for a system to automatically understand and infer an observer's motivations (and the observer themselves may not actually be aware of what these are either).

These studies also give insight into the impact of making work and interaction history available to different types of observers, both inside and outside of the peer production environment. The findings that both employers and project owners assess coders' expertise and interpersonal attitudes imply that making this type of information accessible and interpretable is helpful in these contexts.

At the same time, differences in signals used across Study 1 and Study 2 suggest that not all types of information are relevant and sought after by all types of observers. Although in both hiring and collaborative contexts individuals sought out information that was perceived as reliable, observable, and easy to verify, the specific cues varied depending on who the observer was. The findings from Study 1 also suggest broader concerns about information being used out of the context in which it was generated.

As a result of different audiences with different needs viewing information about a user's activity, designers of systems that display publicly visible activity traces may wish to consider this and dynamically display or present different features of information based on the role of who is observing an individual's profile. The notion of "context collapse" that has been noted in other social media settings (e.g. Marwick, 2011) is also relevant here. The system could either create an "observer only view" if

the observer was a person who was not a member of the site or was not directly connected to the user, or allow users to choose what they saw about a person through customizing their site settings.

Study 3 also provides initial insights into design principles for sites that generate and display activity trace information. The contributions to design include a greater understanding of some of the psychological implications of providing information about people's past behavior in different formats. Specifically, I explored the effects of granularity and abstraction in an interpersonal work evaluation context.

This work provides design contributions to aid designers of future peer production and open collaboration systems, both in open source software development, crowdsourcing, and beyond. By creating and testing activity history presentations that were informed by psychological theory and cognitive principles, the experiments show that seemingly small or simple design choices can spill over into differing perceptions of and actions towards a similar target. We see how granularity of detail and visual abstraction lead to different outcomes such as positivity towards the worker and effort to edit their work. An open question for future work is the degree to which granularity of detail and abstraction of visualizing individuals' interaction histories will show similar patterns in evaluation of the worker.

These findings can be leveraged to optimize various desired outcomes, be these positive feelings towards individuals or ease of processing. Perhaps most interesting is the tradeoff, particular to the peer production setting, between interpersonal relations and sensitivity to editing work. For example, it could be that designers of a peer review system visualizing activity history could emphasize negative aspects of an individual's' work in order to encourage sensitivity towards making fixes. For individuals who are new and lack any prior history, minimizing their new-ness through de-emphasizing the total number of contributions through a visual summary could reduce the negative reaction towards newcomers.

An interesting related consideration is the interplay between what a system showcases about individuals' activities and how this influences individual behavior. For example, making changes in the visual interface of a person's profile to make the percent of

contributions accepted (something that GitHub did not incorporate during my research interviews) more salient or obvious might encourage workers to be more cautious or aware of the quality of the work they are doing. While I did not specifically explore these elements in this work, they provide fruitful directions for future research.

# 10 Conclusion

In conclusion, as the feasibility and utility of doing peer production work online increase ease of contribution by of volunteer contributors, uncertainty about the quality and attitudes of unknown others can (but does not have to) remain high. Designers of sites and systems that collect and display activity trace information about workers make choices about how to visually display this information, and as this thesis suggests, these design choices have both psychological and behavioral implications.

People in a peer production context face conflicting choices about who to encourage and how much effort to expend on editing work, often with limited amounts of time, and as a result often choose signals to help inform their decisions that are reliable or easy to assess. Given that people do not always use details given to them to the best of their abilities, there is great promise to the an idea of a system that can dynamically present appropriate and relevant information to the task at hand in a way that balances reliability and ease of interpretation.

Therefore, understanding the interplay between what information about others' activity histories is displayed, how it is displayed, how it is perceived and interpreted, and how it is used to inform behaviors and actions is important. This thesis provides initial insights into these dynamics and suggests promising future directions to explore in order to help peer production participants work together more effectively.

# 11 References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2–3), 131–152. doi:10.1016/S0360-1315(99)00029-9

Algan, Y., Benkler, Y., Morell, M. F., & Hergueux, J. (2013). Cooperation in a Peer Production Economy Experimental Evidence from Wikipedia. In *Workshop on Information Systems and Economics, Milan, Italy* (pp. 1–31). Retrieved from http://redirect.subscribe.ru/_/-/www.eief.it/files/2013/08/yann-algan.pdf

Ambady, N. (2010). The perils of pondering: Intuition and thin slice judgments. *Psychological Inquiry*, *21*(4), 271–278.

Anderson, C., & Shirako, A. (2008). Are individuals' reputations related to their history of behavior? *Journal of Personality and Social Psychology*, *94*(2), 320.

Antheunis, M. L., Valkenburg, P. M., & Peter, J. (2010). Getting acquainted through social network sites: Testing a model of online uncertainty reduction and social attraction. *Computers in Human Behavior*, *26*(1), 100–109.

Antheunis, & Schouten, A. P. (2011). The Effects of Other-Generated and System-Generated Cues on Adolescents' Perceived Attractiveness on Social Network Sites. *Journal of Computer-Mediated Communication*, *16*(3), 391–406.

Baruh, L., Chisik, Y., Bisson, C., & Senova, B. (2012). More or Less: Amount of Personal Information Displayed in Social Network Site Profiles and Its Impact on Viewers' Intentions to Socialize with the Profile Owner. In *ICWSM*. Retrieved from

http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/viewFile/460
1/5015

Begel, A., DeLine, R., & Zimmermann, T. (2010). Social media for software engineering.
In *Proceedings of the FSE/SDP workshop on Future of software engineering research*
(pp. 33–38).

Begel, A., Khoo, Y. P., & Zimmermann, T. (2010). Codebook: discovering and exploiting
relationships in software repositories. In *Software Engineering, 2010 ACM/IEEE
32nd International Conference on* (Vol. 1, pp. 125–134).

Benkler, & Nissenbaum, H. (2006). Commons-based Peer Production and Virtue*.
*Journal of Political Philosophy*, *14*(4), 394–419.

Benkler, Y. (2002). Coase's Penguin, or, Linux and" The Nature of the Firm." *Yale Law
Journal*, 369–446.

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and
beyond: Toward a developmental theory of interpersonal communication.
*Human Communication Research*, *1*(2), 99–112.

Bohnert, D., & Ross, W. H. (2010). The influence of social networking web sites on the
evaluation of job candidates. *Cyberpsychology, Behavior, and Social Networking*,
*13*(3), 341–347.

boyd, danah, & Ellison, N. (2007). Social network sites: Definition, history, and
scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230.

Carr, C. T., & Walther, J. B. (2014). Increasing Attributional Certainty via Social Media: Learning About Others One Bit at a Time. *Journal of Computer-Mediated Communication*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/jcc4.12072/full

Charmaz, K. (2008). Grounded theory as an emergent method. *Handbook of Emergent Methods*, 155–170.

Chen, S., Shechter, D., & Chaiken, S. (1996). Getting at the truth or getting along: Accuracy-versus impression-motivated heuristic and systematic processing. *Journal of Personality and Social Psychology*, *71*(2), 262.

Chevalier, F., Huot, S., & Fekete, J.-D. (2010). WikipediaViz: Conveying article quality for casual Wikipedia readers. In *Pacific Visualization Symposium (PacificVis), 2010 IEEE* (pp. 49–56). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5429611

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39–67.

Crowston, K., Howison, J., & Annabi, H. (2006). Information systems success in free and open source software development: theory and measures. *Software Process: Improvement and Practice*, *11*(2), 123–148. doi:10.1002/spip.259

Cuddy, A. J., Glick, P., & Beninger, A. (2011). The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in Organizational Behavior*, *31*, 73–98.

Culnan, M. J., & Markus, M. L. (1987). Information technologies. Retrieved from

  http://psycnet.apa.org/psycinfo/1987-98580-013

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social coding in GitHub:

  transparency and collaboration in an open software repository. In *Proceedings of*

  *the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 1277–1286).

  Retrieved from http://dl.acm.org/citation.cfm?id=2145204.2145396

DeAndrea, D. C. (2014). Advancing Warranting Theory. *Communication Theory*, 24(2),

  186–204.

De Bock, T., Pandelaere, M., & Van Kenhove, P. (2013). When colors backfire: The

  impact of color cues on moral judgment. *Journal of Consumer Psychology*, 23(3),

  341–348.

Derthick, K., Tsao, P., Kriplean, T., Borning, A., Zachry, M., & McDonald, D. W. (2011).

  Collaborative Sensemaking during Admin Permission Granting in Wikipedia. In

  A. A. Ozok & P. Zaphiris (Eds.), *Online Communities and Social Computing* (Vol.

  6778, pp. 100–109). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved

  from http://www.springerlink.com/content/f75x002m68651qq0/

DiMicco, J. M., & Millen, D. R. (2007). Identity management: multiple presentations of

  self in facebook (pp. 383–386).

Donath, & boyd, D. (2004). Public displays of connection. *Bt Technology Journal*, 22(4),

  71–82.

Donath, J. (2007). Signals in social supernets. *Journal of Computer-Mediated Communication*, *13*(1), 231–251.

Donath, J. S. (1999). Identity and deception in the virtual community. *Communities in Cyberspace*, *1996*, 29–59.

Ellison, N. B., Hancock, J. T., & Toma, C. L. (2012). Profile as promise: A framework for conceptualizing veracity in online dating self-presentations. *New Media & Society*, *14*(1), 45–62.

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. *Journal of Computer-Mediated Communication*, *11*(2), 415–441.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*(6), 889.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83.

Funder, D. C. (1995). On the accuracy of personality judgment: a realistic approach. *Psychological Review*, *102*(4), 652.

Gazan, R. (2010). Microcollaborations in a social Q&A community. *Information Processing & Management*, *46*(6), 693–702.

Gibbs, J. L., Ellison, N. B., & Lai, C. H. (2011). First comes love, then comes Google: An investigation of uncertainty reduction strategies and self-disclosure in online dating. *Communication Research*, *38*(1), 70–100.

Goffman, E. (1959). The presentation of self in everyday life. Retrieved from

      http://books.google.com/books?hl=en&lr=&id=TlIAzT5uT-

      IC&oi=fnd&pg=PA120&dq=presentation+of+self&ots=IsDgajMrma&sig=euS9o

      w57JdGzOGsHFmd8J1LSb5U

Gosling, S. D., Gaddis, S., Vazire, S., & others. (2007). Personality impressions based on

      Facebook profiles. In *Proceedings of the International Conference on Weblogs and*

      *Social Media* (pp. 26–28).

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue:

      personality judgments based on offices and bedrooms. *Journal of Personality and*

      *Social Psychology*, *82*(3), 379.

Graham, L. T., & Gosling, S. D. (2012). Impressions of world of warcraft players'

      personalities based on their usernames: Interobserver consensus but no accuracy.

      *Journal of Research in Personality*. Retrieved from

      http://www.sciencedirect.com/science/article/pii/S0092656612000797

Grasz, J. (2009). Forty-five percent of employers use social networking sites to research

      job candidates, CareerBuilder survey finds. CareerBuilder Press Releases.

      Retrieved 21 February 2010 from http://www.careerbuilder.com/share/

      aboutus/pressreleasesdetail.aspx?id=pr519&sd=8%2f19%2f2009&ed=

      12%2f31%2f2009&siteid=cbpr&sc_cmp1=cb_pr519_.

Halfaker, A., Geiger, R. S., & Terveen, L. G. (2014). Snuggle: designing for efficient

      socialization and ideological critique. In *Proceedings of the 32nd annual ACM*

*conference on Human factors in computing systems* (pp. 311–320). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2557313

Hamilton, D. L., Katz, L. B., & Leirer, V. O. (1980). Cognitive representation of personality impressions: Organizational processes in first impression formation. *Journal of Personality and Social Psychology*, *39*(6), 1050–1063.

Hancock, J. T., & Dunham, P. J. (2001). Impression Formation in Computer-Mediated Communication Revisited An Analysis of the Breadth and Intensity of Impressions. *Communication Research*, *28*(3), 325–347.

Hansen, D. L., Schone, P. J., Corey, D., Reid, M., & Gehring, J. (2013). Quality control mechanisms for crowdsourcing: peer review, arbitration, & expertise at familysearch indexing. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 649–660). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2441848

Haythornthwaite, C. (2009). Crowds and communities: Light and heavyweight models of peer production. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on* (pp. 1–10). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4755627

Hornbæk, K., & Hertzum, M. (2011). The notion of overview in information visualization. *International Journal of Human-Computer Studies*, *69*(7), 509–525.

Hullman, J., Adar, E., & Shah, P. (2011a). Benefitting infovis with visual difficulties. *Visualization and Computer Graphics, IEEE Transactions on*, *17*(12), 2213–2222.

Hullman, J., Adar, E., & Shah, P. (2011b). The impact of social information on visual judgments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1461–1470). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1979157

Jacobson, D. (1999). Impression Formation in Cyberspace: Online Expectations and Offline Experiences in Text-based Virtual Communities. *Journal of Computer-Mediated Communication*, *5*(1), 0–0. doi:10.1111/j.1083-6101.1999.tb00333.x

Jacovi, M., Guy, I., Kremer-Davidson, S., Porat, S., & Aizenbud-Reshef, N. (2014). The Perception of Others: Inferring Reputation from Social Media in the Enterprise. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work &#38; Social Computing* (pp. 756–766). New York, NY, USA: ACM. doi:10.1145/2531602.2531667

Johri, A. (2012). From a distance: Impression formation and impression accuracy among geographically distributed coworkers. *Computers in Human Behavior*. Retrieved from http://www.sciencedirect.com/science/article/pii/S0747563212001999

Kanouse, D. E. (1984). EXPLAINING NEGATIVITY BIASES IN EVALUATION AND CHOICE BEHAVIOR: THEORY AND RESEARCH. *Advances in Consumer Research*, *11*(1). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site &authtype=crawler&jrnl=00989258&AN=6434234&h=16sNQInlMN9NXc8ovz53I gmKhLHwP2P1FXl9atr8J658psucfcMGIBrQ3lRykP3lMY1xeN%2BtdvoPljilehaR7 g%3D%3D&crl=c

Kim, D., & Shipman, F. M., III. (2012). Visualizing History to Improve Users' Location and Comprehension of Collaborative Work. In *Proceedings of the 17th ACM International Conference on Supporting Group Work* (pp. 11–20). New York, NY, USA: ACM. doi:10.1145/2389176.2389179

Kim, S., Oh, J. S., & Oh, S. (2007). Best-answer selection criteria in a social Q&A site from the user-oriented relevance perspective. *Proceedings of the American Society for Information Science and Technology*, *44*(1), 1–15.

Kristof-Brown, A., Barrick, M. R., & Franke, M. (2002). Applicant impression management: Dispositional influences and consequences for recruiter perceptions of fit and similarity. *Journal of Management*, *28*(1), 27–46.

Kruglanski, A. W. (1989). The psychology of being" right": The problem of accuracy in social perception and cognition. *Psychological Bulletin*, *106*(3), 395.

Lampe, C. A. C., Ellison, N., & Steinfield, C. (2007). A familiar face (book): profile elements as signals in an online social network. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 435–444).

Lea, M., & Spears, R. (1992). Paralanguage and social perception in computer-mediated communication. *Journal of Organizational Computing and Electronic Commerce*, 2(3-4), 321–341.

Lerner, J., & Tirole, J. (2002). Some simple economics of open source. *The Journal of Industrial Economics*, *50*(2), 197–234.

Leskovec, J., Huttenlocher, D. P., & Kleinberg, J. M. (2010). Governance in social media: A case study of the wikipedia promotion process. In *ICWSM*. Retrieved from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1485/1841/

Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*, *91*(1), 111.

Levine, S. S., & Prietula, M. J. (2013). Open Collaboration for Innovation: Principles and Performance. *Organization Science*. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/orsc.2013.0872

Lewicka, M., Czapinski, J., & Peeters, G. (1992). Positive-negative asymmetry or "When the heart needs a reason." *European Journal of Social Psychology*, 22(5), 425–434.

Liao, Q. V., & Fu, W.-T. (2014a). Can you hear me now?: mitigating the echo chamber effect by source position indicators. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 184–196). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2531711

Liao, Q. V., & Fu, W.-T. (2014b). Expert voices in echo chambers: effects of source expertise indicators on exposure to diverse opinions. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 2745–2754). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2557240

Liao, Q. V., Wagner, C., Pirolli, P., & Fu, W. T. (2012). Understanding experts' and novices' expertise judgment of twitter users. In *Proceedings of the 2012 ACM*

*annual conference on Human Factors in Computing Systems* (pp. 2461–2464).

Retrieved from http://dl.acm.org/citation.cfm?id=2208410

Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology*, *25*(2), 115–126.

Lockwood, N. S., & Massey, A. P. (2012). Communicator trust across media: A comparison of audio conferencing, video conferencing, and a 3D virtual environment. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (pp. 839–848). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6148996

Long, J. (2009). Open Source Software Development Experiences on the Students' Resumes: Do They Count?-Insights from the Employers' Perspectives. *Journal of Information Technology Education*, *8*, 229–242.

Lotufo, R., Passos, L., & Czarnecki, K. (n.d.). Towards Improving Bug Tracking Systems with Game Mechanisms.

Luther, K., Caine, K., Ziegler, K., & Bruckman, A. (2010). Why it works (when it works): success factors in online creative collaboration. In *Proceedings of the 16th ACM international conference on Supporting group work* (pp. 1–10). New York, NY, USA: ACM. doi:10.1145/1880071.1880073

Luther, K., Diakopoulos, N., & Bruckman, A. (2010). Edits & credits: exploring integration and attribution in online creative collaboration. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing*

*systems* (pp. 2823–2832). Retrieved from

http://dl.acm.org/citation.cfm?id=1753869

Marlow, J., & Dabbish, L. (2013). Activity traces and signals in software developer
recruitment and hiring. In *Proceedings of the 2013 conference on Computer supported
cooperative work* (pp. 145–156). Retrieved from
http://dl.acm.org/citation.cfm?id=2441794

Marlow, J., Dabbish, L., & Herbsleb, J. (2013). Impression formation in online peer
production: activity traces and personal profiles in github. In *Proceedings of the
2013 conference on Computer supported cooperative work* (pp. 117–128). Retrieved
from http://dl.acm.org/citation.cfm?id=2441792

Marwick, A. E. (2011). I tweet honestly, I tweet passionately: Twitter users, context
collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133.

Mavlanova, T., Benbunan-Fich, R., & Koufaris, M. (2012). Signaling theory and
information asymmetry in online commerce. *Information & Management*, 240–247.

McDonald, D. W., Gokhman, S., & Zachry, M. (2012). Building for social translucence: a
domain analysis and prototype system. In *Proceedings of the ACM 2012 conference
on Computer Supported Cooperative Work* (pp. 637–646). ACM. Retrieved from
http://dl.acm.org/citation.cfm?id=2145301

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches
to credibility evaluation online. *Journal of Communication*, *60*(3), 413–439.

Monroy-Hernández, A., Hill, B. M., Gonzalez-Rivero, J., & others. (2011). Computers can't give credit: how automatic attribution falls short in an online remixing community. In *Proceedings of the 2011 annual conference on Human factors in computing systems* (pp. 3421–3430). Retrieved from http://dl.acm.org/citation.cfm?id=1979452

Morris, M. R., Counts, S., Roseway, A., Hoff, A., & Schwarz, J. (2012). Tweeting is believing?: understanding microblog credibility perceptions. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 441–450). Retrieved from http://dl.acm.org/citation.cfm?id=2145274

Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, *53*(3), 431.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175.

Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. *Applied Cognitive Psychology*, *20*(2), 139–156.

Oreg, S., & Nov, O. (2008). Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values. *Computers in Human Behavior*, *24*(5), 2055–2073.

Paul, S. A., Hong, L., & Chi, E. H. (2012). Who is Authoritative? Understanding Reputation Mechanisms in Quora. *Arxiv Preprint arXiv:1204.3724*. Retrieved from http://arxiv.org/abs/1204.3724

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, *1*(1), 33–60.

Quinn, K. A., Mason, M. F., & Macrae, C. N. (2009). Familiarity and person construal: Individuating knowledge moderates the automaticity of category activation. *European Journal of Social Psychology*, *39*(5), 852–861.

Raban, D. R., Danan, A., Ronen, I., & Guy, I. (2012). Impression formation in corporate people tagging. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (pp. 569–578). New York, NY, USA: ACM. doi:10.1145/2207676.2207756

Ramirez Jr, A., Walther, J. B., Burgoon, J. K., & Sunnafrank, M. (2002). Information-Seeking Strategies, Uncertainty, and Computer-Mediated Communication. *Human Communication Research*, *28*(2), 213–228.

Rao, V. (2012). GitHub and the democratization of programming. Retrieved from http://www.forbes.com/sites/venkateshrao/2012/03/27/github-and-the-democratization-of-programming/

Rigby, P. C., & Storey, M. A. (2011). Understanding broadcast based peer review on open source software projects. In *Proceeding of the 33rd international conference on Software engineering* (pp. 541–550).

Rivera, L. A. (2012). Hiring as cultural matching the case of elite professional service firms. *American Sociological Review*, *77*(6), 999–1022.

Rusman, E., Van Bruggen, J., Sloep, P., Valcke, M., & Koper, R. (2010). Can I Trust You?. Profile Elements that Inform First Impressions of Trustworthiness in Virtual Project Teams.

Shami, N. S., Ehrlich, K., Gay, G., & Hancock, J. T. (2009). Making sense of strangers' expertise from signals in digital artifacts. In *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 69–78).

Singer, L., Figueira Filho, F., Cleary, B., Treude, C., Storey, M.-A., & Schneider, K. (2013). Mutual assessment in the social programmer ecosystem: An empirical investigation of developer profile aggregators. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 103–116). Retrieved from http://dl.acm.org/citation.cfm?id=2441791

Sinha, V. S., Mani, S., & Sinha, S. (2011). Entering the circle of trust: developer initiation as committers in open-source projects. In *Proceeding of the 8th working conference on Mining software repositories* (pp. 133–142).

Skeels, M. M., & Grudin, J. (2009). When social networks cross boundaries: a case study of workplace use of facebook and linkedin. In *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 95–104).

Smith, E. R., & Collins, E. C. (2009). Contextualizing person perception: Distributed social cognition. *Psychological Review*, *116*(2), 343–364.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Stecher, K., & Counts, S. (2008). Thin Slices of Online Profile Attributes. *Proc. ICWSM*.

Stewart, D. (2005). Social status in an open-source community. *American Sociological Review*, *70*(5), 823–842.

Stewart, K. J., & Gosain, S. (2006). The Impact of Ideology on Effectiveness in Open Source Software Development Teams. *MIS Quarterly*, *30*(2), 291–314.

Stuart, H. C., Dabbish, L., Kiesler, S., Kinnaird, P., & Kang, R. (2012). Social transparency in networked information exchange: a theoretical framework. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 451–460). Retrieved from http://dl.acm.org/citation.cfm?id=2145275

Suh, B., Chi, E. H., Kittur, A., & Pendleton, B. A. (2008). Lifting the veil: improving accountability and social transparency in Wikipedia with wikidashboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1037–1040). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1357214

Takhteyev, Y., & Hilts, A. (2010). *Investigating the geography of open source software through GitHub*. Retrieved from http://takhteyev.org/papers/Takhteyev-Hilts-2010.pdf

Tanis, M., & Postmes, T. (2003). Social Cues and Impression Formation in CMC. *The Journal of Communication*, *53*(4), 676–693.

Tausch, N., Kenworthy, J. B., & Hewstone, M. (2007). The confirmability and disconfirmability of trait concepts revisited: Does content matter? *Journal of Personality and Social Psychology*, *92*(3), 542.

Tausczik, Y. R., & Pennebaker, J. W. (2011). Predicting the perceived quality of online mathematics contributions from users' reputations. In *Proceedings of the 2011 annual conference on Human factors in computing systems* (pp. 1885–1888). Retrieved from http://dl.acm.org/citation.cfm?id=1979215

Thompson, R. S., Rantanen, E. M., Yurcik, W., & Bailey, B. P. (2007). Command line or pretty lines?: comparing textual and visual interfaces for intrusion detection. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (p. 1205). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1240807

Tong, S. T., Van Der Heide, B., Langwell, L., & Walther, J. B. (2008). Too much of a good thing? The relationship between number of friends and interpersonal impressions on Facebook. *Journal of Computer-Mediated Communication*, *13*(3), 531–549.

Towne, W. B., Kittur, A., Kinnaird, P., & Herbsleb, J. (2013). Your process is showing: Controversy management and perceived quality in Wikipedia. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 1059–1068). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2441896

Trainer, E. H., Al-Ani, B., & Redmiles, D. F. (2011). Impact of collaborative traces on trustworthiness. In *Proceeding of the 4th international workshop on Cooperative and*

*human aspects of software engineering* (pp. 40–47). Retrieved from

http://dl.acm.org/citation.cfm?id=1984649

Trammell, K. D., & Keshelashvili, A. (2005). Examining the new influencers: A self-presentation study of A-list blogs. *Journalism and Mass Communication Quarterly*, *82*(4), 968.

Treude, C., & Storey, M. A. (2010). Awareness 2.0: staying aware of projects, developers and tasks using dashboards and feeds. In *Software Engineering, 2010 ACM/IEEE 32nd International Conference on* (Vol. 1, pp. 365–374). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6062104

Tsai, C. I., Klayman, J., & Hastie, R. (2008). Effects of amount of information on judgment accuracy and confidence. *Organizational Behavior and Human Decision Processes*, *107*(2), 97–105.

Tsai, C. I., & McGill, A. L. (2011). No pain, no gain? How fluency and construal level affect consumer confidence. *Journal of Consumer Research*, *37*(5), 807–821.

Van Der Heide, B., D'Angelo, J. D., & Schumaker, E. M. (2012). The Effects of Verbal Versus Photographic Self-Presentation on Impression Formation in Facebook. *Journal of Communication*, *62*(1), 98–116. doi:10.1111/j.1460-2466.2011.01617.x

Walther, J. B. (2011). Visual cues in computer-mediated communication: Sometimes less is more. *Face-to-Face Communication Over the Internet: Emotions in a Web of Culture, Language, and Technology*, 17.

Walther, J. B., & Tong, S. T. (2014). Inventing Partners in Computer-Mediated Communication: How CMC Sustains Self-fulfilling Prophecies and Relational Attributions. In *Social Computing and Social Media* (pp. 517–527). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-07632-4_49

Walther, J. B., Van Der Heide, B., Hamel, L. M., & Shulman, H. C. (2009). Self-Generated Versus Other-Generated Statements and Impressions in Computer-Mediated Communication A Test of Warranting Theory Using Facebook. *Communication Research*, *36*(2), 229–253.

Wattenberg, M., Viégas, F. B., & Hollenbach, K. (2007). Visualizing activity on wikipedia with chromograms. In *Human-Computer Interaction–INTERACT 2007* (pp. 272–287). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-74800-7_23

Weiss, B., & Feldman, R. S. (2006). Looking good and lying to do it: Deception as an impression management strategy in job interviews. *Journal of Applied Social Psychology*, *36*(4), 1070–1086.

Westerman, D., Van Der Heide, B., Klein, K. A., & Walther, J. B. (2008). How do people really seek information about others?: Information seeking across Internet and traditional communication channels. *Journal of Computer-Mediated Communication*, *13*(3), 751–767.

Would you put your stackoverflow profile link on your CV / Resume? (n.d.). Retrieved August 23, 2014, from

http://programmers.stackexchange.com/questions/12070/would-you-put-your-stackoverflow-profile-link-on-your-cv-resume

Yang, J., Counts, S., Morris, M. R., & Hoff, A. (2013). Microblog credibility perceptions: comparing the USA and China. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 575–586). Retrieved from http://dl.acm.org/citation.cfm?id=2441841

Yang, J., Tao, K., Bozzon, A., & Houben, G.-J. (2014). Sparrows and Owls: Characterisation of Expert Behaviour in StackOverflow. In *User Modeling, Adaptation, and Personalization* (pp. 266–277). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-08786-3_23

Yarosh, S., Matthews, T., & Zhou, M. (2012). Asking the right person: supporting expertise selection in the enterprise. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (pp. 2247–2256).

Ybarra, O. (2001). When first impressions don't last: The role of isolation and adaptation processes in the revision of evaluative impressions. *Social Cognition*, *19*(5), 491–520.

Zhu, H., Zhang, A., He, J., Kraut, R. E., & Kittur, A. (2013). Effects of peer feedback on contribution: a field experiment in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2253–2262). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2481311

# Appendix: Visualizations of user activity history in GitHub and related sites

Example visualizations of user activity history found on GitHub and similar existing sites that showcase software developers' open source activities. Common themes in terms of different types of visualizations emerged across the sites, which include (1) summary statistic view, (2) text-based lists of activity and (3) graphical depictions such as graphs of activity over time, icon-based representations, or badges.

## Site 1: GitHub
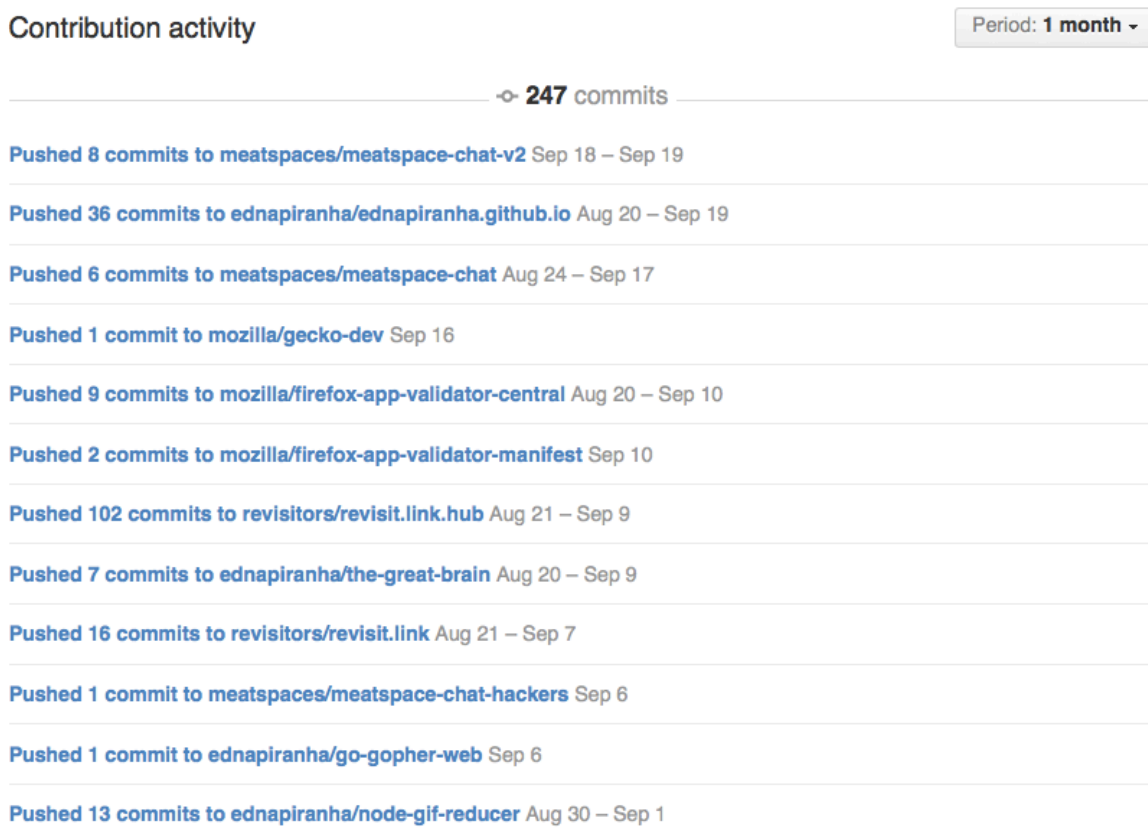


GitHub text-based list visualization of activity history (overview of total number of contributions in a given time period: Below we see the first part of 247 commits made over the past month). Time stamps are provided, but details of the commits are not visible.

**2 days ago**
ednapiranha **commented on pull request** meatspaces/meatspace-chat#241
lol

**3 days ago**
ednapiranha **closed issue** ednapiranha/ednapiranha.github.io#1
Weird overescaped HTML in Interviews + Articles section

**3 days ago**
ednapiranha **commented on issue** ednapiranha/ednapiranha.github.io#1
lol just noticed that now - fixed!

**3 days ago**
ednapiranha **pushed to** gh-pages **at** ednapiranha/ednapiranh...
**gh-pages** is now 1c03157

**3 days ago**
ednapiranha **pushed to** master **at** ednapiranha/ednapiranh...
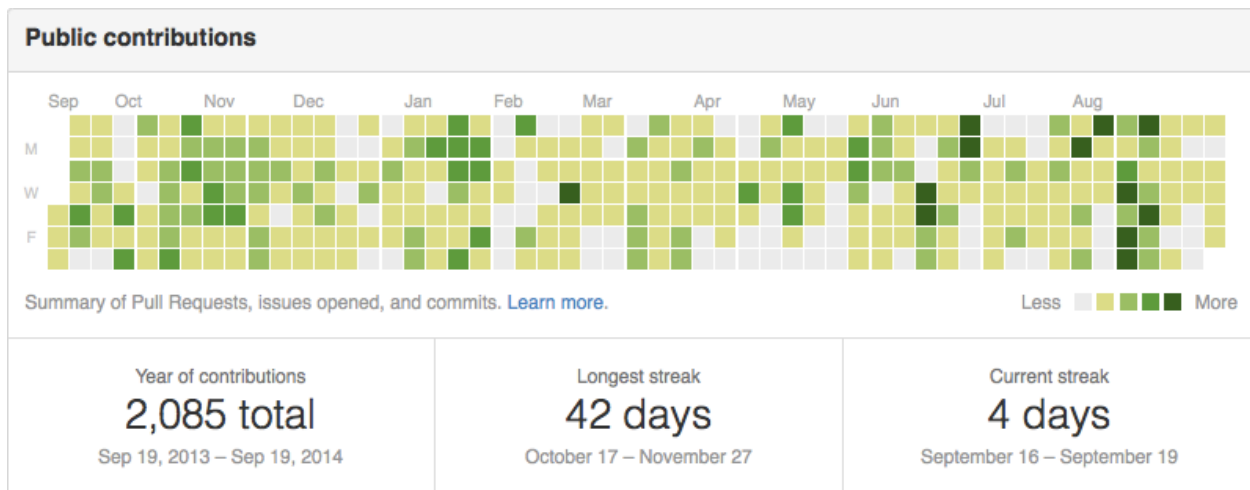0d0781f footer fix
a25d2af footer updates
View comparison for these 2 commits »

GitHub alternative text-based list visualization of activity history from most recent actions (shown here occurring 2 and 3 days ago). Here one can see the actions taken by the user as well as further details of these actions (such as the actual comment text written).

**Public contributions**

|  | Sep | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug |

Summary of Pull Requests, issues opened, and commits. Learn more.      Less ▢▢▢▢ More

| Year of contributions | Longest streak | Current streak |
|---|---|---|
| **2,085 total** | **42 days** | **4 days** |
| Sep 19, 2013 – Sep 19, 2014 | October 17 – November 27 | September 16 – September 19 |

Visualization of activity history over time in chromogram (color-based) format. Below are numbers providing further statistics about contributions made over one year, longest streak of activity to date and current longest streak of activity.

## Site 2: Stack Overflow

Stack Overflow is a question and answer site for software professionals. Members on this site have personal profiles that showcase their activity in a variety of different formats. For example:

| summary | answers | questions | tags | badges | favorites | bounties | reputation | activity |

**30,038 Answers**  votes  activity  newest

5461  Why is subtracting these two times (in 1927) giving a strang…
1561  What's the difference between String and string?
1491  Why is char[] preferred over String for passwords?
1091  What is the difference between Decimal, Float and Double i…
982  What are the correct version numbers for C#?

view more

**710,684 Reputation**  top 0.01% overall

+15  IEqualityComparer for Annoymous Type
+1  What could cause java.lang.reflect.InvocationTargetException?
+15  Class inheritance and retrieval
+15  Return the number of elements in a linked list recursively

view more

**31 Questions**  votes  activity  newest

791  What are the correct version numbers for C#?
363  What's your most controversial programming opinion?
329  Curious null-coalescing operator custom implicit conversion…
322  What's the strangest corner case you've seen in C# or .NE…
283  What's the hardest or most misunderstood aspect of LINQ?…

view more

**4,407 Tags**

128k  c#  × 16349        10k  generics  × 1096
74k  java  × 8872        8k  date  × 370
45k  .net  × 5004        7k  multithreading  × 950
18k  linq  × 2598        7k  timezone  × 219
10k  string  × 873       5k  arrays  × 699

view more

**15 Accounts**

| | | | |
|---|---|---|---|
| Stack Overflow | 710,684 rep | ● 296 ● 4633 ● 5999 |
| Meta Stack Exchange | 68,473 rep | ● 22 ● 152 ● 311 |
| Super User | 4,003 rep | ● 3 ● 20 ● 37 |
| Programmers | 2,841 rep | ● 6 ● 21 ● 25 |
| Server Fault | 2,544 rep | ● 15 ● 15 |

**655 Badges**  recent  class  name

● Good Answer × 1323        ● web-applications
● Guru × 473                ● code-review
● Nice Answer × 5463        ● hashcode
● Enlightened × 2606        ● struct
● jsp                       ● Great Answer × 161

"Summary" tab of user activity across various categories



| summary | answers | questions | tags | badges | favorites | bounties | reputation | activity |

**30,038 Answers**  votes  activity  newest

| 0 | Java compilation error: could not find or load main class Main | 3 hours ago |
| 1 | What character set does encode the "é" character as 130 in decimal? | 5 hours ago |
| 7 | IEqualityComparer for Annoymous Type | 5 hours ago |
| 2 | LINQ to Entities does not recognize the method Double Round(Double, Int32, System.MidpointRounding) method | 5 hours ago |
| 3 | Class inheritance and retrieval | 8 hours ago |
| 8 | Return the number of elements in a linked list recursively | 12 hours ago |
| 3 | Return a value from paralell stream? java | 12 hours ago |
| 2 | Will an inner class get GC'ed if enclosing type reference is maintained? | yesterday |
| 0 | Google Calendar V3 hangs when used outside local environment (asp.net) | yesterday |
| 7 | incompatible types required: Class<T> found: Class<CAP#1> where T is a type-variable | yesterday |

Chronological list of questions answered by user, from most recent first (recency shown by timestamps on the left).

85 **4** 3k **Why does struct alignment depend on whether a field type is primitive or user-defined?**
20 votes answers views

c# .net struct clr memory-alignment

jul 22 at 0:09 Hans Passant 487k

17 **3** 1k **How can I efficiently fetch a Mercurial change log in TeamCity?**
2 votes answers views

mercurial teamcity

apr 24 at 15:39 bigmike7801 1,143

87 **3** 4k **How can I perform a culture-sensitive "starts-with" operation from the middle of a string?**
6 votes answers views

.net string unicode

mar 19 at 17:05 Mårten Wikström 3,287

42 **1** 2k **How can I find out what's causing differences in generated Sandcastle docs?**
3 votes answer views

documentation sandcastle finalize shfb

feb 21 '13 at 11:54 Kharlos Dominguez 1,014

Chronological list of questions asked by user, with most recent first.

3h | comment | **Can I use hashcode of class member for class?**
Note that your `equals` method assumes that `name` isn't null, but your `hashCode` method assumes that it might be. Which is it?

3h | comment | **Java compilation error: could not find or load main class Main**
@aelgoa: Did the compile stage work properly? If you take out the manifest part, does your jar file build properly with everything you need? Why are you setting the `Class-Path` part of the manifest ?

3h | revised | ▶ **Java compilation error: could not find or load main class Main**
deleted 4 characters in body

3h | answered | ▶ **Java compilation error: could not find or load main class Main**

3h | comment | **Uploading file via WebClient**
Hint: catching an exception and not doing anything with it, not even logging it, is rarely a good idea.

3h | comment | **Async Await Without Task<T> method**
Well is `GetMyString` actually asynchronous in any useful way? If it's not, it's not going to work well when you use it with code which is *expecting* appropriate asynchronous code (i.e. it returns pretty much immediately, and the task will complete later).

4h | comment | **Linq Let Object not showing up in resuls - Creating a crosstab query**
Can you put this in the form of a short but complete program? That would really help us to help you.

4h | comment | **Linq Let Object not showing up in resuls - Creating a crosstab query**
You've got two commas in a row in your query in the "New With" part... I'm not a VB person, but I'm surprised that even builds.

4h | awarded | ● Good Answer

Overall activity feed of user, with all most recent actions displayed in a list.

710,684 **Reputation**       top **0.01%** overall

Summary graph of user reputation points gained over time (details to help interpret the graph are not provided, however).

157

## Jon Skeet  less info



# 710,684
reputation

● 296 ● 4633 ● 5999

Basic, aggregated information about user including total sum of reputation points and numbers/counts of various awards/badges (gold, silver, and bronze. As described by the site: "Gold badges recognize important contributions from members of the community. They are rarely awarded." "Silver badges are less common than bronze ones. You'll need to plan your strategy to get one of these." "Bronze badges encourage users to try out new features on the site. They are easy to get if you try!")
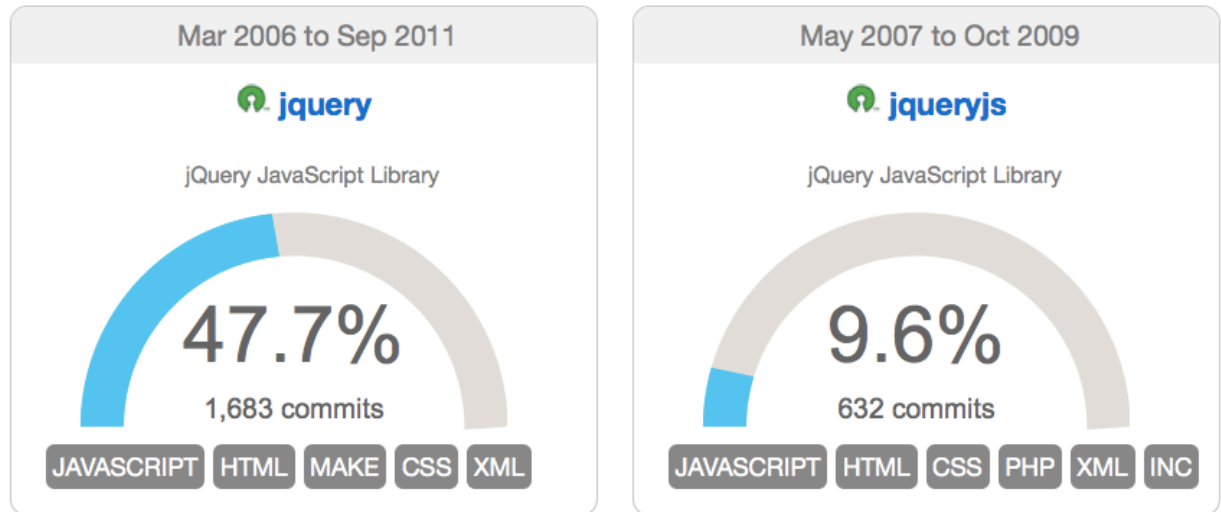
## Site 3: Masterbranch

Masterbranch is known as a "profile aggregator" for developers and allows people to show off their activity across a variety of software related sites.

## Languages & technologies

`JAVASCRIPT` `HTML` `CSS` `MAKE` `PHP` `XML` `PERL` `JQUERY` `SIZZLE` `QUNIT`

## Projects

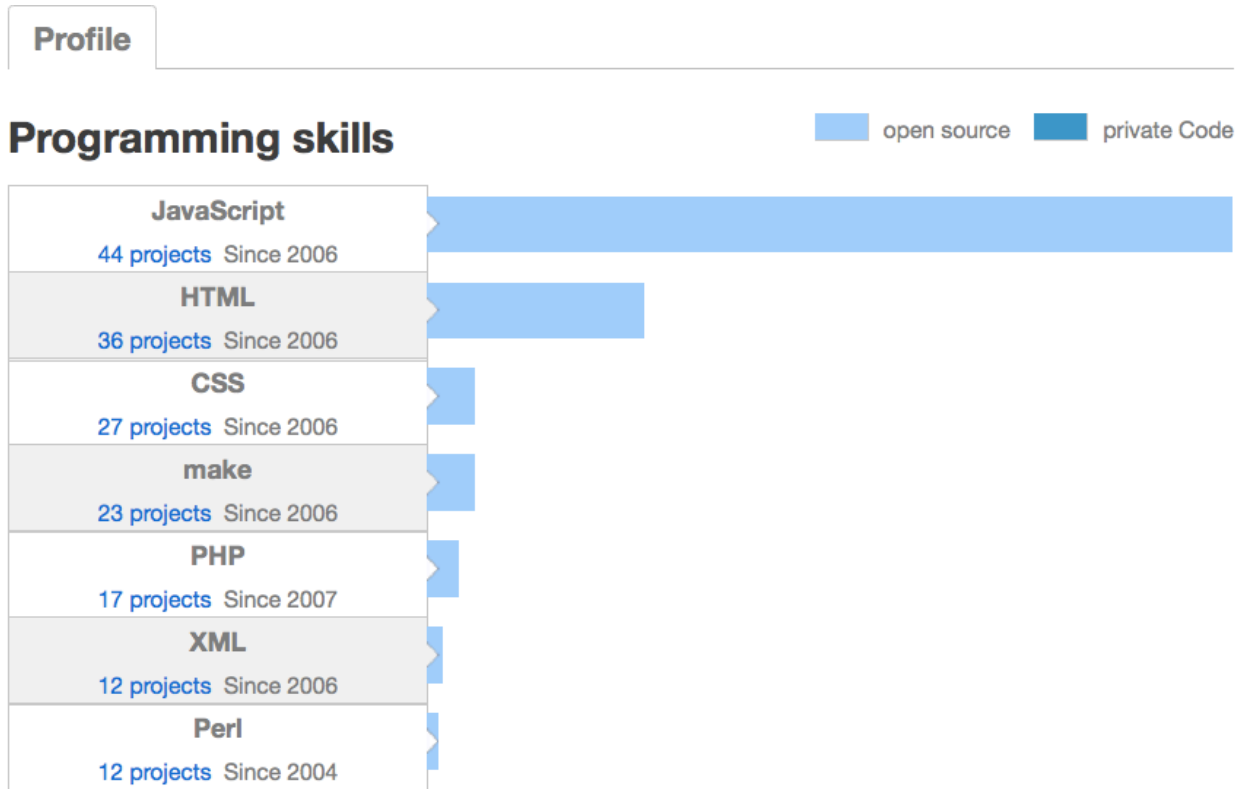| Mar 2006 to Sep 2011 | May 2007 to Oct 2009 |
|---|---|
| **jquery** | **jqueryjs** |
| jQuery JavaScript Library | jQuery JavaScript Library |
| **47.7%** | **9.6%** |
| 1,683 commits | 632 commits |
| `JAVASCRIPT` `HTML` `MAKE` `CSS` `XML` | `JAVASCRIPT` `HTML` `CSS` `PHP` `XML` `INC` |

Text-based tabs listing all languages used, text + graphical format showing percentage of total commits the user has made to various open source projects.

| | | |
|---|---|---|
| canvas-editor | **204** commits | 97.6% |
| khan-mobile | **108** commits | 47% |
| jquery-mobile | **75** commits | 1.3% |
| hdStore | **45** commits | 18.4% |
| wtpa-bot | **36** commits | 97.3% |
| dromaeo | **36** commits | 70.6% |
| fireunit | **33** commits | 44.6% |

Text-only based list showing the same information displayed in the image above, but in a less visual format.



**Profile**

**Programming skills**                                    open source    private Code

**JavaScript**
44 projects  Since 2006

**HTML**
36 projects  Since 2006

**CSS**
27 projects  Since 2006

**make**
23 projects  Since 2006

**PHP**
17 projects  Since 2007

**XML**
12 projects  Since 2006

**Perl**
12 projects  Since 2004

Graph showing behavior across projects (bar graph differentiates between open source and private code), text descriptions under programming languages indicate number of projects involved in and date started.

Summary of user (name and aggregate "DevScore")

## Site 4: OpenHub.net (previously Ohloh)



Overview of user



Basic user account summary (mostly summary, text-based)

## Commits by Language   *Sep 2007 - Present*
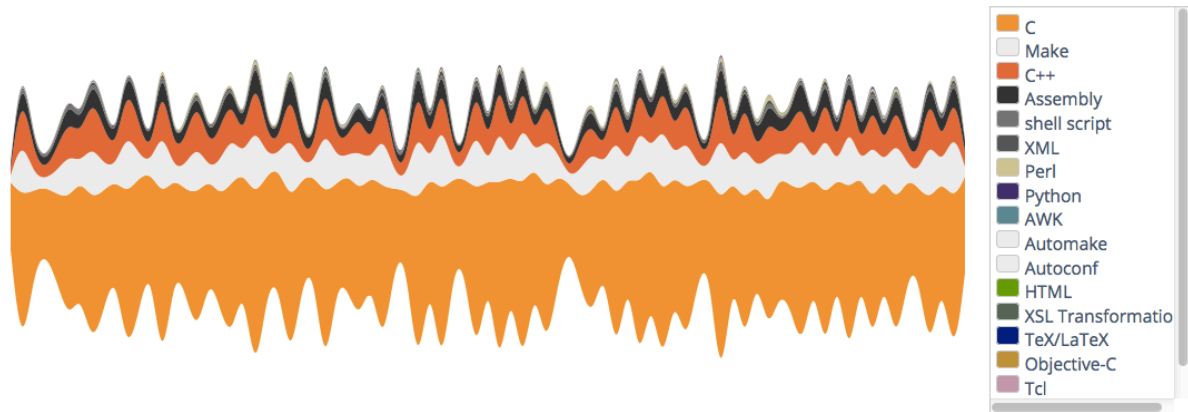


Graph showing commits by language over time

| Language | Aggregate Coding Time | Total Commits | Total Lines Changed | Most Commits: | | Most Recent Commit: | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | | | Project | Commits | Project | Commit Month |
| C | over 12 years | 21,032 | 47,254,663 | Linux Kernel | 21032 | Git | May 2014 |
| Make | over 12 years | 5,897 | 144,526 | Linux Kernel | 5897 | Linux Kernel | Sep 2014 |
| C++ | over 12 years | 4,871 | 717,373 | Linux Kernel | 4871 | Linux Kernel | Sep 2014 |
| Assembly | over 12 years | 2,757 | 930,742 | Linux Kernel | 2757 | Linux Kernel | Sep 2014 |
| shell script | over 10 years | 605 | 28,773 | Linux Kernel | 605 | Git | Jan 2014 |
| Perl | about 8 years | 286 | 29,983 | Linux Kernel | 286 | Linux Kernel | Sep 2014 |
| XML | almost 7 years | 251 | 192,533 | Linux Kernel | 251 | Linux Kernel | Aug 2014 |
| Python | over 2 years | 47 | 7,708 | Linux Kernel | 47 | Linux Kernel | Aug 2014 |
| AWK | almost 2 years | 37 | 1,462 | Linux Kernel | 37 | Linux Kernel | Aug 2014 |
| Tcl | 5 months | 9 | 2,627 | Git | 9 | Git | Feb 2008 |

Alternate view of contributions by language (detailed, text-based tabular format)

## Development History

**Commits by Project**   *Sep 2007 - Present*
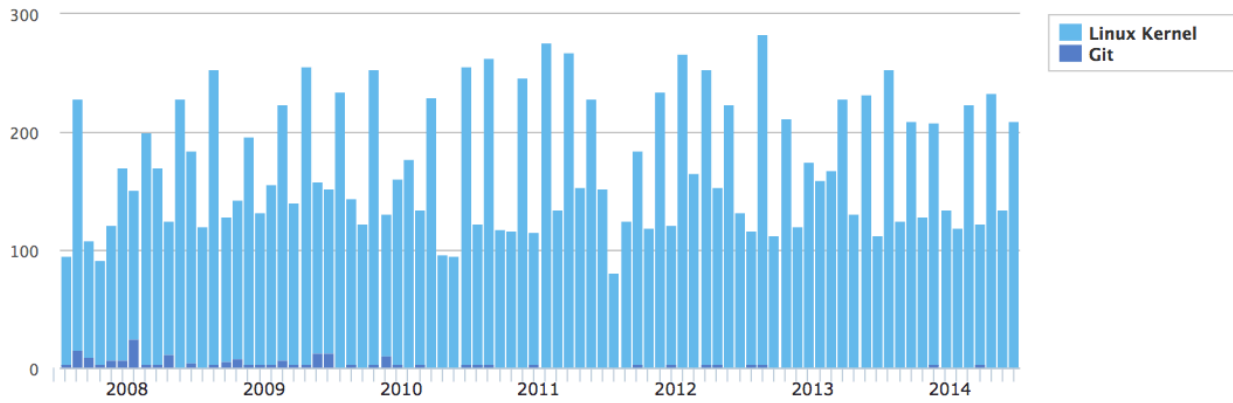


Graph showing user contributions to projects over time

## Activity on Git by Linus.Torvalds (see full profile for Linus.Torvalds)

| | |
|---|---|
| All-time Commits: 1109 | Overall Kudo Rank: |
| 12-Month Commits: 3 | First Commit: 07-Apr-2005 |
| 30-Day Commits: 0 | Last Commit: 29-May-2014 |

Summary view of activity on project (number of commits and time frame)