Mnikr: Reputation Construction Through Human Trading of Distributed Social Identities

Brendan Francis O'Connor
Department of Computer Science
The Johns Hopkins University
Baltimore, Maryland, United States
bfo@ussjoin.com

John Linwood Griffin
Department of Computer Science
The Johns Hopkins University
Baltimore, Maryland, United States
jlg@cs.jhu.edu

ABSTRACT

Reputation forms an important part of how we come to trust people in face-to-face interactions, and thus situations involving trust online have come to realize that reputation is an important characteristic in the digital age. We propose a new holistic and context-free approach to quantifying reputation on the Internet, based upon a stock exchange where users can trade reputation shares of other users and obtain goodwill dividends, including new algorithms for identifying and creating digital identities not inherently tied to a user's personally identifiable information. We developed such a system, named Mnikr, and deployed our system on the Internet for a month to demonstrate and evaluate this approach. Our results suggest that existing public data sources can indeed be used to create an overarching social network whose utility is greater than its number of users would indicate, and in which reputation measurements are generated that are actually indicative of each user's standing in society.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection; J.4 [**Social and Behavioral Sciences**]: Economics

General Terms

Security, Design, Algorithms

Keywords

Reputation, Identity, Social Networking, Stock Market, Digital Persona

1. INTRODUCTION

Reputation is a nebulous construct of our society. On the one hand, it is often easy to say that a person has a "good reputation" or a "bad reputation," based upon our understanding of the person and that person's role within a community. On the other hand, a person's reputation may be quite different between communities; people prized for their programming prowess may be vilified as "griefers" in computer gaming circles, or people known in one town for their kindness and generosity may in fact be escaped convicts known to the law enforcement of another. Such examples point to the subjectivity of reputation, and would seem to suggest that it would be a measure worth avoiding as we move more interactions to a digital world.

At the same time, however, reputation forms an important part of how we come to trust people in face-to-face interactions, and thus situations involving trust online have come to realize that reputation, or some analogue thereof, is an important characteristic in the digital age. Online marketplaces like eBay have created elaborate systems of reputation to allow people to feel that their experiences, good or bad, with merchants can serve as a message to others; whole websites, such as ResellerRatings, exist to serve the same goal across the Internet. In less critical—but no less interesting—circumstances, online destinations like HotOrNot provide a reputation score for a person's physical appearance.

All of these reputation systems, however, have severe shortcomings. Many of them are susceptible to the Sybil attack [6], which involves the creation of trivial user accounts to undermine any multi-user system, rendering such systems subject to the whims of massive and mobile online groups. Other reputation systems—for instance, the eBay auction site—require some sort of monetary transaction to earn the right of commentary. Such systems are subject to extortion because of this property, something rendering their ratings nearly meaningless except as a most rough-hewn marker. All of these systems are subject to the vagaries of the contexts in which they were formed, and are indeed inextricable from those contexts. This might seem like a feature to be desired—one might not care, for the purposes of an eBay transaction, how well one does at a particular game, or about the details of one's physical appearance—but in the quest for a holistic reputation, it is important to realize that these very context-sensitive reputation values are not necessarily related to a person as a whole, only to that person's performance in a particular task. Clearly, some other system needs to exist that can achieve the fairly difficult goals inherent in a reputation system.

Toward this need, we pose the question: Can the idea of buying and selling shares on a theoretical market be used to gain a machine-readable understanding of traditionally unquantifiable data, without resorting to the techniques of machine learning or artificial neural networks? If so, there

are several secondary questions of interest: How effective would such a system be at establishing measurably different reputations? How well would its reputations correlate with real-world perceptions of reputation? Finally, and perhaps most importantly, would such a system truly be useful across multiple fields, rather than having applicability only to one context?

As more transactions move entirely out of a physical realm, it is useful to give people a means by which to authenticate an entire person, rather than just that person's activities in one area; this helps to increase users' trust of digital interactions. In face-to-face transactions, such global validation is accomplished through asking others about a name, but as we move away from given names and verified identity, it is necessary to give people something of the same ability without tying everyone to a given identifier, validated by a central agency such as a government (as names are). This need for independent validation provides the rationale for a context-free reputation system.

Toward answering these questions, we examine the design of a system, called Mnikr (pronounced [ˈmɒn.ɪ.kɚ]), that we created to facilitate the creation of online reputations. Mnikr automatically generates combined digital identities, using publicly-available information sources to find both different parts of one identity, and the links between different identities. Users of Mnikr can then buy and sell shares of these identities on a stock exchange, and this process determines the value of an identity's reputation.

The remainder of this paper is organized as follows. In §2 we define the terminology we use to describe Mnikr. In §3, we discuss the work related to Mnikr and digital reputation generally. In §4 we examine the design of the Mnikr system, including the methods for determining what constitutes a single digital identity. Our implementation notes on this design, including the tradeoffs made to make the system practical, are listed in §5. We exposed the Mnikr system to the public Internet for 32 days and we report on and interpret our preliminary results in §6. In §7 we discuss the implications of context-free reputation, and the directions for future work in this area, and in §8 we conclude.

2. DEFINITIONS

We refer to several components of the Social Web throughout this paper; for clarity, we suggest definitions for these components as follows:

User: An actual human being, or other autonomous entity, capable of using services; in terms of Mnikr, a User represents a human being who has logged in.

Service: A web application, such as Flickr, Twitter, or YouTube, that provides some sort of per-user value.

Persona: The account of one User on one particular Service, represented with a particular username which is only unique to that service. A User may have many Personas on many different Services, and a User may have multiple Personas on one Service.

Identity: The combination of many Personas, all held by the same User. This may not be entirely determinable, but in general, if there exists data linking multiple Personas together, we say that in their entirety they represent one Identity. For instance, if a User has the username `Bob` on Twitter, `Robert` on LinkedIn, and `ChunkyMonkey` on MySpace, and there exists data linking these various Personas, we can say that these Personas form one Identity.

Activity: From the Atom Activity Extensions specification [1], we will define an activity as "a description of an action that was performed at some instant in time by some actor [a Persona]... usually on some social object...." and an Activity Stream as a collection of such actions. This is also known in some groups as an Action Stream. Each Persona, then, will have its own Activity Stream; an Activity Stream for an Identity will consist of an amalgamation of the Activity Streams for all its associated Personas.

Figure 1 illustrates the relationships among these terms.

3. RELATED WORK

When considering related work, it is important first to note that while the idea of a reputation stock market has not been proposed within academic literature, we cannot claim that the idea is ours alone. In the book *Accelerando* [11] by author and futurist Charles Stross, the main character refers to his "publicly traded reputation," and the idea of reputation systems not only being portable across contexts, but being usable to settle disputes; the idea of reputation dividends being considered to be "goodwill" comes from the same work. Stross does not develop the idea further, as it is not particularly important to the core of the book.

The primary idea differentiating our approach from that of the related academic work in this field is the notion that in society, one's reputation does not stand by itself, but rather as part of a community; to put it bluntly, much is based on who you know. This idea can be seen in everything from middle-school dramatics to the system of academic rankings of universities, based on the best-known professors in a subject area as determined by their peers across the country. This idea is not new to the field of computer science; for example, it forms the centerpiece of Google's PageRank [3] algorithm, wherein sites with a high PageRank can contribute more to other sites than sites with a low PageRank, but to the authors' knowledge it has not been applied in systems of reputation applying to humans, rather than websites.

Another significant differentiator from the current state of the art is the central realization that existing context-sensitive reputation data is not able usefully to be decontextualized. Several large projects, such as those of Windley et al. [13] and the OpenPrivacy Project [9], have focused on the aggregation of reputation data created at a variety of existing sources, and its subsequent decontextualization to attempt to remove the effects of those sources, and come to a null result; additional research has gone into attempting to rectify mistakes in these external data sources [12], which, while the corrections to the data might be useful to those data sources, is ultimately, as the authors noted, not helpful in creating a working reputation system that provides a holistic measure of a person. Also in the area of decontextualization, the work of the Google identity team [5] deals with the problem of coalesced social graphs, ultimately deciding that the consequences for such decontextualization need to be better understood; this sort of decontextualization, however, Mnikr does in its Identity coalescing algorithms.

One other significant area in reputation systems, as well as in many other sorts of network-based applications, is in dealing with the Sybil attack [6]—the collusion of many pseudonymous users to create trivial self-representations to destroy whatever particular metric an application seeks to create. Sherchan et al. [10] seek to address this problem through the use of an external artificially intelligent agent
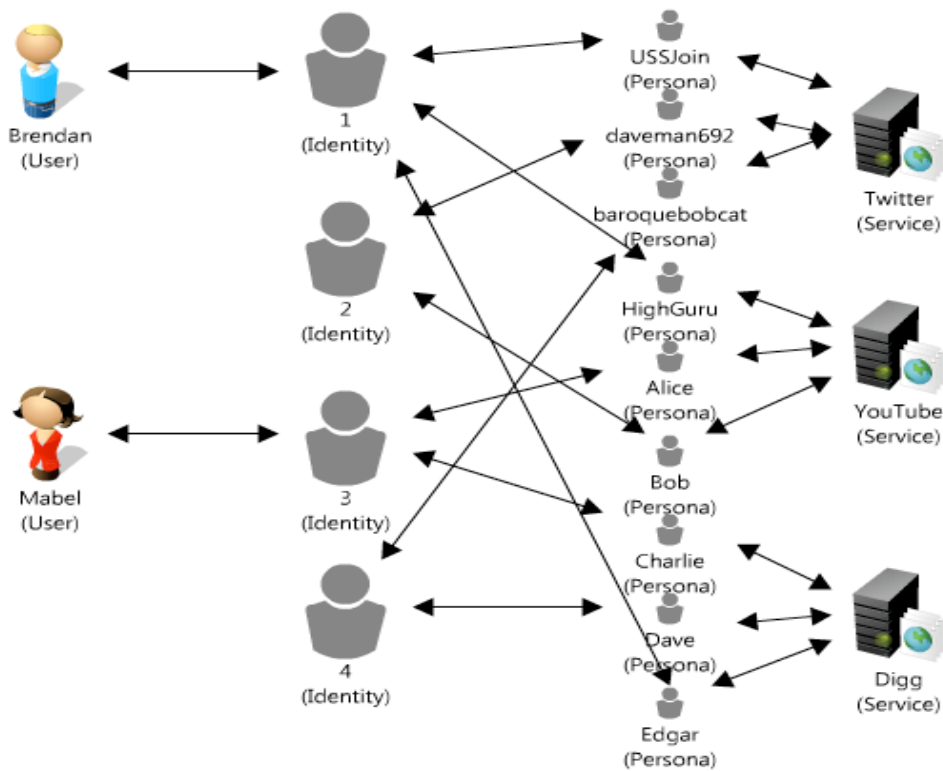
Figure 1: This shows the relationship between Users, Services, Personas, and Identities. Each Service will have multiple Personas on it, and an Identity will consist of Personas from multiple Services; one User generally has only one Identity. Note that the existence of an Identity does not necessarily imply that a Mnikr User exists for it.

to determine whether a user has malicious intent in the manipulation of a reputation system, but they find it to be an extraordinarily difficult problem to solve. Mnikr addresses this problem more as an inherent advantage of its construction; it grants the power to change reputation values only to established reputations—which are themselves defined as established by the fact that other established reputations invest in them. This power is defined in terms of its internal currency of points, which are in general derived from dividends; thus, reputations that are highly valued will have more points to spend on buying shares of other reputations. A tradeoff of this solution is that it requires a bootstrapping process to establish initial Identities with the means to make reputation trades (see §4.2). On the other hand, this solution prevents the need for difficult heuristic solutions to the Sybil problem, while at the same time avoiding most of the original author's proposed solutions of draconian identity verification measures; indeed, Mnikr requires no ties between its coalesced Identities and a user's true identity in real life at all, allowing pseudonymous persons to participate as first-class Users.

Finally, there is much interesting work going on relating economics and reputation, though not in the same vein as Mnikr's reputation trading system. Yan and Van Roy [14] deal with a market for the acquisition of what they call reputation data, but what might be better termed personally identifiable information. Their market-based system, then, determines the truth or fiction level in a particular set of data through demand levels. Mnikr does not attempt to as-

sert truth levels in its found data, simply conveying other parties' assertions of fact; this is beneficial, in that this again allows pseudonymous entities to utilize Mnikr as first-class Users. The other area of inquiry toward economics and reputation is the work of Joseph Blocher [2], which explores whether one's online reputation may be thought of as property which exists in a virtual economy. While the author speaks of a sort of metaphorical economy, Mnikr actually makes real this idea, through a direct system of trades and currency for reputation creation, valuation, and exchange. To the best of our knowledge, this is an approach that has not previously been tried in the area of digital reputation.

4. DESIGN

Having now considered what makes Mnikr different from the work that has come before, we can now examine the makeup of the Mnikr system. Mnikr is a web application, exposed to the public Internet, available for the time being at `http://www.mnikr.com`.

In this section we discuss the high-level design for Mnikr (§4.1), the mechanics of the reputation trading system (§4.2), the data formats chosen for the system (§4.3), the assumptions made in this design and the sacrifices those imply (§4.4), and the algorithms for creating and managing the digital Identities (§4.5).

4.1 Design overview

From the perspective of a User, Mnikr is fairly simply laid out. Aside from standard website staples like a help page, terms of service, etc., Mnikr consists primarily of a set of profile pages, one for each Identity known to the system at a given time.

Each Identity has a numeric identifier, which, although not guaranteed to remain the same over time, does uniquely identify the Identity at a particular point in time. Each Identity's profile page can be reached either through this identifier, or through referencing any Persona; in this latter case, one could access the Identity that contains the Persona with the username `USSJoin` on the Twitter Service through the URL `http://www.mnikr.com/profile/ussjoin@twitter`.

Each profile page contains a unified Activity Stream for all the Personas linked to the Identity (see Figure 2), as well as a list of all such Personas, a list of all the Identities marked on any Service as friends or contacts of the Identity, and the Identity's reputation score. The profile page also contains buttons to buy or sell reputation scores (if the User has logged in).

Users also have access to a list of their most recent reputation trades, as well as their *portfolio*, their set of all reputations owned. The goal with the full design is to allow a relatively low-friction interface for Users to buy and sell reputation shares.

4.2 Reputation Trading System

The central idea in this reputation trading system is that when a user buys a share of an Identity, that user either likes the Identity or believes that they will get high returns from their purchase. Either case is an expression that the Identity's reputation has value; purchasing a share of an Identity is not precisely the same as a recommendation of an Identity, but it does express confidence in the Identity and the person's contributions to society. Dividends earned from this stock market can be seen as reaping the "good will" of others towards those with positive reputations. An important point in Mnikr that differentiates it from a standard stock market is that users cannot make negative value judgements about an Identity. Instead, they can only sell their own shares (if they have any), but they cannot make a "short-selling" scheme where they express confidence that the user is overvalued.

The reputation trading system is set out like a stock market; Users can buy and sell shares of Identities, and this activity determines the Identity's reputation in the Mnikr system. Unlike a stock market, however, where the commodities being traded—for instance, corporations—can determine the number of shares offered, Mnikr instead treats shares as infinite and fungible. This, then, means that since there is unlimited supply, Users will not buy from other Users, as would happen in a traditional stock market; instead, Users buy from and sell to the Mnikr system.

The other consequence of supply being infinite is that scarcity, which (in theory) sets the asking price for purchasing shares in a traditional stock market, is not a factor; we must thus define a different pricing structure for reputations, which will in turn become the value for a reputation. In Mnikr, this is defined to be the number of all shares of a particular Identity owned in the system, plus one; that is, if fifteen shares are held for a particular Identity, that Identity's reputation will be 16, and if a user wishes to buy a share of that identity, they will need to pay 16 points, after which the reputation will be 17; if they then wish to sell a share of that Identity, they will be compensated 16 points, after which the reputation value will once again be 16; while the sales compensation seems incorrectly low, the alternative—to compensate the user at the start of the sale transaction, rather than at its end—generates an infinite amount of money, given unlimited transactions; $X - 16 + 17 = X + 1$, which is undesirable. The plus-one in the calculation of a reputation's current value allows us to avoid having Users be able to purchase a share for no cost, and can be thought of as an Identity owning one share of their own reputation. Note that this system intentionally gives no value to network connections in and of themselves; for a connection to provide value to a reputation score, the connected Identity must purchase reputation shares.

To buy and sell shares, there must first be some sort of currency with which to undertake this activity, and accordingly, we have set up a currency with value only inside Mnikr, which is referred to in the user interface as *points*. This allows us to hand out points with relative impunity at the start of the project, a necessary step to put initial value into the market. After an initial bootstrapping period, giving users points will no longer be necessary, and all users' points can come from the dividends produced by their portfolios.

To bootstrap the system we gave prospective users a starting base of 50 points with which to make their first trades. They were then free to undertake trading, as well as collection of dividends, as normal. We did not adjust starting reputation scores, which meant that each Identity began with a starting reputation score of 1, representing the share each Identity holds of itself, and then moved up from there.

These dividends are calculated as follows: every 24 hours, each Identity receives

$$\frac{reputationValue * numberOfSharesOwnedByThisIdentity}{20}$$

points, rounded down, for each Identity of which it owns a share; it then receives

$$\frac{reputationValue}{5}$$

points, rounded down, for its own reputation value. These numbers are somewhat arbitrary, but are designed to encourage investment through granting dividends for any significant investment in a lower-value Identity, and any investment at all in a higher-value Identity. The effect of changing these constants would be one area for future work (see §7).

Note that we award dividends to each *Identity*, rather than to each User. This means that Identities with high reputation scores, whether or not they are associated with a Mnikr User, can accrue points which will then be available to the appropriate person should they decide to become a User.

There is no concept of interest on points held in reserve, nor is there a penalty for such behavior; points held on an Identity do not change over time. On a related note, there is no opportunity cost for buying or selling shares; that is, a User who buys, then immediately sells a share will have the same amount of points with which he or she started. The effect of introducing market friction in this fashion would be another interesting area for future work, as certainly modern stock markets have this friction.

The other significant and arbitrary choice in the reputation market is that we limit the number of shares of a single

Figure 2: An example Activity Stream from a Profile page on Mnikr. This stream contains only actions from Twitter, but a stream can contain actions from an unlimited number of sources.

Identity held by one other single Identity to five. This is to prevent one User from unduly influencing the market, and is in accordance with the concepts outlined in previous sections of having a reputation derived from an entire community, rather than one person of high reputation, and thus, due to dividends, high means. While five was chosen as an arbitrary number, it is related to the numbers chosen for the dividends, as it will allow a user who invests fully in even an Identity with the lowest possible reputation to receive dividends immediately, which in turn encourages Users to invest even in Identities not yet recognized by the community, and thus in a broader range of Identities.

4.3 Data Formats

On the Web, it would be a luxury to assume that all data exists in a controlled schema, properly formatted and escaped. We have made as an assumption, however, that we can extract sufficient amounts of useful data from microformatted data—i.e., data existing for other purposes on the Internet that has had small tags applied to it to mark it as a specific data type, such as identity or contact data. For more on this assumption and its consequences, see §4.4.

The primary microformat for identity data that we will be using is the XHTML[1] Friends Network, or XFN [4]. This standard was created as part of an effort to give machine-readable semantic connotations to the web without creating whole new formats; accordingly, it works through adding relationship metadata, such as whether a linked-to resource denotes a friend, spouse, sibling, or another location owned by the current document's author, to a standard web link. This technology allows us to find all of the identity information stores (also referred to as *identity silos*) associated with one person, and all of that person's outgoing friendship links, without using semantic context analysis, which is necessarily a very error-prone task. While utilizing XFN, as will be discussed later, does not guarantee an error-free coalescing of an Identity, it greatly simplifies the task.

To obtain information about each Identity's Actions under each Persona, we will be using the techniques outlined in the Activity Extensions specification; as that is a relatively new document, we will in general use the publicly-available per-user streams generated by each Service, and use Activity Streams Verbs only when available, as while they provide additional insight into each action (for instance, they might describe the original source of a web link shared by a user), they are not necessary for the basic gathering of information. These per-user feeds are ordinarily available as either RSS[2] or Atom[3] feeds. There is no standardized location across Services at which to publish such feeds— although some work has gone into a standard for web agents to discover such per-user feeds' locations[4], this standard has not been widely adopted (in contrast to the wide adoption of the idiomatic discovery method for syndication feeds for the content displayed on a particular page)—and so we have used a compiled list of feeds for the services we support.

4.4 Assumptions

The primary simplifying assumption we will make, as we tackle the algorithms for coalescing discrete Personas into Identities, is that the XFN data available accurately represents the links between identity silos. This assumption is not without its problems. One issue is that people may not bother to keep their XFN links up to date, and thus data no longer associated with them may be integrated. However, there is no particular solution to this; as correlating data, such as usernames, may not be the same, or even similar, across services, and not all services have (or choose to expose) other data such as real name, there is not a source of correlating data to match against the XFN. The problem does not, however, appear to be hugely significant; only three Mnikr Users noted any incorrect additions to their profiles during the testing period, out of the entire User base of 41 Users. The other issue with relying upon XFN data is that users may not have linked all their profiles together using XFN data, and thus multiple Identities may represent the same person. This is actually a feature; as Mnikr is concerned with digital Identity, these different Identities can be considered to be discrete people, since they have no digital interaction. This may indeed be the intent: For instance, if a user creates different accounts for work and personal use, and acts separately for each, there is no reason those could or should be construed to be the same. When it is not the

---

[1]The Extensible Hypertext Markup Language, defined by the Worldwide Web Consortium and available at `http://www.w3.org/TR/2001/REC-xhtml11-20010531/`.

[2]Really Simple Syndication 2.0, as defined most currently at `http://cyber.law.harvard.edu/rss/rss.html`

[3]Atom Syndication Format, as promulgated as IETF RFC 4287

[4]The Social Graph Node system, with more information available at `http://code.google.com/p/google-sgnodemapper/`

intent, Users can combine their Identities through claiming Personas found in the other Identities; this claim, once validated through some means (in the current implementation, a successful OAuth[5] validation of the claim) can allow users to unify their Identities.

In general, we assume that the XFN data to which we have access will be good enough for our purposes. Being able to add additional sources of correlative data would be an improvement for future work.

4.5 Algorithms

The most significant algorithms in Mnikr are those dealing with the creation and coalescing of an Identity from distributed Personas.

The lifetime for an Identity begins at its creation, the algorithm for which can be expressed as follows:

1. Determine the username and Service name from the URL, and create a Persona with those characteristics, associated with a new Identity object.

2. Take the URL, and query the XFN data source (more information about this in the next section) for all URLs linked from the given URL by an XFN `"me"` link.

3. For each URL, find the username and Service, and create a new Persona object linked to the current Identity.

This algorithm can lead to duplicates, due to one-way `"me"` links; for instance, if a user has two accounts, one on service A and one on service B, if A has a link to B but not the reverse, then if the algorithm is run on A first, then B, then just one Identity with two Personas will be created. If the algorithm is run on B first, then A, however, two Identities will be created: one with Persona B, and one with Persona A and Persona B. This possibility must exist, as some Personas will exist on multiple Identities; for instance, when two humans both update a shared corporate Twitter account, for public relations (PR) reasons. In such a case, it would be incorrect to combine the two Identities, as they will in all other situations act separately, and represent two different digital people. However, since in some cases this will just be an ordering bug, we must also have an algorithm for determining when duplicates may exist, and if duplicates are found, conditionally to combine them.

The algorithm for determining possible duplicates is as follows:

1. For each Identity I, find all its Personas, and store them in a list L.

2. Find any Persona that has a username and Service combination in L, but whose Identity is not I; store them in list P.

3. If size(P) > 0, then store the Identities related to each Persona in P in a set D; add I to this set.

This algorithm gives us possible duplicates, but not all sets found in this fashion will be true duplicate Identities; for instance, in the situation of the PR account mentioned above, each of the PR workers' Identities would be marked as a possible duplicate. Therefore, to prevent incorrect combining of Identities, we use a fairly restrictive algorithm to determine whether two Identities should be combined:

1. For each pair of Identities, I and J, in a duplicate set D:

2. If for every Persona on I, there exists a Persona on J with the same username and Service, or vice versa (that is, that the set of Personas of I is a subset of the set of Personas of J, or the reverse), then combine the two Identities; they are the same.

There will of course be some duplicate sets left over after this last algorithm runs; this is acceptable, as it is better to have incorrectly separate Identities than incorrectly combined Identities.

5. IMPLEMENTATION

Mnikr is implemented as a Ruby on Rails web application, and in general has the Model-View-Controller (MVC) [7] layout common to applications written in that framework. The primary data models are:

- User: representing an actual person who has logged in to the web application,

- Identity: representing an Identity as defined above,

- Persona: representing a Persona as defined above,

- Investor: representing one Identity buying shares in another's reputation.

These models, as well as other data relevant to the web application, are stored in a MySQL database. The web server used is the Apache HTTPD Server, serving the Rails application through the Phusion Passenger module, with the Ruby code being executed by the Ruby Enterprise Edition interpreter.

In addition to the primary web application, some of the periodic tasks, such as rebuilding Identities, are run asynchronously by a fleet of distributed worker processes, orchestrated by the Gearman server. These workers are written in Ruby, and are run through the Rails Rake system, giving them access to the database and object models of the Mnikr web application. All of the above processes are executed on a Virtual Private Server with appropriate resources and Internet connectivity.

Mnikr does not store Action Stream information; this was attempted, but once the number of Personas approached one million, the amount of bandwidth and processing power necessary to support this effort was unsustainable. Therefore, all Action Stream data is fetched using AJAX[6] when an Identity's profile page is loaded by a user, and only proxied, due to JavaScript security restrictions, though Mnikr.

5.1 XFN Microformat

The XFN microformat is designed to be a very lightweight addition to the data users already put on the Web; as such, it resides entirely within the `"rel"` optional element for links in XHTML.

For instance, to link to the author's home page from some other website, one might ordinarily use the XHTML code:

```
<a href="http://ussjoin.com">Brendan Fran-
cis O'Connor</a>
```

Using XFN, since this document and that page share an author, that relationship would be indicated in the modified code as:

---

[5]OAuth is an open authorization protocol to allow granting permissions without giving the grantee an account password; its primary site is `http://oauth.net/`

[6]Asynchronous JavaScript and XML, referring to the technique of using JavaScript to load data on a website after the initial rendering of a page

```
<a href="http://ussjoin.com" rel="me">Brendan
Francis O'Connor</a>.
```
To indicate a friendship, one would use:
```
rel="friend"
```
For a business contact:
```
rel="contact"
```
And so on. The entire list of possible relationships is not particularly large, but is available in the XFN standard [4]; in Mnikr, we interpret all XFN data as one of two classes, either `"me"` or `"other"`, the latter of which we treat as a friend relationship. Once again, note that these friend relationships have no bearing on reputation scores; we use them solely to begin to display the network of contacts surrounding each user.

5.2 Extant Web Tool Dependencies

XFN data, as used by Mnikr, can be obtained by directly crawling the profiles representing the Personas for every Identity in the system. However, with more than 1.1 million Personas in the Mnikr system (see Results, below), this would quickly become unwieldy; taking the author's Twitter profile page as a representative example (while many profile pages contain more information than a Twitter profile page, few contain significantly less, and in any case, this number is only for illustration), a single Profile page contains (excluding overhead data required to send the HTTP commands, headers, etc.) approximately 11 kilobytes of data to obtain just the raw XHTML—that is, the minimum data necessary to parse XFN information, excluding things that would be necessary to display the page in a web browser, such as CSS[7] information, images, etc. To crawl all the profiles, then, would use approximately 11.5 gigabytes of data (again, excluding overhead); while this might not seem like an incredible amount of bandwidth, one must also realize that since XFN data is constantly changing as users add and remove friends, as well as link other identity silos to their existing ones, this crawl must be done periodically to catch updates. To do this crawl sufficiently often—perhaps multiple times daily, for the highest-traffic sites—would impose a high burden on each Service. Additionally, this crawl would suffer major issues due to latency, as well as the simple transient nature of the Internet—with sites and backbones constantly going up and down, it would be a significant undertaking to do such a crawl even once, let alone with sufficient frequency to offer a good user experience.

To solve all of these issues, rather than crawl all of the sites for XFN data individually, we use the Social Graph API[8], provided by Google, Inc. The SGAPI allows one to make a single query to obtain all the linked identity silos for a given starting profile or profiles, or to obtain all the linked friend or contact profiles for a given starting profile or profiles. The bandwidth savings here are significant; to continue on the same example from before, a call for the author's entire list of identity silo pages, starting from the aforementioned Twitter account, resulted in only approximately 3.5 kilobytes of data—and this number is compared to 11.5 kilobytes for *each* of the identity silos linked for the author's Identity; 27 at the moment. In addition, Google gathers this data during

its regular and frequent crawls of the Internet, meaning that Mnikr's querying for this data puts no additional load on the servers for the individual Services; since Google is extremely large, it is a valid assumption at this time that the traffic of Mnikr's queries is insignificant. In addition, Google as a whole has approximately zero downtime and, due to its geographical distribution, very low latency, meaning that many of the aforementioned problems with a whole-Internet crawl are neutralized. Finally, the SGAPI takes the useful step of normalizing the URLs it serves, which can prevent some sorts of simple errors in the Mnikr system.

Therefore, Mnikr uses the Social Graph API as its exclusive source of identity silo and friend linking data. In practice, Mnikr stores the discovered identity silo information, using it to build Persona and Identity objects, and rebuilds these objects once every 24 hours. The friend connection information, by contrast, is pulled on demand from the SGAPI and parsed to find the relevant Identity objects given the links found, as well as to de-duplicate the list (for instance, if an Identity has made friends with a user on both Twitter and Digg, the SGAPI will return both profiles as friends of the Identity in question; Mnikr then finds that both are Personas of the same Identity, and lists the Identity only once); this is done for the reason that storing and updating a graph of infinite size and complexity is a difficult problem. As noted, however, if one wished to, one could remove entirely the dependency upon the SGAPI, if one was willing to shoulder the burdens thereby imposed, and obtain the same results.
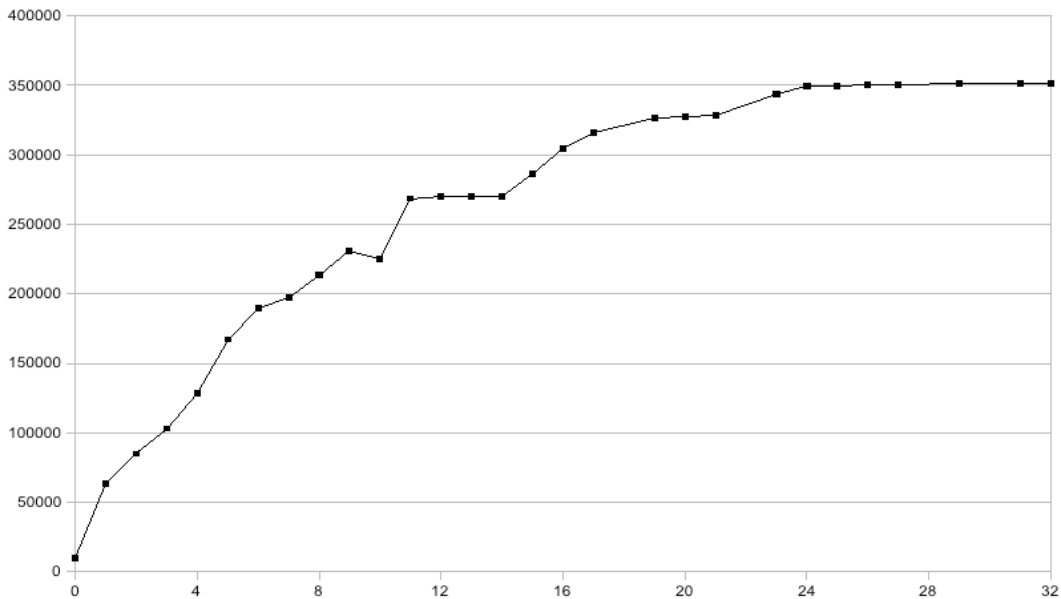
6. RESULTS

As a preliminary exploration of the concept of reputation trading, we deployed Mnikr for a period of 32 days and attempted to solicit users from the community through public blog posts, and posting information about the project to relevant message boards. In 32 days of usage, 41 Users registered with Mnikr and our system collected 351,356 Identities, comprised of 543,902 validated Personas. Figure 3 shows the growth in number of identities in Mnikr over time.

As can be seen from this graph, Mnikr started with just under ten thousand Identities stored from testing work, and grew quite quickly at the outset as users viewed profiles for Identities, triggering the algorithms to coalesce the linked Identities. The dip in Identities visible on the graph on day 10 is due to the first run of the Identity combination algorithm detailed at the end of §4. Given this data set, we can now examine how Mnikr was used, and what this reputation system created.

6.1 The Network Effect

As a social network, one of the goals for Mnikr was that it be free from Metcalfe's Law of network utility [8]; that is, that its utility be greater than the square of its users. Mnikr has the ability to achieve this through its dynamic Identity creation; rather than creating Identities only for Users who take the time to create an account, Mnikr finds Identities as they are referenced by existing Identities, which may or may not be tied to Users. The way this growth is kept from being unchecked is that Identities are only created when an Identity linking to them is *accessed*, rather than created; that is, if an Identity A links to B, and B links to C, when A is first accessed, B will be created. C, however, will not be created until B is first accessed. If we apply Metcalfe's

---
[7]Cascading Style Sheets, the W3C standard for adding styling information to a web page; the current standard, Level 2 Revision 1, is available from `http://www.w3.org/TR/CSS21/`

[8]Information about the Social Graph API (SGAPI) is available from `http://code.google.com/apis/socialgraph/`

Figure 3: This shows the number of Identities known to Mnikr as a function of days since it was deployed to the public Internet.

law to Mnikr as normal, then, with 41 users it would have an objective utility measure of 1,681. If, however, we use Identities as the central measure of utility, objective utility would increase without limit as Users continue to explore; at the end of the data collection period, the utility would be calculated as $351,356^2$, or 123,451,038,736!

Determining utility, however, is more than the possibility of accessing data; while overall utility is perhaps too subjective to measure directly, we can ask the related question: "Are Users able meaningfully to interact with Identities not associated with another User?"

The answer, from the data, is yes. Of the top twenty Identities by reputation value (the top ten of which can be seen in Table 1), only four were associated with Users of Mnikr. At least one other account was related to the same person as a Mnikr User, but that person had created a separate Identity for Mnikr use. Those four Identities were also not inherently the most popular of the top twenty; instead, they were fairly well-distributed throughout, being ranked fourth, sixth, eleventh, and eighteenth in the top twenty. Mnikr seems to provide utility regardless of the number of Users, which is an interesting result for social networks in itself.

6.2 Correlation of Reputation

A primary question for Mnikr is whether the reputation scores produced are actually indicative of the person's standing in society—that is, whether the reputation scores are correlated with their actual reputation. We acknowledge that this question cannot fully be answered with a 32-day sample of trading data. Therefore, as an initial sanity check we examine the highest-ranked Identities on Mnikr to see whether those people could reasonably be expected to be well-thought-of by a typical user of new web applications.

As listed in Table 1, we see that the top Identities are indeed figures of note, for the most part. Chris Messina is a founder of the Distributed Social project, which cre-

Nickname	Reputation Score
Chris Messina	15
Why The Lucky Stiff	13
Brendan O'Connor	12
Yehuda Katz	12
Mnikr	9
David Recordon	9
Nick Howard	9
Chad Fowler	8
Dave Troy	8
Jonathan Coulton	7

Table 1: The top ten Identities at the end of the data collection period.

ates projects that rely upon pulling social data from across the Internet, much like Mnikr. "Why The Lucky Stiff" is a pseudonymous master programmer considered to be one of the greatest, certainly one of the most prolific, Ruby programmers in existence. Yehuda Katz and Chad Fowler are both internationally-known Ruby and Rails experts. David Recordon and Dave Troy are both significant figures in different parts of advocacy for the Open Social Web. Jonathan Coulton is a singer and songwriter who also has made a significant push toward releasing content on the Internet, and Nick Howard is another prolific Ruby programmer and blogger. All of these people, then, could credibly have a high reputation among significant users of the Internet.

These reputations can also serve to validate another correlation concept: that the friends of those with high reputation will also tend to have high reputation. While Mnikr does not have a high enough user group to bring every user above a base reputation level, the data that is available does tend to support this hypothesis, as David Recordon has exposed a friendship relationship with the author and Chris Messina,

Identity ID	Days
22417	29
22433	27
33424	0
93863	0
260004	0
303214	0
337775	7
342490	0
342912	0
357579	1
361191	0
382575	0
384706	0

Table 2: The difference in days between a User's initial reputation trade and their last trade, for all Users who made at least one trade during the data collection period.

the author is friends with six of the other nine people in the top ten, Dave Troy is friends with Chris Messina, Chad Fowler, and Why, and so on. This tends to validate the concept exemplified in Google's PageRank algorithm, of significant entities linking to each other.

6.3 Use Patterns

Finally, then, it would be useful to see whether Users came back, as they accrued dividends, to make additional reputation trades. This can be represented fairly easily as the difference in days between their first and last trades, as shown in Table 2, which demonstrates that while some Users do indeed only create an account and trade on Mnikr in one session, more than 30% of users returned on at least one additional day for additional trading using their earned dividends.

7. DISCUSSION

Our primary goal with the Mnikr system was to gain a machine-readable understanding of the traditionally unquantifiable concept of reputation, that would have a correlation with real-world reputations and be free of contextual bias. The results from the previous section show some level of promise for this idea. The base idea certainly seems to work, and its simple extensions—for instance, trading things such as corporations or software—seem to be possible; indeed, the Identity for Mnikr itself was traded in the top ten reputations, suggesting the possibility of trading people on an equal footing with multinational entities. The philosophical implications of such a system are out of scope for this work.

At the same time, there is ample opportunity for future research in this area, both within the limits of the current Mnikr functionality, and beyond them. First, it would be a useful addition to Mnikr's capabilities to have some sort of correlation for the Identity coalescence; rather than relying solely on the known-unreliable XFN data, it would be beneficial to have some sort of external data source to validate at least a part of the conclusions to which the algorithm comes. It is unknown at this time, however, if such a correl-

ative source exists, or to what degree such a resource could exist.

Several interesting questions are raised by the available trading data that could be answered by a longer-term analysis with more users. First, why is the growth of Identities, as shown in Figure 3, stopping at just over 350,000 Identities collected? One assumes that, given the existence of the closed-data Facebook social network with well over 100 million users, there would be at least an order of magnitude more Identities in open social networks; is there, then, a large-group clique among early adopters, or is there some other force underlying the lack of growth?

In addition, one wonders if given a broader user base, reputations would become vulnerable to the same base excitement and panic that can drive stock markets to sudden swings; for instance, if the "pump and dump" schemes that use spam communications to artificially raise the price of low-value stocks would be as effective for the trading of reputations as they are with corporations.

There are in addition some unexplored possibilities in the mechanisms of trading; for instance, as noted in §4.2, the amount of dividends, as well as the caps for individual influence upon a reputation's value, will likely have dramatic effects upon the reputation market. Similarly, using real, rather than artificial, money could have interesting effects upon users' willingness to buy reputations, as could introducing market friction, such as interest or penalties on points held in reserve, or transaction fees associated with buying and selling.

In a related vein, we did not explore user interface and design issues such as what types of instructions to users are most beneficial in terms of encouraging users to participate in the market. Mnikr has very little in terms of direct instruction to users on what to do to trade reputations. For these experiments we instead relied on users' natural curiosity to explore the system to give them an agenda, as well as the relative tech-savviness of the testers we were able to contact through our promotion methods.

One final area in which more exploration would be greatly beneficial is the concept of derivatives trading of reputations, such as index or mutual funds, that would allow Users to trade in subparts of shares of the reputations of many Identities simultaneously; for instance, one could have a fund of shares of the Time Magazine 100 Most Influential People, or other such known quantities. Additionally, derivatives trading opens other areas such as short selling, or selling futures of reputations—that is, selling based on the assumption that a reputation value will increase or decrease in the future.

Beyond the questions around trading and the reputation system itself, there is another issue worthy of examination: do the privacy implications of a system that autonomously collects, collates, and allows interaction with personally identifiable information, without a specific opt-in mechanism, outweigh the advantages of being able to interact with users through the system who have not yet signed up for it? The answer to this is relatively straightforward; Mnikr's algorithms do not combine personal information in any way that human observers, using—as Mnikr does—only public data, could not undertake. If users wish to keep, for instance, their personal and work-related activities on the Internet separate, they simply need not to link those Personas to each other; if there is no discovery path between them, Mnikr will not link those Personas in a single Identity. If a user wishes

entirely to be invisible to systems like Mnikr, they need only configure their privacy settings on the services they use to disallow public access– the same as they would to prevent, say, access by a search engine. Indeed, Mnikr provides a valuable service to people wishing to safeguard their online privacy, by demonstrating discoverable links between Personas that a user might not previously have noticed.

The ultimate question for this work is its intended utility– the question of why a person would choose to use a holistic reputation system, rather than using the pre-existing contextualized reputation systems for each of the person's chosen contexts; after all, one does not interact with people in every possible context, but only in specific contexts. There are, however, many contexts in which a system for communicating reputation does not exist; for instance, commenting in forums or on weblogs, there is often no system of reputation, and it can be hard to separate comments made by people with great ideas from comments of less value. Even when reputation systems do exist in a narrow context, they are often not portable; there is no way for a person with a high reputation at Slashdot to communicate this fact to another technology forum, for instance, even though they could well be considered to have the same context. The Mnikr project thus provides some idea of how to create a system of reputation usable not just in one context, but in any.

8. CONCLUSION

This work explores a new approach to creating machine-intelligible valuations for reputation through the application of a stock-trading metaphor to buying and selling shares of reputations. Using pseudonymous identification, derived from the publicly-available distributed information silos of Personas on Services across the public Internet, we create entities that represent a person's total presence across the Internet and allow others to value or devalue that person's contributions and existence as a whole.

The data we collected in just a month of exposure of the Mnikr system to the public Internet yields promising results, and indicates that this idea may hold some promise toward solving the challenge of creating holistic reputation measures in place of context-sensitive measures. More data is needed to make firm statements about the ultimate applicability and utility of such a system.

Moving forward, we expect to see additional work in clarifying the effects of variable choices in the dividend and coalescence algorithms, in developing extensions of the trading system to support different types of trades, and in relating the currency to such other systems as world monetary currencies.

9. ACKNOWLEDGEMENTS

10. REFERENCES

[1] M. Atkins, D. Recordon, C. Messina, M. Keller, and A. Steinberg. *Atom Activity Extensions*, April 2009.

[2] J. Blocher. Reputation as property in virtual economies. *Yale L.J. Pocket Part*, 118:120–125, January 2009.

[3] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.*, 30(1-7):107–117, 1998.

[4] T. Çelik, M. Mullenweg, and E. Meyer. *XFN 1.1 relationships meta data profile*, 2009.

[5] M. Chew, D. Balfanz, and B. Laurie. (Under)mining privacy in social networks. In *Web 2.0 Security and Privacy 2008*, May 2008.

[6] J. R. Douceur. The Sybil attack. In *IPTPS '01: Revised Papers from the First International Workshop on Peer-to-Peer Systems*, pages 251–260, London, UK, 2002. Springer-Verlag.

[7] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley, 1995.

[8] G. F. Gilder. *Telecosm: How infinite bandwidth will revolutionize our world*. Free Press, New York :, 2000.

[9] F. Labalme. Reputation capital and exchange mechanisms. Published Draft Working Paper, May 2002.

[10] W. Sherchan, S. W. Loke, and S. Krishnaswamy. A fuzzy model for reasoning about reputation in web services. In *SAC '06: Proceedings of the 2006 ACM symposium on Applied computing*, pages 1886–1892, New York, NY, USA, 2006. ACM.

[11] C. Stross. *Accelerando*. Ace Books, New York, 2005.

[12] W. T. L. Teacy, J. Patel, N. R. Jennings, and M. Luck. Coping with inaccurate reputation sources: Experimental analysis of a probabilistic trust model. In *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous Agents and Multiagent Systems*, pages 997–1004, New York, NY, USA, 2005. ACM.

[13] P. J. Windley, D. Daley, B. Cutler, and K. Tew. Using reputation to augment explicit authorization. In *DIM '07: Proceedings of the 2007 ACM workshop on Digital Identity Management*, pages 72–81, New York, NY, USA, 2007. ACM.

[14] X. Yan and B. Van Roy. Reputation markets. In *NetEcon '08: Proceedings of the 3rd international workshop on Economics of Networked Systems*, pages 79–84, New York, NY, USA, 2008. ACM.