# Facebook: Threats to Privacy

Harvey Jones, José Hiram Soltren

December 14, 2005

**Abstract**

End-users share a wide variety of information on Facebook, but a discussion of the privacy implications of doing so has yet to emerge. We examined how Facebook affects privacy, and found serious flaws in the system. Privacy on Facebook is undermined by three principal factors: users disclose too much, Facebook does not take adequate steps to protect user privacy, and third parties are actively seeking out end-user information using Facebook. We based our end-user findings on a survey of MIT students and statistical analysis of Facebook data from MIT, Harvard, NYU, and the University of Oklahoma. We analyzed the Facebook system in terms of Fair Information Practices as recommended by the Federal Trade Commission. In light of the information available and the system that protects it, we used a threat model to analyze specific privacy risks. Specifically, university administrators are using Facebook for disciplinary purposes, firms are using it for marketing purposes, and intruders are exploiting security holes. For each threat, we analyze the efficacy of the current protection, and where solutions are inadequate, we make recommendations on how to address the issue.

# Contents

# 1  Introduction

Facebook[1] (www.facebook.com) is one of the foremost social networking websites, with over 8 million users spanning 2,000 college campuses. [4] With this much detailed information arranged uniformly and aggregated into one place, there are bound to be risks to privacy. University administrators or police officers may search the site for evidence of students breaking their school's regulations. Users may submit their data without being aware that it may be shared with advertisers. Third parties may build a database of Facebook data to sell. Intruders may steal passwords, or entire databases, from Facebook. We undertook several steps to investigate these privacy risks. Our goal was to first analyze the extent of disclosure of data, then to analyze the steps that the system took to protect that data. Finally, we conducted a "threat model" analysis to investigate ways in which these factors could produce unwanted disclosure of private data. Our analysis found that Facebook was firmly entrenched in college students' lives, but users had not restricted who had access to this portion of their life. We discovered questionable information practices with Facebook, and found that third parties were actively seeking out information.

To analyze the extent of user disclosure, we constructed a spider that "crawls" and indexes Facebook, attempting to download every single profile at a given school. Using this tool, we indexed the entire Facebook accessible to a typical user at Massachusetts Institute of Technology (MIT), Harvard, New York University (NYU), and the University of Oklahoma. To supplement this data, we surveyed the MIT student body to ascertain the level of use of certain Facebook features. Our study found that upwards of 80% of matriculating freshmen join Facebook before even arriving for Orientation, and that these users share significant amounts of personal information. We also found that Facebook's privacy measures are not utilized by the majority of college students. To analyze the Facebook system we investigated the facets of the website, and of the terms of use and compared them against the current standards of "Fair Information Practices" as defined by the Federal Trade Commission, as well as the standards set by competing sites. Although many Facebook features empower users to control their private information, there are still significant shortcomings. Finally, we took the perspective of a third party acting in a self-interested manner, looking either for financial gain or for assistance in the enforcement of university policy. We surveyed news articles on the consequences of Facebook information disclosure, and interviewed students that harvested data, as well as students who were punished for disclosing too much. Given the existing threats to security, we constructed a threat model that attempted to address all possible categories of privacy failures. From a systems perspective, there are a number of changes that can be made, both to give the user a reasonable perception of the level of privacy protection available, and to protect against disclosure to intruders. For each threat, we make recommendations for Facebook, its

---

[1] "Facebook", as opposed to "the Facebook", is how the site's literature refers to itself. We adopt that terminology throughout the paper.

users, and college administrators. These include eliminating the consecutive profile IDs, using SSL for login, extending "My Privacy" to cover photos, and educating end-users about privacy concerns.

# 2 Background

## 2.1 Social Networking and Facebook

Users share a variety of information about themselves on their Facebook profiles, including photos, contact information, and tastes in movies and books. They list their "friends", including friends at other schools. Users can also specify what courses they are taking and join a variety of "groups" of people with similar interests ("Red Sox Nation", "Northern California"). The site is often used to obtain contact information, to match names to faces, and to browse for entertainment. [4]

Facebook was founded in 2004 by Mark Zuckerburg, then a Harvard undergraduate. The site is unique among social networking sites in that it is focused around universities – "Facebook" is actually a collection of sites, each focused on one of 2,000 individual colleges. Users need an @college.edu email address to sign up for a particular college's account, and their privileges on the site are largely limited to browsing the profiles of students of that college.

Over the last two years, Facebook has become fixture at campuses nationwide, and Facebook evolved from a hobby to a full-time job for Zuckerburg and his friends. In May 2005, Facebook received $13 million dollars in venture funding. Facebook sells targeted advertising to users of its site, and parters with firms such as Apple and JetBlue to assist in marketing their products to college students. [14]

## 2.2 Information that Facebook stores

**First-party information** All data fields on Facebook may be left blank, aside from name, e-mail address, and user status (one of: Alumnus/Alumna, Faculty, Grad Student, Staff, Student, and Summer Student). A minimal Facebook profile will only tell a user's name, date of joining, school, status, and e-mail address. Any information posted beyond these basic fields is posted by the will of the end user. Although the required amount of information for a Facebook account is minimal, the total amount of information a user can post is quite large. User-configurable setting on Facebook can be divided into eight basic categories: profile, friends, photos, groups, events, messages, account settings, and privacy settings. For the purposes of this paper, we will investigate profiles, friends, and privacy settings.

Profile information is divided into six basic categories: Basic, Contact Info, Personal, Professional, Courses, and Picture. All six of these categories allow a user to post personally identifiable information to the service. Users can enter information about their home towns, their current residences and other contact information, personal interests, job information, and a descriptive pho-

tograph. We will investigate the amount and kind of information a typical user at a given school is able to see, and look for trends. A major goal of Facebook is to allow users to interact with each other online. Users define each other as friends through the service, creating a visible connection.

| My Profile | Contains "Account Info", "Basic Info", "Contact Info" "Personal Info", "My Groups", and a list of friends |
| --- | --- |
| The Wall | Allows other users to post notes in a space on one's profile |
| My Photos | Allows users to upload photographs and label who is in each one. If a friend lists me as being in a photograph, there is a link added from my profile to that photograph |
| My Groups | Users can form groups with other like-minded users to show support for a cause, use the available message boards, or find people with similar interests. |

Table 1: Facebook Features

**Third-party information**    Two current features of Facebook have to do with third parties associating information with a user's profile. The "Wall" allows other users a bulletin board of sorts on a user's profile page. Other users can leave notes, birthday wishes, and personal messages. The "My Photos" service allows users to upload, store and view photos. Users can append metadata to the photographs that allows other users to see who is in the photographs, and where in the photograph they are located. These tags can be cross-linked to user profiles, and searched from a search dialog. The only recourse a user has against an unwelcome Facebook photo posted by someone else, aside from asking them to remove it, is to manually remove the metadata tag of their name, individually, from each photograph. Users may disable others' access to their Wall, but not to the Photos feature.

**"My Privacy"**    Facebook's privacy features give users a good deal of flexibility in who is allowed to see their information. By default, all other users at a user's school are allowed to see any information a user posts to the service. The privacy settings page allows a user to specify who can see them in searches, who can see their profile, who can see their contact info, and which fields other users can see. In addition, the privacy settings page allows users to block specific people from seeing their profile. As per the usage agreement, a user can request Facebook to not share information with third parties, though the method of specifying this is not located on the privacy settings page.

## 3    Previous Work

No previous academic work specific to Facebook was found on the Lexis databases, Google's database for scholarly papers, the Social Science Research Network, or for "facebook AND journal AND arti-

| Visibility to Search? | **Everyone** |
|---|---|
| | Restricted |
| Profile Visibility | **Everyone at school** |
| | Friends of friends at school |
| | Just friends |
| Contact Info Visibility | **Everyone at school** |
| | Friends of friends at school |
| | Just friends |
| Profile also shows... | **My friends** |
| | **My last login** |
| | **My upcoming events** |
| | **My courses** |
| | **My wall** |
| | **Groups that a lot of my friends are in** |

Table 2: "My Privacy" settings (defaults in **bold**)

cle" and numerous other terms in a general web query. Although no journal articles exist, there are many news articles that have been published about the emergence of Facebook, its incorporation and subsequent venture funding, and recently, the consequences of third parties discovering information that users have made public[14][20][21]. In related fields, the Federal Trade Commission has done research into the area of online privacy practices, and has published several reports on the matter, including the 1998 report to Congress entitled "Privacy Online." [6] Previous work in social networking has included a thorough investigation of "Club Nexus", a site similar to Facebook located at Stanford University[1].

## 4   Principles and Methods of Research

In order to investigate the ways in which Facebook is used, we closely investigated the usage patterns of Facebook. We employ two methods of data collection to learn more about the way users interact with Facebook. First, we conducted a survey of MIT students on the use of Facebook's features. Second, we harvested data from the Facebook site directly.

### 4.1   Usage patterns of interest.

Our main objective in gathering and analyzing Facebook user data was to make statements and generalizations regarding the way users use their Facebook accounts. We investigated when users create their accounts, and which kinds of users create accounts. Though the friending service is of

Figure 1: A sample Facebook page. Note the layout, accessible fields, and formation of URL used to retrieve this page.

great interest to social network research, for the purposes of our paper, we primarily investigated the number of friends users have on the service as an indicator of use, and look for trends.

## 4.2 User surveys

Our direct user data collection procedure employed both paper surveys and Web based forms to ask individual users questions concerning their Facebook practices.

In designing our survey, we aimed for a minimum number of straightforward, multiple choice questions which would serve to reveal usage patterns, familiarity with various aspects of the service, and opinions on the quality of the service. The questions asked about the subject's gender, residence, and status, their date of joining Facebook and utilization thereof. It also asked about their knowledge of Facebook's Terms of Service, Privacy Policy, and privacy features, as well as their familiarity with Facebook's practices. We designed the survey such that it would fit on one printed page, and take approximately three minutes to complete. The complete text of our survey is included as an appendix.

In order to diversify the survey results, we gathered data through four routes. We set up a table in the MIT Student Center, offering students a chocolate-based incentive for completing surveys. We asked classmates in Public Policy, MIT course 17.30J/11.002J, to complete the survey. Via e-mail, we asked the residents of the East Campus, Burton-Conner, Simmons Hall, and Random Hall dormitories to complete the surveys. Finally, we asked all survey takers to notify others of the survey.

## 4.3 Direct data collection

Our collection of data directly from Facebook served two principles. It served as a proof of concept, to demonstrate that it is possible for an individual to automatically gather large amounts of data from Facebook. The collection of data was not entirely trivial, but we were able to produce the scripts necessary to do so within 48 hours. Also, the collection of data from Facebook will provide us with a large, nearly exhaustive and statistically significant data set, from which we can draw valuable conclusions on usage trends.

## 4.4 Obscuring personal data

Before analyzing data, we aggregated it into a spreadsheet. When we considered sets of more than one record, we obscured data we deemed to be personally identifiable – Name, Phone Number, AOL Screenname, High School, and Dormitory. These fields were unchanged if left blank by the user, and replaced by "OBSCURED"[2].

---

[2]Before we developed the software to obscure the data, we did do enough analysis to discover that 48 Facebook users at the schools we studied have the phone number 867-5309

## 4.5   A brief technical description of Facebook from a user perspective

Facebook uses server-side Hypertext Preprocesser (PHP) scripts and applications to host and format the content available on the service. Content is stored centrally on Facebook servers. Scripts and applications at Facebook get, process, and filter information on demand, and deliver it to users in real time, to a Web browser over the Internet. Users begin their Facebook session at the service's top level site, http://www.facebook.com/.

At the main Facebook page, a user can log in to the service, or browse the small amount of information available to the general public. The main page of the service is spartan, and does not provide any personally identifiable information or technical insight. Facebook does require a school e-mail address to use their service.

To log in to Facebook, users enter their username and password into the appropriate fields on the page, and click Login. This sends a special URL to the service:

$$http: //www.facebook.com/login.php?email = USERNAME@SCHOOL.edu\&pass = PASSWORD$$

$$(1)$$

Note that this URL contains a user's login credentials in clear text. This information is vulnerable to detection by a third party. No secure socket layer (SSL) or other encryption is used in logging in tot he service.

During the login process, the service provides the user's web browser with some information, which is stored in the form of a cookie. Some of this information, such as the user's e-mail address, is written to a file so the user does not have to enter his or her e-mail at the next login. Facebook's service creates and gives a user a unique checksum at every login, which the browser stores as a session cookie and generally does not write to a file. This checksum varies from login to login, but other parameters do not.

Once logged in to the service, a user is free to interact with Facebook. The user may edit their profile, look at others' profiles, add or change their friends lost or personally identifiable information, and explore the service.

The majority of features on Facebook are requested via simple, human-readable URLs. For example, profile URLs are retrieved by requesting a URL of the form:

$$http: //SCHOOL.facebook.com/profile.php?id = USERID \qquad (2)$$

Facebook will read the school and user ID, and give the user either the requested user's profile page, filtered for privacy by the user's request before being delivered, or return the user's home page if the profile he requested is blocked or does not exist. The first user at every school is called "The Creator." This profile's USERID is the lowest userid at any given school. The date of its creation is the date which Facebook was opened to that school. User Ids continue to be assigned sequentially from the first valid number, created at the time of creation of each new account.

Facebook's human-readable URLs and regularly formatted HTML make automated acquisition, parsing, and analysis relatively easy. We discuss how we and others have done this in the next section.

Each separate school has its own Facebook "server" for its content. Users with a school e-mail address @SCHOOL.edu will go through http://SCHOOL.facebook.com/. For the most part, many of these "servers" redirect to the same machine. For example, harvard.facebook.com, mit.facebook.com, nyu.facebook.com, and ou.facebook.com all redirect to 204.15.20.25. This architecture allows Facebook to easily move different schools to different servers if necessary.

By default, a new user's profile and all information are fully visible to all other users at the same school, but not visible to anyone at another school. Many users do not change their default settings, making their information accessible.

When a user logs out of Facebook or closes their web browser, the session cookies are lost. This generally means that once a user exits the service, they must enter at least their password to use the service again.

### 4.5.1 Data acquisition

We are not the first to download user profiles from Facebook in large numbers. In the past, others have utilized Facebook's use of predictable, easy to understand URLs to automatically request information and save user information for further analysis. Our approach used the incremental profile identifier to download information in large quantities.

The algorithm we used to gather this data is very straightforward:

1. Log in to Facebook and save session cookies.

2. Load your home page and note the USERID of the page.

3. Decrease the USERID until you find the ID of "The Creator," the first profile at a given school. Save this number as USERID-LOW.

4. Increase the USERID until you find the ID of a user who joined recently, i.e. within the past day. Save this number as USERID-HIGH.

5. For every profile from USERID-LOW to USERID-HIGH at a given school SCHOOL: Get the profile, using URL

$$http://SCHOOL.facebook.com/profile.php?id = USERID \qquad (3)$$

, and save the profile as a file.

To implement our algorithm, we used wget, "the non-interactive network downloader." In addition to implementing the above algorithm, we made wget pretend to be another web browser

by changing its user agent (to avoid potential suspicion at using wget to log in to Facebook). We also had wget randomly insert a delay between requests, to keep load off of Facebook's servers and make our requests less difficult to detect. We took advantage of the fact that logins and passwords are not encrypted, and can be sent as part of the login URL as an email and password pair.

The final application we used to download profiles was a short (five line!) BASH shell script, which we include in the appendix.

We ran this script four times: once for Harvard, MIT, the University of Oklahoma (OU), and New York University (NYU).

## 4.6   Statistical significance

**Survey data**   Over the course of the two weeks we ran the survey, 419 MIT students responded to the questions asked. The users answering our profile questions came from all of campus, with strong concentrations in dorms where we e-mailed the survey. The respondents were mostly undergraduates (90%). There were 224 female respondents and 195 male respondents. Reflecting an MIT student population of 4,000 undergraduates and 6,000 graduate students, we can find the statistical significance of our findings using the results of confidence levels and confidence intervals from statistics.

The sample size of a survey group is related to the confidence value, the percentage picking a choice, and the confidence interval by the formula

$$S = \frac{Z^2 p(1-p)}{c^2} \tag{4}$$

Where $S$ is our sample size, $Z$ is a value proportional to the confidence level (1.96 for a 95% confidence interval), $p$ is the percentage picking a choice, expressed as a decimal (with a worst case value of 0.5), and $c$ is the confidence interval, expressed as a decimal (i.e. $0.04 \pm 0.04$). For small populations, we use the correction

$$S' = \frac{S}{1 + \frac{S-1}{P}} \tag{5}$$

Where $S$ is our original sample size, $S'$ is our new sample size, and $P$ is our sample population. [17]

Our survey results are good enough to make coarse extrapolations to the MIT community in general. At a confidence level of 95%, and a sample size of 419 applying to an MIT student population of 10,000 total undergraduates and graduate students, and a worst case answer uncertainty of 50%, we find our confidence interval to be 4.68%. In other words, we can be 95% certain that our survey responses fall within 4.68% of the true values. At a confidence level of 99%, our uncertainty increases to 6.17%.

**Collected Facebook data**   In general, we were able to collect large numbers of user profiles from Facebook using our information collection system. We exhaustively downloaded every profile

available at our four subject schools, so there is no sampling uncertainty, as long as we limit our conclusions to generalizations about the population of students with accessible Facebook profiles. We will attempt to statistically correlate certain variables to prove hypotheses, and at other points we will show raw data when we want to indicate a trend. The following table summarizes our success in downloading information.

Success Rates In Downloading Profiles

| School | Number Profiles | Number Downloaded | Percentage |
|---|---|---|---|
| MIT | 10063 | 8021 | 79.71% |
| Harvard | 25759 | 17704 | 66.16% |
| Oklahoma U. | 28201 | 24695 | 70.54% |
| NYU | 32250 | 24695 | 77.41% |
| Total | 97273 | 70311 | 72.28% |

**Aggregate Statistics**   We established a "disclosure score" to quantitatively rank the amount of PII disclosed by different colleges, classes, and genders. The overall score is the sum of the percentage disclosure of (Gender, Major, Dorm, High School, AIM Screenname, Mobile Phone, Interests, Clubs, Music, Movies, and Books). From there, we created two sub-scores, one to reflect contact information that could conceivably be used to contact or locate users (Dorm, AIM Screenname, Mobile Phone, and Clubs/Jobs), as well as a sub-score reflecting disclosure of user interests (Interests, Clubs/Jobs, Music, Movies, and Books).

## 5   End-Users' Interaction with Facebook

### 5.1   Major trends

After processing the results of our user survey and downloaded Facebook profiles, we found some general trends in Facebook usage. Facebook is ubiquitous at the schools where it has been established. Users put real time and effort into their profiles. Students tend to join as soon as possible, often before arriving on campus. Users share lots of information but do not guard it. Users give imperfect explicit consent to the distribution and sharing of their information. Privacy concerns differ across genders.

In the following pages, we analyze the collected data along numerous lines, and statistically justify our findings. Our full numerical findings are included in the appendix.
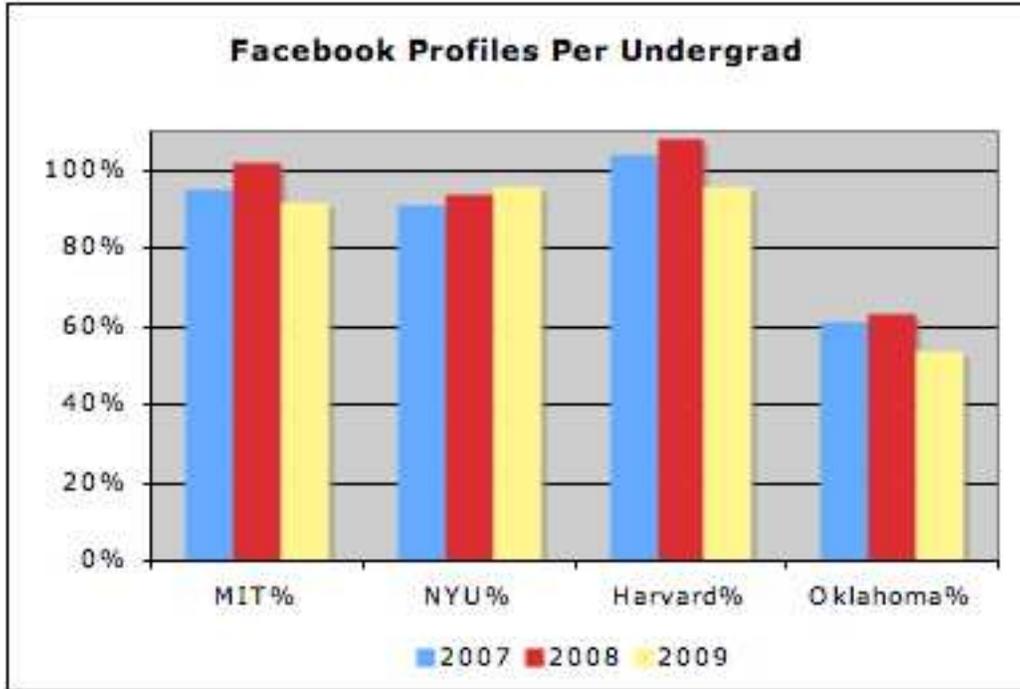
Figure 2: Number of Profiles identifying as a class divided by students in that class

## 5.2 Facebook is ubiquitous

**Possession of a Facebook account**  Survey results indicated that large majority of MIT students have Facebook profiles. Of 413 respondents, 374 (91%) claimed to have Facebook accounts, while only 39 (9%) did not. Indexing the Facebook seemed to indicate a similar result; the vast majority of undergraduates have Facebook accounts. Although fake accounts could bloat the number of accounts, the fact that the Facebook user base is quite similar to the MIT undergraduate population point to the fact that a large percentage of Facebook users are genuine. There are 948, 1016, and 921 accounts that provide the class years of 2007, 2008, and 2009, respectively, compared to a class size of roughly 1,000. As shown below, the majority of Facebook accounts are updated at least monthly, which fits the profile of large numbers of users updating information about themselves. Aside from her romantic attachments perhaps, a Paris Hilton account[3] would not need to be constantly updated. At NYU, where potential pranksters are limited to two e-mail addresses[18], the number of accounts for the classes of 2007-2009 (3850, 4012, 4076) correspond closely to the class sizes of 4,250. [16]

---

[3]Until recently, the Facebook FAQ warned against creating fake accounts, telling users that "Everyone knows that you're not Paris Hilton"

14

| Month | Three Months | Six Months | One Year |
|-------|--------------|------------|----------|
| 53%   | 82%          | 92%        | 98%      |

Figure 3: Virtually all users update profiles often

## 5.3   Users put time and effort into profiles

The vast majority of users update their accounts frequently, with over half updating in November 2005[4]. This indicates that not only do the majority of undergraduates have Facebook accounts, the majority of them also keep them constantly updated.

## 5.4   Students join Facebook before arriving on campus

We looked at the distributions of profile creating dates of members of the classes of 2008, and 2009. The class of 2008 enrolled at MIT admission and had access to MIT server by May of 2004, whereas the class of 2009, the current freshman class, had MIT server accounts by May of 2005[5]. Note that MIT admits classes of approximately 1,000 freshmen.

Members of the MIT class of 2008 tended to create their profiles as soon as they heard about Facebook, which was generally over the summer or during orientation. The majority of the class of 2008 joined Facebook from June 2004 to August 2004. In this time, 699 members of the class of 2008 created their profiles. Approximately 100 created their profiles in May of 2004 (i.e. as soon as they could), and the remainder created their profiles at later times, dropping to approximately 10 per month. We were able to access 1016 members of the class of 2008 with Facebook profiles[6].

The class of 2009 had an even more pronounced spike at matriculation time, indicating the extraordinary draw of the Facebook. During May and June of 2005, 538 members of the class of 2009 created Facebook accounts. At present, 921 members of the class of 2009 have unrestricted Facebook accounts.

At other schools, users exhibit similar behavior in creating their Facebook profiles. Strikingly, over 948 (roughly 60%) Harvard Class of 2009 freshmen created their accounts within a month of getting their email address. Freshmen create their accounts as soon as they can. The Harvard trends are even more pronounced as we can see from the graph, with most 2008 freshmen signing up

---

[4]19% of Harvard profiles, 15% of MIT Facebook profiles, 10% of NYU profiles, and 6% of Oklahoma profiles do not have an update timestamp. Because no update timestamps exist before June 2004, it is probable that the feature was implemented at that point, and all unstamped profiles were last updated before that point. This hypothesis is substantiated by the fact that the number of blank update fields at a school is proportional to the length of time before June 2004 Facebook was available at that school. Given the exponential tail-off of the last update times, it is also likely that this 15% compose users who signed up right at the launch of Facebook for their school and did not update their accounts afterwards.

[5]Our experience is that MIT sends out MIT Server coupons around this time

[6]Note that these numbers may be skewed by accounts for fictional people or celebrities.
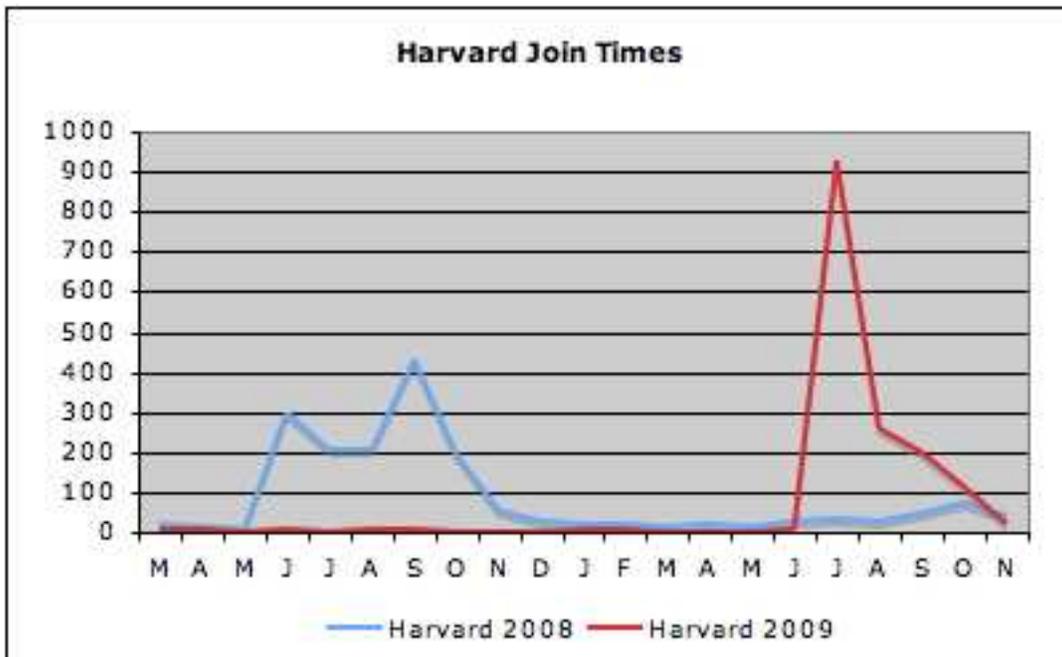
Figure 4: Freshmen create accounts sooner and sooner after matriculation

over a three-month period, while the class of 2009 obtained their Facebook accounts immediately.

## 5.5 A substantial proportion of students share identifiable information

Facebook users at MIT tend to give a large amount of personal information, and tend not to restrict access to it. Furthermore, Facebook users are more wary of some kinds of personal information than others. Users were most willing to indicate their high school, and became increasingly protective of their information regarding residence hall, interests, screen name, music interests, favorite movies, favorite books, clubs and jobs, and mobile telephone number.

## 5.6 The most active users disclose the most

Users who frequently update their profiles tend to be even more open. Of the 5279 MIT profiles updated on or after September 1, 2005, we found that, although the general trends of relative disclosure did not change, the relative willingness to disclose all information increased.

Using another heuristic for determining active users, users with lots of friends tend to be much more forthcoming with their personal information, particularly that which might be valuable to advertisers.

Facebook has grown extremely rapidly, establishing a user base of 8,000,000 users, and close to 100% penetration at certain schools. If Facebook continues to grow in popularity, the average user will likely become more and more like the "well-connected" user. If this trend continues, the level
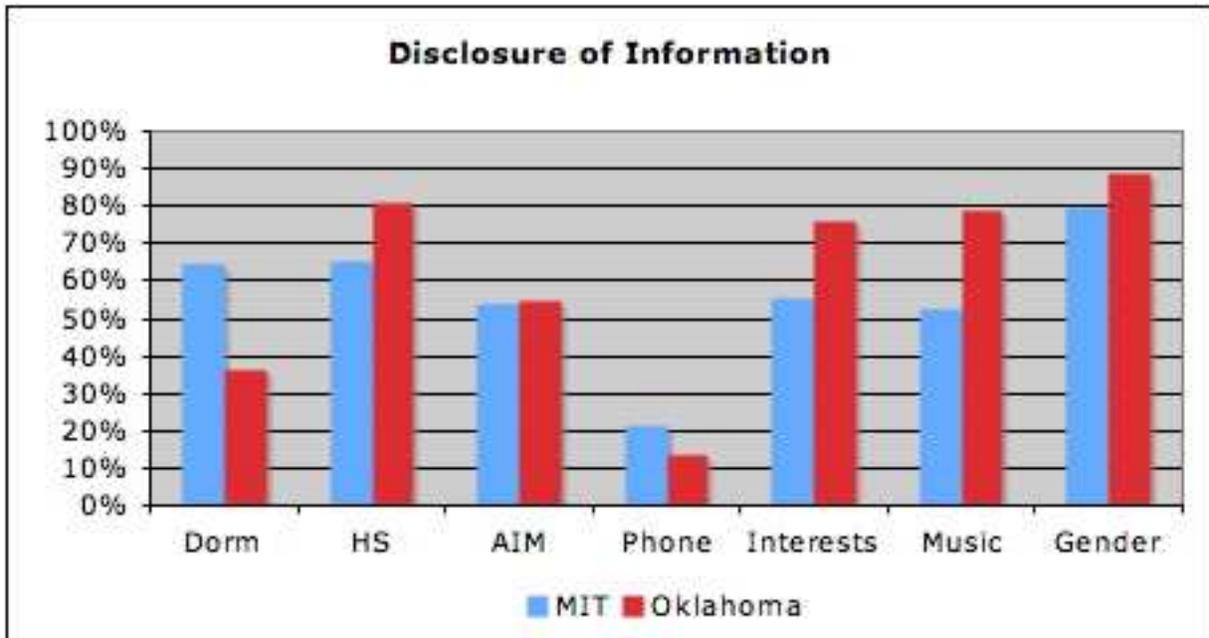
Figure 5: Users disclose personally identifiable information
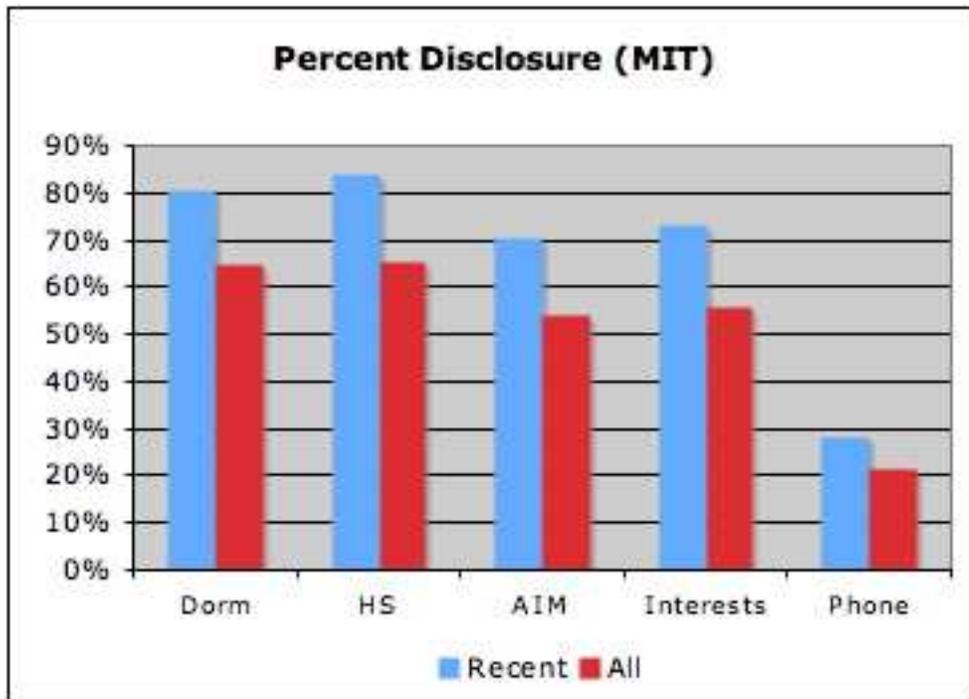


Figure 6: Recent users disclose even more

All Schools: Disclosure of PII

| | Clubs | Interests | Movies | Music | Books | Gender | Mobile |
|---|---|---|---|---|---|---|---|
| 300+ Friends | 81.0% | 85.3% | 81.7% | 82.9% | 76.6% | 92.8% | 25.6% |
| All Users | 51.5% | 64.1% | 62.7% | 64.0% | 59.1% | 82.8% | 17.1% |
| Difference | 29.4% | 21.2% | 19.0% | 18.9% | 17.4% | 10.1% | 8.5% |

Figure 7: Connected users disclose more personal information, especially commercially valuable information

of information disclosure will keep increasing correspondingly.

## 5.7 Undergraduates share the most, and classes keep sharing more

As shown in the table below, undergraduates share much more data than average, in almost every case. As the majority of new registrants for Facebook each year are going to be undergraduates, and the undergraduates most likely to disclose information no less, this is another indication that more and more data will become available on Facebook.

**Difference between classes**   In order to determine if there is a statistically significant difference between courses, we attempted to correlate disclosure scores to class years. We ran a regression of number of years in attendance at the college[7] against the disclosure index, and the contact and interest subscores. We did this at all four schools, and the result was that all disclosure scores were weakly correlated to class year ($r = -.496$ for the overall score, $r = -.151$ for the contact score, and $-.187$ for the interest score.). This means that there is a correlation between being in a younger class and disclosing more information.

## 5.8 Differences among universities

Among the four universities we investigated, we found subtle differences in the way student interact with Facebook. Of the universities, Harvard provided us with the lowest percentage of visible profiles from existing profiles (66%), whereas MIT provided the highest (79%). Students at the University of Oklahoma were much less likely to share contact information (such as residence, screen name, and mobile phone number) than students from any other university in our study. On the other hand, students at Oklahoma were the most forthcoming about their tastes in books, movies, and music.

The differences we found really speak to the notion that Facebook is different at every school it supports. The differences we noted are probably a function of many variables specific to the school, such as the social atmospheres at the school, policies on information sharing, administrative advice on Facebook usage, and so on. Such topics are outside the scope of this paper.

---

[7]0, 1, 2 for the Classes of 2009, 2008, and 2007, respectively.

Difference in Disclosure

|  | Harvard | MIT |
|---|---|---|
| Gender | 22% | 17% |
| Major | -6% | 23% |
| Dorm | 30% | 23% |
| Room? | 23% | 4% |
| High School | 32% | 18% |
| AIM | 26% | 18% |
| Mobile | 3% | 10% |
| Interests | 29% | 16% |
| Clubs/Jobs | 17% | 23% |
| Music | 33% | 18% |
| Movies | 31% | 19% |
| Books | 31% | 17% |

Figure 8: Difference between Class of 2009 exposure and all users

|  | MIT | Harvard | OK | NYU |
|---|---|---|---|---|
| Major | 81% | 64% | 91% | 79% |
| Dorm | 96% | 94% | 85% | 89% |
| AIM | 71% | 72% | 62% | 76% |
| Mobile | 24% | 27% | 17% | 15% |
| Interests | 78% | 81% | 89% | 81% |
| Clubs/Jobs | 49% | 58% | 76% | 50% |
| Music | 77% | 82% | 93% | 84% |
| Movies | 74% | 80% | 90% | 82% |
| Books | 74% | 80% | 81% | 77% |

Figure 9: Disclosure rates of the Class of 2009

## 5.9   Even more students share commercially valuable information

The information most relevant to advertisers would likely be demographic data (age, gender, location), as paired with interests. In general, over 70% of users are willing to disclose both categories of information, making the Facebook a valuable trove of demographic data for marketers. In addition, this database of interests could easily be cross-referenced by a database from a third-party vendor, matching the details about users' interests and current location to addresses, phone numbers, and social security numbers. As shown above, dedicated users have a tendency to disclose this information much more often, which may be a leading indicator of even greater disclosure.

## 5.10   Users are not guarded about who sees their information

**Knowledge and use of "My Privacy" feature**   As a whole, users are familiar with the privacy features Facebook offers, and choose not to use them. Of 389 users indicating familiarity with "My Privacy", 289 (74%) say they are familiar with "My Privacy," while 100 (26%) say they are not. At the same time, of the 380 users who gave information regarding their use of "My Privacy," 234 (62%) said they use the feature, while 146 (38%) said they do not. Actively choosing to not use "My Privacy" indicates that users believe there is a benefit to providing information and allowing others to see it.

**Concerns about Facebook privacy**   As a whole, survey respondents expressly indicated low concern for Facebook's privacy policies. Of 329 respondents, 76 (23%) are not concerned with Facebook privacy, 117 (35.5%) are barely concerned, 104 (31.6%) are somewhat concerned, 20 (6.1%) are quite concerned, and 12 (3.6%) are very concerned.

**Likelihood of "friending" strangers.**   Facebook users at MIT tend to friend people they know, doing so almost exclusively. Of the 383 respondents to this question, 243 people (63.45%) never friend strangers, 110 people (28.72%) friend strangers on occasion, and 30 (7.83%) claim to friend strangers. Although this seems like an intuitive notion, it merits further attention. Only allowing people whom users know in real life to access their information is a good Facebook security strategy when combined with other privacy features and selective posting. This tendency of users is further evidence that Facebook use is more characteristic of physical relationships than that of an exclusively online community, a powerful metaphor that is at the heart of the way users share their information on Facebook. Women and men are equally unlikely to add a stranger to their list of friends.

## 5.11   Users Are Not Fully Informed About Privacy

**Familiarity with the TOS and the Privacy Policy**   We asked Facebook users if they had read Facebook's policies regarding their use of the service. Of 389 respondents, 353 (91%) had not read

the Terms of Service. Of 390 respondents, 347 (89%) had never read the Privacy Policy.

**Understanding of Privacy Policy**   We asked users to guess whether or not Facebook can share your information with other companies. Of 374 respondents, 174 (47%) believed Facebook could not do this, while 200 (53%) believed Facebook could. Facebook can indeed share your information with other companies for advertising or other purposes, as indicated in their privacy policy[8].

## 5.12   As Facebook Expands, More Risks Are Presented

**Familiarity with "My Photos" feature**   The overwhelming majority of Facebook users are familiar with the "My Photo" feature. Of 389 respondents, some 342 (87.9%) were familiar with the feature. Furthermore, although most users are familiar with the feature, few seem to worry about its potential implications. When asked if users have any control over the "My Photo" content of others, specifically, on restricting access to photos posted on the service, 196 users of 416 respondents (47%) said yes, 139 users (33%) said no, and some 84 (20%) did not know, or did not provide an answer.

## 5.13   Women self-censor their data

In addition to the above analysis, we compared the trends of male and female users. Women are more likely to log into Facebook, have more friends, and have a higher percentage of friends from MIT. Both genders are equally unfamiliar with Facebook's Terms of Service and Privacy Policy. Women were more likely to use Facebook's "My Privacy" feature in our survey, but not to a statistically significant level. Women definitely self-censor their Facebook data more than men do. This is pronounced in the number of mobile phone numbers made available to the public, as shown in the table[9].

In addition, we calculated the correlation between self-reported gender percentages at the different universities, and correlated these to the contact information index. We found that schools with more women share proportionately less contact information, with a correlation coefficient r = -.462.

---

[8]The FAQ and Privacy Policy are actually in direct contradiction on this point. The FAQ states that "We don't distribute your user information to third parties." The Privacy Policy, on the other hand, states that "we may share your information with third parties, including responsible companies with which we have a relationship." The Facebook then lists reasons that they may share information, including legal requests and "facilitating their business." Although the policy could be construed to imply they will not share information, it is certainly not clearly stated, and a strict reading would imply that Facebook can share information with third parties.

[9]The correlation coefficient of male to female mobile phone disclosure is .992, indicating an extremely strong link between the behavior of the genders at any particular school.

Disclosure of phone number, by gender

|  | Male | Female |
|---|---|---|
| Harvard | 33% | 26.5% |
| MIT | 29.7% | 20.5% |
| NYU | 22.2% | 11.6% |
| Oklahoma | 21% | 8% |

Figure 10: Women self-censor the information they share

### 5.14 Men talk less about themselves

In contrast, we compared gender ratios to the interest data index (the extent to which users share their interests, clubs, and favorite books, etc.). Here we found that the male-dominated schools tended to share less information, which may indicate that women are more likely to share information about themselves which will not lead to phone calls or unwanted visits. The correlation coefficient between self-reported female percentage and the interest index was r=.625.

### 5.15 General Conclusions

Facebook is an institution at the colleges we surveyed. As time goes on, it is becoming even more entrenched in college life. Although they tend to self-censor, especially women, users still share a lot of personal information that could be valuable to many parties. As Facebook becomes more entrenched, disclosure rates are likely to rise, until Facebook changes the parameters of their system, or there are enough newsworthy privacy stories to change users' perceptions.

## 6 Facebook and "Fair Information Practices"

### 6.1 Overview

In 1998, the Federal Trade Commission published Privacy Online, a report to Congress assessing the state of privacy on the Internet. This report identified the five "widely accepted fair information practices": Notice, Choice, Access, Security, and Redress. These areas cover the basic principles of online privacy, areas Facebook needs to address if they are to protect the privacy of its users. [6]

### 6.2 Notice

Notice is the first and most important requirement of fair information practices. Customers must be aware of information collection and their rights regarding that collection before they can exercise them. The basic "notice" requirements are a clear statement given to the consumer, before data is collected, including, among other things:

- Identification of the entity collecting the data, the uses to which the data will be put, and any potential recipients of the data.

- The nature of the data collected and the means by which it is collected if not obvious (passively, by means of electronic monitoring, or actively, by asking the consumer to provide the information).

- Whether the provision of the requested data is voluntary or required, and the consequences of a refusal to provide the requested information.

- The steps taken by the data collector to ensure the confidentiality, integrity and quality of the data. [6]

The Facebook Privacy Policy aims to fulfill this requirement. It specifies Facebook as the entity collecting the data, and does a good job of identifying which data will be collected in most cases, including non-obvious data such as session data and IP addresses. Parts of the policy are vague, however, and some are seemingly contradictory and confusing, such as "Facebook also collects information about you from other sources, such as newspapers and instant messaging services. This information is gathered regardless of your use of the Web Site. We use the information about you that we have collected from other sources to supplement your profile unless you specify in your privacy settings that you do not want this to be done." This passage is either inaccurate or outdated, as no setting related to this information is available in the "My Privacy" feature.

Even though Facebook accurately addresses what information they will be including on the whole, their Privacy Policy falls short in other areas. The identification of the uses to which the data will be put are nonexistent, and the identification of the targets of potential disclosure is anybody Facebook deems appropriate, including marketing partners. Facebook has close relationships with several corporations, integrating their marketing efforts seamlessly into the site via giving them special "Groups" for interested students. This disclosure is certainly legal, and users are receiving the use of an extremely useful and popular site for free in exchange for it. Unfortunately, not all users understand the terms of the bargain; our survey showed that 46% of Facebook users believed that Facebook could not share their information with third parties.

## 6.3 Choice

"At its simplest, choice means giving consumers options as to how any personal information collected from them may be used. Specifically, choice relates to secondary uses of information – i.e., uses beyond those necessary to complete the contemplated transaction." [6]

Clearly, it is necessary to enter some personal information if one wishes to participate in a social networking website. However, there is large amounts of additional disclosure going on. The two types of disclosure are disclosure to other users of the site, and disclosure to third parties, primarily

advertisers. The privacy features provided by Facebook, to a large extent, allow the interested user to easily control what other users of the site can see about their profile data.

The issue here is that there are virtually no controls on what Facebook can expose to advertisers. The blanket statement regarding disclosure allows Facebook to disclose any personal data to advertisers. It also allows advertisers to set cookies that are not governed by the privacy policy. There is way to request that Facebook not share your information with others, but it is not transparent and there is no evidence that one's request is actually honored. See later in the paper for more details.

## 6.4 Access

"[Access] refers to an individual's ability both to access data about him or herself – i.e., to view the data in an entity's files – and to contest that data's accuracy and completeness. Both are essential to ensuring that data are accurate and complete." [6]

This attribute is more targeted at credit agencies and other organizations which maintain files on users which they may not want to disclose. Because Facebook is based on the sharing of information, and because Facebook provides users with the ability to control this information, Facebook follows this principle fairly well.

## 6.5 Security

Security is the process that ensures data integrity and restricts access to those who have been granted it legitimately. Privacy Online states in part "To assure data integrity, collectors must take reasonable steps, such as using only reputable sources of data and cross-referencing data against multiple sources, providing consumer access to data, and destroying untimely data or converting it to anonymous form."

Although Facebook is certainly vague about the uses to which the data will be put, it gives users control over the existence of information about themselves in the Facebook database. Their terms of service clearly state that "You may remove your Member Content from the site at any time. If you choose to remove your Member Content, the license granted above (that permits Facebook to use the data) will automatically expire."

"Security measures include encryption in the transmission and storage of data; use of passwords; and the storage of data on secure servers or computers that are inaccessible by modem."

By this standard, Facebook falls short. Although Facebook uses passwords to protect accounts and a MD5 hash as authorization, their use of encryption is nonexistent. All authorization information is sent in the clear, including the account passwords, making them exceedingly easy to sniff off of a public network. This is clearly inferior to the current best practices for password protection.

The "My Photos" feature seems to run counter to the Security principle, as third parties can upload pictures and associate them with one's account, without any checks on the accuracy or

appropriateness of the data. Users have no way of preventing pictures of them from being uploaded. Even if users seek to disassociate themselves with any photos, the most they can do is remove the tag that links the photo directly to the user's profile. In addition, there are absolutely no user controls akin to "My Privacy" relating to photos at all. We have found that any Facebook picture is accessible from any Facebook account, with no regard for privacy settings, or even the default Facebook per-university controls. One can ask to see all of the pictures of "Michael Smith" at Stanford and view them, even if one is logged into the MIT facebook.

## 6.6   Redress

"To be effective, self-regulatory regimes should include both mechanisms to ensure compliance (enforcement) and appropriate means of recourse by injured parties (redress)."

Much like the other privacy principles, Redress requires that customers be aware of ways in which they may be harmed. In the case of security breaches, there is no policy for notification of customers. In light of holes such as the "advanced search" hole described below, a clear policy on this matter would have been beneficial for users.

In addition, redress should entail acknowledgment of user requests and transparency in follow-through on them. The "prevent my information from being transmitted to third parties" request would be much improved if one could track the ramifications of that request.

# 7   Threat Model

## 7.1   Security Breach

**Threat and Feasibility**

A security breach at Facebook, either from an outsider locating vulnerability or from a disgruntled insider, would potentially put all 8,000,000 Facebook records at risk. This is not a risk that can be eliminated; no site is perfectly secure. The fear of a security breach is certainly a reasonable one, as large data warehouses are often targets of intruders. For example, ChoicePoint's databases were breached and 145,000 records were compromised. [3] While a Facebook breach would not be sufficient to start performing identity theft, a trove of so much personal information would contain much information that people would not want to make public.

**MySpace: A Comparison**

MySpace has several clauses in its Privacy policy that deal directly with contingencies that are not pleasant for the company to admit. The company tells users that security breaches can never

be completely prevented, even if "reasonable" steps are taken to prevent security breaches. This ensures that an unreasonable expectation of data security is not established[10].

In addition, MySpace confronts the possibility that they will be acquired, and notifies its users that their new owners could be less than scrupulous about using personal data. Their notification requirements regarding changes to their privacy policy appear to be aimed at this contingency. Unfortunately, MySpace does not have a notice requirement in the case of security breaches.

**Recommendation for Facebook: Security Disclosures**    Facebook should have a policy regarding disclosures of private information due to security breaches or unethical employees. A clearly stated requirement in their terms of service that they notify end-users whose privacy was violated would empower end-users.

## 7.2   Commercial Datamining

**Threat**

Companies such as ChoicePoint, Inc. have built billion-dollar business on selling databases of personal information. Facebook has a database on 8 million college students that is far more accurate than the usual commercial data, as users have an incentive to make information accurate. Profiles used for social networking are likely to be 100% accurate, as they are maintained by their subjects. This is in marked contrast to the accuracy of databases such as those maintained by ChoicePoint and Acxiom, which have records of dubious accuracy[15].

**Feasibility**

Using our code, attached as an appendix, we were able to crawl Facebook for four schools, creating a comprehensive data-set spanning all accessible profiles. Thus, we can conclude that it is possible to harvest data from the site. The fact that we (two students) were able to data-mine the Facebook in a week, using the time allotted to us for one class is evidence that data-mining the Facebook is evidence that it is not only possible, but easy.

**Current Precaution**

Facebook's Terms of Service state that using the site for data-harvesting purposes is forbidden. This statement offers no protection, however, if it is possible to use the site for these purposes, and there is no recourse against those who may seek to do so. Our data collection violates the Terms of Service for Facebook, which states that "You further agree not to harvest or collect email addresses or other contact information of members ... for the purposes of sending unsolicited emails or other unsolicited communications. Additionally, you agree not to use automated scripts to collect information from the Web site or for any other purpose." "Clickwrap" licenses like the terms of

service have generally been upheld by courts[10], but the danger posed to a person breaching this contract is uncertain at best. There are no provisions for the violation of the Terms of Service, and the termination of the offending account would not be a sufficient deterrent for those determined to obtain and use this information.

**Recommendations To Facebook: Better URL System**  Because of the method by which Facebook assigns User IDs, one can easily download all accessible profiles. A better system would be to make the profile number space 10 times the number of people eligible for accounts at the university, and assign user IDs randomly out of that. Then, when invalid UIDs are accessed, those IPs/accounts could be monitored for signs of abuse.

## 7.3  Database Reverse-Engineering

**Threat and Feasibility**

Facebook's "advanced search" allows one to query the database of users using any of the fields in a profile. For example, one can search for sophomore males at Duke that enjoy Kurt Vonnegut.

The problem is that when people hide their profile page, they expect the information on it to remain private. An MIT student could write "getting drunk" as an interest and set their profile so that only their friends could see their profile, expecting that this information is secure. This information is not actually secure unless they also exclude their profile from searches. An advanced search for "getting drunk" would still associate the students' name with this string.

The problem was compounded by a security hole that multiple people have discovered. Normally, performing a query at a certain college requires that one be logged in from an @thatcollege.edu account. A high school student at an MIT summer program discovered that by changing the server in the query URL from "mit.facebook.com" to "school.facebook.com", he could perform the query on any school without having a valid account for that school. He also discovered that most fields are indexed by ID number, so he was able to systematically query who lived in dorm "101", "102", etc, until he had a comprehensive list of where everyone said they lived in their profiles. He was only interested in using data on MIT students in an aggregated manner, but with that knowledge, one could easily reconstruct all Facebook profiles regardless of privacy preferences.

Further research found a student that actually employed this strategy to create a database of at other local schools. Up until November 10, 2005, he was able to systematically build up a database from queries on Facebook's database. Over the course of a month, he compiled information on over 82,000 students at 8 Boston-area schools.

---

[10]ProCD v. Zeidenberg, referenced in [19]

**Current Facebook Precaution**

Facebook blocks Advanced Search, except at one's school, which limits the scope of the problem. The "Exclude my name from searches" preference in the "My Privacy" section actually solves the problem. Because an intuitive leap is needed to see how to use the Advanced Search for data-mining, however, it takes the same intuitive leap for users to see the risk and protect themselves from it.

**Recommendation to Facebook: Restricting Search**   When users set their profile to be friends-only, all information save their name should be withheld from being searched by "Advanced Search."

## 7.4   Password Interception

**Threat**

The fact that the username and password were sent in cleartext is a security vulnerability. An adversary could read Facebook user names and passwords off of the Ethernet or unencrypted wireless traffic, obtaining access to users' Facebook passwords, as well as any additional accounts they use those passwords for. Because of the ethical and legal implications of doing so, we did not attempt to steal passwords. It should be noted, however, that MIT cited password theft as a real problem when they maintained telnet servers that had login data sent as cleartext. The University of New Mexico cited this as the main reason they chose to disable Facebook access from their network. Because many many users use their university email passwords as their Facebook passwords, UNM views Facebook as a security liability for their network.

**Current Facebook Precaution**

Facebook currently takes no steps to protect user passwords in transit.

**Recommendation to Facebook: Encrypt the Passwords**   Using SSL for login is the industry best practice for protecting passwords on login. It is used by Google Mail, eBay, MIT WebMail, and countless other sites to protect sensitive information as it is being transferred. It is a simple, cheap solution that would close a major security hole.

## 7.5   Incomplete Access Controls

**Threat and Feasibility**

In searching for user photos on Facebook, the service uses a variant of this URL:

$$http://mit.facebook.com/photo\_search.php\&name = John \tag{6}$$

There is nothing inherently wrong with allowing users to search for photos, but there are no restrictions akin to "My Privacy" for photographs. In addition, the usual access controls do not apply to "My Photos," anyone from any university can search for and see any other photograph by editing the query URL.

The ability of users to upload and tag photographs easily, and the difficulty for a user to de-tag large numbers of photographs, makes it easy for others to find photographs with few restrictions.

### Current Facebook Precaution

Facebook limits photograph searches by profile in the same way they limit regular searches; the problem lies in the additional unrestricted method of searching all photos by name.

**Recommendation to Facebook:  Restrictions on Pictures Search**   This is weaker than any other access controls on the site; by default, users are unable to view others' profiles on other websites, but they can view all pictures. "My Privacy" should extend to the "My Photos" feature as well, and the search by name should be disabled.

## 7.6   University Surveillance

### Threat

Students in many cases are unaware of the complex interactions between university policy and the information they are making available online.  Administrators are using Facebook to learn about their students... and their students' activities. Recent months have seen a rash of incidents coming from students disclosing information that they never thought would end up in deans' offices, but has. These problems are not limited to technical schools like MIT, they exist all over the nation.

### Feasibility

**MIT**   MIT has not had any high-profile Facebook-related cases yet, but there have been smaller incidents, and a growing realization of the importance of Facebook in a college environment. Dean of Residential Life Programs Andrew Ryder has stated that MIT is not actively monitoring Facebook for rule infractions.  He did say, however, that if public or quasi-public Facebook information was brought to his attention, he would have to act on it.  It is also his personal belief that Facebook data would be admissible in Committee on Discipline hearings.  Without detailing specific cases, he alluded to the fact that Facebook incidents that MIT has had to deal with so far have related to a student posting unflattering or untrue information about another student, which generated a complaint to the Department for Student Life. The one other MIT case involved a freshman in the class of 2008 advertising a party in his soon-to-be dorm room on Facebook before he even arrived on campus.

**Cameron Walker and Fisher College**   In October of 2005, Cameron Walker, then a second year student at Fisher College in Boston, MA, was expelled from the school and barred from the campus. The reason for this action given by Fisher College was Walker's creation of a Facebook group committed to the dismissal of a campus security officer believed to regularly overstep the limits of his line of duty. School officials who monitored Facebook, pressured Walker to remove the group, and ultimately canceled Fisher's student status.

   Mr. Walker's expulsion could set a dangerous precedent for university officials. Students believe that the information they post to Facebook should be protected as correspondence, while school officials, particularly at schools with strict codes of discipline, will use evidence posted on Facebook to bring formal disciplinary charges against students. This is the first incident of a student being expelled for actions on Facebook. We conducted a phone interview with Walker in mid-Novemnber. He was a sophomore in the class of 2008 in October 2005, when the events leading to his expulsion occurred. His expulsion demonstrates the issues that can arise from the interactions of Internet publication and "unclear, ambiguous, and vague" (Walker's words) student codes of conduct, especially as they pertain to harassment. Walker claims that his expulsion was an example of a "few administrators doing whatever they wanted", and that he "was naive about Facebook, because it wasn't affiliated with a university."

**News at Other Schools**   In recent weeks, there has been an explosion of articles in college newspapers relating to the privacy concerns of Facebook. The recent expulsion of Cameron Walker may have created a concrete example of the harm that can come from Facebook activity; it is the one case that many news articles mention. Since November 1, cautionary articles have appeared in the newspapers of Emory[21], Georgia College[22], Dartmouth[23], the University of Oregon[24], Trinity College[25], Macalester[26], Syracuse[27], Brown[28], GW, University of Tennessee at Chattanooga[29], UNC Greensboro[30], and UPenn[31].

**Current Facebook Precaution**

The Facebook currently does not take steps to prevent this type of disclosure.

**Recommendation to Universities:**   From a student perspective, Facebook has been an area relatively free of administrative interference until now. University policies are two-fold; there is the letter of the law, and what is actually enforced. The wealth of new information available to administrators pushes the enforceability much closer to the literal readings of school policies, which could have many unintended consequences. On the other hand, administrators are not free to set whatever policies they see fit, and in an age of litigation, they cannot afford to selectively enforce policies. To do so would be to make the university vulnerable to lawsuits in cases where forbidden behavior goes too far undetected.

In addition, Facebook is becoming a key component of college life, and college administrators would not be doing their jobs if they didn't understand and explore how a large portion of their student body was using their spare time and interacting with each other.

Because of this complex interaction, and the differing goals that administrators have, colleges should look at their primary interaction with Facebook an educational one. Students can only claim that they have been treated unfairly if they can establish an expectation of privacy. If universities are going to use this information, they should tell their students this up-front.

**Recommendation to Universities: Educate Students** The university's most important role, however, is that of education. To fulfill this mission, universities should educate their students about the dangers that online disclosure of information can pose. Because students are getting accounts earlier and earlier, a program during Orientation would help students from running afoul of university policy or being harassed.

**Recommendation to Facebook: Warnings Page** In an environment of growing misuse of information made public by Facebook, Facebook would do its users a great service to explain the dangers of security breaches and outside monitoring. Until the societal norms regarding this new use of computers become well-established, Facebook could clearly state that they could provide no guarantees regarding the security of their data, and that if users make their profiles public, all information contained therein may be viewed by job interviewers and college administrators.

**Recommendation to Facebook: Opt-Out Privacy** In a world where a minority of users change software preferences, privacy protection cannot be an "opt-in" option. Facebook faces a tough choice here: their business model is based on many ad views, which requires extended browsing sessions, which requires a relatively open network. Yet, opt-out protection is far more effective, as demonstrated by Shah and Sandvig in "Software Defaults as De Facto Regulation." Their study found that if encryption on WAPs is set by default, 96% of users employ it, 3.4 times the number that do when it is not set by default.

**Recommendation to Facebook: Merge "My Privacy"** Facebook is unique, however, in that users are expected to return often and update their "preferences" (who their friends are, their profile information). Thus, Facebook could leverage this culture by merging the functions of profile updating and privacy settings. One page could contain fields regarding basic profile information as well as privacy settings, thereby greatly increasing the number of views the privacy settings get daily.

## 7.7 Disclosure to Advertisers

**Threat and Feasibility**

Facebook has a relationship with several companies currently. Apple and JetBlue, among others, have their own "groups" that interested users can join, to show their brand loyalty, or for a chance at giveaways. Facebook's privacy policy explicitly says that they may disclose profile information to third parties, so the prospect of them doing so is clearly realistic.

**Current Facebook Precautions**

Facebook offers an "opt out" link on their Privacy Policy page, which, if clicked, means that one can "submit a request" to Facebook to not share information with third parties. They say that they "will make every effort to implement any choice you make as soon as possible." Offering the user choice in this matter is clearly to the user's benefit. However, the feature has no followup or feedback, and is couched in language that does not actually imply any sort of binding agreement.

**Other Services' Precautions**

**Friendster**    Friendster's privacy policy is indicative of a more mature service, with narrower goals, dealing with smaller amounts of personal information than Facebook. Friendster only collects the data you enter into your profile, your name, e-mail address, IP address, and user agent. Unlike Facebook, Friendster agrees to never share your information with any outside agency, unless expressly required to do so by law.

**MySpace**    MySpace also has a much more explicit and user-oriented disclosure policy. The scope of disclosure to third parties is much more explicitly dealt with, and limited to:

- Disclosure to advertisers whom users have "explicitly requested" to receive information from[11].

- The use of cookies by advertisers. [12]

- Disclosures required to enforce their TOS, to protect them legally, or to protect the safety of the public[13].

---

[11]Users may be asked to provide personal information including name, email address or home address or to answer questions in order to participate. We may transfer personal information to certain ad partners that you have explicitly requested to receive information from. It will be clear at the point of collection who is collecting the personal information and whose privacy statement will apply.

[12] "A User is bound by any minor changes to the policy when she or he uses the site after those changes have been posted If, however, we are going to use users' personally identifiable information in a manner materially different from that stated at the time of collection we will notify by posting a notice on our Web site for 30 days."

[13] "Except as otherwise described in this privacy statement, MySpace will not disclose personal information to any

**Recommendation to Facebook: Accountability and Accessibility for Third-Party Opt-Out**
An opt-out feature that guaranteed that the user's information would not be disclosed in the future would allow users much more control over their privacy. If the process is complex, then a method for tracking one's request would increase the transparency of the process. In addition, the link is buried in the privacy policy, which is a legal agreement; users who want to take action would look to "My Privacy." To actually make the option effective, it should be located in "My Privacy."

**Recommendation to Facebook: Privacy Policy Improvements**  Facebook's privacy policy is vague and subject to change at the whim of the owners of the website. The Facebook policy allows any disclosure of information to third parties that Facebook feels is appropriate. Facebook should seek to emulate MySpace in this manner, and perhaps even go farther.

A user-centered Terms of Service would clearly delineate which information is shared with which partners, depending on whether a user clicked on a third party's ad or joined a third party's group. A notice period announcing a change in the Terms of Service is another change that would improve the user experience.

## 7.8   Lack of User Control of Information

**Threat**

Other users can upload and associate information to one's Facebook account. The most prominent feature of this type is the "My Photos" feature, which allows users to upload photos and tag them with the names of the people in the pictures. This functionality has already resulted in trouble for an underage student at University of Missouri-Columbia when college administrators found a picture of her duct-taped to a chair while another student poured beer in her mouth. This was a matter of considerable embarassment as she had just been elected student body vice president. The university is currently considering removing her from that role.

**Current Facebook Precaution**

Facebook allows users to de-associate themselves from unwanted data, but in the case of photographs, the data remains on the server. This is also an "opt-in" function that requires constant monitoring of the system.

---

third party unless we believe that disclosure is necessary: (1) to conform to legal requirements or to respond to a subpoena, search warrant or other legal process received by MySpace.com, whether or not a response is required by applicable law; (2) to enforce the MySpace.com Terms of Use Agreement or to protect our rights; or (3) to protect the safety of members of the public and users of the service."

**Recommendation to Facebook: Better Restrictions on Third-Party Information**   Third parties' ability to submit and associate information about users violates one of the key principles of information practices: the idea that users should have the ability to control and correct the information about them in a particular database. Although Facebook allows users to delete Wall postings and de-associate themselves with photographs, this is an "opt-in" mechanism that requires constant monitoring. Modifying the "My Privacy" feature to allow a blanket disabling of these features for a particular user would help users control their information.

**Recommendation to Users: Exercise Caution**   Users should be aware that there are effectively no access controls on pictures, and that they should only upload the pictures that they would feel comfortable having anybody on the Facebook viewing.

In addition, realize that the photos that you upload of other people may be viewed by their high school friends or their family. Don't post anything of them doing anything that you wouldn't want your parents to see you doing.

## 7.9   Summary and Conclusion

Ultimately, lasting change in online privacy will only come from a gradual development of common sense regarding what is appropriate to post in social networking forums. Unfortunately, this is not an easy fix. Until users view alluding to underage drinking or drug use on their profiles as risky, mistakes regarding privacy will continue to occur. Revealing this sort of information needs to be viewed as the equivalent of going alone to the apartment of a person one met on the Internet.

It is vital that Facebook users everywhere appreciate the potential for use of the system by administrators. We strongly advise all Facebook users to restrict access to their profiles, to not post information of illegal or policy-violating actions to their profiles, and to be cautious with the information they make available.

This lasting change will only come with time and understanding. Nobody can fault Facebook for students making questionable decisions, but the environment that Facebook creates should be one that fosters good decision-making. Privacy should be the default, encryption should be the norm, and Facebook should take strides to inform users of their rights and responsibilities.

# 8   Conclusion

## 8.1   Postscript: What the Facebook does right

A paper that analyzes the threats to privacy a system poses will inevitably adopt a negative tone about the target of its examination. Although Facebook has flaws, there are also areas in which it is a leader among social networking sites. The fact that each university Facebook is effectively its own

site virtually firewalled off from the rest of the network is a much more private-by-default system than Friendster or MySpace, which explicitly notes that there is no way to restrict profile information. This system makes data harvesting much harder, though not impossible. The requirement of having a school email account to sign up is largely effective in preventing fake accounts and what could otherwise be a problem of Facebook "identity theft."

The "My Privacy" settings model is fundamentally sound. The current model would be close to ideal if the defaults and behaviors of settings were changed, which would not require a substantial engineering effort.

Although the flaws with "My Photos" are pronounced, the existing security model is robust enough to solve most of the problems associated with it. If the name search for photos followed "My Privacy" rules, it would be allow users to control their data very easily.

### 8.2   Final Thoughts

Facebook is used by over 8 million college students, but no academic study has been done of its effect on end-users. As with any emerging technology, the common sense regarding its proper use has lagged behind what technology has made possible. Although the Internet has made it possible to publish personal information online for a decade, social networking sites are unique in that they standardize, centralize, and encourage the publication of personal data to an unprecedented extent. The consequences of excessive disclosure of personal information and false senses of security are just beginning to emerge. Although no national attention has been devoted to the issue, more stories of students being disciplined because of Facebook appear in college newspapers every week. As information retrieval and analysis tools become more powerful, the public needs to develop common sense about accepted practices on these sites. Much as it is now common sense to not meet people online without taking significant precautions, a body of common knowledge about disclosing information online would protect the public. This research aims to begin that dialogue. From a technological perspective, there has been little dialogue about investigating the protections put in place at one of the most-visited sites on the internet, which contains detailed files on more than 8 million young adults. Security by obscurity is not the best practice for any system, let alone one used by so many. The user community of this site and future sites will benefit from increased attention to these issues.

## References

[1] Adamic, Lada A., Buyukkotken, Orkut, and Adar, Eytan. 2002. "A Social Network Caught In The Web." http://www.hpl.hp.com/research/idl/papers/social/social.pdf

[2] Sandvig, C. & Shah, R. (2005). Defaults as De Facto Regulation: The Case of Wireless Access Points. Paper presented at the 33rd Telecommunications Policy Research Conference (TPRC) on Communication, Information, and Internet Policy, Arlington, Virginia, USA.

[3] Konrad, Rachel. Associated Press. February 24, 2005, "Burned by ChoicePoint breach, potential ID theft victims face a lifetime of vigilance."

[4] Terremark Worldwide, Inc. "Facebook Expands Operations at Terremark's NAP West Facility" Tuesday November 1, 8:30 am ET.

[5] Newitz, Annalee. "Dangerous Terms: A User's Guide to EULAs." http://www.eff.org/wp/eula.php. Loaded December 14, 2005.

[6] Federal Trade Commission, Privacy Online: Report to Congress, 1999.

[7] Facebook Privacy Policy, available online at http://www.facebook.com/policy.php.

[8] Facebook Terms of Service, available online at http://www.facebook.com/terms.php.

[9] MySpace Terms of Service, available online at http://viewmorepics.myspace.com/misc/terms.html.

[10] MySpace Privacy Policy, available online at http://viewmorepics.myspace.com/misc/privacy.html.

[11] Friendster Terms of Service, available online at http://www.friendster.com/info/tos.php.

[12] Friendster Privacy Policy, available online at http://www.friendster.com/info/privacy.php.

[13] New York Times, August 28, 2005. "Do You MySpace?" By Alex Williams.

[14] Marshall, Matt and Anna Tong. "Palo Alto, Calif.-based Facebook brings social networking online." San Jose Mercury News, August 29, 2005.

[15] Data Aggregators: A Study of Data Quality and Responsiveness. Pierce, Deborah and Linda Ackerman. May 19, 2005 http://www.privacyactivism.org/docs/DataAggregatorsStudy.html

[16] New York University Admissions, "Fast Facts", http://admissions.nyu.edu/fast_facts/

[17] Sample Size Calculator, http://www.surveysystem.com/sscalc.htm

[18] Phone Interview, Daniel Dedap

[19] Contracts, Copyright, and Confusion: Revisiting the Enforceability of 'Shrinkwrap' Licenses. Heath, Steven. Chicago-Kent Intellectual Property Law Society Journal of Intellectual Property.

## 8.3 College Newspaper Articles

[20] Sealy, Will. "What facebook doesnt tell you." The Flat Hat, student newspaper of The College of William and Mary. http://flathat.wm.edu/story.php?issue=2005-11-04&type=2&aid=3. Loaded December 14, 2005.

[21] Zelkowitz, Rachel. " 'Wasted' Facebook group causes controversy." The Emory Wheel Online, November 22, 2005. http://www.emorywheel.com/vnews/display.v/ART/2005/11/22/43829c13eb4d8. Loaded December 14, 2005.

[22] "Public Safety considers Facebook a valuable tool for party busts." The Colonnade, Georgia College and State University. November 4, 2005. http://www.gcsunade.com/media/paper299/news/2005/11/04/CampusNews/ Public.Safety.Considers.Facebook.A.Valuable.Tool.For.Party.Busts-1046210.shtmlLoaded December 14, 2005.

[23] Paquin, Christine. "Administrators advise caution in Facebook postings" The Dartmouth, November 21, 2005. http://www.thedartmouth.com/article.php?aid=2005112101070. Loaded December 14, 2005.

[24] "Facebook could invite more than your friends." Oregon Daily Emerald, November 28, 2005. http://www.dailyemerald.com/vnews/display.v/ART/2005/11/28/438aca3122ba8. Loaded December 14, 2005.

[25] Montermini, Fabrizio. "Facebook Raises Privacy Concerns." The Trinity Tripod, November 29, 2005. http://www.trinitytripod.com/media/paper520/news/2005/11/29/News/ Facebook.Raises.Privacy.Concerns-1115345.shtml. Loaded December 14, 2005.

[26] Martucci, Brian. "As Facebook grows, more than just friends are watching." The Mac Weekly, December 9, 2005. http://www.themacweekly.com/article.php?arid=133. Loaded December 14, 2005.

[27] Shoffel, Jessical. "SUNY-ESF warns students of Facebook content violating conduct codes." The Daily Orange, December 2, 2005. http://www.dailyorange.com/media/paper522/news/2005/12/02/News/ SunyEsf.Warns.Students.Of.Facebook.Content.Violating.Conduct.Codes-1119079.shtml. Loaded December 14, 2005.

[28] Woo, Stu. "The Facebook: not just for students." The Brown Daily Herald, November 3, 2005. http://www.browndailyherald.com/media/paper472/news/2005/11/03/CampusWatch/ The-Facebook.Not.Just.For.Students-1044229.shtml. Loaded December 14, 2005.

[29] Walker, Rachel. "UTC cops check Facebook for underage drinkers." The Echo online, November 10, 2005. http://www.utcecho.com/media/paper483/news/2005/11/10/Culture/Utc-Cops.Check.Facebook.For.Underage.Drinkers-1053481.shtml. Loaded December 14, 2005.

[30] McIntyre, Luke. "FAILURE TO COMMUNICATE: Don't let Facebook land you in jail." The Carolinian Online, November 8, 2005. http://www.carolinianonline.com/media/paper301/news/2005/11/08/Opinions/Failure.To.Communicate.Dont.Let.Facebook.Land.You.In.Jail-1048102.shtml. Loaded December 14, 2005.

[31] Kramer, Melody Joy. "Forfeiting privacy, one post at a time." The Daily Pennsylvanian, November 30, 2005. http://www.dailypennsylvanian.com/vnews/display.v/ART/438d34a676ff6. Loaded December 14, 2005.

[32] Wang, Jiao. "Facebook Profiles Become Handy Tool for Recruiters." The Tech, December 13, 2005. http://www-tech.mit.edu/V125/N61/facebook.html. Loaded December 14, 2005.

# 9 Acknowledgements

## 9.1 Interview subjects

- Andrew Ryder, Assistant Dean, MIT Residential Life Programs

- Sharon Snaggs, Residential Life Associate, MIT

- Christopher Varenhorst, MIT Undergraduate

- Facebook scraper (name withheld)

- Jeff Gassaway, University of New Mexico Security Administrator

- Cameron Walker, Fisher College student

- Daniel Dedap, NYU alumnus, class of 2005

# A    Facebook Privacy Policy

[7] This policy is effective as of June 28, 2005.

**Introduction**    The Facebook Privacy Policy is designed to assist you in understanding how we collect and use the personal information that you provide to us and to assist you in making informed decisions when using the Facebook web site located at www.facebook.com (the "Web Site").

**The Information We Collect**    When you visit the Web Site you may provide us with two types of information: personal information you knowingly choose to disclose that is collected by us and Web Site use information collected by us on an aggregate basis as you and others browse our Web Site.

When you register on the Web Site, you provide us with certain personal information, such as your name, your email address, your telephone number, your address, your gender, schools attended and any other personal or preference information that you provide to us.

When you enter our Web Site, we collect the user's browser type and IP address. This information is gathered for all users to the Web Site. In addition, we store certain information from your browser using "cookies." A cookie is a piece of data stored on the user's computer tied to information about the user. We use session ID cookies to confirm that users are logged in. These cookies terminate once the users close the browser. We do not use cookies to collect private information from any user.

Facebook also collects information about you from other sources, such as newspapers and instant messaging services. This information is gathered regardless of your use of the Web Site.

**Children Under Age 13**    Facebook does not knowingly collect or solicit personal information from anyone under the age of 13 or allow such persons to register. If you are under 13, please do not send any information about yourself to us – including information like your name, address, telephone number, or e-mail address. No one under age 13 is allowed to provide any personal information or use our public forums. In the event that we learn that we have collected personal information from a child under age 13 without verification of parental consent, we will delete that information as quickly as possible. If you believe that we might have any information from or about a child under 13, please contact us at: info@facebook.com.

**Children Between the Ages of 13 and 18**    We recommend that minors over the age of 13 ask their parents for permission before sending any information about themselves to anyone over the Internet.

**Use of Information Obtained by Facebook**    When you register on the Web Site, you create your own profile and privacy settings. Your profile information, as well as your name, email and photo,

are displayed to people in the groups specified in your privacy settings to support the function of the Web Site. In addition, we may use your name and email address to send you notifications regarding the Web Site and, occasionally, new services we think you may find valuable.

No personal information that you submit to Facebook will be available to any user of the Web Site who does not belong to at least one of the groups specified by you in your privacy settings.

We use the information about you that we have collected from other sources to supplement your profile unless you specify in your privacy settings that you do not want this to be done.

**Sharing Your Information with Third Parties**   We may share your information with third parties, including responsible companies with which we have a relationship. For example:

- We may provide information to service providers to help us bring you the services we offer. Specifically, we may use third parties to facilitate our business, such as to send email solicitations. In connection with these offerings and business operations, our service providers may have access to your personal information for use in connection with these business activities.

- We may be required to disclose customer information pursuant to lawful requests, such as subpoenas or court orders, or in compliance with applicable laws. Additionally, we may share account or other information when we believe it is necessary to comply with law or to protect our interests or property. This may include sharing information with other companies, lawyers, agents or government agencies.

- If the ownership of all or substantially all of the Facebook business were to change, your user information would likely be transferred to the new owner. If you do not want to receive promotional email from Facebook and/or do not want us to share your information with third parties for marketing purposes, please submit a request by clicking here http://mit.facebook.com/help.php?add=1. We will make every effort to implement any choice you make as soon as possible.

**Links**   This site may contain links to other websites. Facebook is not responsible for the privacy practices of other web sites. We encourage our users to be aware when they leave our site and to read the privacy statements of each and every web site that collects personally identifiable information. This privacy statement applies solely to information collected by Facebook Web Site.

**Third Party Advertising**   Advertisements that appear on the Web Site are delivered to users by our advertising partners. Our advertising partners may download cookies to your computer. Doing this allows the advertising network to recognize your computer each time they send you an advertisement. In this way, they may compile information about where you, or others who are using your computer, saw their advertisements and determine which advertisements are clicked. This information allows

an advertising network to deliver targeted advertisements that they believe will be of most interest to you. Facebook does not have access to or control of the cookies that may be placed by the third party advertisers.

This privacy statement covers the use of cookies by Facebook and does not cover the use of cookies by any of its advertisers.

**Changing or Removing Information**   Facebook users may modify or remove any of their personal information at any time by logging into their account. Information will be updated immediately.

**Security**   Facebook takes appropriate precautions to protect our users' information. Your account information is located on a secured server behind a firewall. Because email is not recognized as a secure medium of communication, we request that you do not send private information to us by email. If you have any questions about the security of Facebook Web Site, please visit our Help page http://mit.facebook.com/help.php for more information..

**Changes in Our Privacy Policy**   We reserve the right to change our privacy policy at any time. If we do this, we will post the changes to this policy on this page and will indicate at the top of this page the policy's effective date. We therefore encourage you to refer to this policy on an ongoing basis so that you understand our current privacy policy.

**Contacting the Web Site**   If you have any questions about this privacy policy, please visit our Help page http://mit.facebook.com/help.php for more information.

# B   Facebook Terms Of Service

[8] These Terms of Use are effective as of October 3, 2005.

**Introduction**   Welcome to the Facebook, an online directory that connects people through networks of academic and geographic centers. The Facebook service is operated by the Facebook network ("Facebook"). By using the Facebook web site (the "Web site") you signify that you have read, understand and agree to be bound by these Terms of Use (this "Agreement"). We reserve the right, at our sole discretion, to change, modify, add, or delete portions of these Terms of Use at any time without further notice. If we do this, we will post the changes to these Terms of Use on this page and will indicate at the top of this page the Terms of Use's effective date. Your continued use of the Web site after any such changes constitutes your acceptance of the new Terms of Use. If you do not agree to abide by these or any future Terms of Use, please do not use or access Web site. It is your responsibility to regularly review these Terms of Use.

**Eligibility**  You must be thirteen years of age or older to register as a member of Facebook or use the Web site. If you are under the age of 13, you are not allowed to register and become a member of Facebook or access Facebook content, features and services on the Web Site. Membership in the Service is void where prohibited. By using the Web site, you represent and warrant that you agree to and to abide by all of the terms and conditions of this Agreement. Facebook may terminate your membership for any reason, at any time.

**Member Conduct**  You understand that the Web site is available for your personal, non-commercial use only. You agree that no materials of any kind submitted through your account will violate or infringe upon the rights of any third party, including copyright, trademark, privacy or other personal or proprietary rights; or contain libelous, defamatory or otherwise unlawful material. You further agree not to harvest or collect email addresses or other contact information of members from the Web site by electronic or other means for the purposes of sending unsolicited emails or other unsolicited communications. Additionally, you agree not to use automated scripts to collect information from the Web site or for any other purpose. You further agree that you may not use Web site in any unlawful manner or in any other manner that could damage, disable, overburden or impair Web site. In addition, you agree not to use the Web site to:

- upload, post, email, transmit or otherwise make available any content that we deem to be harmful, threatening, abusive, harassing, vulgar, obscene, hateful, or racially, ethnically or otherwise objectionable;

- impersonate any person or entity, or falsely state or otherwise misrepresent yourself or your affiliation with any person or entity;

- upload, post, email, transmit or otherwise make available any unsolicited or unauthorized advertising, promotional materials, "junk mail," "spam," "chain letters," "pyramid schemes," or any other form of solicitation;

- upload, post, email, transmit or otherwise make available any material that contains software viruses or any other computer code, files or programs designed to interrupt, destroy or limit the functionality of any computer software or hardware or telecommunications equipment;

- intimidate or harass another;

- use or attempt to use another's account, service or system without authorization from Web site, or create a false identity on this website.

**Proprietary Rights in Content on Facebook**  All content on Web site, including but not limited to design, text, graphics, other files, and their selection and arrangement (the "Content"), are

the proprietary property of Facebook or its licensors. All rights reserved. The Content may not be modified, copied, distributed, framed, reproduced, republished, downloaded, displayed, posted, transmitted, or sold in any form or by any means, in whole or in part, without Web site's prior written permission. You may download or print a copy of any portion of the Content solely for your personal, non-commercial use, provided that you keep all copyright or other proprietary notices intact. You may not republish Content on any Internet, Intranet or Extranet site or incorporate the information in any other database or compilation. Any other use of the Content is strictly prohibited.

All trademarks, logos, trade dress and service marks on the Web site are either trademarks or registered trademarks of Facebook or its licensors and may not be copied, imitated, or used, in whole or in part, without the prior written permission of Facebook.

**Member Content Posted on the Site** You are solely responsible for the content, photos or profiles Content that you publish or display (hereinafter, "post") on the Service, or transmit to other Members (collectively the "Member Content"). You understand and agree that Facebook may review and delete or remove any Member Content that in the sole judgment of Facebook violate this Agreement or which might be offensive, illegal, or that might violate the rights, harm, or threaten the safety of Members.

By posting Member Content to any part of the Web site, you automatically grant, and you represent and warrant that you have the right to grant, to Facebook an irrevocable, perpetual, non-exclusive, transferable, fully paid, worldwide license (with the right to sublicense) to use, copy, perform, display, reformat, translate, excerpt (in whole or in part) and distribute such information and content and to prepare derivative works of, or incorporate into other works, such information and content, and to grant and authorize sublicenses of the foregoing.

You may remove your Member Content from the site at any time. If you choose to remove your Member Content, the license granted above will automatically expire.

**Copyright Policy** Facebook respects the intellectual property rights of others. If you believe your work has been copied in a way that constitutes copyright infringement or are aware of any infringing material on the Web site, please contact us at copyright@facebook.com and provide us with the following information: an electronic or physical signature of the person authorized to act on behalf of the owner of the copyright interest; a description of the copyrighted work that you claim has been infringed; a description of where the material that you claim is infringing is located on the Web site; your address, telephone number, and email address; a written statement by you that you have a good faith belief that the disputed use is not authorized by the copyright owner, its agent, or the law; a statement by you, made under penalty of perjury, that the above information in your notice is accurate and that you are the copyright owner or authorized to act on the copyright owner's behalf.

**Links to other websites**  The Web site contains links to other web sites.  Facebook is not responsible for the content, accuracy or opinions express in such web sites, and such web sites are not investigated, monitored or checked for accuracy or completeness by us. Inclusion of any linked web site on Facebook Web site does not imply approval or endorsement of the linked web site by Facebook.  If you decide to leave Facebook Web site and access these third-party sites, you do so at your own risk.

**Member Disputes**  You are solely responsible for your interactions with other Facebook Members. Facebook reserves the right, but has no obligation, to monitor disputes between you and other Members.

**Privacy**  Facebook cares about the privacy of its members.  Click here to view the Web site's Privacy Policy.

**Disclaimers**  Facebook is not responsible for any incorrect or inaccurate Content posted on the Web site or in connection with the Service, whether caused by users of the Web site, Members or by any of the equipment or programming associated with or utilized in the Service.  Facebook is not responsible for the conduct, whether online or offline, of any user of the Web site or Member of the Service.  The Service may be temporarily unavailable from time to time for maintenance or other reasons.  Facebook assumes no responsibility for any error, omission, interruption, deletion, defect, delay in operation or transmission, communications line failure, theft or destruction or unauthorized access to, or alteration of, user or Member communications.  Facebook is not responsible for any problems or technical malfunction of any telephone network or lines, computer online systems, servers or providers, computer equipment, software, failure of email or players on account of technical problems or traffic congestion on the Internet or at any web site or combination thereof, including injury or damage to users and/or Members or to any other person's computer related to or resulting from participating or downloading materials in connection with the Web and/or in connection with the Service.  Under no circumstances will Facebook be responsible for any loss or damage, including personal injury or death, resulting from anyone's use of the Web site or the Service, any Content posted on the Web site or transmitted to Members, or any interactions between users of the Web site, whether online or offline.  THE WEB SITE, THE SERVICE AND THE CONTENT ARE PROVIDED "AS-IS" AND FACEBOOK DISCLAIMS ANY AND ALL WARRANTIES, WHETHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION IMPLIED WARRANTIES OF TITLE, MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT. FACEBOOK CANNOT GUARANTEE AND DOES NOT PROMISE ANY SPECIFIC RESULTS FROM USE OF THE WEB SITE AND/OR THE SERVICE.

**Limitation on Liability** EXCEPT IN JURISDICTIONS WHERE SUCH PROVISIONS ARE RE-STRICTED, IN NO EVENT WILL FACEBOOK BE LIABLE TO YOU OR ANY THIRD PERSON FOR ANY INDIRECT, CONSEQUENTIAL, EXEMPLARY, INCIDENTAL, SPECIAL OR PUNITIVE DAMAGES, INCLUDING ALSO LOST PROFITS ARISING FROM YOUR USE OF THE WEB SITE OR THE SERVICE, EVEN IF FACEBOOK HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. NOTWITHSTANDING ANYTHING TO THE CONTRARY CONTAINED HEREIN, FACEBOOK'S LIABILITY TO YOU FOR ANY CAUSE WHATSOEVER, AND REGARDLESS OF THE FORM OF THE ACTION, WILL AT ALL TIMES BE LIMITED TO THE AMOUNT PAID, IF ANY, BY YOU TO FACEBOOK FOR THE SERVICE DURING THE TERM OF MEMBERSHIP.

**Governing Law and Venue** If there is any dispute about or involving the Web site and/or the Service, you agree that the dispute will be governed by the laws of the State of California without regard to its conflict of law provisions. You also agree to the exclusive jurisdiction and venue of the courts of the state and federal courts of Santa Clara County, California and waive all defenses of lack of personal jurisdiction and forum non conveniens. Any cause of action by you with respect to the Web site and/or the Service must be instituted within one (1) year after the cause of action arose or be forever waived and barred.

**Indemnity** You agree to indemnify and hold Facebook, its subsidiaries, affiliates, officers, agents, and other partners and employees, harmless from any loss, liability, claim, or demand, including reasonable attorney's fees, made by any third party due to or arising out of your use of the Service in violation of this Agreement or your violation of any law or the rights of a third party.

**Other** These Terms of Use constitute the entire agreement between you and Facebook regarding the use of the Web site and/or the Service, superseding any prior agreements between you and Facebook relating to your use of the Web site or the Service. The failure of Facebook to exercise or enforce any right or provision of these Terms of Use shall not constitute a waiver of such right or provision. If any provision of this Agreement is held invalid, the remainder of this Agreement shall continue in full force and effect.

**Questions** Please visit our Help page for more information.

# C  Facebook "Spider" Code: Acquisition and Processing

The following code extracts all Facebook accounts from a given school that are accessible given the user account provided.

## C.1   Data Downloading BASH Shell Script

```
wget --cookies=on --user-agent='Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US;
rv:1.7.12) Gecko/20050915 Firefox/1.0.7' --save-cookies=cookies.txt
--keep-session-cookies --load-cookies=cookies.txt
'http://www.facebook.com/login.php?email=LOGIN&pass=PASS'


for ((  COUNT = USERID_LOW ;  COUNT <= USERID_HIGH;  COUNT++  ))


do


wget --cookies=on --wait=12 --random-wait --user-agent='Mozilla/5.0 (Windows;
U; Windows NT 5.1; en-US; rv:1.7.12) Gecko/20050915 Firefox/1.0.7'
--save-cookies=cookies.txt --keep-session-cookies --load-cookies=cookies.txt
http://SCHOOL.facebook.com/profile.php?id=$COUNT


done
```

## C.2   Facebook Profile to Tab Separated Variable Python Script

```python
import string
import sys
import re
import os


htmltag = re.compile('<.*?>')


def make_search(str):
    lam = lambda data: re.search(".*%s\:.*" % str, data)
    return lam


def strip_html(data):
    return htmltag.sub("", data)


attrib=["Name", "Member Since", "Last Update", "School", "Status", "Sex",
"Concentration", "Residence", "Mailbox", "Hometown", "High School",
"Screenname", "Mobile", "Site", "Interests", "Clubs and Jobs", "Favorite
Music", "Favorite Movies", "Favorite Books"]
```

```python
lambdas = map(make_search, attrib)


def process(fname):
    f = open(fname, "r")
    data = f.read()
    dbak = data
    try:
        friendstr = string.split(data, "category_id=2")[1]
        friends = string.split(friendstr, " ")[0][2:]
    except IndexError:
        friends= ""


    try:
        data = string.split(data, "<h2>Information</h2>")[1]
        data = string.split(data, "<!-- userprofile -->")[0]
    except IndexError:
        sys.stderr.write("Error! %s" % fname)
        data = dbak



    if len(string.split(data, "Groups")) == 2:
            data = string.split(data, "Groups")[0]
    data = string.split(data, "\n")
    data = map(strip_html, data)


    fields=[""]*len(attrib)


    for x in range(len(attrib)):
        field = filter(lambdas[x], data)
        if field == []:
            fields[x] = ""
        else:
            fields[x] = string.split(field[0], ":")[1]
        if attrib[x] == "Name":
            fields[x] = string.split(fields[x], "&")[0]
    for f in fields:
```

```
        print f, "\t",
    print friends



for f in os.listdir(sys.argv[1]):
    if f[:5] == "profi":
        process(sys.argv[1]+"/"+ f)
```

## C.3   Data Analysis Scripts

### C.3.1   The after date script.

```
import string
import sys


# usage: python afterdate.py col val
# afterdate prints all records whose column #col is after val
# val is of the form yyyymmdd


col = int(sys.argv[1])
val = string.strip(sys.argv[2])


s = "foo"


month={"January":"01",
       "February":"02",
       "March":"03",
       "April":"04",
       "May":"05",
       "June":"06",
       "July":"07",
       "August":"08",
       "September":"09",
       "October":"10",
       "November":"11",
       "December":"12",}


while True:
```

```
    try:
        s = raw_input()
    except EOFError:
        break


    try:
        field = string.strip(string.split(s, "\t")[col])
    except IndexError:
        sys.stderr.write("PROCESS ERROR\n")
        continue


    fs = string.split(field)


    if len(field) > 2:
        date = int("%s%s%02i" % (fs[2], month[fs[0]], int(fs[1][:-1])))
        if date> int(sys.argv[2]):
            print s
```

## C.3.2   The bin count script.

```
import string
import os
import sys


vals=[0]*150


col = int(sys.argv[1])
bin = int(sys.argv[2])


s = "foo"


while True:
    try:
        s = raw_input()
    except EOFError:
        break
    try:
```

```
            field = string.split(s, "\t")[col]
        except IndexError:
            print "PROCESS ERROR"
            continue


        if field == "one":
            field = "1"
        if field == "":
            continue


        try:
            fval = int(field)
        except ValueError:
            print "ERROR:", field


        try:
            vals[fval/10] += 1
        except IndexError:
            print len(vals)
            print "ERROR:" + str(fval)


if int(sys.argv[2]) == 1:
    for k in vals:
        print k
```

### C.3.3   The bin date script.

```
import string
import sys


# usage: bindate col
# col = number of column to use MUST BE A DATE COLUMN
# bindate prints the number of records where
# column #col = January 2004, then February 2004, etc.


col = int(sys.argv[1])
```

```python
s = "foo"

month={"January":"01",
       "February":"02",
       "March":"03",
       "April":"04",
       "May":"05",
       "June":"06",
       "July":"07",
       "August":"08",
       "September":"09",
       "October":"10",
       "November":"11",
       "December":"12",}

year={
    "2004": 0,
    "2005": 1}

bins=[0]*24

while True:
    try:
        s = raw_input()
    except EOFError:
        break

    try:
        field = string.strip(string.split(s, "\t")[col])
    except IndexError:
        sys.stderr.write("PROCESS ERROR\n")
        continue

    fs = string.split(field)

    if len(field) > 2:
        bins[year[fs[2]]*12 + int(month[fs[0]])-1] += 1
```

```
for x in range(len(bins)):
    y = str(2004 + x/12)
    m = str((x % 12) + 1)
    print bins[x]
#    print "%s/%s\t%i" % (m, y, bins[x])
```

### C.3.4   The count number script.

```
import string
import os
import sys


# countnumber col printall
# Countnumber reads from stdin and generates a histogram of the column
# col = the column to read from
# printall = whether to print each individual value



vals={}

col = int(sys.argv[1])

s = "foo"

n = 0

while True:
    try:
        s = raw_input()
    except EOFError:
        break
    try:
        field = string.split(s, "\t")[col]
    except IndexError:
        print "PROCESS ERROR"
        continue
```

```
    if n % 500 == 0:
        print field
    if field in vals.keys():
        vals[field]+=1
    else:
        vals[field] = 1
    n += 1


if int(sys.argv[2]) == 1:
    for k in vals.keys():
        print k, "\t", vals[k]


if " " in vals.keys():
    print "BLANK : ",  vals[" "]
    print "NOTBLANK : ", n - vals[" "]
    print "TOTAL : ", n
```

### C.3.5 The filter field script.

```
import string
import sys


# usage: python filterfield.py col val
# if col is equal to val, print this record
# otherwise, do nothing


col = int(sys.argv[1])
val = string.strip(sys.argv[2])


s = "foo"


while True:
    try:
        s = raw_input()
    except EOFError:
        break
```

```python
    try:
        field = string.strip(string.split(s, "\t")[col])
    except IndexError:
        sys.stderr.write("PROCESS ERROR\n")
        continue

    if field == val:
        print s
```

### C.3.6   The greater than script.

```python
import string
import os
import sys

vals=[0]*150

col = int(sys.argv[1])
val = int(sys.argv[2])

s = "foo"

while True:
    try:
        s = raw_input()
    except EOFError:
        break
    try:
        field = string.split(s, "\t")[col]
    except IndexError:
        print "PROCESS ERROR"
        continue

    if field == "one":
        field = "1"
    if field == "":
```

```
        continue

try:
    fval = int(field)
except ValueError:
    print "ERROR:", field

try:
    if fval > val:
        print s
except IndexError:
    print len(vals)
    print "ERROR:" + str(fval)
```

| Which gender describes you best? n=419 | | |
|---|---|---|
| | Number | Percentage |
| No Response | 9 | 3% |
| Male | 186 | 44% |
| Female | 224 | 53% |

Which gender describes you best? n=419
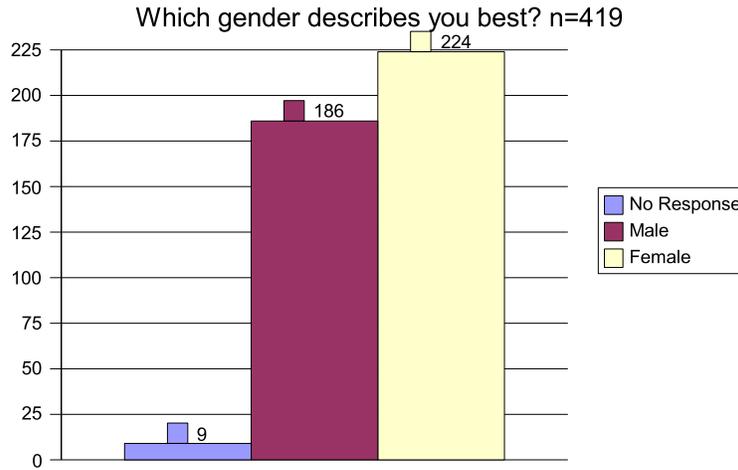


Figure 11: Gender of survey takers

# D    Supplemental Data

In this section, we included the numerical results of the numerous analyses we performed on the data we collected from users and directly from Facebook. We referred to many, but not all, of these figures earlier. This data is useful alone in looking for trends and correlations that did not find their way into this paper.

| Which best describes your living arrangements? n=419 | | |
|---|---|---|
| House | Number Responding | Percentage |
| No Response | 45 | 10.74% |
| Alpha Chi Omega | 1 | 0.24% |
| Alpha Epsilon Phi | 1 | 0.24% |
| Alpha Phi | 4 | 0.95% |
| Baker House | 4 | 0.95% |
| Beta Theta Pi | 1 | 0.24% |
| Bexley Hall | 2 | 0.48% |
| Burton Conner House | 87 | 20.76% |
| Chi Phi | 2 | 0.48% |
| East Campus | 107 | 25.54% |
| Kappa Alpha Theta | 1 | 0.24% |
| Kappa Sigma | 2 | 0.48% |
| Lambda Chi Alpha | 2 | 0.48% |
| MacGregor House | 9 | 2.15% |
| McCormick Hall | 2 | 0.48% |
| New House | 3 | 0.72% |
| Next House | 4 | 0.95% |
| No. 6 | 1 | 0.24% |
| Phi Delta Theta | 2 | 0.48% |
| Phi Kappa Sigma | 1 | 0.24% |
| Phi Kappa Theta | 1 | 0.24% |
| Pi Lambda Phi | 1 | 0.24% |
| Pika | 1 | 0.24% |
| Random Hall | 42 | 10.02% |
| Senior House | 6 | 1.43% |
| Sidney-Pacific | 1 | 0.24% |
| Sigma Alpha Epsilon | 1 | 0.24% |
| Sigma Chi | 1 | 0.24% |
| Sigma Kappa | 1 | 0.24% |
| Sigma Nu | 1 | 0.24% |
| Simmons Hall | 63 | 15.04% |
| Tau Epsilon Phi | 7 | 1.67% |
| Theta Xi | 1 | 0.24% |
| WILG | 10 | 2.39% |
| Zeta Beta Tau | 1 | 0.24% |

Figure 12: Chart of survey takers over dorms and ILGs.
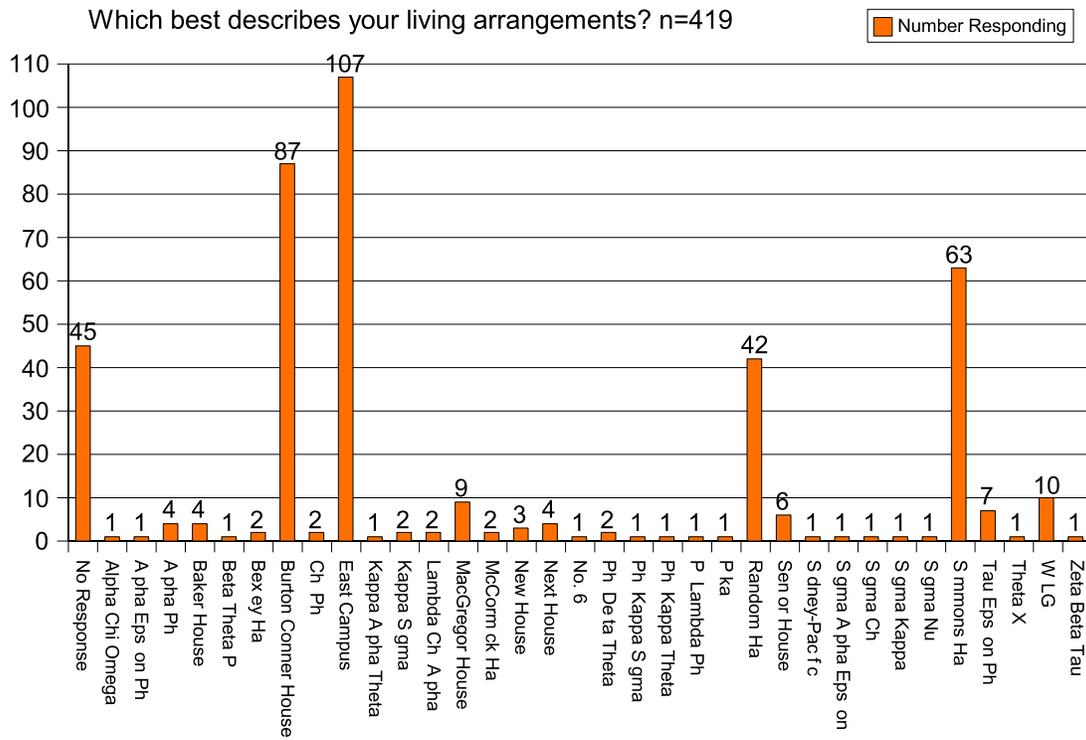
Figure 13: Distribution of survey takers over dorms and ILGs.

### What is your student status? n=419

|              | Number | Percentage |
| ------------ | ------ | ---------- |
| No Answer    | 10     | 2.39%      |
| Undergrad    | 380    | 90.69%     |
| Grad Student | 13     | 3.1%       |
| Alumnus      | 14     | 3.34%      |

Figure 14: Status of survey takers

| Facebook Logins Per Week n=371 | | | | |
|---|---|---|---|---|
| | Number | Percentage | Number Male | Number Female |
| 1 to 3 | 139 | 37.47% | 66 | 70 |
| 4 to 8 | 95 | 25.61% | 36 | 57 |
| 9 to 15 | 64 | 17.25% | 27 | 37 |
| 20 to 30 | 40 | 10.78% | 22 | 18 |
| 31 or more | 33 | 8.89% | 11 | 10 |

Approximately how many times a week do
you log into Facebook? n=371
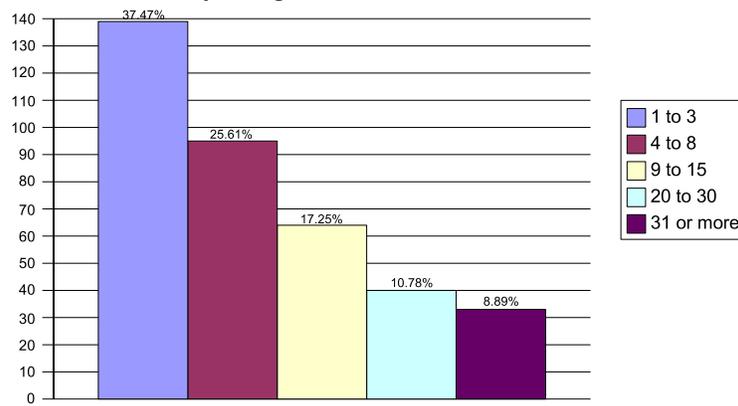
Logins per week, by gender
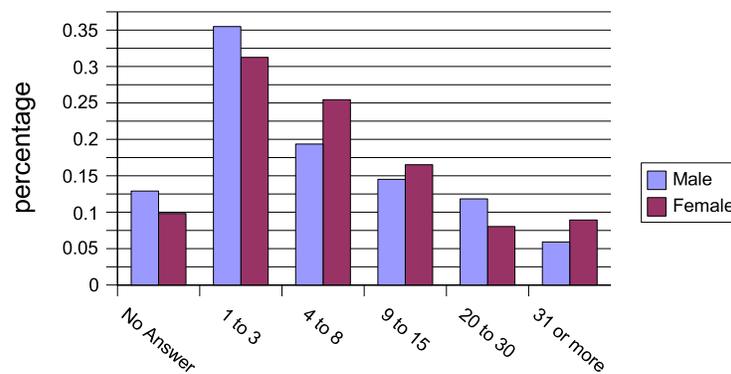
186 males, 224 females

Figure 15: Logins per week

| Number of friends n=378 | | | | |
| --- | --- | --- | --- | --- |
| | Number | Percentage | Males | Females |
| 1 to 10 | 5 | 1.32% | 3 | 2 |
| 11 to 50 | 56 | 14.81% | 31 | 23 |
| 51 to 100 | 117 | 30.95% | 54 | 62 |
| 101 to 200 | 143 | 37.83% | 58 | 84 |
| 201 to 349 | 49 | 12.96% | 15 | 33 |
| 350 or more | 8 | 2.12% | 4 | 2 |

Approximately how many Facebook friends do you have from your school? n=378
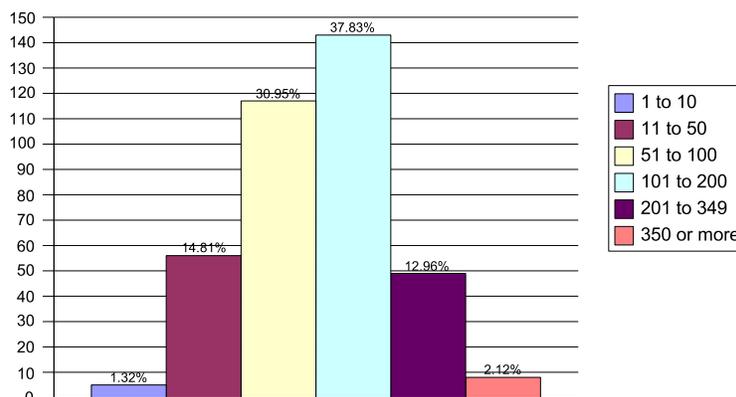


Number of Friends, by gender
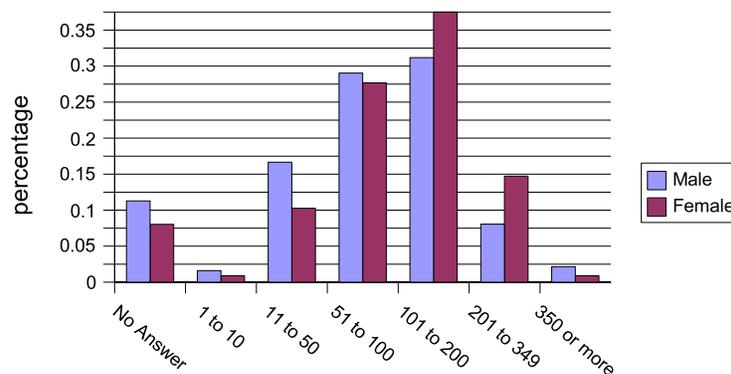


186 males, 224 females

Figure 16: Number of Friends at MIT

| Percentage of friends from MIT n=372 | | | | |
|---|---|---|---|---|
| | Number | Percentage | Males | Females |
| 1-15% | 5 | 1.34% | 2 | 3 |
| 16-33% | 43 | 11.56% | 20 | 23 |
| 34-50% | 107 | 28.76% | 56 | 49 |
| 51-75% | 174 | 46.77% | 72 | 101 |
| 76-100% | 43 | 11.56% | 12 | 28 |

Approximately what percentage of all of your friends
are from your school? n=372
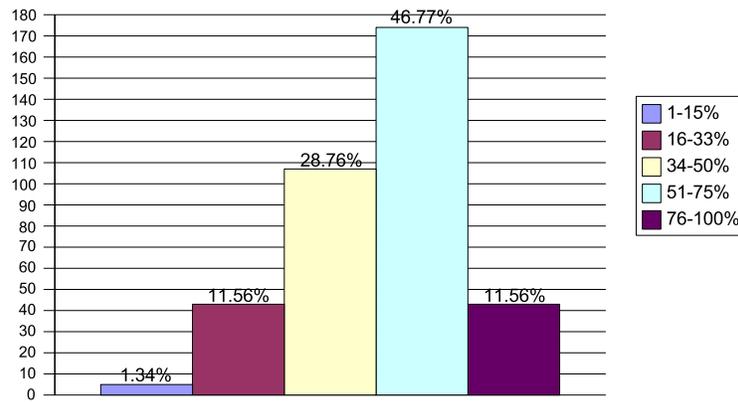


Percentage of fr ends from M T, by gender
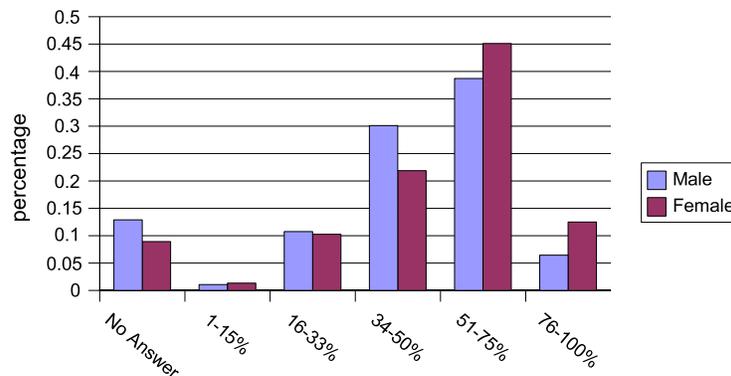


186 males, 224 females

Figure 17: Percentage of Friends from MIT

| Number Allowing Strangers To Friend n=383 | | | | |
|---|---|---|---|---|
| | Number | Percentage | Males | Females |
| No | 243 | 63.45% | 109 | 129 |
| Yes | 30 | 7.83% | 17 | 12 |
| Sometimes | 110 | 28.72% | 44 | 65 |

Do you ever friend people whom you
have never met in person? n=383



Percentage Friending Strangers, by gender
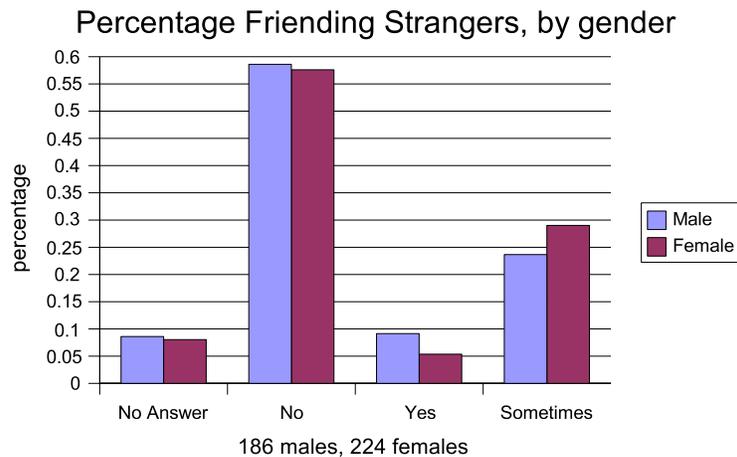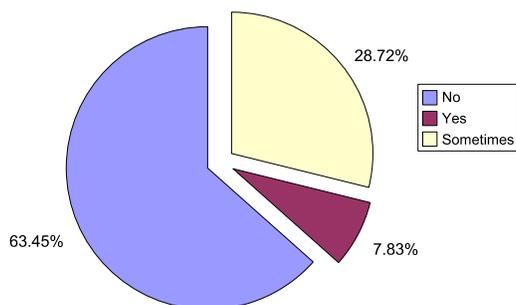


Figure 18: Analysis of users friending strangers on Facebook

| Facebook and My Privacy: Familiarity and Utilization n=419 | | | | | | |
|---|---|---|---|---|---|---|
| | Number Familiar | Males | Females | Number Using | Males | Females |
| No Answer | 30 | 15 | 33 | 39 | 18 | 19 |
| No | 100 | 38 | 59 | 234 | 111 | 119 |
| Yes | 289 | 133 | 152 | 146 | 57 | 86 |

Are you familiar with Facebook's "My Privacy" feature, that lets
you control who may view your profile? Do you use the
"My Privacy" feature to control who may view your profile? n=419
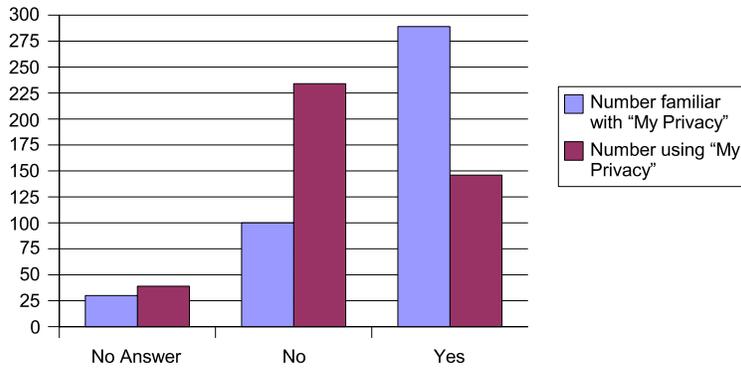


Fam  ar ty w th My Pr vacy, by gender
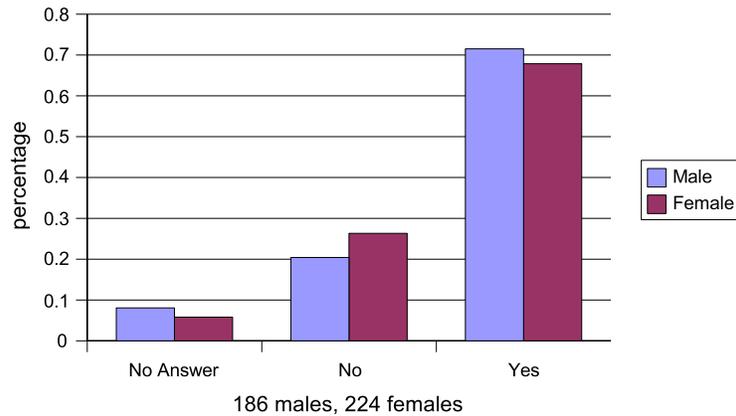


Utilization of My Privacy, by gender
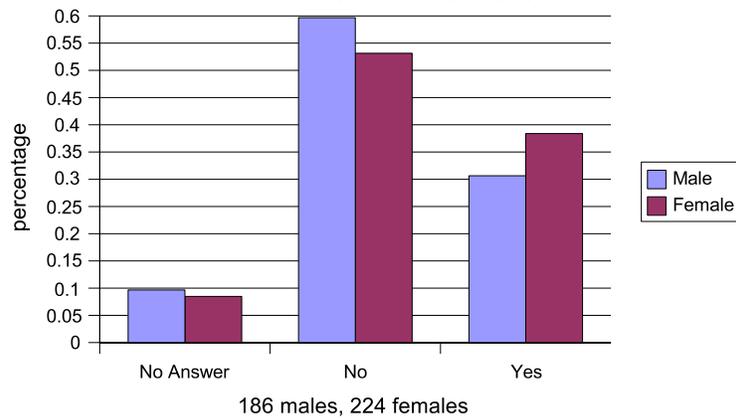


Figure 19: My Privacy, and knoweldge and utilization thereof

| How concerned are you about Facebook and privacy?  n=329 | | | | |
|---|---|---|---|---|
| | Number | Percentage | Males | Females |
| Not at all | 76 | 23.1% | 43 | 31 |
| Barely | 117 | 35.56% | 43 | 71 |
| Somewhat | 104 | 31.61% | 39 | 64 |
| Quite | 20 | 6.08% | 7 | 12 |
| Very Concerned | 12 | 3.65% | 7 | 5 |

How concerned are you about the privacy of
your data on Facebook? n=329

Concerns with Facebook Security, by gender

186 males, 224 females
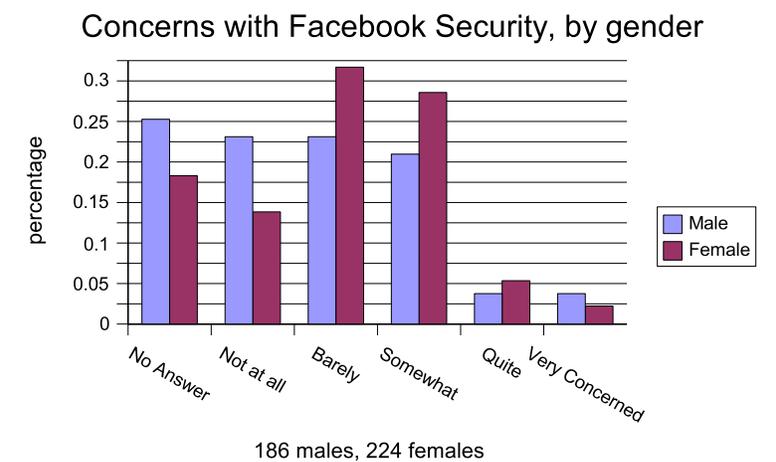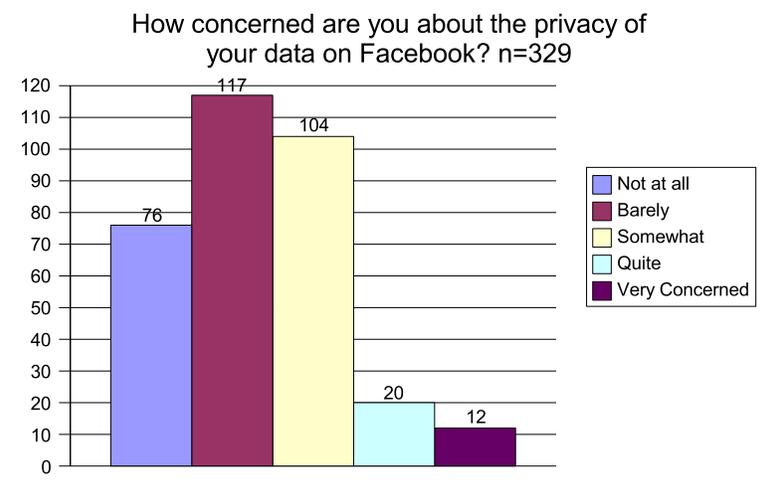
Figure 20: Concern for Facebook Privacy

| Reading of Facebook Terms of Service and Privacy Policy n=419 | | | | |
|---|---|---|---|---|
|  | Read TOS? | Percentage | Read PP? | Percentage |
| No Answer | 30 | 7.16 % | 29 | 6.92 % |
| No | 353 | 84.25 % | 347 | 82.82 % |
| Yes | 36 | 8.59 % | 43 | 10.26 % |

Have you read Facebook's
Terms of Service? n=389

9.25%

No
Yes

90.75%
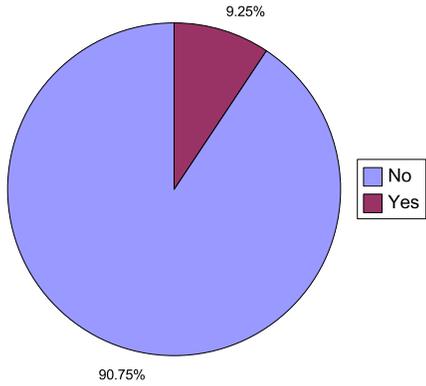
Have you read Facebook's
Pr vacy Po cy? n  390

11.03%

No
Yes

88.97%
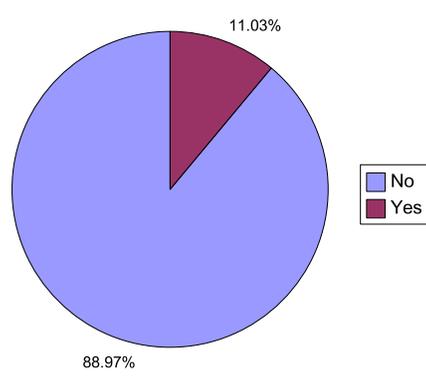
Figure 21: Most users do not read the policies that regulate their Facebook use.

Have you read Facebook's
Terms of Service? n=389

9.25%

No
Yes

90.75%

Can Facebook Share Information? n=419

|  | Number Responding | Percentage |
|---|---|---|
| No Answer | 45 | 10.74 % |
| No | 174 | 41.53 % |
| Yes | 200 | 47.73 % |

Can Facebook share your information
with other companies? n=419

10.74%

47.73%

No Answer
No
Yes

41.53%
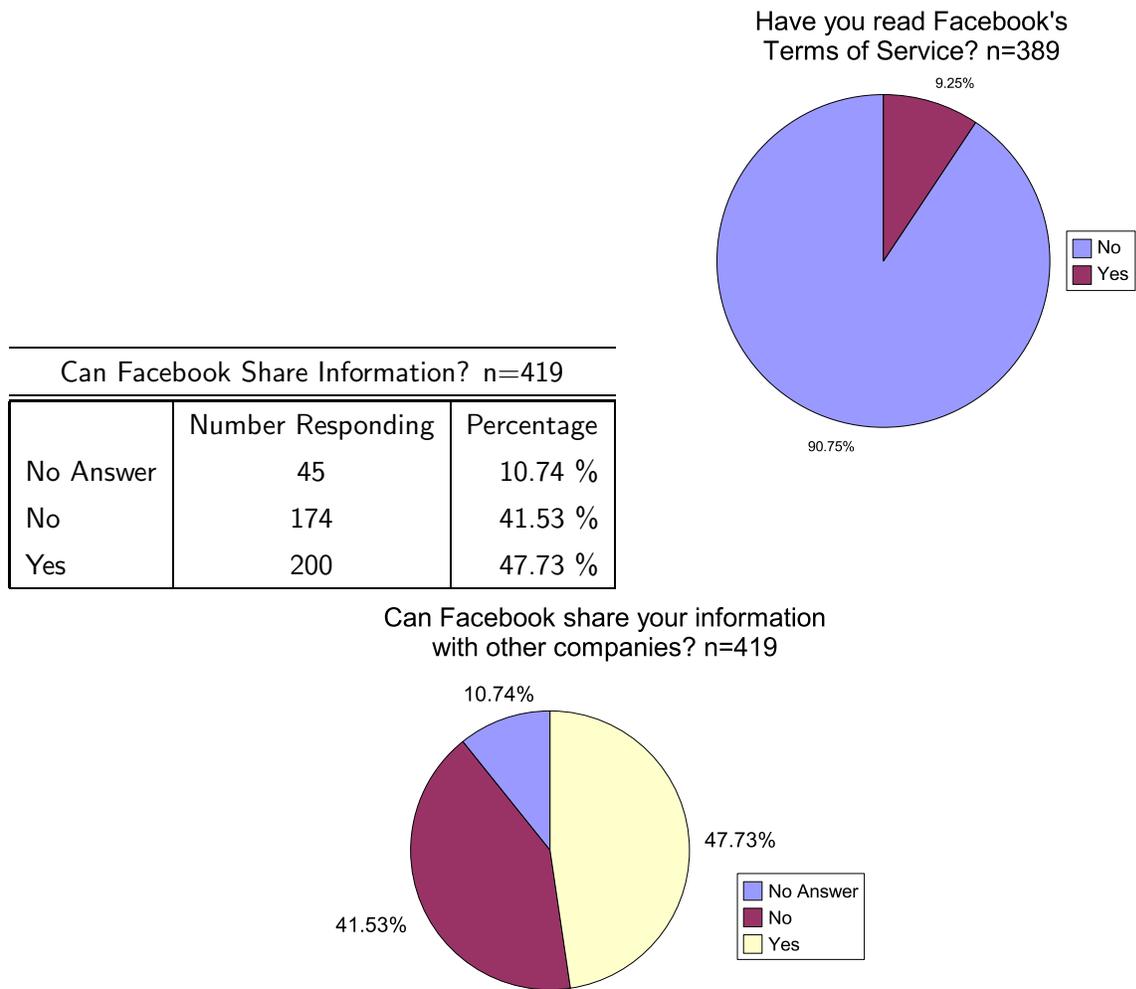
Figure 22: Users are split on whether or not Facebook can share your information with other companies, indicating a guess.

| Familiarity with "My Photo" feature and policies. n=419 | | | | |
|---|---|---|---|---|
|  | Familiar | Percentage | Can you restrict access? | Percentage |
| No Answer | 30 | 7.16% | 84 | 20.05% |
| No | 47 | 11.22% | 139 | 33.17% |
| Yes | 342 | 81.62% | 196 | 46.78% |

Figure 23: Are you familiar with "My Photo?" Can you restrict access to it?

| Does Facebook do an adequate job in protecting your privacy? n=419 | | | | |
|---|---|---|---|---|
| | Number | Percentage | Males | Females |
| No Answer | 102 | 24.34% | 48 | 50 |
| No | 139 | 33.17% | 67 | 68 |
| Yes | 177 | 42.24% | 70 | 106 |

### Does Facebook do an adequate job in protecting your privacy? n=419



### Sat sfact on w th Facebook Pr vacy, by gender



186 males, 224 females
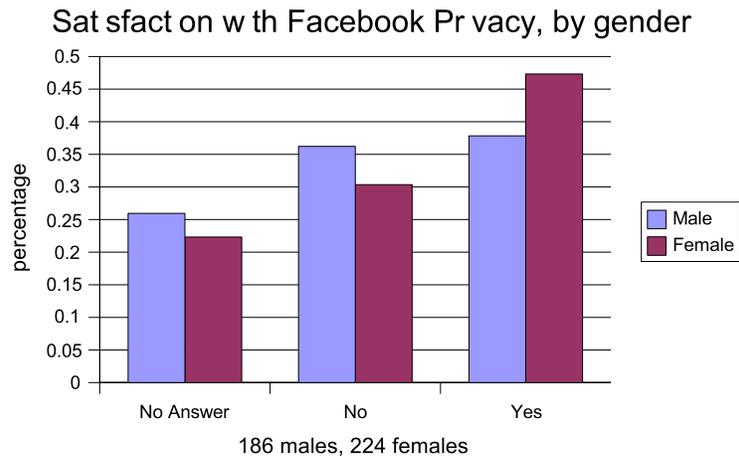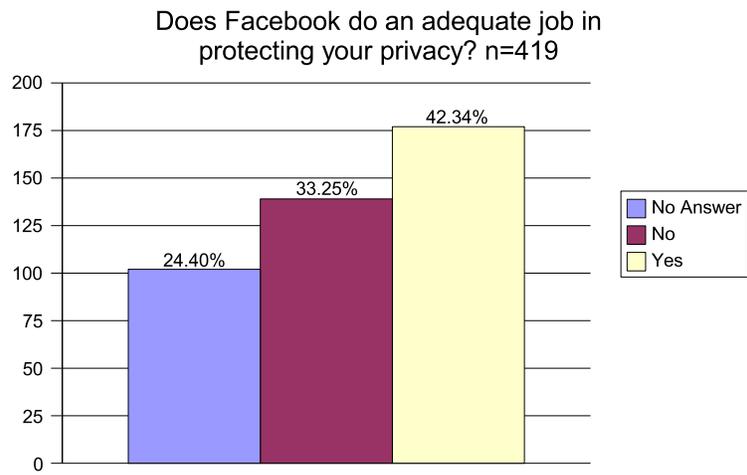
Figure 24: Users show indifference and approval for Facebook's security practices.

| Distributions Of Facebook User Categories At Four Universities | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | MIT | | Oklahoma | | NYU | | Harvard | |
| Size | 8023 | | 19910 | | 24696 | | 17750 | |
| **Number Reporting Gender: Distribution** | | | | | | | | |
| Males | 3868 | 48.21% | 8863 | 44.52% | 8689 | 35.18% | 7461 | 42.03% |
| Females | 2483 | 30.95% | 8814 | 44.27% | 12118 | 49.07% | 5940 | 33.46% |
| **Class Distribution: Graduating class of year indicated, self reported.** | | | | | | | | |
| 2003 | 189 | 2.36% | 78 | 0.39% | 200 | 0.81% | 876 | 4.94% |
| 2004 | 539 | 6.72% | 630 | 3.16% | 961 | 3.89% | 1351 | 7.61% |
| 2005 | 762 | 9.5% | 2224 | 11.17% | 2643 | 10.7% | 1605 | 9.04% |
| 2006 | 878 | 10.94% | 2952 | 14.83% | 3353 | 13.58% | 1657 | 9.34% |
| 2007 | 948 | 11.82% | 3039 | 15.26% | 3850 | 15.59% | 1710 | 9.63% |
| 2008 | 1016 | 12.66% | 3151 | 15.83% | 4012 | 16.25% | 1785 | 10.06% |
| 2009 | 921 | 11.48% | 2690 | 13.51% | 4076 | 16.5% | 1583 | 8.92% |
| 2010 | 93 | 1.16% | 162 | 0.81% | 60 | 0.24% | 132 | 0.74% |
| Other | 2677 | 33.37% | 4984 | 25.03% | 5541 | 22.44% | 7051 | 39.72% |
| **User Distribution: Kinds of Users at each school. ("Undergraduate" unique to OU.)** | | | | | | | | |
| Alumnus/Alumna | 2226 | 27.75% | 2662 | 13.37% | 4730 | 19.15% | 7010 | 39.49% |
| Faculty | 76 | 0.95% | 81 | 0.41% | 183 | 0.74% | 208 | 1.17% |
| Grad Student | 845 | 10.53% | 1312 | 6.59% | 1511 | 6.12% | 1933 | 10.89% |
| Staff | 161 | 2.01% | 188 | 0.94% | 187 | 0.76% | 438 | 2.47% |
| Student | 4702 | 58.61% | 10406 | 52.27% | 18055 | 73.11% | 8085 | 45.55% |
| Summer Student | 10 | 0.12% | 4 | 0.02% | 26 | 0.11% | 27 | 0.15% |
| Undergraduate | – | – | 5239 | 26.31% | – | – | – | – |

Figure 25: Summary of Facebook usage statistics at four schools: the Massachusetts Institute of Technology, University of Oklahoma, New York University, and Harvard University.

| Willingness to Share Personal Information at each school. | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| All Students | MIT | | Oklahoma | | NYU | | Harvard | |
| Residence | 5172 | 64.46 % | 7190 | 36.11 % | 11582 | 46.9 % | 4260 | 24 % |
| High School | 5252 | 65.46 % | 16133 | 81.03 % | 18359 | 74.34 % | 7270 | 40.96 % |
| Screen Name | 4341 | 54.11 % | 10860 | 54.55 % | 16157 | 65.42 % | 8186 | 46.12 % |
| Mobile | 1700 | 21.19 % | 2637 | 13.24 % | 3443 | 13.94 % | 8582 | 48.35 % |
| Interests | 4453 | 55.5 % | 15099 | 75.84 % | 16473 | 66.7 % | 8607 | 48.49 % |
| Clubs/Jobs | 3400 | 42.38 % | 13170 | 66.15 % | 12426 | 50.32 % | 8758 | 49.34 % |
| Music | 4236 | 52.8 % | 15608 | 78.39 % | 16470 | 66.69 % | 9116 | 51.36 % |
| Movies | 4084 | 50.9 % | 15255 | 76.62 % | 16218 | 65.67 % | 10694 | 60.25 % |
| Books | 3956 | 49.31 % | 13626 | 68.44 % | 15427 | 62.47 % | 11271 | 63.5 % |
| Gender | 6351 | 79.16 % | 17677 | 88.78 % | 20807 | 84.25 % | 13401 | 75.5 % |
| After 10/1/05 | MIT | | Oklahoma | | NYU | | Harvard | |
| Residence | 3309 | 80.71 % | 6316 | 40.48 % | 9601 | 58.59 % | 7466 | 79.5 % |
| High School | 3433 | 83.73 % | 13841 | 88.71 % | 14341 | 87.51 % | 7613 | 81.07 % |
| Screen Name | 2890 | 70.49 % | 9396 | 60.22 % | 12627 | 77.05 % | 5965 | 63.52 % |
| Mobile | 1159 | 28.27 % | 2228 | 14.28 % | 2698 | 16.46 % | 3100 | 33.01 % |
| Interests | 2996 | 73.07 % | 13075 | 83.8 % | 13047 | 79.62 % | 6661 | 70.93 % |
| Clubs/Jobs | 2373 | 57.88 % | 11562 | 74.1 % | 9839 | 60.04 % | 5452 | 58.06 % |
| Music | 2894 | 70.59 % | 13564 | 86.93 % | 13091 | 79.89 % | 6457 | 68.76 % |
| Movies | 2808 | 68.49 % | 13251 | 84.93 % | 16387 | 100 % | 6295 | 67.03 % |
| Books | 2710 | 66.1 % | 11848 | 75.93 % | 12216 | 74.55 % | 6293 | 67.01 % |
| Gender | 3817 | 93.1 % | 14906 | 95.53 % | 15479 | 94.46 % | 8497 | 90.48 % |
| Total | 4100 | 100 % | 15603 | 100 % | 16387 | 100 % | 9391 | 100 % |

Figure 26: Willingness of Facebook users to disclose personal information on the service, at four schools, showing all users and only those who have updated their profiles on or after October 1, 2005.

| Willingness to Share Personal Information at each school, by gender. | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Males | MIT | | Oklahoma | | NYU | | Harvard | |
| Residence | 3005 | 77.69 % | 3377 | 38.1 % | 4536 | 52.2 % | 5804 | 77.79 % |
| High School | 2979 | 77.02 % | 7661 | 86.44 % | 7066 | 81.32 % | 5479 | 73.44 % |
| Screen Name | 2514 | 64.99 % | 5309 | 59.9 % | 6374 | 73.36 % | 4224 | 56.61 % |
| Mobile | 1147 | 29.65 % | 1859 | 20.97 % | 1930 | 22.21 % | 2461 | 32.98 % |
| Interests | 2580 | 66.7 % | 7888 | 88.99 % | 6468 | 74.44 % | 4680 | 62.73 % |
| Clubs/Jobs | 1941 | 50.18 % | 6168 | 69.59 % | 4897 | 56.36 % | 3770 | 50.53 % |
| Music | 2470 | 63.86 % | 7471 | 84.29 % | 6513 | 74.96 % | 4572 | 61.28 % |
| Movies | 2335 | 60.37 % | 7223 | 81.5 % | 6369 | 73.3 % | 4439 | 59.5 % |
| Books | 2244 | 58.01 % | 6418 | 72.41 % | 5960 | 68.59 % | 4410 | 59.11 % |
| Gender | 3868 | 100 % | 8863 | 100 % | 8689 | 100 % | 7461 | 100 % |
| Females | MIT | | Oklahoma | | NYU | | Harvard | |
| Residence | 2003 | 80.67 % | 3609 | 40.95 % | 6736 | 55.59 % | 4852 | 81.68 % |
| High School | 2083 | 83.89 % | 7964 | 90.36 % | 10631 | 87.73 % | 4577 | 77.05 % |
| Screen Name | 1667 | 67.14 % | 5200 | 59 % | 9103 | 75.12 % | 3474 | 58.48 % |
| Mobile | 510 | 20.54 % | 710 | 8.06 % | 1407 | 11.61 % | 1577 | 26.55 % |
| Interests | 1661 | 66.89 % | 7211 | 81.81 % | 9276 | 76.55 % | 3763 | 63.35 % |
| Clubs/Jobs | 1325 | 53.36 % | 6497 | 73.71 % | 7032 | 58.03 % | 3064 | 51.58 % |
| Music | 1595 | 64.24 % | 7540 | 85.55 % | 9289 | 76.65 % | 3624 | 61.01 % |
| Movies | 1594 | 64.2 % | 7447 | 84.49 % | 9233 | 76.19 % | 3599 | 60.59 % |
| Books | 1550 | 62.42 % | 6693 | 75.94 % | 8846 | 73 % | 3635 | 61.2 % |
| Gender | 2483 | 100 % | 8814 | 100 % | 12118 | 100 % | 5940 | 100 % |

Figure 27: Willingness of Facebook users to disclose personal information on the service, at four schools, by gender.

| When Users Join And Update Facebook at MIT | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Month Of | Join | | Update | | 2007 Join | | 2008 Join | | 2009 Join | |
| Mar 1, 04 | 1087 | 13.55 % | 0 | 0 % | 320 | 33.76 % | 3 | 0.3 % | 0 | 0 % |
| Apr 1, 04 | 879 | 10.96 % | 0 | 0 % | 195 | 20.57 % | 9 | 0.89 % | 0 | 0 % |
| May 1, 04 | 601 | 7.49 % | 0 | 0 % | 83 | 8.76 % | 98 | 9.65 % | 0 | 0 % |
| Jun 1, 04 | 329 | 4.1 % | 0 | 0 % | 21 | 2.22 % | 143 | 14.07 % | 1 | 0.11 % |
| Jul 1, 04 | 340 | 4.24 % | 18 | 0.26 % | 18 | 1.9 % | 198 | 19.49 % | 4 | 0.43 % |
| Aug 1, 04 | 392 | 4.89 % | 22 | 0.32 % | 37 | 3.9 % | 196 | 19.29 % | 2 | 0.22 % |
| Sep 1, 04 | 403 | 5.02 % | 39 | 0.57 % | 27 | 2.85 % | 165 | 16.24 % | 1 | 0.11 % |
| Oct 1, 04 | 274 | 3.42 % | 51 | 0.75 % | 26 | 2.74 % | 64 | 6.3 % | 1 | 0.11 % |
| Nov 1, 04 | 240 | 2.99 % | 60 | 0.88 % | 20 | 2.11 % | 30 | 2.95 % | 0 | 0 % |
| Dec 1, 04 | 230 | 2.87 % | 67 | 0.98 % | 21 | 2.22 % | 21 | 2.07 % | 3 | 0.33 % |
| Jan 1, 05 | 245 | 3.05 % | 62 | 0.91 % | 27 | 2.85 % | 5 | 0.49 % | 2 | 0.22 % |
| Feb 1, 05 | 226 | 2.82 % | 99 | 1.45 % | 21 | 2.22 % | 10 | 0.98 % | 1 | 0.11 % |
| Mar 1, 05 | 196 | 2.44 % | 94 | 1.38 % | 14 | 1.48 % | 9 | 0.89 % | 1 | 0.11 % |
| Apr 1, 05 | 184 | 2.29 % | 101 | 1.48 % | 12 | 1.27 % | 11 | 1.08 % | 5 | 0.54 % |
| May 1, 05 | 515 | 6.42 % | 185 | 2.71 % | 13 | 1.37 % | 7 | 0.69 % | 322 | 34.96 % |
| Jun 1, 05 | 400 | 4.99 % | 250 | 3.67 % | 15 | 1.58 % | 5 | 0.49 % | 211 | 22.91 % |
| Jul 1, 05 | 336 | 4.19 % | 252 | 3.7 % | 11 | 1.16 % | 2 | 0.2 % | 142 | 15.42 % |
| Aug 1, 05 | 378 | 4.71 % | 482 | 7.07 % | 12 | 1.27 % | 14 | 1.38 % | 155 | 16.83 % |
| Sep 1, 05 | 335 | 4.18 % | 907 | 13.3 % | 24 | 2.53 % | 16 | 1.57 % | 44 | 4.78 % |
| Oct 1, 05 | 285 | 3.55 % | 1638 | 24.02 % | 21 | 2.22 % | 7 | 0.69 % | 22 | 2.39 % |
| Nov 1, 05 | 146 | 1.82 % | 2493 | 36.55 % | 10 | 1.05 % | 3 | 0.3 % | 4 | 0.43 % |
| Total | 8021 | 100 % | 6820 | 85.03 % | 948 | 11.82 % | 1016 | 12.67 % | 921 | 11.48 % |

Figure 28: Facebook usage data for the Massachusetts Institute of Technology.

| When Users Join And Update Facebook at U. Oklahoma | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Month Of | Join | | Update | | 2007 Join | | 2008 Join | | 2009 Join | |
| Aug 1, 04 | 1 | 0.01 % | 0 | 0 % | 1 | 0.03 % | 0 | 0 % | 0 | 0 % |
| Sep 1, 04 | 448 | 2.25 % | 5 | 0.03 % | 141 | 4.64 % | 131 | 4.16 % | 3 | 0.11 % |
| Oct 1, 04 | 966 | 4.86 % | 4 | 0.02 % | 254 | 8.36 % | 316 | 10.03 % | 3 | 0.11 % |
| Nov 1, 04 | 3908 | 19.65 % | 38 | 0.2 % | 813 | 26.75 % | 1089 | 34.56 % | 24 | 0.89 % |
| Dec 1, 04 | 2723 | 13.69 % | 79 | 0.42 % | 458 | 15.07 % | 432 | 13.71 % | 21 | 0.78 % |
| Jan 1, 05 | 1388 | 6.98 % | 68 | 0.36 % | 218 | 7.17 % | 188 | 5.97 % | 24 | 0.89 % |
| Feb 1, 05 | 1411 | 7.09 % | 95 | 0.51 % | 208 | 6.84 % | 183 | 5.81 % | 40 | 1.49 % |
| Mar 1, 05 | 836 | 4.2 % | 122 | 0.65 % | 107 | 3.52 % | 86 | 2.73 % | 37 | 1.38 % |
| Apr 1, 05 | 1008 | 5.07 % | 151 | 0.81 % | 122 | 4.01 % | 109 | 3.46 % | 97 | 3.61 % |
| May 1, 05 | 862 | 4.33 % | 223 | 1.19 % | 103 | 3.39 % | 83 | 2.63 % | 196 | 7.29 % |
| Jun 1, 05 | 905 | 4.55 % | 179 | 0.96 % | 71 | 2.34 % | 71 | 2.25 % | 414 | 15.39 % |
| Jul 1, 05 | 1117 | 5.62 % | 274 | 1.47 % | 75 | 2.47 % | 73 | 2.32 % | 650 | 24.16 % |
| Aug 1, 05 | 1631 | 8.2 % | 564 | 3.02 % | 127 | 4.18 % | 131 | 4.16 % | 805 | 29.93 % |
| Sep 1, 05 | 1237 | 6.22 % | 1242 | 6.65 % | 174 | 5.73 % | 134 | 4.25 % | 259 | 9.63 % |
| Oct 1, 05 | 1083 | 5.44 % | 3329 | 17.82 % | 130 | 4.28 % | 99 | 3.14 % | 96 | 3.57 % |
| Nov 1, 05 | 369 | 1.85 % | 12311 | 65.89 % | 37 | 1.22 % | 26 | 0.83 % | 21 | 0.78 % |
| Total | 19893 | 100 % | 18684 | 93.92 % | 3039 | 15.28 % | 3151 | 15.84 % | 2690 | 13.52 % |

Figure 29: Facebook usage data for the University of Oklahoma.

| When Users Join And Update Facebook at NYU | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Month Of | Join | | Update | | 2007 Join | | 2008 Join | | 2009 Join | |
| Mar 1, 04 | 667 | 2.7 % | 0 | 0 % | 348 | 9.04 % | 3 | 0.07 % | 0 | 0 % |
| Apr 1, 04 | 3350 | 13.57 % | 0 | 0 % | 1287 | 33.43 % | 18 | 0.45 % | 5 | 0.12 % |
| May 1, 04 | 1868 | 7.56 % | 0 | 0 % | 338 | 8.78 % | 218 | 5.43 % | 3 | 0.07 % |
| Jun 1, 04 | 785 | 3.18 % | 3 | 0.01 % | 75 | 1.95 % | 230 | 5.73 % | 1 | 0.02 % |
| Jul 1, 04 | 968 | 3.92 % | 18 | 0.08 % | 72 | 1.87 % | 566 | 14.11 % | 1 | 0.02 % |
| Aug 1, 04 | 1509 | 6.11 % | 24 | 0.11 % | 138 | 3.58 % | 957 | 23.85 % | 3 | 0.07 % |
| Sep 1, 04 | 1672 | 6.77 % | 54 | 0.24 % | 229 | 5.95 % | 736 | 18.34 % | 1 | 0.02 % |
| Oct 1, 04 | 1396 | 5.65 % | 98 | 0.44 % | 217 | 5.64 % | 382 | 9.52 % | 3 | 0.07 % |
| Nov 1, 04 | 1236 | 5.01 % | 143 | 0.64 % | 142 | 3.69 % | 209 | 5.21 % | 4 | 0.1 % |
| Dec 1, 04 | 958 | 3.88 % | 161 | 0.72 % | 111 | 2.88 % | 96 | 2.39 % | 3 | 0.07 % |
| Jan 1, 05 | 813 | 3.29 % | 169 | 0.76 % | 132 | 3.43 % | 69 | 1.72 % | 2 | 0.05 % |
| Feb 1, 05 | 692 | 2.8 % | 177 | 0.8 % | 82 | 2.13 % | 58 | 1.45 % | 0 | 0 % |
| Mar 1, 05 | 769 | 3.11 % | 222 | 1 % | 63 | 1.64 % | 46 | 1.15 % | 179 | 4.39 % |
| Apr 1, 05 | 1019 | 4.13 % | 278 | 1.25 % | 73 | 1.9 % | 52 | 1.3 % | 429 | 10.53 % |
| May 1, 05 | 1489 | 6.03 % | 477 | 2.15 % | 89 | 2.31 % | 82 | 2.04 % | 839 | 20.58 % |
| Jun 1, 05 | 1319 | 5.34 % | 480 | 2.16 % | 79 | 2.05 % | 60 | 1.5 % | 850 | 20.85 % |
| Jul 1, 05 | 1248 | 5.05 % | 526 | 2.37 % | 60 | 1.56 % | 51 | 1.27 % | 800 | 19.63 % |
| Aug 1, 05 | 1187 | 4.81 % | 998 | 4.49 % | 106 | 2.75 % | 71 | 1.77 % | 621 | 15.24 % |
| Sep 1, 05 | 955 | 3.87 % | 1923 | 8.66 % | 127 | 3.3 % | 65 | 1.62 % | 251 | 6.16 % |
| Oct 1, 05 | 664 | 2.69 % | 4776 | 21.5 % | 71 | 1.84 % | 36 | 0.9 % | 69 | 1.69 % |
| Nov 1, 05 | 131 | 0.53 % | 11686 | 52.61 % | 11 | 0.29 % | 7 | 0.17 % | 12 | 0.29 % |
| Total | 24695 | 100 % | 22213 | 89.95 % | 3850 | 15.59 % | 4012 | 16.25 % | 4076 | 16.51 % |

Figure 30: Facebook usage data for New York University.

# E  Selected Survey Comments

The paper and web form survey we gave to users provided space for user feedback. The feedback we received was insightful. Of 441 respondents, 129 (29%) found the need to tell us their thoughts. We strongly recommend that Facebook read and consider this valuable user feedback.

All included feedback results are as entered by the users.

## E.1  User Feedback

- Facebook doesn't really secure your data... but then again you're putting it up for the world to see.

- give me a break. all of this information is readily available to anyone will to put 15 minutes into stalking a person. Facebook is not a tool of big brother.

- I don't give them much personal data anyway.

| When Users Join And Update Facebook at Harvard | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Month Of | Join | | Update | | 2007 Join | | 2008 Join | | 2009 Join | |
| Mar 1, 04 | 5698 | 32.18 % | 0 | 0 % | 1065 | 62.28 % | 21 | 1.18 % | 9 | 0.57 % |
| Apr 1, 04 | 1387 | 7.83 % | 0 | 0 % | 80 | 4.68 % | 14 | 0.78 % | 4 | 0.25 % |
| May 1, 04 | 698 | 3.94 % | 0 | 0 % | 71 | 4.15 % | 9 | 0.5 % | 0 | 0 % |
| Jun 1, 04 | 850 | 4.8 % | 0 | 0 % | 31 | 1.81 % | 298 | 16.69 % | 7 | 0.44 % |
| Jul 1, 04 | 491 | 2.77 % | 2 | 0.01 % | 16 | 0.94 % | 206 | 11.54 % | 3 | 0.19 % |
| Aug 1, 04 | 410 | 2.32 % | 30 | 0.21 % | 10 | 0.58 % | 204 | 11.43 % | 4 | 0.25 % |
| Sep 1, 04 | 711 | 4.02 % | 52 | 0.36 % | 38 | 2.22 % | 431 | 24.15 % | 4 | 0.25 % |
| Oct 1, 04 | 556 | 3.14 % | 70 | 0.49 % | 33 | 1.93 % | 195 | 10.92 % | 1 | 0.06 % |
| Nov 1, 04 | 387 | 2.19 % | 110 | 0.77 % | 32 | 1.87 % | 51 | 2.86 % | 1 | 0.06 % |
| Dec 1, 04 | 394 | 2.23 % | 145 | 1.01 % | 32 | 1.87 % | 27 | 1.51 % | 0 | 0 % |
| Jan 1, 05 | 380 | 2.15 % | 138 | 0.96 % | 26 | 1.52 % | 19 | 1.06 % | 4 | 0.25 % |
| Feb 1, 05 | 417 | 2.36 % | 173 | 1.21 % | 19 | 1.11 % | 22 | 1.23 % | 5 | 0.32 % |
| Mar 1, 05 | 402 | 2.27 % | 192 | 1.34 % | 28 | 1.64 % | 15 | 0.84 % | 3 | 0.19 % |
| Apr 1, 05 | 324 | 1.83 % | 209 | 1.46 % | 11 | 0.64 % | 19 | 1.06 % | 2 | 0.13 % |
| May 1, 05 | 285 | 1.61 % | 237 | 1.65 % | 13 | 0.76 % | 14 | 0.78 % | 2 | 0.13 % |
| Jun 1, 05 | 346 | 1.95 % | 382 | 2.67 % | 18 | 1.05 % | 24 | 1.34 % | 6 | 0.38 % |
| Jul 1, 05 | 1261 | 7.12 % | 480 | 3.35 % | 32 | 1.87 % | 31 | 1.74 % | 930 | 58.75 % |
| Aug 1, 05 | 594 | 3.36 % | 462 | 3.22 % | 21 | 1.23 % | 25 | 1.4 % | 255 | 16.11 % |
| Sep 1, 05 | 620 | 3.5 % | 840 | 5.86 % | 36 | 2.11 % | 47 | 2.63 % | 197 | 12.44 % |
| Oct 1, 05 | 636 | 3.59 % | 1419 | 9.9 % | 35 | 2.05 % | 71 | 3.98 % | 115 | 7.26 % |
| Nov 1, 05 | 538 | 3.04 % | 2887 | 20.15 % | 37 | 2.16 % | 37 | 2.07 % | 22 | 1.39 % |
| Dec 1, 05 | 319 | 1.8 % | 6564 | 45.81 % | 26 | 1.52 % | 5 | 0.28 % | 9 | 0.57 % |
| Total | 17704 | 100 % | 14392 | 81.29 % | 1710 | 9.66 % | 1785 | 10.08 % | 1583 | 8.94 % |

Figure 31: Facebook usage data for Harvard University.

- I dont really care about my privacy on the facebook because i lie in my profile a lot

- I set the option that prevents non-friends from seeing my cell phone number.

- I think people need to be aware that anything they put on Facebook is public domain. Even though I'm not sure of the legalities, I don't put information up that is too personal (phone numbers, etc.)

- I think that it is primarily the users' responsibility to be careful what is placed up on the facebook; not the other way around.

- I think you should have to approve a tagged pictured before it goes up rather than having to check periodically to see if any pictures are not something you want up, having to untag it and possibly report it.

- I wish I could automatically block all photo "tags"

- it is hard to tell whether ppl take facebook seriously or goof off with it, the my photo is nice but needs a seurity on it as well - asking permission of the people in it ahead of time etc.

- Since you willingly submit information to Facebook - such as your name, age, gender, etc. - you should be fully aware that practically anyone from your school can view your personal information if you do not change your privacy settings; that Facebook can share your information with third-party companies is somewhat alarming, but there is an option to request that your information is not shared with third-parties.

- the photo feature is highly questionable, especially since users other than yourself can "tag" you in their photos.

- There are appropriate options, but only if you take advantage/know about them

- They need to support SSL.

- To clarify my privacy concerns, I treat Facebook like any other open internet forum, and filter things through the concern that anyone may view the information. Since my peers have such easy access to the data and can be sure it actually belongs to me, I am even more careful about posting information (such as my sexuality) that I might not want acquaintances from high school asking about. Basically, I put the burden of protecting my privacy on myself via posting responsibly, not on Facebook via restricting access to what I choose to post.

- what i think is interesting is that third parties can post photos of you and link them to you and it is unclear to me if you have any control over that or who can view those.

- When I place information on thefacebook, I do so specifically because I want it to be in the public domain. There is obviously information that I would like to keep private, but I don't place it on thefacebook.

# F  Paper Survey

The paper survey follows. The web form survey asked the same questions, plus an additional question: " How concerned are you about the privacy of your data on the Facebook?" Possible answers here were: N/A, Not, Barely, Somewhat, Quite, Very.

Facebook Privacy Study: Survey                                    FREE CANDY

*Instructions*: Please circle your answers honestly. You may skip any question you do not wish to answer. Please DO NOT write name, e-mail, or other contact information on this form.

1.  Which gender describes you best? (please circle one)...........MALE          FEMALE
2.  Which category describes you best? (circle one)

     UNDERGRADUATE          GRAD STUDENT          POSTDOC
     FACULTY          STAFF          OTHER

3.  Which best describes your current living situation?

     DORMITORY: (please specify)
     FSILG: (please specify)
     OFF CAMPUS / NON-MIT HOUSING          OTHER

4.  Do you currently have a Facebook account on www.facebook.com?

     NO – Did you ever?          NO     YES
     YES – When did you create your account?          DON'T KNOW / FORGOT
     JAN   FEB   MAR   APR   MAY   JUN   JUL   AUG   SEP   OCT   NOV   DEC
          2003          2004          2005

5.  Approximately how many times a week do you log in to the Facebook?

     0     1-3     4-8     9-15     20-30          31+   DON'T KNOW

6.  Approximately how many Facebook friends do you have from MIT?

     1-10   11-50   51-100          101-200          200-350          >350

7.  Approximately what percentage of all of your friends are from MIT?

     0-15%          16-33%          34-50%          50-75%          75-100%

8.  Do you ever friend people whom you have never met in person?

     NEVER          SOMETIMES          ALWAYS

9.  Are you familiar with Facebook's "My Privacy" feature, that lets you control who may view your profile?......................................................................          NO          YES
10. Do you use the "My Privacy" feature to control who may view your profile?

                                          NO          YES
11. Have you ever read Facebook's Terms of Service in full?..............  NO          YES
12. Have you ever read Facebook's Privacy Policy in full?..................  NO          YES
13. Can Facebook share your information with other companies?.......  NO          YES
14. Are you familiar with Facebook's "My Photo" feature?.................  NO          YES
15. Can you prevent other Facebook users from seeing your photos?...NO          YES
16. Does Facebook do an adequate job in securing your personal data?NO          YES

Please use the back of this form to make any additional comments you may have.

THANK YOU!