# Privacy in Social Networks

Carlos Ordonez

David Matusevich

### Outline

#### I. Overview

- What is a Social network?
- Prominent social networks
- What is Privacy?

### II. Privacy Issues

- Social; Legal
- Differences with data security
- Commercial advantage

### III. Controlling privacy

- What Data is Collected
- Managing Privacy Settings

### IV. Strategies for Safe Sharing:

- Protecting your online "brand"
- I. Safe Online Socializing

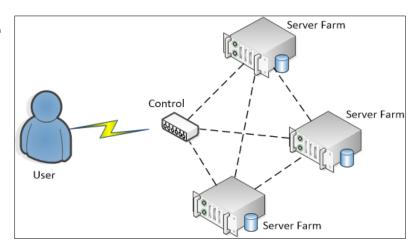
## I - Overview

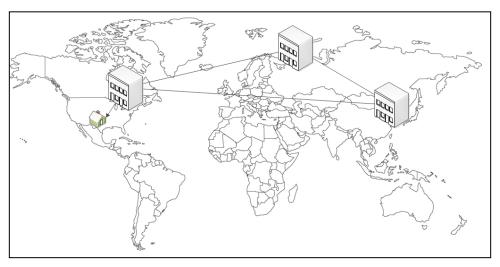
### What is a Social Network?

- A Social Network Site is a Web-based service that allows individuals to "construct a public or semipublic profile within a bounded system; articulate a list of other users with whom they share a connection; and view and traverse their list of connections and those made by others within the system", therefore increasing their social capital.
- We will not consider other kinds of social networks that don't rely on a social network provider.

# Social Network Data Storage

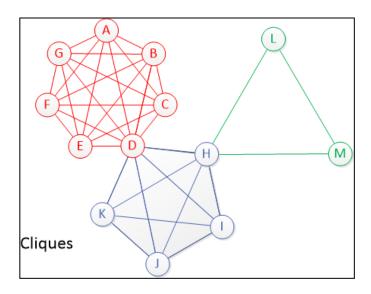
- USA/Canada and Europe
  - Local
  - External
- At multiple sites
- Transferred and shared

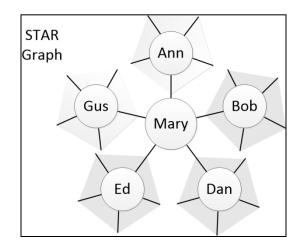


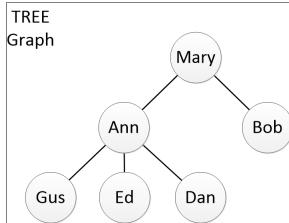


# Dissemination of information Shape of friends graph

- Star
- Tree
- Interconnected circles
- Cliques







### Social Network

- A Social network is a central place that combines
  - Entertainment
  - Social interactions
  - Communication facilities
- Social network operators (users) build profiles that can be seen by other users. The user manages the amount of information others can see.
- Personal data is now considered the new "oil", and companies are eager to cash in on this new resource.

### Social Network Elements

- A bounded set of users
- Public or semipublic personal profiles
- Definition of a set of people related to a person (friends, relatives)
- Freedom to traverse lists of connections (their own and others)
- Social Network Capital: The expected collective or economic benefits derived from the preferential treatment and cooperation between individuals and groups

# Social Networks vs. Internet Communities

- Internet Communities: Similar to social networks but there are no explicit interactions between users and no connections.
- YouTube, Amazon, eBay are communities, but since there is no set of connections, they cannot be considered networks.
- As time goes by, the lines between communities and networks are becoming more blurred.

# The Social Aspect

- Social networks are used to connect with people met offline or online
- To a lesser extent investigate people (as a primitive background check)
- Colleagues, classmates and friends in general, may share connections online, but not necessarily offline

### Motivations for Joining a Social Network

### People join social networks to:

- Create and share content about themselves
- To connect with others (either old acquaintances or new)
- To meet people with similar interests
- Financial Motivations

In order to achieve these goals, there must be a measure of voluntary disclosure among multiple users

### Top Ten Social Networks by Users

Rank	Name	Active user accounts	Site Country of origin
1	<u>Facebook</u>	1 billion <sup>[1]</sup>	<u>United States</u>
2	Tencent QQ	712 million <sup>[3]</sup>	China
3	<u>Qzone</u>	400+ million <sup>[5]</sup>	<u>China</u>
4	Sina Weibo	300+ million <sup>[7]</sup>	China
5	Google+	235 million <sup>[8]</sup>	<u>United States</u>
6	Twitter	200+ million <sup>[10]</sup>	<u>United States</u>
7	<u>VK</u>	190+ million <sup>[11]</sup>	<u>Russia</u>
8	LinkedIn	160 million <sup>[12]</sup>	<u>United States</u>
9	Renren	160+ million <sup>[13]</sup>	China
10	Skype	145+ million <sup>[14]</sup>	<u>Estonia</u>

# Social nets becoming more common

- Social networks are an increasingly ubiquitous part of Americans' daily lives;
- Recent data shows that 65% of Internet-using U.S. adults maintain a profile on an SNS
- This figure is increased to 81% when considering teens

# Facebook phenomenon

- More than 1 Billion active users.
- 50% percent of users log in daily.
- The average user has 130 friends,
- Average user is a member of 12 groups, and spends more than 55 minutes per day on the

site



### Facebook in more detail

- 2.5 billion photos uploaded each month, with more than
- 3.5 billion pieces of content shared each week
- There are currently more than 70 translations of the site available
- 70% of Facebook users coming from outside of the United States (Facebook, 2010).

# Privacy Overview: Privacy Definition

- Privacy is the ability of an individual or group to seclude themselves or information about themselves and thereby reveal themselves selectively.
- Westin 1967: "the claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others"
- Altman 1975: selective control of access to the self



# There is NO privacy any more!

- In essence the European Community considers privacy a Human Right, not something that is granted by the government.
- Article 8 of the European Convention on Human Rights provides a right to respect for one's "private and family life, his home and his correspondence", subject to certain restrictions that are "in accordance with law" and "necessary in a democratic society".
- This view is not universally accepted. In 1999 Sun Microsystems
   CEO Scott McNealy called privacy a "red herring". "You have zero
   privacy, get over it!", he said.
- Google CEO Eric Schmidt said that "If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place", when asked if users should be sharing their information with Google.

# Principles governing the European Community (OECD) recommendations for protection of personal data

- Notice—data subjects should be given notice when their data is being collected;
- 2. Purpose—data should only be used for the purpose stated and not for any other purposes;
- Consent—data should not be disclosed without the data subject's consent;
- 4. Security—collected data should be kept secure from any potential abuses;
- 5. Disclosure—data subjects should be informed as to who is collecting their data;
- **6. Access**—data subjects should be allowed to access their data and make corrections to any inaccurate data; and
- 7. Accountability—data subjects should have a method available to them to hold data collectors accountable for following the above principles

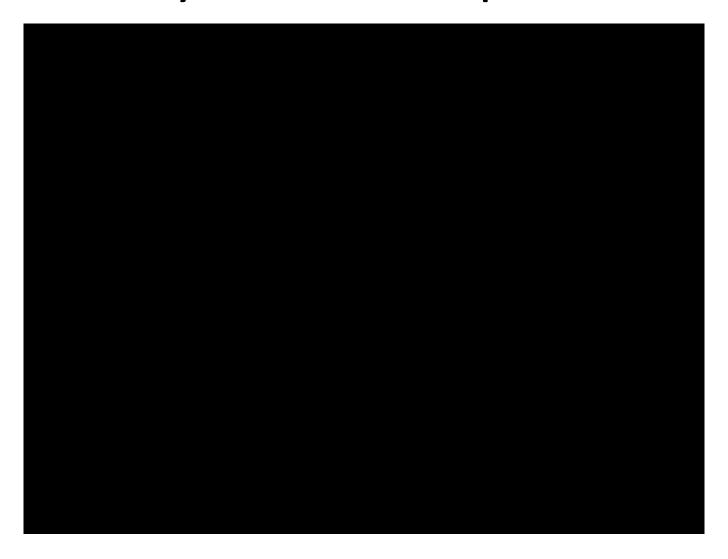
# Online vs. offline privacy

- We are weary of people that might approach us in our daily life, but we react different to strangers we meet online.
- This behavior is seen on people from all ages, from children to adults.
- There is a real disconnect between online and offline notions of privacy.

### Hardware

- Social networks can be accessed by a large number of different devices.
- Laptop and desktop computers give access to better "views" of the network site, but limit the spontaneity of sharing.
- Cellphones and tablets provide easy access to the social networks and enable quick sharing of photos and geo-location.

# **Privacy Invasion Experiment**



# Information privacy

• Information privacy, or data privacy (or data protection), is the relationship between collection and dissemination of data, technology, the public expectation of privacy, and the legal and political issues surrounding them.

## Privacy types

- Social privacy: how people protect themselves from other users
- Institutional privacy: how the company that runs the social network uses people's data
- Concern: the heightened visibility that is the result of having a large number of friends, including people in different ages → social surveillance and social control

# Privacy issues of social networking sites

- Social networks keep track of all interactions used on their sites and save them for later use.
- Issues include:
  - Cyber-stalking,
  - location disclosure,
  - social profiling,
  - data leakage and information integration,
  - 3rd party personal information disclosure,
  - government use of social network websites in investigations without the safeguard of a search warrant.

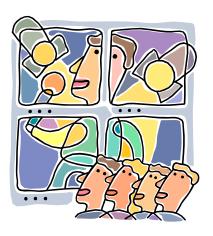
## **Impact**

- The privacy impact of social networks should not be underestimated.
- Many users do not seem to realize that their free use of social networks has an indirect but steep effect through the exposure of their own personal data.
- In addition, many users do not realize which impact they have on the privacy of their friends and families when they publish information about them.

# Unbounded audience

Marwick and Boyd (2011)

"We may understand that the Twitter or Facebook audience is potentially limitless, but we often act as if it were bounded"



## **Privacy Awareness**

- Privacy issues often only become apparent when it is already too late.
- It is practically impossible to predict (all) negative consequences of the use of personal data.
- Even if one can foresee a few, they are very abstract, distant and uncertain.

# Example: Consequences for College Applications

- Colleges currently use social media sites to recruit new students.
- Of those admission officers that visited applicants SNS, 35% discovered something negative about the applicant.

(Kaplan's College Admissions Survey)

# Example: Consequences in the Job Hunting

- 93% of recruiters review a candidate's online presence as part of the screening process.
- 42% have reconsidered candidates based on their online presence (both negatively and positively)
- Even spelling and grammatical errors influence recruiters negatively (61%)
- Posts/Tweets about volunteering and charity donations influence recruiters positively (65%)

(Jobvite Social Recruiting Survey, 2013)

# Means to achieve privacy

- Withdrawal from society activities
- With physical or psychological means, in solitude or in a small group of people
- Anonymity: the Dark Web; disable cookies; IP hiding.
- Live "off-line"

## Less privacy in the Future

- Even if an individual might know intellectually that the usage might have negative consequences, this is not going to change behavior that much.
- Our search-history, location-data, browsing-habits, reading-behavior and much more, is collected and/or used to a degree we can barely imagine.
- Technology, nowadays, allows for unprecedented forms of data-matching, de-anonimization and data mining, all contributing to extensive 'digital dossiers'.

# II - Privacy Issues

# The Economics of Privacy

- Companies can determine what adds you see online, what products to recommend you, even what articles to read, based on your previous behavior.
- Companies adopt a "collect first, ask questions later" policy.
- Some are selling consumer-specific data for purposes that fall right on the boundaries of the Fair Credit Reporting Act and other laws.

## The Economics of Privacy

 CampaignGrid (Republicans) and Precision Network (Democrats) have political information on 150 million American Internet users, or roughly 80 percent of the nation's registered voters.

# Main causes of privacy issues

- Data publically available
- Blurred or no personal boundaries
- User has limited control over information dissemination or transfer
- For a long period of time; forever?
- Hard to remove a derogatory post or comment
- Net etiquette different from face to face etiquette
- New cases not considered by existing law

# Information Leakage and Linkage

- Information leakage happens whenever a system that is designed to be closed to an eavesdropper reveals some information to unauthorized parties nonetheless.
- Information linkage is joining together of two datasets to produce one single dataset. In short it is possible to use information leaked from social networks to sniff out information private to the user, such as email addresses, ID numbers, etc.
- Linkage cannot happen without Leakage. Some measure of leakage is unavoidable.

### Identity in social networks

- True
- Partially concealed
- Anonymous
- 2<sup>nd</sup> life; alter ego

### Apps and Information Leakage

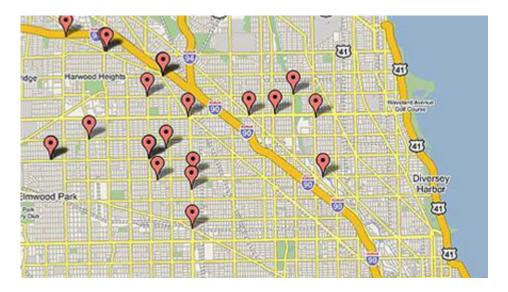
- Apps within social networks (games, messengers, utilities, music apps, etc.) are an important source of leakage.
- People will share address books, phone numbers, credit card numbers, etc. with applications that have little or no security and may even be malicious.
- This behaviour is the same in phone apps.

### The Internet does not forget.

- 'Right to be Forgotten': The right of individuals to have their data no longer processed and deleted when they are no longer needed for legitimate purposes.
- The 'right to be forgotten' clearly takes a proprietary approach to privacy protection. Its scope, therefore, strongly depends on a clear and consistent definition of 'personal data'.

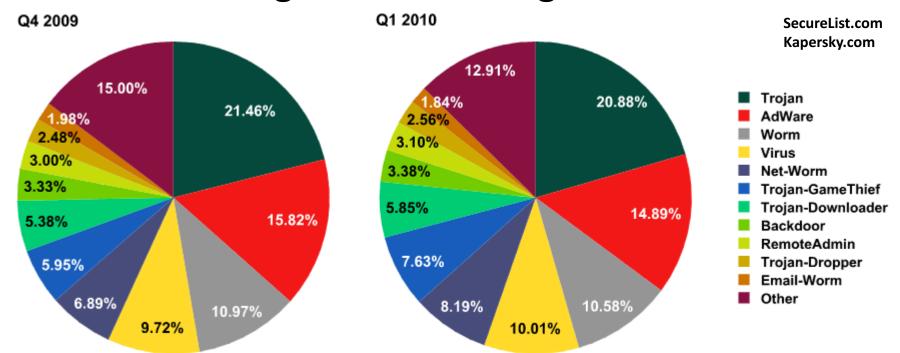
# Compromised Privacy: Finding user identity

- Generic Searches (Google, Yahoo, etc.)
- From user public profile
- Matching data across sites
- Exploit photos and video tags and geo-tags.



### Security Threats of unsecured access

- Hackers
- Identity thieves
- Government global knowledge



### Cookies

- Cookies are a way of storing persistent client data so that a site can maintain information on a user across HTTP connections (text files).
- Information stored ranges from
  - Shopping Carts,
  - Forms, Addresses and Personal information (usernames and passwords),
  - Login information
- Main culprit of information leakage.

### Issues beyond security

- Bad user behavior (bad language, cyberbullying, anonymous threats)
- Inability to control social spheres
- Blurred boundaries between acquaintances, friends, relatives
- The user is responsible for managing what is disclosed, not an organization

# Further privacy issues even when user willingly agrees to disclosure

- Open discussion of personal information among contacts,
- The posting and tagging of photographs that identify other users,
- Disclosure of demographic data,
- Posting personal information on profile pages that implicates other users

### Who can disclose data?

- Person himself/herself
- A friend or relative
- A 3<sup>rd</sup> party



### Social aspects

Social Networks are a recent phenomenon and as such there are no existing, clear social conventions about their use. Example: ignore "friend" requests.

Other users consider the number of friends as a status symbol, effectively causing the boundaries between private life and professional life to become increasingly blurred.

### Social Context Collapse

- Social network collapse is the flattening out of multiple distinct audiences in one's social network.
- People from different contexts become part
   of a singular group of message recipients.
   Users can quickly diffuse information across
   their entire network and facilitate interaction
   across diverse groups of individuals who
   would otherwise be unlikely to communicate.

## Legal Cause Privacy Policies and Users Rights

- Written in vague legalese
- People do not read them
- Network externalities, lock-in and the lack of valid alternatives often force people into consenting.

### Legal Issues Regulation

Low: 3<sup>rd</sup> world countries

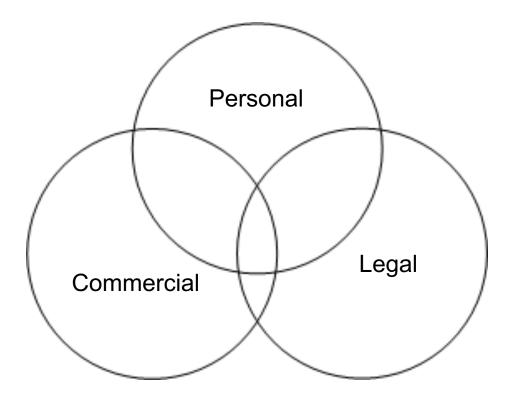
Medium: US

• High: Europe



# Legal issues boundary blurred

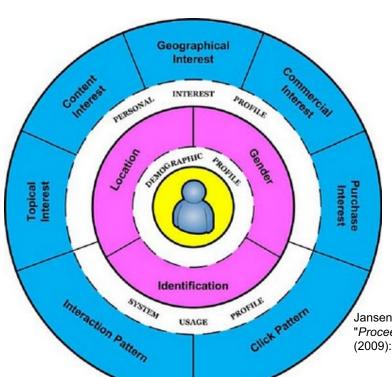
- Personal
- Business
- Government



### Commercial advantage

Targeted advertising

User profiling





Jansen, Bernard, et al. "To what degree can log data profile a web searcher?. "Proceedings of the American Society for Information Science and Technology46.1 (2009): 1-19.

### Commercial advantage: Facebook

- Selling targeted advertising
- Virtual currency (Facebook credits)
- Facebook apps and games collect information about your habits and about your friends, without your knowledge or consent.

### The Situation in USA

There is currently no federal online privacy law, which makes it essentially impossible for government agencies like the Federal Trade Commission to go after Internet companies unless they violate their own published privacy policies.

### The Situation in USA

- There are some rules in place to deal with Privacy in regards to children under the age of 13.
- The Children's Online Privacy Protection Act (or COPPA) was passed in 1998.
- A new set of rules was published by the FTC (Dec 2012) clarifying what is or isn't allowed.

#### **New FTC Rules**

- Make clear that the "personal information" that can't be collected without parental consent includes geo-location information, photographs, and videos
- Make clear that third parties (like advertising networks) must also comply with COPPA
- Close a loophole that allowed kids' information to be collected via plug-ins without parental notice
- Clarify that "persistent identifiers" are also protected information, like IP addresses and mobile device IDs
- Require that websites aimed at kids have "reasonable procedures" for data retention and deletion

### The Situation in USA (Cont)

- Legislation has been proposed to include "Do Not Track" options on web browsers.
- This legislation is not politically viable due to opposition from the business community.



### III - Controlling Privacy

## Automatically Collected Computer Data

- IP address
- Computer name
- Linking data across different sites
- Time, date
- Location
- Mechanism: Cookies

### Information collected with cookies

- Geographical Location
- Detailed date/time
- Computer name, id
- IP address, MAC
- Logged in user name
- Other web pages visited
- Form data

### Computer Data Manually Entered

- Personal information
- Comments
- Photos, Video



### Social networks: users and data

- Parties
  - Social network operators
  - Users
  - Application providers
- Roles to manage data
  - Data controller
  - Data processor
  - Data provider



### Palliative: Privacy Settings

- Extensive sets of privacy controls different levels of sociability
- Shield content sharing
- Potential problem: users are not able to properly utilize the privacy settings provided by SNSs
  - Controls are difficult to understand and most users just leave the recommended settings (preferred by the SSN)

### Privacy control

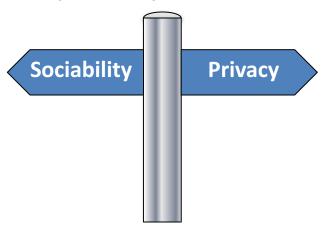
- Who in the network can access information in your personal Facebook profile?
- Can you find the minimum age for using Facebook?
- How can you change your Facebook settings to restrict visibility to your profile?
- How can you change your Facebook settings so that you are alerted when you are tagged in a photo?

### Controlling access to data

- Limited
- Site-dependent
- Difficult to understand legal language
- Impossible to know if other person discloses data
- Transferrable

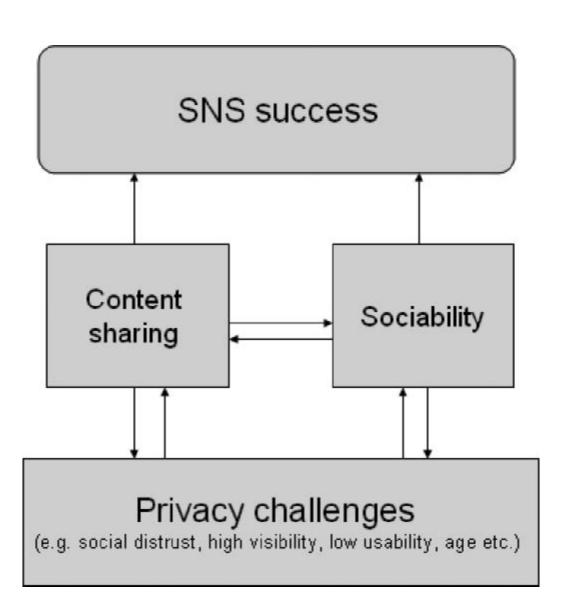
### Dilemma

- In either case below, the consequence is undesirable:
  - if privacy is protected, then sociability and content sharing will be compromised,
  - whereas if sociability and content sharing are promoted, then privacy will suffer.



### **Implications**

- Increased social utility and a growing social diversity of the user population, which help users to be readily available and visible to a lot of people: "all friends in one-place solution."
- SNS profiles mix friends, family, co-workers, and business contacts
- No simple and adequate way to separate them and keep some parts of the information private



### Comparison of the Younger Adult Sample and the Older Adult Sample in Regard to Social Practices

#### **Young Adult**

- Uses SNS for short periods of time, but many times a day
- Mainly contacts with friends that they see every day
- Uses SNS for coordination with friends, flirting and photo sharing
- Usually share large amounts of photos and videos from social gatherings
- Infrequent status updates

#### **Older Adult**

- Uses SNS for fewer longer sessions
- Mainly contact with family and old friends that are not seen often
- Uses them for getting in touch with old friends, nostalgia
- Shares photos less often (rarely videos) mainly of vacations
- Frequent status updates

### Comparison of the Younger Adult Sample and the Older Adult Sample in Regard to Privacy

#### **Young Adult**

- Confident in the usage and knowledge of privacy controls
- Thinks other people are more likely to have problems with privacy
- Concerned about privacy in the context of job hunting.
   Not concerned with the use of information by the SNS

#### **Older Adult**

- Less confident. Usually ask for help from the younger adults in the household
- Many privacy concerns, in particular regarding the younger generations
- Concerned but less aware of privacy issue. May think that burglars might use SNS to case their homes, for instance

## Lowest Common Denominator Strategy

- Individuals for whom a message is not intended but would receive the message nonetheless.
- Err on the side of caution: If any of these individuals would find the message problematic, it should not be posted.



### **Privacy Preservation Costs**

- Required from the user in order to make use of the site's privacy features:
  - Time required to understand and operate the myriad of different user settings.
  - Knowledge of the intricacies of the particular social network.
  - The time and knowledge invested in one network is not transferrable to another.

### Limitation principle

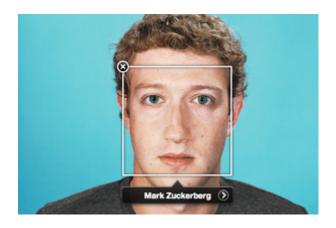
- Confining data processing to a previously defined scope: might seem to restrict the amount of potential harm in theory.
- But in an ever-increasing personalized web (where every piece of personal data can be considered as 'useful'), the value of this principle has become questionable too.

#### Settings

- Users may not be well-versed in privacy settings or unwilling to take the time to change settings.
- Distributing content to one's entire network appears to carry a lower cost in terms of time, knowledge, and skills. However, such strategies may negatively impact relationships on the site, especially if the majority of posts are relevant to a minority of Friends.
- While individuals choosing a lowest common denominator approach may avoid alienating friends with irrelevant content, they may also miss the benefits derived from interactions with all members of their networks.

# Additional automated sources of information

- Recognition
  - face
  - voice
- "sway" user into tagging



# Concern: Status Public Channels

 Status updates provide the quickest method through which one can distribute messages to a wide audience

 It may be more likely to be used even when the message is only relevant to a subset of

Friends.



#### Control Data Transfer

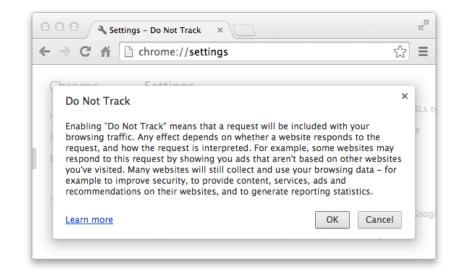
- Request 'personal data' to be deleted on one site
- Deletion may imply just hiding data; not shredding it
- Meanwhile the information might have been copied and/or 'anonymized' already.

# Solutions to control data sharing and transfer

- awareness-raising,
- transparency,
- clearer privacy notices,
- data-minimization,
- stricter control on the purpose limitation principle, 'anonymisation',
- transparency,
- encryption,

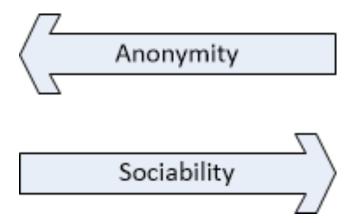
#### Anonymization

- Major line of defense.
- The right does not offer any solution.
- Individuals may be profiled/targeted extensively and their data might (in)directly be used for comprehensive datamining,
- The individual cannot have a 'right to be forgotten' with regard to this information.
- An Extreme Example: The Tor Network: The Deep Web.



## **Anonymization Solutions: Catch 22**

- Anonymization measures prevent (potentially harmful) information to be shared
- But, in an ever-increasing social Internet, many features depend on disclosing personal data.



## **Examples: Privacy Control Features**

- Facebook
- Twitter
- Snapchat
- Foursquare

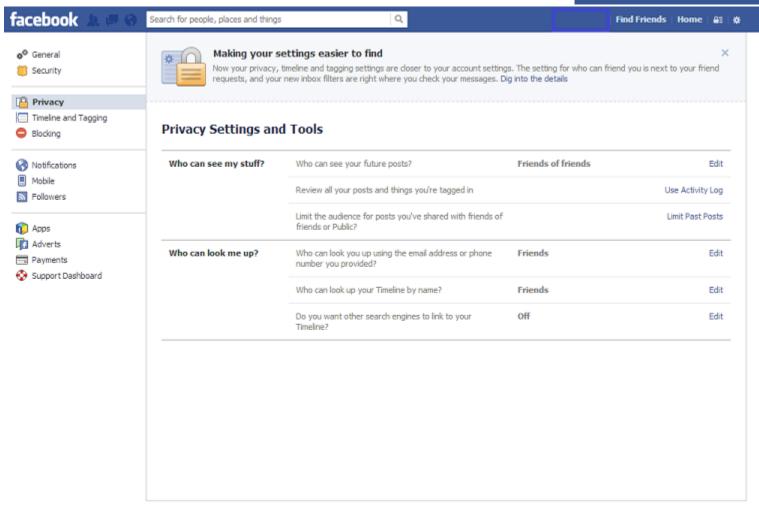


# facebook

#### Facebook

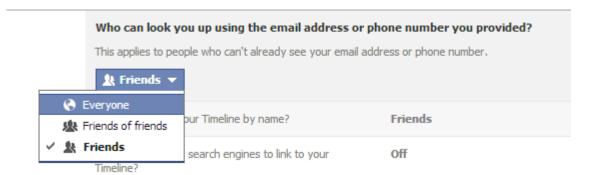
- Overwhelming
- 5 groups
- Fine grained
- Constantly changing
- Pre-defined options: Friends, Public, Custom
- Learning curve

# facebook



About Create an Advert Create a Page Developers Careers Privacy Cookies Terms Help





# Do you want other search engines to link to your Timeline? Please note: When this setting is on, it is easier for other search engines to link to your timeline in search results. If you turn off this setting, it may take a while for search engines to stop showing the link to your timeline in their results. To turn on this setting, first you need to set your Who can look up your Timeline by name? setting to Everyone. Let other search engines link to your timeline

#### **Twitter**



- The privacy settings are basic
- Much simpler than FB.
- Profiles can be public or private.
- Your Bio, name and Twitter handle always visible
- Email address is private
- Guards against identity theft





**Snapchat** is a photo messaging application. Using the app, users can take photos, record videos, add text and drawings, and send them to a controlled list of recipients. These sent photographs and videos are known as "Snaps". Users set a time limit for how long recipients can view their Snaps (as of December 2013, the range is from 1 to 10 seconds), after which they will be hidden from the recipient's device and kept on Snapchat's servers forever

## Snapchat



- Snapchat has two privacy settings, one for who can send you Snaps and another for who can see your Stories. Both have two options "Everyone" and "My Friends."
- By default, only users you add to your friends list can send you Snaps. If a Snapchatter you haven't added as a friend tries to send you a Snap, you'll receive a notification that they added you, but you will not receive the Snap they sent unless you add them to your friends list.
- Usernames and personal phone numbers of millions of users have been stolen and posted online through a website entitled SnapchatDB.



#### Foursquare

There are certain data that will always be public. This includes your name, your hometown (i.e. "location" in your profile), your bio, your profile picture and other public photos, your likes, your tips, your lists, and your friends. The only way to hide this info is to either not include it, or to change it so it doesn't actually reveal any personal information about you. If you don't, this information can easily be found through a simple Google search for your name, or other online outlets.

#### Your "Online Brand"

- In business terms, a brand comprises all of the things that make up a company's identity to customers, from its corporate logo to the names for its products. Because a strong brand is so important to a company's reputation and success, executives take great pains to protect their brand
- With high-powered search engines like Google and Bing, finding information about a potential job applicant, business partner, or date, is easier than ever. By applying the principles of business branding and online reputation management to yourself, you can make sure that your name looks good in search results.

# IV - Strategies for Safe Sharing

Tips and advice

#### How to protect your "Online Brand"

- Take charge of your "online reputation"
  - Find out what is on the Internet about you
    - Use search engines
    - Search blogs and social networks
  - Evaluate your online reputation
    - Does the information about you reflect how you want others to perceive you?
  - Protect your online reputation
    - Think before you share
    - Treat others as you would like to be treated
    - Stay vigilant about what the Internet is saying about you



#### Restore your Online Reputation

- If you find information about yourself that does not fit the reputation you want, act quickly. The longer it stays public, the greater the chance that it will be spread or archived.
- In a respectful way, ask the person who posted it to remove it or correct an error. If it is a correction, ask him or her to include a notice (CORRECTION or UPDATED) right next to the original (incorrect) material.
- If the person does not respond or refuses to help, ask the website administrator to remove the digital damage.
- If you feel a public correction is necessary, present your case simply and politely without attacking the person.

(<a href="http://www.microsoft.com/security/online-privacy/reputation.aspx">http://www.microsoft.com/security/online-privacy/reputation.aspx</a>)

## Safer Online Socializing

- Set your boundaries:
  - Think carefully about how public you want your profile or blog to be
  - Evaluate the social site before you use it
- Be selective about friends:
  - Think twice about who you accept as a friend
  - Periodically reassess who has access
  - Review what your friends write about you

## Safer Online Socializing

- Think before you post
  - Choose a user name that doesn't attract unwanted attention or help someone find you
  - Do not over share
  - Treat others as you would like to be treated
  - Think about the future of your information on the web
- Defend your computer against online threats
  - Be wary about clicking links
  - Build up your computer's defences and keep them up to date
  - Be careful about installing add-on apps

## Safer Online Socializing

#### Report Issues:

#### No one has the right to threaten or upset you. Report:

- Any negative incidents to the Web service, including content that exploits minors, obscene or hateful material, inappropriate behaviour, or theft of your account.
- Continued harassment or physical threats to local law enforcement.
- Identity theft to the U.S. Federal Trade Commission (FTC) at <a href="ftc.gov/idtheft">ftc.gov/idtheft</a> or call toll free: (877) 438-4338.
- Scams or fraud to the FTC. Go to <a href="ftc.gov/bcp/consumer.shtm">ftc.gov/bcp/consumer.shtm</a> and click
   File a Complaint, or call toll free: (877) 382-4357.

(http://go.microsoft.com/?linkid=9708812)

#### Social Networks and Gaming

- Social gaming networks are a proof of that:
  - Social networks fostered by the gaming companies: Xbox Live, Playstation Network, Nintendo Network, etc.
  - Social networks maintained by users: Raptr,
     Playfire, Duxter, etc.

## Safe Gaming

- Educate yourself and your kids about the risks:
  - Kids play alone, with others in the room, or online. They
    play against the game itself or another person, with a
    team of several players, or in games which may have
    hundreds of thousands playing at any one time.
  - The bad may download with the good: some "free" games require extensive profiles, then illegally sell your data.
  - Online bullying: Some gamers play simply to harass and taunt other players.
  - Bad people may befriend kids, and through these social gaming sites obtain personal information that might lead to harm.

#### Conclusions

- Privacy no longer viable
- Companies hungry for data
- We should become aware of privacy issues
- Goal: safety+privacy, preserving "personal brand"