Specific Purpose: Get more in depth understanding of how much information they are gathering and how much the user is actually giving away.

Introduction
I. In 2012, Target used data gathered from women in their baby registry to discover a pattern in shopping habits of pregnant women. Using this data, they identified women who might be pregnant and sent ads for baby supplies to them. This led to them sending coupons to a teenage girl, that were discovered by her father who didn’t know.

II. There is a difference between privacy and confidentiality. The confidential treatment of information is not only applicable to private and personal information, but it may refer to any category of information, such as, inter alia, trade secrets. So confidentiality applies to a bigger scope than privacy does, but we are going to just focus on the privacy aspect of internet marketing.

III. (Thesis) The internet led a huge socio-economic effect where information sharing became a source of income for businesses, and in return the consumer gets free service. But, is giving away personal information in return for services beneficial or ethical?

[Transition-Let’s start peel off a layer of the internet and see what and how data is collected and transfer through the internet]

Body
I. Notice and Choice versus Free and Informed Consent is a difference that the consumer should know about because it is what creates this whole issue of informational gathering and usage.

   A. “Informational privacy is the ability to determine for yourself when others may collect and use any of your personal information for any purpose. To determine this for yourself you must give…free and informed consent to how others process your information.” p78 of our CS377 Textbook.
B. Federal Trade Commission Fair Information Practice Principles state that the consumers should be given notice of an entity’s information practices before any personal information is collected from them and secondly, the company must give options as to how any personal information collected from them may be used.

C. Notice and choice does not ensure consent because consumers do not read contracts or privacy policies, and those policies and agreements change periodically, so the consumer has to keep up to know what is being used.

II. Contextual Advertising versus Behavioral Advertising are the two types on the internet.

A. In a contextual advertising scheme, ads are displayed to the user based on content currently being viewed on the website by the user. These ads are chosen by an automated system.

B. In behavioral advertising, consumer activities online are tracked and advertising is tailored to the user via profiling, ideally in real-time (perhaps even based on a store you are physically standing in, if GPS info is available).

a. Entities involved in delivering behavioral ads:

i. Profilers “create focused descriptions that segment buyers into groups in order to predict their willingness to buy specific types of products.” (Sloan and Warner 275). What do they base profiles on; “age, sex , ethnicity, marital status, profession, Internet search information, and sites visited.” (Sloan and Warner 276)

ii. Advertising Agencies design the ad campaigns for client businesses.

iii. Advertising Exchanges “deliver display advertisements to websites that display them. When a buyer visits a website, the exchange retargets advertisements by combining a buyer’s profile with information about the buyer’s current website activity. The exchange then conducts an auction in which businesses bid for the
opportunity to present their targeted advertisements. The whole process takes milliseconds.” (Sloan and Warner 276)

III. Specifically how are they collecting the data? Let’s talk about Big Data. There are three types of data being collected: Social Data (Facebook, Twitter, Google+), Machine Data (RFID, GPS), and Transactional Data (Paypal, Amazon, eBay, Walmart, etc). Here are some facts about the data that gets collected:

A. Every day consumers make around 11.5 million payments by using Paypal,
B. every hour, Walmart (chain of discount department stores) handles more than 1 million customer transactions
C. 510 comments, 293000 status and 136000 updates are posted on Facebook every minute, and
D. Every second, ~7000 tweets are made on Twitter.
E. Let’s pick Internet Search Engines. How do they collect data?
   a. There are three key factors in determining a search engine’s privacy practices:
      i. data retention length,
      ii. how that data is deleted and
      iii. whether such data can be tied to an individual user.
   b. Search engines make money almost completely through advertising. Advertisers pay the engines to appear on search results pages when certain keywords they have paid for are searched. These advertisements, not natural algorithmic results, are usually labeled “sponsored links” or “sponsor results.”
   c. Most major search engines may keep track of the searches you enter. If you have a toolbar, email address or other account with a search engine, your search and browsing history will be keyed to the account. If you do not have a toolbar or account, your history may be tied to a cookie on your Web browser.
d. An example of this would be when CNET News.com explains the major privacy gaffe that occurred when AOL publicly exposed millions of users’ searches connected with their IP addresses. Within a matter of days, journalists were able to connect searches on private issues to specific individuals.

F. Speaking of cookies, what are cookies (the non-edible kind) and how are they used?
   a. When you visit different websites, many of the sites deposit data about **1st party Cookies**: Legitimate websites use cookies to make special offers to returning users and to track the results of their advertising.
   b. **3rd-Party cookies**: communicate data about you to an advertising clearinghouse which in turn shares that data with other online marketers.
   c. **Flash Cookies**: your visit, called "cookies," on your hard drive. Cookies are pieces of information sent by a web server to a user's browser. Cookies may include information such as login or registration identification, user preferences, online "shopping cart" information, and so on.
      i. Flash cookies are not stored with other cookies in your browser. To delete them you need to utilize a separate control panel. Users who block and delete cookies in their browser may be unaware that they are not taking steps to do the same to cookies delivered alongside flash content.

G. When they collect all this information, there is whole new world out there to analyze and sort the data.
   a. **Cluster analysis** is when the data groups people by properties, attributes, and/or behaviors.
   b. **Data mining** is when the company discovers new behaviors based on large data sets.
   c. **Predictive modeling** is when the company can predict the outcome of actions taken.
d. **Textual/sentiment analysis** is understanding natural language and attributing it to positive, negative, or neutral reactions.

IV. At this point, we will ask the audience on the ethical issue of information gathering and use, and pros and cons of this concept.

V. Solutions and Obstacles with the Ethical Issue of Informational Gathering and Usage

A. Legislation, an imperfect solution. Why?
   a. Legislation moves at the speed of Government; it cannot keep up with tech advances.
   b. Companies have a strong, profit driven incentive to cheat or find ways around regulations.
   c. Regulation is more powerful when it is imposed from within (when businesses want to make the customer happy, not when they are only worried about incurring a fine).

B. Creating Value Optimal Norms/Best Practices at a societal level of what is ok and what is not as regards a contract between service providers and users is better. You don’t have to read the warranty that came with your water heater, you already know what it says; it’s pretty standard. And because protections in such documents are accepted societal norms you can expect courts to be on the side of consumers when businesses violate those norms. This, however, has only come about over an extended period of time.

C. Educating the consumer on the resources available to prevent undesired information leakage, and understanding what the services the consumer is providing is taking in return.
   a. Knowing how the consumer’s computer and how the internet works is important as well such as knowing if your IP address is static or dynamic
   b. Tor ([https://www.torproject.org/](https://www.torproject.org/)) is an example of tracking prevention software that can potentially anonymize your web-traffic. However, it is neither fool-proof nor bulletproof.
   c. Use a Virtual Private Network (VPN) to be more secure. A VPN replaces your IP address with one from the VPN provider and, thus reduces traceability.
d. For Mobile Devices, your carrier will collect the data so see what they do with it. Their policies tend to be clear cut and straight forward, so people can read them clearly, but they know that no one is going to look it up or ask questions…

e. Avoid using Wi-fi Hotspots because they come with little to no security. It is one of the most insecure ways to access the internet.

f. For cookies that will not go away, use Ghostery. Another example of tracking preventions software, Ghostery is a browser tool that scans webpages for trackers (including cookies) and notifies you of the companies whose code is present on the page you are visiting.

g. For social media and search engine privacy there are some tips that can help prevent too much unwanted informational gathering.

i. It's a good idea to avoid using the same website for both your web-based email and as your search engine. Web email accounts will always require some type of a login, so if you use the same site as your search engine, your searches can be connected to your email account.

ii. Avoid downloading search engine toolbars (for example, the Google toolbar or Yahoo toolbar).

iii. Google combines information about you from most of its services, including its search engine, Gmail, and YouTube. Be sure to disable automatic sign-ins

VI. Pros and benefits to the existence of Big Data

A. The more information you share, the better personalized services you get from the companies you shared with.

a. YouTube will give recommended videos based on the ones you watched

b. Amazon will offer similar products or ones that go with the products you’ve bought or are looking at

B. Some free services function based on advertisements from companies who use the big data

a. 90% the money Google gets to pay for its services comes from specialized ads
i. Gmail is a free email service offered by Google  
ii. YouTube is site full of free videos owned by Google  
iii. Google Drive is Google’s way of giving free data storage for everyone  
iv. Google Maps is a collection of map images that is used for navigation and direction services.  
v. Search Engines is the user’s encyclopedia on the internet.  
b. Music/Radio Apps (Spotify, Pandora, etc) will play ads to pay for services and allow you to listen to music for free  
c. Social Networks (Facebook, Twitter, LinkedIn) allow for connection with many people and get paid by companies that research and advertise to the users.

[Transition-To sum up, ]

Conclusion  
I. Summary: Your information is out there on the internet and someone is using it.  
II. Clincher: More and more services are becoming available to us all the time. But the key to the quality of those services are quickly being tied to how much they know about us. How much is too much? Where is the balance?

Reference List  


