Getting to Know You, Getting to Know All About You:
Big Data—Friend or Foe?
Your Panel

• **Kerry Childe**
  Senior Corporate Counsel, Best Buy

• **Harold Federow**
  Contract Vendor & IP Manager, Port of Seattle

• **Jennifer B. Schlosstein**
  Associate General Counsel, Wal-Mart Stores, Inc.

• **Alan Sutin**
  Partner, Greenberg Traurig, LLP
What We Will Cover

• Defining Big Data
• How companies currently are using Big Data and what issues this presents
• What the law currently says about Big Data and a framework for analyzing Big Data issues
• Practical tips and consideration for in-house counsel
Quick Poll

• How many have been to a presentation on big data within the past year?

• How many feel they could define it?
Another Poll

• How many have bought a crib in the past year?
• How many have bought diapers in the past year?
• How many have purchased baby food in the past year?
• How many have bought toys in the past year?
How Big Data Can Work in Retail

• Can put your purchase data together with purchase method, Facebook, or other information

• Targeted ads for age appropriate diapers, baby food, toys, etc.
What is Big Data?
What is big data?

• Advances in technology and the proliferation of mobile devices and sensors has resulted in an exponential growth in the amount and type of data generated

• In the past, the high cost of data storage required extracting the data required, and even then sampling was necessary and the data had to be structured

• Today it is economical to store what used to be considered “data exhaust”, and vast bodies of data can be examined
What is big data?

Not just retail customer analytics:

- Weather
- Space Exploration
- Public Health
- Manufacturing
- Air Travel
- Academia
- Supply chain

- Science
- Engineering
- Education
- Transportation
- Entertainment
- Multi-Media (Marketing/Advertising)
What is big data?

Characteristics

• Can extract knowledge from data long after collection
• Can mine and analyze unstructured data
• Can merge datasets from many sources
• Extremely adept at determining correlation (but not causation)
Cost of Data Storage

- Computing resources are becoming much cheaper and more powerful
- In 1980 a terabyte of disk storage cost $14 million; now it costs about $30

640K ought to be enough for anybody.
Three Dimensions of Big Data
Dimensions of Big Data: Volume

• 90% of the data in the world today has been created in the last two years
  – This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, cell phone GPS signal

• Lower storage costs have made it practical to store more data
Dimensions of Big Data: Velocity

The velocity of data creation is astonishing

• From the beginning of recorded time until 2003, we created 5 billion gigabytes (5 exabytes) of data

• In 2011, the same amount was created every two days

• In 2014, the same amount of data is created every 10 minutes
Dimensions of Big Data: Velocity
Dimensions of Big Data: Variety

• Data analysis traditionally required a structured data set

• Big Data technologies allow analytics to be run across databases of different structures or even unstructured data consisting of text, sensor data, audio, video, click streams, log files and more

• New insights are found when analyzing these data types together
## Where is all this data Being Collected?

<table>
<thead>
<tr>
<th>Web Browsers</th>
<th>Search Engines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft’s Internet Explorer</td>
<td>Google’s</td>
</tr>
<tr>
<td>Mozilla’s FireFox</td>
<td>Microsoft’s Bing</td>
</tr>
<tr>
<td>Google’s Chrome</td>
<td>Yahoo’s</td>
</tr>
<tr>
<td>Apple’s Safari</td>
<td>IAC Search’s Ask.com</td>
</tr>
</tbody>
</table>
Where is all this data Being Collected?

Multiplicity of devices, applications and networks collecting location and other data

- **Devices**: RIM, Nokia, Android
- **Apps**: Socialight, Loopt, Google Maps, Foursquare
- **Networks/Data**: Sprint, Nextel, Vodafone, AT&T, Verizon Wireless, NAVIZON

Location and Other Data
Where is all this data Being Collected?

Apps!

What are they collecting?

• Restaurant reservations (Open Table)
• Weather in L.A. in 3 days (Weather+)
• Side effects of medications (MedWatcher)
• 3-star hotels in New Orleans (Priceline)
• Movies watched (Netflix)
• Products Purchased (Amazon)
Where is all this data Being Collected?

Games consoles

GPS Systems

“Smart” TVs and Blu-Ray Players

Sensors
Where is all this data Being Collected?

• Purchase transaction records
• Smartphone GPS sensors
• Computers connected to the Internet
• Voice calls over VoIP or traditional telephony
• Social media sites
• Passive sensors (automobiles, RFID tags, etc.)
• Video recording devices (facial recognition, etc.)
Where is all this data Being Collected?

Hospitals & Medical Systems

- Pharmacies
- Laboratories
- Imaging Centers
- Emergency Medical Services (EMS)
- Hospital Information Systems
- Doc-in-a-Box
- Electronic Medical Records
- Blood Banks
- Birth & Death Records

Banking and Financial Services

- Branches
- ATMs
- Call centers
- Internet
- Relationship managers/agents

- VISA
- MasterCard
- American Express
- Discover
... And Let’s Not Forget Uncle Sam!

- Six of the ten most powerful supercomputers in the world are owned by the US Government

- In March 2012, the Obama Administration announced the Big Data Research and Development Initiative

- $200 million in new R&D investments, which will explore how Big Data could be used to address important problems facing the government. The initiative was composed of 84 different Big Data programs spread across six departments

- On January 17, 2014, President Obama announced an initiative to review the privacy implications of Big Data

- In May 2014, the FTC issued a report entitled “Data Brokers: A Call for Transparency and Accountability”
The ACLU View of Big Data’s Future
How Is Big Data Being Used And What Issues Will This Present?
Big Data Can Add Value to Society

- Google claims can predict influenza and other epidemics faster than the CDC
- The World Economic Forum promotes Big Data as a means for evaluating and responding to populations that previously were unmeasured
- Big Data can offer better fraud protection, crime prevention, and other benefits
Big Data Sometimes Has Unintended Consequences

- Major retailer inadvertently discloses 15-year old’s pregnancy
- Orbitz charges more to users who log on from Apple devices
- Facebook and OKCupid engaged in studies without informing their users
- NYC using eviction data to help social workers line up services
Considerations

• Just because you can may not mean you should
  – London rubbish bins are not expected to collect location information

• Customers may balk
  – Nordstrom

• Regulators have Opinions about Big Data
Why All The Attention to Big Data?

Researchers and policymakers are beginning to realize the potential for channeling these torrents of data into actionable information that can be used to identify needs, provide services, and predict and prevent crises for the benefit of low-income populations.

How Companies Are Using Big Data

• Serving the Customer Better
  • Companies with a history of fact-based decision-making for product and advertising decisions (Proctor & Gamble, UPS, various retailers)

• Part of the Business Model
  • Web-Native Companies (Google, Netflix)

• Creating Better Products
  • Engineering and Research Companies (Lockheed Martin, RAND Corporation)
What effect does this have?

• Business Benefit:
  • Price transparency for consumers
  • Business success and growth (knowing the characteristics and value of customers)
  • Anticipating customers’ needs and serving the customer: Previous behaviors, similar customer behaviors, emerging buying behaviors

• Other Considerations:
  • Data retention expenses and legal implications
  • Freedom and privacy (individuals, children)
... [F]undamentally, I believe we should be concerned about the damage that is done to our sense of privacy and autonomy in a society in which information about some of the most sensitive aspects of our lives is available for analysts to examine without our knowledge or consent, and for anyone to buy if they are willing to pay the going price.

“Not everything that can be counted counts, and not everything that counts can be counted.”

-- Albert Einstein
Actual Facebook Graph Searches
Is “Anonymizing” Data a Solution?

• AOL and Netflix each have released “anonymized” or “de-identified” data sets to the community for research

• In each case, researchers were able to ascertain the personal identity of data subjects through analytics
What the Law Currently Says About Big Data and A Framework For Analyzing Issues
The Regulation of Privacy Around the World

- United States
  - Historically has had minimal privacy regulation...
  - *BUT*, has been at the forefront of technological innovation and commercial implementation, AND
  - Moving rapidly toward increased oversight and regulation
- Europe
  - Strict Privacy Regime adopted early...
  - *BUT*, bureaucratic, slow to respond to rapidly-developing technology and changing business models, and inconsistent regulation and enforcement at the national level
- Canada
  - Generally follows the European model but more commercially focused
- Rest of World
  - Playing catch up and learn how to adopt regulation without stifling online commercial development
An Analytical Framework for Thinking about International Privacy

• When thinking about privacy and data protection in an international context:
  – Looking at the United States is important because the early adoption and commercial exploitation of new technology here raises issues not previously considered
  – Looking at the EU is important because the principles on which the EU regulation of privacy is based are gaining broad acceptance around the world (even if the implementation details vary)
  – A realistic assessment of your company’s risk tolerance and ability to comply is important, because complete compliance is difficult to achieve and often impractical
Regulation of Privacy in the U.S.

- Patchwork quilt of state and federal laws
- Existing Federal laws and regulations largely cover specific types of data relating to finance, health or children (e.g., GLB, HIPAA, FCRA, COPPA)
- About a decade ago the FTC began bringing enforcement actions in the privacy context using its authority under Section 5 of the FTC Act
  - Initial focus was on deceptive acts inconsistent with representations made in a company’s privacy notice
FTC Report – Issued March 2012
FTC Report: Privacy By Design

- Companies should adopt “privacy by design" by building Fair Information Practices into everyday business operations:
  - Provide reasonable security for consumer data
  - Collect only the data needed for a specific business purpose
  - Retain data only as long as necessary to fulfill that purpose
  - Safely dispose of data no longer being used
  - Implement reasonable procedures to promote data accuracy
FTC Report: Simplified Choice

• Companies should provide choices to consumers about their data practices in a simpler, more streamlined way than has been used in the past

• Inferred Consent for Obvious Practices:
  – Companies can engage in certain commonly accepted or obvious practices where consent is properly inferred

• Examples cited by FTC include product and service fulfillment, internal operations such as improving services offered, fraud prevention, legal compliance, and first-party marketing
  – However, the FTC has not provided guidelines for defining what practices are “obvious or commonly accepted”
FTC Report: Simplified Choice

• No Inferred Consent Where Practices Are Not Obvious:
  – For data practices that are not “commonly accepted,” consumers should be able to make informed and meaningful choices
  – Choices should be clearly and concisely described and offered – and at the point the consumer is making a decision about data
FTC Report: Transparency

• Companies should make their data practices more transparent to consumers
  – Improve their privacy policies and make them more consistent and easier to read so parties can better compare data practices and choices
  – Provide consumers with reasonable access to the data that companies maintain about them, particularly for companies that do not interact with consumers directly, such as data brokers
  – The extent of access should be proportional to both the sensitivity of the data and its intended use
  – All entities must provide robust notice and obtain affirmative consent for material, retroactive changes to data policies
Current Regulation of Privacy Doesn’t Always Work for Big Data

- Obtain consent of data subject (or at least notify and permit opt-out)
  - Since data can be stored longer and combined with other data, and since technologies are constantly evolving to permit new forms of analysis, notice to the consumer at the time of collection is becoming difficult
  - Big Data often involves multiple data sets collected at different times for different reasons, so even ascertaining the scope of “notice” or “consent” is difficult
  - Since Big Data analytics often reveal correlations that were not contemplated, it is not even possible to state the purpose
Current Regulation of Privacy Doesn’t Always Work for Big Data

• Collect only the information needed for the immediate purpose
  – The power of Big Data is its ability:
    • Perform analytics across data sets and see correlations you were not looking for
    • Thus, the assumption that data is collected for a single purpose may no longer be useful

• Retain only as long as necessary
  – Cheaper storage permits longer retention and historical analytics
  – Retention for a “necessary” time assumes collection for a particular purpose
Current Regulation of Privacy Doesn’t Always Work for Big Data

• Provide consumers with reasonable access to the data that companies maintain about them, particularly for companies that do not interact with consumers directly, such as data brokers
  – Big Data analytics often are performed against multiple data sets collected at different times by different parties for different purposes
  – Providing access to the consumer is often not practical

• All entities must provide robust notice and obtain affirmative consent for material, retroactive changes to data policies
  – Any given data set may have been collected under multiple privacy notices, and it is not always possible to track the policies under which data sets were collected
Practical Tips for In-house Counsel
Advising your Project Team

- What is your goal? What are you trying to accomplish?
- What data collection will happen? Is it personally identifiable? How will you define “personally identifiable”?
- If collecting personally identifiable information, then consider your privacy policy:
  - What does your policy say you collect? What does your policy say you will do with the data you collect?
  - How will you address notice and choice? Do you have regulatory obligations that overlay your activity?
  - Are you part of an enforcement, conciliation, or settlement agreement that gives you heightened obligations?
- How long will this information be helpful?
  - How will you get rid of it when it isn’t helpful anymore?
Analyzing Big Data Efforts

• Do you need personal data at all? If so, can it be anonymized?

• Privacy Impact Assessment
  – How will this affect people concerned?
  – Are you after general trends or will what you are doing affect individuals?

Source: “Big Data and Data Protection,” (UK) Information Commissioner’s Office, 2014
Analyzing Big Data Efforts (Cont’d)

• Repurposing Data
  – Is the new purpose compatible with the original one? Do you need consent again?
  – If you are acquiring the data from elsewhere what due diligence is needed?

• Keeping Data
  – Do you really need to stockpile the data?
  – What long term uses might justify keeping it?
Analyzing Big Data Efforts (Cont’d)

• Transparency
  – Be as open as possible about what you are doing
  – Explain purposes, implication, and benefits
  – Convey information to the people concerned in innovative ways

• Subject access to data
  – Will people be able to see the data you are processing about them?
  – What will you do to make this simple for them?
Practical Tips for In-House Counsel

• Some Contracts that Can Come into Play
  • Technology contracts
  • Cloud computing (SaaS, IaaS, PaaS)
  • Acquisitions (data scientists, algorithms)
  • Data licenses
  • Software licenses
  • Data analytics service agreements
  • Data management and storage contracts
  • Data center agreements
  • Service agreements (data privileges, data access)
Practical Tips for In-House Counsel

**Build a Team:** Privacy and security lawyers, IT security professionals, big data experts, international data export lawyers (as needed), sourcing/procurement professionals, data scientists, and others

**Know Your Policies & Customer Terms:** Company definitions of confidential and “highly sensitive” information, policies on use of customer data, record retention policies, security policies, and T&Cs with customers (contracts, online terms, web site terms, privacy policies, and more)

**Vendor Due Diligence:** Before entering into a contract
10 CONTRACT CONSIDERATIONS

1. DATA AND INFORMATION SECURITY (IS) TERMS
   - Define “Company Information” and state ownership
   - Vendor must comply with all Company rules, security policies, and information security terms (robust terms)
     - IS Program with minimum requirements
     - Immediate notice to Company when data incidents
     - Use of Company Information is restricted
     - Mobility and transfer of data
     - Level of security certification required
Practical Tips for In-House Counsel

– *No Aggregated Use*: Vendor cannot reconfigure, de-identify, or aggregate Company Information

– *Segmentation*: Logical separation of Company data

– Location and use of Company Information
  • Access and control of data
  • *Third Party Access*: No subcontracting, unless...
  • Copies, removal, and secure return of Company Information

2. CONFIDENTIALITY TERMS

3. SERVICE LEVEL AGREEMENTS ("SLAs")

– Availability of service and data (cloud) (hosting and up-time commitments)
Practical Tips for In-House Counsel

4. USER ACCEPTANCE TESTING (Company Tests)

5. SECURITY CERTIFICATIONS AND SECURITY AUDITS
   - Vendor must audit the security of computers and computing environment used in processing Company’s data
   - Audit report that verifies compliance (remediate)
   - Vendor will continue to improve, customize, and enhance the technology, at its expense, to keep pace with technological advances in the industry.

6. INTERNATIONAL LEGAL TERMS
   - Local country addendum (as needed)

7. WARRANTIES (with meaningful remedies)
Practical Tips for In-House Counsel

8. VENDOR’S INDEMNIFICATION OBLIGATIONS
   – full indemnity for Company for any Vendor breach of IS terms, including any “data incidents”

9. LIMITATIONS OF LIABILITY
   – no cap on Vendor’s liability for breach of security terms
   – no disclaimer of consequential damages for breach of IS

10. INSURANCE
    – Network security insurance required of Vendor
Thank You!