Information Governance Meets Big Data Analytics

PRESENTED BY

LINDA SHARP

BENNETT BORDEN

ZLTECH

Drinker Biddle®
Speakers

**Linda Sharp** – Associate General Counsel, *ZL Technologies*

**Bennett Borden** – Chief Data Scientist and Chair of the Information Governance Group, *Drinker Biddle & Reath, LLP*
Information Governance: Defined

“Information Governance is the activities and technologies that organizations employ to maximize the value of their information while minimizing associated risks and costs.”

- IG is a coordinating function across information stakeholders
- Every business/organization is different
- IG is an ongoing activity, not a one-time project

Source: Information Governance Initiative Annual Report, August 2014
There are two main areas where information governance aligned with Big Data bring value to organizations:

**Consumer data, client data, prospects**
- Provide better more efficient products
- Timely meet customer expectations
- Drive the industry or market

**Enterprise data stores**
- Reduce overhead to the organization
- Gain greater insight to create economies of scale
- Improves profit margins
Every Company is a Data Company
Information is Overwhelmingly Electronic
And data is coming from more sources than ever
Insight into Human Conduct to Unparalleled Degree
Unparalleled Market Efficiency

- The Right (Public or Private) Product or Service
  - Right time
  - Right place
  - Right consumer
  - Right cost
  - Right price
Unparalleled Potential for Disruption and Misuse
The Future of Analytics
Analytics in the Market Place
Cumulative Economic Impact
Putting people into groups and offering them stuff

This is the definition of **discrimination**
Algorithmic Bullying

United Airlines Passenger Is Dragged From an Overbooked Flight

By DANIEL VICTOR and MATT STEVENS  APRIL 10, 2017
The Role of Algorithms

- United boards a plane, it isn’t overbooked, everyone is seated.
- A deadheading flight crew of four arrives at the gate.
- A scheduling algorithm decides that the crew has priority over passenger revenue.
- A financial algorithm authorizes an offer of $800 (and no more) to volunteers to leave the plane. No one takes the offer.
- A customer value algorithm determines which four passengers will cause the least lifetime revenue loss.
- Stock price drop: $225M

Credit: John Robb
How Do We Decide Who Gets What?

- College admissions
- Financial Aid
- Street Repair
- Snow and Trash Removal
- Electricity Distribution
- Voting Districts
- Health Care Resources
- And many many more....
Current Laws and Regulations Don’t Fit Current Applications of Analytics

In 2012 Facebook conducted an “emotional contagion” study, manipulating the display of happy and sad content in the news feeds of 150,000 users to see if they would share happy or sad content.
Notice and Consent
Data Ownership
Acting in a Legal Greenfield
Hello Barbie vs. Amazon Echo

Understand the risks you are creating and act reasonably to mitigate them

\textit{value to advertisers.} - Angela Campbell, Esq, Director of Communications and Technology Clinic, Georgetown Law

#HellNoBarbie
The Way Forward

- The law is based on reasonable conduct
- Identify and quantify risks in analytics projects
- Identify mitigation strategies
- Include a diversity of opinion
The Future

Analytics

- Fast, Accurate Answers
- Unparalleled Sociological Record
- Unparalleled Level of Certainty
- Unparalleled Potential for Abuse
- Like every advent of new technology, Analytics are Highly Disruptive

disruption
Enterprise Data Stores
Why is Data Analysis a Significant Part of IG?

“Dark data is the information assets organizations collect, process and store during regular business activities, but generally fail to use for other purposes (for example, analytics, business relationships and direct monetizing). Similar to dark matter in physics, dark data often comprises most organizations’ universe of information assets.”

—Gartner
What Data Are You Targeting?

- Customer or consumer centric data – most common use, several analytics tools in the market today

- Employee centric or business data
  - **Structured data**
    - Sits in databases
    - Easily searchable
  - **Unstructured data** – the Wild Wild West
    - Email systems
    - File Shares
    - SharePoint sights
    - Social Media
What is the Defined Use for the Data?

- Regulatory Requirements – HSR, EU Requirements
- Divestures, Mergers/Acquisition
- eDiscovery
- Repurpose and reuse of information by business teams and executive teams
- Conduct analytics to gain a greater understanding of what you have in your systems
General Process for Analyzing Enterprise Content and Pitfalls

- Identify data required
- Make a copy
- Export to a third party provider to massage the data and glean reports
- Delay in knowledge assimilation and transfer
  - Decisions are not based on current state of affairs
Failure to Clean Up Data Stores

- Increased eDiscovery costs
- Effects data analytics and viability of information effects on business decisions
- Increased risk due to potential cyber attacks, if you don’t have it, they cannot get it
- Overall storage costs
- Ability to repurpose and reuse the information
What’s Abandoned Data Costing You?

written by Jim McGann October 2, 2014

“The average corporate turnover rate for employees is 15.1 percent across all industries, with some specific verticals experiencing as high as 30%. For an organization with 10,000 employees this can account for 1,500 to 3,000 people annually (Compensation Force: 2013 Turnover Rates by Industry).”

“When an employee leaves an organization the IT department will typically wipe or recycle their hard drive, containing their digital files and email, however, they neglect to clean and manage former employees’ data on corporate networks and servers.”
Your “Mashup” of eDiscovery Market Estimates is Even Earlier This Year: eDiscovery Trends

“The eDiscovery Software and Services market is expected to grow an estimated 15.31% Compound Annual Growth Rate (CAGR) per year from 2016 to 2021 from $7.997 billion to $16.304 billion per year. Services will comprise approximately 69.9% of the market and software will comprise approximately 30.1% by 2021.”

Doug Austin, 3/6/17, CloudNine
http://www.jdsupra.com/legalnews/your-mashup-of-ediscovery-market-22808/

At whose expense??
Benefits of an Organized Environment

- Reduction in storage costs
- Reduction in risk, if you don’t have it, they can’t get it
- Ease in locating information when needed for eDiscovery or by the business units
- Ability to repurpose and reuse prior art
Multi-Jurisdictional Issues Around Data Management

- Understanding issues around creating and manipulating data about and of minors
- Rights of minors to require that information regarding them be removed
- Right to be forgotten, etc.
More to Consider

“...In Depth Into The $519 Million Teva FCPA Enforcement Action
Mike Koehler  December 22, 2016
DOJ and SEC announced... a $519 million enforcement action against Teva Pharmaceutical Industries Ltd. (an Israeli company with American Depository Receipts traded in the U.S.) and a related entity. The settlement amount included a $283 million DOJ component and a related $236 million SEC component.”

http://fcpaprofessor.com/category/teva-pharmaceuticals/

4.5 years of investigations, attorneys fees, executive level terminations and negotiations with the Feds.
Key Takeaways

1. **Identify Project Drivers to Secure Leadership Buy-In**
   - Cost
   - Security
   - Legal/governance

2. **Take an Internal Inventory**
   - What data types need to be analyzed?
   - Which locations/departments/etc. have the most high risk data?

3. **Understand Available Technologies**
   - What technologies support a long-term approach and proactive information governance?
We Live in a Virtual World!
GDPR Compounds the Problem

- May 25, 2018 – is rapidly approaching.
- Who is affected? Article 3 of GDPR provides that any companies that process “personal data of data subjects located in the EU”, regardless of where they are headquartered, established or their venue, the Regulation applies.
- It affects ANY organization that markets goods or services to EU residents.
- Applies regardless of payment.

What if you don’t formally market to them, but they order from your website???
GDPR

- The GDPR - 4 May 2016, and goes into effect 25 May 2018 targets any organization that operates in the EU and processes the personal data of EU subjects.
- Non-compliance may lead to huge fines of up to €20m or 4% of total annual worldwide turnover, whichever is higher.
- Applies to companies established in the EU who controls or processes customers, employees, or other personal data.
- Applies to companies established outside the EU that offer goods or services to EU residents or monitor the behavior of EU residents.

- Does GDPR affect your company? Why not?
- What are you doing to prepare the organization?
- Privacy Shield
- Right to be forgotten
What Information Do You Divulge?

In creating an account for almost any kind of website. They use this information to allow you to “reset your password”.

- Name your first dog,
- What high school did you go to,
- What is the favorite place you lived.

Certain of this information can also be used to look at character traits.

How many people that visit the website selected the “first dog” question. Perhaps they then make sure that dog products pop up when you visit their site, assuming that you must be a dog person.
Potential Violations – help create an ROI

- Article 79 – any company that violates certain provisions under GDPR, i.e. Processing (Article 5) or cross border data transfers – subject to fines up to 4% of the “company’s total worldwide turnover.”
Questions?

Bennett B. Borden  
*Chief Data Scientist*  
**Drinker Biddle & Reath LLP**  
1500 K Street, N.W.  
Washington, DC 20005  
(202) 230-5194  
*Bennett.Borden@DBR.com*

Linda G. Sharp, Esq., MBA, CLSP  
*Associate General Counsel*  
**ZL Technologies**  
860 N. McCarthy Blvd., Ste. 100  
Milpitas, CA 95035  
(323) 333-0782  
*lisharp@zlti.com*