



The MIT 2008 Information Quality Industry Symposium



## 'Fit For Use' to a Fault

Presented by

**Deborah Henderson, Tandum Lett,  
Anne Marie Smith and Cora Zeeman**

**July 2008**



The MIT 2008 Information Quality Industry Symposium



<b>Authors</b>	<b>Affiliations</b>
Deborah Henderson	CapGemini
Tandum Lett	Sullivan & Cromwell, LLP
Anne Marie Smith	EWSolutions
Cora Zeeman	University of Toronto



The MIT 2008 Information Quality Industry Symposium



## Presentation Objectives

- Introduce the concept of professional ethics in data management
- Inspect the behaviour of corporations in the drive for increased data quality and possible ethical implications
- Suggest a dimensional model of ethical challenges
- Present example ethical dilemmas
- Propose professional advocacy in the workplace around ethical handling of data.



The MIT 2008 Information Quality Industry Symposium



## Defining Ethics

- Compelled ethics - voluntary code, personal responsibility, signed code of ethics as a condition of professional affiliation
  - DAMA
  - ACM
- Imposed ethics – regulatory, cultural, legal framework
  - A comparison of approaches on privacy; Canada and the United States



The MIT 2008 Information Quality Industry Symposium



## US Privacy Law – Self Regulatory

US privacy law is self-regulatory

- Under the self regulatory regimes organizations design and implement their own privacy programs based on the criteria set down by the FTC :
- **Notice:** data collectors must disclose their information practices before collecting personal information from consumers;
- **Choice:** consumers must be given options with respect to whether and how personal information collected from them may be used for purposes beyond those for which the information was provided;
- **Access:** consumers should be able to view and contest the accuracy and completeness of data collected about them;
- **Security:** data collectors must take reasonable steps to assure that information collected from consumers is accurate and secure from unauthorized use; and
- **Enforcement:** the use of a reliable mechanism to impose sanctions for noncompliance with these fair information practices.

CZ



The MIT 2008 Information Quality Industry Symposium



## Canadian Privacy Law – the regulated approach

Canadian privacy law is a hybrid between a comprehensive regime of privacy protection with industry self regulation. **PIPEDA (personal information protection and electronic documents act)** covers all businesses who collect, use and disseminate personal information in the course of commercial activities and stipulates rules with exceptions that organizations must follow in the collection, use and dissemination thereof. The Act codifies the industry created privacy guidelines of the CSA Model Code for the Protection of Personal Information. The 10 guidelines set out therein are statutory obligations that all organizations that collect, use and disseminate personal information must follow.

- **Accountability:** An organization is responsible for personal information under its control and must designate an individual to be accountable for the organization's compliance with the principles;
- **Identifying Purposes:** An organization must identify the purposes for which personal information is collected at or before the time the information is collected;

CZ



The MIT 2008 Information Quality Industry Symposium



## Canadian Privacy Law – the regulated approach

- **Consent:** An organization must obtain the knowledge and consent of the individual for the collection, use or disclosure of personal information, except where inappropriate;
- **Limiting Collection, Use, Disclosure and Retention:** The collection of personal information must be limited to that which is necessary for the purposes identified by the organization. Information shall be collected by fair and lawful means. Personal information shall not be used or disclosed for purposes other than those for which it was collected, except with the consent of the individual or as required by law. Personal information shall be retained only as long as necessary for the fulfillment of those purposes;
- **Accuracy:** Personal information must be as accurate, complete, and up-to-date as is necessary for the purposes for which it is to be used;
- **Safeguards:** Personal information must be protected by security safeguards appropriate to the sensitivity of the information;

CZ



The MIT 2008 Information Quality Industry Symposium



## Canadian Privacy Law – the regulated approach

- **Openness:** An organization must make specific information about its policies and practices relating to the management of their personal information readily available to individuals;
- **Individual Access:** Upon request, an individual shall be informed of the existence, use and disclosure of his or her personal information and shall be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate, and;
- **Challenging Compliance:** An individual shall be able to address a challenge concerning compliance with the above principles to the designated individual or individuals accountable for the organization's compliance.)

The federal privacy commissioner has the sole responsibility for handling privacy complaints against organizations. However, she fills an ombuds role wherein her decisions are recommendations and **not legally binding** and her decisions have **no precedential value, even within her own office.**

CZ



The MIT 2008 Information Quality Industry Symposium



## IT Compliance

Organizations often belong to privacy programs like **BBBonline** and **eTrust** which, along with providing organizations with privacy guidelines, monitor their compliance with the guidelines.

Complaints against any organization for violation of privacy can be brought to the FTC under their control over unfair and deceptive commercial practices.

CZ



The MIT 2008 Information Quality Industry Symposium



## Crossing the Line on Ethical Use

- Ubiquitous collection and use data
- Younger demographic don't see this as an issue
  - Facebook, YouTube, 15 minutes of fame
  - My privacy is not an issue ...as long as it doesn't affect me negatively
  - Understanding the issues

Ethical issues can only partially be monitored by automated means

CZ



The MIT 2008 Information Quality Industry Symposium



## Ethics and Data Management: an Analysis of Ethical Risk

Putting yourself in the hands of Corporate data and  
business intelligence analysts – what they do :

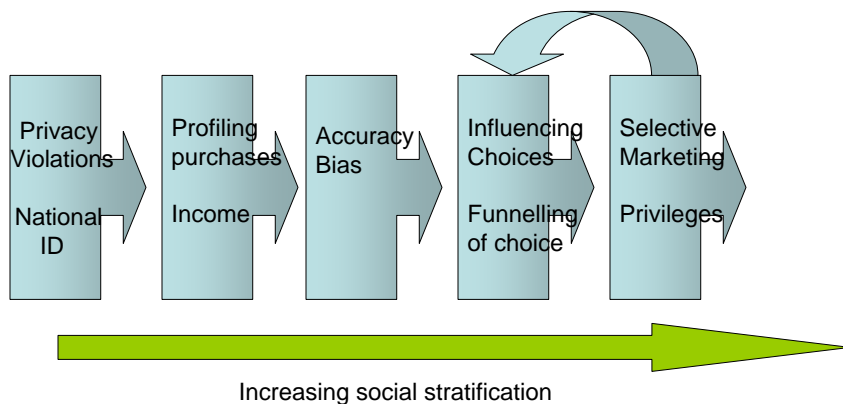
- WHO you are - Identity and theft, terrorist and criminal identification - **National Identity Schemes**
- WHAT you do – Profiling - Big Brother IS watching !
- WHEN you do it – Timing of data analysis, Bias, accuracy
- WHERE you do it – profiling and corralling your choices
- HOW are you treated - Credit scoring, preference tracking – ultimately privileged or not



The MIT 2008 Information Quality Industry Symposium



## Ethics and Data Management: an Analysis of Ethical Risk



**DO YOU WANT TO LIVE HERE?**



The MIT 2008 Information Quality Industry Symposium



## Who is Collecting Information on You?

- Government – census, taxes, any funded program you participate in
- Anywhere you go on-line – cookie, re-selling
- How you spend your money – RFID tagging connected to your credit card information?
- Employers and any organization you belong to

AS



The MIT 2008 Information Quality Industry Symposium



## Protection We Like

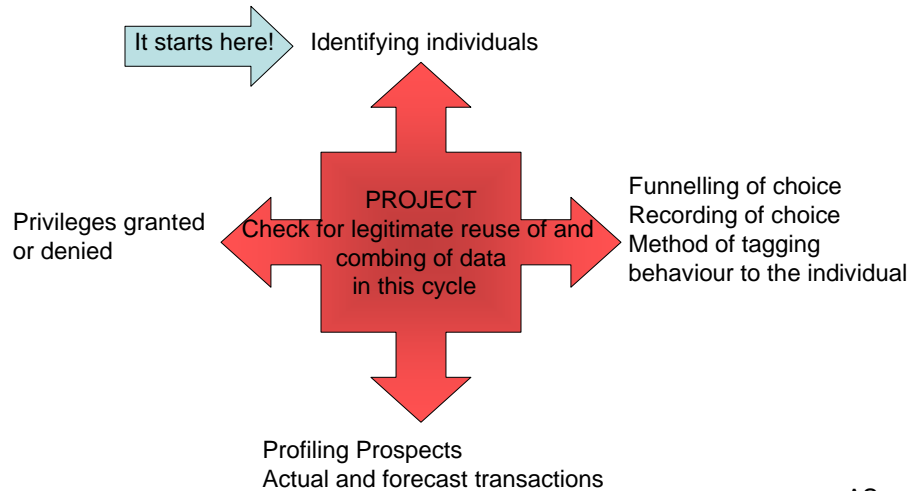
### We will give up our privacy for these types of concerns:

- **Personal protection** - Credit card real time tracking of purchases protects us from fraud
- **Greater good** - Money laundering analysis by the Treasury (reporting of all large monetary transactions)
- **Greater good** – “No Fly” list - Protection from terror

AS



## Evaluating Ethical Risks – a Model



## Ethics and Data Management: some Examples of Ethical Risk

- Data Aggregation and Reconstitution
- Profiling - Categorizing into Strata
- Biased data collection and use
- Data Cleaning and standardization
- International movement of personal data
- 'Chinese wall' (non-porous) for information inside companies





The MIT 2008 Information Quality Industry Symposium



## Categorizing into Strata

MarketMath in Canada has statistically analyzed the entire population into 52 psychographic segments.

- Gray-haired small town
- Executive burbs
- Francophone upwardly mobile

They have bought and combined datasets from the Census, private companies, liquor stores, pollsters

- combines with a geographic information system
- Sells this package for target marketing

***I get the offers for service, bargains, products  
YOU don't !***

AS



The MIT 2008 Information Quality Industry Symposium



## Bias Data Collection and Use

- Hunch and Search: Analyst has a hunch and wants to satisfy hunch; only uses data that satisfies their hunch
- Data Collection for Pre-defined result: Analyst is pressured to collect data and produce results based on pre-defined desires
- Biased use of data collected: Data is used to satisfy a chosen approach, data is manipulated for a chosen approach

AS



The MIT 2008 Information Quality Industry Symposium



## Data Cleaning & Standardization

- 30% of us move households every year
- Billions of transactions for analysis
- Organizations that are data cleaning – to what standards? And selling the latest and greatest information on you for profit
- Who assesses accuracy?

TQM Cop-out we have all heard:

***“We have no separate quality control, the quality is baked-in to the process”***



The MIT 2008 Information Quality Industry Symposium



## Data Cleaning & Standardization in Identity – We give up privacy for security – Right ?

### Steps Involved

- Registration
  - Biometric sample taken, stored and compared
  - ID token issued, based on existing records
- Data-matching and profiling, quality
  - Ongoing, behind the scenes
  - Data gathering + database linkages
- Authentication (at control points)
  - Identity match between body and ID token
  - Database checks (personal data, watch list)
  - Request denied or approved

AS



The MIT 2008 Information Quality Industry Symposium



## Securely and Reliably Identify Everyone? False Sense of Security

Everyone with a 'clean' record passes

- Most 9/11 attackers had NO record of suspicion
- Terrorist training manual: "fit in" as "normal"
- Terrorists can repeatedly test screening system, then only need to pass once!

***"The positive identification of individuals does not equate to trustworthiness or lack of criminal intent."***  
(emphasis in original)

(Ben Shneiderman, USACM testimony at the Congressional Hearings on National Identification Card Systems, Nov 2001)

AS



The MIT 2008 Information Quality Industry Symposium



## Movement of Data, internationally

- Privacy laws differ worldwide
- Corporations cannot move employee personal data (identification) across borders without reviewing laws of sending and receiving countries.
- May need to set expectations of 'single sign-on' in large multi-nationals



The MIT 2008 Information Quality Industry Symposium



## **‘Chinese Walls’ in Corporations**

- Legal firms may represent both the plaintiff and the defendant
- IT policies need to be implemented and monitored to protect the privacy of both parties.

Recombination of information in the external business world may embarrass the firm

TL

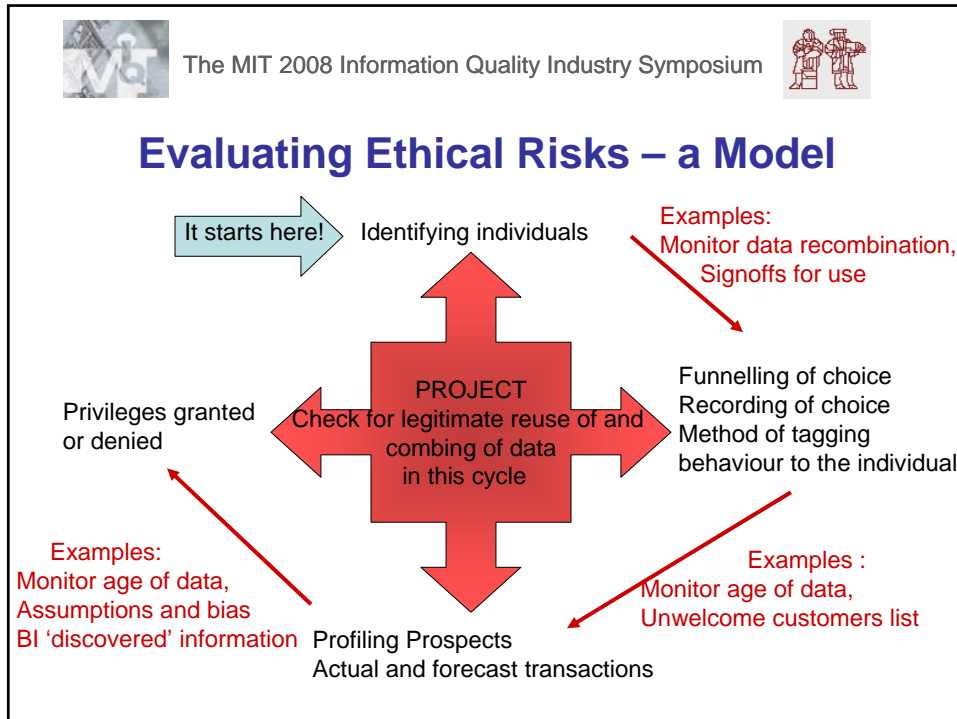


The MIT 2008 Information Quality Industry Symposium



## **Call to Action**

- Business is not aware of ‘where the data comes from’ and ethical issues may not be obvious to them
- Automated monitoring is not sufficient protection
- Cultural norms and ethics in the workplace influence corporate behavior
- Taking a Professional stand
- Acting ~! Evaluate the risk/benefit
  - Take and “index” of Ethical Risk based on our model



- 
- ## Conclusion - What do we\* do now?
- \* **We = Professionals/Associations + citizens**
- Convene and participate in public data ethics forums
    - DAMA, MITIQ!.. Get more going!
  - Resist emphasis on overly costly, unreliable, narrowly technological approaches
    - What are the purposes? Would it be effective?
    - Who is being served? Disadvantaged?
    - What are the alternatives?
  - Demand social and political accountability
- TL



The MIT 2008 Information Quality Industry Symposium



## Questions



The MIT 2008 Information Quality Industry Symposium



## Contacts

- Deborah Henderson:  
[deborah.henderson@inergi.ca](mailto:deborah.henderson@inergi.ca)
- Tamdum Lett:  
[LettT@sullcrom.com](mailto:LettT@sullcrom.com)
- Anne Marie Smith:  
[AMSmith@ewsolutions.com](mailto:AMSmith@ewsolutions.com)
- Cora Zeeman:  
[kitytje@hotmail.com](mailto:kitytje@hotmail.com)