LEGACY DILEMMA

• **Yesterday**
  - Main frames would create both tangible and intangible risks to individuals and a free society
  - Early thinkers thought the best way to protect individuals was by assuring individual autonomy
  - Led to OECD Guidelines and eventually Canadian law
  - Made sense based on pre-database mainframe technology
    - Data provided mostly by the individual
    - Systems and data were the same

• **Today**
  - Generation after generation of new technological shocks
    - Today it is observational world – think Internet-of-Things
  - Advanced analytics drive insights which become inferred data
  - Autonomy --- individual consent --- alone no longer protects against harms
THE ISSUE

• Predicting future outcomes – no matter the endeavor – is compelling
  • Big data is making predictions based on non-intuitive data relationships

• There is exponentially more data in an observational world to drive predictions

• This creates the dilemma of whether the processing is legal, fair, and just?

• Resolving that dilemma requires new ways of assessing processing and managing mitigation
IF ISSUE IS NOT RESOLVED

• A lack of clear thinking leads first and foremost to reticence risk
  • Your organization doesn’t use data aggressively enough to drive innovation
  • Reticence risk at others will drive norms hurting whole industries

• Legal and regulatory bright lines that prohibit data creation and use

• Lost trust – a killer to brands
WHY IS THE BIG DATA DIFFERENT?

• **Volume, velocity, diversity** – blending data robustly
  • More room for processing errors
  • Diverse data sets stretch compatibility to stated purpose

• **Goes beyond testing best intuition** – data drives questions
  • Data correlations drives further questions and then insights

• **Correlation not causation**
  • Is the correlation meaningful
SO WHY BIG DATA?

• Big data sees things that go beyond intuition, and that vision improves outcomes
  • Health
  • Disaster relief
  • Improved markets

• The improvements are compelling and inevitable

• But potential discrimination and mistakes must be identified

• Governance must be in place to facilitate the compelling and prevent inappropriate results
THE DILEMMA EXISTS FOR ALL STAKEHOLDERS

• Individuals have interests that go well beyond compliance, including better health, education, opportunities and choices

• Companies can only do with data what is acceptable, and acceptable does not equal compliant

• Enforcement agencies are tasked with protecting individuals while facilitating fair data driven innovation
POSSIBLE ANSWERS

• Depend on very long complex notices and implied consent
  • Is this really purpose specification?

• Exemptions to consent established by law
  • Not very flexible

• Just process and hope the regulator never finds out
  • Works for some but not very compliant

• Processing based on demonstrable governance
  • How can one trust the organization?
LEGITIMATE INTEREST

As described in EU Directive

• processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed

• except where such interests are overridden by risks to fundamental rights and freedoms of the data subject ---
BIG DATA MUST BE LEGAL, FAIR AND JUST

CIVIL LAW DATA PROTECTION
• All processing must have a legal basis
  • Where consent is ineffective one must demonstrate
    • Fair
    • Just

U.S. COMMON LAW
• Processing must not be illegal
  • But absence of legal prohibition is not enough, must be:
    • Not unfair
    • Just
IAF AND BIG DATA WORK

• 2013 – Two phase (discovery and application) approach by Abrams, Bruening and Leta

• 2014 began exploring assessment process for big data
DIFFERENTIATING PRIVACY AND DATA PROTECTION

• Many definitions of data protection and privacy

• For purposes of this discussion
  • Privacy relates to the protection of individual autonomy and family life
  • Data protection is foundational for the protection of the full range of individual interests in an information society

• Data compatibility
  • Within context
  • Not incompatible with individual’s understanding
KEY GOVERNANCE CONCEPTS

- Data protection assures the full range of individual interests, not just a narrow definition of privacy
- Reticence risk is meaningful and real
- A fair and just assessment is necessary – what is the cost of not processing
- This takes us beyond compliance to ethics
A Unified Ethical Frame for Big Data

Big Data provides **unprecedented opportunities** to drive innovation in economies, healthcare, public safety, education, transportation, and almost every human endeavor.

Big Data also **creates risk** to both individuals and society unless effective governance is in place.

Governance must be sensitive to the full range of interests. **Ethics must form the basis of decisions** balancing the incentives that come with free enterprise and the breadth of human interests.
Big Data Code of Ethical Practice – Structure

Part A
- Description of the unified ethical framework.
- Creates a basis for the interrogation guidance.

Part B
- Interrogation guidance for implementing the code.
  - Illuminates the key issues that must be considered in making a judgment on whether a Big Data project is fair, responsible and ethical

Part C
- Mechanisms for enforceability.
  - Ensures compliance with code.

Part D
- Contextual interrogation questionnaire.
  - Customized for organizations, industries, mediums.
Big Data Code of Ethical Practice – Part A

- Big Data requires an assessment against a variety of factors, including existing laws as well as the ethical frames of the stakeholders impacted by the processing.

- This requires an assessment of the intersection between the various ethical frames; resulting in a more united ethical frame (an approximation).

- To create that approximation, the Foundation adopted five key values that when taken as a whole assist in assessing this approximation.

- The unified ethical code makes use of the full range of fundamental rights as recognized by data protection law.
### Values for an Ethical Frame

<table>
<thead>
<tr>
<th>Values</th>
<th>Criteria</th>
</tr>
</thead>
</table>
| Beneficial | • Sense of purpose  
• Define the benefits  
• Parties who gain benefit  
• Risk-Benefit balance |
| Progressive| • Materially better  
• Measures  
• Support innovation  
• Consider risks |
| Sustainable| • Legal basis/permissions  
• Influence of model(s)  
• Longevity of insights  
• Ability to refine/correct |
| Respectful | • Data origination/restrictions  
• Context/Purpose  
• Interests of all parties  
• Expectations/Rights |
| Fair       | • Insights/Applications  
• Inferences/Discrimination  
• Labeling/Predestination  
• Compatibility/Legal Basis |
PART B – ASSESSMENT FRAMEWORK

- Full project description with clear intents
- Questions to identify issues related to processing and accountability
- Clear description of stakeholders
- A description of intended benefits and possible risks
- A final assessment of fairness based on facts
WHEN SHOULD ASSESSMENT TAKE PLACE

• Discovery
  • Project scoping
  • Discovery processing

• Application
  • Using insights to predict behavior
  • Reviewing for effectiveness
Big data enforcement - Part C

1. Linkage to DPA's Legal Mandate

2. Code Conduct or Practice


4. Mechanism for DPA Receiving Attestation

5. Mechanism for Data Sub Complaints

6. Mechanism for DPA Spot Checks

7. Mechanism for Re-Attestations
MARKETING SUBGROUP CHALLENGE PART D CHALLENGE

• Create an assessment tool for marketing
  • Based on the unified ethical frame
  • Would be approachable for organization both facilitating and using big data for digital marketing
  • Would enhance decision making
  • Be describable to stakeholders
  • Demonstrable to regulators
Big Data Analytics – Is the Project Fair?

Fairness

Project Purpose

Sources ➔ Prep (ETL) ➔ Insights ➔ Application

Legal Obligations ➔ Accountability ➔ Stakeholders ➔ Benefits/Risks/Mitigation
### Big Data Analytics – Ethical Frame

<table>
<thead>
<tr>
<th>Purpose Interrogation</th>
<th>Source Interrogation</th>
<th>Preparation Interrogation</th>
<th>Insights Interrogation</th>
<th>Application Interrogation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Overview</td>
<td>All Sources</td>
<td>Consistent Formatting</td>
<td>Expected</td>
<td>Better than Current</td>
</tr>
<tr>
<td>Project Sensitivity</td>
<td>All Elements &amp; Sensitivity</td>
<td>Other Synthesizing</td>
<td>Unexpected</td>
<td>Positive Stakeholder Outcomes</td>
</tr>
<tr>
<td>New or Expanded Insights</td>
<td>Origin of Data</td>
<td>Integration Technique</td>
<td>Use(s)</td>
<td>Negative Stakeholder Outcomes</td>
</tr>
<tr>
<td></td>
<td>PII Linkability</td>
<td>Integration Accuracy</td>
<td>Acceptable Accuracy/Use</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data Structure</td>
<td></td>
<td>Sensitivity/Use</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Source Transparency for Use</td>
<td></td>
<td>Useful Life</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Source Accuracy</td>
<td></td>
<td>Value Over Time</td>
<td></td>
</tr>
</tbody>
</table>

- **Purpose Interrogation**
  - Project Overview
  - Project Sensitivity
  - New or Expanded Insights

- **Source Interrogation**
  - All Sources
  - All Elements & Sensitivity
  - Origin of Data
  - PII Linkability
  - Data Structure
  - Source Transparency for Use
  - Source Accuracy

- **Preparation Interrogation**
  - Consistent Formatting
  - Other Synthesizing
  - Integration Technique
  - Integration Accuracy

- **Insights Interrogation**
  - Expected
  - Unexpected
  - Use(s)
  - Acceptable Accuracy/Use
  - Sensitivity/Use
  - Useful Life
  - Value Over Time

- **Application Interrogation**
  - Better than Current
  - Positive Stakeholder Outcomes
  - Negative Stakeholder Outcomes
  - Public Policy Impact
Big Data Analytics – Ethical Frame

<table>
<thead>
<tr>
<th>Contractual and Legal Interrogation</th>
<th>Accountability Interrogation</th>
<th>Stakeholder Interrogation</th>
<th>Benefits &amp; Risks Interrogation</th>
<th>Fairness Interrogation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Obligations</td>
<td>Accountable Senior Leader</td>
<td>Source Stakeholders</td>
<td>Benefits for Stakeholders</td>
<td>Unfairness to Individuals</td>
</tr>
<tr>
<td>Application Obligations</td>
<td>Other Project Leaders</td>
<td>Insight Stakeholders</td>
<td>Risks for Stakeholders</td>
<td>Issues from Project</td>
</tr>
<tr>
<td>Security Obligations</td>
<td>All Leaders Comfort with Project</td>
<td>Use Stakeholders</td>
<td>Mitigations for Each Risk</td>
<td>Balance of Residual Risks and Benefits to Individual and Society</td>
</tr>
<tr>
<td></td>
<td>All Stakeholder Concerns</td>
<td>All Stakeholder Concerns</td>
<td>Residual Risks</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Risk/Benefit Balance</td>
<td></td>
</tr>
</tbody>
</table>
CANADIAN CHALLENGE – WHAT IS LEGITIMATE?

• Canadian law has room for “reasonable” uses of data where consent is not fully effective?
• Is legal, fair and just a proxy for “reasonable”?
• Could a balancing process be demonstrated?
• Would a demonstrable process be trusted?
• Using authority to encourage the market could the regulator issue guidance which is an extension of accountability guidance?
SO WHAT IS NEXT?