Contextual Assessment Worksheet
for the Marketing Industry

IAF Big Data Ethics Initiative, Part D

Draft
08 October 2015

Introduction

The purpose of interrogation is to identify the issues that must be resolved to assure a big data project is fair to the full range of stakeholders (see Part A, Unified Ethical Frame). The questions below have been designed to illuminate those issues for decision makers and create a record for review. For consistency purposes, it is best to stick as close as possible to these questions. However, these questions may be modified if different wording would be better understood within an industry. As long as the revised questions still identify the key issues so fair decisions will be made, changed wording is appropriate.

There are points where an ethical analysis might be appropriate. Some of these are:
1. Scoping when a big data project is first conceived but before analysis is conducted to create insights;
2. Discovery where data is processed actually to develop insights that might be applied going forward;
3. Application where insights are applied to generate results;
4. Review where the big data project is assessed to see where the expected results are actually being achieved.
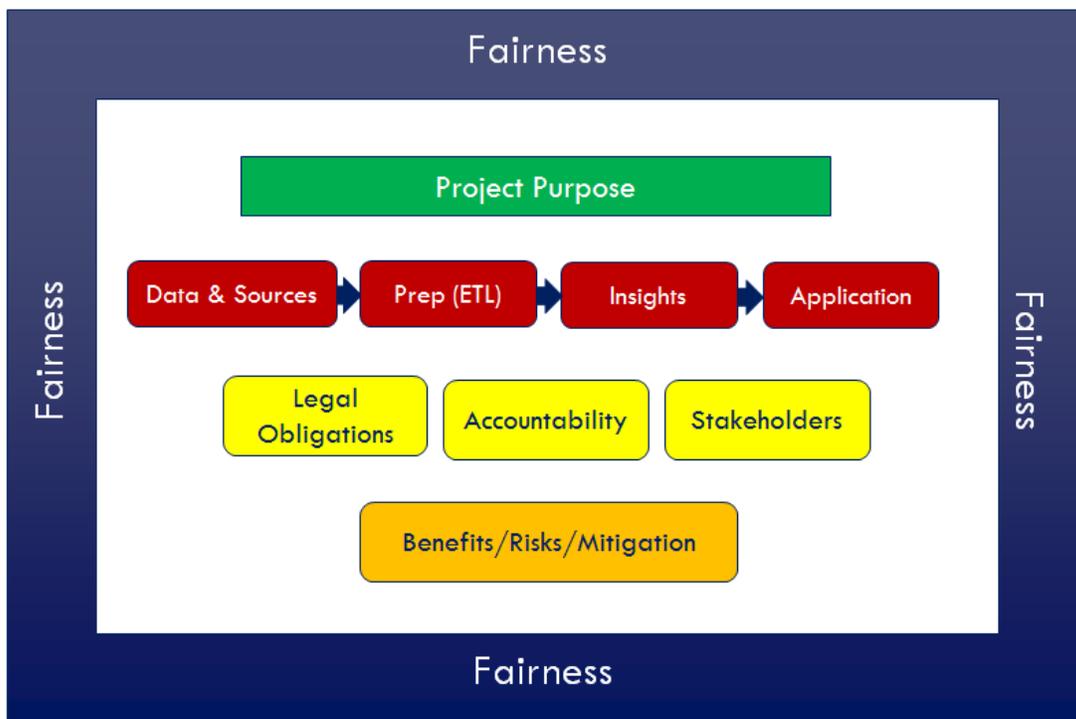
Some level of review is appropriate at each phase. Questions need not be repeated in later phases if underlying conditions have not changed. If there have been changes to the project that impact answers, they should be repeated.

While developed primarily for big data processes, this tool may be used to assess any practices within an organisation where data may be used in a manner that goes beyond the original understanding of the individuals to whom the data pertains.

Charts

The following three charts give an assessor a workflow for this assessment process. Chart 1 provides an overview of the process as a whole. The ultimate objective is an assessment that is fair to all stakeholders, with a particular emphasis on individuals. The chart's outer boundary is the fairness goal. To measure fairness, the assessor needs to understand the project's purpose (green); the processing process (red); the obligations, to whom and how they are protected (yellow); and lastly, risks, benefits and mitigation (orange).

Chart 1: Big Data Analytics – Project Assessment?

Fairness

Fairness

Project Purpose

Data & Sources → Prep (ETL) → Insights → Application

Legal Obligations   Accountability   Stakeholders

Benefits/Risks/Mitigation

Fairness

Fairness

Charts 2 and 3 provide the assessor with a snapshot of the issues to be resolved in each of the sections.

Chart 2: Big Data Analytics – Ethical Frame

Purpose Interrogation	Source Interrogation	Preparation Interrogation	Contractual and Legal Interrogation	Accuracy Interrogation	Insights Interrogation
Project Overview	All Sources	Consistent Formatting	Source Obligations	Accuracy of Sourced Data	Expected
	All Elements & Sensitivity	Other Synthesizing	Application Obligations	Quality of Final Data	Unexpected
New or Expanded Insights	Origin of Data	Integration Technique		Source Transformation	Use(s)
	PII Linkability		Security Obligations	Impact of Time	Sensitivity/Use
	Data Sensitivity	Integration Accuracy			Useful Life
	Data Structure				Value Over Time
	Source Transparency for Use				
	Obligations				
	Source Accuracy				

Chart 3: Big Data Analytics – Ethical Frame

Outcomes Interrogation	Accountability Interrogation	Stakeholder Interrogation	Benefits & Risks Interrogation	Fairness Interrogation
Better than Current	Accountable Senior Leader	Source Stakeholders	Benefits for Stakeholders	Unfairness to Individuals
Positive Stakeholder Outcomes	Other Project Leaders	Insight Stakeholders	Risks for Stakeholders	Issues from Project
Negative Stakeholder Outcomes	All Leaders Comfort with Project	Use Stakeholders	Mitigations for Each Risk	Balance of Residual Risks and Benefits to Individual and Society
Outcome Predictability		All Stakeholder Concerns	Residual Risks	
			Risk/Benefit Balance	

Contextual Assessment Worksheet for the Marketing Industry
Big Data Ethics Initiative, Part D

Version 2.0, April 2015

Questions	Explanatory Commentary	Answers
CHARACTERISING The PROJECT	These should be answered by the development team.	
<u>Purpose:</u> **Understand the purpose and intended outcomes of the project.** Provide a project overview that describes the main purpose of the project. Is the primary purpose of the project to generate new insights or to expand on insights from a previous project or previous work?	Consider such purposes as: • Marketing or risk management • Solution and product capability • Distribution network • Brand enhancement • Marketing: traditional direct mail, email, telemarketing, digital advertising. If expanding on an insight from previous project, include any previous analysis done. (Note: Data flow mapping may be a technique that can help answer these questions.)	

<u>Sources:</u> **Understand the sources of data to be used in the project.**
 	
What are all the sources of the data?
 	
What actual data elements are found in the data?
 	
How was the data from each source originated?
	Data Origins: Data may come directly from the individual to either the organisation or a third party, or may be data that was created by other simple calculations or the result of statistical analysis.
	• Provided by the individual
	• Scraped from the web
	  - Is this data allowed by the terms of use or website policy, i.e., scraping or commercialisation prohibited by website, sensor or derived data?
	• Obtained from public sources
	  - Are there limitations for data uses from public sources?
	• Provided by third-party aggregator
	• Observed in some other fashion
	• Derived from data (i.e., transformation/manipulation)	

	▪ Inferred from analytics ▪ Provided by vendor Linkability: See glossary for definitions related to linkability. ▪ Personally Identifiable Information ▪ Pseudonymous ▪ Device Identifiable Information ▪ De-Identified ▪ Aggregated See Direct Marketing Association sensitive definition. Glossary includes definition of unstructured data.	
Is the data linkable to a particular individual? If so, please indicate whether the individual is identified or not.		
	Transparency and Choice: The relationship between the individual and where the data first originates impacts both transparency and choice. If data is collected by an organisation in a direct relationship with the individual, transparency and choice both are direct and future marketing use may be expected. Future marketing use may be expected if data is collected by a third party for use by the organisation with which the individual has a relationship but may not be expected if the data collected by the third party is for use by a different marketer.	
Is sensitive data involved?		
Is the source data structured or unstructured?		
What kind of transparency and choice was given to use the data as intended in this analysis?		
Are there legal, policy, contractual, industry or other obligations that apply to the data?	Industry obligations include codes of conduct.	
How accurate is the data?	Data accuracy: Data accuracy may be directly related to how data was sourced. Data directly observed may be more precise than data inferred from an algorithm. Data	

	directly observed may be more accurate than data volunteered by individuals.	
Preparation: Understand the pre-processing that will be done before the analysis. What work will be done to put the data in a consistent format? How will errors and redundancy in the data be identified and dealt with? How will the data sources be consolidated for analysis? Will further synthesising of the data be necessary?	See glossary for definition of pre-processing of data. Companies may have standard processes to manage data preparation. If so, this section may not be necessary once the standard process has been reviewed. Steps in Preparation: ▪ Data standardisation ▪ Data hygiene ▪ Data integration (consolidation) Synthesising means deriving different data elements from various source elements. ▪ What is state of data security? ▪ Is obscuring of data necessary?	
Contractual and legal conditions: All processing and applications should be within the context of the conditions associated with the data. Has there been a listing of all obligations associated with the data? Is the data being used within the context of its origination? If the data is originated by others, are conditions on the data being respected?	This links to the respectful value. See glossary for definition or Part A for context. As a general obligation, marketing uses of data should be open and above board. Organisations should have means of making their use of data for marketing visible to individuals, thus making related choices easy to exercise. The answers to these questions relate to transparency. Obligations associated with the data include: ▪ Laws ▪ Policies ▪ Contracts ▪ Industry codes	

If the project moves forward, will the project security be adequate/proportional to the risks related to the data? Would the application of insights be seen as ethical and respectful if publicly exposed?	Based on how the data was originated, is it reasonable that the data would be used in this way? Conditions associated with data should be respected as the analysis progresses.	
Accuracy: Evaluate the accuracy of the consolidated data. What is the accuracy of sourced data? Are there concerns about the quality of the final data set to be analysed relative to the use? Have the sources been consolidated/transformed in such a way that accuracy is affected? Will time impact the accuracy of the combined data set?	Data Transformation: Data from diverse sources in diverse formats must be put into a form where the data can be analysed. This process can impact the accuracy of the data itself and the insights that might come from big data. ▪ Bring forward the analysis conducted on the accuracy of the data. What is the timeframe for accuracy of data?	
Insights: Understand what insights (outputs) are expected from the analysis. What is the output from the analysis?	The questions to the left link to the sustainable value. Outputs reflect status at a particular point in time. Status can naturally change over time. The actual application of an insight may begin the process of change. Many of the questions to the left are designed to understand both	

How will the insights from the analysis be used?	expected outputs and how durable those outputs might be when applied as insights. (NOTE: A demo can be useful in helping to understand the insights.)	
How long might an insight endure? Do insights degrade over time? Are the insights repeatable and for how long? Can the application of the insights impact behaviour in a manner that could reduce the predictive value of the insights over time?	Insight (output) can be a report. Are uses of the insight internal or external? Insights lose predictiveness over time. When do insights age past usefulness? An insight may be problematic. Application of an insight may be problematic. Actions by a marketer may lead to changed behaviour by consumers. This changed behaviour may impact predictiveness.	
Outcomes: Will the project result in better outcomes than currently available? Which stakeholders have positive outcomes? How significant are the positive outcomes? Can the same or similar outcomes be achieved with fewer risks (e.g., possibly done with less robust data)? What are the downstream risks?	The questions on the left link to the progressive value. Progressive means better outcomes than existing means. Better outcomes for marketers include: ▪ Improved profitability ▪ Enhanced customer relationship ▪ Undamaged brand/reputation Better outcomes for individuals include: ▪ Better product selection ▪ Significant discounts ▪ Appropriately linked to other choices, etc. All processing creates risks. If using more data and robust processing does not yield more than a marginal gain, it should be avoided. Is the gain progressive?	

Can the outcomes be hypothesised, anticipated or achieved?	Evaluating expected (versus actual) outcomes.	
Accountability: Identify the individuals who are responsible for the project. Who has ultimate project accountability? Who is accountable for the various phases of the project? Do the insights contemplated by the project seem inappropriate, creepy, intrusive or rude by the person ultimately accountable for the project?	Project team includes: - Data capture/acquisition - Data preparation/ management - Oversight for restrictions (legal or contractual) - Appropriate application of the analysis/insights	
Stakeholders: Identify all the stakeholders and their concerns. Who are all the possible stakeholders related to both the analysis and the use of the resulting insights? What stakeholder concerns may arise? Are there other factors that should be taken into account?	Possible stakeholders include: - Individuals/data subjects (consumers, customers and prospects) - Organisations (including businesses and non-governmental organisations) - Political entities/government - Society/public-at-large/community - Others Possible stakeholder concerns include: - Data use out of context - Contract, legal, privacy or security obligations - Data sensitivity - Revenue/business needs Other factors include: - Cultural differences - Commonly held societal values - Compatibility with organisational values	

	• Compatibility with social norms regarding the use of sensitive information	

BENEFICIAL		
Benefits: Describe the benefits in detail for each stakeholder identified above that are expected to come from the analysis.	There may be more than one benefit for a stakeholder. Describe them. Obvious benefits can include: • Personalisation • Health • Education • Economic opportunity • Society as a whole • Access to new products or services • Consumer/customer engagement • Brand building • Increased revenue • Convenience • Price • Other (specify)	
Risks: Describe the risks in detail associated with each stakeholder identified above that are anticipated from the analysis. For each risk, identify if it is likely to happen and, if so, what the severity is.	Risks to stakeholders take into account: potential impacts of false positives or negatives. Possible risks to individuals include: • Physical harm • Financial harm • Heath • Reputational • Embarrassment • Shock or surprise • Inappropriate discrimination • Misuse of data	

	Possible organisational risks include: ▪ Negative media attention ▪ Negative regulatory impact ▪ Compliance/reputational ▪ Business continuity	
Mitigations: How is each affected risk tested and mitigated? What are the residual risks after mitigation? Are the mitigated risks sufficiently balanced by the benefits?	The risk/benefit analysis should be documented. Risk mitigation includes outcomes. ▪ Evaluate expected (versus actual) outcomes. ▪ Are there false positives, false negatives and security risks? Geolocation Transformation	
FAIR		
Are there reasonable arguments about why this application might be considered unfair to individuals? If so, how? Are there other issues that could arise from this project? Will the residual risks and benefits balance individual and societal interests? From your perspective as the person accountable for the project, are you confident that the interests of stakeholders are balanced in a fair fashion?	The analysis in this section should be based, for the most part, on the answers in previous sections. Factors not considered before but considered in this section should be noted. Fairness issues include: ▪ Regulatory ▪ Media ▪ Public backlash ▪ Discriminatory affects such as economic opportunity, physical security, physical wellbeing and limiting self-determination.	