ETM

ENGINEERING & TECHNOLOGY MANAGEMENT

Evaluating Big Data Projects Probability of Success: A Hierarchical Decision Model

Course Title: Decision Making

Course Number: ETM 530/630

Instructor: Prof. Tugrul Daim

Term: Winter 2017

Author(s): Abdulhakim Giadedi
Ali Eljayar
Husam Barham
Priyanka Patil
Shreyas Vasanth

Table of Contents

Abstract: **This paper represents a team project conducted as part of ETM 530/630 Decision Making class in Winter-2017. In this project, team 4 investigated a known challenge related to big data; big data projects have high failure percentage, causing firms to lose time, money, and resources in futile efforts to gain advantages from big data insights and analytics.**

In this project, the reasons behind big data projects failure were explored. Leading to the development of an HDM model that can be used by firms to evaluate readiness to implement this kind of projects, and highlight/address probable causes of failure before the project even start. Hence, increasing the chances of implementing a successful project that will lead to a big data system that can deliver value to firm and provide insights and analytics that will significantly help in addressing the problem it is built to help solve. The model was evaluated by experts from the industry, and then tested against a hypothetical case, in which Portland State University readiness to implement a big data project to address a main problem facing the university was conducted. Finally, a discussion about the results of the model, experts' evaluation, and case study were offered.

1. Introduction

Big data is one of the leading technologies in the last few years [1]. Firms use big data to support decision making on the strategic, operational, and product levels, by leveraging on big data insights and analytics [2]. A 2015 survey of the 1000 fortunate firms' CEOs found that: 70% of the CEOs reported that big data is of critical importance to their firms, up from 21% in 2012. And 63% of the firms reported having Big Data in production, up from just 5% in 2012 [3].

However, studies indicate that more than half of big data projects fail. It either never finish or do not generate the expected outcome [3][4] [5]. The reasons behind this high percent of failure were the subject of many studies in the past few years [6][7].

So, the focus of this paper is about finding out what are the main challenges facing big data projects. And how firms can prepare in advance to deal with those challenges.

1.1 Problem Statement

The objective of this project is to develop an HDM decision model that can be used to evaluate firm's readiness to implement big data projects. So, after a firm decides to implement a big data project, they can use the model to evaluate the firm's readiness before starting the project, and determine what are the weak areas, within the firm, that might cause the project to fail, and consequently, address those areas before the project even started.

2. Literature Review

In this section, we will first define big data, and then we will review common challenges known in big data projects.

2.1 What is Big Data?

Organizations collect/generate data from various sources, they collect more than what they know about or can process. So, big data is a holistic information management approach to consume and integrate data, whether the data is structured (e.g. transactional records) or unstructured (e.g. social media and web behaviors) from multiple internal and external sources. Then, identifies relations among them, and creates insights that allow for complex analysis and future predictions, which ultimately will result in higher probability of making the right decision, and hence leapfrog competitors and lead the market [8].

Big data is fast becoming a tool that not only analyzes patterns but also can provide the predictive likelihood of an event and help decision makers to take action. Organizations are exploring how large value data can usefully be deployed to create and capture value for individuals, business, communities, and government [9]. In August 2010, the White House, OMB, and OSTP proclaimed that Big Data is a national challenge and priority along with healthcare and national security. The National Science Foundation, the National Institutes of Health, the U.S. Geological Survey, the Departments of Defense and Energy, and the Defense Advanced Research Projects Agency announced a joint R&D initiative in March 2012 that will invest more than $200 million to develop new big data tools and techniques [10].

2.2 Big Data Challenges

A survey from Infochimps showed that 55 percent of big data projects are never finished. Another research conducted by Gartner showed also that many data analytics projects aren't successful. A project fails when it isn't completed within the set amount of budget, it isn't finished on the established timetable, or it doesn't have the benefits and features that were promised when the project began [11]. The literature review revealed several known challenges that cause big data projects to fail, we divide those issues under five perspectives: Personal, Technical, Political, Economic, and Management. Following is a review of challenges under each perspective.

2.2.1 Personal Perspective:

Data Scientist: A Data Scientist is the person who systematically studies the organization and brings structure to large quantities of formless data to determine significance in its value, and systematic relationships between the variables[12][13]. A recent survey of C-suite executives by KPMG found 99% of respondents thought analysis of big data was important to their strategy next year. In an age where enterprise data is expected to exceed 240 Exabyte per day by 2020, the need for data scientists with the skills to extract valuable insights from this data is more important than ever. A growing demand for people trained in data science has caused the shortage of these people to balloon [14]. A study by McKinsey projects that "by 2018, the US alone may face a 50 percent to 60 percent gap between supply and requisite demand of deep analytic talent." So, employing good data scientist is a big problem who is integral part of a Big Data Project [12].

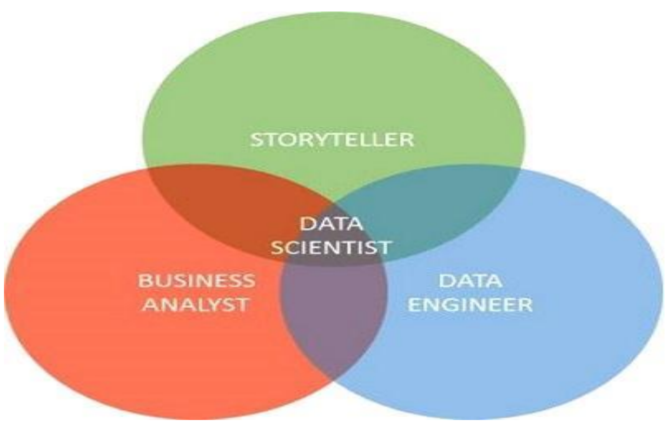


Figure 1. Data Scientist Skillsets [14]

Employees' Skills: The skills of other employees are also important for the success of big data projects. They play a major role in pushing and pulling data into and from the system.
A small miss of information entered to the system will generate results that don't make sense. Also, management needs to have certain level of analytical skills to be able to use and make sense of the analytics generated by the system, failing to do so will make the system useless [15][16].

2.2.2 Technical Perspective:

Data integration Complexities: One factor for big data to offer real value is its ability to aggregate and analyze data from various sources. Overall, data integration and data interoperability influence the organization's performance. The data integration is a complex challenge for the organizations deploying big data architectures due to the heterogeneous nature of data used by it. Therefore, it requires a comprehensive approach to negotiate the challenges in integration and interoperability. Data integration plays a key role in determining the efficiency of a big data project, be it at the level of backend systems integration or integration of processes, administrative tasks, and databases [17].The complexity of data integration and interoperability emphasizes on the levels of data storage, structure and the levels at which the data can be integrated and operated as a single entity. Collecting and maintaining the large data sets is costly, therefore some organizations tend to adapt to cloud methodologies for storing the data and reuse [17]. As the sizes of data set are often very huge, sometimes several Terabytes or more, and their origin from heterogeneous sources, current real-world databases are severely susceptible to inconsistent, incomplete, and noisy data. Therefore, data integration techniques can be applied to remove noise and correct inconsistencies [18].

Data Availability: Availability refers to the resources of the system accessible on demand by an authorized individual [19] [20]. However, there are four important factors that have a significant impact on data availability including:

(1) Volume refers to the amount of all types of data generated from different sources and continues to expand. The benefit of gathering large amounts of data includes the creation of hidden information and patterns through data analysis [21] [20].

(2) Variety refers to the different types of data collected via sensors, smartphones, or social networks. Such data types include video, image, text, audio, and data logs, in either structured or unstructured format [20].

(3) Velocity refers to the speed of data transfer. The contents of data constantly change because of the absorption of complementary data collections, introduction of previously archived data or

legacy collections, and streamed data arriving from multiple sources [22] [20].
(4) Value is the most important aspect of big data; it refers to the process of discovering huge hidden values from large datasets with various types and rapid generation [23] [20].

Technology Solutions Complexities: In big data projects, the goals the project is trying to achieve and the nature and sources of data related to the project dictate what type of tools needed. When the mix includes several tools, the probability of incompatibility becomes higher leading to software bugs and technical issues, not to mention the challenge of finding software engineers that that are experts in all the tools needed [10][22][24].

2.2.3 Political Perspective:

There are certain political challenges that, if not addressed properly, might undermine any project in general. Here we investigate some of those challenges that have high impact when it comes to big data projects.

External Sources of Data: An important aspect that affects big data ability to offer real value is getting data from different internal and external sources [24]. To gain access to external sources of data, including, data available at clients, suppliers, and other entities ends, firms depend on other entities willing to share their data. Firm's management should negotiate with those entities to get access to their data. Making sure to understand their needs and concerns. The best way to get external entities to share their data is by looking for mutual benefits and creating win-win situations [24] [25].

Data Ownership: Big data generates statistics and analysis based on data coming from internal and external sources, and that creates challenges regarding the ownership of the data and the analysis created based on it, and what degree of freedom a firm has in sharing the data. For example, can a firm share analysis, including one supplier data, with another competing supplier? or who can access this information within the firm itself?

So, firms need to balance between sharing the analysis results, in order to achieve the required goals, and avoid conflicts with data sources' owners that could result in losing access to these sources. [26][27].

Data security, privacy, and governance: Even after making beneficial agreements about data ownership and sharing, firms still face challenges related to make sure the data is secure, privacy is considered properly, and the way the data is being handled is not breaking any related regulations. And external entities are more willing to share their data, if they know that their data is well protected. Moreover, any breach in data privacy could result in legal and public image damages with severe consequences [28][29]. Following resource has many known such cases [30].

2.2.4 Economic Perspective:

The Total Cost Concept: Building a Big Data analytical solution is somewhat like building your own house. Surely, building a house according to one's design would be the most customized option. It is a dream comes true. But more important criteria are undermined when dreaming. As with analytics, building by oneself doesn't mean carrying the bricks and laying the tiles on the roof. But one must choose the right professionals and know enough about the project in order to supervise and manage it well. The single, most common thing in common to both cases of building is that one has no idea what the true and final cost will be [31] [32].
The expenses for building true cost of big data analytics solution are divided to three main categories: infrastructure, software and human resources, the latter being the most demanding.

Table 1: Big Data Costs [32]

	1TB PER MONTH	1TB PER QUARTER	3TB PER MONTH
Infrastructure	$1,900	$6,400	$16,000
Software	$1,600	$2,500	$8,000
Human Resources	$13,000	$22,000	$35,000

Total Monthly	**$16,500**	**$10,300**	**$59,000**
Total Annual	**$198,000**	**$370,800**	**$708,000**

The infrastructure of an analytics solution consists of data storage, servers, network and monitoring tools. All costs are proportional to the platform's size.

The major software expense when building analytics is the analytical database. Using 3TB per month as an example, based on leading platform providers, an analytical database is likely to cost more than $134,000. Additional required tools are an ETL (Extraction, Transform and Load) or Hadoop, real-time database and visualization tools – presenting the data graphically so that it can be shared with other people, make changes, adding and shifting data around.

In the above case of 3TB per month, infrastructure and software amount close to $180,000. Buying an end-to end-solution not only includes this cost, but also saves the time and effort invested in evaluating, testing, deploying and integrating through a long process of trial and error [31].

Having said all this, the most significant cost of building a Big Data analytics solution is human resources. The solution is complex, requires real know-how and involves numerous specialists. All need to be engineers who are experienced with Big Data, which is a rather scarce resource nowadays, and an expensive one at that. A partial list of the experts the system requires: ETL developers, cloud infra experts, Java/Python developers, database administrators (DBAs), data analysts, dashboard developers and so forth. All in all, a system for 3TB per month requires about eight data engineers at a cost of roughly $800,000. An acquired solution cuts these costs down substantially.

Equally important is that a bought solution allows staff to keep on doing their job. Building a solution, on the other hand, diverts company personnel to a field of expertise that is not their innate domain. This harms not only the analytics solution, but company revenue as well due to time, money and human resources getting diverted away from a core business. This opportunity cost of a built solution must be taken into account [32].

A common estimate for the list price to acquire a Hadoop cluster, including hardware and open source software, is less than $1,000 US per TB of data stored. This is several times less than the average list price of widely-used, highly-parallel data warehouse platforms. Thus, many people assume that an analytic data solution implemented on Hadoop will always cost less than the same solution on a data warehouse platform. Some might imagine that system acquisition price is the only important factor.

However, getting the business value from data is not only about “system” price. To understand the actual cost of any business solution, one needs to determine the cost of constructing and maintaining the solution as well as the ongoing cost of using the data for its business purpose.

To capture the total cost of a big data solution, we propose a framework for estimating the total cost of data, or TCOD. In addition to system cost, TCOD takes into consideration the cost of using the data over a period of time. This includes the cost of developing and maintaining the business solutions – complex queries, analyses and other analytic applications – built on top of the system. These solutions deliver the real value of data to the organization. Thus, it is critical for an organization to identify and understand the total cost picture before selecting and implementing any technology for a big data analytic solution. The major cost components of TCOD (see Figure 2) are: [31]

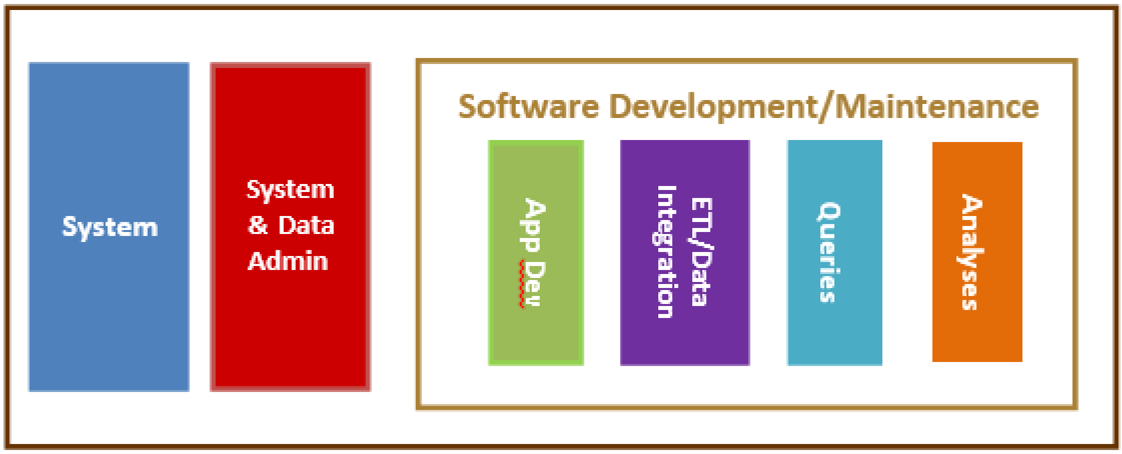


Figure 2. Components of Total Cost of Data (TCOD) [31]

- System cost – The cost to acquire, maintain/support and upgrade the hardware and system software plus the cost of space, power and cooling. In the case of a data warehouse, this also includes the cost of the database management software.
- Cost of system and data administration – The cost of expert staff to administer the system and the data it stores.

- Cost of data integration – The cost of developing or acquiring an ETL (extract, transform and load) or ELT solution to prepare data for analytic use. This is the cost of developing a process to cleanse the source data, reorganize it as necessary and store it in the database in accordance with an integrated database design.
- Cost to develop queries – The cost of developing queries that can be expressed in SQL.
- Cost to develop analyses – The cost of developing procedural programs that perform data analyses too complex to express in SQL.
- Cost to develop analytic applications – The cost of developing substantial application programs that use the data to support repeatable processes such as campaign management or loan approval in a financial institution.

2.2.5 Management Perspective:

Management issues might be the most difficult problem to address with big data. This issue first surfaced a decade ago in the UK eScience initiatives where data was distributed geographically and “owned” and “managed” by multiple entities. There is no perfect big data management solution yet, and this represents a significant gap in the research literature on big data that needs to be filled [10]. It is critical for the managers to know how to use big data to reinforce organizational sustainability and how different operational, strategic, and corporate activities are affected in this process [28]. Many factors that can influence big data project under management perspective. In this report, we focus on three important factors; management support, data strategies, and clarity of scope.

Management Support: Few big data projects that develop new processes retrain their staff and successfully integrate with existing big data operations are still rarely seen as true successes because of inconsistent oversight. Most projects have poor operational control because big data projects are considered high risk. Most project leaders don’t want to set expectations too high or have too much visibility into their projects because they believe the projects are likely to fail. As a result, new projects don’t get the same performance measures required to show company leadership the positive impact of their success. Even if leaders are told that a project is in production, they don’t know why that project is necessary to the business. As with the lack of integration, when it comes time to budget cuts, new projects are the first to go, not because they’re not impactful, but because it’s not clear to that decision makers what impact these projects have [33]. An important study was conducted by MIT in collaboration with IBM to understand how all organization deal with big data projects and apply analytics. This study includes a survey of nearly 3,000 executive managers worldwide, 30 industries and 100 countries. A part of this survey as shown in Figure (1) illustrate that approximately 38% of the respondents said that lack of understanding of how to use analytics to improve the business is the top challenge followed by 35% lack of management bandwidth due to competing priorities [34].

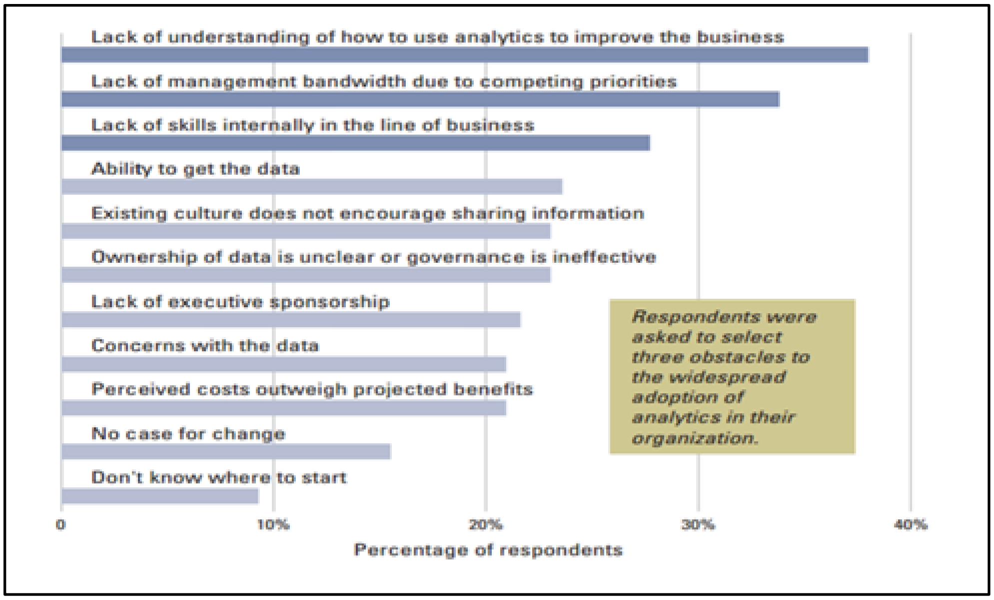


Figure 3. Part of the survey conducted by MIT and IBM [34]

Data strategies: The first step to define a strategy for any Big Data project is to develop a framework to understand big data and select the best strategies and the techniques to deal with big data projects. This framework has two dimensions; the first dimension is the business objective: When developing big data capabilities, companies try to measure or experiment. When measuring data, organizations know exactly what they are looking for and look to see what the values of the measures are. The second dimension is the data type; companies collect data that has structure or scheme for their operations and this is called transactional data or data come from sources other than transactions are typically unstructured. This combination results in four quadrants as shown in Figure 4, each representing a different strategy: performance management, data exploration, social analytics, and decision science [35] [36].

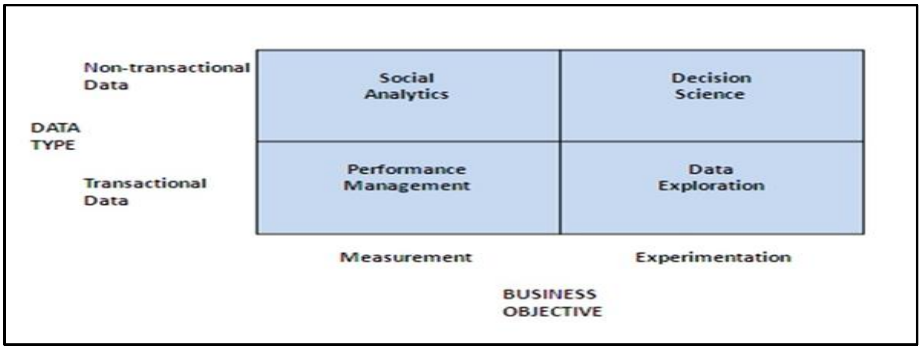


Figure 4. Big Data Framework [36]

Clarity of Scope: One of the primary reasons for big data project failure is a lack of planning. Companies often go into a project completely underestimating just how influential and large it can turn out to be. This underestimation can lead to mismanagement of time and improper allocation of resources. Proper planning requires experienced management and oversight along with project leaders that can clearly communicate to the whole team what is needed and how to accomplish specific tasks [28]. Before embarking on any new big data projects, leaders should consider the operations implications and map out the systems required for their new project. The key to success is to start planning for operations before starting any new big data project. Make sure to have the supporting processes, skills, integrations and oversight to get their projects into production [33].

3. Methodology

3.1 The Hierarchical Decision Model (HDM)

The methodology that has been used in this research is the Hierarchical Decision Model (HDM) that was introduced by Cleland and Kocaoglu in 1987 [37]. This Model has abbreviation (MOGSA), which is invented by Dr. Kocaoglu. This abbreviation came from the first letter of the following terms Mission, Objectives, Goals, Strategies and Actions. The approach presented in this paper is based on pairwise comparisons among alternatives by a respondent. The alternatives are presented, two at a time, for a measure of relative weights with respect to each other. The respondent divides 100 points between the pair to reflect his judgment of each element's relative importance in comparison with the other element of the pair. In HDM, the subjective judgments expressed in pairwise comparisons are converted to relative weights in ratio scale. This is done by a series of mathematical operations on three matrices. The methodology can be used for quantifying the judgment of a single decision maker, or multiple decision makers. When multiple decision makers are involved, the HDM approach is an effective way to form consensus among decision makers where the members of the group have different goals. HDM links the decision elements at multiple levels of organizational entities, in which decisions at the operational level are made in support of higher level goals and objectives, and when the objectives are met, the final results of the operational decisions are transformed into benefits for the organization. This is a systematic process, but it is difficult to quantify the direct relationships between the benefits at the top of decision hierarchy and the operational decisions at the bottom without dividing the space between the top and bottom of decision hierarchy into intermediate levels. That is what the HDM does [37]. However, hierarchical decision model has been applied to a wide range of problems with success. A representative list of select applications to date is given below [38]:

1. Relative priorities for police calls.
2. Allocation of patrol resources to all precincts.
3. High school selection.
4. Evaluation of R&D programs.
5. Personnel allocation.
6. Medical Care evaluation.
7. Higher education scenarios.
8. Transportation planning.
9. Energy policies.

3.2 Case Study

According to Eisenhardt [39] "case study is a research strategy which focuses on understanding the dynamics present within single settings". There are many applications for case study in academia including testing theories, describing phenomena, or even building theories. [39][40]. In this paper, a case study will be conducted to test the model validity in real situations.

4. The HDM Model

Based on literature review and consulting with experts, an HDM model was constructed (Figure 5). The model was created using ETM HDM© software tool.

In this section an explanation for the model is offered. As well as, details about the expert panel and how the data was collected.

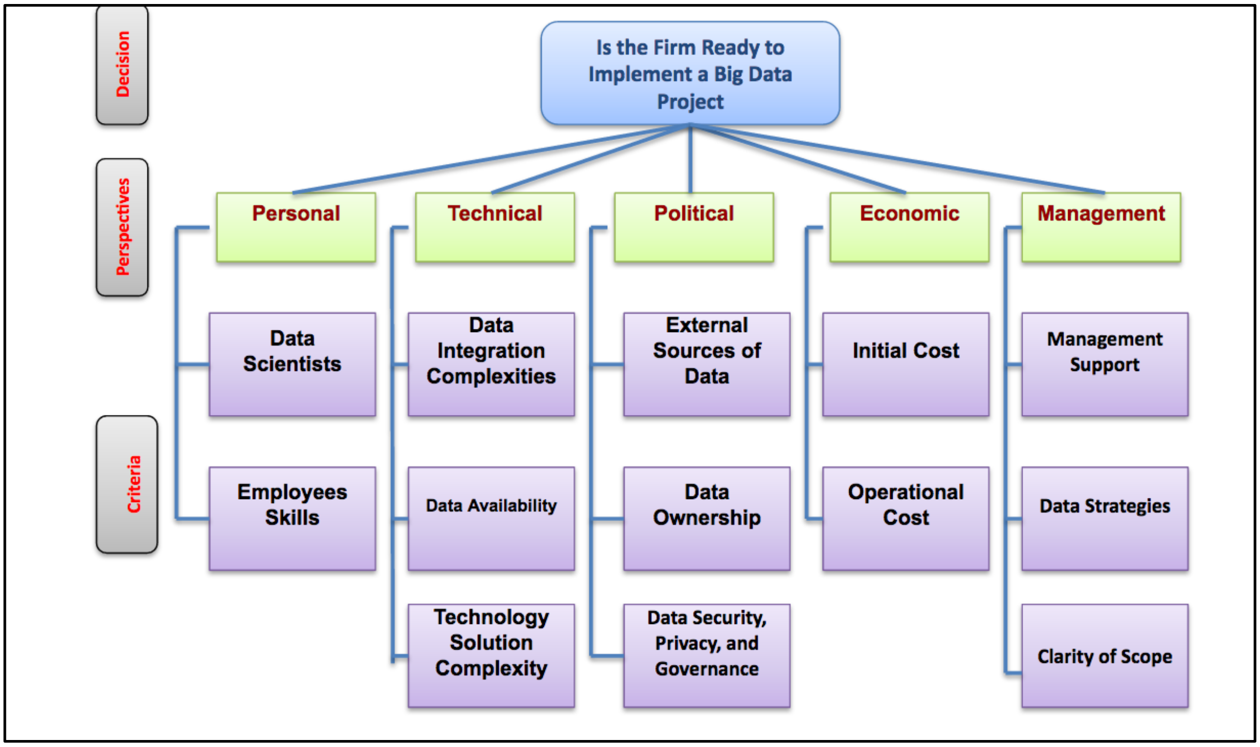


Figure 5. HDM model to assess firm's readiness to implement big data project

4.1 Criteria Selection and Model Building

Personal Perspective:

1. **Data scientists:** Data scientists are the minds that can realize what type, size, and frequency of data need to be capture; they device the predictive analysis algorithms that maximize data value, with realization of organizational goals, as well as, internal and external factors around the firm.
   This criterion evaluates firm's data scientists and their level of experience
2. **Employees' Skills:** This criterion evaluates employees technical skills and their ability to operate complex software systems

Technical Perspective:

3. **Data integration Complexities:** One factor for big data to offer real value, is its ability to aggregate and analyze data from various sources.
   This criterion evaluates the challenges in integrating the data from the sources that will be

used in the project

4. **Data Availability:** This criterion evaluates various issues related to the data including: volume, velocity, Quality, Collecting the right kind of data
5. **Technology Solutions Complexities:** In big data projects, several software tools are used together to achieve the project goals
   This criterion evaluates how complex is the mix of tools to be used in the project

Political Perspective:

6. **External Sources of Data:** This criterion evaluates firm's accessibility to external sources of data needed for the project, such data is available at clients, suppliers, and other stakeholders ends, are they willing to share it or not?
7. **Data Ownership:** This criterion evaluates how much freedom the firm has in disseminating analysis generated by big data system based on data coming from external resources
8. **Data security, privacy, and governance:** This criterion evaluates the level of security and privacy the system must have versus the value it can generate

Economic Perspective:

9. **Initial Cost:** This criterion evaluates whether the initial cost is justified in reference of the expected value of the system.
10. **Operational Cost:** This criterion evaluates operation cost of running the big data system in reference of the expected value of the system.

Management Perspective:

11. **Management Support** Leadership and support for any project plays significant role in the success chances of implementing the project. This even more true in case of big data projects that requires a lot of changes within the organization.
    This criterion evaluates management level of support for the big data project and the goals it is expected to achieve.
12. **Data strategies:** This criterion evaluates whether the firm has any strategies, when it comes to handling data
13. **Clarity of Scope:** Defining scope is perhaps the most important part of the upfront process of defining a project. In fact, if top management doesn't know for sure what they are delivering and what the boundaries of the project are, they have no chance for success.

This criterion evaluates how well-defined the project is.

4.2 Experts Panel

To build and evaluate the model, two experts panel were established, an expert panel, consists of experts in the field of big data, who were asked to evaluate and weight the model itself. And a second experts panel consists of this paper authors, who built the desirability curves and evaluated the case study.

4.2.1 Model Experts Panel

The first panel experts were asked to evaluate the model. The panel consisted of seven experts in the field of big data, with more than 10 years of expertise in the following areas: project management, big data software development, and academic big data research. Table 2 offer more details about the experts.

Table 2: Experts Panel

Expert			**Big Data Area of Expertise**		
	Name	**Current Position**	**Project Management**	**Software development**	**Academic Research**
Expert 1	Aisha Al-Qasab	Head of E-Services Dept.			
Expert 2	Amjad Sandouqa	Senior Programmer			
Expert 3	Baker Shawkat	IT Project Manager			
Expert 4	Dr. Hani Omar	Researcher			Holds a Ph.D in Big Data
Expert 5	Moayyad AlFaris	Solutions			

		Architect			
Expert 6	Mohammad AbdulSalam	IT Project Manager			
Expert 7	Wasim AlAyoubi	Middleware Architect			

The first panel experts are previous colleagues of one of the authors; he worked with each of them on various projects in the last decade.

The experts were contacted via Whatsapp texting app and e-mail to explain the project to them and what is required of them. Then each of them got an email with the details of the model and how to evaluate it. Finally, the experts used ETM HDM tool to evaluate the model. Appendix A contains a sample of the letter sent to experts.

Experts offered comments on the model and evaluated the criteria, the results section has more detailed about their evaluation and analysis of the results.

4.2.2 Desirability Curves and Case Study Experts Panel

The second panel consisted of this project's team. One team member, Husam Barham, has sound background in big data (with more than 10 years of experience both technical and project management areas of big data), the rest of the team gained a lot of knowledge about big data from doing the literature review.

The second panel, constructed the desirability curves matrix, assigned values for each criterion matrix, and evaluated the case study against those desirability curves, by conducting a series of brainstorm sessions for that purpose.

5. Case Study - PSU

To test the model, a hypothetical case was introduced. In this case, Portland State University (PSU) decided to tackle an admission problem facing Oregon universities and colleges using big data.

The problem is that the number of students finishing their post high school degree, in Oregon, in the first five years, is less than 50%. In fact, a staggering number of 550,000 Oregonians (almost quarter of the population) started some post high school degree but never finished [41][42]. Oregon State dedicated a $60 million in 2015-2017 budget for the public universities, including PSU, to address this issue [41].

So, to test the HDM model, the authors suggested a hypothetical case, where PSU decided to use big data to address the college dropout phenomena, by getting advantage of the state's funding. Hence, the readiness assessment HDM model, developed by this paper project, will be used to evaluate PSU readiness to implement this big data project.

6. Data Analysis and Results

6.1 HDM Model Results:

After analyzing the model evaluation results, that have been collected from the first panel, which include seven experts who have solid knowledge and long experience of managing Big Data, it can be concluded by the final calculation results shown in Table 3, and Table 4 the following:

Table 3 illustrates the relative weight of perspectives towards objective. Higher weight represents more important issue in satisfying the decision level. From the table we can notice that management perspective is the most important factor with weight of 0.26 followed by both technical and economic based on the experts' opinions.

Table 3: Relative value of each Perspective

Perspectives Level	Personal	Technical	Political	Economic	Management	Inconsistency
Expert 1	0.22	0.18	0.12	0.16	0.31	0.04
Expert 2	0.22	0.09	0.17	0.12	0.4	0.09
Expert 3	0.17	0.27	0.23	0.14	0.19	0.01
Expert 4	0.25	0.25	0.19	0.17	0.14	0.01
Expert 5	0.2	0.39	0.14	0.11	0.15	0.07
Expert 6	0.14	0.07	0.32	0.18	0.29	0.03

Expert 7	0.05	0.1	0.03	0.45	0.37	0.01
Mean	0.18	0.19	0.17	0.19	0.26	
Minimum	0.05	0.07	0.03	0.11	0.14	
Maximum	0.25	0.39	0.32	0.45	0.4	
Std. Deviation	0.07	0.12	0.09	0.12	0.11	

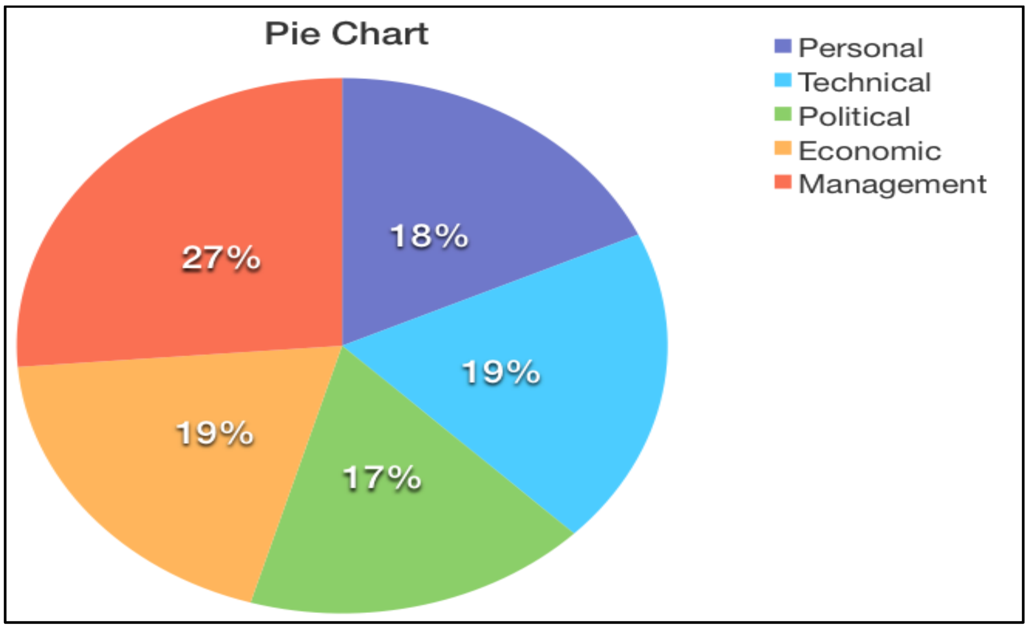


Figure 6. Amount of percentage for each perspective towards objective

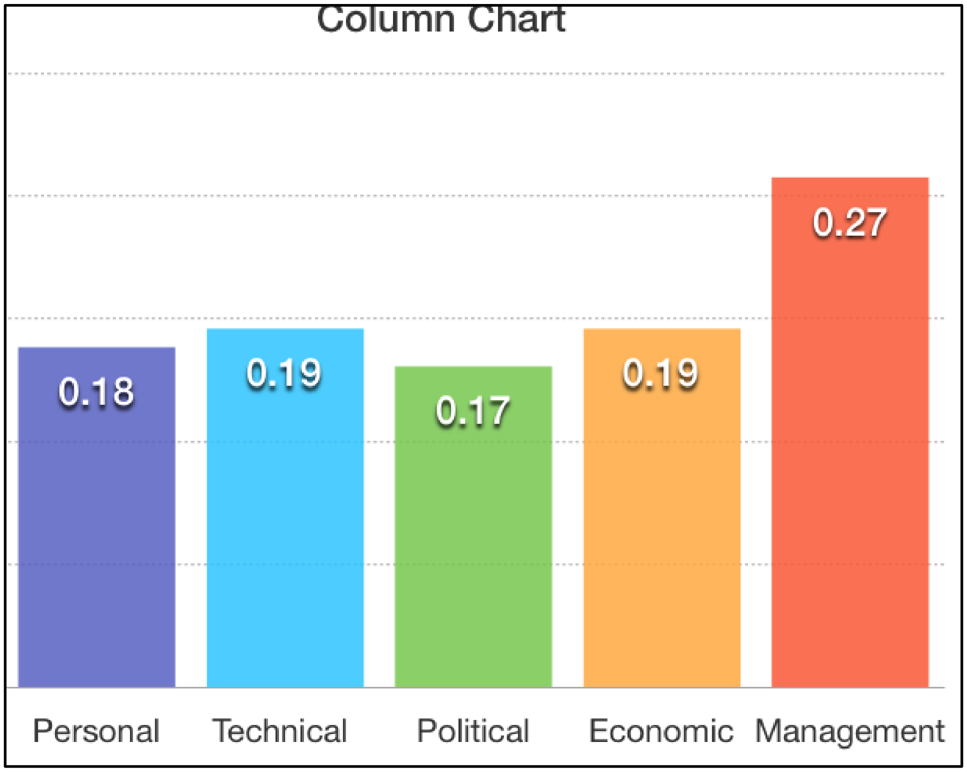


Figure 7. Relative weight of perspectives towards objective

Table 4 illustrates s the contribution of each criterion to satisfy the decision through perspectives. Multiplying the contribution of each criterion by the relative weight of corresponding perspective gives the relative value of each criterion towards the objective. However, from the table we can conclude that Management Support, Initial Cost and Data scientists are the most important criterias with the same value of contribution in order to satisfy the objective.

Table 4: Contribution of each criteria to the objective through Perspectives (Relative value of each criteria)

Is the Firm Ready to Implement a Big Data Project	Data scientists	Employees' Skills	Data integration Complexities	Data Availability	Technology Solutions Complexities	External Sources of Data	Data Ownership	Data security, privacy, and governance	Initial Cost	Operational Cost	Management Support	Data strategies	Clarity of Scope	Inconsistency
Expert 1	0.13	0.09	0.05	0.07	0.06	0.03	0.05	0.04	0.05	0.11	0.1	0.1	0.1	0.01
Expert 2	0.11	0.11	0.02	0.06	0.02	0.04	0.08	0.06	0.07	0.05	0.09	0.05	0.26	0.02

Expert 3	0.1	0.07	0.09	0.07	0.11	0.1	0.06	0.07	0.06	0.08	0.06	0.05	0.08	0
Expert 4	0.19	0.06	0.09	0.11	0.05	0.04	0.07	0.07	0.09	0.09	0.06	0.04	0.04	0.02
Expert 5	0.03	0.17	0.21	0.09	0.09	0.02	0.05	0.07	0.08	0.03	0.04	0.07	0.04	0.02
Expert 6	0.1	0.03	0.01	0.04	0.01	0.11	0.11	0.11	0.12	0.06	0.09	0.14	0.06	0
Expert 7	0.01	0.04	0.02	0.06	0.02	0.02	0	0.01	0.25	0.2	0.28	0.03	0.05	0.01
Mean	0.1	0.08	0.07	0.07	0.05	0.05	0.06	0.06	0.1	0.09	0.1	0.07	0.09	
Minimum	0.01	0.03	0.01	0.04	0.01	0.02	0	0.01	0.05	0.03	0.04	0.03	0.04	
Maximum	0.19	0.17	0.21	0.11	0.11	0.11	0.11	0.11	0.25	0.2	0.28	0.14	0.26	
Std. Deviation	0.06	0.04	0.06	0.02	0.04	0.03	0.03	0.03	0.06	0.05	0.07	0.04	0.07	
Disagreement														0.047

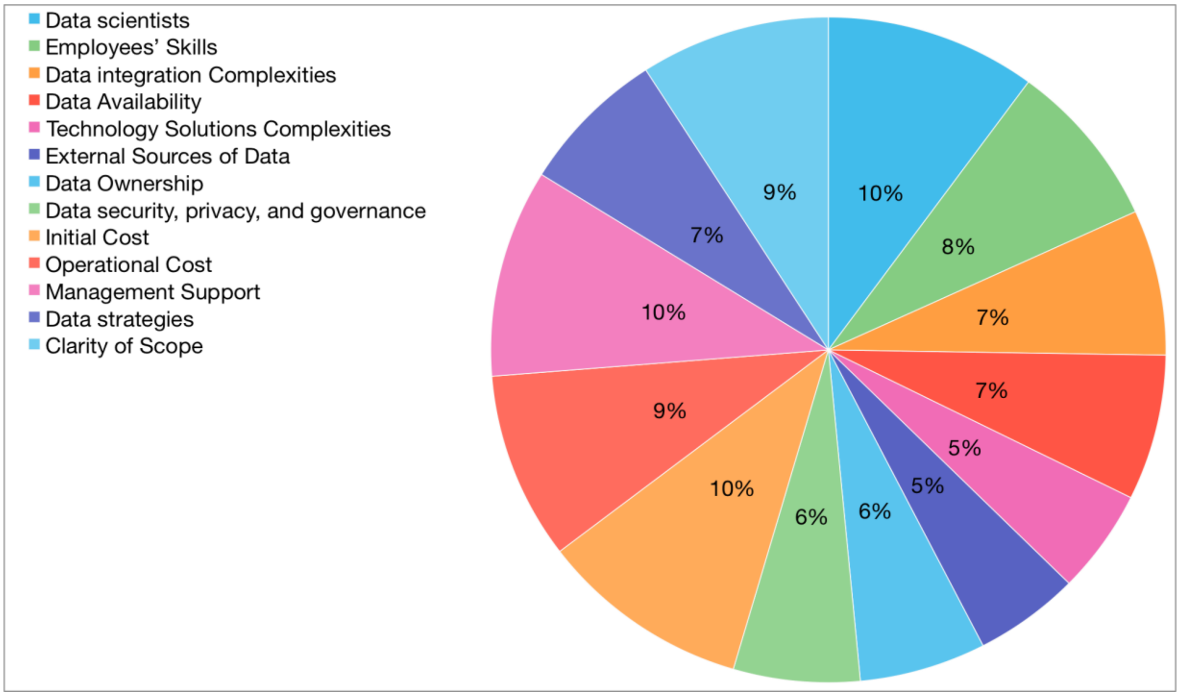


Figure 8. Amount of percentage for each criterion towards objective

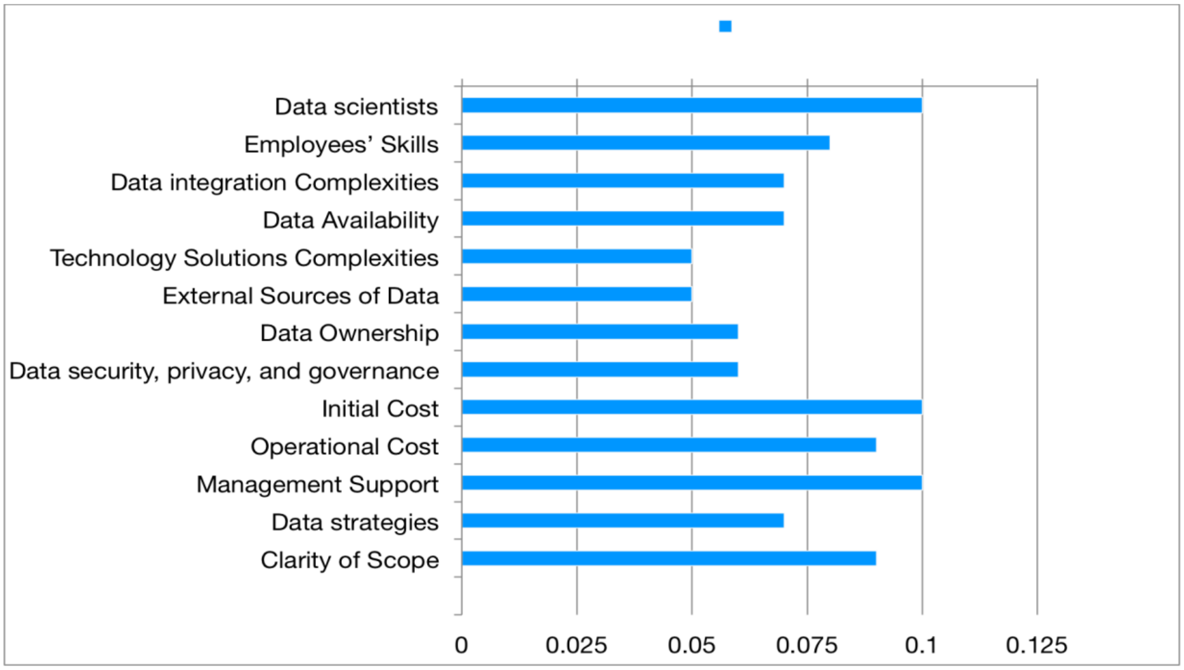


Figure 9. Relative weight of each criterion towards objective

The inconsistency in individual expert's judgmental value is within limit. For each expert the inconsistency is < 0.1. A value less than 0.1 is acceptable for inconsistency. Disagreement is defined as the way to identify commonality among experts in pairwise comparing. A value near zero indicates that the experts were close to consensus. Lower values of both disagreement and inconsistency indicate reliable assessment. However, it is noteworthy that the data obtained from our panel of experts did not show a high disagreement value. The disagreement value shown, just 0.047, gives us a good indication that the experts' opinions about the decision were very close. Moreover, the model illustrates a small amount of inconsistency that each expert has and two experts have almost zero inconsistency. The disagreement and the inconsistency results are highly support our model and illustrate its reliability.

It worth mentioning here that two of the five perspectives have only two sub criteria. Hence, there would never be inconsistency under those perspectives, which might 'dilute' the accuracy of measuring inconsistency across all criteria, but a quick-and-dirty check of the other perspectives separately showed that the inconsistency is still acceptable.

6.2 Desirability Curves Development

The desirability curves were built by the second panel, which includes the authors of this paper. The desirability curves possess was completed through brainstorm sessions to derive the results. And the graphical method was used to develop the curves, by ranking a score of 0-100 to each criterion, based on the desirability of a success attribute. Table 5 shows the desirability curves matrix.

Table 5: Desirability Curves Matrix

Perspective	Criterion	Unit of Measurements
Personal Perspective	**Data scientists**	What data science qualifications the firm has?
	Employees Skills	What are the technical capabilities for the firm's employees on average?
Technical Perspective	**Data integration Complexities**	How integrable are the data sources?
	Data Availability	Availability of data needed in term of volume, velocity, and quality
	Technology Solutions Complexities	The mix of tools needed to achieve the project goals
Political Perspective	**External Sources of Data**	The willing of external entities to share data
	Data Ownership	Do we have full ownership and control over the data
	Data security, privacy, and governance	The Level of data security, privacy and governance needed
Economic Perspective	**Initial Cost**	The level of cost in compare with the risk/impact of project's failure
	Operational Cost	The level of cost in compare with the risk/impact of project's failure
Management Perspective	**Management Support**	The degree of top management support

	Data strategies	Is there a data strategy in the client site
	Clarity of Scope	Does client has clear objective

Appendix B contains a list of all the desirability curves.

Here are two examples showing the desirability curves of Data Scientists and Technology Solutions Complexities criteria.

Example 1:

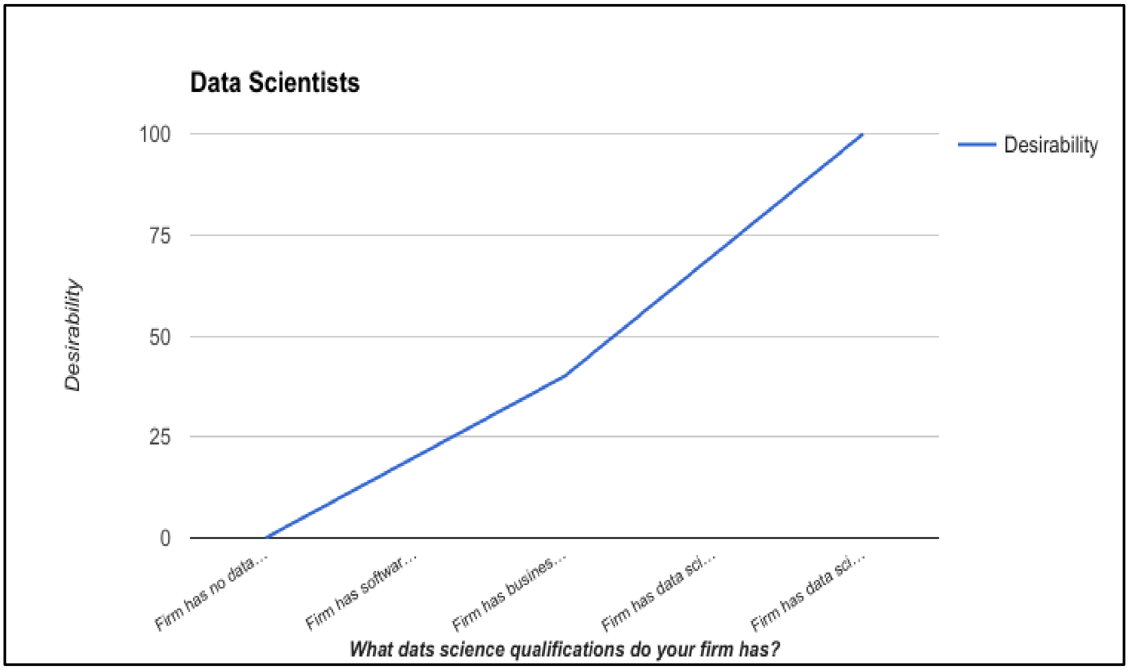


Figure 10. Desirability Curves Example 1 - Data Scientists

Table 5: Desirability Curves Example 1 - Data Scientists

Data Scientists	
Description	**Desirability**
Firm has no data scientists	0
Firm has software engineers with statistics skills	20
Firm has business analysts with statistics skills and some IT background	40
Firm has data scientists who are not strongly relates to the project goals	70
Firm has data scientists who are strongly related to the project goals	100

Example 2:

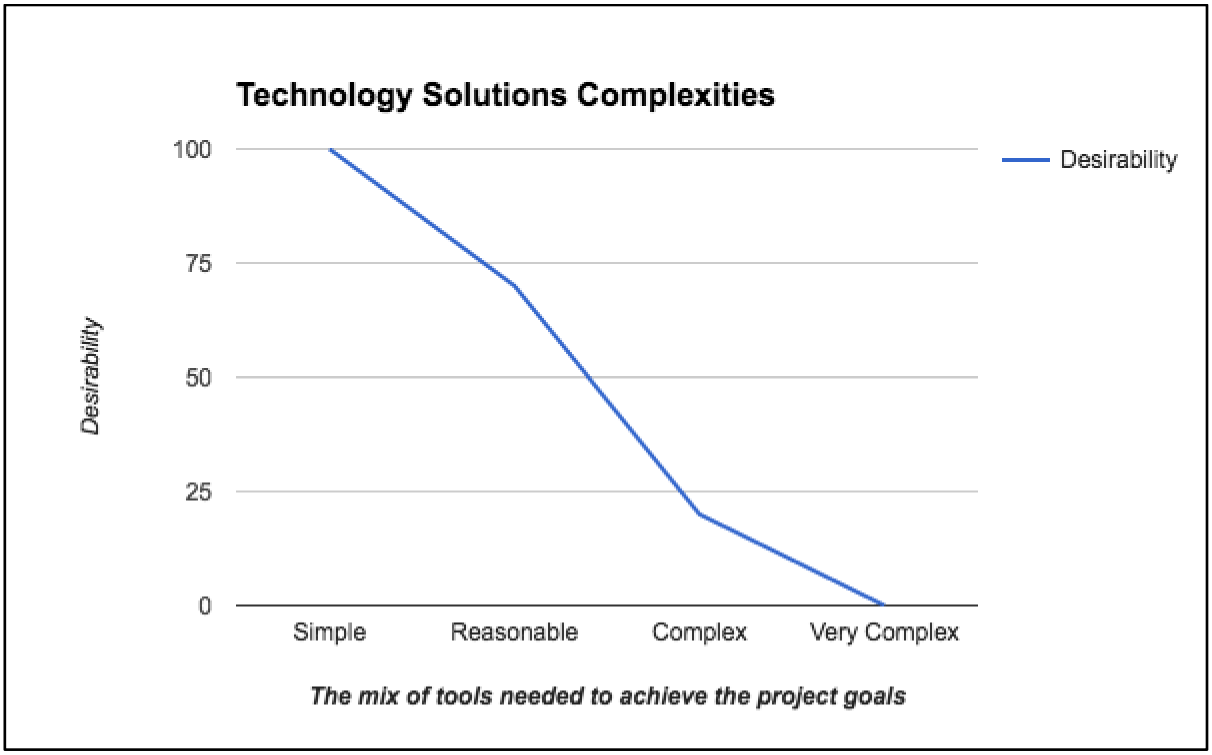


Figure 11. Desirability Curves Example 1 - Technology Solutions Complexity

Table 6: Desirability Curves Example 1 - Technology Solutions Complexity

Technology Solutions Complexities	
Description	**Desirability**
Simple	100
Reasonable	70
Complex	20
Very Complex	0

6.3 Case Study Results:

Table 12 shows the results of evaluating PSU readiness to conduct a big data project, and Figure 7 shows those results on the HDM model itself.

PSU total score is 55.95 (on a scale of 100). And since the HDM model objective is the assessment of readiness, following is a look at the criteria where PSU scored low and what were the reasons:

Data scientists: PSU has many renowned data scientists. However, they are professors who pursue academic interests in several departments, e.g. computer science, sociology, urban studies. And it would be a challenge to gather them to work on this project, as it is not part of their current research interests.

Technology Solution Complexities: To address the college drop out problem, there is a need to get data related to schools, job market, socio-economic status that are available in different formats and will need a complex mix of big data tools to collect, cleans, aggregate, and build analysis on top of these data.

External sources of data: As mentioned in the previous point, there is a need for data from several outside sources. And in many cases, it would be a challenging task to convince the owners of those external data sources to share.

Operation Cost: While the initial cost of the project will be covered from the State's budget [41], as explained in the case study section, the long-term costs of running and maintaining the big data system might not be available causing the system to be shut down after 2017.

Management Support: PSU president is retiring in 2017 [43], and the new president will spend most of the year trying to understand PSU challenges and needs. Hence, he/she might not have the time and focus needed to support this system.

Table 12: Case Study evaluation by the second experts panel

PSU Case Study Evaluation				
Perspective	**Criterion**	**Weight**	**Score**	**Results**
Personal Perspective	**Data scientists**	**0.1**	**20**	**2**
	Employees Skills	**0.08**	**100**	**8**
Technical Perspective	**Data integration Complexities**	**0.07**	**65**	**4.55**
	Data Availability	**0.07**	**60**	**4.2**

	Technology Solutions Complexities	**0.05**	**20**	**1**
Political Perspective	**External Sources of Data**	**0.05**	**20**	**1**
	Data Ownership	**0.06**	**100**	**6**
	Data security, privacy, and governance	**0.06**	**40**	**2.4**
Economic Perspective	**Initial Cost**	**0.1**	**100**	**10**
	Operational Cost	**0.09**	**10**	**0.9**
Management Perspective	**Management Support**	**0.1**	**20**	**2**
	Data strategies	**0.07**	**70**	**4.9**
	Clarity of Scope	**0.09**	**100**	**9**
		Final Result		**55.95**

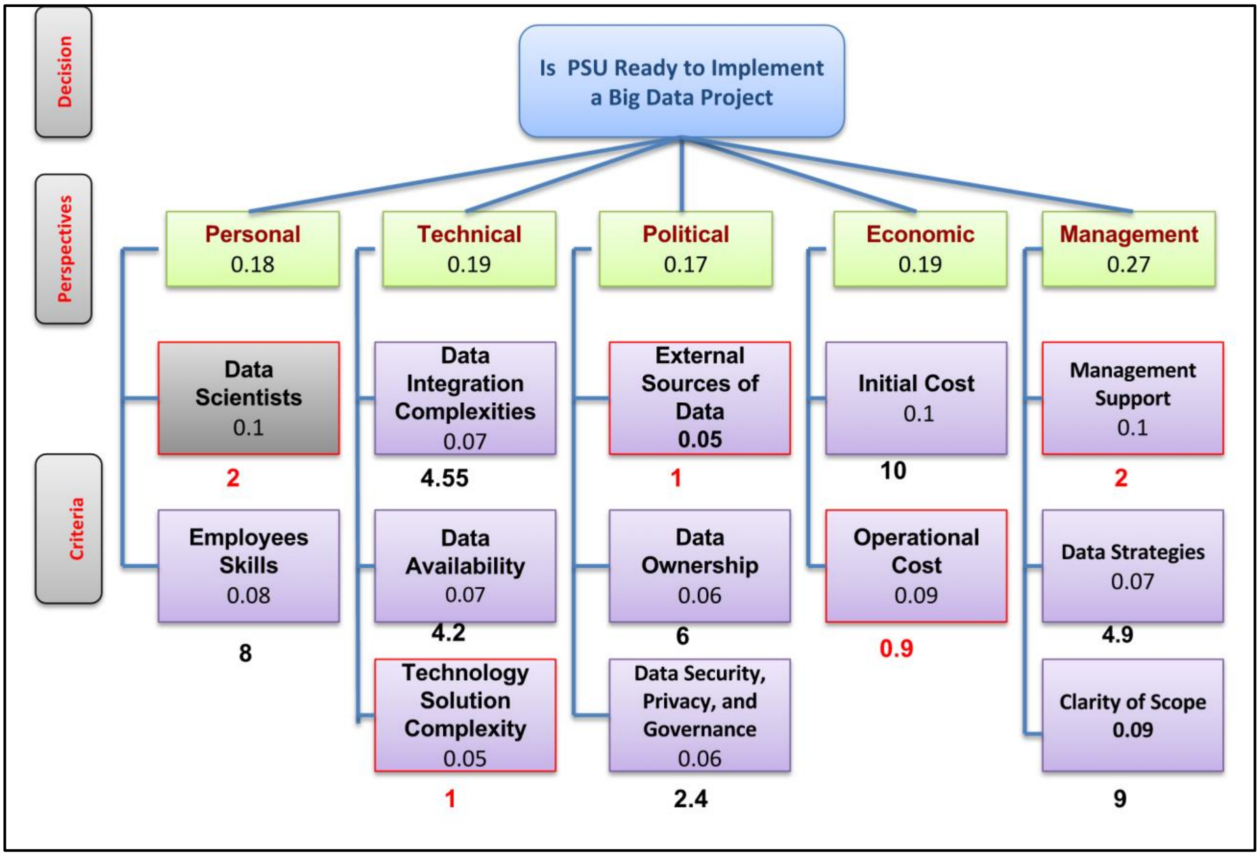


Figure 7. PSU readiness assessment HDM model

6.4 Discussion:

6.3.1 The Model:

The results of evaluating the model by the experts indicated that the most important criteria affecting the success of big data projects are management support, data scientists, initial cost, and clarity of scope, following is a discussion of those criteria based on literature review and feedback from experts:

Management support is important for any type of projects, but it is even more important for big data projects; as the big data system to be build in the project, is usually seen by employees as an extra task to work on, and hence, they resist to use it, leading to the failure of the project. So, management support has a crucial domino effect role to address this challenge. for example, when a CEO is supportive to the project, she will ask direct subordinates (like CFO, CMO, COO) to make sure to use big data analytics in any report presented to her. Those managers in turn will ask their subordinates (middle management) to use big data as part of any report they made to them, leading to wide use and acceptance of big data.

Moreover, big data projects result in major changes within the organization, and management support is important to address employees tendency to resist change.

Data scientists are another very important factor for the success of big data projects, as indicated by the experts. Since they are ‘the brain’ that identify and create the analytics needed to address the firm’s problems, and how to use it to address those problems.

Initial cost is also an important factor affecting big data. Experts indicated that such projects are usually expensive and if the outcome can’t justify the cost, management support will be lost and as explained earlier that will lead most likely to the project failure.

Finally, experts sought lack of clarity of scope as an important reason for big data projects failure. If the problem that big data is supposed to address is not clear, then big data most likely will end up offering useless complex analytics. Making a waste of all the related efforts, resources, cost, and time needed to build it.

6.3.1 The Case Study:

The results indicated several areas where PSU have serious weaknesses. Notably among these are the ‘data scientists’, ‘management support’, and ‘operational cost’ as those criteria have high weight and PSU scored badly in each of them. So, before PSU can start a big data project to address the college dropout problem, they need to make sure those problems are taking care of.

Data scientist could be the easiest one, as the experts are available. PSU has academic professors who are skilled data scientists. So, PSU need to reach out for them and create incentives for them to participate in this project.

Management support could be addressed in two ways; first, finding a top management person, reporting directly to the president, and have proper authority and interest in the project to be the project hero. Alternatively, the project could be postponed until the new president take office, and show interest in the project.

To handle operational cost, it would be helpful to conduct financial analysis and find out if PSU can spare funding for the ongoing operations of the big data system. Or alternatively, working with the state to extend the 2015-2017 budget funding. By communicating to the state officials how the big data system is helping in addressing the college dropout problem, and why it could keep adding value on the long term.

7. Conclusions

This paper addressed the problem of big data projects tendency to fail. By first, identifying challenges (main reasons for failure) for such projects. Then, developing an HDM model that can be used to assess firms' readiness against those challenges. The model was evaluated by experts in the industry. And then put to test, by using it to evaluate PSU readiness to conduct a big data project to address one of PSU's main challenges, students dropout of college.

The results and experts feedback show that the HDM model can offer great value to firms, by allowing firms to prepare properly before conducting big data projects, and hence reduce the risk of project failure.

Following are the recommendations based on this paper's findings:

- Certain criteria (like management support, data scientists) are very important for the success of big data project and should addressed properly
- Internal culture toward big data is vital for the success of such project
- The key to success is to start planning for operations before starting any new big data project, and making sure that the needed supporting processes, skills, integrations, and oversight are in place before starting the project
- Assessing and evaluating the current situation before starting any type of projects is a significant factor for the success of that project
- HDM model can be used as a preventive tool by using it to assess the readiness of an organization before conducting any type of projects.

8. Limitations and Future Research

Following points are some limitations to this paper that can be addressed in future work:

- The experts offered feedback on the model, including suggesting more criteria to be added, time limitation prevented the inclusion of those valuable suggestions to the model, So that should be addressed in any future work.
  Following are some criteria suggested by the experts:
  - Firm's IT Infrastructure Capabilities
  - Cross-Borders issues

- Using industry experts to evaluate the desirability curves, and case study
- Using multiple expert panels
- Considered real case studies to better evaluating the reliability of the model.

9. Acknowledgement

The authors are grateful for the time and efforts the experts made in evaluating the HDM model.

References

[1] D. Cearley, "Gartner Inc.: Top 10 Technology Trends for 2016," *Forbes*, 15-Jan-2016.
[2] International Institute for Analytics, "Advanced Analytics & Big Data Adoption Report 2016 - International Institute for Analytics," 2014.
[3] R. Bean, "Just Using Big Data Isn't Enough Anymore," *Harvard Business Review*, 09-Feb-2016. [Online]. Available: https://hbr.org/2016/02/just-using-big-data-isnt-enough-anymore. [Accessed: 21-Apr-2016].
[4] S. K. White, "Study reveals that most companies are failing at big data," *CIO*, 10-Nov-2015. [Online]. Available: http://www.cio.com/article/3003538/big-data/study-reveals-that-most-companies-are-failing-at-big-data.html. [Accessed: 21-Apr-2016].
[5] B. Marr, "Where Big Data Projects Fail," *Forbes*, 17-Mar-2015.
[6] InfoChimps, "CIOs & Big Data | Visual.ly," 04-Jan-2013. [Online]. Available: /cios-big-data. [Accessed: 19-Mar-2017].
[7] A. McAfee and E. Brynjolfsson, "Big Data: The Management Revolution," *Harvard Business Review*, no. October 2012 Issue, 01-Oct-2012.
[8] H. Barham, "Achieving Competitive Advantage Through Big Data: A Literature Review," 2017.
[9] G. George, M. R. Haas, and A. Pentland, "From The Editors: Big Data and Management," *Acad. Manage. J.*, vol. 57, no. 2, pp. 321–326, Apr. 2014.
[10] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big Data: Issues and Challenges Moving Forward," 2013, pp. 995–1004.
[11] J. Buckley, "Why the Majority of Big Data Projects Fail," *Qubole*, 18-Jun-2015. .
[12] G. Strawn, "Data Scientist," *IT Prof.*, vol. 18, no. 3, pp. 55–57, May 2016.
[13] M. Theuwissen, "The different data science roles in the industry," Nov-2015. [Online]. Available: http://www.kdnuggets.com/2015/11/different-data-science-roles-industry.html. [Accessed: 19-Mar-2017].
[14] E. Jones, "The Role of a Data Scientist in 2016," *InfoQ*, 27-Mar-2016. [Online]. Available: https://www.infoq.com/articles/role-of-a-data-scientist-in-2016. [Accessed: 19-Mar-2017].
[15] N. T. Tariq RS, "Big Data Challenges," *Comput. Eng. Inf. Technol.*, vol. 4, no. 3, 2015.
[16] J. Manyika *et al.*, "Big data: The next frontier for innovation, competition, and productivity | McKinsey & Company." [Online]. Available: http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation. [Accessed: 19-Mar-2017].
[17] A. Kadadi, R. Agrawal, C. Nyamful, and R. Atiq, "Challenges of data integration and interoperability in big data," 2014, pp. 38–40.
[18] C. L. Philip Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," *Inf. Sci.*, vol. 275, pp. 314–347, Aug. 2014.
[19] D. Zissis and D. Lekkas, "Addressing cloud computing security issues," *Future Gener. Comput. Syst.*, vol. 28,

no. 3, pp. 583–592, Mar. 2012.
[20] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. Ullah Khan, "The rise of 'big data' on cloud computing: Review and open research issues," *Inf. Syst.*, vol. 47, pp. 98–115, Jan. 2015.
[21] J. K. Laurila *et al.*, "The Mobile Data Challenge: Big Data for Mobile Computing Research," presented at the Pervasive Computing, 2012.
[22] J. J. Berman, *Principles of big data: preparing, sharing, and analyzing complex information*. Amsterdam: Elsevier, Morgan Kaufmann, 2013.
[23] M. Chen, S. Mao, and Y. Liu, "Big Data: A Survey," *Mob. Netw. Appl.*, vol. 19, no. 2, pp. 171–209, Apr. 2014.
[24] S. Jernigan, S. Ransbotham, and D. Kiron, "Data Sharing and Analytics Drive Success With IoT," *MIT Sloan Management Review*, 08-Sep-2016.
[25] B. Mathieu, "How to use Big Data Analytics as a Competitive Advantage," *GoodData*, 16-Dec-2015. [Online]. Available: http://www.gooddata.com/blog/how-use-big-data-analytics-competitive-advantage. [Accessed: 01-Jun-2016].
[26] J. Harris, "Who owns big data?," *The Data Roundtable*, 26-Mar-2015. .
[27]M. Andrejevic, "Big Data, Big Questions| The Big Data Divide," *Int. J. Commun.*, vol. 8, no. 0, p. 17, Jun. 2014.
[28]D. Etzion and J. A. Aragon-Correa, "Big Data, Management, and Sustainability: Strategic Opportunities Ahead," *Organ. Environ.*, vol. 29, no. 2, pp. 147–155, Jun. 2016.
[29] J. Klein, "When Big Data goes bad - Fortune," 05-Nov-2013. [Online]. Available: http://fortune.com/2013/11/05/when-big-data-goes-bad/. [Accessed: 02-Jun-2016].
[30]T. Reddy, "7 Big Data Blunders You're Thankful Your Company Didn't Make," *Umbel*, 22-Oct-2014. [Online]. Available: https://www.umbel.com/blog/big-data/7-big-data-blunders/. [Accessed: 19-Mar-2017].
[31] R. Winter, R. Gilbert, and J. Davis, "SPECIAL REPORT. Big Data. What Does It Really Cost?," 2013. [Online]. Available: http://docplayer.net/1374678-Special-report-big-data-what-does-it-really-cost-r-t-s-m-a-n-a-g-e-m-e-n-t-expe-d-a-t-a-s-c-a-l-e-l-a-r-g-e-t-h-e.html. [Accessed: 19-Mar-2017].
[32] CoolaData, "The True Cost of Building a Big Data Solution," *CoolaData*, 31-Dec-2014. .
[33] O. Trajman, "Most Big Data Projects Fail: 4 Ways to Make Yours Succeed!," *Data Informed*, 05-Jan-2017. [Online]. Available: http://data-informed.com/most-big-data-projects-fail-4-ways-to-make-yours-succeed/. [Accessed: 19-Mar-2017].
[34] S. LaValle, M. S. Hopkins, E. Lesser, R. Shockley, and N. Kruschwitz, "Analytics: The New Path to Value," *MIT Sloan Management Review*, 24-Oct-2010.
[35] S. Parise, B. Iyer, and D. Vesset, "Four strategies to capture and create value from big data •," *Richard Ivey School of Business Foundation*, Aug-2012.
[36] K. Ebner, T. Buhnen, and N. Urbach, "Think Big with Big Data: Identifying Suitable Big Data Strategies in Corporate Environments," 2014, pp. 3748–3757.
[37]T. U. Daim and D. F. Kocaoglu, Eds., *Hierarchical decision modeling: essays in Honor of Dundar F. Kocaoglu*. Cham: Springer, 2016.
[38] D. F. Kocaoglu, "A participative approach to program evaluation," *IEEE Trans. Eng. Manag.*, vol. EM-30, no. 3, pp. 112–118, 1983.
[39] K. M. Eisenhardt, "Building Theories from Case Study Research.," *Acad. Manage. Rev.*, vol. 14, no. 4, pp. 532–550, Oct. 1989.
[40] R. K. Yin, *Case study research: design and methods*, 4th ed. Los Angeles, Calif: Sage Publications, 2009.
[41] A. Theen, "Report: More Oregonians attend college but don't finish degrees than national average," *OregonLive.com*, 11-Apr-2016. [Online]. Available: http://www.oregonlive.com/education/index.ssf/2016/04/report_more_oregonians_attend.html. [Accessed: 19-Mar-2017].
[42] College Completion Project, "Graduation rates by state | College Completion," 2013. [Online]. Available: http://collegecompletion.chronicle.com/state/, http://collegecompletion.chronicle.com/state/. [Accessed: 19-

Mar-2017].
[43] S. K. Moore, “Portland Tribune: Retiring PSU president reshaped institution,” 02-Aug-2016. [Online]. Available: https://www.pdx.edu/president/news/portland-tribune-retiring-psu-president-reshaped-institution. [Accessed: 19-Mar-2017].

Appendix A: A Sample of the Letter Sent to Experts

Dear X,

Based on our earlier discussion, please find the attached document.
It includes details about the research and how to do the evaluation.

Many thanks for your time and support.

The Document:

Research:
Evaluating Firm's Readiness to Implement a Big Data Projects: A Hierarchical Decision Model

Model Objective

To be used by consultation companies that evaluate firm's readiness to do big data projects.
After a firm decides to implement a big data project, they call the consultant to evaluate the firm's readiness before starting the project.
The consultant first step is to use this model to verify firm's readiness against well-known reasons of failure for such projects.

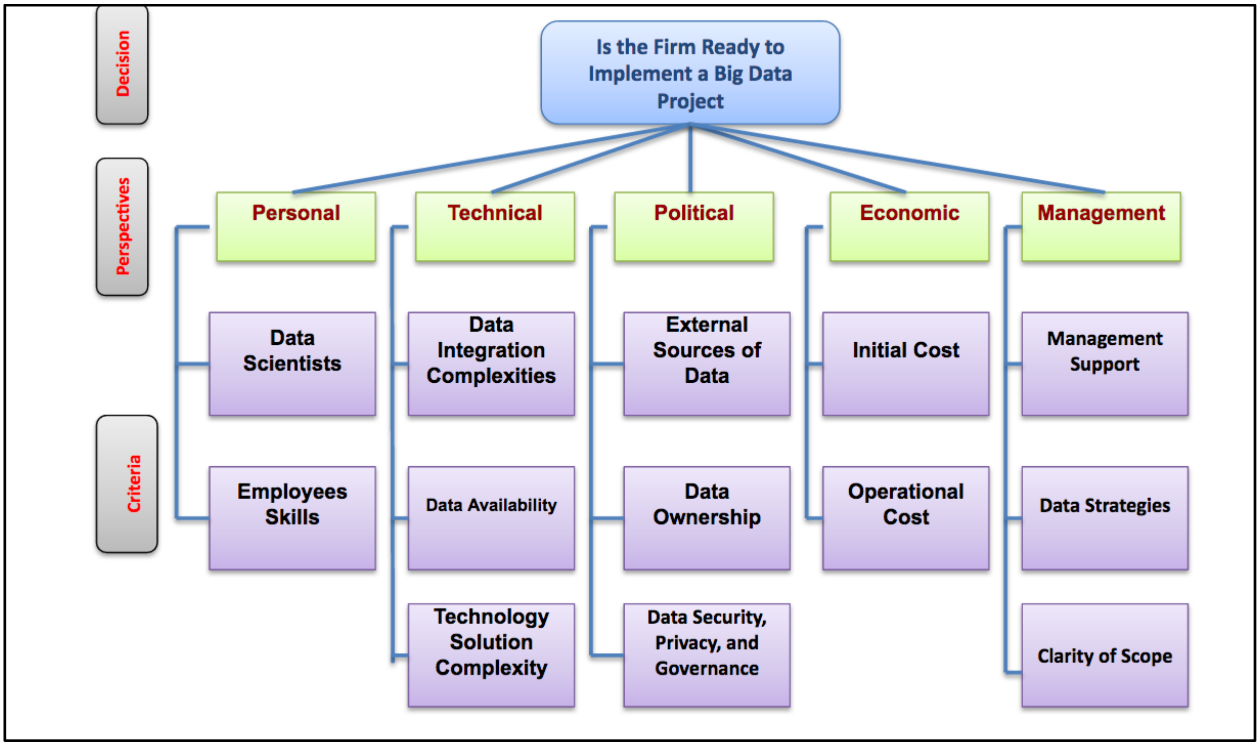
Experts Goal:

Evaluate the model:

- Identifying the priorities of criteria by weighting it using the "pairwise comparison" approach.
    - Compare criteria in pairs by distributing 100 points between each two criterions with the criterion having higher priority getting more points
    - Doing this on the Perspectives and Criteria levels
- Identify any missing criteria, the expert thinks are important, and are not covered in the model

Note: details on how to use the evaluation tool are provided in the "Tool's Tips" section of this document

The Model:



The Criteria Explanation:

Personal Perspective:

1. **Data scientists:** *Data scientists are the minds that can realize what type, size, and frequency of data need to be capture; they device the predictive analysis algorithms that maximize data value, with realization of organizational goals, as well as, internal and external factors around the firm.*
   This criterion evaluates firm's data scientists and their level of experience
2. **Employees' Skills:** *This criterion evaluates employees technical skills and their ability to operate complex software systems*

Technical Perspective:

3. **Data integration Complexities:** *One factor for big data to offer real value, is its ability*

to aggregate and analyse data from various sources.

This criterion evaluates the challenges in integrating the data from the sources that will be used in the project

4. **Data Availability:** *This criterion evaluates various issues related to the data including: volume, velocity, Quality, Collecting the right kind of data*
5. **Technology Solutions Complexities:** *In big data projects, several software tools are used together to achieve the project goals*
   This criterion evaluates how complex is the mix of tools to be used in the project

Political Perspective:

6. **External Sources of Data:** *This criterion evaluates firm's accessibility to external sources of data needed for the project, such data is available at clients, suppliers, and other stakeholders ends, are they willing to share it or not?*
7. **Data Ownership:** *This criterion evaluates how much freedom the firm has in disseminating analysis generated by big data system based on data coming from external resources*
8. **Data security, privacy, and governance:** *This criterion evaluates the level of security and privacy the system must have versus the value it can generate*

Economic Perspective:

9. **Initial Cost:** *This criterion evaluates whether the initial cost is justified in reference of the expected value of the system.*
10. **Operational Cost:** *This criterion evaluates operation cost of running the big data system in reference of the expected value of the system.*

Management Perspective:

11. **Management Support** *Leadership and support for any project plays significant role in the success chances of implementing the project. This even more true in case of big data projects that requires a lot of changes within the organization.*
    This criterion evaluates management level of support for the big data project and the goals it is expected to achieve.
12. **Data strategies:** *This criterion evaluates whether the firm has any strategies, when it comes to handling data*
13. **Clarity of Scope:** *Defining scope is perhaps the most important part of the upfront*

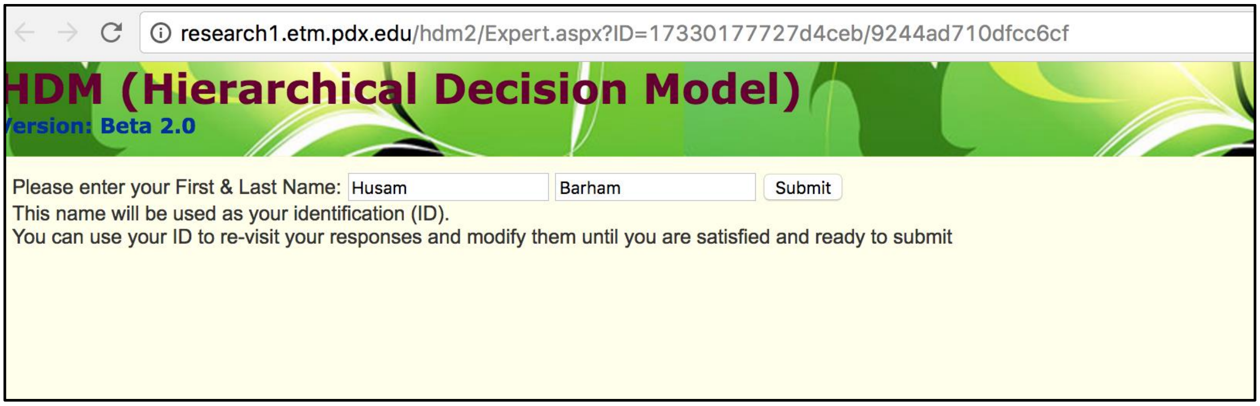
process of defining a project. In fact, if top management doesn't know for sure what they are delivering and what the boundaries of the project are, they have no chance for success.

This criterion evaluates how well-defined the project is

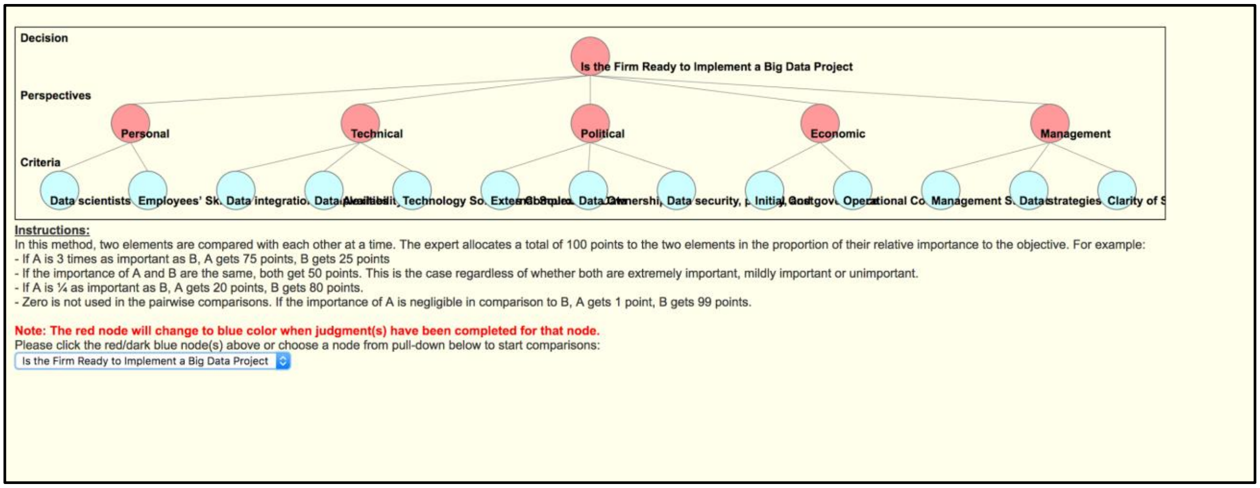
Tool's Tips

- Open the following link in your browser: **http://research1.etm.pdx.edu/hdm2/Expert.aspx?ID=17330177727d4ceb/9244ad710dfcc6cf**

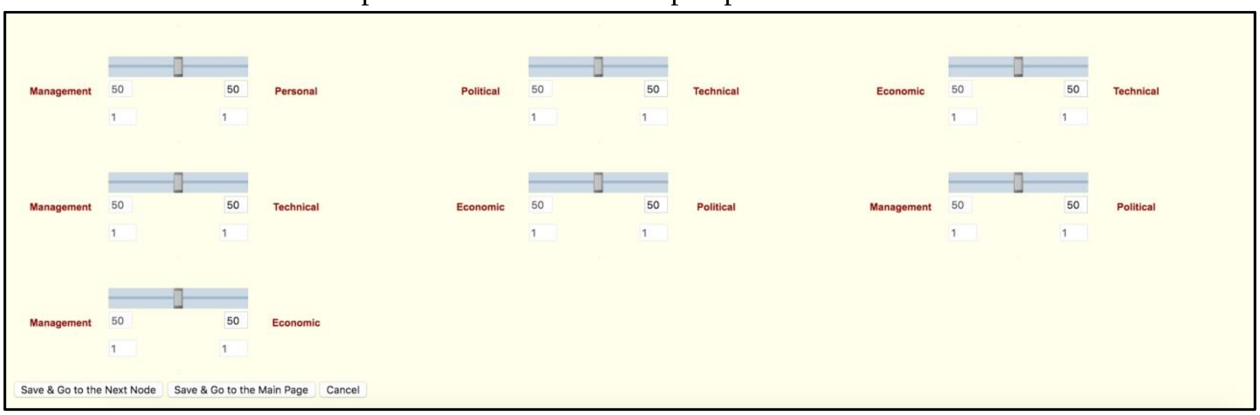
- Enter your first and last names and hit "Submit"



- In the next page, choose "Is the firm ready to implement a Big Data project" from the drop list:

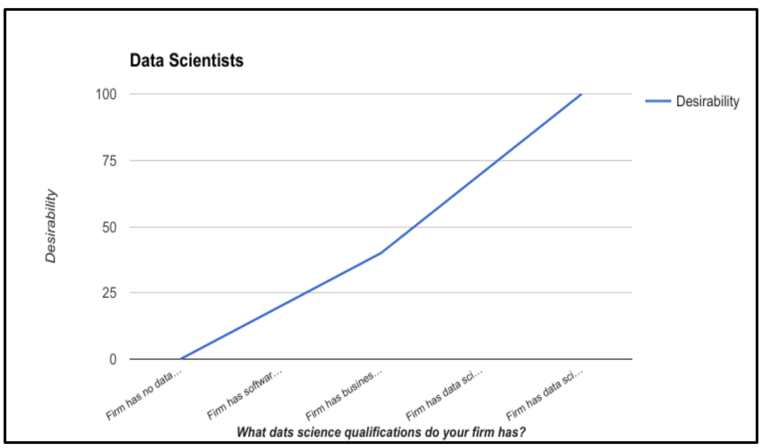


- In the next page start to do "pairwise comparisons" by distributing 100 point for each comparison, the more important criterion gets more points:
  This will be done on two levels:
  - First: the perspectives level: compare the five perspectives against each other
  - Second: compare criteria under each perspective



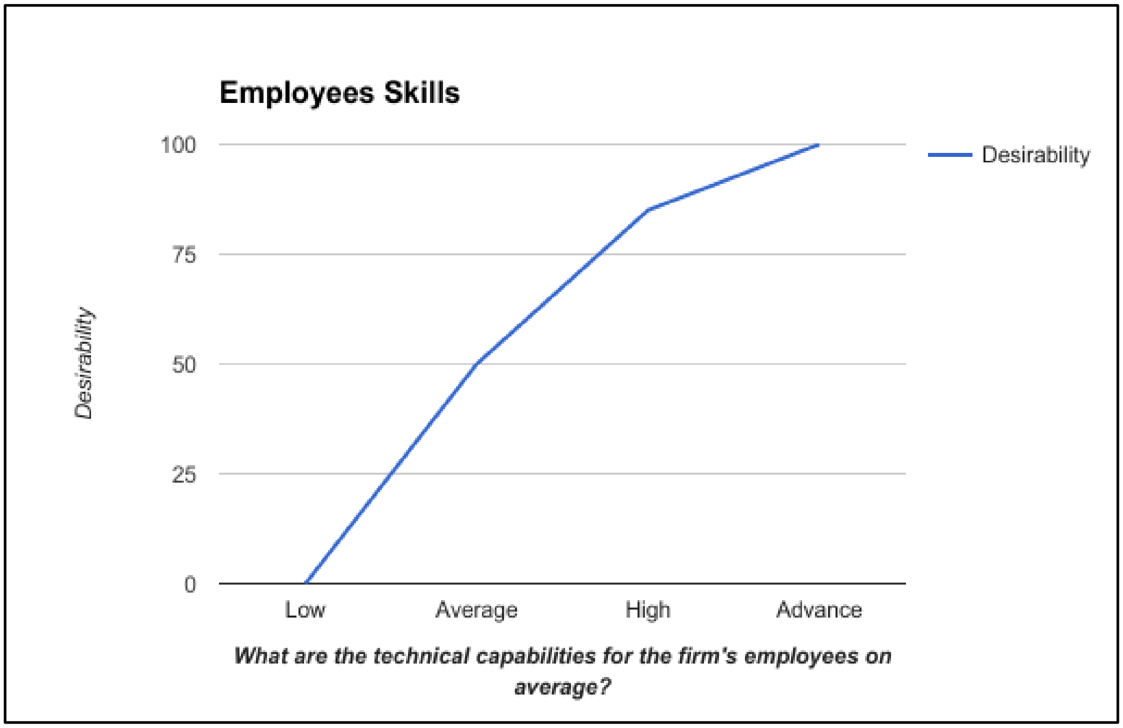
- When you are done, click "Save & Go to the Next Node", and keep doing this until whole model is done, at which stage you will be asked to add any comments and submit the final results.

Appendix B: Desirability Curves



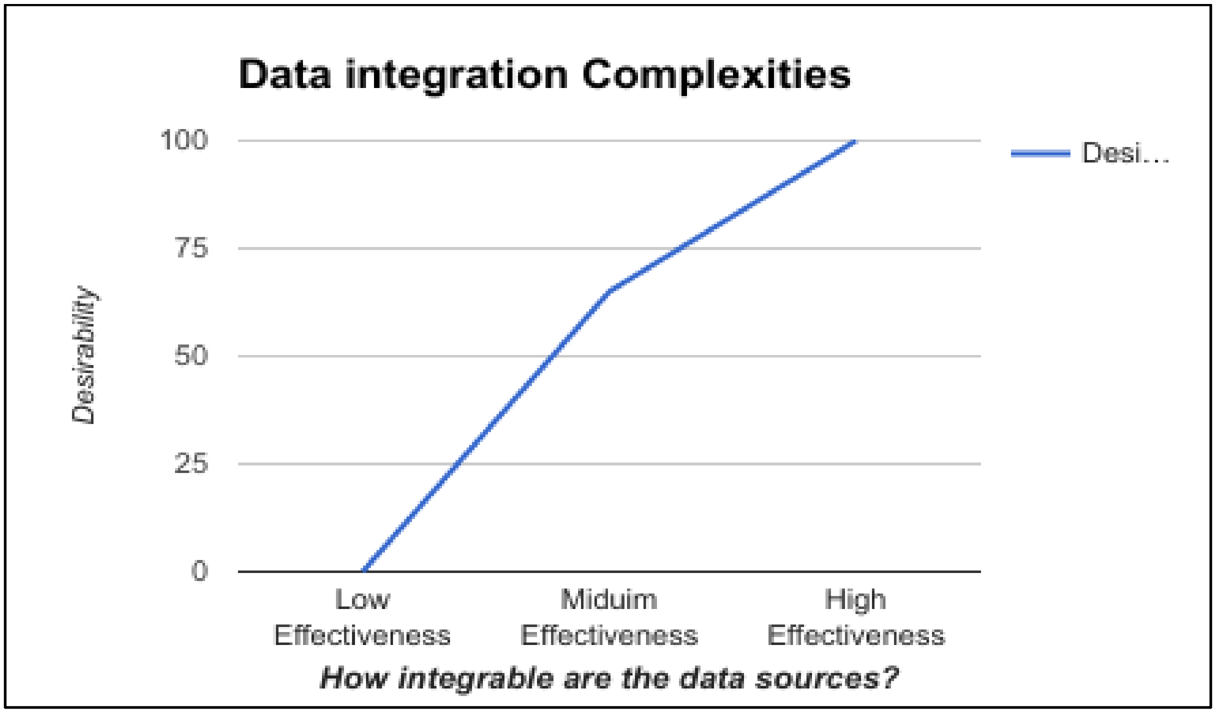
Data Scientists	

Description	Desirability
Firm has no data scientists	0
Firm has software engineers with statistics skills	20
Firm has business analysts with statistics skills and some IT background	40
Firm has data scientists who are not strongly relates to the project goals	70
Firm has data scientists who are strongly related to the project goals	100

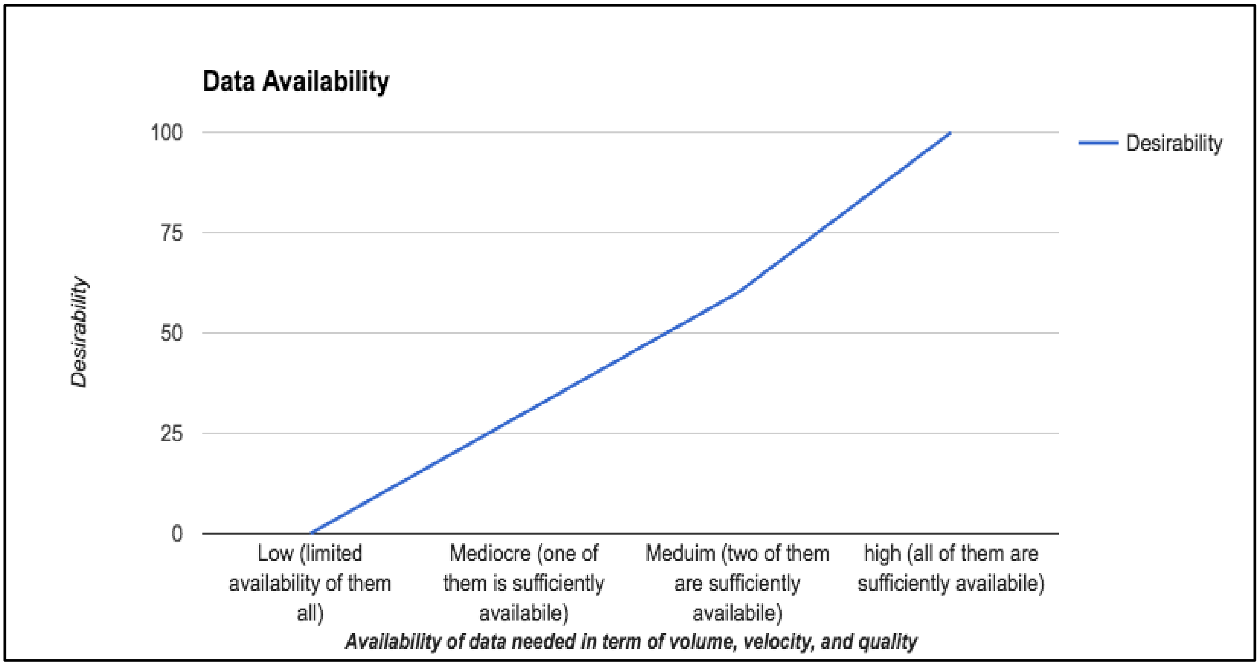


Employees Skills	
Description	**Desirability**
Low	0

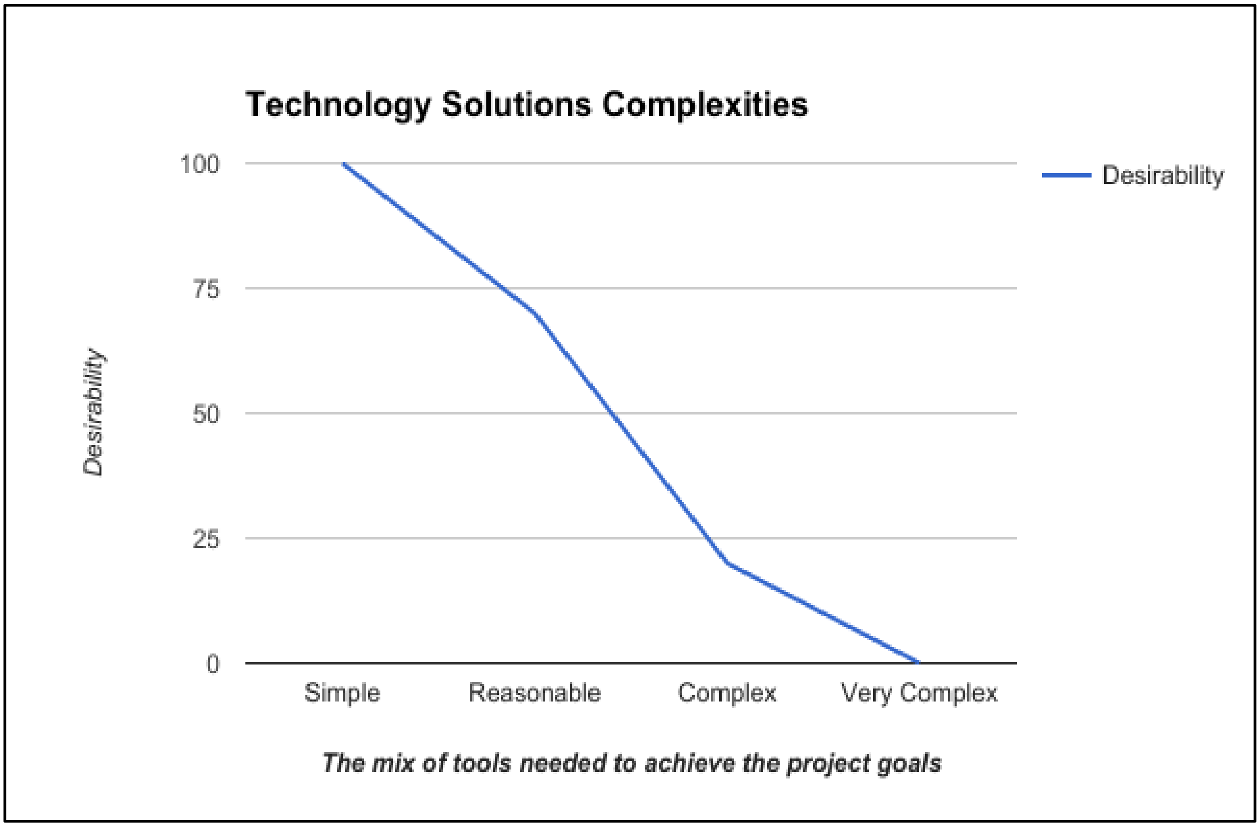
Average	50
High	85
Advance	100



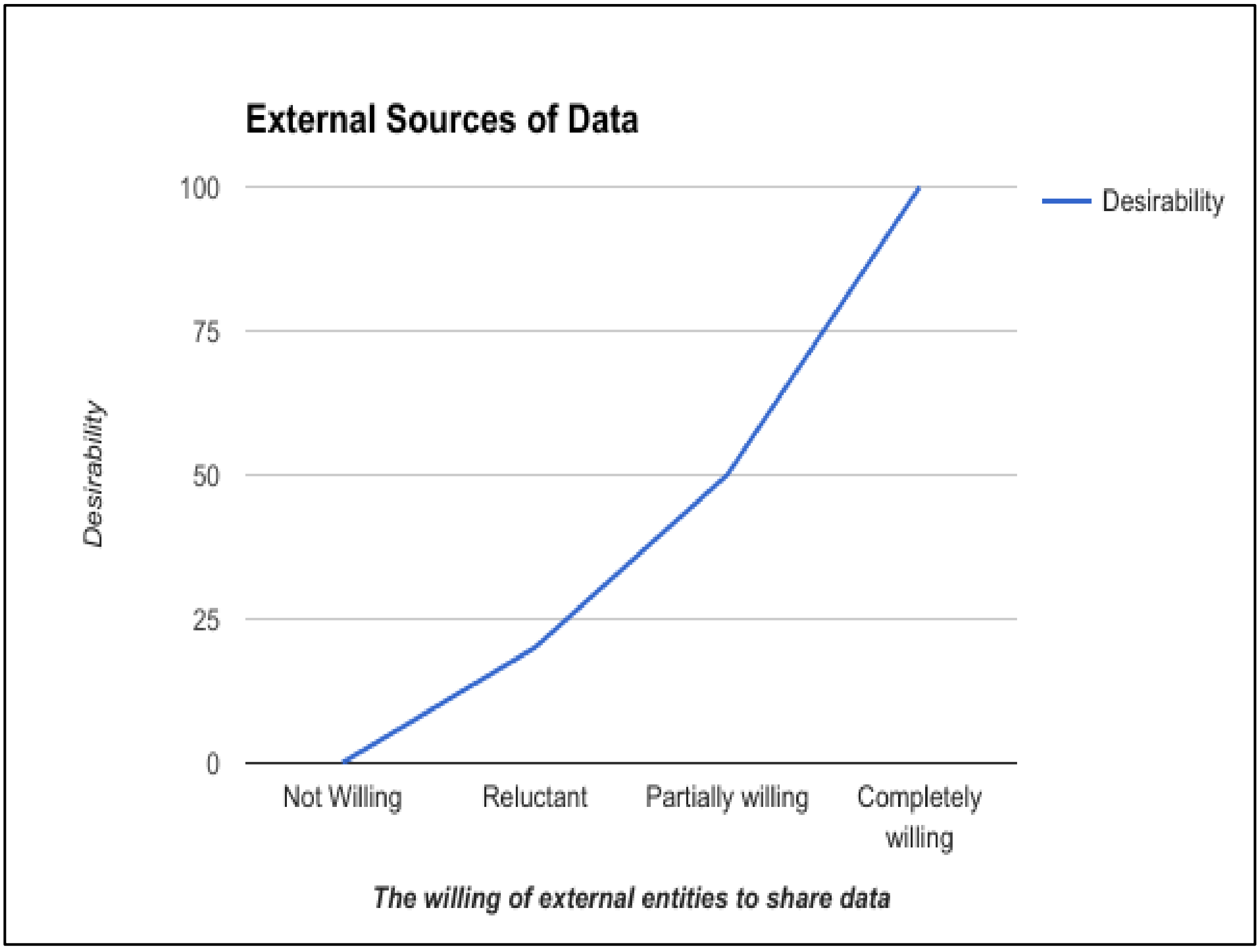
Data integration Complexities	
Description	**Desirability**
Low Effectiveness	0
Medium Effectiveness	65
High Effectiveness	100



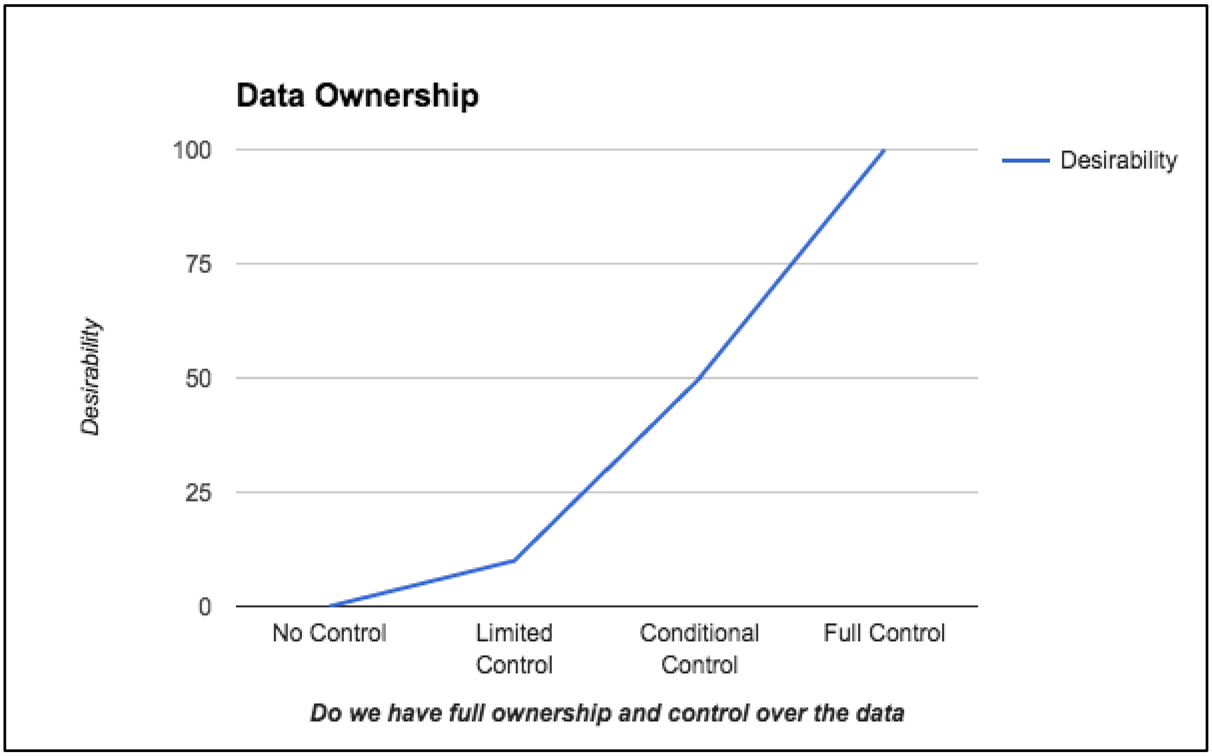
Data Availability	
Description	**Desirability**
Low (limited availability of them all)	0
Mediocre (one of them is sufficiently available)	30
Medium (two of them are sufficiently available)	60
high (all of them are sufficiently available)	100



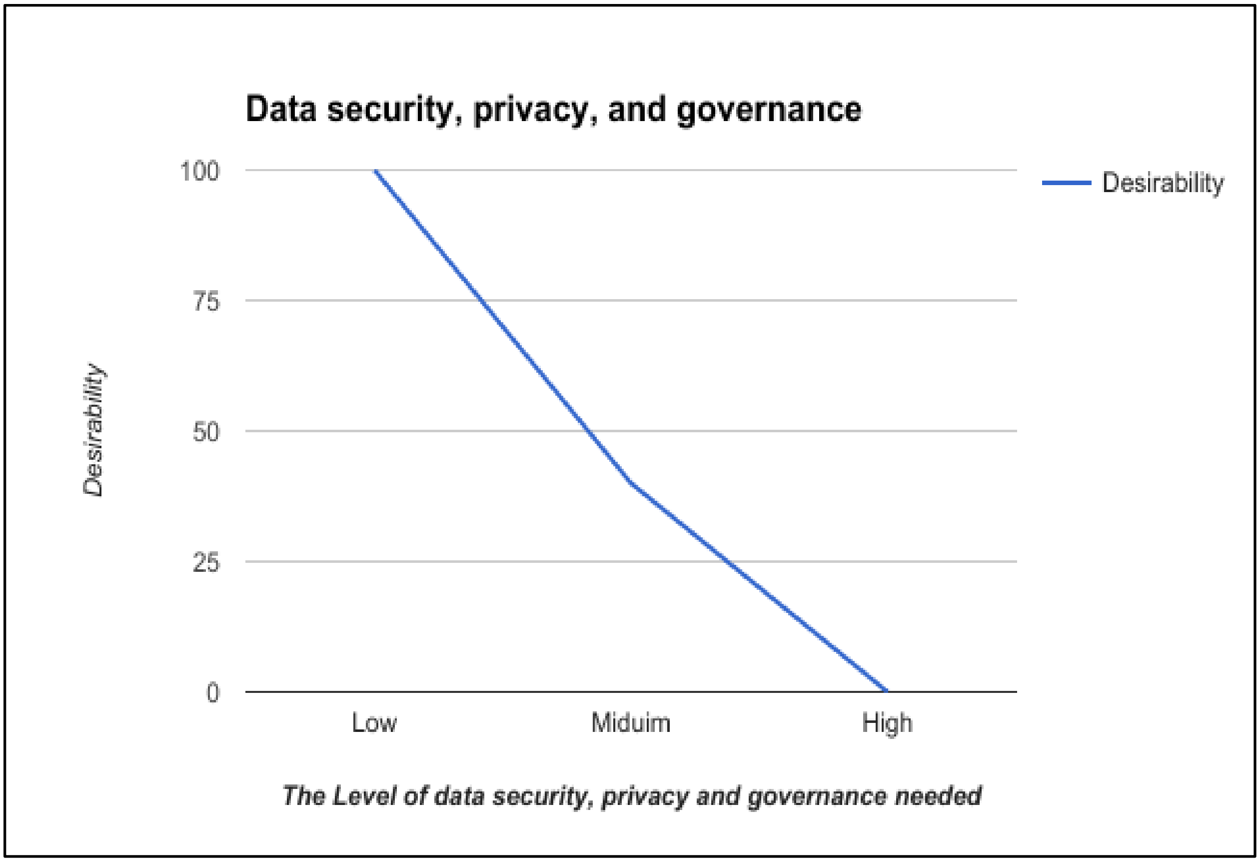
Technology Solutions Complexities	
Description	**Desirability**
Simple	100
Reasonable	70
Complex	20
Very Complex	0



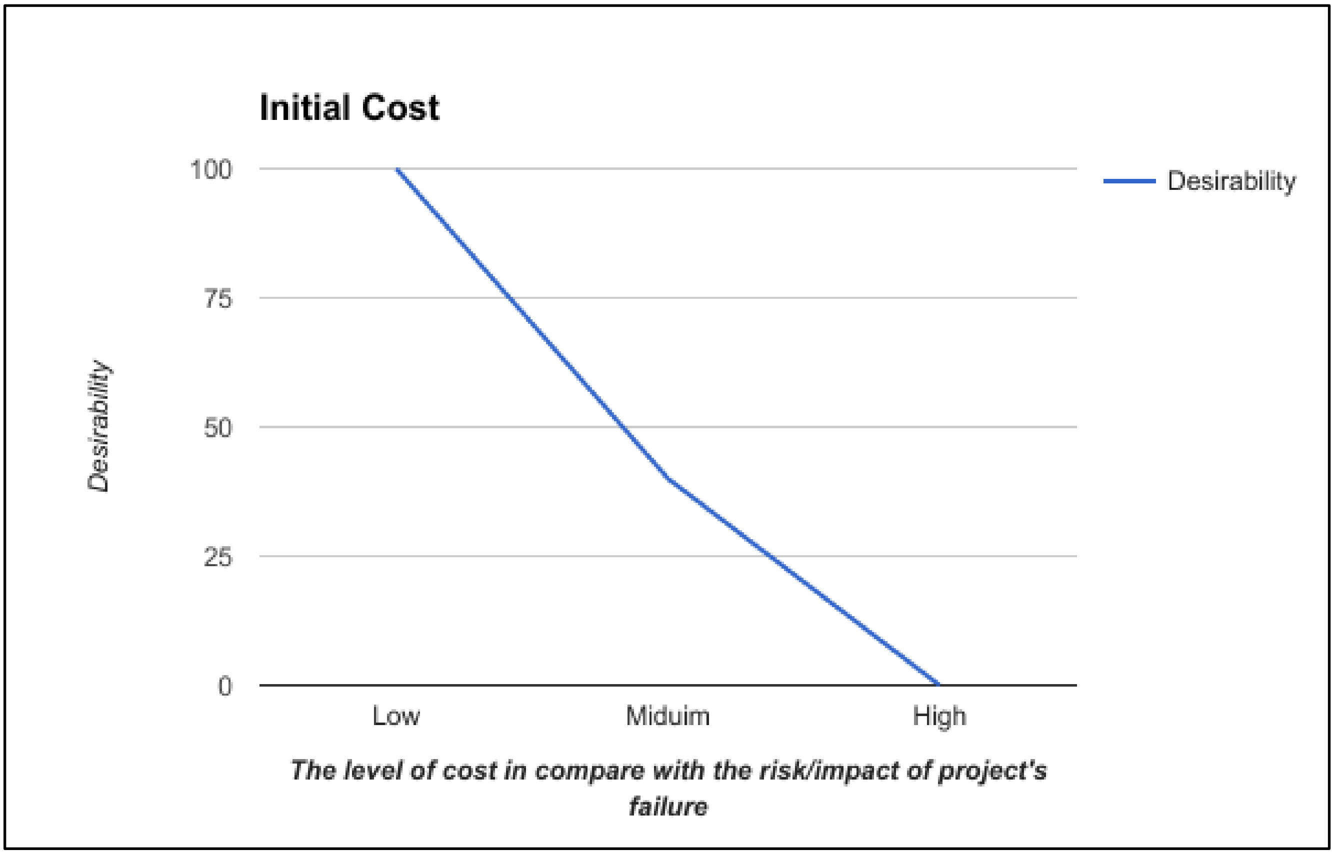
External Sources of Data	
Description	**Desirability**
Not Willing	0
Reluctant	20
Partially willing	50
Completely willing	100



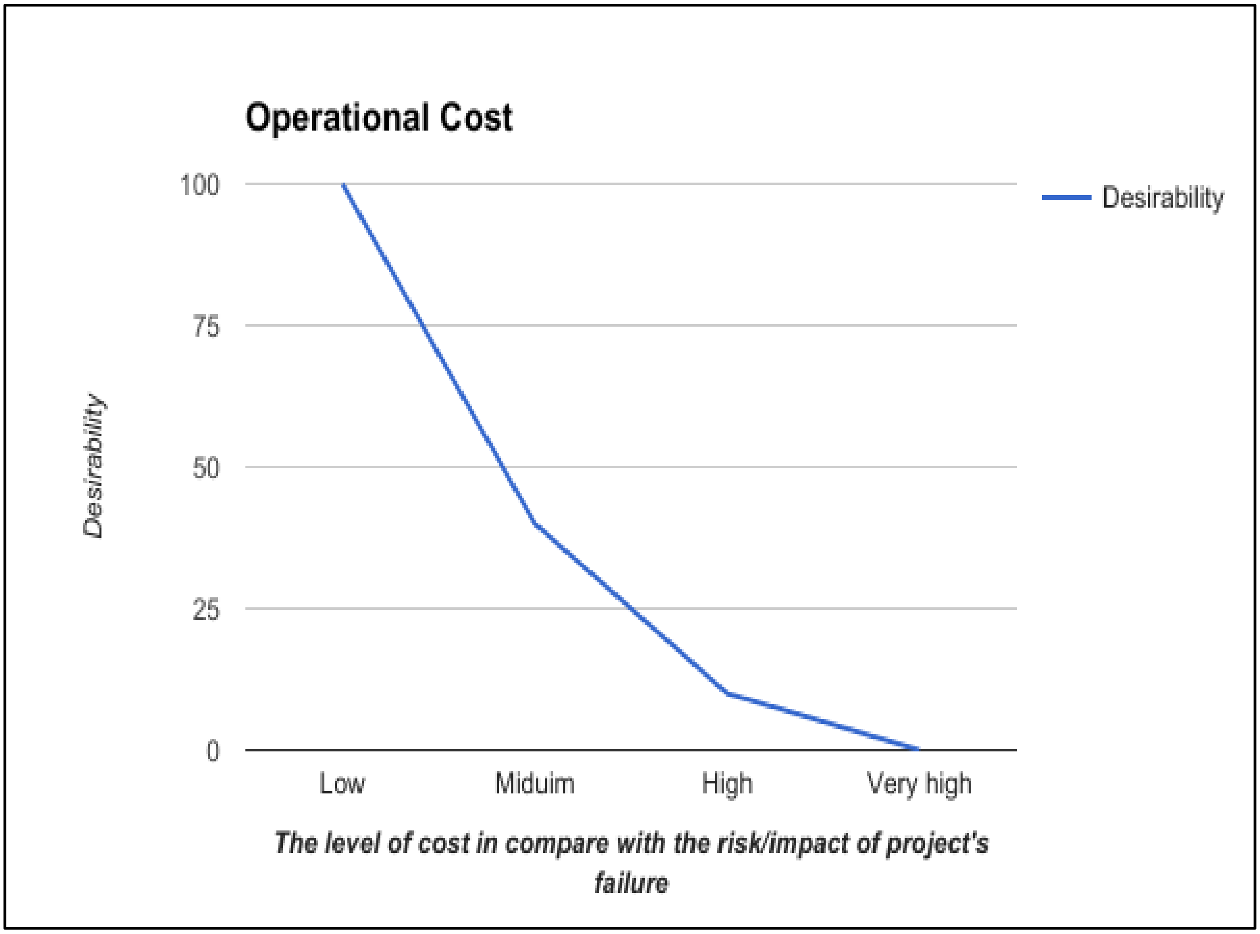
Data Ownership	
Description	**Desirability**
No Control	0
Limited Control	10
Conditional Control	50
Full Control	100



Data security, privacy, and governance	
Description	**Desirability**
Low	100
Miduim	40
High	0

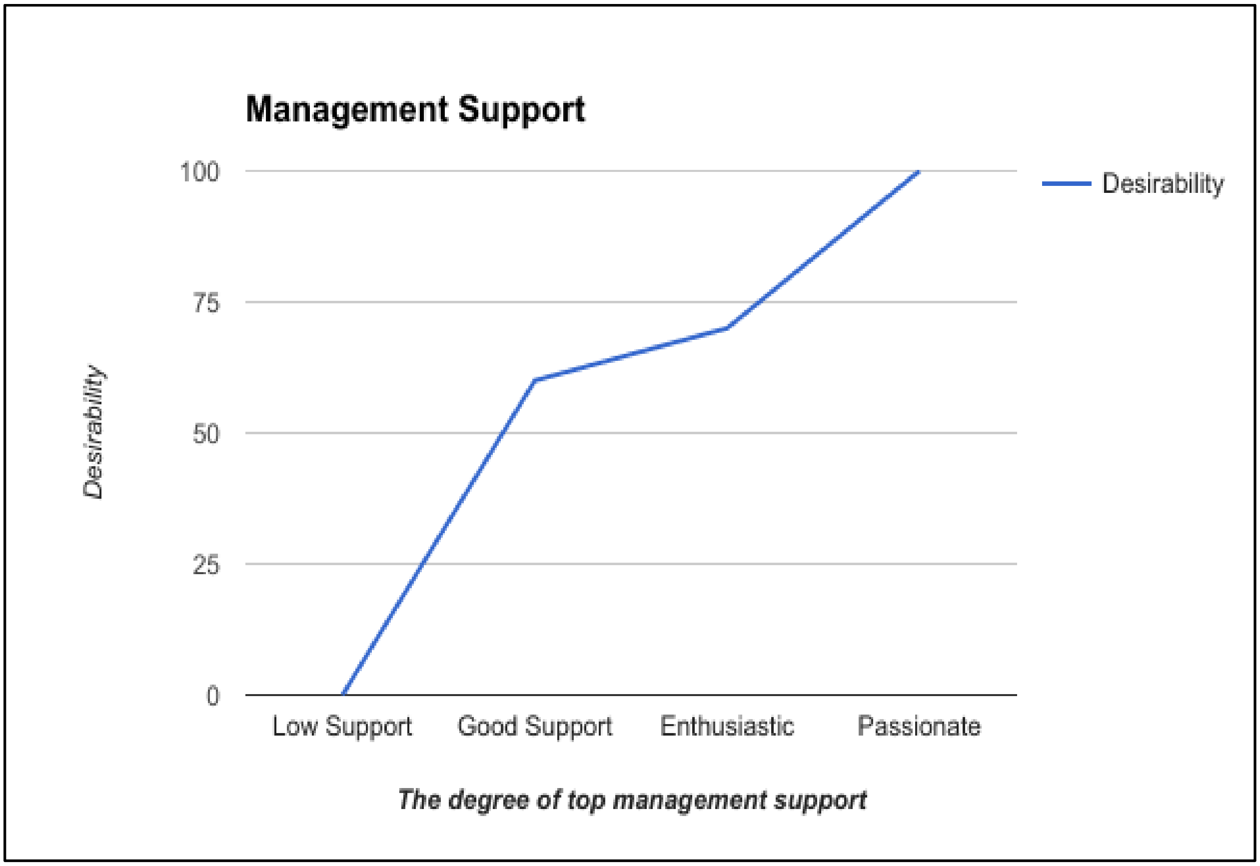


Initial Cost	
Description	**Desirability**
Low	100
Miduim	40
High	0



Operational Cost	
Description	**Desirability**
Low	100
Miduim	40

High	10
Very high	0



Management Support	
Description	**Desirability**
Low Support	0
Good Support	60
Enthusiastic	70

Passionate	100

Data strategies	
Description	**Desirability**
No Strategy	0

simple	20
Mature	70
Advance	100

Clarity of Scope	
Description	**Desirability**
Low	0
Medium	40
High	70
Advance	100