

A New Approach to Business Value Driven Planning for Data Projects

Work in Progress

Authors:

Tadhg Nagle, Room 2.125, Business Information Systems, O' Rahilly Building, University College Cork, Cork, Ireland. t.nagle@ucc.ie (021-4205181)

David Sammon, Room 3.72, Business Information Systems, O'Rahilly Building, University College Cork, dsammon@ucc.ie (0214903836)

Walter Cleary, Irish Management Institute, Dublin, Ireland. waltcleary@gmail.com

Introduction

With the advent of new technology and digital trends, realizing value from data is a top priority for organizations. Along with this, the increased awareness that every business is a data business is beginning to take hold, especially in organizations engaging in data projects through the use of technologies such as Big Data, the Internet of Things and Advanced Analytics. However, it has been shown that there is a lack of understanding on how these projects will deliver value or benefit for the organization (LaValle et al. 2011). Or indeed, there is a lack of understanding on how to effectively manage and govern such projects and capabilities (Tallon 2013). In other words, implementing a data project does not automatically deliver business value, execute as expected, or make your organization data driven. To make your organization more effective when implementing data projects and developing a mature data capability, conversations need to be initiated between stakeholders and focus on the key problem to be solved by the data project. This focus is provided by answering six simple questions: *why, what, when, who, where, and how*. Yet, given the multitude of conversations that need to take place around the problem there are a lack of appropriate tools that can enable stakeholders to reach a shared understanding when planning data projects.

Our research with major public and private sector organizations over the past three years has resulted in the development of a new ‘discursive template’ (c.f. Tsoukas and Chia 2002) (namely the Information Supply Chain – ISC) to promote new transformative conversations within data projects while also producing a more rigorous and robust validation of the potential value of those projects. Moreover, this new approach is the output of four studies, which include: (i) a survey of over 50 organizations worldwide which examined the drivers, goals and barriers of data analytics, (ii) an analysis of 18 projects focused on developing data solutions, (iii) an analysis of over 100 implementations of the ISC, and (iv) one in-depth case study with multiple implementations. The objective of this paper is to present both our new approach for planning data projects along with the insights gained from these studies.

Problems with Planning Data Projects

The necessity of Information Technology in enabling digitization and appreciating the value of data as an asset is undoubted. Yet from our research it has been found that technology can become the key focus of conversations in the planning of data projects to the detriment of the project itself. This mindless behavior is not uncommon especially when the project involves innovative technologies and organizations jump on the ‘bandwagon’ of that technology (Swanson and Ramiller 2004). However, in the case of data projects the focus on technology provides a level of abstraction that assumes the embedded data processes in the technology are aligned to the needs of the organization. Yet, in a climate where technologies are providing a ‘one size fits all’ solution, that assumption can be very costly. Reason being, it postpones the tough data conversations to a later stage in the project when problems begin to emerge, at which stage they become increasingly expensive to fix. Nevertheless, the speed at which some projects need to be implemented necessitates the purposeful postponing of these data conversations. In these cases the future expense is noted under necessary technical debt that will have to be paid at a later stage in the project implementation (Kruchten et al. 2012). However, while this is recognized as typical practice, it does not mean that it is an advisable strategy. Another issue that occurs when there is an overemphasis on technology is that data problems can be mistakenly attributed to technology. By not logically separating technology from data or indeed by not giving data the primary focus of key conversations it can become very easy to label all data problems as technology problems. Naturally, the next rational step in this path is to source a technological solution to a ‘technological’ problem. However, this overlooks the fact that people and organizational processes play just as much a role (if not more) in realizing the potential value from data.

Difficulties with planning data projects can also be attributed to the liquid nature of data and resulting flow across all boundaries (functional and managerial). This would suggest that (i) all organizational parties have a collective responsibility to ensure data quality (and potential value) is kept high, (ii) its flow is not obstructed through the organization, and (iii) there is a collective ownership from which all parties should benefit. However, ownership of data is often abdicated by those who have most to benefit (business owners) and transferred to those who lack the knowledge/domain expertise (technologists) to ensure its quality is maintained and value realized. More importantly, it is communication between all data stakeholders that provide the foundation for realizing the value of data rather than the type of technology used (Redman 2013). However, with a divergent range of mindsets, agendas and languages, the ability to have a clear conversation between these stakeholders becomes much more difficult. This

clearly indicates a need for an approach to enable the development of a shared language and shared vision from which all the stakeholders can communicate effectively and align with organizational value driven objectives (Agarwal et al. 2009). While these problems may be well known, if not self-evident (Nelson and Coopridge 1996) it is still a major problem for data projects (Redman 2013).

Research Grounding

This article, aimed at practitioners undertaking data projects is based on data collected from four linked studies over the course of 3 years. As well as providing insight into the delivery of data projects, the studies also track the iterative development of the Information Supply Chain (ISC) from an analysis framework to a project planning approach.

The first two studies focused on examining the issues in delivering data projects, while the latter two studies provided insights from the implementation of the new approach. Included in the first two studies is: (i) a survey of that examined organizations' data analytics capability, and (ii) an analysis of 18 data projects. The survey focused on: the drivers behind, the goals/benefits of, and the barriers to data analytics. In addition, the maturity of various data capabilities and whether a data governance program has been implemented was also examined. As part of the study, the ISC was used to structure and later analyze the data collected. In total, the survey collected data from 51 organizations globally. Fifty-five percent responding organizations are headquartered in North America, another 22% in Asia/Australia/Pacific, 14% in Europe, and 6% in Africa, with the remainder headquartered in South America and the Middle East. The second study involved the analysis of 18 Design Research projects that specifically focused on the development of data solutions. In particular, each of the projects lasted over 18 months with the main objective of designing, building and evaluating an artefact to solve a key data problem for their organization. All the organizations were either large multi-national corporations or large national entities and included both public and private sectors. Each project was documented in a report of up to 16000 words document which were analyzed using the phases of the ISC.

The latter two studies involved: (i) the analysis of over 100 ISC implementations across 58 organizations, and (ii) a longitudinal case study of the implementation of the ISC approach in detailing a data strategy for an organization. Each of the 41 ISC implementations involved separate individuals on separate projects across 21 organizations. The analysis provided useful insights for improving the approach as well as highlighting some of the general trends across each of the phases, namely: acquisition, integration, analysis, delivery, governance, and value. Lastly, the final study involved a longitudinal case study of the implementation of the ISC approach in detailing a data strategy for an organization. The case study of a data services provider lasted for over 12 months during which multiple ISC instances were created. The impacts from these instantiations provided rich insights into the use of the ISC as a planning approach for data projects as well as the key benefits from its utilization.

The Information Supply Chain

In this article, we present a new approach to facilitate value driven planning for data projects that is based on a rigorous and systematic exploration of the flow of data through four phases: (i) acquisition, (ii) integration, (iii) analysis, and (iv) delivery. Underpinning these four phases is data governance, which along with the four phases is ultimately converted into value (see Figure 1). Providing a holistic view on a potential data project, the approach draws on concepts from the areas of Information Systems and Design Thinking. From Information Systems the components of *'people, process and technology'* are incorporated, as it is the fusion of these three components that not only make up an Information System but also dictate its success (Keen 1981; Keen 1993).

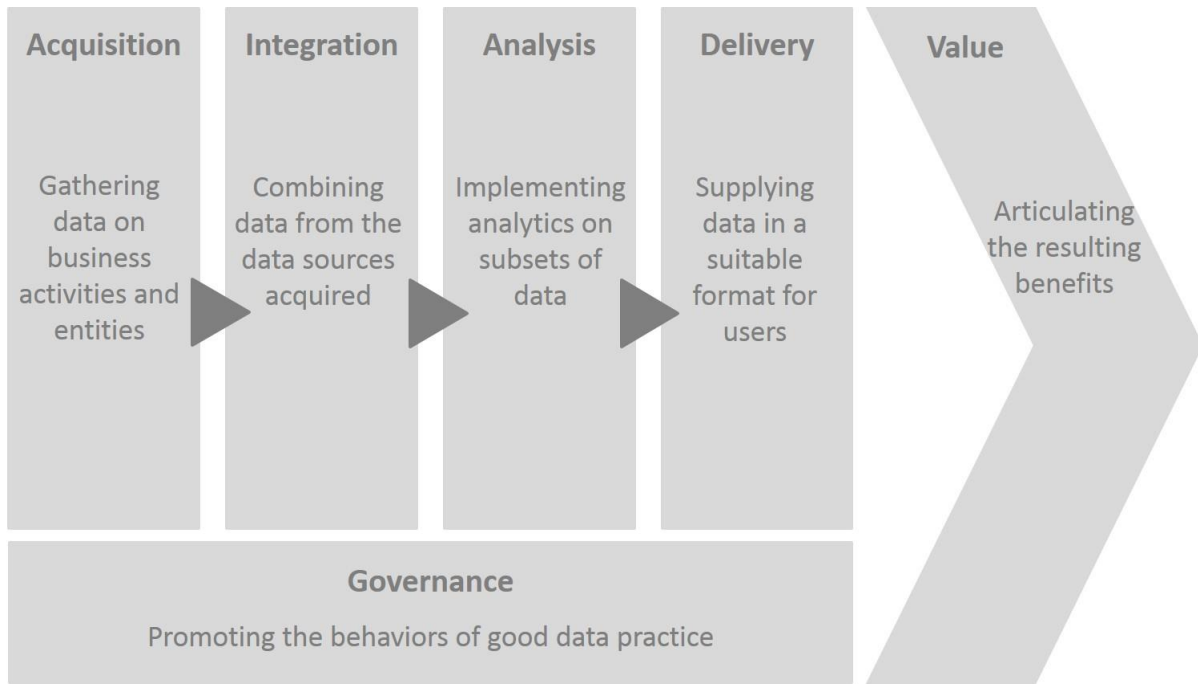


Figure 1: The Information Supply Chain (ISC)

With data being the key focus of the ISC, the alignment of people, process, and technology with the capability to “organize, govern and share data to achieve business ends” (Information Builder, 2014 p.1), is explicitly demonstrated in this paper to promote conversations. To aid the detailing of people, process, technology and data aspects of a data project six very simple questions (Why, What, When, Who, Where and How) ensure all underlying project assumptions are at least questioned if not fully examined. This forms the basis of the ISC card (see Table 1), which is then customized to each of the first four phases of the ISC (see Tables 2-5).

Data	Why is it ...? This ensures the underlying objectives and motivations for doing the data task is clearly detailed.
	What data is....? This ensures the actual data involved in the operation is defined.
	When is the data.....? This ensures the temporal aspect of data is not overlooked. Very important in keeping the quality of data high.
People	Who ... the data? This ensures all stakeholders are identified for all data tasks outlined. If possible these stakeholders should be included in project discussions.
Technology	Where is data....? Details the technologies that are needed for each of the data tasks.
Process	How is it....? Outlines the actual processes needed to move the data from one phase to the next in the ISC.

Table 1: Basic Structure of an ISC Card

With the basic structure of an ISC Card outlined, the key in implementing the ISC approach is completing a card for each: data source, data integration point, data analysis process, and data delivery type. Once these are completed they are visualized as abstracted nodes on the ISC itself to effectively communicate the output (see Figure 2 for a completed example). At first the thought of completing each set of questions for every aspect in the ISC may seem repetitive if not total overkill. However, the very act of going through these questions each time makes the process rigorous with every implicit assumption being questioned. Finally, for the approach to be effective in enabling value driven conversations and successful data projects; it needs to facilitate cross-functional collaboration and also

challenge the status quo, to get people thinking outside the box and beyond data silos. To achieve this objective the approach needs to be applied in the spirit of Design Thinking, a human centered approach to solving problems, which embraces integrative thinking, optimism, experimentation, collaboration and stakeholder empathy (Brown 2008). By incorporating these techniques and perspectives the ISC provides the basis of a very strong approach to the planning of data projects that also differs from other approaches in the following ways:

- Outlines four logical phases that data flows through and provides specific questions to explore and detail key aspects at each phase.
- Positions value as the key output at every phase and ensures this value is explicitly outlined through a description of the benefit, measure of the benefit, and identification of owner.
- Enables a level of abstraction that allows the flow of data to be visualized and patterns to be identified.
- Provides a framework for unambiguous data conversations as technology, people and process are accurately positioned in relation to data, giving a clear context to such conversations.
- With the notion of an ISC Card detailed questions are supplied to ensure a rigorous and collaborative approach to the planning process.
- Integrates data governance as part of project planning while also providing a more aligned bottom up approach to data governance implementation. This approach is set out in terms of data behaviors, benchmarks for such behaviors and the owners that are accountable for promoting the behaviors.

The following section defines each phase of the ISC and outlines the associated ISC cards (and incorporated questions) for each phase. In addition, findings from our studies are highlighted within each phase to provide context to the Cards and insight into completing a value driven planning process for data projects.

Acquisition

The acquisition phase of the ISC details the gathering of data on business activities and entities. The rigorous examination of the data sources (see Table 2) is especially important as potential data issues should be identified as early as possible in both the planning process and in the ISC. Many of these issues arise as many organizations do not know what data they have, how critical it is, where the data is stored or the degree of redundancy of that data (Khatri and Brown, 2010). In addition, the rigorous examination of why the data needs to be acquired will make the gathering of data much more focused on the value the data source will bring.

One example of uncovering a data issue in the acquisition phase highlighted some unusual characteristics around the sourcing of data. A senior manager of a customer services group who needed to track the group's performance found himself sourcing data from an internal analytics package and then storing the data in an excel spreadsheet to be modified into a state from which it could be integrated with other data sets. Before mapping this sequence through the ISC, the unorthodox behavior of taking data that had already been analyzed for a very specific purpose and reverse engineering it into a raw spreadsheet format was never questioned or seen as anything but the right thing to do. This example shows the power of visualization in uncovering mindless behaviors while also providing a platform for transformative conversations around these behaviors. From the cases that have used the ISC, the widespread use of spreadsheets is evident across all the four phases. However, like the example above, the approach quickly highlights the issues of using spreadsheets in phases outside of Delivery and the need to have strong governance behaviors to ensure high data quality where spreadsheets are involved.

Data	Why is it acquired? This seeks to uncover the motivation as to why the data is collected and provides a basis from which its alignment with the overall objective of the information supply chain can be judged.
	What data is acquired? This asks to describe the data itself. This provides a label for the data (eg employee data) and if needed the exact fields in the data set acquired.
	When is it acquired? This outlines the temporal aspect of the acquired data and can highlight potential data quality issues or bottlenecks in the data flow.
People	Who is it acquired from? To ensure all stakeholders are identified and the fact that data can be acquired from both internal and external sources, it is important that these people are identified if not fully included in the planning process.
	Who acquires it? This outlines the people who are involved in the acquisition of the data. This can be frontline employees that manually acquire the data or technical staff that support automated acquisitions.
Technology	Where is it acquired from? The need to detail the technologies/applications used by the sources is also important. Changes in these technologies, while possibly outside of an organizations' control can have substantial impacts on the data flow.
	Where is it stored? Once it is acquired the data will be stored in a particular technology/application, which will also need to be detailed.
Process	How is it acquired? Finally, the process in which it is acquired is also needed. At a very simplistic level this can be either manual or automated but can be described in more detail if required.

Table 2: Acquisition Card.

Integration

The next phase of the ISC is integration and describes the combining of data from the data sources acquired. This is indeed an important aspect in the process and was noted by 45% of companies in our survey as being their biggest pressure in driving analytics investments, due to their data being stored in silos. While the storing of data in one location can resolve silos, it will do little to enable an organization to become more data driven. Activities normally associated with data are primarily associated with ETL (extract, transform, and load) technologies. However, ETL at a conceptual level denotes that data will be translated from its 'source' structure to a new 'target' structure. This translation from source to target requires us to think 'ontology' with regards to our data. In fact, thinking ontology translates as thinking data models. To date the data models that dominate in the integration phase are those most often associated with a data warehouse. These data models are predominantly relational, with 80% being dimensional (e.g. star schema) and only 20% being normalised. Data warehouse designed to conform to those data models are known as being a Kimball or Inmon architecture, respectively. So the value of using the ISC to visualise the integration phase is that the organization gets to understand how integrated their data actually is, and indeed how close they may be to the 'holy grail' of a single version of the truth (SVOTT), which from our only 20% of organisations have.

Two examples of a more rigorous examination of the integration phase highlight the value of such an activity as both had financial impacts over a million euros. The first example involved the Logistics function of a large multinational that handled over 2.25 million shipments annually. In this project, a more rigorous examination of the integration phase uncovered a multimillion dollar exposure of defective returns that had not been accurately accounted for, due to datasets not being shared with the Logistics team. The second example for a multinational in the Food and Drinks industry, demonstrated that a keener exploration shifted the organizations focus from being technology centric to data centric. With this change in focus, the organization stopped an asset management project and avoided a very expensive IT implementation that would have not met business expectations. This exploration into the data integration phase of the ISC highlighted that while the IT system would form part of the solution, it would not have been enough to bring about a resolution to the business problem. In essence, both of these examples illustrate the power of having open and clear 'data' conversations, which have real financial impact. For the complete list of questions to be completed on each data integration point in the ISC, see Table 3.

Data	Why is it integrated? Not all data needs to be integrated. A keener focus on why data must be integrated will make the project more effective and utilization of the data more efficient as the unnecessary integration of data will slow system performance.
	What data is integrated? An understanding of what data is to be integrated may highlight the possibility of joining common datasets to the benefit of the organization as a whole.
	When is it integrated? The timeliness of data across all the ISC has a big impact on data quality and as a result must be well documented.
People	Who integrates the data? This is important as the people integrating the data will need a strong understanding of the business to ensure the integrated data will provide a solid platform for valuable analysis. Also in the planning stage of data integration having business people involved has shown to increase the longevity and quality of the end result.
Technology	Where is data integrated? For more traditional systems you mainly have Relational Databases or Data Marts and Data Warehouses that integrate structured data. More recently there is the advent NoSql databases and Data Lakes that integrate structured and unstructured data.
Process	How is it integrated? Currently Extract Transform and Load (ETL) process are the most common in organizations. Other processes could include, Master Data Management, Semantic Modelling or Data Modelling.

Table 3: Integration Card.

Analysis

The data analysis phase describes the implementation of analytics on subsets of data and can be fully detailed by the questions asked in Table 4. It has been shown that through the development of an analytics capability and putting the capability to work in the right way (the subprime financial crisis is an example of where analytics were used incorrectly), organizations have been able to gain real benefits such as: better strategic decisions, improved asset management, and better risk mitigation (Davenport and Harris 2007; Davenport et al. 2010). While there are a number of ways to do analytics, they mainly come under the headings of descriptive, predictive and prescriptive. Descriptive analytical models describe what happened in the past, predictive models utilize a variety of statistical techniques and are applied to historic or current data to make a range of predictive judgements, and prescriptive models go beyond predictive analytics and detail a course of action along with predictive statements. In addition, analytics can come under the remit of a variety of roles, such as: Business Intelligence Developers, Data Scientists, and Data Miners. Notwithstanding the diversity in roles, they all need to collaborate with a range of stakeholders to bridge their knowledge of the business, with their capability to derive valuable insights.

Data	Why is the data analyzed? Keeping a strong focus on the objective of the analysis will enable more efficient value realization rather than efficiently executed analytics.
	What data is analyzed? A key understanding of the data analyzed will facilitate a basis for a stronger alignment in choice of analytical techniques.
	When is it analyzed? The timeliness of analytics and their resulting actions.
People	Who analyses the data? Diverse stakeholders and roles are involved in the delivery of analytical insights.
Technology	Where is the data analyzed? As the volume of data has increased, so too has the volume of analytical applications and technologies. Technologies range from OLAP, Data mining, Statistical/Mathematical, Advanced Analytics and Business Intelligence packages.
Process	How is the data analyzed? Within the different types of analytical models (descriptive, predictive and prescriptive) a number of statistical techniques can also be utilized.

Table 4: Analysis Card.

During the development of the predictive model of customer churn for a Cloud SaaS vendor, it became evident that the emphasis switched from the initial focus on factors that make customers churn to a focus on customer retention factors. From a business perspective the two mean the same thing but from an analysis perspective they are very different. During the project it was found that customer support

was a key aspect in customer retention and in particular, customers that had opened a case with customer support had a lifetime 6 times longer than those that never opened a call. This enabled the organization to be more proactive in their customer support, viewing this strategy as revenue generation rather than cost incurring activity. While the shift in focus was subtle (from why do customers leave to why do customers stay) the impact it had on the project was substantial. This example demonstrates the need to constantly focus attention on the overall objective (why) rather than the actual predictive model developed (how).

Delivery

The delivery of data focuses on the supplying of data in a suitable format. Delivery is a key aspect of the ISC as our study indicates that 40% of the respondents highlight the need to provide interactive/drill-down BI dashboards to more employees, as being a key motivation for data projects. This is not surprising as the symbiotic relationship between the format of data and the actual content has been widely acknowledged (McLuhan and Fiore 1967). This sentiment was also strongly recognized in a number of the projects where the 'look and feel' of the analytics were deemed just as important as the analytics themselves, as an unappealing appearance can quickly turn users off. On the other hand, good data presentation can make analytics (and the potential insights) accessible to a wider audience.

For instance, one data visualization project demonstrated that data does not always have to be delivered as a report or dashboard. In this case an infographic (in poster format) was created of the cost of maintaining the features of a globally successful software product and the corresponding utilization of these features by customers. The infographic highlighted that it cost over €500,000 to maintain the bottom 10 features ranked by customer utilization. Two of the features were utilized by a single customer, which ultimately highlighted the need for a more focused feature retirement process. Moreover, the delivery of data can also have a major impact on the efficiency of decision making. Firstly, for the manufacturing function of a tech multinational, the planning of an operations cockpit detailed savings of almost €750,000 from (i) time saved by managers in managing operations, (ii) increased productivity from staff due to production information being dispatched earlier, and (iii) the more efficient allocation of staff. Finally, efficiencies can also be gained through improving the delivery format itself. Secondly, for a customer service department, the organizing of a customer report for a high value customer would take 4 people 3 working days, from which the output would be a 50-page presentation. That was converted into a Customer Service executive dashboard, which could deliver that same information instantly. What these examples illustrate is the benefit of being mindful of the format of data when it is delivered and value that can be realized when the delivery is aligned with the needs of data consumers. To support this alignment, the questions in Delivery Card (Table 5) need to be completed for each type of delivery.

Data	Why is the data delivered? What are the key decisions to be made or what key objectives in delivering analytics will play a major role in forming the answers to the proceeding questions.
	What data is delivered? The analytical datasets are needed to facilitate the attainment of the objective.
	When is the data delivered? The timeliness needed to ensure the data delivered will be of good quality.
People	Who receives the data delivery? The key stakeholders that requested or need the data.
Technology	Where is it delivered? The technologies used can be a part of the analytics packages but can also be separate depending on the delivery requirements.
Process	How is the data delivered? The process would determine the type of format, such as: an infographic, dashboard, or report.

Table 5: Delivery Card.

Governance

Underpinning the four phases of the ISC is data governance, which is defined as the promotion of behaviors for good data practice. While being a very important aspect of the ISC our research has shown that data governance is a problem area for the majority of organizations. From our survey there is recognition that the implementation of a data governance policy does not take that long with 53% of respondents indicating that a data governance policy can be implemented in less than 3 months;

however, only 10% actually have a policy in place. In detailing the behaviors, a benchmark for the behavior as well as an owner or person accountable for the promotion of the behavior is sought (see Table 6). The types of such behaviors can be categorized under the following headings taken from Khatri and Brown (2010): (i) data principles – clarifying the role of data as an asset, (ii) data quality – establishing the requirements of intended data use, (iii) metadata – establishing the semantics or ‘content’ of data so it is interpretable by users, (iv) data access – specifying access requirements of data, and (v) data lifecycle – determining the definition, production, retention and retirement of data.

For certain behaviors there are documented measures and activities available to develop behaviors and benchmarks (Redman 2008). However, there will always be the need to develop such measures and activities that align to the unique nature of organizations. For example, in our research with a large organization that provides statistical reports as their core value proposition, the need for governance around the use of SAS analytics was highlighted. Yet, no benchmark was available for the task. As a result, the organization focused their effort in building a maturity model for SAS analytics end-user computing. This enabled the organization to firstly measure the level of governance and also develop benchmarks to guide its maturity. From documenting the behaviors, benchmarks and owners it is possible to get an overview of the level of governance employed along the ISC and the potential to identify areas that also require attention. For one project in the study, the detailing of governance behaviors in a project by project approach, provided a more realistic and cost effective way of developing a comprehensive data governance program. This aligns with recent research on big data that noted, the identification of data governance practices that maintain a balance between value creation and risk exposure is the new organizational imperative for unlocking competitive advantage and maximizing value (Tallon 2013).

Governance Category	What data governance factor is the behavior directed towards? This can be one of the following: data principles, data quality, metadata, data access, data lifecycle.
Behavior	What is the behavior? This describes the behavior and should align with the factor detailed.
Benchmark	What is the benchmark defined for the behavior? This defines the level of acceptability or indeed a goal to achieve for the organization. It also provides insight into how the benchmark is measured.
Owner	Who is responsible for the promoting the behavior? For some organizations it may be their Data Governor, for others it may be the business or line manager.

Table 6: Governance Card.

Value

The difficulty in measuring and communicating the value of IT in general has been previously highlighted and tackled by previous studies (Mitra et al. 2011; Ward et al. 2008). Leveraging these studies and in particular the IT Business Case from Ward et al. (2008), the ISC approach enables stakeholders to explicitly detail data as an asset but also highlight areas in which poor practices are costing the organization. The need to help organizations focus on detailing benefits as part of the ISC is particularly evident in our research with 58% of our survey respondents never calculating the cost of poor quality data. Moreover, the difficulty in being able to make a business case for a data project has been highlighted as a key issue for a number of organizations involved in our study. Implementing the ISC, it is advised that the key value trigger of a data project is outlined up-front. You do not need to exhaustively define all the benefits before you start planning but you do need a value trigger to justify the planning effort. During the planning process a more complete picture of the benefits can be detailed.

In line with Ward et al. (2008) the ISC prompts the detailing of benefits with corresponding measures and owners (see Table 7). In our research, a project focusing on visualizing key data points for individual customers was triggered by the need to increase the sales/value of their current customers. However, once a detailed analysis was completed, 24 benefits in total were outlined for the project. These benefits included: time (reduction in the amount of time needed for customer research by sales reps), process effectiveness (increased ability to priorities potential leads), and financial (increase of net new sales pipeline). As with this project, the key goal of good business cases is that the more benefits identified the more complete a picture of the potential value is created. In addition, to ensure the value focus is aligned with the strategic objectives of the organization, the type of project can be defined under the headings: (i) transformational, (ii) renewal, (iii) experimental, or (iv) process improvement. Having this classification will ensure a strategic input into the planning process and also help with the prioritization of such projects (Ross and Beath 2002).

Benefit	What is the benefit? This describes the benefit that will accrue from the implementation of the data project.
Measure	How is the benefit measured? This will provide a method and metric which will determine the successful attainment of the benefit.
Owner	Who is responsible for the benefit? This identifies the person responsible for driving the benefit during and after the project implementation.

Table 7: The Value Card.

Case Study – Applying the New Planning Approach

To illustrate the application of the ISC, Figure 2 visualizes all the data nodes with their associated ISC cards outlined in Appendix Tables A-1 to A-6. Within each of the nodes are numbers which link to the corresponding ISC cards that provide the detail around the abstracted data node. The case on which the approach is based is a global provider of data services to the travel industry. The high-level organizational impact of using the ISC was that it “built an awareness of data” in a very short timeframe. It was even noted that it became much easier to get people to attend meetings when the ISC was the focus of the meeting.

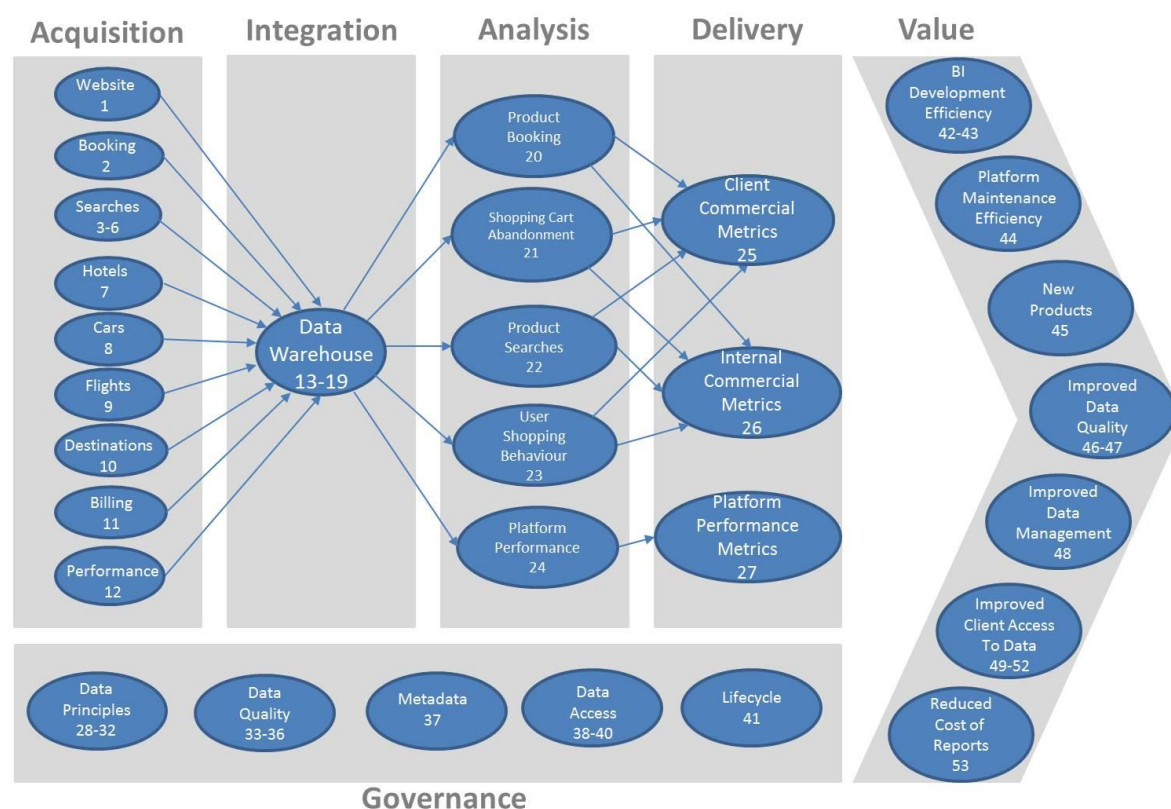


Figure 2: Working example of the ISC applied to a data services company

In more detail, the organization used the ISC to detail plans for six initiatives derived from their data strategy, which focused on maturing their data service products. This strategy was the accumulation of a two-year process from which the value of the organizations data was fully examined. However, the strategy never gained much traction until the introduction of the ISC. The ISC below (Figure 2) details the fourth plan, which was triggered by the identified need to reduce the maintenance costs of their data platform. However, during the use of the ISC several other potential benefits were outlined for their customers as well as the company itself. Delivery of the data required the development of an internal and external product dashboard as well as an internal only platform performance dashboard. At this stage of the strategy, descriptive analytics were enough to fulfil their needs. However, plans for later stages of the strategy included prescriptive analytics. While a key focus of the data project was to integrate all data sources to a single data warehouse, the investigation of the acquisition phase highlighted valuable data sources that were not previously used. The examination also prompted the

control of the data integration function to be transferred to the business stakeholders not the IT department. This increased the priority and value of the data itself above aspects such as data access security; a key IT function. Finally, the governance section outlined a number of practices that needed to be implemented but also initiated the development of automated alerts for increasing the quality of data.

Contributions

As with the organization in question the projects in the study demonstrated evidence of the intrinsic benefits from the planning process as well as the extrinsic benefit of the plan itself. Moreover, these benefits point to the effectiveness of the ISC as an effective sensemaking device for new data projects/technologies during the early stages of organizational innovation through data namely: *assimilation* – developing an organising vision for the new data project/technology, and *adoption* – developing a clear understanding to why the data project/technology should be implemented (Swanson and Ramiller 2004).

These benefits include:

Ability to learn from experimentation and iteration – one of the key intrinsic benefits of going through the ISC planning process is that it will take a number of iterations to reach a level of shared understanding and abstraction that will enable it to become a useful artefact. This is part of the process but enables the users of the approach to learn from each iteration. From the case study, each version of the ISC uncovered aspects of the project that needed to be further discussed. Be it the difficulty of sourcing data or new ideas on how the data project could be governed after implementation.

Balance complexity and simplicity – for some the completion of the ISC was the first time they attempted to visualize a data project/process. The complex and interconnected nature of data, processes and technology make the mapping and communication of such complexity a real challenge. The ability to abstract this complexity without comprising the reality of the analysis is a key affordance of the ISC approach that is also embodied in the output – the completed ISC. In addition, evidence from the case-study highlighted the ease that the ISC could be comprehended. One senior manager from a project management background noted in the case-study, “I don’t understand technical diagrams but I understand this...it is something we can talk about in meetings”.

A better understanding of how the organization operates – through the process of exploring the ISC and engagement with different stakeholders, the parties involved in completing the approach developed a better insight into their own organization and as a result a better chance of an aligned project. During the case study, the ISC was used to present the project and overarching strategy to an international office where communication was noted as already being difficult due to language and cultural issues. However, using the ISC the international office immediately understood the strategy, especially as value was the central focus of the presentation.

Access to Expertise – in completing the ISC, it becomes evident who are experts in the business and technological aspects of the data flow. These are the people that bring additional clarity to the use of data. In addition, the ISC itself demonstrated the ability to be an approach for building shared understanding across functional boundaries and across levels of management. In the case study, the ISC quickly highlighted gaps in the skills of the organization to fully implement the project. Primarily this skills shortage was identified in the analysis aspect of the data project and longer term data strategy.

Ability to challenge the status quo and call out problems/opportunities – with human centered design being a core part of the ISC, the challenging of the status quo and root cause analysis is naturally incorporated in the prescribed planning process. However, the ISC itself has been demonstrated as a communication tool to highlight key opportunities and the potential value associated with them. On the flipside it can also be utilized to depict problems and the value that can be derived from solving those problems. In the case study the ISC was itself a challenge to the way the organization viewed their data services. In particular, it challenged the cost avoidance perspective held by the organization and ultimately provided a more positive revenue generation potential that could be achieved.

References

- Agarwal, R., Sambamurthy, V., and C.V., B. 2009. "Editors' comments-Special issue on IT-business alignment," *MIS Quarterly Executive* (8:1).
- Brown, T. 2008. "Design thinking," *Harvard business review* (86:6), p 84.

- Davenport, T. H., and Harris, J. G. 2007. *Competing on analytics: The new science of winning*, (Harvard Business Press).
- Davenport, T. H., Harris, J. G., and Morison, R. 2010. *Analytics at work: Smarter decisions, better results*, (Harvard Business Press).
- Information Building (2014), "Driving Better Business Performance With a Practical Data Strategy", White Paper, pp. 1-14.
- Keen, P. G. W. 1981. "Information systems and organizational change," *Communications of the ACM* (24:1), pp 24-33.
- Keen, P. G. W. 1993. "Information Technology and the Management Difference - a Fusion Map," *IBM Systems Journal* (32:1), pp 17-39.
- Khatri, V., and Brown, C. V. 2010. "Designing data governance," *Communications of the ACM* (53:1), pp 148-152.
- Kruchten, P., Nord, R. L., and Ozkaya, I. 2012. "Technical debt: from metaphor to theory and practice," *IEEE Software* (0:6), pp 18-21.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., and Kruschwitz, N. 2011. "Big data, analytics and the path from insights to value," *Mit Sloan Management Review* (52:2), pp 21-31.
- McLuhan, M., and Fiore, Q. 1967. "The medium is the message," *New York* (123), pp 126-128.
- Mitra, S., Sambamurthy, V., and Westerman, G. 2011. "Measuring IT Performance and Communicating Value," *MIS Quarterly Executive* (10:1) March, pp 47-59.
- Nelson, K. M., and Coopridge, J. G. 1996. "The contribution of shared knowledge to IS group performance," *MIS Quarterly* (20:4) Dec, pp 409-432.
- Redman, T. C. 2008. *Data driven: profiting from your most important business asset*, (Harvard Business Press).
- Redman, T. C. 2013. "Data's credibility problem," *Harvard Business Review* (91:12), pp 84-88.
- Ross, J. W., and Beath, C. M. 2002. "Beyond the business case: New approaches to IT investment," *MIT Sloan Management Review* (43:2) Win, pp 51-+.
- Swanson, E. B., and Ramiller, N. C. 2004. "Innovating mindfully with information technology," *MIS Quarterly* (28:4), pp 553-583.
- Tallon, P. P. 2013. "Corporate governance of big data: Perspectives on value, risk, and cost," *Computer* (46:6), pp 32-38.
- Tsoukas, H. and Chia, R. (2002), "On organizational becoming: Rethinking organizational change", *Organization Science*, Vol. 13 No. 5, pp. 567-582.
- Ward, J., Daniel, E., and Peppard, J. 2008. "Building better business cases for IT investments," *MIS Quarterly Executive* (7:1) March, pp 1-15.

Appendix 1 – ISC Case Study Cards Detail

Card No.	Data (Acquisition)			People		Technology		Process
	What data is acquired?	Why is it acquired?	When is it acquired?	Who is it acquired from?	Who acquires it?	Where is it acquired from?	Where is it stored?	How is it acquired?
1	Web Traffic data - Audience, Acquisition, Behavior and Conversion data	To record user behavior on the site, how many came onto the site, what they did when they were on the site, where they came from, where they left the site, how they accessed the site etc.	Daily - Real time	Online users	Project Development Team	Web Analytics Platform - Google Analytics	Google Analytics Platform	Automated - Google Analytics content tracking enabled on the platform
2	Product Booking details (for all online products - Flight, Hotel, Car, Insurance)	To record product bookings made by users on the site and to enable users to manage and service those bookings (update, cancel etc.)	Daily - Real time	Online users & Client call centre staff	Project Development Team	Client Booking Website and Call Centre Module	Call Centre System Database	Automated - Booking Platform Submit process
3	Product Search details (for all online products - Flight, Hotel, Car, Insurance)	To record the intention and interest of online users (what specific products, destinations, time periods etc. that they are interested in; travel party details - family, couples, single etc.)	Daily - Real time	Online users & Client call centre staff	Project Development Team	Client Booking Website and Call Centre Module	Search Data Database	Automated - Search data logging process
4	Product Search Results details (for all online products - Flight, Hotel, Car, Insurance)	To record the options returned to the user from the search - products, product levels, prices etc. Enables the company to analyse the suitability of the options returned.	Daily - Real time	Product Suppliers and Content providers	Project Development Team	Suppliers content systems and client booking platform	Search Data Database	Automated - Search data logging process
5	Product Search Items viewed by users (Product search items returned to and viewed by the user)	The record the product items that users were interested in i.e. those items that they had reviewed (product types, levels, pricing, locations, search result page/position). Enables analysis of users selection criteria and preferences.	Daily - Real time	Online users	Project Development Team	Client Booking Website	Search Data Database	Automated - Product Item Review data logging process

Table A- 1: Sample of Acquisition Cards for Worked Example.

Data (Integration)				People	Technology	Process
Card No.	What data is integrated?	Why is it integrated?	When is it integrated?	Who integrates it?	Where is it integrated?	How is it integrated?
13	Product Booking Data	To analyze specific details of products booked, modified and cancelled on the platform.	Daily	Data Team	Single tenant data warehouse	ETL Process
14	Product Search Data	To analyze specific details of the products searched for and reviewed by the user	Daily	Data Team	Single tenant data warehouse	ETL Process
15	Destination Data	To determine the performance of specific locations in terms of both user bookings and searches	Daily	Data Team	Single tenant data warehouse	ETL Process
16	Product Content Data	To analyze and to deep-dive into the specifics of the products which may/may not be performing well. E.g. most/least booked products, the products most viewed but never booked etc. Determine what the product long-tail is - for each product type etc.	Daily	Data Team	Single tenant data warehouse	ETL Process
17	Shopping Cart Abandonment Data	To analyze the trends and details of shopping cart abandonment (abandonment stage, products selected, user issues encountered e.g. technical, navigational, payment, supplier)	Daily	Data Team	Single tenant data warehouse	ETL Process
18	Performance Data	To analyze booking platform performance and to identify/investigate any performance issues/trends encountered. Also to establish potential impact on performance from factors such as web traffic, user behavior, supplier type and process type.	Daily	Data Team	Single tenant data warehouse	ETL Process
19	Web Traffic Data	To analyze user behavior and activity on the booking platform	Daily	Data Team	Single tenant data warehouse	ETL Process

Table A- 2: Integration Cards for worked example.

Data (Analysis)				People	Technology	Process
Card No	What data is analyzed?	Why is it analyzed?	When is it analyzed?	Who analyses it?	Where is the data analyzed?	How is the data analyzed?
20	Product Booking Data	To give insights into the channels, markets, products, booking-personas, product mix, location, duration, time-periods and price ranges for bookings on the platform.	Daily	Data Team	BI Platform - Tableau/Microstrategy	Descriptive analysis
21	Shopping Cart Abandonment Data	To give insights into the user transactions that did not complete (i.e. abandoned carts) and to help determine the reasons why. Establish any commonalities in the factors associated with the issues and help identify solutions to reduce this in the future.	Daily	Data Team	BI Platform - Tableau/Microstrategy	Descriptive analysis

22	Product Search Data	To give insights into the channels, markets, products, search-personas, product mix, location, duration, time-periods, and pricing that users are searching for.	Daily	Data Team	BI Platform - Tableau/Microstrategy	Descriptive analysis
23	User Shopping Behaviour Data	To identify how the users use and navigate through the booking platform, searches per session, travel detail changes, exit and entry points, activity periods, navigation and flow preferences.	Daily	Data Team	BI Platform - Tableau/Microstrategy Google Analytics	Descriptive analysis
24	Performance Data	To identify any platform bottlenecks or issues. Also to analyze any identifiable factors that may be contributing to performance issues.	Daily	Data Team	BI Platform - Tableau/Microstrategy	Descriptive analysis

Table A- 3: Analysis Cards for Worked Example.

Data (Delivery)			People	Technology	Process	
Card No.	What data is delivered?	Why is it delivered?	When is it delivered?	Who receives the data delivery?	Where is the data delivered?	How is the data delivered?
25	Client Commercial Metrics, Client specified KPI's	So that clients can use the data to review commercial performance, optimize their selling proposition and product offerings, determine whether existing promotions and marketing programs are working, what new promotions they should consider and other areas to focus on to improve conversion rates etc.	Daily	External Clients	BI Server, BI Platform	Drill-down dashboard
26	Internal Commercial Metrics, Internal specified KPI's	So that the business can use the data to analyze commercial performance on the platform across client accounts, identify areas for improvement, make recommendations to improve conversion rates and measure impact of said recommendations.	Daily	Internal Business Customers - Project Managers, Department heads, Senior Mgmt.	BI Server, BI Platform	Drill-down dashboard
27	Platform Performance Metrics & KPI's	So that the technical managers can use the data to optimize the platforms performance, highlight areas of concern and understand the factors and circumstances that impact performance.	Daily	Internal Technical Customers - System Architects, Senior Technical Managers	BI Server, BI Platform	Drill-down dashboard

Table A- 4: Delivery Cards for Worked Example.

Card No.	Category	Behavior	Benchmark	Owner
28	Data Principles	Awareness Building - Communication from senior management on the importance of data to the organization and the priority that should be afforded to data related elements in each project.	Communication plan developed and rolled out. Project Managers include data related tasks in their project planning and backlogs.	Senior Management team Project Management Office
30	Data Principles	Develop a standard process for data and supplier content loading to ensure consistency across projects.	All project data and supplier content loading tasks adhere to the process.	Technical Managers
36	Data Quality	Automated checks are developed to ensure accuracy of the data that is processed in the Data Warehouse and to ensure consistency with the production data.	Automated checks developed and results delivered to the stakeholders each morning - identifying any issues encountered.	Data Manager
37	Metadata	Develop data dictionaries to ensure an understanding of the data across projects and teams.	Data dictionaries are developed and maintained for each data category.	Data Manager
38	Data Access	Client Production Data storage is governed under a defined set of security policies as agreed with the client security /audit team	0 security breaches	IT Operations, Infrastructure Manager
41	Data Lifecycle	A formal plan is created for Data lifecycles and a review of data is undertaken based on the plan. Also, the team responsible for maintaining this is involved in its creation and sign off on same.	Plan is created and signed-off. Adherence to the plan. Ongoing review and check-point scheduled created.	IT Operations, Infrastructure Manager and Data Manager

Table A- 5: Sample of Governance Cards for Worked Example.

Card No.	Benefit	Measure	Owner
42	Reduced BI Development Costs	60% reduction in man days	Data Manager
43	Reduced Senior Engineer time on BI related tasks	50% reduction in senior engineers time on BI related tasks	Data Manager
44	Reduced Maintenance Development Time	20% reduction in man days	System Technical Manager
45	Increased revenue opportunities through the introduction of new data products & services	2 new/revised data products or services introduced per annum	Sales Manager

Table A- 6: Sample of Value Cards for Worked Example