A review and future direction of agile, business intelligence, analytics and data science

Deanne Larson [a], Victor Chang [b],*

[a] Larson & Associates, LLC, USA
[b] Xi'an Jiaotong Liverpool University, Suzhou, China

ARTICLE INFO

ABSTRACT

Agile methodologies were introduced in 2001. Since this time, practitioners have applied Agile methodologies to many delivery disciplines. This article explores the application of Agile methodologies and principles to business intelligence delivery and how Agile has changed with the evolution of business intelligence. Business intelligence has evolved because the amount of data generated through the internet and smart devices has grown exponentially altering how organizations and individuals use information. The practice of business intelligence delivery with an Agile methodology has matured; however, business intelligence has evolved altering the use of Agile principles and practices. The Big Data phenomenon, the volume, variety, and velocity of data, has impacted business intelligence and the use of information. New trends such as fast analytics and data science have emerged as part of business intelligence. This paper addresses how Agile principles and practices have evolved with business intelligence, as well as its challenges and future directions.

1. Introduction

The manifesto and principles for Agile Software Development (ASD) were published in 2001, and since then, the objectives and principles have been interpreted and applied to Business Intelligence (BI). The application to BI is natural, because of the iterative and incremental nature of BI development. The intent of this article is to provide practitioners an understanding of how the Agile principles can be applied to BI delivery, fast analytics, and data science. Beck et al. (2001) outlined the core ideals of the manifesto: individuals and interactions over processes and tools; working software over comprehensive documentation; customer collaboration over contract negotiation; and responding to change over following a plan. The result of following these ideals, software development becomes less formal, more dynamic, and customer focused.

Information Technology (IT) departments are faced with maintaining a competitive edge, which, in turn, increases pressure to deliver high quality technology solutions faster. Under these circumstances, the value of technology efforts are determined based on how soon payback and return on investment occur. BI initiatives require significant upfront and ongoing investment to maintain value, inviting constant scrutiny on whether business value occurs. Measuring BI value continues to be a struggle for organizations, mainly due to the challenge of directly attributing return to the investment in BI. BI plays the role of an enabler – enabling the organization to become smarter, work smarter, and make better decisions through the use of information. The enabler role makes it difficult to directly attribute a return on investment and over time, the use of information becomes routine and expected.

The information value chain is the process used to derive value from information and information from data; BI delivery is centered on the information value chain. Collecting raw data is the first step in the value chain; applying logic and business context to the data creates information; information is then consumed by BI users; decisions and actions are a result of the consumption of data; resulting in decisions and actions that provide business value.

The information value chain is an important concept in understanding the benefits of Agile principles applied to BI delivery. BI delivery is not accomplished via traditional waterfall software development (although some organizations attempt this); it is more focused on data discovery and understanding how information is going to be used. This perspective drives how Agile principles should be applied to BI delivery – less focus on software development and more focus on information use. The need to delivery faster has increased over the last 5 years due to the demand of real-time data analysis (Halpern, 2015). The Internet of Things (IoT),

* Corresponding author.
  E-mail addresses: larsonelink@aol.com (D. Larson), ic.victor.chang@gmail.com (V. Chang).

where data collection is embedded into devices, contributes to this demand for fresher data. Monitoring equipment failures, for example, will be possible with data that is seconds old versus data that is hours or days old (Halpern, 2015).

The objectives of this article are fourfold. First, revisit the alignment between Agile principles and BI delivery, fast analytics, and data science. Second, analyze Agile methodologies and how they have been applied with BI and are emerging with Big Data. Third, review the components and best practices of Agile BI delivery considering the impact of Big Data. Last, propose an Agile framework for BI delivery, fast analytics, and data science; fast analytics and data science are the emerging data analysis trends due to Big Data (Fig. 1).

2. Background

Business Intelligence (BI) is defined by literature and scholars in similar ways. Noble (2006) defines BI as the ability to provide the business an information advantage; business doing what it has always done, but more efficient. Singer (2001) described BI as the value proposition that helps organizations tap into decision-making information that regular reporting does not provide. Singer outlined that BI requires tools, applications, and technologies focused on enhanced decision-making and is commonly used in supply chain, sales, finance, and marketing. Negash and Gray (2008) outlined BI more comprehensively. BI is a data driven process that combines data storage and gathering with knowledge management to provide input into the business decision making process. BI enables organizations to enhance the decision making process and requires processes, skills, technology, and data. More recently, Gartner (2013) and Halpern (2015) have extended BI to be an umbrella term which includes applications, tools, infrastructure, and practices to enable access and analysis of information to optimize performance and decision-making.

The challenges In BI delivery include business and IT collaboration that results in data become information. Delivery of BI is accomplished via a methodology. Creswell (2005) outlined that a methodology is set of processes, methods, and rules applied within a discipline. Successful BI methodology should focus on the information value chain and less on the development of software as is the focus of traditional information technology (IT) development. Research has demonstrated that waterfall lifecycles and traditional software development practices are not successful in BI. Software and hardware do not provide organizations value pertaining to BI; it is the use of information (Larson, 2009).

Common stumbling blocks traditionally experienced in BI projects included: fuzzy requirements; lacking an understanding about how data is created and used; data quality is not measured or known; source system constraints dictate design and service levels; developing based on perceptions of data; results are not demonstrated in a timely manner; and working with a lack of trust between IT and business stakeholders (TDWI, 2009). While these challenges still remain, the need to have information sooner has been influenced by the phenomenon of "Big Data". Big Data is a broad term used to describe data sets that are large, complex, and cannot be addressed by traditional IT methodologies and applications (Davenport, 2013).

2.1. Big data

Traditional data processing has transformed because how data is generated today has transformed (Davenport, 2013). Historically IT departments managed transaction processing systems. Business was about transactions – orders, sales, shipments, inventory, and accounting, to list a few examples. Transactional data is structured, stable, and understood by organizations. Structured data is formatted in rows and columns. Transactional data is primarily used for decision support. The differentiating points between transactional data and Big Data are volume, variety, and velocity. Volume refers to the amount of data, variety is based on the types of data sources, and the velocity represents the age of data. Volume, variety, and velocity are referred to as the "3 V's" (Davenport, 2014). Transactional data is structured, well understood, and volume is tens of terabytes or less (Davenport, 2014). Transactional data is used in decision support analysis. Decision support analysis focuses on "what has happened" or retrospective analysis.

Volume for Big Data is measured in excess of 100 TB or petabytes (although research in this threshold varies). Big Data is characterized by data types considered unstructured – not predefined or known – thus a high degree of variety. Velocity of Big Data can be measured in less than a few minutes. Big Data is used in machine learning and predictive analysis where organizations focus on "what will happen" versus "what has happened" (Davenport, 2014). The emergence of Big Data has changed the face of BI.

2.2. Analytics

Analytics have been around since the 1950s and are not a new idea (Davenport, 2014). Analytics started with a limited number of data sources which came from internal systems and data was stored in a repository such as a data mart or data warehouse—defined as traditional BI. Most analysis was descriptive and BI consisted primarily of reporting. In 2003, Big Data started emerging when high technology firms such as Google and Yahoo began using Big Data for internal analytics and customer-focused processes. The velocity of Big Data changed traditional BI as data had to be stored and processed quickly. Predictive and prescriptive analytics started to emerge, but visual analytics of descriptive data was still the prominent form of analytics (Davenport, 2014). Big Data became "bigger" – more volume, variety, and velocity – and organizations began focusing on the data-driven economy. Organizations of all types are developing data-based offerings to remain competitive (Davenport, 2014).

With the advent of Big Data and analytics evolving, BI delivery has been impacted. Data has to be turned into information quickly for analysis. Organizations are focusing more on prescriptive and predictive analysis that utilize machine learning as well as fast analytics through visualization. Fast analytics refers to the ability to acquire and visualize data quickly (Halpern, 2015; Jarr, 2015). The velocity increase in data has accelerated the need for IT departments to acquire data and transform it into information. Table 1 illustrates the characteristic differences between traditional BI and fast analytics with Big Data.

3. Application of agile to BI and Big Data

Agile ideals and principles were published by Beck et al. (2001) and since this time, practitioners have focused on applying an Agile approach to BI. The challenges that BI projects face make the Agile approach an attractive answer due to the parallels that exist between them. By using an Agile approach, means the methodology is less formal, more dynamic, and customer focused. The dynamics required in BI delivery make Agile approach a good fit with BI; however, practice with Agile methodologies have resulted in adjustments to the well-known Agile methodologies to focus on the utility of information versus primarily software development. Agile methodologies align well also with Big Data where little time is spent defining requirements up front and the emphasis is on developing small projects quickly. Agile methodologies will align well
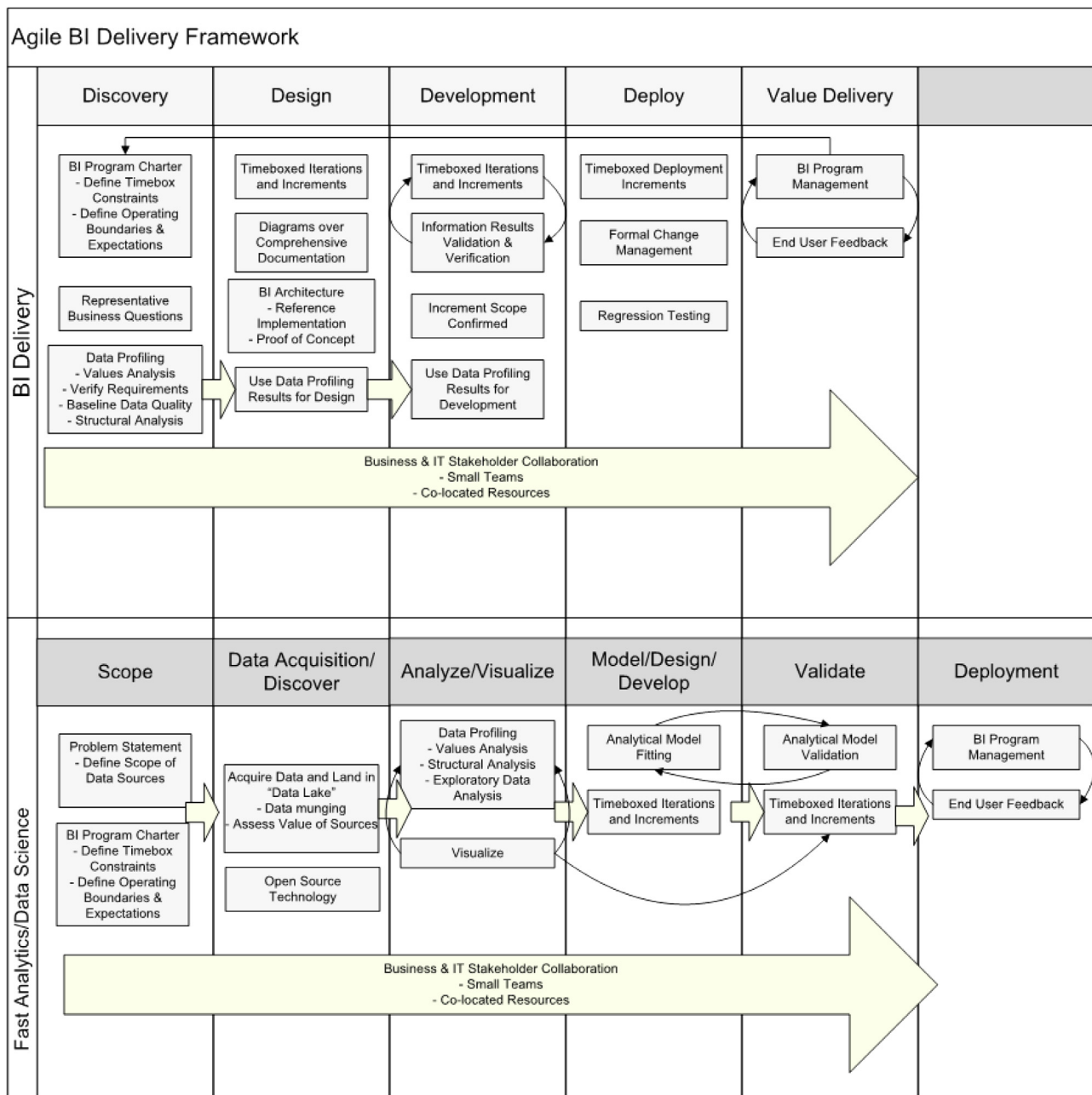
Fig. 1. Proposed agile delivery framework.

with iterative discovery and validation which support prescriptive and predictive analytics (Ambler & Lines, 2016).

3.1. Agile principles

Analyzing the Agile principles provides an understanding of how using an Agile approach matches well with BI delivery. To reiterate the principles: individuals and interactions over processes and tools; working software over comprehensive documentation; customer collaboration over contract negotiation; and responding to change over following a plan. Beck et al. (2001) outlined that an Agile approach focuses more on the left side of the principle; however, the right side is not ignored.

3.2. Individuals and interactions over processes and tools

Experienced individuals working together are more effective and build better systems than less experienced individuals using structured process and tools (Ambler, 2006). With BI, the system includes multiple components such as source systems, Extract, Transformation, and Load (ETL) processes, databases, and front-end tools. The infrastructure of a BI system is the enabler to gaining value from organizational data. BI is less about the process and tools and more about the utility of information. Although the ideal discussed here emphasizes individuals and interactions over processes and tools, processes and tools are not eliminated from an Agile approach.

Faster analytics that have emerged with Big Data are needed to align with faster technology. Analysis that may take hours or days

with transactional data and descriptive approaches are done in seconds with Big Data technology. New roles such as the data scientist have emerged to use machine learning algorithms. Machine learning processes rely on processes like Agile where frequent delivery of information to stakeholders are required (Davenport, 2014). The challenge is adapting Agile approaches to match the speed of the new technologies. The process to deliver faster analytics also emphasizes the Agile principle of individuals and interactions over process and tools. Faster analytics and machine learning focus on discovery and iteration up front which means more interaction and collaboration is required for discovery and insight.

3.3. Working software over comprehensive documentation

Documentation is valuable; however, the value is not the issue. Documentation has an inherent problem – usability. Documentation has been a dreaded aspect of traditional development methodologies. Documentation takes too much time to complete, tends to be out-of-date, and is rarely used after the initial deployment. Creating comprehensive documentation does not allow for quick delivery; however, not producing documentation can be more detrimental. For Agile, documentation needs be usable and add value. Documentation should less textual and more visible. Development artifacts in BI such as source to target mappings, diagrams, and models, are examples of valuable artifacts that are easy to use and maintain. Diagrams can provide a level of documentation that is adequate to support requirements, design, and development and are easy to maintain. A picture is worth a thousand words (Larson, 2009).

Documentation with faster analytics and Big Data aligns with Agile principles and comprehensive documentation is not the norm. Documentation is a lower priority and the priority with faster analytics is insight from information. Due to the speed of data availability, documentation is primarily ignored (Adamson, 2015).

3.4. Customer collaboration over contract negotiation

Practicing ongoing collaboration throughout any process adds value – communication is increased, expectations are consistently reaffirmed, and ownership of the end product is shared. Collaboration is emphasized in "interaction and individuals over process and tools" and fundamental to the success of Agile. Without pre-determined expectations, contracts can frame expectations but allow refinement and change. The details surrounding requirements are not often known in enough detail to document. Collaboration between stakeholders addresses this via delivery by determining what the expectations are and increasing communication between stakeholders (Larson, 2009).

Faster analytics places an urgency on collaboration between technical and customer resources. To facilitate discovery, technical resources such as the data scientist and developers work together to determine the data sources to work with. As a data scientist processes data and produces interim results, constant interaction with stakeholders is needed to validate results and direction. Visual analytics used by analysts requires an understanding of multiple data sources in different formats to produce charts and diagrams. Collaboration continues to be a priority in BI whether focusing on descriptive analysis or faster analytics (Davenport, 2014; Schutt & O'Neil, 2013).

3.5. Responding to change over following a plan

A change in project requirements means a change in scope, which impacts time, resources, and budget, the foundational aspects of project management. The traditional approach to managing a project is to follow the plan and discourage change. Change in

traditional methodology is the exception and not the rule (Larson, 2009). One of the objectives of Agile principles is removing bureaucracy from delivery of working software. With Agile, the approach is to be prepared for change and respond accordingly.

Change is inherent in BI and the use of Big Data. The primary resource used in both is data. Data is organic and constantly changing and with Big Data sources, even more so as data is unstructured and undefined. Responding to change continues to be a need in BI and Big Data development processes (Davenport, 2014).

3.6. Agile methodologies

The manifesto and principles for Agile Software Development (ASD) were published in 2001, and since then, the objectives and principles have been interpreted and applied to new Agile methodologies. The popular approaches from which the manifesto and principles were derived – Extreme Programming (XP) and Scrum – are in practice today with success and are considered standard development methodologies (Hsieh & Chen, 2015). Agile principles have been applied to other disciplines such as project management with success (Kaleshovska, Josimovski, Pulevska-Ivanovska, Postolov, & Janevski, 2015). Success with Agile methodologies include reduced cycle time, higher quality, increased requirements clarify, increased flexibility, and a higher overall stakeholder satisfaction rate when compared to similar projects using different project or software development methodologies (Hsieh & Chen, 2015; Kaleshovska et al., 2015). The core practices of Agile methodologies include: small, short releases; stakeholders physically located together; and a time-boxed project cycle (typically 60–90 days, although the cycle may be shorter depending on the deliverable) (Kendall & Kendall, 2004). These practices continue to contribute to the success of Agile projects (Hsieh & Chen, 2015; Kaleshovska et al., 2015).

3.7. Agile and business intelligence

The primary goal of a BI project is to enable the use of information. If the primary goal of BI is enabling the use of information, then scope of the BI project focuses on turning data into information. Software development is part of the data to information process; however, software development in BI is less about creating a working program and more about application of business context to data. Software used in BI includes database management systems, data cleansing, data transformation, and analytical systems. The scope of development in BI includes more configuration and application of logic versus programmatic coding. In order to understand how to apply logic and configure the software, IT will need to comprehend the business use of data (Larson, 2009). Big Data technology has includes the scope of software and hardware used in BI (Davenport, 2014).

BI delivery tends to be a process where customer expectations are a cycle of discovery and refinement, hence the problem of fuzzy requirements. Turning data into information is not a simple process nor are requirements easy to determine even with the use of subject matter experts. BI begins with some key questions: What business questions need to be answered? What data sources qualify as the system of record? How will data be used? These questions are addressed through a discovery process that examines how data is created and how data is becomes information. BI systems include multiple components such as source systems, ETL, databases, and front-end tools. The infrastructure of a BI system is the enabler to gaining value from organizational data. "Individuals over interactions over processes and tools" support discovery (Larson, 2009). The real requirements are discovered through the sharing of knowledge versus relying solely on stakeholders' experience to define requirements (Larson, 2009). Collaboration is a success require-

Table 1
Comparison of traditional business intelligence systems and fast analytics with big data.

Criteria	Traditional Business Intelligence	Fast Analytics with Big Data
Analytics Type	Descriptive, Predictive	Predictive, Prescriptive
Analytics Objective	Decision Support, Performance Management	Drive the Business
Data Type	Structured and Defined	Unstructured, Undefined
Data Age	>24 h	<Min

ment for implementing BI which is emphasized in the ideal of "interaction and individuals over process and tools".

As mentioned prior, BI requires a discovery process where customer expectations are determined. Without pre-determined expectations, using contracts in BI would be challenging. BI projects need a framework of expectations which allow refinement and change. The objective is to focus more on collaboration versus spending time completing a detailed plan. Detailed plans are often difficult to create since only high-level planning information is known. Collaboration helps resolve this through determining what the expectations are and increasing communication between stakeholders. The Agile principles of "customer collaboration over contract negotiation" and "responding to change over following a plan" address challenges of BI systems. Fast analytics projects align with these same Agile principles as the initial scope of these projects is quick data acquisition to be used for discovery; the value of information is yet to be determined (Davenport, 2014).

The most popular Agile development methodologies for BI are Agile Data Warehousing, Extreme Scoping, and Scrum (Muntean & Surcel, 2013). Agile Data Warehousing and Extreme Scoping work well with BI if a data warehouse is involved. Data warehouses are central to a BI architecture providing a central repository with integrated data for analysis.

Extreme Scoping is an Agile methodology specifically focused on the data integration aspect of BI projects. Data integration, the acquisition and transformation of data sources into the data warehouse, primarily focuses on data management activities. The focus of development activities is acquiring and understanding data from data sources, data cleansing, data modeling, and preparing the data for loading. Extreme Scoping is "data-centric" and includes all data management activities (Powell, 2014). Extreme scoping is broken down into seven steps where the first step identifies data deliverables, the second step breaks the deliverables into small releases, and in the remaining steps focus on the business value of the deliverables, the effort, identifying new technology, and defining the various data development tracks and resources to complete the work (Powell, 2014). Small releases then are planned to deliver the project.

Agile Data Warehousing is a broader category where several similar methodologies are present. One example methodology contains and end-to-end approach to Agile data warehouse delivery (Hughes, 2013) and another contains Agile approaches for every aspect of BI delivery including a specific Agile approach for data modeling (Amble & Lines, 2016). The commonalities of these methodologies include architecture vision, model and prototype the data throughout the project, expect multiple definitions for data, organize work by requirements, and stakeholder involvement (Hughes, 2013; Ambler & Lines, 2016). Agile Data Warehousing uses many of the same concepts as Scrum for BI project delivery.

Scrum, while not limited to BI, is the most popular Agile approach used in Agile software development and BI (Muntean & Surcel, 2013). The concepts of Scrum primarily used in the BI are the user story, sprint backlog, product backlog, the sprint, and the daily scrum. BI requirements are divided into small stories which then are packaged into a collection of stories to comprise a BI project. Each story is designed, developed, tested, and released. A sprint lasts from one to two weeks and contains a cycle requirements,

analysis, design, development, and end user testing. Stories can be either grouped into product or sprint backlog. Sprint backlog refers to the work the development team completes during the spring. Product backlog is a list of all stories ordered by priority to be considered for the next spring. Users are involved in all of the sprint steps. Daily meetings which are less than 15 min are held to review status (Muntean & Surcel, 2013).

While recent research shows that applying Agile methodologies to BI projects can increase quality, reduce cycle time, and improve stakeholder satisfaction, 60–70% of BI projects (including Big Data) still fail (Gartner, 2015). While some success of Agile has increased its adoption in BI, Agile has become more of a focus with Big Data and the increased focus on analytics.

3.8. Agile and Big Data

Big Data is a phrase coined to describe the changing technology landscape that resulted in large volumes of data, continuous flow of data, multiple data sources, and multiple data formats (Davenport, 2014). The input to a BI project is data and the output is information. With the data landscape changing so rapidly, BI projects and the methodologies used are also changing.

Big Data is a fairly new phenomenon, thus research is limited. Some research suggest Big Data is a media term or vendor term used to describe hardware, software, and analysis connected with non-traditional data sources (Davenport, 2014). Several analysis trends have emerged (in some cases re-emerged) such as predictive and prescriptive analysis. Data science, which is defined as an interdisciplinary field focusing on extracting insights from data in various forms, is the next generation of data analysis fields such as statistics and data mining. With Data Science, analysis is completed using statistics and machine learning algorithms to produce data products or models that perform descriptive, predictive, or prescriptive analysis (Schutt & O'Neil, 2013). As the 3 V's describing Big Data suggest, data freshness, data variety, and the speed of analysis have changed how BI projects have traditionally been approached.

Fast analytics and data science principles exasperate some of the Agile principles to the point where the results of an Agile methodology need to be produced as if "on steroids". Fast analytics and data science use unstructured data that is acquired quickly and stored for analysis, eliminating the traditional steps required for design. Discovery, completed as part of design and development, is moved to the front of the development cycle where data analysis starts as soon as data is acquired. Visualization of data occurs interactively and iteratively to support the discovery of insights. Little research exists on application of Agile principles in this manner; however, the research available suggests that Agile would align well but would need to be "short-cycle Agile" suggesting faster results are needed (Davenport, 2014).

4. Business intelligence delivery, analytics, and Big Data

4.1. Goals of BI delivery

Yeoh and Koronios (2010) posited that a BI system is not a conventional IT system (i.e. transactional system); however, BI systems

have similar characteristics to enterprise systems or infrastructure projects. BI system implementation is a complex activity involving hardware, software, and resources over the life of the system. The complexity of the BI system infrastructure increases with the scope. An enterprise BI system can include a data warehouse, integrated data structures, source systems, and large data volumes.

BI success focuses on a few main Critical Success Factors (CSF) categorized by organization, process, and technology (Yeoh & Koronios, 2010). Organizational CSFs consists of establishing a vision, outlining a business case for BI, and gaining leadership support for BI as a priority. Process CSFs focus on managing BI as an evolving ongoing program. Process CSFs include having a dedicated, skilled BI team for technical delivery as well as ongoing program and change management that focus on aligning BI with business goals. The technical category centers on two areas – data and infrastructure. Data and infrastructure CSFs consist of many factors related to stability and quality since these two areas are the major technical components of the BI systems (Yeoh & Koronios, 2010).

Mungree, Rudra and Morien (2013) completed a 15-year review on BI CSFs which included the research by Yeoh and Koronios (2010). Mungree et al., identified the following top ten BI CSFs: committed management; appropriate team skills, appropriate technical framework and resources; align BI projects with business strategy; clear vision and requirements; user-oriented change management; effective data management; committed executive sponsor; project scope management. The results have commonality with the research by Yeoh and Koronios and reinforce that BI projects require alignment with business strategy, collaboration with business users, require flexibility, and should be of manageable scope.

Analyzing the success factors defines the goals of BI delivery. BI delivery consists of practices, methods, skills, and competencies required to create, implement, and sustain BI systems. The success factors guide best practices in BI delivery. In simpler terms, BI delivery needs to support organic and evolutionary change, driven by the constant evaluation of information and user feedback. BI systems would be constantly optimized and improved based on an ongoing feedback loop. Kenney (2015) provided a link between CSFs and Agile methodology where several of the success factors aligned with best practices in BI Agile projects. Best practices outlined included: small projects to see value quickly, involvement of a dedicated sponsor, and close collaboration between development and user teams (Kenney, 2015; Couture, 2013).

Fast analytics and data science projects align with Agile principles in that the shorter the scope and cycle, the faster results. Little research exists at this time to highlight the results of using Agile methodologies with fast analytics and data science projects; however, some cases studies infer that short-cycle Agile approaches are used due to the success of Agile methodologies in BI (Davenport, 2014). Short-cycle Agile refers to faster and more flexible sprints.

4.2. Iteration and incremental

One of the synergies that Agile has with BI is the short, small release and experts' recommendation that BI, fast analytics, and data science projects work best delivered in increments (Davenport, 2014; Yoeh & Koronios, 2010; Mungree, Rudra, & Morien (2013)). This incremental approach supports that fact that businesses and technology change quickly and want to evaluate the impact of these changes. An incremental approach allows for management of risk, allows for more control, and enables customers to see tangible results.

Correct use of increments and iterations in BI begins with understanding that these concepts are not the same. Both concepts apply to BI delivery but in a different way. BI literature tends to use these concepts interchangeably. Iteration refers to the cyclic process of

Table 2
Comparison of the business intelligence lifecycle and fast analytics/data science project lifecycle.

Business Intelligence Lifecycle	Fast Analytics/Data Science Lifecycle
Discovery	Scope
Design	Data Acquisition/Discovery
Development	Analyze/Visualize
Test	Model/Design/Development
Deploy	Validate
Support	Deployment
Value	Support/Feedback

refinement to get to the best solution. Incremental is a staging and scheduling strategy where the scope of delivery is adjusted as necessary (Cockburn, 2008).

Increments deal with the staging and scheduling of deliverables which may occur at different rates. Iterations are cycles to revise and improve the deliverable. Increments are scheduled as part of a roadmap or release plan tied to an overall BI strategy that outlines what information capabilities are needed and when. Iterations will happen within the increment. Increments are time-boxed, therefore the results can be less or more than expected. If less than expected is delivered, increments are adjusted accordingly. Simply, increments manage the scope of the delivery and iterations are used to refine the quality of the deliverable. Deliverables can be code, models, diagrams, or any artifact created as part of the cycle.

4.3. The BI and fast analytics lifecycle

A lifecycle is the progression of something from conception to end of life or when something no longer provides value. Lifecycles have phases that comprise the progression of conception to end; the BI lifecycle is no different. The BI lifecycle parallels the SDLC with similar phases and is centered on the utility of information versus the development of software. Fast analytics and data science projects take on a different approach due to the speed of technology and the acquisition of data. In BI, discovery focuses on requirements, design on defining and working with structured data, development on creating code and databases, testing on validating development, deployment to production, then support and value measurement. With fast analytics and data science, design is encapsulated into development, because data is acquired too quickly to analyze and it is unstructured and dynamic. Fast analytics can involve iteration and visualization of data to understand and define. Data science involves iterative development of analytical models where models are created, validated, and altered until the desired results are achieved (Schutt & O'Neil, 2013). Table 2 compares the two different lifecycles.

4.4. BI lifecycle

4.4.1. Discovery

During the discovery phase, the expectations of BI projects are not initially clear to stakeholders. Business users begin with the knowledge that information and analysis capabilities are needed, and IT professionals are ready to take down requirements without a clear starting point. For these reasons, the first phase is the discovery phase where stakeholders determine information requirements. Information requirements begin with defining business questions which provide insight into data sources, dimensions, and facts needed.

4.4.2. Design

Design in BI focuses heavily on modeling, but may start with establishing the architecture of the system. Architecture in BI is more than hardware infrastructure. BI architecture includes busi-

ness, technical, process, data, and project components. BI business architecture centers on defining the drivers, goals, and strategy of the organization that drive information needs. BI project architecture describes the incremental methodology used for short, small releases. Process architecture includes the framework for data acquisition to data presentation. Data architecture address how data will be structured in data repositories, such as a data mart or warehouse. BI technology architecture includes hardware, BI software, and networks required to deliver BI projects. If the BI architecture is established, design will center on modeling data and processes to support information needs. Models created in this phase could include conceptual, logical, and physical data models as well as process models for ETL.

4.4.3. Development

BI development may include a wide array of activities. The primary focus of the development phase is to produce a working system that applies business context to data and presents information in a way that enables end users to analyze actionable information. Activities could include coding ETL, configuring logical layers in a BI tool, or scripting scheduling jobs. The scope of development can involve data acquisition to staging, staging to presentation, and presentation to the access and delivery of information.

4.4.4. Deploy

BI systems tend to be complex for many reasons. One reason for complexity is that BI systems have many independent components that require integration. Another reason is BI systems are impacted by continuous change. Because of this complexity, the BI deployment phase is formal and controlled. Activities in this phase focus on integration of new functionality and capability into production, and regression testing to verify that previously working functionality is not impacted. Deployment focuses on introduction of new components and maintaining the stability of the production BI system.

4.4.5. Value delivery

The value delivery phase includes stabilization, maintenance, change management, and end user feedback. Successful BI systems generally have a long life and require program management to address change and maintain ongoing value. Due to continuous change and the dynamic uses of information, BI system value requires constant attention. Change impacting a BI system can initiate from source systems, business processes, software upgrades, new data integration, and organizational strategy. End user feedback provides an understanding of how information is used and the overall value gained from the BI system.

4.5. Fast analytics/data science lifecycle

4.5.1. Scope

Fast analytics and data science emergence is due to Big Data; however, it is important to point out that forms of fast analytics and data science have existed for some time. Visual analytics is considered synonymous with fast analytics and data mining is also used synonymously with data science. Both fast analytics and data science are newer versions of known data analysis methods (Davenport, 2014; Jarr, 2015; Kiem, Kohlhammer, & Ellis, 2010; Schutt & O'Neil, 2013). Visual analytics and data mining became complementary techniques where visualization was used during the discovery phase of data mining (Kiem et al., 2010). As Big Data was used more in analysis, visualization tools were used to explore raw data to support exploratory data analysis in the data science process. New tools emerged in the BI industry for visualization that included new functionality such as complex graphs and charts and

the ability to connect to many different data sources (Davenport, 2014).

The scope of fast analytics and data science is to acquire data quickly to analyze. Fast analytics is more about discovery and data science uses fast analytics as part of its process. As a result of the data science process, a data product such as a prediction engine, a classifier, or recommendation engine is created (Schutt & O'Neil, 2013). The scope of fast analytics and data science will depend on the problem statement of the analysis. Many data sources could be included in the scope of analysis. Data sources may not be limited to unstructured data. Here BI program management can have value as a charter for the analytical model can define the problem statement and objectives as well as include operating boundaries and expectations.

4.5.2. Data Acquisition/Discovery

New technologies have made it possible to acquire data without a full understanding of its structure or meaning which is the opposite of what occurs in the BI lifecycle where data is profiled and analyzed to understand its meaning before loaded into a data repository for use. Hadoop or the Hadoop File System (HDFS) originated at Google and is now used in an open-source format by organizations to land data without the need for data modeling (Davenport, 2014). Analysts use fast analytics to access, assess, and visualize to discover the value and use of data sources. New data repositories such as the "Data Lake" have emerged where technology enables storage and processing power to support analyzing large unstructured data sets (Davenport, 2014).

4.5.3. Analyze/Visualize

For both fast analytics and data science, analysis and visualization are an iterative process. With fast analytics the primary goal is visual analytics to support analysis. Fast analytics can produce new knowledge that creates a refinement of the visual product. Fast analytic can iteratively produce new dashboards or scorecards to be used in ongoing BI or produce one-time analysis tools to support new knowledge gain. With data science, fast analytics and visualization is completed as part of the exploratory data analysis phase where descriptive analysis is used to highlight variable relationships and identify parameters to be used in analytical models (Schutt & O'Neil, 2013). If fast analytics and visualization produces a BI product such a dashboard or scorecard, the BI product is then validated. It is possible that fast analytics is primarily focused on discover, and a BI product is not produced.

4.5.4. Model/Design/Development

Modeling is used two ways in this phase: analytical modeling in data science and data modeling to describe data used in fast analytics. Analytical modeling include descriptive, predictive, and prescriptive analysis using machine learning algorithms such as regression, clustering, or classification (Schutt & O'Neil, 2013). In fast analytics, data is modeled after analysis to document data structures and association for future use (Adamson, 2015).

4.5.5. Validate

The validation phase is representing the data science process of validating the analytic model iteratively to the point where the error of the modeling is minimalized. This process is referred to as "fitting" the model. Additionally, fast analytics can be used to identify new parameters to incorporate into the analytical modeling process (Schutt & O'Neil, 2013). In this phase, new data sources may also be incorporated.

4.5.6. Deployment

As with BI products and systems, analytical models, dashboards, scorecards and other visualization tools have little value unless they

are used. These analytical products are added to the production environment to provide new functionality to the environment, just like the BI deployment.

4.5.7. Support/feedback

Analytical products need to be supported and revised as the organizational environment changes. The life cycle of an analytical model depends on the rate of change in the organization and the industry the organization operates within. Analytical models lose value an applicability over time and ongoing feedback from users and analysis determines how the analytical models should be adjusted.

4.5.8. Synthesis of the fast analytics/data science lifecycle and agile

Three phases of the BI Lifecycle have characteristics where using an Agile approach fits. The discovery, design, and development phases benefit from iterative cycles, stakeholder collaboration, small time-boxed increments, and co-located resources (Ambler & Lines, 2016; Hughes, 2013; Muntean & Surcel, 2013; Powell, 2014).

Fast analytics and data science are more fluid and iterative than BI due to the discovery involved in investigating a problem statement. Fast analytics and data science are inherently agile as each follows iterations, use small teams, and require collaboration between business subject matter experts and technical resources. Time-boxed increments can be applied, but may or may not be used as both processes are focused on discovery and data science has the objective of creating an analytical model that produces the best results (Mohanty, Jagadeesh, & Srivatsa, 2013; Schutt & O'Neil, 2013).

4.5.9. Agile delivery framework for BI, fast analytics and data science

This author compiled a BI Delivery Framework based on research and experience, which synthesized Agile practices with BI delivery practices (Larson, 2012). Since the development of this framework, additional research has been published that supports the value of using Agile methodologies with BI projects. Fast analytics and data science has become prominent practices in data analysis due to the emergence of Big Data. Organizations are expanded competencies in the BI field to include data scientists (Davenport, 2014; Schutt & O'Neil, 2013; Mohanty et al., 2013).

4.6. Valuable practices from agile principles

4.6.1. Discovery

The expectations of BI projects are not always clear to stakeholders. End users know they need information and analysis capabilities and IT knows they need to deliver something. This phase is where discovery is highlighted the most. Outlining business questions are a best practice in gathering BI requirements. These questions immediately provide insight into data sources, dimensions, and facts needed.

Most of what can and cannot be delivered is determined by data quality and availability. Once data sources have been identified, the next step requires gaining an understanding of the data. Data profiling focuses on two phases – values analysis and structure analysis. Data profiling provides data demographics and descriptive statistics such as: frequency distribution, high and low values, blank attributes and records, exceptions to domain values, dependencies between attributes, unknown constraints, mean, median, mode, and standard deviation. The knowledge gained from analyzing data demographics provides the basis for data quality metrics and can be used later in the lifecycle for modeling, development, and testing. Most importantly, assumptions about the data and

information capabilities are removed. With this knowledge, information needs can be prioritized and increments planned (Larson, 2009).

With fast analytics and data science, exploratory data analysis is a required process to review data demographics, explore variable relationships such as linearity, and select parameters to be considered in an analytical model. Data profiling fits well here as summary demographics provide a data scientist a comprehensive view of data. The fast analytics and data science phases of scope, data acquisition/discovery, and analyze/visualize align with this phase of BI Delivery.

4.6.2. Architecture

At the beginning of a BI program, the architecture needs to be established. Creating a flexible, scalable architecture is essential to supporting growth. Envisioning the architecture is the first step in Agile BI (Ambler, 2006). As mentioned in the BI lifecycle section, BI architecture includes the business, technical, process, data, and project architecture.

Envisioning the architecture begins with diagramming. Diagrams work well in Agile as they are easily altered and maintained versus text-based documents. Diagrams include data models, data flows, process flows, and infrastructure diagrams. With technical architecture, the deliverable can be a diagram outlining the different technologies required. A conceptual subject-level model can be the beginnings of the data architecture.

Diagrams are a beginning, but they do not prove out the architectural vision. Architecture decisions are ones that cannot be easily reversed once implemented. The approach of a reference implementation works well in the Agile paradigm. Like a prototype, a reference implementation is a working model but focuses on proving out the architecture. Reference implementations for ETL architecture, for example can demonstrate if service levels are possible and remove assumptions about the technology. A proof of concept (POC) is also another approach used in validating architectural decisions. POCs are often used in BI due to organizations using the best of breed approach. The best of breed approach is defined as organizations choosing independent tools, such as ETL and databases, which need to be integrated as part of the technical architecture. Although reference implementations and POCs are used in traditional software development, in Agile BI they become the rule (Larson, 2009).

Architecture for fast analytics and data science primarily focuses on technical architecture. Technical architecture trends include open-source options like Hadoop and R for analytical modeling (Davenport, 2014; Schutt & O'Neil, 2013). Hadoop is considered the best practice for Data Lake environments due to high volumes of unstructured data. BI technical architecture and Data Lake architecture have to integrate and share data. Reference implementations and POCs are applicable for fast analytics and data science.

4.6.3. Design

The activities completed in the design phase of the BI framework are modeling and mapping. These activities are iterative in nature and use the output of the discovery phase. Data profiling analysis and high-level architectural diagrams provide the context for design.

Modeling in this framework is focused on prioritized requirements, data demographics, and a stable scope for the increment. Business questions provide an understanding of how data will be used and data demographics assist the modeler in identifying business transactions, uniqueness, and primary/foreign key relationships. The modeling iteration is shortened through the use of data discovery early in the project. The modeling iteration may include a completed logical or physical model; however, due to the iterative cycles, the models may be a first pass. At a minimum,

models will demonstrate behavioral and informational semantics. Models can represent sources and targets.

Mapping the data between source and target is an essential design activity. The source to target mapping will be evolutionary within the scope of the increment. The exercise of mapping confirms data understanding and discovers business, transformation, and cleansing rules.

By having models and source to target mappings, development on ETL and end user capabilities can begin. Refinements to the design can occur via development iterations. Subject matter experts from the business and IT collaborate to clarify and refine design throughout the increment.

As outlined, in fast analytics and data science, data is not modeled before used in analysis. Iterations are completed as part of discovery to further analysis and understand data and determine what data products are to be created (Davenport, 2014; Schutt & O'Neil, 2013). Design activities are fluid and very informal. Data models if created are done at the conceptual level due to unstructured data (Adamson, 2015).

4.6.4. Development

In an Agile environment, the goal of development is to deliver working software regularly. In BI, development deliverables can include ETL processes, analysis, or reporting capabilities. Different approaches to ETL exist such as Enterprise Application Integration (EAI), Enterprise Information Integration (EII), and Extract, Load, and Transform (ELT) which are out of scope for this research. Regardless of ETL approach, BI development includes an ETL deliverable.

Development iterations focus on the delivery of requirements; however, the requirements are not delivered the first cycle. Throughout the requirements and design iterations, stakeholders are working with the data to confirm understanding and remove assumptions. Development will produce software that enriches the data. The development iteration refines requirements and design through stakeholder collaboration. Stakeholders can confirm information results through validation of business rules and verification of output to alternate sources. Through development iterations, the scope that can be delivered in the allotted timeframe becomes clear. At the conclusion of the development phase, requirements and design are concluded for the increment and the development deliverables are ready to be tested.

Fast analytics and data science are farther removed from traditional development and producing software than BI. The products that come from fast analytics and data science can be one-time visualizations or analytical models. Data munging, the process of parsing, scraping, and formatting data, requires development; however, development is not focused on meeting customer requirements (Davenport, 2014; Schutt & O'Neil, 2013). Analytical models require development, but often utilize packaged functions to apply machine learning algorithms. The fast analytics and data science phase of model/design/development aligns with the development phase of BI.

4.6.5. Test

With an Agile approach, testing occurs constantly through the interactions of stakeholders. Collaboration with stakeholders ensures results are verified during the lifecycle to produce higher quality results. Since BI systems tend to be complex, a formal change control process is recommended. Additionally, a regression test suite for the BI system is essential. With the fuzzy nature of information, it is possible to impact prior working functionality and not see the impact until after deployment.

Testing for fast analytics and data science is included in the model/design/development phase and is not separated out. Analytical models go through the fitting process where models are refined

and tuned to minimize overall error (Davenport, 2014; Schutt & O'Neil, 2013).

4.6.6. Deploy

Complex BI systems require formal support and maintenance procedures to ensure the overall health of the system. This is where the flexible nature of Agile ends. New increments need a formal process to ensure existing operations and service levels are not impacted. Without a formal process, the risk of failure increases. Using an incremental approach allows a gradual and controlled deployment. In addition, introducing new functionality sooner allows stakeholders to recognize value and lessens the complexity of deployment. Formal support and feedback procedures are needed for analytical models that are deployed. Analytical models lose effectiveness due to environmental change (Davenport, 2014; Schutt & O'Neil, 2013).

5. Summary of the agile delivery framework for BI, fast analytics, and data science

The basis for the framework has been established through the analysis and synthesis of Agile principles with BI, fast analytics and data science. Agile practices have been adopted by BI practitioners and recent research has supported success in this area. BI has been impacted by Big Data and analysis has shifted from descriptive analysis to predictive and prescriptive analysis. Based on these changes, additional content has been added to extend the framework to include the new analysis trends – fast analytics and data science.

The following diagram depicts the framework for Agile BI delivery, fast analytics and data science. There are two layers of strategic tasks that go hand in hand in the Agile BI Delivery Framework. The top layer includes BI Delivery and the bottom layer includes Fast Analytics/Data Science. In the top layer, there are five sequential steps involved: Discovery, Design, Development, Deploy and Value Delivery. In each step, there are specific tasks to be completed and are related to the goals of achieving business and IT stakeholder collaboration. The bottom layer includes six sequential steps: Scope, Data Acquisition/Discover, Analyze/Visualize, Validate and Deployment. Similarly, all these steps work towards successful business and IT stakeholder collaboration. The alignment, integration and streamlining both layers can ensure the execution and management for the Agile BI Delivery Framework.

6. Discussion

The scope of this paper focused on an Agile delivery framework addressing the influence of Big Data on Business Intelligence. Fast analytics and data science are not the only emerging trends that should be considered in Business Intelligence and Big Data adoption. Emerging trends that warrant further research not addressed in the scope of the paper include Emerging Services and Analytics, cloud delivery and storage models, and security implications with Big Data.

BIaaS is a Software as a Service (SaaS) approach to BI where applications are offered as a service where computing occurs in a public or a private cloud (Chang, 2014). BIaaS emerged after applications began migrating to the cloud; however, BIaaS has taking time to become mainstream. Organizations take advantage of BIaaS primarily to lessen the delivery time of analytics. Improving time to value and minimizing the cost of implementation and infrastructure are the primary drivers for organizations to move to BIaaS (Chang, 2014). Everything as a Service (EaaS) approaches are still evolving and it is not clear on how Big Data will influence BIaaS. Similarly, Newman et al. (2016) explain that the use of Business

Data Science (BDS) approach can enhance business performance. Since there are no clear guidelines in BDS, they focus on the introduction of their system design and experimental approach to allow organizations to consider adopting their proposal.

BIaaS is only one area that is emerging as a service. Chang (2016) introduces Emerging Services and Analytics where services are delivered using a combination of Platform as a Service (PaaS) and SaaS. Emerging Services and Analytics are presented as one integrated service which merges different technologies to work easily with users who use the service. Emerging Services and Analytics include merging technologies such as the data warehouse, machine learning, visualization, analytical models, software, and Big Data (Chang, 2016). Visualization, while not a new concept, is a key component of fast analytics. Visualization as a part of a service allows users to quickly understand complex data sets created from statistical analysis or analytical models. Big Data is often defined not only by velocity, variety, and volume, but also by validity, veracity, value, and visibility – implying the importance of visualization.

At the heart of services is cloud computing. Cloud computing provides shared processing resources as a service to organizations and individuals. Services utilize cloud computing and Big Data is an area that is being researched due to MapReduce, which is a framework used to process Big Data quickly (Chang & Wills, 2016). MapReduce maps and categorizes data together, thereby reducing the data to be processed. Big Data in the cloud offers opportunities for data scientists and other communities to process and analyze data quickly using terabytes of data and many different data types including images. Both cloud and non-cloud solutions are still being researched as to which is the best option. Chang and Willis (2016) used Organizational Sustainability Modeling (OSM) to measure performance between cloud and non-cloud platforms. They proposed that non-cloud solutions may be best for secure and private data, but may not provide the most consistent performance and that cloud solutions provide the most consistent performance.

Other considerations with Big Data adoption include disaster recovery and security. Privacy and security are concerns for organizations that use cloud computing as processing platforms are outside the realm of control. Big Data is a concern due to new categories of sensitive data used for analysis including biomedical data and new forms of personal identifiable information (PII). Combine Big Data with cloud computing and security and privacy concerns increase. Chang, Kuo, & Ramachandran, 2016 outline that security and privacy are challenges for organizations that use cloud computing and Big Data. Challenges include ensuring current security practices are applied as part of the system development. Additionally, they recommend a risk-based approach to ensure data is maintained with integrity and confidentiality. Approaches such as Cloud Computing Adoption Framework (CCAF) are being developed to provide greater security in the cloud. CCAF is multi-layered security framework that includes identity management, firewall, and encryption (Chang et al., 2016). Since security for businesses pose challenges for growing numbers of organizations, ethical hacking and large scale penetration testing will be essential to test the robustness of secure systems. Chang and Ramachandran (2016) use their CCAF framework to test with 10,000 known viruses and trojans in 2013 and results showed that their work can block and kill more than 99% of viruses and trojans that has been injected every hour. In the event of having the data centers compromised, they identified that all rescued and recovery should be completed within 125 h. They also tested with 10 petabytes of data to verify the robustness of resiliency of their approaches. In summary, BI services and platforms should be resilient to all types of attacks and can provide a guarantee of secure services as close as to 100%.

As part of security, data integrity is also a concern. Veracity and value of Big Data can be impacted if a solid disaster recovery process is not in place. There is much focus on security and privacy of Big Data, but not much research on disaster recovery solutions for Big Data. Chang (2015) outlines that Big Data poses a disaster recovery challenge because of the quantity, size and complexity of the data. Chang (2015) outlines challenges with backing up virtual machines, using multiple sites, and replication. Issues with Big Data since as data ownership and governance also contribute to disaster recovery challenges.

7. Conclusion

This paper focused on the recent developments in adoption of Agile principles to BI delivery and how Agile has changed with the face of BI. Fast analytics and data science have been included under the umbrella of BI. Agile ideals fit well into the BI world and research on successful application has emerged. Agile addresses many of the common problems found in BI projects by promoting interaction and collaboration between stakeholders. Close collaboration between parties ensures clearer requirements, an understanding of data, joint accountability, and higher quality results. Less time is spent attempting to determine information requirements, and more time is devoted to discovering what is possible. Future research opportunities are abundant as the landscape of BI and data analysis is transforming with Big Data. Topics in discussion have addressed the current challenges and future directions for adopting business intelligence platforms, applications and services for all types of organizations.

Keywords and definitions

Analytics is the information gained from the systematic analysis of data and statistics.

Business Intelligence is a data driven process that combines data storage and gathering with knowledge management to provide input into the business decision making process. Business Intelligence enables organizations to enhance the decision making process.

The Cloud or Cloud computing, also on-demand computing, is a kind of Internet-based computing that provides shared processing resources and data to computers and other devices on demand

Data Warehouse is an integrated, subject-oriented, non-volatile, time-variant data store. A data warehouse can also be considered a union of all data marts. Data warehouses are used to support business intelligence

Extract, Transformation, and Load is the process used to capture data from sources, apply transformation rules (filter, select, derive, translate, convert), and load data into a target data store.

Information Technology is the use of software, hardware, and infrastructure to manage and deliver information. Organizations have information technology departments that focus on managing information as an asset. Information technology departments manage the infrastructure used to deliver information Information Utility is the process of using information for decision making, knowledge extraction, or for other management activities.

Metadata are all the information that is used to define and describe contents, definitions, operations, and structures within an organization's system architecture. Metadata are required by users to understand data meaning and context. Metadata can be categorized into business, technical, and process metadata.

Proof of Concept (POC) is demonstration in principle, whose purpose is to verify that a concept or theory is feasible.

References

Adamson, C. (2015). *The age of big data modeling.* www.tdwi.org

Ambler, S. (2006). *Agile database techniques: effective strategies for the agile developer.* John Wiley & Sons.

Ambler, S. W., & Lines, M. (2016). The Disciplined Agile Process Decision Framework.

Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., et al. (2001). *Manifesto for agile software development.*. Available on: http://agilemanifesto.org/

Chang, V. (2014). The business intelligence as a service in the cloud. *Future Generation Computer Systems, 37*(July), 512–514.

Chang, V. (2015). Towards a big data system disaster recovery in a private cloud. *Ad Hoc Networks, 35,* 65–82.

Chang, V. (2016). An overview, examples and impacts offered by Emerging Services and Analytics in Cloud Computing. *International Journal of Information Management,* in press.

Chang, V., Kuo, Y., & Ramachandran, M. (2016). Cloud Computing Adoption Framework—a security framework for business clouds. *Future Generation Computer Systems, 57*, 24–41.

Chang, V., & Ramachandran, M. (2016). Towards achieving data security with the cloud computing adoption framework. *IEEE Transactions on Services Computing, 9*(1), 138–151.

Chang, V., & Wills, G. (2016). A model to compare cloud and non-Cloud storage of big data. *Future Generation Computer Systems, 57*, 56–76.

Cockburn, A. (2008). Using both incremental and iterative development. *The Journal of Defense Software Engineering.*

Couture, N. (2013). Best practices for adopting agile methodologies for data warehouse development. *Business Intelligence Journal, 18*(2), 8–17.

Creswell, J. W. (2005). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Davenport, T. H. (2013). Analytics 3.0. *Harvard Business Review, 91*(12), 64–72.

Davenport, T. H. (2014). *Big data @ work: dispelling the myths, uncovering the opportunities.* Harvard Business Review Press.

Gartner IT Glossary. (2013). *Definition of business intelligence..* Retrieved from: http://www.gartner.com/it-glossary/business-intelligence-bi/ (accessed on 15.04.16.)

Gartner Research. (2015). Gartner says business intelligence and analytics leaders must focus on mindsets and culture to kick start advanced analytics. *Gartner business intelligence & analytics summit 2015.*

Halpern, F. (2015). *Next-generation analytics and platforms for business success: tDWI research report.* Available on: www.tdwi.org

Hsieh, C. Y., & Chen, C. T. (2015). Patterns for continuous integration builds in cross-platform agile software development. *Journal of Information Science & Engineering, 31*(3), 897–924.

Hughes, R. (2013). *Agile data warehousing project management: business intelligence systems using Scrum and XP.* Waltham, MA: Morgan Kaufmann.

Jarr, S. (2015). *Fast data and the new enterprise data architecture.* O'Reilly Publishing.

Kaleshovska, N., Josimovski, S., Pulevska-Ivanovska, L., Postolov, K., & Janevski, Z. (2015). The contribution of scrum in managing successful software development projects. *Economic Development/Ekonomiski Razvoj, 17*(1/2), 175–194.

Kenney, T. (2015). How to achieve Agile business intelligence success. *CIO, 1* (13284045).

Kendall, J. E., & Kendall, K. E. (2004). Agile methodologies and the lone systems analyst: when individual creativity and organizational goals collide in the global IT environment. *Journal of Individual Employment Rights, 11*(4), 333–347.

Kiem, D., Kohlhammer, J., & Ellis, G. (2010). *Mastering the information age: solving problems with visual analytics.* Eurographics Association.

Larson, D. (2009). BI principles for agile development: keeping focused. *Business Intelligence Journal, 14*(4), 36–41. Retrieved from Business Source Complete database.

Mohanty, S., Jagadeesh, M., & Srivatsa, H. (2013). Big data imperatives: enterprise big data warehouse. *BI Implementations and Analytics.*

Muntean, M., & Surcel, T. (2013). Agile BI—the future of BI. *Informatica Economica, 17*(3), 114–124. http://dx.doi.org/10.12948/issn14531305/17.3.2013.10

Mungree, D., Rudra, A., & Morien, D. (2013). *Proceedings of the Nineteenth Americas Conference on Information Systems.*

Negash, S., & Gray, P. (2008). Business intelligence. In F. Burstein, & C. W. Holsapple (Eds.), *Handbook on decision support systems 2* (pp. 175–193). Berlin, Heidelberg: Springer.

Newman, R., Chang, V., Walters, R. J., & Wills, G. B. (2016). Model and experimental development for business data science. *International Journal of Information Management, 36*(4), 607–617.

Noble, J. (2006). *The core of IT.* pp. 15–17. CIO Insight.

Powell, J. (2014). *BI this week: agile basics and best practices..* Available on: www.tdwi.org

Schutt, R., & O'Neil, C. (2013). *Doing data science: straight talk from the frontline.* O'Reilly Media, Inc.

Singer, T. (2001). Information engineering: the search for business intelligence. *Plant Engineering,* 34–36.

The Data Warehouse Institute. (2009). *TDWI requirements gathering: getting correct and complete requirements for BI systems.* The Data Warehouse Institute. www.tdwi.org

Yeoh, W., & Koronios, A. (2010). Critical succes factors for business inteligence systems. *Journal of Computer Information Systems, 50*(3), 23–32. Retrieved from: Business Source Complete database.

Futher reading

Abrahamsson, P., Salo, J., & Ronkainen, J. (2008). Agile software development methods: review and analysis. *VTT Technical Report.*

Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: the new science of winning.* Boston: Mass Harvard Business School Press.

Dyba, T., & Dingsøyr, T. (2008). Empirical studies of agile software development: a systematic review. *Information and Software Technology,* http://dx.doi.org/10.1016/j.infsof.2008.01.006

Hedgebeth, D. (2007). Data-driven decision making for the enterprise: an overview of business intelligence applications. *VINE, 37*(4), 414–420. http://dx.doi.org/10.1108/03055720710838498

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: a revolution that will transform how we live, work and think.* John Murray: London.

Marr, B. (2015). Big data: Using smart big data, analytics and metrics to make better decisions and improve performance.

Sriram, R. S. (2008). Business intelligence in context of global environment. *Journal of Global Information Technology Management, 11*(2), 1.

Wang, Xiaofeng, Lane, Michael, Conboy, Kieran, & Pikkarainen, Minna. (2009). Where agile research goes: starting from a 7-year retrospective (report on agile research workshop at XP2009). *SIGSOFT Software Engineering Notes, 34*(October (5)), 28–30. http://dx.doi.org/10.1145/1598732.1598755 http://doi.acm.org/10.1145/1598732.1598755