# Big Data and Analytics in the Modern Audit Engagement: Research Needs

**Article** *in* Auditing A Journal of Practice & Theory · February 2017

DOI: 10.2308/ajpt-51684



Deniz A Appelbaum
Montclair State University



Miklos A. Vasarhelyi
Rutgers Business School

# Big Data and Analytics in the Modern Audit Engagement: Research Needs

Deniz Appelbaum
PhD Candidate
Rutgers, the State University of New Jersey, Newark
deniz.appelbaum@rutgers.edu
cell: 202-425-3633

Dr. Alexander Kogan
Professor
Rutgers, the State University of New Jersey, Newark
kogan@business.rutgers.edu
office:

Dr. Miklos A. Vasarhelyi
KPMG Distinguished Professor of AIS
Director of CARLab
Rutgers, the State University of New Jersey, Newark
miklosv@andromeda.rutgers.edu
cell: 201-454-4377

# Big Data and Analytics in the Modern Audit Engagement: Research Needs

***Abstract:***

*Modern audit engagements often involve examination of clients that are using big data and analytics to remain competitive and relevant in today's business environment. Client systems now are integrated with the cloud, the Internet of Things, and external data sources such as social media. Furthermore, many engagement clients are now integrating this big data with new and complex business analytical approaches to generate intelligence for decision making. This scenario provides almost limitless opportunities and also the urgency for the external auditor to utilize advanced analytics. This paper first positions the need for the external audit profession to move towards big data and audit analytics. It then reviews the regulations regarding audit evidence and analytical procedures, in contrast to the emerging environment of big data and advanced analytics. In a big data environment, the audit profession has the potential to undertake more advanced predictive and prescriptive oriented analytics. The next section proposes and discusses six key research questions and ideas followed with particular emphasis on the research needs of quantification of measurement and reporting. This paper provides a synthesis and review of the concerns facing the audit community with the growing use of big data and complex analytics by their clients. It contributes to the literature by expanding upon these emerging concerns and providing opportunities for future research.*



Keywords: Audit Analytics, big data, external audit, audit evidence

# Big Data and Analytics in the Modern Audit Engagement: Research Needs

## Introduction/Discussion of the Current External Audit Environment

There is an increasing recognition in the audit profession that the emergence of big data (Vasarhelyi, Kogan, and Tuttle 2015) as well as growing use of data analytics in business processes has brought a set of new concerns to the audit community. Accountants[1], Large Audit Firms[2], Standard Setters[3], and Academics [4] have been progressively raising many issues, among which we find:

1. Should new (modern) analytics methods be used in the audit process?
2. Which of these methods are the most promising?
3. Where in the audit are these applicable?
4. Should auditing standards be changed to allow / facilitate these methods?



[1] The AICPA's Assurance Services Committee (ASEC) has met three times over the last three years with the Auditing Standards Board (ASB) to discuss audit analytics, and how the use of analytical tools and techniques fit within the current standards. As a result, the ASEC is developing a new Audit Data Analytics guide that will replace the current Analytical Procedures guide. The Audit Data Analytics guide will update and carry forward much of the content found in the Analytical Procedures guide, and will also include discussions around Audit Data Analytics and how they can fit within the current audit process. ASEC's Emerging Assurance Technologies task force is also working on a document that will map the traditional audit procedures to current analytical tools available today and elements of continuous audit.

[2] Every one of the "Big Four" has publicly announced efforts in the area of data analytics. Some have published white papers on the matter (e.g. Deloitte, "Adding insight to audit – Transforming Internal Audit through data analytics"; PwC, "The Internal Audit Analytics Conundrum—Finding your path through data"; KPMG, "Leveraging data analytics and continuous auditing processes for improved audit planning, effectiveness, and efficiency"; EY, "Big data and analytics in the audit process: mitigating risk and unlocking value").

[3] In April 2015, the IAASB started a subcommittee on analytic methods and heard presentations on the matter (e.g., Dohrer, Vasarhelyi, and McCollough 2015). The objectives of the subcommittee are to explore developments in audit data analytics and how the IAASB will respond to these developments. Also, the PCAOB has approached the "Big Four" to discuss the usage of analytics.

[4] A special section of Accounting Horizons with 7 articles (see Vasarhelyi, Kogan, and Tuttle 2015) has been dedicated to big data. An increasing number of articles in the accounting literature (see ensuing sections) have emerged proposing and illustrating analytic methods.

5. Should the auditor report be more informative?[5]

6. What are the competencies needed by auditors in this environment?

These concerns have emerged even though analytical procedures in general have been addressed by the American Institute of Certified Public Accountants (AICPA) guidelines of 1972 and in numerous academic papers since 1955. The Statement on Auditing Standards (SAS) No. #1, states:

> *"The evidential matter required by the third standard (of field work) is obtained through two general classes of auditing procedures: (a) tests of details of transactions and balances, and (b) analytical review procedures applied to financial information (AICPA 1972 par. 320.70)."*



There is a fine balance in every audit engagement between detailed evidence collection and analytical procedures (Yoon 2016). Detailed evidence collection can be quite costly yet deemed more reliable according to the standards, while analytical procedures are widely viewed as being less costly and believed less reliable by regulators (Daroca and Holder 1985; Tabor and Willis 1985). Both processes are allowed by the standards; their degree of utilization depends on auditor professional judgment. While the requirement of tests of details of transactions and balances is somewhat defined, the second requirement of analytical review procedures is completely undefined, except that it should be applied to financial data (Tabor and Willis 1985).

More recently, according to AU-C Section 520 about Analytical Procedures (AICPA 2012a), to conduct substantive analytical procedures the auditor should:

- determine the suitability of a certain substantive procedure, given the account;
- evaluate the reliability of the data from which these ratios are developed;

[5] The PCAOB issued Release No. 2016-003 on May 11, 2016 re-proposing new standards for the audit report in which in addition to the traditional pass/fail model "critical audit matters" (CAM) would be disclosed.

- develop an expectation of recorded amounts and ratios and whether these are accurate, and finally
- determine the amount of difference (if any) between the recorded amounts and the auditor's expected values and
- decide if the difference is significant or not.

The lack of detailed recommendations in this age of automation and big data regarding which analytical procedures to undertake in the external audit engagement has inspired considerable discussion. Although the internal audit environment is increasingly using analytics (Vasarhelyi et al. 2015; Perols and Lougee 2011; Dilla et al. 2010; Yue et al. 2007; Alles et al. 2006; Church et al. 2001), the external audit field has not responded to the same degree. The regulations, such as the guidance for sampling, have remained unchanged even though many audit clients automate the collection and analysis of 100% of their transactions (Schneider et al. 2015; Zhang et al. 2015).



This paper provides a synthesis and review of the concerns facing the audit community with the growing use of big data and complex analytics by their clients. It contributes to the literature by clarifying and expanding upon these emerging concerns and by suggesting opportunities for future research. This paper first reviews the current standards regarding evidence collection and analytical procedures as currently understood in the profession, before discussing big data and business analytics. The role of big data and business analytics and their implications for the audit profession should first be understood in the context of current practice. Then this paper broadly reviews each of these six concerns emerging in the profession as a result of the use of big data and analytics by engagement clients. These concerns are subsequently followed with an elaboration of additional forward looking research issues, with special emphasis on the quantification of audit processes and judgments.

## Background: Current Practice and the Standards

It is essential to understand the current scope and constraints of the public audit profession before envisioning the role of more complex analytics and big data in the engagement. Since auditing is largely a regulation driven profession, the expectations regarding evidence collection and analytical procedures should be considered. The auditor still needs to test for basic assertions to make sure that the objectives of the audit are fulfilled regardless of the nature of the evidence and the way the evidence is being collected. The tests for certain assertions may change in the current new environment with its different nature of evidence and the way this evidence is collected and analyzed. However, even if the tests of assertions were to be altered, the assertions themselves wouldn't change and neither would the fundamental objective of the public auditor – to provide opinion on the financial statements as to whether they represent the financial position of the client in accordance with the generally accepted accounting principles.



### *Evidence Collection and the Standards*

The main purpose of the work conducted by an auditor in an external engagement is to obtain reasonable assurance that the client's financial statements are free from material misstatements and to subsequently express an opinion regarding these financial statements and the client's internal controls in the auditor's report. To accomplish this task, the auditor must design and perform audit procedures to obtain sufficient appropriate evidence; furthermore, the Audit Standards require auditors to examine physical evidence as part of the risk assessment process (PCAOB 2010, AS 1105; AICPA 2012, SAS 122; IAASB 2009, ISA 500). Audit evidence is all the information (whether obtained from audit procedures or other sources) that either confirms or contradicts or is neutral about management's assertions on the financial statements or internal controls.

Additionally, the Sarbanes-Oxley Act (SOX) demands that auditors verify the accuracy of the information or evidence that forms the basis of their audit opinion. Since SOX, audit firms have relied more heavily on detailed audit examination and scanning for substantive tests as these are regarded to be "harder" audit evidence formats than regression and other "softer" analytical techniques (Glover et al 2014). The impact of this legislation on the profession's analytical procedures choices should not be ignored. However, as mentioned and footnoted in the Introduction, every one of the "Big Four" has recently publicly announced efforts in the domain of data analytics for assurance services.



Since audit evidence is all the information used by the auditors to form the audit opinion (PCAOB, 2010, AS 1105), it should be both sufficient and appropriate. Basically, if the underlying information is not reliable or strong enough and its origin isn't verifiable, then more evidence will need to be collected and reviewed (Appelbaum, 2016). Poor quality evidence cannot be compensated for by collecting a larger amount of data (PCAOB 2010, AS 1105).



However, in today's complex IT and big data environment, the nature and competence of this audit evidence has changed (Brown-Liburd and Vasarhelyi 2015; Warren et al. 2015; Nearon 2005). With big data, quantity of evidence is hardly an attribute with which to be concerned. However, quality of electronic evidence becomes even more dominant in the equation and may be more challenging to verify. Most stages of a transaction can be computer generated and recorded and can only be verified electronically. For example, with additional information available from external big data, intangible assets might be partially valued by the client from information derived from text analysis of aggregated tweets and web scraping of social media.

Unfortunately, the reliability of these tweets and social media is hard to verify (Appelbaum 2016).

The issues for electronic accounting data and electronic audit evidence are drastically different from that of manual and paper-based examination. Many of the characteristics that are strengths with paper-based evidence pose issues for electronic evidence. Where paper documentation is regarded as not easily altered, electronic data may be easily changed and these alterations might not be detected, absent the appropriate controls. In paper-based evidence collection, sources that are verified external to the client are considered to be highly reliable (PCAOB 2010, AS 1105), whereas external electronic evidence is difficult to verify for origin and reliability. Paper-based evidence is easy to evaluate and understand, whereas electronic data and evidence may require a high level of technical expertise of the auditor. Since big data is electronic data, big data presents a scenario where these complexities are magnified greatly. Furthermore, the types of tests that should be undertaken by auditors to examine basic assertions may change.



***Analytical Procedures and the Standards***

Analytical procedures are required by the Public Company Accounting Oversight Board (PCAOB) in the planning phase (PCAOB 2010, AS No. #2110) and review phase (PCAOB 2010, AS No. #2810), but are undertaken according to auditor judgement in the substantive procedures phase (PCAOB 2010, AS No. #2305).

The purpose of analytical procedures is different for each audit phase. For the risk assessment/planning phase, analytical procedures should enhance the auditor’s understanding of the client's business and its transactions or events, and identify areas that may indicate particular risks to the audit. The auditor is expected to perform analytical procedures for the revenue accounts, to reveal unusual relationships indicative of possible material misstatements. The auditor should also use his or her knowledge of the client and its industry to develop expectations. The standards admit that the data may be at a more aggregated level and result in a less precise analytical procedure which is still acceptable at this phase.

Accordingly, in AS No. #2305.04 analytical procedures are used in the substantive testing phase to obtain evidence about certain assertions related to certain accounts or business cycles. Analytical procedures may be more effective than tests of details in some circumstances (Yoon 2016). In AS No. #2305.09, the PCAOB states that "the decision about which procedure or procedures to use to achieve a particular audit objective is based on the auditor's judgement on the expected effectiveness and efficiency of the available procedures." The main limitations appear to be the "availability" of certain procedures and the auditor's judgement on the expected effectiveness of certain analytical methods. The latter condition would appear to reflect the auditor's level of familiarity with certain analytical methods.



For the review phase of the audit engagement, analytical procedures are required to evaluate the auditor's conclusions regarding significant accounts and to assist in the formation of the audit opinion (PCAOB 2010, AS No. #2810.05-.10). Similarly, in the planning phase the auditor is required to perform analytical procedures related to revenue during the relevant period. In this section, there is no mention of any one analytical approach, except that this phase typically is similar to the planning phase. As such, it is expected that the more complex exploratory or confirmatory techniques are not excluded here either (Liu 2014).

## Background: Current Business/Client Environment and Its Challenges

Auditors are required to conduct the audit engagement within the parameters of the regulations, regardless of the IT or accounting complexity of the client. It is highly probable that the client may be undergoing processes with advanced analytical techniques and new sources of data. The newest challenges facing the auditor are the increasing use of big data and the subsequent application of more advanced analytics by clients. After gaining an understanding of this current audit environment of big data and advanced analytics, what follows are immediate research questions that should be addressed if the profession is to integrate itself within this new business paradigm.



### *Big Data*

Many client systems now are increasingly integrated with the cloud, the Internet of Things, and external data sources such as social media. Client data may exhibit large variety, high velocity, and enormous volume – big data (Cukier and Mayer-Schoenberger 2013). This data may originate from sensors, videos, audio files, tweets and other textual social media – all data types typically unfamiliar to an auditor (Warren et al. 2015). However, this big data provides almost limitless opportunities to the external auditor to utilize advanced analytics. According to extant analytics research (Holsapple, Lee-Post, and Pakath 2014; Lee et al. 2014; Delen and Demirkan 2012), big data should provide auditors the opportunity to conduct prescriptive analytics – that is, to apply techniques that computationally determine available actions and their consequences and/or alternatives, given the engagement's complexities, rules, and constraints (Lee et al. 2014).

Furthermore, this environment of Big Data (Vasarhelyi, Kogan, and Tuttle 2015), personal devices and the Internet of Things (IoT) (Atzori, Lera, and Morabito 2010; Domingos 2011; Dai and Vasarhelyi 2016) is progressively interconnecting with corporate systems.[6] The economics of hardware and software development are of very different nature than traditional systems. It is not inconceivable that analytic methods such as regression may be built into chips, including powerful explanatory software[7] that would provide interpretations of the results and recommend decisions for the user, in this case an auditor.



Advances in text interpretation, voice recognition, and video (picture) recognition would additionally expand the interconnected environment previously described. On another dimension, the latency of information and its processing systems are progressively reduced, mainly as the result of faster chips, interconnected devices, and automatic sensing of information. The traditional annual audit, or even quarterly report evaluation would have limited meaning in this world of real-time measurement. A progressive audit[8] by exception methodology would be required in this type of environment.



In this big data environment with its many sources of information that would be novel for the audit profession to include in the examination, the standards regarding audit evidence may need to be discussed and possibly re-examined in the context of big data. Regardless of the source, the data should be reliable and verifiable. Table One outlines the challenges that big data poses to the current audit profession and suggests avenues of research:

[6] It is not surprising that this hybrid environment with numerous points of access and interconnections is a fertile ground for cyber-intrusion.

[7] Byrnes (2015) has developed a clustering decision aid that can make decisions in the clustering interpretation process without human intervention. More sophisticated devices can be built into chips to accelerate and formalize this process and can benefit from standard interfaces and protocols.

[8] Montgomery (1913) already argued for a "continuous audit" that would provide progressive review results instead of the final audit opinion.

(Insert Table One here)

How can the availability of big data sets, both internally and externally to the enterprise, be utilized to enhance analytics? Can the extremely large amounts of data compensate for uncertain or, at times, lower quality of such data? There are some that argue that big data is meant to be messy (Cukier and Mayer-Schoenberger 2013). In cases where big data is of dubious origins or lacking audit trails (Appelbaum 2016), the standards currently would indicate that no amount could compensate for being poor, unreliable data.

Consider for example the Jans et al. (2014) paper with the application of process mining. This paper details the use of process mining on a 100% test of the transactions to find the anomalies in the sample where controls fail in the processing of 26,185 POs. Basically, the audit trails are problematic. A series of process mining tests (a type of ADA) narrows the sample of anomalies down to the highest risk scenarios which exemplify high rates of violations among individuals and small networks of people working together. It seems that this is the perfect example of how ADAs can be used for more efficient audit testing.



Fraud related issues may be as challenging, if not more, to the audit team in a big data environment. More data does not necessarily equal more effective information, and the added complexity of the big data could complicate the assessment of audit evidence for fraud (Srivastava et al. 2009; Srivastava et al. 2011; and Fukukawa et al. 2014). Fraud detection also focuses on the assessment of internal controls, regardless of whether the analytics are based on sampling or on processing 100% of the population. It is important to point out that no matter how strong the internal control system, management can still perpetrate fraud by over-riding the internal controls. In a big data environment, it is quite possible that the volume and complexity of the data might actually hinder what is already a troublesome task for many engagement teams – the determination of the probability that fraud has occurred.

Furthermore, how can the amount of audit evidence provided by analytics in a big data context be measured? How can this evidence be aggregated with other types of audit evidence in a methodologically sound way? How can such quantitative measures be used to provide support for the auditor's judgement about the sufficiency of audit evidence? The entire standards of audit evidence may need to be reassessed and subsequently revised in this age of electronic and big data evidence (Appelbaum 2016; Brown-Liburd and Vasarhelyi 2015). Electronic and big data evidence often raise issues opposite of those assumed by the standards for paper-based documentation. As business processes now are very infrequently paper-driven, the standards on reliable evidence, which are derived from quality evidence of sufficient amount, may need to be revised to provide a more quantitative measure of quality vs. quantity in an IT audit.



***Business Analytics***

Care should be exercised when discussing analytical procedures and business analytics (BA) in the public audit engagement context because the two terms might not be completely interchangeable. Analytical procedures, according to AS 2305 (PCAOB, AS 2305 2016), are an important part of the audit process and mainly consists of an analysis of financial information made by a study of believable or plausible relationships among both financial and non-financial data. These analytical procedures could be as basic as scanning (viewing the data for abnormal events or items for further examination) to more complex approaches (not clarified by the standards, except that the approach should enable the auditor to appropriately develop an expectation and subsequently examine these expectations to the reported results).

Business Analytics (BA) that is utilized by client management and their accountants has been defined as "*the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their operations, and make better, fact-based decisions*" (Davenport and Harris 2007). BA may be further conceptualized with the three dimensions of Domain, Orientation, and Technique. (Holsapple et al 2014). Domain represents the context or environment for the analytics. Orientation describes the vision or focus of the analysis – either descriptive, predictive, or prescriptive. Descriptive orientation answers what happened and is backward looking. Its techniques convert this analysis into useful information via visualization, graphs, and descriptive statistics. Predictive orientation then takes the descriptive information of what happened and hypothesizes what could happen. Predictive analysis is the process of developing expectation models, with which auditors are quite familiar. Basically, predictive analysis uses data from the past and the present to generate relevant predictions (many logical, statistical, machine learning approaches). Prescriptive orientations take predictions further. Based on what happened and using experimental design, this mode presents an optimization analysis to identify the best possible alternative. The techniques define the actual method or approach for analysis. (Holsapple et al 2014; Davenport and Kim 2013; Evans 2012).



The focus or context of BA for management would be somewhat different from that of the auditor. Management accountants are seeking to extract and develop insightful knowledge to enhance efficiency and effectiveness of operations, in addition to providing forecasts to enhance management decision-making. Internal auditors are seeking to verify the effectiveness and accuracy of this information. External auditors are concerned with BA as they relate to verification of the veracity of the financial statements. However, both audit tasks involve generating expectation models as well as confirmatory models. Since auditors examine business financial data, their work is affected by business analytics.

Techniques are the analytical approaches that can be described as either descriptive, predictive, or prescriptive, depending on the task of the analysis and the type of data. The more forward looking the task and the more varied and voluminous the data (big data), the more likely the analysis will be prescriptive or at the very least, predictive. Advanced or more complex BA may be defined as *"Any solution that supports the identification of meaningful patterns and correlations among variables in complex, structured and unstructured, historical, and potential future data sets for the purposes of predicting future events and assessing the attractiveness of various courses of action. Advanced analytics typically incorporate such functionality as data mining, descriptive modeling, econometrics, forecasting, operations research, optimization, predictive modeling, simulation, statistics, and text analysis"* (Kobelius 2010).



If audit clients are utilizing these more advanced BA techniques operation wide, is the auditor conducting an effective and efficient engagement by utilizing ratio and trend analysis and scanning, which are the techniques typically used and with which the auditor is comfortable (Glover et al. 2014)? When would the auditor rely more on analytical procedures over substantive detailed testing? Or, is there room in the current understanding and regulations of analytical procedures for these more complex approaches? Can analytical procedures be regarded as Audit Data Analytics?

Stewart (2015) defines: "Audit Data Analytics (ADA) is the analysis of data underlying financial statements, together with related financial or non-financial information, for the purpose of identifying potential misstatements or risks of material misstatement." This definition is illustrated by linking analytical procedures with traditional data procedures (Figure One). ADA encompasses both the traditional file interrogation with which auditors are quite familiar as well as analytical procedures and analytics, some of which auditors may be less acquainted with. Both may be more easily understood by obtaining an understanding of the modes of ADA. Traditional file interrogation and analytical procedures are subsets of the larger field of ADA. If ADA is understood as exploratory or confirmatory in task, this task oriented approach "allows" the auditor to utilize other techniques.

(Insert Figure One here)

Liu (2014) has proposed the use of Exploratory Data Analysis (EDA) (Tukey 1977, 1980) in the audit process to generate more directed and risk sensitive audit assertions for their ensuing usage through Confirmatory Data Analysis (CDA). Furthermore, Liu (2014) examined where these applications could be used in the audit process as well as their placement in extant audit standards (see Appendix A). Liu (2014) and Stewart (2015) placed EDA and CDA into the context of audit data analytics and argued for its usage as parts of audit standards. To this definition, Stewart (2015) and Liu (2014) add that ADA can be exploratory and confirmatory and illustrate its functionalities. Although new or more complex methods can be proposed and even adopted by firms, it does not mean that these methods are being promoted by the standards – instead, these new methods are simply not precluded. For instance, while regression was incorporated in the Deloitte, Haskins and Sells methodology (Stringer and Stewart 1966), its use today is marginal at best.

Since that time, audit researchers are revisiting Bayesian (Dutta and Srivastava 1993; Srivastava 1996; Srivastava 2011; Srivastava et al, 2012; and Srivastava, Wright, and Mock

2012) and Dempster-Shafer (Gordon and Shortliffe 1985; Perl 1986; Shafer and Srivastava 1990; Srivastava 1995; Srivastava and Shafer 1992; Sun, Srivastava, and Mock 2006; and Srivastava 2011) Frameworks of Belief Functions to assist with analysis of audit evidence uncertainties. The strength of evidence that supports various assertions should be measured and aggregated to finally determine if the assertions are true. This requirement holds true even in a big data environment.

While measurement is one issue, the structure of evidence is another concern because some items of evidence support one assertion or one account while others may provide support to more than one assertion or accounts. Thus, the audit judgment is basically reasoning with evidence in a network of variables, variables being the assertions and accounts, which is also called in the Artificial Intelligence literature "Evidential Reasoning". Gordon and Shortliffe (1985) discuss this approach of the Dempster–Shafer theory of evidence in Rule-Based Expert Systems and Pearl (1986) uses this approach for analyzing causal models under the Bayesian framework, while more recently Srivastava (2011) applies this technique for aggregating audit evidence both under Bayesian and Dempster-Shafer Theory.



More recently, Dempster Shafer theory is applied to assist auditors with the aggregation of evidence to obtain judgements about measuring risks and strengths (Fukukawa and Mock, 2011; and Fukukawa, Mock, and Srivastava 2014). However, it is not clear to what degree the profession feels comfortable and confident with implementing these approaches in the engagement given the prevailing competitive and regulatory pressures. If the PCAOB were to issue guidelines and best practices for applying Belief Functions and Dempster Shafer Probability Theories for the risk assessment phase, then perhaps the engagement team, if familiar with these techniques, would implement them.

In summary, the standards define the task for analytical procedures in each of the three phases, but are non-committal about which techniques auditors should undertake to achieve these objectives. Hence, whether an auditor employs more complex analytics such as Belief Functions or "traditional analytical procedure" techniques such as ratio analysis would seem to depend on the auditor's own knowledge and less so on the standards. It has also been proposed that any adoption by the external audit profession of either advanced analytics or big data would be due to market or business forces exogenous to the firms (Alles 2015). The recent revival of interest in ADA by the firms may be due to these forces.

This brief discussion of BA in contrast to the analytical procedures utilized by auditors in engagements provides many areas for future debate and research. These areas are broadly summarized in the six concerns that follow.



## Six Concerns Relative to Advanced Analytics in the Modern Engagement

The advent of computers, large storage systems, and integrated software has transformed business processes in the first wave of the information age. Their availability has brought to the front the potential of a large number of analytic methods progressively being used in business but still emerging in the external audit domain. The six questions enumerated in the Introduction are discussed in detail in this section.

### 1. *Should New Analytics Be Used in the Audit Process?*

Perhaps this research question could be rephrased as: Should auditors expand their use of analytical procedures beyond that of scanning, ratio and time series analysis, and detailed examination? Are these techniques effective and efficient in a big data context? Basically, these questions emerge and are summarized in Table Two: Should there be more guidance regarding analytic methods in the audit? Do we know enough about these methods that this guidance can be issued? What are the tradeoffs between 100% population tests, sampling, and ad hoc analytics? The standards (PCAOB 2010, AS 1105) suggest that 100% testing would only apply in certain situations, such as: the population consists of a small number of high value elements; the audit procedure that is designed to respond to a significant risk and other means of testing do not provide sufficient evidence; and finally, the audit procedure can be automated effectively and applied to the entire population. The last condition is noteworthy, as current technologies can support automation of basic audit tests such as three-way matching and sampling, in addition to handling fairly large data sets.



The strong emphasis on judgment that exists in auditing is justified by the enormous variety of situations that complex businesses, different industries, international locations, and data structures present to the engagement team, limiting their ability to narrowly pre-set audit rules. Do modern statistical and machine learning methodologies make it possible to automate pre-set rules in many situations in order to perform procedures, derive results, and integrate these in a larger judgment? Can audit findings and judgments be disclosed in more disaggregate manner with the usage of drill-down technologies where the opinion would be rendered and broken down into sub-opinions and quantified in terms of probabilistic estimates (Chesley 1977, 1978)? Can the above be stated in terms of rules implementable in automated audit systems to continuously monitor and drive Audit by Exception (ABE) (Vasarhelyi and Halper 1991)?

(Insert Table Two here)

### 2. *Which of These Methods are the Most Promising?*

The literature on Big Data and Analytics methods applied to business is rich in detail. These methods suggest different staging of the audit (audit re-modularization), changed organization (separate analytic function), changed sequencing, changed tasks, changed timing (continuous, agent driven, exception driven) (Vasarhelyi and Halper 1991) and changed personnel (more literate in IT and data; specialized) making it difficult to evaluate the literature in the context of the external audit. Appelbaum, Kogan, and Vasarhelyi (2016) have recently organized, examined and categorized this body of external audit literature. That study covers more than 300 papers published since the mid-1950's that discuss analytics in at least one phase of the audit. Due to the standards requiring analytical procedures in both the planning and review stages, these two phases are the predominant focus in the literature as is substantive testing and sampling (Appelbaum et al. 2016). Many different analytical techniques are utilized at all phases of the audit, but in an inconsistent manner. Methods that are most promising are categorized as follows:



1) Audit Examinations: transaction tests, ratio analysis, sampling, confirmations, re-performance, CAATS automation;
2) Unsupervised[9]: Clustering, Text Mining, Visualizations, and Process Mining (discovery models);
3) Supervised[10]: Process Mining (process optimization), SVM, ANN, Genetic Algorithms, Expert Systems, Decision Aids, Bagging, Boosting, C4.5 classifiers, Bayesian Theory, Bayesian Belief Networks, Dempster-Shafer Theory Models, Probability theory models;
4) Regression: Logistic, Linear, Time Series, ARIMA, Univariate, Multivariate;

[9] Unsupervised approaches are those techniques that draw inferences from unlabeled or unknown datasets since there is minimal hypothesis of the results based on labeled responses

[10] Supervised approaches are those techniques that draw inferences from labeled or known dataset types, otherwise known as training data

5) Other Statistics: Multi-Criteria Decision Aid, Benford's Law. Descriptive Statistics, Structural Models, AHP, Spearman Rank Correlation Measurements, Hypothesis Evaluations, and Monte Carlo Study/Simulation.

These analytical models range from very simple substantive tests and routines to more complex and predictive techniques requiring significant auditor judgement. The auditor will need to determine what type of analysis gives the best quality and most efficient audit, given the audit task, the assessed audit risk, and the available data. Ideally, the auditor should be able to perform most if not all procedures to more exacting standards in a big data and continuous auditing or monitoring environment using a variety of analytical approaches. Using targeted techniques, auditors would spend less time navigating through insufficient samples and instead, identify and almost immediately examine the transactions of high risk.



Auditors selecting these more complex techniques need to understand them in terms of their benefits and limitations. Furthermore, the tasks of risk assessment, substantive procedures and tests of controls may be different when 100% of the data is examined (Yoon 2016). For example, if auditors are examining 100% of items in the population (PCAOB 2010, AS No. #1105.24), the emphasis and reason for testing internal controls should change. Internal Control testing has been prescribed in the regulations (American Institute of Certified Public Accountants [AICPA] 1997, SAS No. #80) to supplement substantive testing for validating sampling results when auditors have limited access to data. It has been suggested (IAAE 2016 p. 18) that internal controls testing in an Audit by Exception type of environment could provide some assurance regarding data quality.

To summarize the issues of which methods are the most promising (Table Three) given the audit task as defined by the standards: A new environment of assurance is emerging where automation of controls, full population testing, and analytic methods will interplay. Research is needed on modern analytic methods to establish: their applicability in different instances, their cumulative effect, their ability to be formalized, their classification (creation of taxonomies of analytic methods and data structures[11]), and their quantification.

A set of questions arises with the application of analytics that must be tested in the field. Would a safe harbor experimentation (a la XBRL) process be needed for the testing of approaches? Although in the traditional environment a yes/maybe/no attestation is provided, the new proposal provides information of audit results in at least five areas where needed. How would these results be disclosed?

(Insert Table Three here)



### 3. *Where in the Audit Are These Applicable?*

The traditional organization and processes of the audit as defined in the current standards will be affected in many ways by the emerging environment and its disruptive technologies. If some form of Audit by Exception (ABE) (Vasarhelyi and Halper 1991) emerges whereby the audit process is activated by alarms triggered in data streams, and a plethora of new analytics emerge, clearly the sequence of events will be transformed and the applicability of analytic methods expanded. Furthermore, there will be ubiquitous use of techniques such as visualization, and multi-complementary use of many analytic methods. Visualizations are used heavily in business management to explain the results of analysis (Dilla et al. 2010; Kohavi et al, 2004). Many techniques exhibit varying strengths and weaknesses and are more beneficial when applied in

[11] The AICPA has created the Audit Data Standard (Zhang et al. 2012) to guide in the formalization of data to be received in the audit, its classification (into cycles), and its measurement.

combination rather than separately. The sequencing (or simultaneity) of events will change as automated use of data analytics will precede / or coincide with the more traditional audit examination which may progressively be reduced. For example, today the audit engagement typically progresses as shown in Figure Two but is envisioned to eventually innovate to a more Audit By Exception (ABE) approach (Figure Three).

(Insert Figure Two here)

The above process, which drives most current engagements, is sample driven; in a more data driven environment the examination process would be analytically reviewed, audited automatically, and exceptions or outliers would be subsequently examined in detail (Figure Three).



(Insert Figure Three here)



However, in this ABE approach the auditors may face a different challenge: testing all of the transactions may produce thousands of exceptions (Dohrer, McCullough, and Vasarhelyi 2015) if the threshold definition of a material deviation is set too high. That is, the threshold approach for sampling most likely will not work in ABE; the threshold should be more precise to eliminate the “false positive” exceptions. The standards require that all exceptions should be examined (PCAOB 2010, AS No. #2305, AS No. #2315), but this was mandated for sampling (IAAE 2016 p. 17). In an ABE context, if the tests were not configured correctly, there could be an unreasonable number of exceptions to investigate as required. Some auditors have performed additional tests to “explain away” many of exceptions and categorize the resulting few as “Exceptional Exceptions” (Issa et al. 2016). Clearly auditors will need to possess a broad and comprehensive knowledge of analytical techniques in an ABE environment. Furthermore, ABE may be applied to non-financial data as well. Brazel et al (2009) combine financial and non-financial ABEs to assess fraud risk.

The level of automation of the audit, and as discussed before, the availability and comfort with analytical techniques, the competences of the auditor, and the circumstances and assertions of the specific audit process will guide the locus of the application. As such, ABE is a more advanced audit approach, reflecting the confluence of automation, advanced analytics, and revised regulations. Issues that may emerge during this process could be as follows (Table Four): How different are the objectives of Internal Audit and External Audit in the current context (Li et al. 2016)? Isn't there a substantive overlap between business monitoring and real time assurance?



Considering that there is substantive overlap in data analytic needs, are the traditional three lines of defense (Freeman 2015; Chambers 2014) still relevant[12]? Traditional auditing has a retrospective approach, as traditional technologies did not allow for other approaches - can the current environment allow for a prospective look and to what extent? What parts / procedures of the audit are fully or partially automatable? Will the disruptive changes (Christensen 2013) be allowed by the leading audit firms?



Can the key contingencies such as risk assessment and opinion formation in the audit be formalized? In the same line, but extending expanded testing and reporting, should quantitative guidelines be issued for ABE and its structures, and should within period results be disclosed as part of the auditor's report? The succinctness of the traditional report is not necessary any more,

[12] There should be effective risk management functions within a company. These monitoring and assurance functions have been modeled as the "Three Lines of Defense" by the IIA. This model serves as an example, where: 1) the first line of defense represents functions that own or manage the risk; 2) the second line of defense, where there are functions that specialize in risk management and compliance; and 3) the third line of defense, where there are functions that provide assurance

and drill downs on the results of Critical Audit Matters (CAM) examination, their details, and other information is possible.

(Insert Table Four here)

### 4. *Should Auditing Standards Be Changed to Allow / Facilitate These Methods?*

In general, the aforementioned meetings between the AICPA's ASB and the ASEC committee have concluded that the standards do not forbid the usage of analytics, but it can be argued that the standards, and the economics of external audit, make analytics more difficult or in some instances impractical if not nearly impossible to use. For example, audits of financial and insurance industry clients are quite complex and the engagement team may find it impractical within the budgeted hours to conduct any additional analytical techniques beyond the acceptable ratio analysis and sampling. The lack of a more detailed discussion of appropriate analytical techniques within the standards, when placed in the context of a highly competitive business environment, does not encourage the profession to explore new techniques even in the face of big data and automation. The use of more automation and analytics in the engagement, particularly in a big data environment, generates these additional issues (Table Five):



- The economics of the audit is encumbered by a series of anachronistic requirements that are still being enforced by the PCAOB. Consequently, the pricing of the audit, in a competitive environment, leaves little space for additional analytics even if these give stronger assurance of fair representation. Furthermore, what would be the cost versus benefit trade-off with the usage of analytics? Or, would there be a point where the cost of conducting a sample driven audit exceeds that of ABE audit? When would the additional assurance derived from the analytic results justify the cost of their application? Even further, if a certain analytics method is more powerful and uncovers issues that were not previously detected, what would be the liability of the accounting firm, particularly if these issues were also present in the prior years? (Krahel and Titera 2015, p. 418)

- Sampling requires laborious follow ups on abnormalities detected, but in a population of millions or hundreds of thousands there is little to be gained from picking 25 transactions and reviewing them (Dohrer, McCullough, and Vasarhelyi 2015). Do any areas of the modern audit exist where these small judgmental samples still make sense (Elder at al. 2013)? In juxtaposition to the current requirements, would the auditor then need to justify the use of sampling in circumstances where 100% of the data would be available for testing?



The audit research literature itself has been scant regarding auditors' sampling decisions in the context of economic and competitive pressures, regulations about statistical sampling, as well as how to effectively extract meaningful results from the sampling (Elder et al. 2013, p. 103). Auditing standards (PCAOB 2010, AS No. #2315) define sampling as "the application of an audit procedure to less than 100% of the items in an account balance or class of transactions for the purpose of evaluating some characteristic of the balance or class.", The auditor may choose to select all items for testing if the level of sample risk from possible erroneous decisions is too high (AS No. #2315.07).



There is little guidance as to when 100 percent testing would be more appropriate than selecting specific items. In the standards about Audit Evidence (PCAOB 2010, AS No. #1105.22-.29), sampling is not recommended when the data population is small and/or not homogeneous, when there appears to be significant risk, when there are key items that should be examined, when threshold tests should be applied, nor is it suggested when audit procedures can be automated effectively and applied to the whole population. In the standards regarding sampling (PCAOB 2010, AS No. #2315.07), the auditor should weigh the cost and time to examine all of the data versus the perceived degree of uncertainty from sampling and non-sampling risks, and judge accordingly. Consequently, the practice of sampling has become embedded in basic public auditing practice. PCAOB examinations have been very strict favoring sampling against analytical methods.

- Furthermore, Elder et al. (2013) were unaware of any literature that addresses the auditor's decision to use audit sampling of any type (Elder et al. 2013, p. 111) and suggested that future research should address the issues of when sampling would be appropriate and when other types of tests would negate the need for sampling. In response, Yoon (2016) discussed how substantive analytical procedures (SAPs) applied to 100 percent of the data (with the use of computer assisted auditing techniques) could potentially provide a more efficient and effective audit evidence than sampling, particularly in a big data environment. Perhaps for audit engagements where the client is collecting or analyzing all of the transactions and the auditor is using automated audit software, the standards could more clearly establish that 100 percent tests using substantive analytical procedures would provide efficient, sufficient, and appropriate audit evidence.



For example, three way matches used to be performed manually and reviewed manually. Now advanced accounting systems and ERPs perform these automatically. Is this performance audit evidence, new analytics, or just automation? If considered automation, how do the audit standards take this into consideration? Is there a difference between automation and analytic methods? (Dohrer, McCullough, and Vasarhelyi 2015) If such automation is viewed as preventive internal control, then how does it change the balance between control testing and substantive testing in auditing the modern highly automated enterprise environments? Furthermore, the situation will change if fraud is suspected. Simply automating a process does not necessarily mean that the transactions have been correctly processed and that internal controls are operating effectively. The auditor may still need to test the automated system for its reliability by using test data.

- In highly automated accounting systems many analytics or pre-programmed apps will depend on some form of "audit data standard" (Zhang et al. 2012)[13]. These apps will run frequently or constantly (Vasarhelyi and Hoitash 2005). This form of evidence may use external and internal data (Brown-Liburd and Vasarhelyi, 2015) potentially from external sources like social media, thus providing valuable tertiary audit evidence that may be used to complement / replace current tests. Would these need new guidance? Are the current guidelines for traditional audit evidence the same for external or internal big data, particularly social media? What qualities should these data possess in order to provide reliable audit evidence?
- It has been shown (see e.g., Hoitash et al. 2006) that the performance of audit analytics is significantly improved if the models incorporate contemporaneous peer company data. Conceivably, contemporaneous peer company data should be considered as legitimate sources of information for obtaining an understanding of the relevant industry and the client's position, as outlined in the standards for risk assessment and review (PCAOB 2010, AS No. #2110, AS No. #2810). Large public accounting firms typically audit multiple peers in the same industry, and they could create large internal data warehouses to share such data among the engagement teams during the audit. The current strict interpretation of audit client confidentiality rules causes the firms to err on the side of



[13] The AICPA has published online a series of voluntary suggested audit data standards: http://www.aicpa.org/InterestAreas/FRC/AssuranceAdvisoryServices/Pages/AuditDataStandardWorkingGroup.aspx

caution and disallow any sharing of data even though such data would never leave the confines of the firms. New guidance interpreting client data confidentiality as being safeguarded within a firm (and not within an engagement team) and specifically allowing audit client data sharing among different engagement teams would greatly enhance the performance of audit.

(Insert Table Five here)

### 5. *Should the Audit Report Be More Informative?*



PCAOB Release No. 2016-003 proposes, concerning an unqualified opinion, that the audit report disclose "Critical Audit Matters" (if any) in areas such as estimates, audit judgments, areas of special risk, unusual transactions, and other significant changes in the financial statements. This proposal [14]poses a series of interesting questions worthwhile of research (Table 6): Is the level of proposed disclosure adequate in terms of quantification of these critical audit matters or is it falling back into the comfort zone of the traditional auditor? After all, substantive industry resistance was found to the initial proposal (PCAOB, 2013[15] ). Would some of these Critical Audit Matters (CAMs) provide disclosures that are more disaggregate, or more informative than the traditional audit reports?

Could there be preferable schemata of quantification, or quantitative guidelines for estimates, audit judgments, areas of special risk, unusual transactions, or other significant changes in the financial statements? Should these schemata be defined by the standard setters? On a longer

[14] See also Lynne Turner's comments (https://pcaobus.org//Rulemaking/Docket034/ps_Turner.pdf).

[15] PCAOB Release No. 2013-005, August 13, 2013, Docket Matter No. 034, The Auditor's Report on an audit of Financial Statements When the Auditor expresses an Unqualified Opinion. This report discusses the auditor's responsibilities regarding certain other information in certain documents containing audited financial statements and the related auditor's reports and related amendments to the PCAOB standards.

range, if the auditor is using/ relying on ABE should there be a real-time seal or similar device that would allow investors to know on an immediate basis that auditors are monitoring systems and they seem to be doing well[16]?

(Insert Table Six here)

### 6. *What are the Competencies Needed by Auditors in This Environment?*

As mentioned above, the application of analytics in the external audit is attracting substantial attention from practice and academia. EY[17] and the AAA[18] among several others have brought together these two groups for constructive dialogues. Auditor education and familiarity with analytics has been positioned by the standards as a limiting factor regarding which techniques to apply in the engagement (PCAOB 2010, AS No. #2305). Papers such as Tschakert, Kokina, Kozlowski, and Vasarhelyi (2016) and Appelbaum, Schowalter, Sun, and Vasarhelyi (2015) have discussed the issues facing audit education. In general, some conclusions could be drawn:

- Accounting faculties tend not to be prepared to teach analytics.
- There is a widespread general feeling that students are not receptive to learning analytics (however, the feeling is not pervasive – there are some anecdotal reports to the contrary).
- The accounting curriculum is too full to add more IT, statistics, and modeling.
- As the CPA exam does not include these topics, there is little motivation by students for their addition to the curriculum of study.
- Firms will tend / or already have hired specialist groups from non-accounting backgrounds. These groups, as in IT audits (Vasarhelyi and Romero 2014) will be

[16] This type of continuous assurance would work better with some form of more frequent/ continuous reporting.
[17] EYARC 2015, June 17/18 2015, Dallas Texas.
[18] AAA, Accounting is Big Data, September ¾ 2015, New York, New York.

external to the audit team and brought in if the manager of the engagement setting up the audit plan sees fit.

- Practitioners are also not prepared and their internal audit practices have not caught up properly with these issues.
- Firms have been developing software to improve their processes but feel curtailed by the PCAOB examination process.

These factors lead to a series of educational research questions and potential projects that are paradigm changing (Table Seven): If the curriculum is too full, if memorization in the age of google is a different consideration, and if the domain of coverage is too large, then what educational structures and what types of certificates should be used /developed?



Should the CPA profession expand competencies or progressively rely more and more on specialists from other domains, potentially using other (non-CPA firms) to provide these competencies? Should the set of CPE requirements of the profession be reformulated in terms of a life-long-learning approach where new required skills are defined and progressively required in the accountants learning/ competency profile? Who should manage this learning profile, and who should set the requirements? Should there be a much wider set of accounting specializations with coordinated competencies? Should there be quantification of the different types of accountant skills? And some of these acquired through on the job activities and related experience?

(Insert Table Seven here)

## Technology Adoption Issue: Evolution Towards a New Audit Environment of Big Data and Audit Analytics

"It has also been shown that many internal audit procedures can be automated, thus saving costs, allowing for more frequent audits and freeing up the audit staff for tasks that require human judgment." (AICPA, 2015)

It has been proposed in other technology adoption settings that such automation changes are best considered as *evolutionary* instead of *revolutionary* (Kuhn and Sutton 2010). The topics and suggestions mentioned in this paper may seem extensive in scope and massive in undertaking. These issues could serve as either motivators or impediments to the use of big data and audit data analytics (BD/ADA) by the external audit profession.



Ideally, it would seem that the goal for BD/ADA adoption by the profession would be to save costs and attain greater efficiencies and effectiveness in the audit process. However, it is conceivable that impediments exist that would dampen enthusiasm for BD/ADA adoption and these conflicts may be similar to those of other technology initiatives. Here are just a few of the issues that are proposed as being relevant to BD/ADA adoption (Table Eight):



(Insert Table Eight here)

The literature regarding technology adoption is huge in the audit, accounting, and AIS disciplines. This paper does not attempt to synthesize this literature in support of this discussion; instead, a few select papers are highlighted and a very scant outline for BD/ADA adoption is suggested for future research. For instance, the Information Fusion process that Perols and Murthy (2012) propose could be applicable here in the context of BD/ADA adoption. Kuhn and Sutton (2010) present research challenges that could correspond with BD/ADA in the area of regulatory/adoption/judgment and decision making challenges. Likewise, the "messy matters of

Continuous Auditing (CA) adoption" which Hardy (2015) presents may be applicable to ADA/BD.

It has been suggested (Alles et al 2008; Geerts et al 2013) that the transformation of manual processes to that of automation is best accomplished incrementally. Geerts et al (2013) and Dzuranin and Malaescu (2016) provide a framework based on Design Science for such an integration. Vasarhelyi (2013) proposes a four-step process based on the work of Parasurman et al (2000). According to Parasurman et al (2000), human information processing and its evolution from man to machine can be divided into four phases: 1) information acquisition; 2) information analysis; 3) decision selection; and 4) action implementation. In the Alles et al (2008) proposal, each such successive step should be undertaken methodically once benefits from the previous steps have been realized (Figure Four).

(Insert Figure Four about here)

Furthermore, in the Alles et al. (2008) and Dzuranin and Malescu (2016) frameworks, successful change is more likely to occur if the manual process is re-engineered first to support the eventual automation. In the Alles et al. (2008) proposal, the first step of the process cycle is the consideration of the drivers of change and endorsement by management; the second step in the process is the development and the actual implementation of the components that would enable this change; the third step consists of management, or baseline measurement and evaluation of the solution. This process cycle is repeated for every level of automation transformation in an incremental fashion. Such a process cycle approach could also apply as an incremental use of analytics and big data by the public audit profession.

The initial drivers for the use of analytics and big data by external auditors are already in place, with the increasing complexity of client transactions, analytics, and data sources and the subsequent increase of audit risk to the engagement team if analytical procedures are manual and overly simplistic (Alles 2015; Bedard et al, 2008). Firms are already embracing diverse descriptive approaches (Dilla et al, 2010); it could be argued that some practitioners are about to embark on the next phase, the adoption of more predictive analytics. Basically, firms are discovering that manual and simplistic analytical procedures and data sources create an audit which is more likely than not to be inefficient and ineffective in a big data context. Many firms are investigating ways to integrate more advanced analytics in their engagements, but this initiative is progressing cautiously (Alles 2015). It is suggested that many of the research issues discussed here in this paper will need to be examined in the context of an incremental approach, as illustrated in Figure Five. Figure Five illustrates how the process flow as depicted in Figure Four could be integrated incrementally to incorporate advanced analytics and big data into practice.



(Insert Figure Five about here)

This incremental approach may already be observed to some degree in the audit process – while some manual procedures have been automated, other audit procedures have not. Many audit tests may be conducted on 100% of the test population using Computer Assisted Auditing Techniques (CAATs) software packages (Wang and Cuthbertson 2015). These CAATs can perform analytics very efficiently and quickly and can interface and link easily to the client's system. Although not all CAATs software packages are equipped to handle big data, this limitation will eventually be solved. CAATs are used by auditors on many engagements for GL tests, three way matches, detail tests, and sampling for example. However, these tests do not run automatically but are manually selected by the engagement team. The auditor selects which analytical procedures or tests to run and attributes to examine in the tests of assertions for a particular audit objective.

## Potential Research Issues and Opportunities

Modern audit engagements often involve examination of clients that are using big data and analytics to remain competitive and relevant in today's business environment. Client systems now create and acquire big data and apply advanced analytics to generate intelligence for decision making. However, the public accounting profession is still bound by regulations that may have been applicable years ago but whose relevance should be re-examined today in this modern business environment. There are numerous issues surrounding the standards, practice, and theory of audit data analytics that have emerged from these rapidly evolving different corporate systems and which have not been addressed. This paper highlights six general areas of such concerns and now provides a broad review and collection of additional critical ADA issues that challenge the public auditing profession today.



### *Research Questions*

Many of the issues and sections reiterated similar research questions. Additional research questions are now presented that seem to be also important to answer for audit data analytics to succeed in gaining widespread practical acceptance. Also, quantification of many audit processes and judgements may be called for with the heightened use of advanced analytics and big data.

1. How can analytics methods be used to create accurate expectation models for generating predictions to compare with actual accounting numbers? How should allowable variances of predictions be chosen (Bumgarner and Vasarhelyi 2015)?

Expectation models should be examined in greater depth with the application of more advanced analytics. These more advanced approaches, combined with big data, may establish a narrower variance of prediction.

2. What properties make a particular ADA technique more or less appropriate for a particular audit function? There is a wide range of techniques appropriate for each audit phase, given the client particularities, environment, and industry. The categorization of appropriate techniques given certain client conditions is proposed as an External Audit Framework (EAA) in Appelbaum, Kogan, and Vasarhelyi 2016.
3. What types of "suspicion functions"[19] should be utilized in a preventive audit[20] as contrasted with transaction or account reviews? The weighting of characteristics of variables in linear suspicion functions may be impacted by ADAs such as expert systems, Bayesian Belief systems, probability models and Exceptional Exceptions (Issa, Brown-Liburd, and Kogan 2016).
4. How should the assurance function be reorganized to better use ADA? The assurance function is broader than that of financial statement auditing. Since assurance services should improve the quality of information for decision makers, the quality (relevance and reliability) of data is still paramount. The assurance function may be reorganized in a broader format than the engagement, but standards must continue to be issued.
5. How should audit standards and processes be modified to enable and encourage the utilization of ADA? The standards should be modified to suggest techniques that are acceptable for each phase of the audit, given certain engagement contexts. For



[19] A "suspicion function" is a linear multivariate equation that gives weights to characteristics of variables and analytical evidence to estimate its probability of being fallacious.
[20] Bumgarner and Vasarhelyi (2015) break down audit as retroactive and predictive. A predictive audit may be preventive (when a suspicion score is large, a transaction is blocked for review), or just predictive to set up a standard of comparison.

example, perhaps sampling should be modified for client engagements where 100% of the data is electronically collected and available to the auditor. In this context, ABE or Exceptional Exceptions (Issa, Brown-Liburd, and Kogan 2016) should be acceptable by the standard setters in lieu of sampling where appropriate. Additionally, the standards regarding data as audit evidence should also be examined in the context of electronic data and big data – external evidence may not be as reliable in this case (Appelbaum 2016; Brown-Liburd and Vasarhelyi 2015).

6. What is the proper way of validating expectation models for ADA? Should this validation be carried out for each audit client separately, or can it be extrapolated from one client to all the other clients in the same industry? Validation of models may be established over time by auditors for continuing clients and also for the auditors' own industry expertise. As part of interim activities, updated information could be fed into prescriptive analytical models that over time attain greater accuracy. The standards could also feasibly provide guidance specific for certain industries.
7. What additional verification processes would be desirable with the extant analytic technology? Verification processes and validation remain as open issues with ADA integration in the engagement. Over time, with continuing audit clients, it is likely that prescriptive analytics will become more precise.
8. How can the concept of "accuracy[21]" be defined for ADA? Is it necessary to encourage the use of substantive audit analytics? The concept of accuracy may be formally and quantitatively defined with the use of ADA. Auditor judgement is still necessary, even with advanced analytical techniques.



[21] Acceptable relative error in engineering, materiality in accounting.

*Evolution Towards Quantification of the Audit*

Radical changes in analytics, information processing, and information distribution technologies have allowed assurance that can be continuous (Vasarhelyi and Halper 1991), predictive (Kuenkaikaew and Vasarhelyi 2013), prescriptive (Holsapple et al. 2014), and even facilitate automatic data correction (Kogan et al. 2014). These techniques are intrusive, create transparency, and maybe also some competitive impairment if all the details are disclosed, and generate substantive concerns by the auditee. The public good tradeoff of increased information disclosure versus economic interest of agents is a complex issue and its equilibrium may take many years to be reached, just to be disturbed by additional disruptive technologies.



The increased amount of data available and the progressive ability to discover variances, understand aggregate content, and to predict trends has clearly created an equilibrium misbalance that is becoming larger and larger. Quantification can increase the value of information both internally and externally, but it decreases information asymmetry which is very threatening for agents (managers) and principals. A common thread of research questions relative to quantification were raised throughout this paper and are elaborated upon here:

- Do modern disclosure and statistical methodologies make it possible to, in certain cases, automate pre-set rules in order to perform procedures, derive results, and integrate these in a larger judgment? Such an approach is necessary for “close to the event continuous auditing” (Vasarhelyi and Halper, 1991) that is progressively been made necessary due to large electronic data streams exogenous and endogenous to the company.
- Research is needed on modern analytic methods, their applicability in different instances, their cumulative effect, their ability to be formalized, their classification (creation of

taxonomies of analytic methods and data structures )[22], and their quantification. Traditional audit is backward looking due to the limitations of manual review and storage procedures. These modern analytic methods allow for the detection and prevention of propagation along downstream systems of potential faults (Kogan et al., 2014). These characteristics would force new corporate procedures of timely midstream error correction that do not exist in extant systems. These emerging procedures will be difficult to conceptualize from the point of view of "lines of defense" (IIA, 2013[23]; Freeman 2015; Chambers 2014), as they potentially make such lines blurred.

  - If a midstream process detects faults and activates an error correction process that is a mix of human judgment and automatic correction, is this an audit or a control process? Does this distinction make sense in the modern world of automation?
  - If a continuous audit layer detects "serious faults" (Vasarhelyi and Halper, 1991) and stops a system, is this layer a part of operations, control, or audit?

- Can audit findings and judgments be disclosed in more disaggregate manner with the use of drill-down technologies where the opinion would be rendered and broken down into sub-opinions and quantified in terms of probabilistic estimates (Chesley 1975, 1976, 1977)[24]. The issue of additional information disclosure in audit opinion is considered in the new PCAOB proposal and does not directly address the type of precision that disaggregation would allow. Turner (2014, p5) in the aforementioned comments to the



---

[22] The AICPA has created the Audit Data Standard (Zhang, Yang, & Appelbaum, 2015) to guide in the formalization of data to be received in the audit, its classification (into cycles), and its measurement.

[23] "The tree lines of defense in effective risk management and control", White paper, The Institute of Internal auditors, January 2013.

[24] More detailed and quantitative audit reports are being progressively disclosed. For example, in the Netherlands (annual report of Aegon NV, 2015, p309) there is disclosure of the threshold of materiality EUR 65 million and the statement that "We agree with the audit committee that we would report to the misstatements identified during the audit about EUR 4 million (2014: EUR 4 million) as well as misstatements below that amount that, in our view, warranted reporting for qualitative reasons." Quantitative assessments are also made of coverage and other variables as well as a much more detailed discussion of governance controls and procedures.

PCAOB states "*it is clear that some oppose any disclosure of information not previously disclosed by management. But such an approach defies common sense and is intended to obfuscate and avoid disclosing the information investors want. I urge the Board to reject such an approach as it will result in disclosures that are not worth the time or cost… investors wanted…information that is not "filtered through management" (adapted).*" Improved stochastic estimates in disclosure, although not deterministic statements that create illusory comfort for the readers, may be the solution for this dilemma. Research here is urgently needed.

- Should quantitative guidelines be issued for ABE and its structures, and should within period results be disclosed as part of the auditor's report? A technological continuous audit allows for continuous monitoring and remarkable (not necessarily material) exception reporting. Should these exceptions be reported to all stakeholders (e.g. investors, suppliers, etc.) or only to select stakeholders? Should some of these exceptions be linked to smart contracts (Kosba et al. 2015) that automatically would execute a pre-agreed (e.g. covenant condition) action? A continuous assurance environment requires that events of substance, that can be predicted, be diagnosed and some action executed. As the combinatorics of these events is almost infinite, progressively more and more complex audit (and operational) judgments will be necessary, occupying auditors but necessarily changing their skill requirements (Tschakert et al. 2016).

## Conclusion

This paper contributes to the literature by discussing the concerns facing the external audit profession as business moves towards big data and advanced analytics for many aspects of operations and decision making. These suggested research issues, along with various proposals towards greater use of big data and analytics will hopefully encourage and inspire ideas and research that is useful for professionals, regulators, and researchers. Although many concerns are reviewed, many are also not mentioned. It is expected that as research and findings evolve in this domain, some concerns will become less important while others many unexpectedly gain urgency. However, the emerging overall importance that big data and advanced analytics present to the public audit profession cannot be ignored.

Most of the research discussion is focused within audit standards setting, audit practice issues, and the development of better audit data analytics. While these areas all lend themselves to empirical research in auditing, this paper has been oriented more towards theory and practice. Theoretical proposals and questions as to how analytics and big data will be affecting the external audit have been discussed. Future empirical research is required to validate these theoretical approaches before their adoption by the audit profession.



In part, this paper is motivated by a vision as to how audit data analytics could enhance or replace certain auditor conducted procedures. But, perhaps there are other views that could be regarded as more research friendly and perhaps more realistic to a more real-time use of audit data analytics. Two specific areas seem to present easily integrated opportunities.

First, in the background discussion of analytical procedures and the standards, AS No. #2305.04 mentions how analytical procedures are used in the substantive testing phase to obtain evidence. The discussion focuses on how ADA might replace substantive testing and then is elaborated on at later points by focusing on adjusting audit standards to substitute substantive tests with ADAs. What seems more reasonable in the current PCAOB/legal liability environment is that perhaps ADAs are better used to focus auditors' substantive testing.

Consider for example the Jans et al. (2014) paper with the application of process mining. This paper details the use of process mining on a 100% test of the transactions to find the anomalies in the sample where controls fail in the processing of 26,185 POs. A series of process mining tests (a type of ADA) narrows the sample of anomalies down to the highest risk scenarios which exemplify high rates of violations among individuals and small networks of people working together. Possibly this could be regarded is the perfect example of how ADAs can be used for more focused audit testing. This is but a small example of how archival researchers may be able to contribute to the research stream through analysis of big data sets with statistical procedures and/or machine learning techniques to improve the efficiency in targeting substantive audit tests to better identify high risk areas.



The second area presents another aspect to the general discussion on education issues . Possibly additional attention should be focused on what competencies auditors need in this new environment and how the auditor potentially can be a valuable partner in the use of ADA/BA. There is a rich body of literature on industry knowledge, auditors' abilities to recognize patterns and potential irregularities, and on expertise in general.

It seems that a major research thrust should perhaps be how this expertise and professional judgment can be leveraged to develop and use more effective ADA/BA strategies during an audit and to keep the auditor relevant in the tailoring of ADA processes to a given client's business processes. Ultimately the research focus would be more on the development of audit experts that are both good auditors and good data scientists. Is this possible? Can a good audit-focused data scientist produce better results than standardized ADAs? These ideas may perhaps provide behavioral empiricists with additional potential research opportunities to pursue.

In conclusion, big data and business analytics are dramatically changing the business environment and the capabilities of business processes. Business functions are changing, business capabilities are being added, anachronistic business functions are being eliminated, and most of all, processes are being substantially accelerated. The same should occur within the external audit or assurance function, its rules need to be changed, its steps evolved, automation integrated into its basic processes, and its timing should become almost instantaneous in predictive, prescriptive and preventive analytical modes.

## References

Alles, M. G. 2015. Drivers of the use and facilitators and obstacles of the evolution of Big Data by the audit profession. *Accounting Horizons* 29 (2): 439-449.

Alles, M., G. Brennan, A. Kogan, and M.A. Vasarhelyi. 2006. "Continuous monitoring of business process controls: A pilot implementation of a continuous auditing system at Siemens", *International Journal of Accounting Information Systems* 7, no. 2, pp. 137-161.

Alles, M. G., A. Kogan, and M.A. Vasarhelyi. 2008. Audit Automation for Implementing Continuous Auditing: Principles and Problems. Working paper, Rutgers University. Newark, N.J. USA.

American Institute of Certified Public Accountants (AICPA). 1972. *Responsibilities and Functions of the Independent Auditor*. SAS No. 1, Section 110. New York, NY. Available at: http://www.aicpa.org/Research/Standards/AuditAttest/DownloadableDocuments/AU-00110.pdf

American Institute of Certified Public Accountants (AICPA). 1997. *Amendment to SAS no. 31, Evidential Matter*. SAS No. 80. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2012a. *Analytical Procedures.* AU-C Section 520, Source: SAS No. 122. New York, NY. Available at: http://www.aicpa.org/Research/Standards/AuditAttest/DownloadableDocuments/AU-C-00520.pdf

American Institute of Certified Public Accountants. (AICPA). Audit Sampling Committee. 2012b. *Audit Sampling: Audit Guide*. New York, NY: AICPA.

American Institute of Certified Professional Accountants (AICPA). 2015. *Audit Analytics and Continuous Audit: Looking Toward the Future,* New York, NY

Appelbaum,D., D.S.Showalter, T. Sun, and M.A. Vasarhelyi. 2015. “Analytics Knowledge Required of a Modern CPA in this Real-Time Economy: A Normative Position”. *Presented at the Accounting Information Systems Educator Conference, June 25-28, 2015.* Colorado Springs

Appelbaum,D., A. Kogan, and M. A. Vasarhelyi. 2016. “Analytics in External Auditing: A Literature Review”. Working paper, Rutgers Business School, Newark N.J. USA.

Appelbaum, D. 2016. Securing Big Data Provenance for Auditors: The Big Data Provenance Black Box as Reliable Evidence. Working paper, Rutgers Business School, Newark N.J.

Atzori, L., A. Lera, and G. Morabito. 2010. The Internet of things: A survey. Computer Networks 54 (15): 2787–2805.

Bedard, J.C., D.R. Deis, M.B. Curtis, and J.G. Jenkins. 2008. Risk Monitoring and control in audit firms: A research synthesis. *Auditing; A Journal of Practice & Theory*, Vol. 27, No. 1 (May), 187-218.

Brazel, Joe., Keith Jones, and Mark Zimbelman. 2009. Using Nonfinancial Measures to Assess Fraud Risk. *Journal of Accounting Research* (December): pp. 1135-1166.

Brown-Liburd, H., and M.A. Vasarhelyi. 2015. Big Data and Audit Analytics. *Journal of Emerging technologies in Accounting*, Forthcoming.

Bumgarner,N. and M.A Vasarhelyi. 2015. Auditing—A New View. *AUDIT ANALYTICS*, p.3.

Byrnes, P. 2014. Developing Automated Applications for Clustering and Outlier Detection: Data Mining Implications for Auditing Practice, PhD Dissertation, Rutgers Business School, Continuous Audit and reporting Lab, Newark, NJ, 2014

Chambers, A., 2014. New guidance on internal audit–an analysis and appraisal of recent developments. *Managerial Auditing Journal*, *29*(2), pp.196-218.

Chesley, G. R. (1975). Elicitation of subjective probabilities: a review. *Accounting Review*, 325-337.

Chesley, G. R. (1977). Subjective Probability Elicitation: The Effect of Congruity of Datum and Response Mode on Performance. *Journal of Accounting Research*, 1-11.

Chesley, G. R. (1978). Subjective probability elicitation techniques: A performance comparison. *Journal of Accounting Research*, 225-241

Christensen, C. 2013. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Review Press.

Church, B. K., J.J.McMillan and A. Schneider. 2001. Factors affecting internal auditors' consideration of fraudulent financial reporting during analytical procedures. *Auditing: A Journal of Practice & Theory*, 20(1), 65-80

Cukier,K., and V. Mayer-Schoenberger. 2013. Rise of Big Data: How it's Changing the Way We Think about the World, The. *Foreign Affairs*, *92*, p.28

Daroca,F. P., and W.W. Holder. 1985. The use of analytical procedures in review and audit engagements. *Auditing-A Journal of Practice & Theory*, 4(2), 80-92

Davenport, T. H., and L. K. Johnson. 2008. Competing on Analytics: The New Science of Winning. *Harvard Business School Press.*

Davenport, T. H., and J. Kim. 2013. Keeping Up with the Quants. *Harvard Business Review Press,* USA.

Delen,D., and H. Demirkan. 2013. Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359-363

Dilla, W., Janvrin, D.J. and Raschke, R., 2010. Interactive data visualization: New directions for accounting information systems research. *Journal of Information Systems*, *24*(2), pp.1-37.

Dohrer, R., M.A. Vasarhelyi, and P.McCollough. 2015. Audit Data Analytics. Presentation delivered to the IAASB, September 23.

Domingos, P. 2012. A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78-87

Dutta, S. K., and R. P. Srivastava. 1993. Aggregation of Evidence in Auditing: A Likelihood Perspective. *Auditing: A Journal of Practice and Theory*, Vol. 12, Supplement: 137-160.

Dzuranin, A. C., and I. Malaescu. 2016. The Current State and Future Direction of IT Audit: Challenges and Opportunities. *Journal of Information Systems.* Vol. 30, No. 1. Spring 2016, pp. 7-20.

Elder, R. J., A.D. Akresh, S.M. Glover, J.L. Higgs, and J. Liljegren. 2013. Audit sampling research: A synthesis and implications for future research. *Auditing: A Journal of Practice & Theory*, *32*(sp1), 99-129.

Evans, J. R., and C. H. Lindner. 2012. Business Analytics: The Next Frontier for Decision Sciences. *Decision Line,* 43 (2), pp. 4-6.

Freeman, S. 2015. Special report: Engaging lines of defense. *Freeman*, *31*(4), 21.

Fukukawa, H. and T. J. Mock. 2011. Audit Risk Assessments Using Belief versus Probability. Auditing: A Journal of Practice & Theory, Vol. 30, No. 1, February 2011: pp. 75–99.

Fukukawa, H, T. J. Mock, and R. P. Srivastava. 2014. Assessing the Risk of Fraud at Olympus and Identifying an Effective Audit Plan. *The Japanese Accounting Review,* Vol. 4, pp. 1-27 (Invited).

Geerts, G. L., L. E. Graham, E. G. Mauldin, W. E. McCarthy, and V. J. Richardson. 2013. Integrating information technology into accounting and practice. *Accounting Horizons.* 27 (4): 815-840.

Glover, S.M., Prawitt, D.F. and Drake, M.S., 2014. Between a Rock and a Hard Place: A Path Forward for Using Substantive Analytical Procedures in Auditing Large P&L Accounts: Commentary and Analysis. *Auditing: A Journal of Practice & Theory, 34*(3), pp.161-179.

Gordon J, Shortliffe EH. Method for managing evidential reasoning in a hierarchical hypothesis space. Artificial Intelligence; 1985, pp. 323–57.

Hardy, C.A. and Laslett, G., 2014. Continuous Auditing and Monitoring in Practice: Lessons from Metcash's Business Assurance Group. *Journal of Information Systems*, *29*(2), pp.183-194.

Hoitash, R., Kogan, A. and Vasarhelyi, M.A., 2006. Peer-based approach for analytical procedures. *Auditing: A Journal of Practice & Theory*, *25*(2): pp.53-84.

Holsapple,C., A.Lee-Post, and R. Pakath. 2014. A unified foundation for business analytics. *Decision Support Systems*, 64: 130-141

Issa,H., H.Brown-Liburd, and A.Kogan. 2016. *Identifying and Prioritizing Control Deviations Using A Model Derived from Experts' Knowledge* (Working Paper). Rutgers, NJ: Rutgers University.

Jans, M., Alles, M.G. and Vasarhelyi, M.A., 2014. A field study on the use of process mining of event logs as an analytical procedure in auditing. *The Accounting Review*, *89*(5), pp.1751-1773.

Kobelius, J. 2010. The Forrester Wave: Predictive Analytics and Data Mining Solutions. Forrester Research, Inc. USA.

Kogan, A., Alles, M. G., Vasarhelyi, M. A., & Wu, J. (2014). Design and evaluation of a continuous data level auditing system. *Auditing: A Journal of Practice & Theory*, *33*(4): 221-245.

Kohavi, R., L. Mason, R. Parekh, and Z. Zheng. 2004. Lessons and Challenges from Mining E-Commerce Data. *Machine Learning*, 57 (1-2): 83-113.

Kosba, A., Miller, A., Shi, E., Wen, Z., & Papamanthou, C. (2015). *Hawk: The blockchain model of cryptography and privacy-preserving smart contracts*. Cryptology ePrint Archive, Report 2015/675, 2015. http://eprint. iacr. org.

Krahel J.P. and W.R. Titera. 2015. Consequences of Big Data and Formalization on Accounting and Auditing Standards. *Accounting Horizons*, 29 (2): pp 409-422.

Kuenkaikaew, S., & Vasarhelyi, M. A. (2013). The Predictive Audit Framework. *International Journal of Digital Accounting Research*. Vol. 13: 37-71.

Kuhn, J. R., and S. G. Sutton. 2010. Continuous Auditing in ERP System Environments: The Current State and Future Directions. *Journal of Information Systems,* Vol. 24, No. 1: pp. 91-112.

Lee,M., M.Cho, J.Gim, D.H.Jeong, and Jung, H. 2014. Prescriptive Analytics System for Scholar Research Performance Enhancement. In *HCI International 2014-Posters' Extended Abstracts* (pp. 186-190). Springer International Publishing.

Li,H., J.Dai, T.Gershberg, and M.A.Vasarhelyi. 2015. Understanding Usage and Value of Audit Analytics in the Internal Audit: An Organizational Approach. Working paper, Continuous Auditing and Reporting Laboratory, 2013.

Liddy,J. P. 2015. The Future of Audit. Available at: www.forbes.com/sites/realspin/2014/08/04/the-future-of-audit/

Liu, Q .2014. The application of exploratory Data Analysis in Auditing. PhD Dissertation, Rutgers Business School, Continuous Audit and reporting Lab, Newark, NJ, 2014

Montgomery, R.H. 1919. Auditing Theory and Practice, The Ronald Press, 1919 2nd edition, New York

Nearon, B.H., 2005. Foundations in auditing and digital evidence. *The CPA Journal*, *75*(1): p.32.
Nelson, K.M., Kogan, A., Srivastava, R.P., Vasarhelyi, M.A. and Lu, H., 2000. Virtual auditing agents: the EDGAR Agent challenge. *Decision Support Systems*, *28*(3), pp.241-253.

Pearl J. Evidential reasoning using stochastic simulation of causal models. Artificial Intelligence; 1986. P. 245–57.

Perols,J., and B. Lougee. 2011. The Relation between earnings management and financial statement fraud. *Advances in Accounting*, incorporating Advances in Accounting 27 (2011): pp 19-53

Perols, J.L. and Murthy, U.S., 2012. Information fusion in continuous assurance. *Journal of Information Systems*, *26*(2), pp.35-52.

Public Company Accounting Oversight Board (PCAOB). 2010. *Audit Evidence.* Auditing Standard (AS) No. 1105. Washington, D.C: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2010. *Identifying and Assessing Risks of Material Misstatement.* Auditing Standard (AS) No. 2110, Washington D.C.: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2010. *Substantive Analytical Procedures.* Auditing Standard (AS) No. 2305. Washington D.C.: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2010. *Audit Sampling.* Auditing Standard (AS) No. 2315. Washington D.C.: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2010. *Evaluating Audit Results.* Auditing Standard (AS) No. 2810. Washington, D.C.: PCAOB.

Schneider,G., J.Dai, D.Janvrin, K.Ajayi, and R.L. Raschke. 2015. Infer, Predict, and Assure: Accounting Opportunities in Data Analytics. *Accounting Horizons* Vol. 29, No. 3: 719-742.

Shafer, G. and R. P. Srivastava. 1990. The Bayesian And Belief-Function Formalisms: A General Perspective for Auditing. *Auditing: A Journal of Practice and Theory*, Supplement: 110-148.

Song,S. K., D.J. Kim, M. Hwang, J. Kim, D.H. Jeong, S. Lee, and W.Sung. 2013. Prescriptive Analytics System for Improving Research Power. In *Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on* (pp. 1144-1145). IEEE.

Song,S. K., D.H. Jeong, J.Kim, M. Hwang, J. Gim, and H. Jung. 2014. Research Advising System based on Prescriptive Analytics. In *Future Information Technology* (pp. 569-574). Springer Berlin Heidelberg.

Srivastava, R. P. 1995. The Belief-Function Approach to Aggregating Audit Evidence. *International Journal of Intelligent Systems*, Vol. 10, No. 3, March: 329-356.

Srivastava, R. P. 1996. Discussion of An Empirical Test of Bentham's Theory of the Persuasiveness of Evidence," *Auditing: A Journal of Practice and Theory*, Supplement, Vol. 15: 23-32.

Srivastava, R. P. 2011. An Introduction to Evidential Reasoning for Decision Making under Uncertainty: Bayesian and Belief Functions Perspectives. *International Journal of Accounting Information Systems*, Vol. 12: 126–135.

Srivastava, R. P., and G. Shafer. 1992. Belief-Function Formulas for Audit Risk. *The Accounting Review*, April: 249-283.

Srivastava, R. P., T. J. Mock, and L. Gao. 2011. The Dempster-Shafer Theory of Belief Functions for Managing Uncertainties: An Introduction and Fraud Risk Assessment Illustration. *Australian Accounting Review*, Volume 21, Issue 3, pp. 282–291.

Srivastava, R. P., T. J. Mock, and J. Turner. 2009. Bayesian Fraud Risk Formula for Financial Statement Audits. *ABACUS*, Vol. 45, No. 1, pp. 66-87

Srivastava, R. P., T. J. Mock, K. Pincus, and A. Wright. 2012. Causal inference in auditing: A framework. *Auditing: A Journal of Practice and Theory,* Vol. 31, No. 3, pp. 177-201.

Srivastava, R.P., A. Wright, and T. Mock. 2002. Multiple Hypothesis Evaluation in Auditing. *Journal of Accounting and Finance* (Australian), Volume 42, No. 3, November: 251-277.

Stewart, T. 2015. Data analytics for financial-statement Audits, Chapter 5 in AICPA, Audit Analytics and Continuous Audit: Looking Toward the Future, American Institute of Certified Public accountants. New York, N.Y.

Stringer, K.W., and T.R.Stewart. 1986. *Statistical techniques for analytical review in auditing*. Ronald Press.

Sun, L., R. P. Srivastava, and T. Mock. 2006. An Information Systems Security Risk Assessment Model under Dempster-Shafer Theory of Belief Functions. *Journal of Management Information Systems,* Vol. 22, No. 4: 109-142.

Tabor,R. H., and J.T. Willis.1985.Empirical evidence on the changing role of analytical review procedures. *Auditing-a Journal of Practice & Theory*, 4(2), 93-109

Tschakert, N., J. Kokina, S.Kozlowski, and Vasarhelyi, M.A. 2016. CPAs and Data Analytics. The Journal of Accountancy, forthcoming.

Tukey, J. W. 1977. Exploratory data analysis. Reading, MA: Addison-Wesley.

Tukey, J. W. 1980. We need both exploratory and confirmatory. *The American Statistician*, 34: 23-25

Vasarhelyi,M.A., and F.B.Halper. 1991. The continuous audit of online systems. *Auditing: A Journal of Practice & Theory*, *10*(1), 110-125

Vasarhelyi, M.A. 2013. Formalization of Standards, Automation, Robots, and IT Governance. *Journal of Information Systems,* Vol. 27, No. 1 Spring 2013, pp. 1-11.

Vasarhelyi, M. A. and S. Romero. 2014. Technology in audit engagements: a case study. *Managerial Auditing Journal*, *29*(4), pp.350-365

Vasarhelyi, M.A. 2015. The new scenario of business processes and applications on the digital world. Working Paper, CarLab.

Vasarhelyi, M.A., A. Kogan, and B.M. Tuttle. 2015. Big data in accounting: An overview. *Accounting Horizons*, *29* (2): 381-396.

Wang, T., and R. Cuthbertson. 2015. Eight Issues on audit data analytics we would like researched. *Journal of Information Systems* 29 (1): 155-162.

Warren Jr, J.D., Moffitt, K.C. and Byrnes, P., 2015. How Big Data will change accounting. *Accounting Horizons*, *29*(2), pp.397-407.

Yoon, K. (2016). “Big Data as Audit Evidence: Utilizing Weather Indicators.” Chapter 3 of the dissertation titled *Three Essays on Unorthodox Audit Evidence*, Rutgers University, Newark N.J.

Yue,D., X.Wu, Y.Wang, Y.Li, and C.H.Chu. 2007. A review of data mining-based financial fraud detection research. In Wireless Communications, Networking and Mobile Computing, 2007. WiCom. International Conference on (pp. 5519-5522). IEEE.

Zhang,L., A.R.Pawlicki, D.McQuilken, and W.R.Titera. 2012. The AICPA assurance services executive committee emerging assurance technologies task force: The audit data standards (ADS) initiative. *Journal of Information Systems*, *26*(1), 199-205.

Zhang,J., X.Yang, and D.Appelbaum. 2015. Toward Effective Big Data Analysis in Continuous Auditing. *Accounting Horizons* 29 (2) 469-476.

# APPENDIX A:

## Potential application areas of EDA in Clarified Statements on Audit Standards issued by AICPA

| Audit Standard | Potential application areas of EDA |
|---|---|
| AU-C sec. 240<br><br>Consideration of Fraud in a Financial Statement Audit | **.22** Based on analytical procedures performed as part of risk assessment procedures,[8] the auditor should evaluate whether ***unusual or unexpected*** relationships that have been identified indicate risks of material misstatement due to fraud. To the extent not already included, the analytical procedures, and evaluation thereof, should include procedures relating to revenue accounts. (Ref: par. A24–. A26 and .A46)<br><br>**.27** The auditor should treat those assessed risks of material misstatement due to fraud as significant risks and, accordingly, to the extent not already done so, the auditor should ***obtain an understanding*** of the entity's related controls, including control activities, relevant to such risks, including the evaluation of whether such controls have been suitably designed and implemented to mitigate such fraud risks. (Ref: par.A36–.A37)<br><br>**.32** Even if specific risks of material misstatement due to fraud are not identified by the auditor, a possibility exists that management override of controls could occur. Accordingly, the auditor should address the risk of management override of controls apart from any conclusions regarding the existence of more specifically identifiable risks by designing and performing audit procedures to, etc.<br><br>**a.** test the appropriateness of journal entries recorded in the general ledger and other adjustments made in the preparation of the financial statements, including entries posted directly to financial statement drafts. In designing and performing audit procedures for such tests, the auditor should (Ref: par. .A47–.A50 and .A55)<br>i. obtain an ***understanding*** of the entity's financial reporting process and controls over journal entries and other adjustments, [12] and the suitability of design and implementation of such controls;<br>ii. make inquiries of individuals involved in the financial reporting process about inappropriate or ***unusual*** activity relating to the processing of journal entries and other adjustments; , etc..<br>**c.** evaluate, for significant transactions that are ***outside the normal*** course of business for the entity or that otherwise appear to be ***unusual*** given the auditor's ***understanding*** of the entity and its environment and other information obtained during the audit, whether the business |

rationale (or the lack thereof) of the transactions suggests that they may have been entered into to engage in fraudulent financial reporting or to conceal misappropriation of assets. (Ref: par. .A54)

**.A21** Those charged with governance of an entity oversee the entity's systems for monitoring risk, financial control, and compliance with the law. In some circumstances, governance practices are well developed, and those charged with governance play an active role in oversight of the entity's assessment of the risks of fraud and of the relevant internal control. Because the responsibilities of those charged with governance and management may vary by entity, it is important that the auditor ***understands*** the respective responsibilities of those charged with governance and management to enable the auditor to obtain an understanding of the oversight exercised by the appropriate individuals.

**.A37** It is, therefore, important for the auditor to ***obtain an understanding*** of the controls that management has designed, implemented, and maintained to prevent and detect fraud.

**.A49** When identifying and selecting journal entries and other adjustments for testing and determining the appropriate method of examining the underlying support for the items selected, the following matters may be relevant:

- *The characteristics of fraudulent journal entries or other adjustments*. Inappropriate journal entries or other adjustments often have unique identifying characteristics. Such characteristics may include entries (*a*) made to ***unrelated, unusual, or seldom-used*** accounts***;*** (*b*) made by individuals who typically do not make journal entries; (*c*) recorded at the end of the period or as post closing entries that have little or no explanation or description; (*d*) made either before or during the preparation of the financial statements that do not have account numbers; or (*e*) containing round numbers or consistent ending numbers.
- *The nature and complexity of the accounts*. Inappropriate journal entries or adjustments may be applied to accounts that (*a*) contain transactions that are ***complex or unusual*** in nature, (*b*) contain significant estimates and period-end adjustments, (*c*) have been prone to misstatements in the past, (*d*) have not been reconciled on a timely basis or contain unreconciled differences, (*e*) contain intercompany transactions, or (*f*) are otherwise associated with an identified risk of material misstatement due to fraud. In audits of entities that have several locations or components, consideration is given to the need to select journal entries from multiple locations.

**Tables and Figures**

Table One:

**ISSUES REGARDING BIG DATA AS AUDIT EVIDENCE**

| **Challenge of Big Data** | **Recommendation** |
|---|---|
| How can the availability of big data sets be used to enhance analytics? | Research can suggest analytical techniques that take advantage of big data and evaluate how they improve audit effectiveness and/or efficiency. |
| Can the volume of data compensate for uncertain or lower quality of data? | Studies should be conducted that determine if there exists an upper threshold of data volume, exceeding which could compensate for lower data quality. A framework for data value should be generated. |
| How can the amount of audit evidence provided by analytics in a big data context be measured? | Research should re-examine the concept of whether evidence derived from analytics is “soft” and a quantitative reliability scoring system developed for all types of audit evidence. This score could then be integrated in the overall risk assessment. |
| How can big data evidence be aggregated with other types of audit evidence in a methodologically sound way? | This research question can be integrated with that of the data measurement system. |
| How can quantitative measures be used to provide support for the auditor’s judgment about the sufficiency of audit evidence? | This research question can be integrated with that of the data measurement system. |
| Alterability: How can the auditor be assured that the data has not been altered? | Research examining various tests for the assertion of accuracy in a big data context should be conducted. |
| Credibility: How can the auditor be assured of the controls surrounding the generation of big data external to the client? | Research examining/suggesting certain verifications of controls should be undertaken. |
| Completeness: How can the auditor verify that the big data is complete? | Research should be undertaken that can provide suggestions as to the verification of big data for the assertion of completeness. |
| Approvals: Should big data provide evidence of approvals/controls validations? Is this viable? | Studies of controls measurements of big data at all levels of generation and extraction should be conducted. For |

| | |
|---|---|
| | example process mining techniques (Jans et al, , 2014) can be used. |
| Ease of Use: Will big data require expertise to understand and extract and prepare for analysis? | What level of expertise should engagement staff attain to be competent in the modern audit engagement? This question is addressed later in this paper. |
| Clarity: Can this big data be replicated/re-performed/recalculated by the auditor? | Research should examine whether this is a viable test in a big data context and if so, how to perform it? This is the level of accuracy to be demanded from big data analytics. The concepts of materiality and relative error in the context of big data audit analytics should be examined in research |

*Table One: Issues regarding big data as audit evidence, expanded from Brown-Liburd and Vasarhelyi, 2015*

Figure One**:**

**LINKING ANALYTICAL PROCEDURES TO TRADITIONAL FILE INTERROGATION**

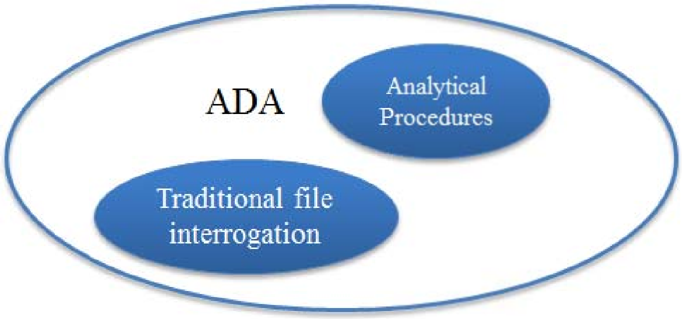


*Figure One: Linking Analytical Procedures to traditional file interrogation (Stewart, 2015)*

Table Two:

**<u>ISSUES OF NEW ANALYTICS IN THE AUDIT</u>**

| **Issue of New Analytics in the Audit** | **Recommendations** |
|---|---|
| Should there be more guidance in the standards regarding analytical methods? | This issue should be debated amongst practitioners, academics, and regulators. Perhaps the PCAOB should open commentary. |
| Do we know enough about these BA methods to issue guidance? | More careful research should be conducted about which methods would be more appropriate for the assertion and audit task before guidance can be issued. |
| What are the trade-offs between 100% population tests, sampling, and ad hoc analytics? | This issue is discussed in depth and recommendations provided later in this paper. Also see Brown-Liburd et el (2015). |
| Does analytics allow for automation of many judgment oriented audit procedures? | More experimental research is needed to evaluate the possibility of automation of many judgment oriented audit processes. |
| Can the audit opinion be disclosed in a more quantified and probabilistic manner? | This issue is discussed in depth and recommendations provided later in this paper. |
| Can the above be stated in terms of rules implementable in automated audit systems to continuously monitor and drive audit by exception (ABE)? | A framework for an automated ABE system should be proposed which takes advantage of the big data processing and business analytics capacities of modern enterprise systems. |

*Table Two: Summary of the Issues regarding New Analytics in the Audit and Recommendations for Future Research*

Table Three:

**WHICH METHODS ARE MOST PROMISING**

| **Issue regarding Which Methods are most promising** | **Recommendations** |
|---|---|
| Under what circumstances would modern analytical or more complex methods be appropriate? | Research should examine if the current standards regarding sampling, selection of specific items, or 100% tests could be expanded. |
| What would be the effect on the engagement, the firm, the standards? | This question could be incorporated in the same research above. |
| Could these approaches be formalized, if not industry wide at least internal to the firm? | This question could be incorporated in the same research question above. |
| Who would classify or standardize these approaches (create a taxonomy of methods and data structures for defined audit tasks)? | Perhaps this process could evolve under the guidance of the AICPA in collaboration with academics and practitioners. |
| How would these approaches be quantified? | A quantification framework could be proposed and demonstrated. |
| How would these approaches be tested in the field? Sand box approaches accompanied with successive levels of adoption? Would these be provided a safe harbor? | This could be part of the AICPA initiative with firm support and academic input. |
| Again, how would this affect the audit opinions? Could these modern analytical methods facilitate more transparent and quantitative disclosure? | A framework or guidance for a more detailed and quantitative opinion disclosure should be developed and proposed. |

*Table Three: Summary of issues regarding which methods are most promising*

Figure Two:

**THE CURRENT TYPICAL AUDIT PLAN**

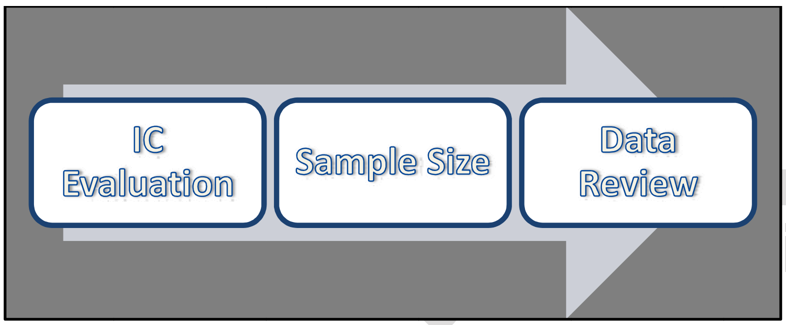


*Figure Two: The current typical audit plan*

Figure Three:

**AUDIT BY EXCEPTION**

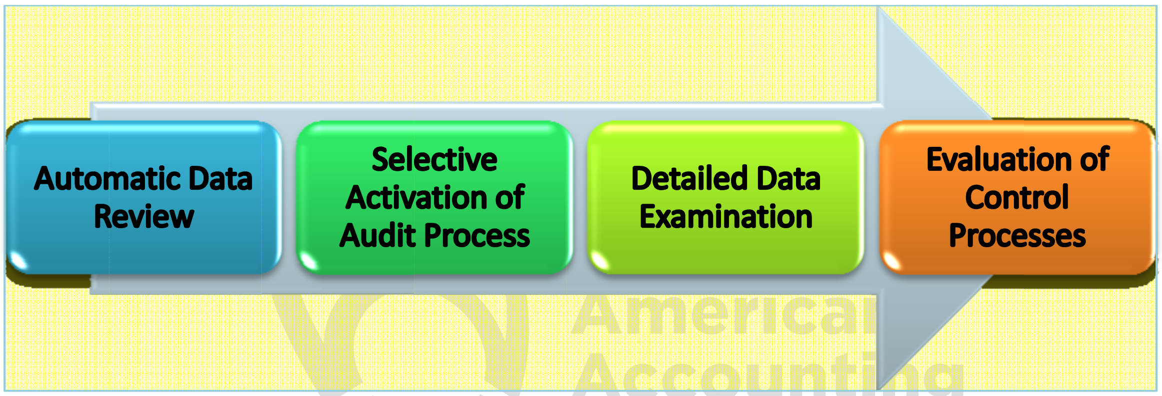


*Figure Three: In a more Data Driven Process, Audit by Exception (ABE) of audit examination*

Table Four:

**WHERE IN THE AUDIT THESE ANALYTICS WOULD BE APPLICABLE**

| **Issues about where in the audit these analytics would be applicable** | **Recommendations** |
|---|---|
| How would the objectives of internal and external audit differ in this context? | Research should examine the areas of convergence and separation in the context of integrated enterprise systems, analytics, and big data |
| Isn't there a substantive overlap between business monitoring and real time assurance? | This has been alluded to in earlier research but should be re-examined if the assurance process changes |
| Considering that there is an overlap in data analytic needs between different functions, how relevant are the three lines of defense? | Recent works by COSO have questioned the feasibility of the three lines of defense – however, the independence of assurance must be maintained, which is an area for future research. There are many possibilities for the three lines of defense. |
| What parts of the audit engagement are fully or partially automatable? Would auditor judgment eventually be replaced with prescriptive analytical algorithms? | This area could be examined at depth with varying levels and moments of audit automation, factoring such variables as judgement and interim testing |
| Would leading audit firms allow such disruptive changes in engagement practice, absent regulation changes? | Would these firms be willing to be key innovators in the assurance side? (Perhaps if they were to be allowed a sandbox or safe harbor? ) |
| Can the key contingencies in the audit be formalized? | These should be examined and articulated with frameworks/guidelines embedded in an expert system |
| If the annual audit opinion can become more informative, as per recent CAM reviews, why stop there? Why not issue CAM level quarterly reports and reports on demand? | The recommendations regarding this issue are discussed later in this paper. CAM reviews could serve as the foundation of a more quantitative opinion report. Other possibilities evolve for an immutable real-time seal of the data and its assurance |

*Table Four: Issues regarding where in the audit these methods would be applicable*

Table Five:

**SHOULD THE STANDARDS CHANGE TO FACILITATE THESE METHODS?**

| **Should the standards change to facilitate these methods?** | **Recommendations** |
|---|---|
| What would be the cost versus benefit trade-off with the usage of analytics in the current regulatory environment? | This issue should be examined as the cost benefit of more advanced analytics may be a major variable affecting the use by firms |
| What would be the breaking point of sample driven audits versus 100% tests resulting in ABE? | The effectiveness and efficiency of the two audit approaches should be examined in future research. This issue has been conceptually addressed in Yoon (2016) |
| When would the value derived from the additional assurance provided by analytical results justify their incremental cost? | Collaborative research efforts between academics and firms would be appropriate to address this issue |
| If more powerful analytics uncovers issues that were not previously detected, what would be the liability of the audit firm, particularly if these issues have been on-going? | This is an issue that the regulators should address, with input from the firms and researchers. This may relate also to earlier "safe harbor" questions |
| If the auditor has access and ability to test 100% of the dataset, would there still be justification for the use of sampling? | This is an issue that research should address, allowing for time, accuracy, and cost calculations for sampling versus 100% tests |
| Is there a way to quantify the evaluation of the cost and time to run 100% tests versus the perceived liability of sampling risk and judge accordingly? | This is an issue that the regulators should address as part of the preceding question |
| Are 100% tests new type of audit evidence or just automation? | This question could be examined along with other issues relevant to big data |
| If these tests are considered automation, how do the standards take this into consideration? Should the current solution of greater reliance on internal controls be quantified? | This is an issue that the regulators should address, with input from the firms and researchers. The controls testing and verification process as it relates to an IT audit and the reliability of information generated within a system may need clarification/quantification. |
| Is there a difference between automation and analytic methods? Isn't automation basically the automated application of analytics? | This is an issue to be considered in future research efforts by academics, as part of a scoring framework for audit evidence |
| If such an automation is viewed as a preventative internal control, then how does | This is an issue that the regulators should address, with input from the firms and |

| | |
|---|---|
| it change the balance between control testing in auditing the modern highly automated enterprise system? | researchers. |
| Would evidence from external sources such as social media require new guidance? | This question should be examined in detail given the veracity issue with external big data. Guidelines regarding normative expectations should be established – this evidence should be scored as part of the quantitative evidence framework.<br>Also, detailed examples/case studies should be discussed regarding the various sources of social media |
| What qualities should this data possess in order to provide reliable audit evidence? | This query can follow the recommendations proposed previously in the big data external evidence guidance discussion |
| Could the standards allow firm industry knowledge to be supplemented with anonymized confidential peer company data? | This is an issue that the regulators should address, with input from the firms and researchers. |
| Could new guidance be offered that defines client confidentiality as being firm wide in scope and not limited to an engagement team? | This is an issue that the regulators should address, with input from the firms and researchers. |

*Table Five: Where should the standards be changed to allow/facilitate these methods?*

Table Six:

**SHOULD THE AUDIT REPORT BE MORE INFORMATIVE?**

| **Should the audit report be more informative?** | **Recommendations** |
|---|---|
| Is the level of disclosure appropriate for more advanced analytics and quantification of critical audit matters (CAMs)? | A framework for appropriate disclosure should be developed |
| Would some of these CAMs provide disclosures that are more disaggregate or more informative than the traditional audit reports? | This is an issue that researchers and regulators should examine as a more informative CAM component of the audit opinion is formulated |
| Should there be quantitative guidelines for estimates, audit judgments, areas of special risk, unusual transactions, or other significant changes to the financial statements, and if so, by whom? Regulators? Researchers? | This is an issue that the regulators should address, with input from the firms and researchers. |
| Or projecting in the future, if the auditor is relying on an ABE assurance protocol, why shouldn't audit reports be generated more frequently or on a just-in-time/on demand basis? | This could be one aspect of a forward-looking paper by academics that conceptualizes a grand vision of the future public audit.<br>This could be a new form of service by auditors that probably now is forbidden by SOX. |

*Table Six: Should the audit report be more informative of Critical Audit Matters (CAMs)?*

Table Seven:

**AUDITOR COMPETENCIES**

| **Issues about auditor competencies** | **Recommendations** |
|---|---|
| In this day of Google and other IT tools, should the curriculum be filled with rote memorization tasks? | This topic should be examined and developed by academics with guidance from the AICPA |
| What types of education requirements, structures, and certification should be developed? | This topic should be examined and developed by academics with guidance from the AICPA |
| Should the audit profession move more towards the use of IT and analytics specialists in the engagement or is there room for this additional knowledge? | This topic should be examined by practitioners and academics in a behavioral study setting |
| Should the CPE requirements of the profession be reformulated to reflect these new learning skills/requirements? | This topic should be examined and developed by academics with guidance from the AICPA |
| Should there be a much wider set of accounting specializations with coordinated competencies? | This topic should be examined and developed by academics with guidance from the AICPA |
| Should there be quantification of the different types of auditing skills? | This topic should be examined and developed by academics with guidance from the AICPA |

*Table Seven: What are the competencies needed by auditors in this environment?*

Table Eight:

**ISSUES FOR BD/ADA ADOPTION**

| **Issues for BD/ADA adoption** | **Recommendations** |
|---|---|
| What are the goals/benefits/costs for each stakeholder/involved party? | Key drivers and motivating factors should be identified by firms, regulators, and clients. These should be discussed in terms of cost benefit analysis and effectiveness |
| Who should be the champions for this change? | To what degree and when would auditors use BD/ADA and who decides this? Who would be the main champions for this change? |
| How would this process develop? | To what degree and when would auditors use BD/ADA and who decides this? Should current audit procedures and regulations be changed prior to use of BD/ADA? |
| Who measures the effectiveness of using BD/ADA vs. not using and by what metrics? | Effectiveness and cost benefit analysis evaluation results may differ between stakeholders. Process of measurement metrics and expectations should be developed. |
| How would BD/ADA adoption take place at the firm level and regulatory level? | This question ties in with the process development (third) question |
| Would audit procedures need to be re-aligned to fit this new engagement environment? | Should current audit procedures and regulations be changed prior to use of BD/ADA? |
| How would auditors best prepare for these tasks that require more judgment and less routine work? | How would firms and regulators go about best preparing practitioners to transition to more judgement based and analytical approaches? |

*Table Eight: Issues that might impact BD/ADA adoption*

Figure Four:

**DIFFERENT STAGES OF ONE PROCESS CYCLE OF INCREMENTAL CHANGE**

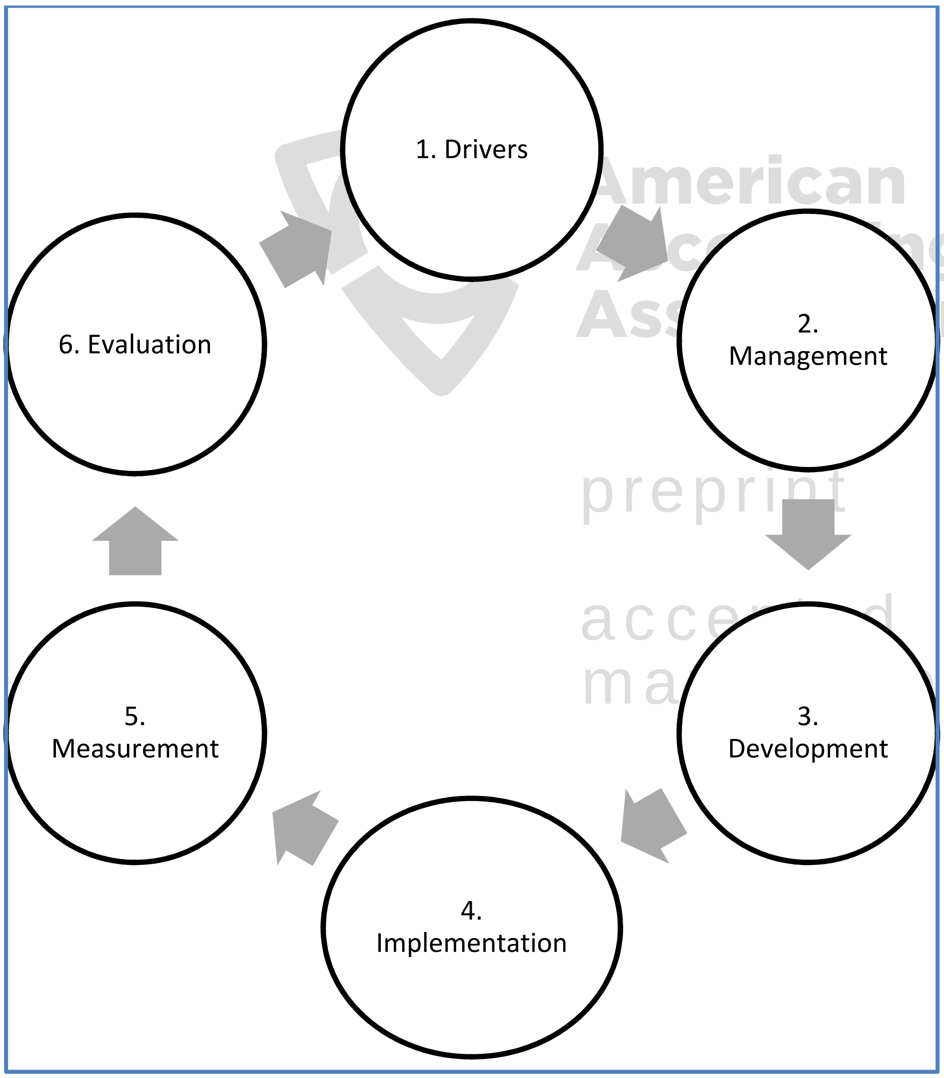


*Figure Four: Different stages of one process cycle of incremental change*

Figure Five:

**POSSIBLE CYCLES OF ADOPTION**

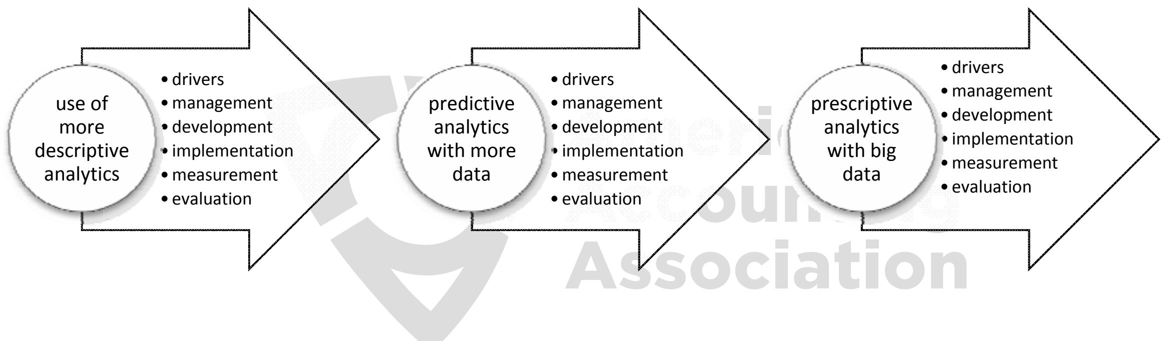


*Figure Five: Three possible cycles of adoption for the use of more advanced analytics and big data by the public audit profession*