An Analysis of Information Systems Literature: Contributions to Fraud Research

Jaime L. Grandstaff[1] & Lori L. Solsma[1]

[1]College of Business & Public Administration, Drake University, Des Moines, IA, USA.

Correspondence: Jaime Grandstaff, College of Business & Public Administration, Drake University, Des Moines, IA 50311, USA. E-mail: jaime.grandstaff@drake.edu

Received: October 17, 2019 Accepted: November 7, 2019 Online Published: November 12, 2019

doi:10.5430/afr.v8n4p219 URL: https://doi.org/10.5430/afr.v8n4p219

Abstract

This study analyzes the knowledge and methods used in information systems (IS) journals in the area of financial statement fraud. The purpose of this analysis is to provide tools and ideas to support interdisciplinary research in accounting and information systems for financial statement fraud topics. The study presents an analysis of five top ranking IS journals (MIS Quarterly, Information Systems Research, Communications of the ACM, Management Science, and Journal of MIS) and five top ranking IS conferences [International Conference on Information Systems (ICIS), Hawaii International Conference on System Sciences (HICSS), International Federation for Information Processing (IFIP), International Conference on Decision Support Systems (DSS), and Decision Sciences Institute National Conference (DSI)]. The literature found from these sources are categorized and presented by year, journal, contribution, type of study, methodology, data set usage, and research design. Although the literature varies, a common thread in many studies is the use of data mining and/or machine learning models to detect fraud.

Keywords: financial statement fraud, predictive modeling, and information systems literature

1. Introduction

The purpose of this study is to support interdisciplinary research regarding financial statement fraud in the accounting and information systems disciplines. The study begins by highlighting the limited amount of accounting research on financial statement fraud that cites information systems literature and similarly the limited amount of information systems research that cites accounting literature. Extant literature reviews from these disciplines provide evidence that the two disciplines referenced the other's research but to a limited extent, providing an indication that the research streams do not build on one another. In this study, literature reviews in the two disciplines are compared to observe cross-referencing citations from 2008 to 2018.

To accomplish the research objective, supporting interdisciplinary research on financial statement fraud, this study also analyzes and presents fraud-related literature found in information systems journals that can be useful in accounting or interdisciplinary research. The literature review began with the five top ranking Information Systems journals (MIS Quarterly, Information Systems Research, Communications of the ACM, Management Science, and Journal of MIS) and the five top ranking IS conferences [International Conference on Information Systems (ICIS), Hawaii International Conference on System Sciences (HICSS), International Federation for Information Processing (IFIP), International Conference on Decision Support Systems (DSS), and Decision Sciences Institute National Conference (DSI)]. The process for retrieving, analyzing, and classifying the information systems studies is based on processes used by Lev & Ellis (2006), Webster & Watson (2002), and Gab et al. (2015). This involves a key word search followed by a forward-looking process (review articles that have cited those from the initial search list, thus published subsequently) and a backward-looking process (review articles that were cited by those from the initial search list, thus published in the past).

This process uncovers and showcases published research in top information systems journals and conferences and other articles linked by citation that may be useful to accounting scholars in their research of financial statement fraud. The articles are categorized by research methods, data source, and contribution. A short summary of each article is provided to help guide future interdisciplinary research.

The remainder of the paper is organized as follows. Section two discusses the background and literature review for financial statement fraud. Section three presents the methodology and the process for locating information systems research for this study. Section four documents the article characteristics and results of the findings. Lastly, section five offers a discussion regarding future research and conclusions.

2. Background and Literature Review

The focus in this study is the research intersect between information systems research and accounting research relating to fraudulent financial statements. The principal statutory regulation for fraudulent financial reporting in publicly traded companies in the U.S. is 17 CFR 240.10b-5, a provision of the Securities Exchange Act of 1934. This law makes it unlawful to "engage in any act, practice or course of business, which operates or would operate as a fraud or deceit upon any person, in connection with a purchase or sale of any security." In general, the perpetrator must know the financial statements contain a false statement, thus showing an intent to deceive. This separates fraud from cases of errors or misunderstandings. However, intent to deceive does not separate fraud acts from earnings management. The primary difference between earnings management and fraud is that fraud is generally outside of generally accepted accounting principles (GAAP) and earnings management is within GAAP (Perols & Lougee, 2011).

Research in the information systems field includes different types of financial fraud (West & Bhattacharya, 2016; Ngai, Hu, Wong, Chen, & Sun, 2011). In this study, we limit our focus to financial statement fraud. Therefore, we exclude all of the financial fraud categories in Table 1, with the exception of financial statement fraud, located in the corporate fraud category.

Table 1. Types of Financial Fraud in Information Systems Literature

Bank Fraud	Insurance Fraud	Corporate Fraud
Credit Card Fraud	Automobile Insurance	Financial Statement Fraud
Mortgage Fraud	Health Insurance	Securities and Commodities
Money Laundering	Crop Insurance	Mass Marketing Fraud

Figure 1 provides a timeline of the publication of several prominent fraud-related literature reviews scanning a ten-year period (from 2008 to 2018). Ngai, Hu, Wong, Chen & Sun (2011) and West & Bhattacharya (2016) published fraud-related literature reviews in Decision Support Systems and Computers & Security, respectively. Both are information systems journals. The other literature reviews (Hogan, Rezaee, Riley, & Velury, 2008; Trompeter, Carpenter, Desai, Jones, & Riley, 2013; Trompeter, Carpenter, Jones, & Riley, 2014; Amiram, Bozanic, Cox, Dupont, Karpoff & Sloan, 2018) are published in the following accounting journals: Auditing: A Journal of Practice and Theory, Accounting Horizons, and Review of Accounting Studies.

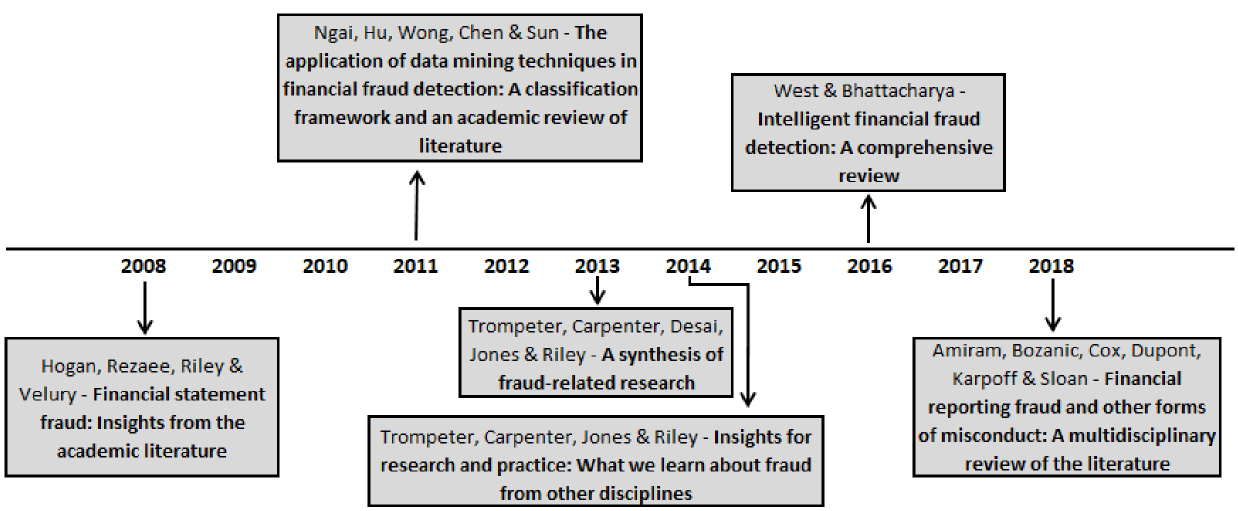

Figure 1. Timeline of Fraud Literature Reviews

As shown in Figure 1, articles from Information Systems journals are located above the timeline and articles from Accounting journals are located below the timeline.

Surprisingly, in the four comprehensive literature reviews of fraud in accounting journals only two information systems journals were cited. This is not because relevant topics were not discussed. For example, Trompeter et al. (2013) discussed analytical procedures used in detecting fraud but did not include the data mining applications and techniques used to extract relevant relationships in the data as discussed in Ngai et al.'s (2011) literature review. In another example, Trompeter et al. (2014) published a literature review that expanded the scope of the Trompeter et al. (2013) article to specifically include research from other disciplines and is titled "Insights for research and practice: What we learn about fraud from other disciplines." They surveyed academic literature from non-accounting publications related to fraud and financial crimes. This article brought in literature from criminology, ethics, psychology, and sociology. Again, the information system's field of literature regarding financial statement fraud was not present in the article despite the opportunity to include this research in a section titled "Computer Analytics" that discusses fraud detection. Similarly, the Hogan et al. (2008) literature review of financial statement frauds contained a review of fraud analytical procedures with no reference to data mining applications and techniques that can be found in information systems publications. Lastly, Amiram et al.'s (2018) published a multidisciplinary review of fraud literature from the legal, accounting, and finance disciplines. The section titled "Recent methodological advances in misconduct prediction" references only one machine learning journal.

In the literature reviews of fraud in the information systems journals examined in this study, the authors were more likely to reference accounting journals than were accounting journal authors to reference information systems journals (as shown in Table 2). Yet, the number of citations were few. Ngai et al. (2011) classified fraud literature by type and by data mining technique and included reference to eight accounting articles. West & Bhattacharya (2016) referred to four accounting articles when describing statistical methods used for fraud detection and noted that the methods of fraud detection in recent studies are far more varied than in earlier years.

Table 2. Accounting Journal Citations in Information Systems Literature Reviews of Fraud

Journal Title	Number of Accounting Journal Citations	
	Ngai et al. (2011)	West & Bhattacharya (2016)
Accounting Review		1
Auditing: A Journal of Practice & Theory	3	
British Accounting Review		1
European Accounting Review	1	1
International Journal of Accounting Information Systems		1
Managerial Auditing Journal	4	
Total	8	4

As evidenced by the limited amount of cross-referencing between literature reviews in the accounting and information systems disciples, there may be a shortage of financial statement fraud cross-disciplinary research. With the increased use of data analytics in the accounting field, the connection between the two disciplines is gaining strength. Therefore, this study focuses on financial statement fraud research in the information systems discipline to provide a resource to accounting researchers seeking to integrate more information systems contributions, methodologies, and research designs.

3. Methodology

Based on rakings from Levy & Ellis (2006) and Bhattacharjee et al. (2004), this study begins with an analysis of extant literature in five top ranking IS journals (MIS Quarterly, Information Systems Research, Communications of the ACM, Management Science, and Journal of MIS). Five top ranking IS conferences [International Conference on Information Systems (ICIS), Hawaii International Conference on System Sciences (HICSS), International Federation for Information Processing (IFIP), International Conference on Decision Support Systems (DSS), and Decision Sciences Institute National Conference (DSI)] are also included for recently presented manuscripts that are unpublished. The process for article retrieval, analysis, and classification schemes follow those used by Levy & Ellis (2006), Webster & Watson (2002), and Gab et al. (2015).

In the first stage, research articles, titles, and abstracts from the journals and conferences outlined previously were searched using databases such as EBSCOhost, JSTOR, Google Scholar and conference websites. The search terms "fraud," "financial statement fraud," and "fraudulent financial reporting" were used to extract relevant articles. The search process occurred in the first quarter of 2019 and spanned a 42-year period from 1977 to 2019 (based on database access). The initial search of the electronic databases resulted in approximately 490 "hits." Several articles were duplicates due to the overlapping nature of the search terms. The search of conference websites using the same search terms resulted in the discovery of 23 working papers.

In the second stage, the articles and working papers are screened to include only those articles that are within the scope of this review. The focus for this review is to analyze the research contributions regarding financial statement fraud from an IS perspective. Therefore, to be included in this study, two criteria must be met. The publication must be from a literature source targeted at the IS community and the topic must be fraudulent financial statement reporting.

After screening the articles and working papers to include only those with the characteristics described, a forward-looking process and backward-looking process was used to uncover more articles that meet the previously stated criteria. A forward-looking process is a search for articles that cited the articles found in the initial search results (articles published subsequently). A backward-looking process is a review of articles that were cited by the articles found in the initial search results (articles published in the past). After completing the screening techniques and the forward- and backward-looking process, 28 IS articles were found. The characteristics of these articles are analyzed and presented in the sections that follow.

4. Results and Article Characteristics

4.1 Articles Published by Year

The number of fraud articles published by year in information systems journals is presented in Figure 2. There appears to be an increasing trend in in the number of fraud-related articles in recent years. The spike in 2011 occurred partially due to the publication of a special issue from Decision Support Systems, called "Quantitative Methods for Detection of Financial Fraud." This issue published six fraud articles.

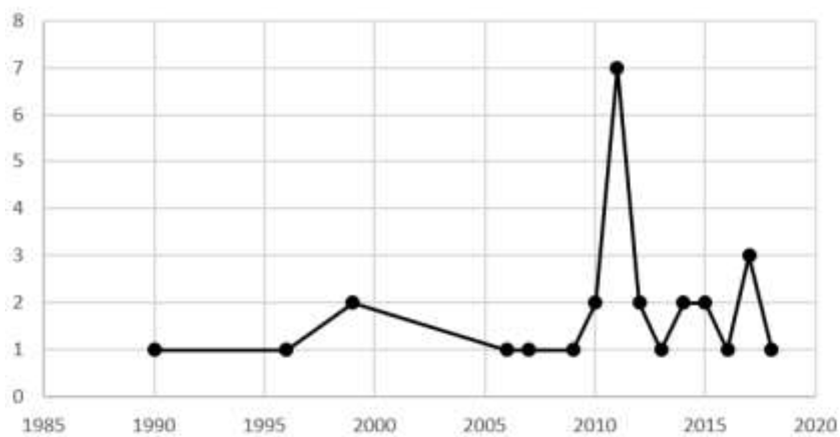

Figure 2. Articles by Year of Publication

4.2 Articles Published by Journal

Information Systems journal outlets (including 2018 impact factors) with the number of fraud articles and publication dates are presented in Table 3. Journals with the highest number of fraud articles include Decision Support Systems, with eight articles and Expert Systems with Applications, with five articles.

Table 3. Articles by Journal

Journal Title	Count	Publication Date(s)	Impact Factor*
Decision Support Systems	8	2009-2015	3.847
Expert Systems with Applications	5	2007-2017	4.292
Computer Fraud & Security	2	1999	n/a
Knowledge-Based Systems	2	2015-2017	5.101
Management Science	2	1996-2010	4.219
Computers & Security	1	1990	3.062
Information Systems Frontiers	1	2017	2.539
International Journal of Computational Intelligence	1	2006	n/a
International Journal of Digital Content Technology and its Application	1	2013	n/a
International Journal of Enterprise Information Systems	1	2012	n/a
Journal of Management Information Systems	1	2018	3.013
Knowledge and Information Systems	1	2011	2.397
MIS Quarterly	1	2012	4.373
The Scientific World Journal	1	2014	n/a

* 2018 Impact factor from Web of Science Journal Citation Reports

n/a = not available

4.3 Article Contributions

4.3.1 Empirical Studies Using Archival Data

The information systems articles in Table 4 use archival data to analyze fraud, in a manner similar to many accounting fraud studies. The articles are categorized by research purpose in the following categories: fraud detection studies, evaluating different technologies to detect fraud, and evaluating the incremental benefit of using both financial and textual data. Data types include financial, textual, audit risk factors, and other quantitative data. The sections that follow describe the articles in Table 4 and are organized according to the previously outlined categories.

Table 4. Characteristics of Studies using Archival Data

Authors	Times Cited*	Data Type	Matching – Sample Size	Fraud Detection Only	Technology Comparison	Financial vs. Textual Data
Abbasi et al. (2012)	44	F	N - 307	✓		
Cecchini et al. (2010) a	35	F	U - 122:3,187	✓		
Cecchini et al. (2010) b	54	F, T	E - 61	✓		
Dong et al. (2018)	8	F, T	E - 64	✓		✓
Glancy & Yadav (2011)	43	T	N - 69	✓		
Hansen et al. (1996)	44	A	U - 77: 305	✓		
Holton (2009)	46	T	E - 40	✓		
Huang et al. (2014)	13	F	E - 72	✓		
Humpherys et al. (2011)	84	T	E - 101	✓		
Throckmorton et al. (2015)	10	F, T	U - 41:1,531	✓		✓
Yeonkook et al. (2016)	15	F	U - 788:2,156	✓		
Chen et al. (2014)	32	F, T	E - 66		✓	
Dutta et al. (2017)	7	F	N - 3,513		✓	
Hajek & Henriques (2017)	13	F, T	E - 311		✓	
Huang (2013)	5	F	U - 1:4 ratio		✓	
Kirkos et al. (2007)	186	F	E - 38		✓	
Kotsiantis et al. (2006)	40	F	U - 41:123		✓	
Lin et al. (2015)	17	F	U - 129:447		✓	
Ravisankar et al. (2011)	117	F	E - 101		✓	
Wozniak (2011)	35	O	N - Varied		✓	

*Times Cited source is Web of Science (search performed in the fourth quarter of 2019).

F = financial; T = text; A = audit risk factors; O = other

E = Equal; U = Unequal; N = No matching

Detecting Fraud using Financial Data. Cecchini et al. (Detecting management fraud in public companies, 2010) used 29 financial variables from publicly available information (from 122 fraud firms and 3,187 non-fraud firms) to test a machine learning/data mining method using support vector machines to detect fraud. They found support vector machines useful in discriminating between fraud and non-fraud firms. Similarly, using 12 financial ratios (sample size of 307 firms) Abbasi et al. (2012) tested the effectiveness of a meta-learning framework (using a design science approach) to detect fraud. They found that each component of the framework contributed to its effectiveness. Huang (2014) used 24 financial variables (from 72 matched firms) in a dual growing hierarchical self-organizing map (GHSOM) approach to find topological patterns of fraud in financial statements. The study found that the patterns follow spatial relationships and therefore, show promise for fraud detection. Yeonkook et al. (2016) used financial data (788 fraud and 2,156 non-fraud yearly and quarterly results) to assess whether their multi-class fraud prediction model, using multinomial logistic regression, a support vector machine, and Bayesian networks, is able to distinguish intentional misstatements from unintentional misstatements. The study reported an 88 percent detection rate. These articles are impactful with times cited ranging from 13 to 44 for the Abbasi et al. (2012) article, as illustrated in Table 4.

Detecting Fraud using Textual or Other Types of Data. In an early study, Hansen et al. (1996) used the audit risk factors of 77 fraud firms and a set of 305 non-fraud firms in a generalized qualitative-response model (EGB2) to predict fraud. Over 20 trials, the model's predictive accuracy was 89.3 percent. Holton's (2009) approach was different than others. She used a matched set of 40 messages (from Vault.com and Yahoo! discussion groups) from disgruntled and non-disgruntled employees to predict fraud using a Naïve Bayes model. The model was successful in identifying a dissatisfied employee, a key fraud risk factor. Using MD&A text from 69 firms, Glancy & Yadav (2011) used a computational fraud detection model that employed text-mining techniques to detect fraud. They found that their model is able to discriminate fraud firms from non-fraud firms. Humpherys et al. (2011) also used MD&A text but used a matched set of 101 firms to investigate the use of deceptive language. The study used the Agent99 Analyzer, the Naïve Bayes classifier, and C4.5 decision classifiers to examine language characteristics. The study found more active language, lower content word diversity, and more complexity (in the form of longer words and more pauses with punctuation) in the MD&A sections of fraud firms than in non-fraud firms. These articles are also impactful with times cited ranging from 43 to 84 for the Humpherys et al. (2011) article.

Detecting Fraud using Financial and Textual Data. Using both financial ratios and text data from the MD&A's of 61 matched firms, Cecchini et al. (Making words work: Using financial text as a predictor of financial events, 2010) developed a methodology for analyzing text using an automating ontology creation. The dictionaries that were created were able to discriminate fraud from non-fraud firms. This article was cited 54 times.

Comparing Methods of Detection using Financial Data. Several studies used financial data to compare different technologies for use in fraud detection. Kotsiantis et al. (2006) used financial variables from 41 Greek fraud firms and 123 non-fraud firms to explore the effectiveness of machine learning techniques in detecting fraud. This study compared a stacking variant methodology to simple and ensemble methods finding that a stacking variant provides better performance in fraud detection. Kirkos et al. (2007) is another study that used financial data from Greek firms (38 matched firms). The study compared the effectiveness of data mining classification techniques using a Decision Tree model, a Neural Network model, and the Bayesian Belief Network model. The fraud detection performance rates were 90.3 percent for the Bayesian Belief Network model, 80 percent for the Neural Network model and 73.6 percent for the Decision Tree model.

Ravisankar et al. (2011) used the financial data of a matched set of 101 Chinese firms. This study compares the following data mining techniques: multilayer feed forward neural network (MLFF), support vector machines (SVM), genetic programming (GP), group method of data handling (GMDH), logistic regression (LR), and probabilistic neural network (PNN) to predict of occurrence of financial statement fraud. This study finds that PNN was the top performer followed by GP. Huang (2013) uses five financial variables in a 1:4 firm matched sample design of Taiwanese firms to compare two data mining methods, a support vector machine techniques and logistic regression. The study found that support vector machines are better at detecting financial statement fraud two years in advance of the fraud event.

Lin et al. (2015) used 32 financial variables in an approximate 1 to 4 (129:447) matched set of Taiwanese firms to compare data mining techniques and expert decisions in classifying the presence of fraud. The artificial neural networks (ANNs) and decision trees (CART) had higher classification rates than logistic regression. Expert decisions were consistent with the data mining techniques. Dutta et al. (2017) used 15 financial variables in 3,513 restatements to compare the following data mining techniques: decision tree (DT), artificial neural network (ANN), Naïve Bayes (NB), support vector machine (SVM), and Bayesian Belief Network (BBN) classifier. The study found that two classifiers, the ANN and DT, are useful predictors of intention (fraudulent) and unintentional (erroneous) financial restatements.

As outlined above, several studies compared fraud detection methods. The impact (based on number of times cited) of these articles ranges from 5 to 186. The Kirkos et al. (2007) article (186 citations) and the Ravisankar et al. (2011) article (117 citations) were the most impactful articles.

Comparing Methods of Detection using Other Data. Wozniak (2011) used publicly available data sources such as classes of waves, heart disease diagnosis, and others to propose an algorithm for uses such as detecting fraud that is able to co-train decision trees to consider possible concept drift (a changing model). The study compared the use of

WEKA and Matlab (with the PRTools toolbox) with the C4.5 algorithm then co-trained using the iDTt-NGE algorithm. The study found that the model with co-training is slightly worse (yet more stable with a smaller standard deviation) and the computational cost is lower than training a new tree. This article was cited 35 times.

Comparing Methods of Detection using Financial and Textual Data. Chen et al. (2014) conducted a study of 24 financial variables and 5 non-financial variables from a matched set of 66 Taiwanese firms to compare classification models used to forecast fraudulent financial statements. The study compares a hybrid approach of stepwise regression, logistic regression, support vector machine, and decision tree and finds that the decision tree C5.0 is the best classifier of fraud and non-fraud firms. Hajek & Henriques (2017) used MD&A text and financial ratios from a matched set of 311 firms to compare the following machine learning methods: logistic regression, Bayesian methods, decision trees, SVM, neural networks, and ensemble methods. The study finds that ensemble methods are best at detecting fraud firms while Bayesian belief networks (BBN) are best at finding non-fraud firms. These articles were published more recently and number of times cited are 32 for the Chen et al. (2014) study and 13 for the Hajek & Henriques study (2017).

Evaluating the Incremental Benefit of using Both Financial and Textual Data. Throckmorton et al. (2015) used 41 quarterly conference call audio files of fraud firms and 1,531 audio files of non-fraud firms to develop a tool to detect fraud using financial numbers, linguistic behavior, and non-verbal cues. The study used a classifier approach called the generalized likelihood ratio test (GLRT) and a Bayesian-based application, and they found that the financial numbers, linguistic behavior, and non-verbal cues provide complementary information for fraud detection. Dong et al. (2018) used unstructured data (from social media sources Seeking Alpha and Yahoo Finance) in a 64-firm matched set to propose a framework to detect fraud. They found that latent features in social media data are lead indicators of fraud, better than MD&A sections and financial ratios. When financial ratios, language-based features, and social media features were combined, they found a prediction rate of 80 percent. These articles were also published more recently. Throckmorton et al. (2015) was cited ten times and Dong et al. (2018) was cited eight times.

4.3.2 Other Fraud Studies

From the search previously described in section 3, the remaining eight articles are displayed in Table 5. The article types are: analytical, case study, conceptual and survey.

Table 5. Other Fraud Studies

Authors	Times Cited*	Type	Data Source
Bhattacharya et al. (2011)	11	Analytical	Self-generated data
Jans et al. (2011)	50	Case study	10,000 purchase orders
Cerullo & Cerullo (1999) a	4	Conceptual	Not applicable
Cerullo & Cerullo (1999) b	n/a	Conceptual	Not applicable
Mercer (1990)	n/a	Conceptual	Not applicable
Zhou & Kapoor (2011)	39	Conceptual	Not applicable
Askary et al. (2012)	3	Survey	522 survey responses
Huang et al. (2017)	2	Survey	11 expert questionnaires

*Times Cited source is Web of Science (search performed in the fourth quarter of 2019). *Analytical.*

Bhattacharya et al. (2011) built an artificial neural network (ANN)-based decision support system to classify data based on Benford's law. The study used one data set constructed from the Benford distribution and 800 other data sets generated by Monte Carlo simulation. Data sets were manipulated at a 10 percent, 20 percent and 50 percent level (representing sophistication in data manipulation) and also for involvement levels (representing collusion). As expected, the study found that the ANN performed better for low and moderate levels of manipulation sophistication across all three levels of involvement. As shown in Table 5, this article was cited 11 times.

Case Study. Jans et al. (2011) conducted a case study of an anonymous top-20 European firm to test whether data mining a process (event logs /the audit trail for purchase orders) is effective for fraud detection. A text dump of 10,000 purchase orders was mined for patterns and less frequent flows using Fuzzy Miner and ProM. The study reports that process mining tools are premature and need to be enhanced before being used to automate the audit process. This article was impactful being cited 50 times.

Conceptual Studies. The first three conceptual articles are dated from the 1990s. The Mercer (1990) article describes how established audit techniques can be used in detecting fraud using computerized techniques. The two Cerullo & Cerullo (Using neural networks to predict financial reporting fraud: Part 1, 1999; Using neural networks to predict financial reporting fraud: Part 2, 1999) studies describe fraud, the applications of neural networks, and illustrates how to create a database, construct a model, and evaluate performance. In another conceptual article, Zhou & Kapoor (2011), the authors propose a methodology (a self-adaptive framework based on a response surface model) with domain knowledge such as, industry, economic conditions, management's choice, and timing considerations to detect fraud. The first three conceptual studies are dated, but the Zhou & Kapoor article is more recent and was cited 39 times.

Survey. Askary et al. (2012) use the findings of the 2008 Computer Crime and Security Survey (CSI), consisting of 522 survey responses, to describe trends in information technology fraud on audit risk models. The study also reports improvements in audit risk. Huang et al. (2017) sent questionnaires to 11 experts to identify financial statement fraud factors. After using a survey and analytic hierarchy process (AHP) to filter and categorize responses to 32 fraud factors in 20 categories, the survey is redistributed to eight experts. The survey results indicate that pressure/incentive is the highest-ranking fraud indicator with the top five measurements being: (1) poor performance, (2) the need for external financing, (3) financial distress, (4) insufficient board oversight, and (5) competition or market saturation. Times cited was three for the Askary et al. (2012) study, and the more recent Huang et al. (2017) study was cited twice.

5. Discussion and Conclusion

In this study, fraud-related literature is obtained through a search of top information systems journals and conferences, including articles cited in this literature and articles that cite this literature. The articles found from the information systems journals are categorized by type of study, data source, and contribution. A short description of each article is also provided. The Appendix provides a list of the articles in descending chronological order.

The largest category of articles is empirical studies that used archival data (primarily financial and textual data). They consist of two broad categories, those that seek to predict or classify firms as fraudulent or non-fraudulent and those that compare detection tools to determine the best model to use in fraud detection. Within these broad categories, two studies evaluate the benefits of using multiple types of data (financial and textual) for fraud detection. These studies have the commonality of using data mining and/or machine learning models to either detect fraud or compare the efficacy of these types of models in fraud detection. The analytical article and case study were similar in their use of data mining or machine learning to detect fraud or data manipulation. The survey article differed in its purpose, to identify factors indicating financial statement fraud. The conceptual articles offered the authors' perspectives on various topics with the most current involving a methodology that uses domain knowledge to detect fraud.

In general, the studies presented in this article appear to be successful in fraud detection, impactful in terms of number of citations, and may prove to be beneficial for cross-disciplinary research involving accounting and information systems researchers. The common goal of preventing and detecting fraud is shared across many business disciplines and practitioners, and there is a natural harmonization for the studies from information systems to enhance accounting research.

For future studies of fraud in the accounting discipline, this study urges researchers to consider incorporating the findings and methodologies used in information systems articles to enhance diversity of thought and quality. Data mining and machine learning models may have a place in the following areas of research on financial statement fraud: risk identification, brainstorming, strategic auditing, client acceptance decisions and others. In order to incorporate ideas from the information systems literature, consider the following questions.

1. Can information systems literature support my hypothesis development?
2. Can I use data mining or machine learning to increase the robustness of my research?
3. Are information technology models available that provide "new" ways to test data?
4. Can I identify someone from the information systems discipline with expertise that can add value to my research?

The contributions of the financial statement fraud articles found (in information systems journals) in this study have two limitations. First, similar to other literature reviews, this review relies on previously published research and the availability of these articles using the steps outlined in the methodology section. Second, the number of articles found in the search is limited as a result of the research objective that presents and evaluates only financial statement frauds rather than all frauds. Opening the scope to more types of fraud provides an avenue for future research that can provide further benefit to accounting researchers.

References

Abbasi, A., Albrecht, C., Vance, A., & Hansen, J. (2012). MetaFraud: A Meta-Learning Framework for Detecting Financial Fraud. *MIS Quarterly*, *36*(4), 1293-1327. https://doi.org/10.2307/41703508

Amiram, D., Bozanic, Z., Cox, J. D., Dupont, Q., Karpoff, J. M., & Sloan, R. (2018). Financial reporting fraud and other forms of misconduct: *A multidisciplinary review of the literature, 23*(2), 732-783. https://doi.org/10.1007/s11142-017-9435-x

Askary, S., Goodwin, D., & Lanis, R. (2012). Improvements in audit risks related to information technology frauds. *International Journal of Enterprise Information Systems, 8*(2), 52-63. https://doi.org/10.4018/jeis.2012040104

Bhattacharjee, S., Tung, Y. A., & Pathak, B. (2004). Author experiences with the IS journal review process. *Communications of the association for Information Systems, 13*, 629-653. https://doi.org/10.17705/1CAIS.01337

Bhattacharya, S., Xu, D., & Kumar, K. (2011). An ANN-based auditor decision support system using Benford's law. *Decision Support Systems, 50*, 576-584. https://doi.org/10.1016/j.dss.2010.08.011

Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010) a. Detecting management fraud in public companies. *Management Science, 56*(7), 1146-1160. https://doi.org/10.1287/mnsc.1100.1174

Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010) b. Making words work: Using financial text as a predictor of financial events. *Decision Support Systems, 50,* 164-175. https://doi.org/10.1016/j.dss.2010.07.012

Cerullo, M. J., & Cerullo, V. (1999). Using neural networks to predict financial reporting fraud: Part 1. *Computer Fraud & Security,* 14-17. https://doi.org/10.1016/S1361-3723(99)80015-3

Cerullo, M. J., & Cerullo, V. (1999). Using neural networks to predict financial reporting fraud: Part 2. *Computer Fraud & Security*, 14-17. https://doi.org/10.1016/S1361-3723(99)80035-9

Chen, S., Goo, Y.-J. J., & Shen, Z.-D. (2014). A hybrid approach of stepwise regression, logistic regression, support vector machine, and decision tree for forecasting fraudulent financial statements. *The Scientific World Journal*, 1-9. https://doi.org/10.1155/2014/968712

Dong, W., Liao, S., & Zhang, Z. (2018). Leveraging financial social media data for corporate fraud detection. *Journal of Management Information Systems, 35*(2), 461-487. https://doi.org/10.1080/07421222.2018.1451954

Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. *Expert Systems With Applications, 90,* 374-393. https://doi.org/10.1016/j.eswa.2017.08.030

Gab, O., Ortbach, K., Kretzer, M., Maedche, A., & Niehaves, B. (2015). Conceptualizing individualization in information systems - A literature review. *Communications of the Association for Information Systems, 37*(3), 64-88. https://doi.org/10.17705/1CAIS.03703

Glancy, F. H., & Yadav, S. B. (2011). A computational model for financial reporting fraud detection. *Decision Support Systems, 50*, 595-601. https://doi.org/10.1016/j.dss.2010.08.010

Hajek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods. *Knowledge-Based Systems, 128*, 139-152. https://doi.org/10.1016/j.knosys.2017.05.001

Hansen, J. V., McDonald, J. B., Messier, Jr., W. F., & Bell, T. B. (1996). A generalized qualitative-response model and the analysis of management fraud. *Management Science, 42*(7), 1022-1032. https://doi.org/10.1287/mnsc.42.7.1022

Hogan, C. E., Rezaee, Z., Riley, R. A., & Velury, U. (2008). Financial statement fraud: Insights from the academic literature. *Auditing: A Journal of Practice & Theory, 27*(2), 231-252. https://doi.org/10.2308/aud.2008.27.2.231

Holton, C. (2009). Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multi-billion dollar problem. *Decision Support Systems, 46*, 853-864. https://doi.org/10.1016/j.dss.2008.11.013

Huang, S. Y. (2013). Fraud detection model by using support vector machine techniquest. *International Journal of Digital Content Technology and its Applications, 7*(2), 32-42. https://doi.org/10.4156/jdcta.vol7.issue2.5

Huang, S. Y., Lin, C. C., Chiu, A. A., & Yen, D. C. (2017). Fraud detection using fraud triangle risk factors. *Information Systems Frontiers*, *19*, 1343-1356. https://doi.org/10.1007/s10796-016-9647-9

Huang, S. Y., Tsaih, R. H., & Yu, F. (2014). Topological pattern discovery and feature extraction for fraudulent financial reporting. *Expert Systems with Applications, 41*, 4360-4372. https://doi.org/10.1016/j.eswa.2014.01.012

Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., & Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems, 50*, 585-594. https://doi.org/10.1016/j.dss.2010.08.009

Jans, M., van der Werf, J. M., Lybaert, N., & Vanhoof, K. (2011). A business process mining application for internal transaction fraud mitigation. *Expert Systems with Applications, 38*, 13351-13359. https://doi.org/10.1016/j.eswa.2011.04.159

Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data Mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications, 32*, 995-1003. https://doi.org/10.1016/j.eswa.2006.02.016

Kotsiantis, S., Koumanakos, E., Tzelepis, D., & Tampakas, V. (2006). Forecasting fraudulent financial statements using data mining. *International Journal of Computational Intelligence, 3*(2). https://doi.org/10.1007/11752912_63

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science Journal, 9*, 181-212. https://doi.org/10.28945/479

Li, Z. F., & Dang, C. (2018). Drivers of Research Impact: Evidence from the Top Three Finance Journals. *Accounting & Finance*, Forthcoming. Retrieved from Available at SSRN: https://ssrn.com/abstract=3123121

Lin, C. C., Chiu, A. A., Huang, S. Y., & Yen, D. C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. *Knowledge-Based Systems, 89*, 459-470. https://doi.org/10.1016/j.knosys.2015.08.011

Mercer, L. C. (1990). Tailor-made auditing of information systems for the detection of fraud. *Computers & Security, 9*(1), 59-66. https://doi.org/10.1016/0167-4048(90)90158-P

Ngai, E., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic. *Decision Support Systems, 50*, 559-569. https://doi.org/10.1016/j.dss.2010.08.006

Perols, J. L., & Lougee, B. A. (2011). The Relation Between Earnings Management and Financial Statement Fraud. *Advances in Accounting, 27*(1), 39-53. https://doi.org/10.1016/j.adiac.2010.10.004

Ravisankar, P., Ravi, V., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems, 50*, 491-500. https://doi.org/10.1016/j.dss.2010.11.006

Throckmorton, C. S., Mayew, W. J., Venkatachalam, M., & Collins, L. M. (2015). Financial fraud detection using vocal, linguistic and financial cues. *Decision Support Systems, 74*, 78-87. https://doi.org/10.1016/j.dss.2015.04.006

Trompeter, G. M., Carpenter, T. D., Desai, N., Jones, K. L., & Riley, R. A. (2013). A synthesis of fraud-related research. *Auditing: A Journal of Practice & Theory, 32*(1), 287-321. https://doi.org/10.2308/ajpt-50360

Trompeter, G. M., Carpenter, T. D., Jones, K. L., & Riley, R. A. (2014). Insights for research and practice: What we learn about fraud from other disciplines. *Accounting Horizons, 28*(4), 769-804. https://doi.org/10.2308/acch-50816

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, 26*(2), xiii-xxiii.

West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security, 57*, 47-66. https://doi.org/10.1016/j.cose.2015.09.005

Wozniak, M. (2011). A hybrid decision tree training method using data streams. *Knowledge and Information Systems, 29*, 335-347. https://doi.org/10.1007/s10115-010-0345-5

Yeonkook, J. K., Baik, B., & Cho, S. (2016). Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning. *Expert Systems With Applications*, *62*, 32-43. https://doi.org/10.1016/j.eswa.2016.06.016

Zhou, W., & Kapoor, G. (2011). Detecting evolutionary financial statement fraud. *Decision Support Systems*, *50*, 570-575. https://doi.org/10.1016/j.dss.2010.08.007

Appendix

The following contains a list of the articles in descending chronological order.

Year	Authors	Article Title	Times Cited*
2018	Dong et al.	Leveraging financial social media data for corporate fraud detection.	8 [8]
2017	Dutta et al.	Detecting financial restatements using data mining techniques.	7 [3.5]
2017	Hajek & Henriques	Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods.	13 [6.5]
2017	Huang et al.	Fraud detection using fraud triangle risk factors.	2 [1]
2016	Yeonkook et al.	Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning.	15 [5]
2015	Lin et al.	Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments.	17 [4.3]
2015	Throckmorton et al.	Financial fraud detection using vocal, linguistic and financial cues.	10 [2.5]
2014	Chen et al.	A hybrid approach of stepwise regression, logistic regression, support vector machine, and decision tree for forecasting fraudulent financial statements.	32 [6.4]
2014	Huang et al.	Topological pattern discovery and feature extraction for fraudulent financial reporting.	13 [2.6]
2013	Huang	Fraud detection model by using support vector machine techniques.	5 [0.8]
2012	Abbasi et al.	MetaFraud: A Meta-Learning Framework for Detecting Financial Fraud.	44 [6.3]
2012	Askary et al.	Improvements in audit risks related to information technology frauds.	3 [0.4]
2011	Bhattacharya et al.	An ANN-based auditor decision support system using Benford's law.	11 [1.4]
2011	Glancy & Yadav	A computational model for financial reporting fraud detection.	43 [5.4]
2011	Humpherys et al.	Identification of fraudulent financial statements using linguistic credibility analysis.	84 [10.5]
2011	Jans et al.	A business process mining application for internal transaction fraud mitigation.	50 [6.3]
2011	Ravisankar et al.	Detection of financial statement fraud and feature selection using data mining techniques.	117 [14.6]
2011	Wozniak	A hybrid decision tree training method using data streams.	35 [4.4]
2011	Zhou & Kapoor	Detecting evolutionary financial statement fraud.	39 [4.9]
2010	Cecchini et al.	Detecting management fraud in public companies.	35 [3.9]
2010	Cecchini et al.	Making words work: Using financial text as a predictor of financial events.	54 [6]
2009	Holton	Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multi-billion dollar problem.	46 [4.6]
2007	Kirkos et al.	Data Mining techniques for the detection of fraudulent financial statements.	186 [15.5]
2006	Kotsiantis et al.	Forecasting fraudulent financial statements using data mining.	40 [3.1]
1999	Cerullo & Cerullo	Using neural networks to predict financial reporting fraud: Part 1.	4 [0.2]
1999	Cerullo & Cerullo	Using neural networks to predict financial reporting fraud: Part 2.	n/a
1996	Hansen et al.	A generalized qualitative-response model and the analysis of management fraud.	44 [1.9]
1990	Mercer	Tailor-made auditing of information systems for the detection of fraud.	n/a

*Times Cited source is Web of Science (search performed in the fourth quarter of 2019).

n/a = not available

[bracket] = annual citations, calculated as times cited divided by paper age. Paper age is calculated as 2019 minus publication year (Li & Dang, 2018).