

IDENTIFYING THE INTELLECTUAL CORES OF AIS RESEARCH IN IS DISCIPLINE THROUGH A LATENT SEMANTIC ANALYSIS

Yibai Li

Washington State University
Todd 442, Pullman, WA 99164
Email: yibai@wsu.edu
Phone: 509-335-7822

Xuequn Wang

Dongbei University of Finance & Economics
217 Jian Shan Street, Sha He Kou District, Dalian, China, 116025
Email: xuequnwang1600@gmail.com

Todd Christopher Martin

Washington State University
Todd 442, Pullman, WA 99164
Email: todd.c.martin@wsu.edu
Phone: 509-335-5319

Xiaolin Lin

Washington State University
Todd 442, Pullman, WA 99164
Email: xiaolin@wsu.edu
Phone: 509-335-7822

ABSTRACT

Accounting Information Systems research has existed since the birth of the IS field, being of importance to both practice and education. Extensive literature reviews show that AIS research as a whole has been improved in scientific rigor while maintaining the leadership role in practice in the past a few decades. Meanwhile, the AIS field is still facing some challenges. Recent literature reviews report that AIS field remains highly diverse, often considered as “a research discipline with no boundaries”. As a result the definition of AIS research remains unclear. In order to provide some clarity, this study applied a Latent Semantic Analysis on all published AIS research within the IS Discipline and identified 42 research themes. We further classify the 42 research themes into 7 categories using the stakeholder framework. The evolution of these 7 research categories over the past 45 years is discussed.

Keywords: AIS, Accounting Information Systems, Latent Semantic Analysis, LSA, Literature Review

INTRODUCTION

Accounting Information Systems (AIS), recognized as the intersection of the accounting and information systems fields (McCarthy 1990; Poston and Grabski 2000; Sutton 1992, 2010), has

been an important component of IS research since its formation. Keen's pioneering speech (1980) in the first International Conference on Information Systems (ICIS) recognized the significance of AIS research in Information Systems (IS) field. Since then, the AIS area has retained its significant status because of its importance to practice and education (Poston et al. 2000). As firms increase their utilization of accounting software packages such as SAP and Oracle Financials, which aid in revenue generation and the execution of global expansion strategies, there is an increased demand for employees with AIS knowledge and skills. This ultimately leads to more students in business schools being attracted to enroll in AIS courses.

As the AIS area remains highly relevant to practice and education, it also produces research with high scientific rigor (Sutton 2010). Recent literature reviews (Ferguson and Seow 2011; Sutton 2010) reported that AIS research published in top AIS journals consistently has stronger underlying theories and research methods than it did several decades ago. Sutton (2010) stated that the comparative advantage of AIS over its sister discipline of accounting is their strength to keep the rigor and relevance at the same time.

Despite the achievements that AIS research accomplished over the past decades, some of the issues that appeared in AIS decades ago are still present. One of the most salient issues is the high degree of diversity of AIS research. On one hand, diversity is an indicator of the prosperity and creativity of a research field. On the other hand, the excess degree of diversity may blur the boundaries of the area, dilute its intellectual cores and eventually negatively impact the identity, the legitimacy, and the image of a research area from its stakeholders' perspective. The diversity issue of AIS can be traced back to Sutton's paper (1992) titled "Can we research a field we cannot define?" In the year 2000, Poston and Grabski (2000) found that there was still "*no specific definition of what constitutes AIS research, there seems to be consensus...*" Up until the year 2010, Sutton still (2010) referred to AIS as "*a research discipline with no boundaries.*"

What defines AIS research? As an attempt to answer this question, this study takes an inductive approach. We argue that as a socially constructed field, the core aspects of AIS can be identified from the work conducted and published by members of the AIS community. It is the research cores of AIS that define the AIS research. In this paper, guided by the stakeholder framework, we seek to define the boundaries of AIS research within the IS discipline by distilling the intellectual core from the multitude of individual research papers using latent semantic analysis (LSA).

As discussed above, AIS is an interdisciplinary area. It is positioned at the intersection of the accounting and IS disciplines. Therefore, the outlets that publish AIS research often cross the disciplines of IS, accounting, computer science, as well as others. The goal of this study doesn't ambitiously intend to identify the intellectual cores of AIS from all these disciplines. Instead, we examine all the published AIS research just within IS discipline. More specifically, we collected all AIS literature published in 9 top IS journals, 4 IS conferences, 4 crossover IS journals and 2 practitioners' journals. Previous studies (Adams and Johnson 2008; Dubé and Paré 2003) show that these outlets are a comprehensive representation of the whole IS discipline.

This study then conducted Latent Semantic Analysis (LSA) on the collected literature. We found that (1) AIS is indeed a very diverse research area as 42 research themes were identified. Each of

the research themes derived from LSA conceptually represents a unique or orthogonal dimension of the area and has been labeled by the researchers. (2) AIS research is fragmented. The largest one among these 42 research themes only accounts for 2.7% of the entire AIS area. (3) The AIS area is becoming mature and stable. As we further classified the 42 research themes into the 7 categories under the stakeholder framework, it shows that AIS research converges to these 7 categories with a different emphasis on each.

Although extensive literature reviews for the AIS field at different time frames are available in existing literature (Ferguson et al. 2011; Granlund 2010; Poston et al. 2000; Sutton 1992, 2010), this study makes a novel contribution in terms of (1) the time frame being examined, (2) the objectivity and (3) the purity of the inductive method. By adopting the Latent Semantic Analysis, an objective and mathematical natural language processing method, this study is able to synthesize the published AIS literature within the IS field. LSA also allows this study to synthesize the research themes in a more inductive manner, because important research themes will cluster together by the factor loadings and manifest themselves by their eigenvalues.

The rest of the paper is organized as follow: First, we introduce the theoretical framework that guides our study. Then, The LSA method and the operationalization of the LSA are introduced. Finally, we report the result from LSA and conclude.

THE THEROTICAL FRAMEWORK

AIS research can be regarded as an area that deals with the entire information lifecycle that focuses on transaction processing and accountability (Borthick 1992; McCarthy 1990; Poston et al. 2000). In the information lifecycle from data entry, data storage, and to data views, several stakeholders are involved. A prior study (Xu Nord Nord and Lin 2003) identified five stakeholders of the accounting information systems: Information Producers, Information Custodians, Information Consumers, Information Managers, and Internal Auditors. In addition to these stakeholders, we identified one additional stakeholder, the scholars and educators within the AIS discipline. Additionally, we differentiated internal and external information consumers. We propose a 7 stakeholder framework of the AIS area with the definitions of each stakeholder given in Table 1.

TABLE 1. THE STAKEHOLDERS OF AIS FIELD (ADOPTED AND REVISED FROM XU ET AL. (2003))

Stakeholders of AIS Field	Definitions
Information Producers	Information producers are those who create or collect information for the accounting information system (AIS).
Information Custodians	Information custodians are those who design, develop, and operate the AIS. System developers, software package vendors fall into this category.
Internal Information Consumers	Information consumers are those who use the accounting information in their work activities within an organization.

External Information Consumers	Information consumers are those who use the accounting information in their work activities outside an organization.
Information Managers	Information managers are those who are responsible for managing the information quality in AIS.
Internal Auditors	Internal auditors are those who monitor the information quality in AIS.
AIS Scholars and Educators	Scholars who are research AIS and educators who teach AIS courses and develop the curriculum of the AIS major.

The stakeholder framework not only provides us a theoretical lens to profile AIS as a whole, but also helps us examine how AIS research serves each involved stakeholders. According to the organizational identity literature (Gioia Schultz and Corley 2000; Scott and Lane 2000), as an organizational image is developed, it leads key internal and external stakeholders to reexamine their perception of the organization, acting as a catalyst for change in organizational identity (Gioia et al. 2000). The image of AIS research in the external stakeholders' view is very important because they selectively provide research and funding opportunities for the field. As suggested by Allen Lee (1999), "all science is socially constructed." AIS area is constructed through a process of negotiation among its various stakeholders. Internal stakeholders such as AIS scholars and educators construct the desired image through their publications and the class content they teach. This 7 stakeholder framework is used to make sense of the results from the Latent Semantic Analysis conducted in the study.

METHOD

Data Collection

The objective of this study is to synthesize the AIS research within IS discipline. To serve this objective, we need to first answer two questions. (1) What outlets are a good representation of IS discipline? (2) What qualifies as AIS research? Based upon previous discussion of IS publication outlet lists (Adams et al. 2008; Dubé et al. 2003), we compiled our list which includes three categories of outlets: IS academic, IS crossover, and IS practitioner (shown in Table 2). The IS academic outlets include 9 top IS journals and 4 major conference proceedings. In addition, we added 3 crossover journals and 2 practitioners' publications.

TABLE 2. PUBLICATION OUTLETS INCLUDED IN THE LITERATURE SEARCH

IS Journals	Decision Support Systems
	European Journal of Information Systems (EJIS)
	Information & Management (I&M)
	Information Systems Research (ISR)
	Journal of Information Technology (JIT)
	Journal of Management Information Systems (JMIS)
	International Journal of Electronic Commerce (IJEE)
	Journal of Strategic Information Systems (JSIS)

	Management Information Systems Quarterly (MISQ)
IS Conferences	Americas Conference on Information Systems (AMCIS)
	International Conference on Information Systems (ICIS)
	PACIS Pacific Asia Conference on Information Systems (PACIS)
	European Conference on Information Systems (ECIS)
IS Crossovers	Communications of the ACM (CACM)
	Communications of the Association for Information Systems (CAIS)
	IEEE Software (IEEE S/W)
IS Practitioners	CIO Magazine (CIO)
	Computerworld Magazine (CW)

In terms of what constitute a qualified AIS research, various researchers have given their criteria (Borthick 1992; McCarthy 1990; Sutton 1992). However, “*no widely accepted and agreed upon definition of AIS research exists.*” (Poston et al. 2000). Different from the common practice used in prior AIS literature reviews (Ferguson et al. 2011; Granlund 2010; Poston et al. 2000; Sutton 1992, 2010) in which the researchers’ subjective judgment was utilized to determine whether a particular paper should be an AIS study, we used a strict criterion to finalize out literature database. We only selected papers that contained the term “accounting” in either their author supplied key words or titles. The rationale behind this manner is that (1) the search term “accounting” will return all the literature that contains accounting related terms; (2) if one doesn’t claim itself as accounting related research, we should not count it as one; (3) this method is not subject to researchers’ bias. Although this criterion is still arguable and returns much less literature from the database, it provides us a sample containing “pure” AIS research. We searched the EBSCO database and the AIS Electronic Library (AISeL), using this criterion and retrieved 91 papers. All the abstracts from these papers were consolidated into a spreadsheet which was used as the dataset in the data analysis.

Data Analysis

The dataset was analyzed with Latent Semantic Analysis (LSA), a well-accepted text mining technique (Han Kamber and Pei 2011), utilizing one of the leading data mining tools – the Rapidminer 5.0 (Jungermann 2009). In recent years, text mining has been increasingly used for knowledge discovery from scholarly literature (Delen and Crossland 2008; Jensen Saric and Bork 2006; Mei and Zhai 2005; Turban Sharda Aronson and King 2008). Among all kinds of text mining techniques, Latent Semantic Analysis (LSA) is a special mathematical and statistical method used to identify the latent concepts within the textual data at the semantic level (Hossain Prybutok and Evangelopoulos 2011). In contrast to many other text mining techniques which analyze textual data at the syntax level by simply counting the occurrence of particular words, LSA is a methodology that can extract the contextual-usage meaning of words and obtain approximate estimates of meaning similarities among words within the given textual data, thus providing the information at the semantic level (Hossain et al. 2011). LSA has numerous applications in natural language processing, search engine and library indexing and many other areas (Hossain et al. 2011).

LSA is based on the fact that multiple words may share the same meaning and one word may mean different things in different contexts. The main concept behind LSA is similar to that of factor analysis. The words that share the same meaning will “load” to their common underlying concept, just like multiple manifest variables will load to their latent factor in factor analysis; one word may “load” to multiple latent concepts other than its main underlying concept, just like one manifest variable may have cross-loadings with multiple latent factors in the factor analysis. Therefore, we can easily draw an analogy between LSA and factor analysis.

The interpretation of LSA results is also similar to the interpretation of factor analysis. LSA generates two sets of loadings, one for the terms (or words) and one for the documents (the definitions or abstracts in this study). The term loading shows how individual terms or words load to different latent concepts. Higher term loading reflects the greater chance that the particular term is truly associated with a certain latent concept. Likewise, the document loadings show how different documents load to different latent concepts. Higher document loading means a greater likelihood that the particular document is truly talking about a certain latent concept. LSA also generates an eigenvalue matrix which shows the importance of all identified latent concepts. A higher eigenvalue is associated with a greater importance of particular latent concepts. A mathematical explanation for LSA can be found in previous studies (Sidorova Evangelopoulos Valacich and Ramakrishnan 2008).

This study follows the well-established text mining procedures implemented and discussed in prior studies (Delen et al. 2008; Fox 1992; Han et al. 2011; Harman 1992; Hossain et al. 2011; Sidorova et al. 2008; Turban et al. 2008). First, the spreadsheet containing all the abstracts was converted into a document object in Rapidminer 5.0. Then the following pre-processing procedures were taken. (1) All letters in these documents were transformed into lower case. (2) The documents then were tokenized with non-letter separators. (3) We removed the “stopwords” such as “and,” “the,” “is,” “a,” “an” and so on. (4) All the tokens that are less than two letters (i.e. “s,” “x,” and so on) were removed. (5) We also removed the words or tokens that appear only in one document, because these tokens are associated only with the specific study and shouldn’t be considered as a reflection of any research theme. (6) We applied term stemming techniques to word list. Terms stemming will identify the root of the words and regard all words with the same root as one token. For example, “collaborate,” “collaborating,” “collaboration,” and “collaborative” will be regarded as a single token, the “collabor-.” By doing so, different variants of the same word are combined and the dimensionality is further decreased. (7) We removed “author,” “paper,” “conclusion” and some other words that are associated solely with the writing style of scholarly articles and don’t provide additional information about the content. (8) Finally, we use n-gram algorithm to identify the terms that contain multiple words. All these term reduction steps eventually resulted in a word list with 487 tokens.

Following the pre-processing procedures, all documents are converted into a term frequency by document matrix. We further transformed the values in the matrix using TF-IDF (term frequency – inverse document frequency) weighting method (Han et al. 2011; Harman 1992; Husbands Simon and Ding 2001; Salton and Buckley 1988; Salton Wong and Yang 1975). This approach puts more weight on the rare terms and discounts the weight of the common terms such as “accounting” that appears in every document (Sidorova et al. 2008).

We then applied singular value decomposition to convert the TF-IDF weighted term matrix into the production of three matrices, the term-by-factor matrix, the singular value matrix (square roots of eigenvalues), and the document-by-factor matrix. The first two matrices are more relevant to the objective of this study. The discussion about them is given in the results section.

RESULTS AND DISCUSSION

We first analyze the number of AIS publication in IS over time. Figure 1 illustrates AIS publications at five year intervals from 1966 to 2010 for IS journals, conferences, cross-overs and practitioners' outlets. The total line presents the sum of all AIS publications in IS at each period. It can be seen that, although AIS research in the IS discipline can be traced back to 1960s; the prosperity of the area began in the early 1990s. Through separating the publications by different categories of outlets, we found that AIS research first appeared in the IS crossover outlets in the period from 1966 to 1970 and then disseminated to major IS journals in the period from 1976 to 1980. In the same period, the IS field was formally established. From 1991 to 2005, the practitioners' outlets and conferences seemed to be leading the discussion of the AIS and produced a relatively greater number of publications. Since 2006, the practitioners have seemingly lost some interest in AIS but the conferences have still actively published AIS research. Overall, the AIS research went through the peak era from 1991 to 2000 and remains at a relative high level of activity.

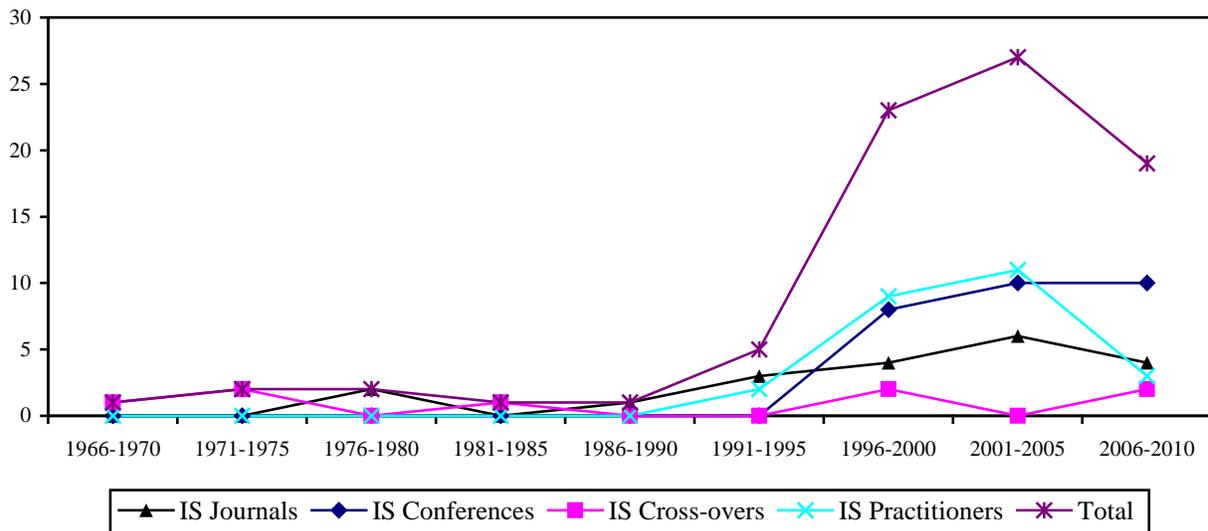


FIGURE 1. THE NUMBER OF PUBLICATIONS ON SOCIAL COMPUTING FROM 1966-2010

In order to identify the major research themes of AIS research IS discipline we further analyzed the results derived from LSA. We utilized the 0.8 as the cut-off point as suggested by prior studies (Sidorova et al. 2008) when judging the importance of each research theme and obtained 42 major factors in AIS research. These 42 factors are treated as major factors as they account for 80% of the variance within all the content of published AIS research. The remaining 20% of the variance is often considered as just noise or trivial meanings that are associated only with

different writing styles of different papers (Sidorova et al. 2008). The factors listed in Table 3 are ranked in order of their importance from high to low. The importance of each research theme or factor is delineated by their singular value and variance explained.

This result confirms that AIS is indeed a very diverse research area in which 42 unique research themes exist. Each theme is derived from an orthogonal matrix in LSA; therefore it conceptually represents a unique dimension of the AIS research. To interpret these factors or research themes, we follow the guidance from prior study (Sidorova et al. 2008) to label the factors using the top loading terms. Two raters labeled the factors independently based upon the high-loading terms. Consensus was reached and the finalized labels are listed in Table 3. The LSA also shows that AIS research is fragmented. Table 3 shows that the largest one among these 42 research themes only accounts for 2.7% of the entire area.

TABLE 3. 42 RESEARCH THEMES FOUND IN LITERATURE

Factors	Singular Values	Variance Explained	Factor Labels
Factor 1	1.289	2.7%	Firm Performance
Factor 2	1.223	2.6%	Non-financial Performance
Factor 3	1.151	2.5%	Accounting standards link clients and firms
Factor 4	1.122	2.4%	The quality of accounting information within intranet
Factor 5	1.077	2.3%	Production cost of an internet based system in a hospital environment
Factor 6	1.057	2.3%	ERP solutions
Factor 7	1.027	2.2%	Accounting standards
Factor 8	0.998	2.1%	Software tools
Factor 9	0.982	2.1%	Measurement or method reliability
Factor 10	0.968	2.1%	ERP usefulness for productivity
Factor 11	0.96	2%	Financial information visualization
Factor 12	0.927	2%	Accounting system implementation
Factor 13	0.917	2%	AIS curriculum
Factor 14	0.913	1.9%	AIS discipline in business school
Factor 15	0.904	1.9%	Automated external reporting
Factor 16	0.894	1.9%	Automated problem solving
Factor 17	0.891	1.9%	Internet based application architecture
Factor 18	0.871	1.9%	Reporting to vendor and customers
Factor 19	0.866	1.8%	The quality of AIS programs in business school
Factor 20	0.847	1.8%	Automated prediction to customers and system users
Factor 21	0.84	1.8%	User participation
Factor 22	0.828	1.8%	Automated recording
Factor 23	0.82	1.8%	User and manager participation to the design of the curriculum in AIS
Factor 24	0.817	1.7%	A Framework for an Accounting Information

			Systems Curriculum
Factor 25	0.801	1.7%	Implementation management
Factor 26	0.793	1.7%	Project management
Factor 27	0.787	1.7%	The relation between AIS discipline to practitioners
Factor 28	0.764	1.6%	The reliability of the computer generated financial statement
Factor 29	0.75	1.6%	The learning challenges of AIS education
Factor 30	0.741	1.6%	Accounting practice
Factor 31	0.725	1.5%	Web-based automated tools
Factor 32	0.709	1.5%	The reliability of web-based tools
Factor 33	0.706	1.5%	The aggregated uncertainty of information technology
Factor 34	0.699	1.5%	The level of reliability of measurement tools
Factor 35	0.691	1.5%	The quality of the customized web-based applications
Factor 36	0.683	1.5%	Web-based applications and accounting practice
Factor 37	0.672	1.4%	Expert systems
Factor 38	0.66	1.4%	Web-based accounting practice
Factor 39	0.653	1.4%	Automated source document entries
Factor 40	0.648	1.4%	Supply chain strategy
Factor 41	0.631	1.3%	The challenges of supply chain system implementation
Factor 42	0.625	1.3%	Panel discussions and AIS students

We then further classified the 42 research themes under the 7-category stakeholder framework proposed previously. Two raters conducted the classification independently based on the definitions given in Table 1. The framework helps us examine how the AIS literature has contributed to and served different stakeholders over the past 45 years. We found that AIS literature did serve every single stakeholder but with different emphasis.

TABLE 4. THE RESEARCH THEMES CLASSIFICATION UNDER STAKEHOLDER FRAMEWORK

Stakeholders Categories	42 Research Theme Categorization	High Loading Terms
Information Producers	22, 30, 36, 38, 39,	Automatic, Record, Accounting Practice, Measure, Web, Entry and etc.
Information Custodians	6, 8, 10, 12, 16, 17, 21, 25, 26, 31, 32, 34, 35	ERP, Solution, Software, Tool, Problem Solving, Implementation and etc.
Internal Information Consumers	1, 2, 5, 6, 11, 20, 33, 40,	Firm Performance, Production, Cost, Strategy and etc.
External Information Consumers	11, 18, 20, 33, 37,	Financial Information, Reporting, Vendor, And Customers and etc.
Information	3, 4, 7, 9, 28, 35,	Accounting Standard, Client, Firm

Managers		Production, Quality and etc.
Internal Auditors	3, 7, 9, 30, 32, 33, 34,	Standard, Reliable, Accounting Practice, Quality and etc.
AIS Scholars and Educators	13, 14, 19, 23, 24, 29, 42	Discipline, Curriculum, Learn, and etc.

Figure 2 shows how AIS literature is distributed into each stakeholder category. Because there are 7 categories, any category that accounts for more than 14.28% is considered above average. The distribution shows that AIS literature has definitely been biased towards the information custodians, internal consumers, and the information managers while much less attention has been given to the research for information producers, external information consumers, scholars and educators.

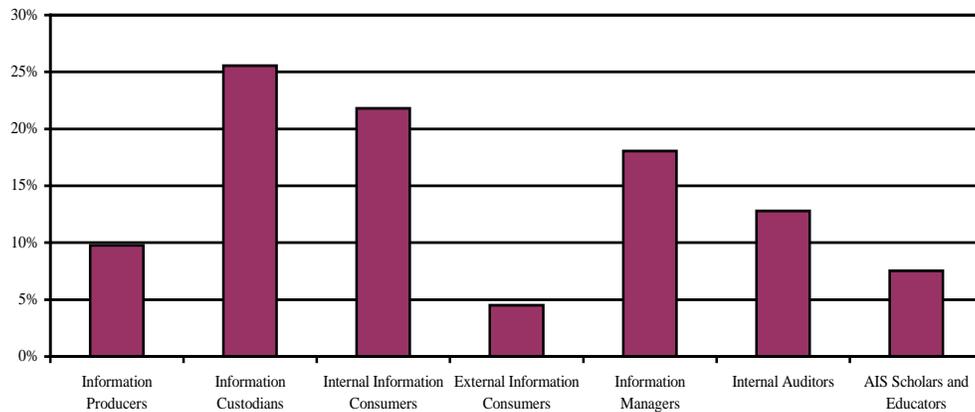


FIGURE 2. DISTRIBUTION OF AIS RESEARCH UNDER THE STAKEHOLDER FRAMEWORK

The evolution of AIS research under the 7-stakeholder framework is depicted in Figure 3. It shows that AIS literature does have different emphasis. The number of the publications that deal with different stakeholders fluctuates over the time. Before 1995, the number of publications is few in all categories. From 1996 to 2000, the information custodian category was the “super star” and enjoyed much more attentions from scholars than other categories did. From 2001 to 2005, internal information consumers and information managers surpassed information custodians and became the most important stakeholder in AIS literature. From 2006 to 2010, AIS area became much more stable and distributed quite even effort on all stakeholders except for the external information consumers. This confirms the finding from prior study (Sutton 2010) which states that AIS area is becoming a more mature research area.

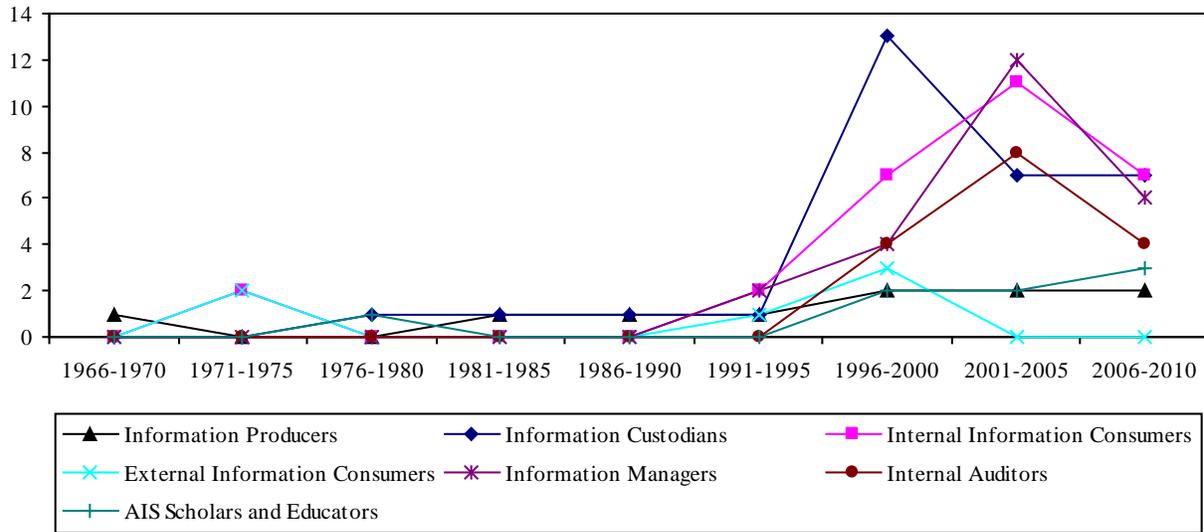


FIGURE 3. THE EVOLVEMENT OF THE AIS UNDER THE STAKEHOLDER FRAMEWORK FROM 1966 TO 2010

LIMITATIONS

We acknowledge that there are limitations in our study. First, we collected AIS publications from only the 18 outlets listed in Table 2 as the sample. Clearly, there will be other AIS-related articles published in other outlets. However, we think that, given that the objective of this study is to uncover the intellectual cores of AIS research within the IS discipline, this sample serves this objective quite well. Second, the results of this study are confined by the key words used to search and collect the papers. We only selected the papers that had the term “accounting” either in their author-supplied key words and titles. There are definitely many papers can be regarded as AIS research without specifically mentioning “accounting.” However, given that AIS research is an area with no consensus on its definition (Poston et al. 2000), we used the rationale that if one doesn’t claim itself to be AIS research, we shouldn’t use our own judgment to label it as AIS research. Third, although this study uses an advanced method, Latent Semantic Analysis, to conduct a systematic and extensive analysis of publications in the AIS area, it still needs researchers to label the research themes, in which researchers’ bias may be involved.

CONCLUSIONS

To uncover the intellectual cores of AIS research within the IS discipline and to partially address the diversity and identity problem of the AIS field, this study takes an inductive approach using Latent Semantic Analysis (LSA) to synthesize the AIS literature published in the IS discipline from all time periods. LSA enjoys the advantage that it is an objective way of performing a literature review. We found from the results that (1) AIS is indeed a very diverse research area. The entire area has 42 research themes. Each of which has been labeled by the researchers. (2) AIS research is fragmented. (3) AIS area is becoming mature and stable in the regards to its contribution to the stakeholders. Future research might examine the AIS literature in the Accounting discipline.

REFERENCES

- Adams, D., and Johnson, N. (2008) The journal list and its use: motivation, perceptions, and reality, *European Journal of Information Systems*, 17, 2, 158-162.
- Borthick, A. (1992) Helping users get the information they want, when they want it, in the form they want it: Integrating the choice and use of information, *Journal of Information Systems*, 6, 2, 75-85.
- Delen, D., and Crossland, M.D. (2008) Seeding the survey and analysis of research literature with text mining, *Expert Systems with Applications*, 34, 3, 1707-1720.
- Dubé, L., and Paré, G. (2003) Rigor in information systems positivist case research: Current practices, trends, and recommendations, *Mis Quarterly*, 597-636.
- Ferguson, C., and Seow, P.S. (2011) Accounting information systems research over the past decade: Past and future trends, *Accounting & Finance*.
- Fox, C. (1992) Lexical analysis and stoplists, *Information retrieval: Data structures and algorithms*, 7, 102-130.
- Gioia, D.A., Schultz, M., and Corley, K.G. (2000) Organizational identity, image, and adaptive instability, *Academy of Management Review*, 63-81.
- Granlund, M. (2010) Extending AIS research to management accounting and control issues: A research note, *International Journal of Accounting Information Systems*.
- Han, J., Kamber, M., and Pei, J. (2011) Data mining: concepts and techniques Morgan Kaufmann.
- Harman, D. (1992) Ranking algorithms, *Information retrieval: Data structures and algorithms*, 363-392.
- Hossain, M.M., Prybutok, V., and Evangelopoulos, N. (2011) Causal Latent Semantic Analysis (cLSA): An Illustration, *International Business Research*, 4, 2, p38.
- Husbands, P., Simon, H., and Ding, C.H.Q. (2001) On the use of the singular value decomposition for text retrieval, *Computational information retrieval*, 145-156.
- Jensen, L.J., Saric, J., and Bork, P. (2006) Literature mining for the biologist: from information retrieval to biological discovery, *Nature reviews genetics*, 7, 2, 119-129.
- Jungermann, F. (2009) Information extraction with rapidminer, Citeseer, 50-61.
- Keen, P.G.W. (1980) MIS research: reference disciplines and a cumulative tradition Center for Information Systems Research, Alfred P. Sloan School of Management.

Lee, A.S. (1999) *Researching mis* Oxford University Press, Oxford.

McCarthy, W. (1990) The journal of information systems editorial guidelines, *Journal of Information Systems*, 4, 2.

Mei, Q., and Zhai, C. (2005) Discovering evolutionary theme patterns from text: an exploration of temporal text mining, in: *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, ACM, Chicago, Illinois, USA, 198-207.

Poston, R.S., and Grabski, S.V. (2000) Accounting information systems research: Is it another QWERTY?, *International Journal of Accounting Information Systems*, 1, 1, 9-53.

Salton, G., and Buckley, C. (1988) Term-weighting approaches in automatic text retrieval, *Information processing & management*, 24, 5, 513-523.

Salton, G., Wong, A., and Yang, C.S. (1975) A vector space model for automatic indexing, *Communications of the ACM*, 18, 11, 613-620.

Scott, S.G., and Lane, V.R. (2000) A stakeholder approach to organizational identity, *Academy of Management Review*, 43-62.

Sidorova, A., Evangelopoulos, N., Valacich, J.S., and Ramakrishnan, T. (2008) Uncovering the intellectual core of the information systems discipline, *Mis Quarterly*, 32, 3, 467-482.

Sutton, S.G. (1992) Can we research a field we cannot define, *Advances in accounting information systems*, 1, 1, 1-13.

Sutton, S.G. (2010) A research discipline with no boundaries: Reflections on 20years of defining AIS research, *International Journal of Accounting Information Systems*, 11, 4, 289-296.

Turban, E., Sharda, R., Aronson, J.E., and King, D.N. (2008) *Business intelligence: a managerial approach* Pearson Prentice Hall.

Xu, H., Nord, J.H., Nord, G.D., and Lin, B. (2003) Key issues of accounting information quality management: Australian case studies, *Industrial Management & Data Systems*, 103, 7, 461-470.