The Intellectual Core of the IS Field:
A Systematic Exploration of Theories in Our Top Journals

Daniel S. Soper and Ofir Turel
Information Systems and Decision Sciences Department
California State University, Fullerton
dsoper@fullerton.edu   oturel@fullerton.edu

Nitza Geri
Department of Management and Economics
The Open University of Israel
nitzage@openu.ac.il

Abstract

The information systems (IS) field has relied on a broad and varied collection of theories. The core of this theoretical landscape is an important determinant of the identity of the IS field, and has hence been discussed extensively in recent years. Nevertheless, little is known regarding the composition, consistency, and evolution of this theoretical core. Here we address these issues empirically by identifying the theories and levels of analysis which constitute the core of the IS field, and by quantifying the degree of theoretical uniformity among the field's top journals. Using a set of 87 theories in conjunction with n-gram analyses, we analyze the full text of every research article published in three top IS journals over a 22-year period in order to identify the field's most frequently and consistently used theories and theory clusters. We further identify substantial overlap among the theoretical core of the field's leading journals, but also note differences in the theoretical character of each journal.

1. Introduction

One of the most interesting and distinguishing traits of the information systems (IS) field[1] is its reliance on a vast patchwork of theories drawn from virtually all realms of scientific and managerial inquiry [1-3]. This situation is perhaps not surprising, especially when one considers the immense variety of behavioral, technological, and organizational phenomena that have come to be arrayed under the expansive and ever-growing banner of IS research. Although this theoretical diversity may in certain ways be advantageous for the field (e.g., by fostering creativity or innovative modes of thought among IS researchers) [4], it also presents a serious challenge in that it makes the intellectual core and boundaries of the field (i.e., the field's identity) difficult to describe [2]. Whereas many disciplines and fields are defined by a relatively small set of macro-theoretic paradigms, the fragmentation and lack of cohesion among the myriad theories used in IS research makes defining the field quite difficult, and complicates not only efforts aimed at understanding how the field has evolved over time, but also attempts to make informed predictions about the directions in which the field might be moving in the future [2, 5, 6].

We do not take a position in this study regarding the debate surrounding the merits of theoretical diversity in the IS field, and neither do we endeavor to support or refute claims made by past research on this topic. Rather, we seek here to describe the current state of theoretical diversity in the field through a systematic, data-driven exploration of the theories appearing in three of the field's top research journals. To this end, we employ advanced computational techniques in order to quantify the extent to which different theories and theory clusters have appeared in our top journals over time. Specifically, we describe herein a study designed to answer three key questions that have certainly crossed the minds of many IS scholars and that have important implications for IS identity research, but which have largely gone unanswered in the IS research literature. To wit:

1) Which theories and theory clusters comprise the intellectual core of the IS field?
2) How have the intellectual core of the IS field and the levels of analysis of its theories evolved over time?
3) To what extent do leading IS journals exhibit theoretical uniformity (i.e., a similar theoretical orientation toward the intellectual core of the field)?

---

[1] Consistent with Hassan (2011) we use the term "field" rather than "discipline", since the former is more appropriate for the current state of affairs in IS studies.

IEEE computer society

Answering the first two of these three questions will reveal deep insights into the current state and historical evolution of the IS field and its theoretical core, while answering the last of these questions will reveal not only the degree of theoretical heterogeneity (or homogeneity) among the field's top journals, but also by extension each journal's unique theoretical persona.

In the sections that follow, we describe the methodology for our study and present and discuss the answers to the above research questions. Before reporting these details, however, it is first necessary to expand upon the foundations of our research by exploring the philosophical and conceptual links between theory and identity.

2. C between Theory and Identity

While the debates regarding the boundaries of the IS field and their associated exclusion and inclusion criteria may be ongoing, research in this realm largely agrees that a key component of the field's identity is its intellectual core, including particularly the primary theories used and advanced by the field [1, 3, 5, 6]. Not only do the theories used by a field signal the field's boundaries [2], but they also help to define the cohesiveness and rigidity of the field (or discipline) [3], and delimit the field's knowledge and academic domains [6].

Given the vital role of theory in defining the identity of the IS field, it is critical to understand which theories and theory clusters have been used in the field, the extent to which those theories have been used, and how their patterns of use have changed over time. Moreover, since consistency in the theoretical core and heterogeneity in the theoretical core have alternately been argued to be both a necessity and an impediment to the field [1, 3, 5, 6], it is important to gain objective insights regarding the homogeneity of the theoretical core of the field as a whole, and in specific journals in particular. Not only can addressing these issues provide insights for the study of the identity of the IS field, but the results can also serve as defensible scientific input for arguments regarding the field's intellectual core.

Certain studies conducted hitherto have acknowledged these needs, and have endeavored to use conceptual or empirical methods to identify elements of the theoretical core of the IS field. Some studies, for example, have used anecdotal evidence [2, 4], while others have employed citation analyses [7], manual literature reviews [1], or manual analyses of titles and abstracts [5] to identify various aspects of the theoretical core of the IS field. Latent categorization has also been applied to abstracts to identify intellectual communities within the IS field [8]. While such efforts serve as important preliminary steps in mapping the theoretical core of the field and its evolution over time, we believe that a computational, automated, and systematic approach has the potential to generate a more complete and accurate picture of at least one aspect of the theoretical core – the theories that are mentioned and used in our leading IS journals. Inasmuch as they define the directions and core topics upon which the field focuses its attention, these journals arguably represent the mainstream IS field as a whole [9, 10]. To that end, the findings of studies such as that described herein can serve as an objective, quantitative basis upon which the exploration of the identity of the IS field can proceed.

3. Method

Our general strategy for inquiring into the extent to which different theories have influenced the IS field was to first identify a large set of theories used in IS research, and then assess the frequency with which those theories appear in the IS literature over time by applying a computational text mining technique known as *n-gram analysis*.

As used in natural language processing, an *n-gram* is a sequence of words of length *n* that is extracted from a larger sequence of words [11]. The unpunctuated national motto of France "liberté égalité fraternité" [12], for example, can be subdivided into three 1-grams (*liberté, égalité,* and *fraternité*), two 2-grams (*liberté égalité* and *égalité fraternité*), and one 3-gram (*liberté égalité fraternité*). The central theoretical tenant underlying n-gram analysis is that the frequency with which a concept is mentioned in a corpus of text provides a clue into the relative importance of that concept within its parent domain at the time when the text was written [13]. For example, if the n-gram "mobile phone" appears 500% oftener than the n-gram "fax machine" during a particular year, one might reasonably infer that the concept of a mobile phone was more important or influential during that year than was the concept of a fax machine.

The true power of n-gram analysis, however, becomes evident only when one considers trends over time. By standardizing each n-gram frequency according to the total quantity of text published during a given year, it becomes possible to identify trends and patterns that reveal powerful insights into the underlying domain of interest. N-gram analysis has been successfully used as the basis of culturomic and scientometric inquiry in many different fields [13, 14], including the IS field [15, 16].

To identify the set of theories that would be used in the analysis, we relied upon the *Theories Used in IS Research* website, which is affiliated with the Association for Information Systems (AIS) and hosted by the Marriott School of Management at Brigham Young University [17]. At the time when theories were being identified for the current project, this resource listed and described 87 unique theories, all of which were adopted for the study. As with most scientific fields, IS researchers frequently use acronyms and multiple names to refer to the same underlying theory (e.g., *Theory of Reasoned Action* and *TRA*). With a view toward ensuring that our analytic results would be as accurate as possible, it was therefore necessary to identify the collection of common names and acronyms that have been used for each theory. The website noted above maintains a list of acronyms and alternative names for many IS theories, and these labels, supplemented with additional synonyms and acronyms obtained from the literature, were hence adopted for use in the current study.

Having identified a set of theories, we next constructed a corpus containing the complete text of every research article published in Management Information Systems Quarterly (MISQ), Information Systems Research (ISR), and the Journal of Management Information Systems (JMIS) during the 22-year period between 1990 and 2011. MISQ, ISR, and JMIS were chosen for inclusion in the corpus because they are consistently ranked among the highest-quality IS journals in the world, and hence serve as a proxy for the mainstream field as a whole [9, 10]. The corpus included articles from the years 1990 through 2011 because 1990 was the first year in which all three journals concurrently existed, while 2011 was the last year for which complete data were available at the time when the corpus was constructed. In total, the corpus spanned 22 years and contained 1,886 articles, together comprising more than 26.4 million words. To put the size of the corpus in perspective, consider that if one were to spend 40 hours per week reading at the average adult rate of 250 words per minute, more than ten months would be required in order to read every research article published by these three journals between 1990 and 2011.

With our corpus complete, we next constructed a custom software system to tokenize the text of each article into a series of n-grams. The number of n-grams that can theoretically be extracted from a large corpus of text greatly exceeds the number of words in the corpus itself, and this situation thus presents serious scaling and performance implications for a corpus containing millions of words. In congruence with past research [13], our analysis was therefore constrained to include n-grams with a maximum length of n = 5. Accepting for acronyms, all of the words in the corpus were converted to lowercase prior to tokenization so as to eliminate problems that might arise due to capitalization. Using this strategy, *Theory of Reasoned Action* would thus be viewed by the system as equivalent to *theory of reasoned action*, while an acronym such as *IS* would be viewed as distinct from the word *is*, thereby ensuring the accuracy of the results.

Broadly speaking, our analysis is predicated on the notion that the degree of influence of each theory is reflected in the relative frequency with which the theory appears in the IS field's top journals over time. It was therefore necessary to compute the frequencies with which the n-grams in the corpus appeared in MISQ, ISR, and JMIS during each year of the analysis. Directly comparing raw frequency counts for the same n-gram across journals or across time would be misleading, however, since doing so would ignore differences in the number of words published by each journal from year to year. It was therefore necessary to calculate relative frequencies for each n-gram by dividing their respective raw frequency counts within each journal for a given year by the total number of words published by those journals during the year in question. This approach was consistent with past research, and yielded a standardized measure of frequency which would allow valid comparisons to be made between n-grams across journals and across time [13]. The standardized frequency values resulting from this process thus indicated how often a particular n-gram appeared in a particular journal during a specific year, relative to the total quantity of text published in that journal during that year.

The result of all of these data extraction and processing tasks was a large database containing nearly 30 million unique n-grams. Since a standardized frequency measure for each of these n-grams was computed for each of the 3 journals and for each of the 22 years of the analysis, the final dataset contained more than 1.96 billion n-gram frequency records. With a simple database query, then, it was possible to identify the relative frequency with which any n-gram of length $n = \{1,2,3,4,5\}$ appeared in MISQ, ISR, or JMIS during a particular year. As a final computational task, we computed the average of the relative frequencies across all three journals for each combination of n-gram and year. This metric served as a mean-stabilized proxy for the relative frequency with which a given n-gram appeared in the IS literature during a particular year.

Finally, we constructed a custom web-based system that enabled us to query, graph, and explore our n-grams database. This system allowed us not only to plot and analyze multiple n-grams simultaneously, but also to combine related search terms into a single result. For example, the search phrase "theory of reasoned action + TRA" would produce output representing the combined frequencies of the n-grams "theory of reasoned action" and "TRA" over time. This approach was used to combine the relative frequencies of all of the labels associated with each theory (i.e., each theory's name, alternative names, and acronyms) into a single result which would accurately represent the overall relative frequency with which a particular theory appeared in the IS literature during a specific year.

4. Results and Discussion

Our first research question focused on identifying the theories and theory clusters that have comprised the theoretical core of the IS field over time. To that end, we begin the presentation of our findings with Table 1 below, which lists the 25 most commonly mentioned theories in the IS field as measured by the average overall relative frequency with which each theory appeared in MISQ, ISR, and JMIS between 1990 and 2011. To aid in interpretation, values in the table are presented as percentages of all of the text appearing within research articles in those three journals during the timeframe in question.

Rank	Theory Name	Overall Average Frequency	Overall Standard Deviation
1	Technology acceptance model (TAM)	0.003045%	0.003882%
2	SERVQUAL	0.002937%	0.008506%
3	Agency theory	0.002611%	0.002667%
4	Transaction cost economics	0.001858%	0.002420%
5	Theory of planned behavior (TPB)	0.001850%	0.003535%
6	Media richness theory	0.001507%	0.003842%
7	Theory of reasoned action (TRA)	0.001013%	0.001218%
8	Real options theory	0.000824%	0.001883%
9	Adaptive structuration theory	0.000753%	0.001412%
10	Critical success factors, theory of	0.000730%	0.001327%
11	Cognitive fit theory	0.000650%	0.001536%
12	Cognitive dissonance theory	0.000649%	0.002080%
13	Social cognitive theory	0.000633%	0.001379%
14	Contingency theory	0.000629%	0.000698%
15	Delone and McLean IS success model	0.000543%	0.001290%
16	Structuration theory	0.000521%	0.001220%
17	Diffusion of innovations theory	0.000514%	0.000704%
18	Absorptive capacity theory	0.000501%	0.001024%
19	Resource-based view of the firm	0.000496%	0.001154%
20	Dynamic capabilities	0.000473%	0.001399%
21	Social exchange theory	0.000457%	0.000817%
22	Hermeneutics	0.000424%	0.001173%
23	Institutional theory	0.000361%	0.000794%
24	Transactive memory theory	0.000352%	0.001431%
25	Chaos theory	0.000327%	0.000615%

Table 1. The theoretical core of the IS field from 1990 through 2011.

As shown in the table, Fred Davis' Technology Acceptance Model (TAM) [18] has the distinction of being the most frequently mentioned theory overall in the IS field between 1990 and 2011, followed closely by the conceptual model of service quality, SERVQUAL [19], and then by Agency Theory [20, 21]. When considering the table, it becomes readily apparent that the field has a highly diverse theoretical foundation. Consistent with the perspectives espoused by Agarwal and Lucas [4], it seems clear that the field has not focused exclusively on user interactions with an IT artifact, but also on the IS function and the organizational impact of IS (as implied by the commonality of the agent-theoretic and transaction cost perspectives). Indeed, the field does not seem to have a single unifying theoretical theme, but rather appears to be comprised of a patchwork of micro and macro perspectives on the uses and impacts of IS.

It is also important to note that there is nearly a full order of magnitude difference between the highest and lowest-ranked theories in the table, thus indicating that the lowest-ranked theories are not just slightly, but rather exponentially less impactful than the highest-ranked theories. This phenomenon is illustrated in Figure 1 below, which depicts the average relative frequency of each theory according to its ordinal impact ranking in Table 1. The figure below also includes a trend line that shows how the comparative impacts of the various theories decline according to an exponential decay function ($R^2 = 0.948$). From this we can conclude that only a handful of theories exert a substantial influence on the IS field, and perhaps are candidates for inclusion in the field's theoretical core. The degree of influence of most theories appearing in the field's top journals appears by contrast to be comparatively marginal.
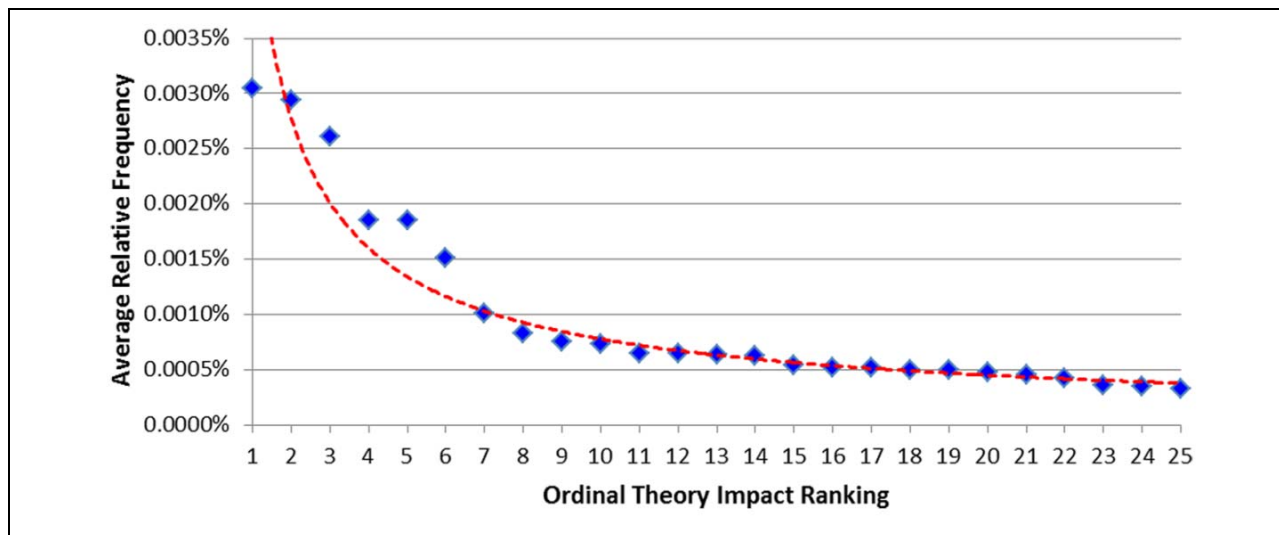

Figure 1. Comparative impact of theories by ordinal ranking in the IS field.

Examining the standard deviations for each theory's overall average relative frequency reveals additional interesting insights (*vide supra*, Table 1), since the standard deviations of the frequencies with which the theories have been mentioned over time can be a good measure of heterogeneity [22]. In contrast to theoretical "fads", theories that appear both frequently in our top journals and in a manner that is relatively stable and consistent over time may be the strongest candidates for inclusion in the theoretical core of the field, and a comparatively small standard deviation may be a good measure of such stability. The standard deviation of TAM, for example, is much smaller than that of SERVQUAL, thus indicating that the frequency with which TAM

has been mentioned in the IS literature has been more stable and consistent over time. A graphical comparison of the longitudinal impacts of these two theoretical models to the IS field is provided as an exemplar in Figure 2 on the following page.

An examination of the trend lines shown in the figure confirms that the impact of TAM has been more consistent over time than that of SERVQUAL, which experienced a spike in interest in 1997. As an aside, note that the overall impact of a given theory can be conceptualized as the area under its associated curve when its relative frequencies are plotted in a fashion similar to that shown in Figure 2. The figure reveals two additional insights to which we would like to draw the reader's attention. First, the relative

impacts of both TAM and SERVQUAL can be seen to ebb and flow in a sinusoidal cycle whose duration is approximately 2 to 3 years. Although this appears to be a common pattern among many of the theories used in the IS literature, the cause of this phenomenon is currently unknown.
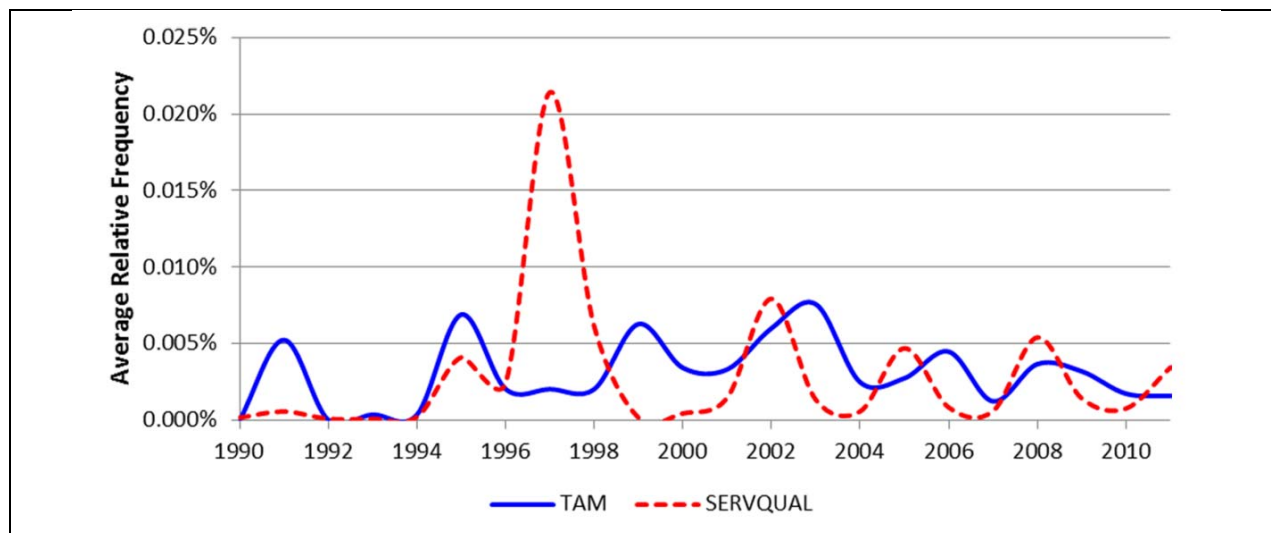

Figure 2. Comparative impact of TAM and SERVQUAL over time.

Second, Figure 2 shows that the relative frequencies with which TAM and SERVQUAL are mentioned in the literature have been declining steadily in recent years. This may indicate that after a period of growth and maturity, these models may now be in the "decline" stage of a theory life cycle that has endured for more than two decades. Further study of these longer-term theoretical trends is clearly needed.

It was noted earlier in the discussion of Table 1 that the IS field has used a combination of theoretical perspectives which employ both micro- and macro-level orientations toward IS phenomena. But to what extent has the IS field relied upon theories at these different levels of analysis? To answer this question, the primary constructs of each of the 87 theories used in the study were evaluated, and each theory was subsequently classified as applying to an individual level of analysis, an organizational level of analysis, or both. It was thence possible to aggregate the relative frequencies of appearance of each of the theories according to their respective levels of analysis. The results of these activities are summarized in Table 2 below, and speak directly to the distribution of theory-driven research within the IS field according to level of analysis.

Level of Analysis	ISR	JMIS	MISQ	All Journals
Individual	39.23%	31.21%	43.39%	39.17%
Organizational	32.90%	39.25%	31.56%	33.84%
Both	27.87%	29.54%	25.06%	26.99%

Table 2. Distribution of theoretical levels of analysis within IS journals from 1990 through 2011.

As shown in the table, research in the field as a whole has used more theories which apply to the individual level of analysis (39.17%) than theories which apply to the organizational level of analysis (33.84%). Moreover, considering each journal individually reveals an element of its unique persona. Whereas the evidence in the table suggests that MISQ and ISR have been more oriented toward the individual level of analysis over time, JMIS has by contrast been more oriented toward the organizational level of analysis. Additional analyses revealed a great deal of statistically random variance in the theoretical levels of analysis from year to year, with interaction tests revealing no statistically significant differences among the slopes of the three levels of analysis over time within ISR (model $p =$

0.984), JMIS (model $p = 0.108$), MISQ (model $p = 0.249$), or within the combined theory frequencies of all journals (model $p = 0.374$). Put another way, there is no statistical evidence to support the notion that the distributions of the three levels of analysis within each journal or across all journals have changed significantly between 1990 and 2011.

Although knowledge of how specific theories and levels of analysis have shaped the IS field is indeed very valuable, deeper insights into the nature and evolution of the field can be gained by examining the theory frequency data at the individual article level. To this end, we began by computing simple descriptive statistics regarding the number of theories appearing in each IS research article during the timeframe of our inquiry, as well as an analysis of how the average number of theories per article has changed over time. The results of these activities are presented in Table 3 and Figure 3.

	n	Theories per Article			
	n	Mean	Std Dev	Min	Max
ISR	513	1.979	1.903	0	11
JMIS	784	2.094	1.950	0	12
MISQ	589	3.192	2.915	0	17
All journals	1886	2.406	2.345	0	17

Table 3. Number of theories per article within IS journals from 1990 through 2011.
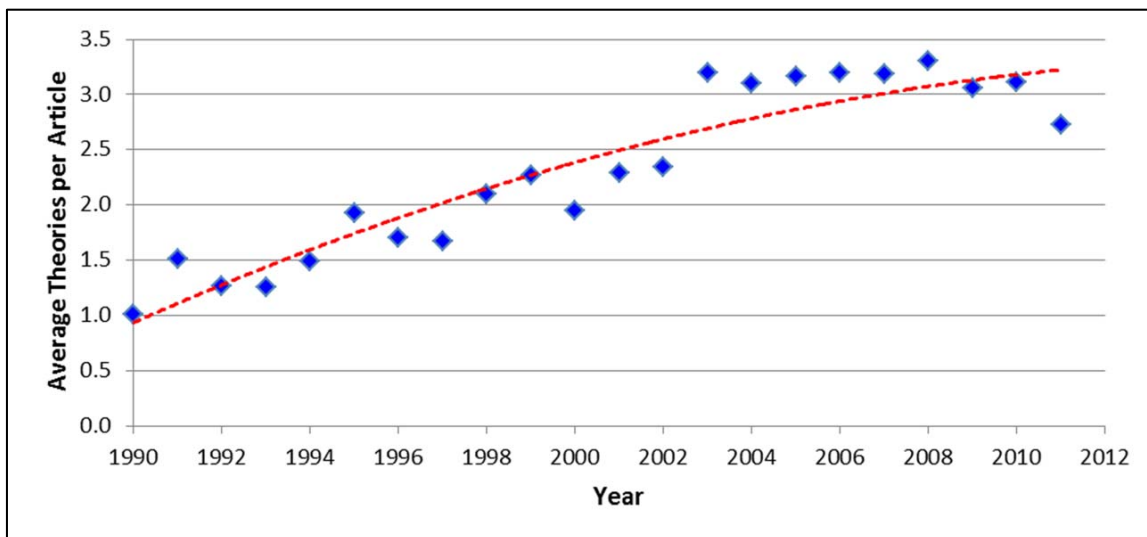

Figure 3. Average number of theories per IS research article over time.

As shown in Table 3 above, IS research articles published between 1990 and 2011 relied on an average of approximately 2.4 unique theories per article. One-way analyses of variance revealed that articles published in MISQ relied on more theories per article than did JMIS ($p < 0.001$) or ISR ($p < 0.001$), while no statistically significant difference was present between the number of theories used per article in ISR and JMIS ($p = 0.295$). Figure 3 reveals a strong upward trend in the average number of theories appearing in IS research articles over time. Notably, the average number of theories per article in top IS journals has tripled since 1990. Shifting editorial expectations notwithstanding, one possible explanation for this remarkable growth is that when the field was in a more nascent stage of its development, phenomena of interest to IS researchers were studied in a comparatively isolated or insulated manner. As the field has evolved, the knowledge it has produced and the expansion of its scope of inquiry have led researchers to consider increasingly complex and nuanced phenomena, thus necessitating the adoption of a larger and more diverse theoretical ecosystem. Note also that Figure 3 includes a second-order polynomial fit line ($R^2 = 0.875$) which seems to indicate that the rate of growth in this trend is decreasing, and the average number of theories per IS research article may be beginning to stabilize.

The rapid growth in the number of theories per article motivates another interesting question; namely, which theories commonly appear together within IS research articles? To address this question we conducted an affinity analysis on the article-level theory frequency data in order to identify the

commonest theory dyads and triads. The results of these analyses are respectively presented in Tables 4 and 5 below, sorted according to their levels of statistical support (i.e., the percentage of all IS research articles published between 1990 and 2011 in which the theory dyads or triads were present).

Rank	Theory Dyad	Support
1	Technology acceptance model, Theory of planned behavior	5.78%
2	Technology acceptance model, Theory of reasoned action	5.09%
3	Agency theory, Transaction cost economics	4.88%
4	Theory of planned behavior, Theory of reasoned action	4.61%
5	Agency theory, Critical realism theory	4.45%

Table 4. Commonest theory dyads within IS research articles from 1990 through 2011.

Rank	Theory Triad	Support
1	Technology acceptance model, Theory of planned behavior, Theory of reasoned action	4.03%
2	Diffusion of innovations theory, Technology acceptance model, Theory of planned behavior	2.07%
3	Diffusion of innovations theory, Technology acceptance model, Theory of reasoned action	2.01%
4	Diffusion of innovations theory, Theory of planned behavior, Theory of reasoned action	1.86%
5	Social cognitive theory, Technology acceptance model, Theory of planned behavior	1.70%

Table 5. Commonest theory triads within IS research articles from 1990 through 2011.

The information shown in the tables above is very revealing with respect to the combinations of theories that have been necessary in order to allow researchers to adequately describe and address complex IS-related phenomena during the past few decades. In considering this information, one quickly notes that it is very common for TAM, TRA, TPB, and diffusion of innovations theory to appear in some combination within the same article. This may suggest that each of these theories is but a facet of a larger and more general theory (such as the Unified Theory of Acceptance and Use of Technology, for example) which would subsume all of the predictive and explanatory power of its sub-theories. When used in conjunction with construct associations among theories – such as those developed by Hovorka et al. [23] – information such as that reported in Tables 4 and 5 above may be the cornerstone of efforts aimed at theory integration and multilevel theory development within the IS field.

Beyond considerations of the overall impacts of different theories and theory clusters to the IS field, our third research question focused on the degree of theoretical uniformity among the IS field's leading journals. To that end, the 10 most frequently mentioned theories in each journal from 1990

through 2011 are shown in Table 6 on the following page.

Of the 17 unique theories appearing in the table, six of the theories (35.3%) appear in the top 10 lists of all three journals (i.e., TAM, SERVQUAL, Agency Theory, Transaction Cost Economics, Media Richness Theory, and the Theory of Planned Behavior), while seven of the theories (41.2%) appear in the top 10 lists of two or more journals. The remaining ten theories in the table appear in the top 10 list of only one journal. With respect to pairwise comparisons between journals, there is a 70% overlap among the top 10 theories appearing in MISQ and ISR, a 60% overlap among the top 10 theories appearing in MISQ and JMIS, and a similar 60% overlap among the top 10 theories appearing in ISR and JMIS.

When considered together, these results imply that although there exists a substantial degree of theoretical uniformity among the IS field's top journals, each journal also exhibits a degree of uniqueness in its personality. Like children from the same family, the leading journals in the IS field share a great deal of theoretical DNA, but there are nevertheless differences in the genetic code of each journal which makes it unique among its siblings.

Rank	MISQ	ISR	JMIS
1	SERVQUAL	Technology acceptance model	Agency theory
2	Technology acceptance model	Theory of planned behavior	Technology acceptance model
3	Agency theory	Agency theory	Transaction cost economics
4	Theory of planned behavior	Media richness theory	SERVQUAL
5	Transaction cost economics	SERVQUAL	Real options theory
6	Media richness theory	Transaction cost economics	Critical success factors, theory of
7	Cognitive dissonance theory	Theory of reasoned action	Contingency theory
8	Theory of reasoned action	Dynamic capabilities	Media richness theory
9	Hermeneutics	Delone and McLean IS success model	Adaptive structuration theory
10	Structuration theory	Cognitive fit theory	Theory of planned behavior

Table 6. The relative frequency of different theories by journal from 1990 through 2011.

5. Summary, Limitations, & Conclusions

The primary objectives of the research described herein were to identify and gain systematic and scientific insights into the theoretical core of the IS field, as reflected by the theories and theory clusters that appear most commonly in our top journals. We further sought to quantify the distribution of research articles according to their levels of analysis within the IS field's top journals, as well as the degree of theoretical uniformity among those journals. To achieve these objectives, we applied a computational text mining technique known as n-gram analysis to every research article published in Management Information Systems Quarterly, Information Systems Research, and the Journal of Management Information Systems between 1990 and 2011. Using an established set of 87 different theories as input, our analyses revealed that a small number of theories have exerted a great deal of influence on IS research over time, and we believe that these theories arguably comprise the theoretical core of the IS field. The disciplinary influence exerted by other theories used in IS research is by comparison relatively marginal.

Our analyses further revealed that the distribution of research articles in the IS field according to the levels of analysis of their theoretical constructs has remained relatively stable over time, while the average number of unique theories appearing in each research article has grown dramatically since 1990. Additionally, a comparatively small number of unique theories were observed to dominate the commonest combinations of theories appearing in IS research articles, potentially suggesting the presence of underlying, more general theories. A great deal of overlap was also observed among the theories which dominate the field's leading journals, suggesting that these journals share a great deal of theoretical DNA. Nevertheless, each journal was seen to exhibit at least a modicum of individuality when characterized by the theories which appear most frequently within its pages.

Further major contributions of this work include the information presented in Tables 1 and 6, which respectively provide a ranked list of the IS field's most commonly used theories, and ranked lists of the theories which have been most dominant in MISQ, ISR, and JMIS over the 22-year span of our analyses. By these means, this paper sheds light on the theoretical core of the field – which has been an elusive concept – and can serve as a basis for more precise argumentation regarding the theories which comprise the core of the IS field.

As with all research which aims to shed light on questions of theory and identity, our work has several limitations which should be acknowledged. To begin, our measurement model is predicated on the notion that the relative frequency with which a theory is mentioned in a field's research literature reflects the level of importance or influence of that theory to the field at the point in time when the research was written. Although this notion is supported by theoretical and applied work in the area of computational linguistics [14-16], there may be other scientometric methods of measuring the impact of a theory on a scientific field.

Next, we relied upon a predefined set of theories as input for our analyses. This strategy, while both convenient and methodologically justifiable, is imperfect in at least two ways. First, the possibility exists that one or more theories appearing in MISQ, ISR, or JMIS did not appear in the set of theories used in our study, thus making the results incomplete. Second, one or more theories used as input for our analyses might never have appeared in MISQ, ISR,

or JMIS. Indeed, our analyses revealed that of the 87 theories described on the *Theories Used in IS Research* website [17], 11 did not appear even a single time in MISQ, ISR, or JMIS over the 22-year span of our investigation. If we accept the premise that the three journals used in our study represent mainstream IS research, then the 11 missing theories might reasonably be considered peripheral to the mainstream IS field. Finally, our work is limited insofar as articles from only three IS journals were included in our analyses. While past research has concluded that MISQ, ISR, and JMIS are among the highest-quality scholarly journals in the IS field [9, 10], we acknowledge that these three journals are an imperfect proxy for the field as a whole.

Although the results reported here provide many new and interesting insights into the theoretical core of the IS field, as well as into the use of different theories over time and in different journals, much remains to be learned. Our future research will therefore seek to identify theories whose influence on the field appears to be growing or declining, and will further endeavor to quantify the extent to which other fields have influenced the identity of the IS field.

In the interim, we would like to leave the reader with a final philosophical thought. To wit, perhaps the unstable and fragmented theoretical landscape within the IS field is in and of itself a core component of the field's identity. If other fields are indeed defined by a comparatively small set of macro-theoretic paradigms, then the fact that the IS field is not similarly defined may be a key element of what makes it unique.

6. References

[1] Hirschheim, R., and Klein, H.K., "A Glorious and Not-So-Short History of the Information Systems Field", Journal of the Association for Information Systems, 13(4), 2012, pp. 188-235.

[2] Benbasat, I., and Zmud, R.W., "The identity crisis within the IS discipline: Defining and communicating the discipline's core properties", MIS Quarterly, 27(2), 2003, pp. 183-194.

[3] Hassan, N.R., "Is information systems a discipline? Foucauldian and Toulminian insights", European Journal of Information Systems, 20(4), 2011, pp. 456-476.

[4] Agarwal, R., and Lucas, H.C., "The Information Systems Identity Crisis: Focusing on High-Visibility and High-Impact Research", MIS Quarterly, 29(3), 2005, pp. 381-398.

[5] Neufeld, D., Fang, Y., and Huff, S., "The IS identity crisis", Communications of the Association for Information Systems, 19, 2007, pp. 447-464.

[6] Somers, M.J., "Using the theory of the professions to understand the IS identity crisis", European Journal of Information Systems, 19(4), 2010, pp. 382-388.

[7] Moody, D., Iacob, M.-E., and Amrit, C., "In Search of Paradigms: Identifying the Theoretical Foundations of the Information Systems Field", 18th European Conference on Information Systems (ECIS), 2010

[8] Larsen, K.R., Monarchi, D.E., Hovorka, D.S., and Bailey, C.N., "Analyzing unstructured text data: Using latent categorization to identify intellectual communities in information systems", Decision Support Systems, 45(4), 2008, pp. 884-896.

[9] Ferratt, T.W., Gorman, M.F., Kanet, J.J., and Salisbury, W.D., "IS Journal Quality Assessment Using the Author Affiliation Index", Communications of the Association for Information Systems, 19, 2007, pp. 710-724.

[10] Rainer, K., and Miller, M., "Examining differences across journal rankings", Communications of the ACM, 48(2), 2005, pp. 91-94.

[11] Manning, C.D., and Schütze, H., Foundations of statistical natural language processing, MIT Press, Cambridge, MA, 1999.

[12] Barante, A.G.P.B.D., Histoire de la Convention Nationale, Langlois et Leclercq, Paris, 1851.

[13] Michel, J.-B., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M.K., Team, T.G.B., Pickett, J.P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M.A., and Aiden, E.L., "Quantitative Analysis of Culture Using Millions of Digitized Books", Science, 331(6014), 2011, pp. 176.

[14] Bohannon, J., "Google Books, Wikipedia, and the Future of Culturomics", Science, 331(6014), 2011, pp. 135.

[15] Soper, D.S., and Turel, O., "An n-Gram Analysis of Communications: 2000-2010", Communications of the ACM, 55(5), 2012, pp. 81-87.

[16] Soper, D.S., and Turel, O., "Who Are We? Mining Institutional Identities Using n-grams", 45th Hawaii International Conference on System Sciences (HICSS), 2012

[17] Schneberger, S., and Wade, M., Theories Used in IS Research. http://istheory.byu.edu, accessed 01 May 2013.

[18] Davis, F.D., "Perceived usefulness, perceived ease of use, and user acceptance of information technology", MIS Quarterly, 13(3), 1989, pp. 319-339.

[19] Parasuraman, A., Berry, L.L., and Zeithaml, V.A., "A Conceptual Model of Service Quality and Its Implications for Future Research", Journal of Marketing, 49(4), 1985, pp. 41-50.

[20] Alchian, A., and Demsetz, H., "Production, information costs, and economic organization", American Economic Review, 62(5), 1972, pp. 777-795.

[21] Eisenhardt, K.M., "Agency theory: An assessment and review", Academy of Management Review, 14(1), 1989, pp. 57-74.

[22] Harrison, D.A., and Klein, K.J., "What's the difference? Diversity constructs as separation, variety, or disparity in organizations", Academy of Management Review, 32(4), 2007, pp. 1199-1228.

[23] Hovorka, D.S., Larsen, K.R., Birt, J., and Finnie, G., "A Meta-theoretic Approach to Theory Integration in Information Systems", 46th Hawaii International Conference on System Sciences (HICSS), 2013.