Diffusion of Business Intelligence and Data Warehousing: An Exploratory Investigation of Research and Practice

Michael L. Gonzales  
University of Texas at El Paso  
mlgonzales2@utep.edu  
Godwin Udo  
University of Texas at El Paso  
gudo@utep.edu  
Kallol Bagchi  
University of Texas at El Paso  
kbagchi@utep.edu  
Peeter Kirs  
University of Texas at El Paso  
pkirs@utep.edu

Abstract
This empirical study extends current Business Intelligence (BI) and Data Warehousing (DW) research by studying the discourse life cycle of IS fashion waves. Using bibliographic methodology and applying Diffusion of Innovation and Management Fashion theories, BI/DW related papers and articles were gathered from both, academic research and practitioner journals, published from 1995 to 2009. Formal diffusion models were employed to examine the level of adoption of BI/DW based on these papers and articles. Our findings demonstrate mixed-influence fashion waves of BI/DW across the academic and private sector communities, with the practitioner’s research being influenced more by external factors compared to academic and academic research diffusion progressing differently than practitioner’s literature diffusion and the latter has comparatively slowed down.

1. Introduction
Justifying and demonstrating the effectiveness of information systems (IS) is increasingly critical as firms look to become more competitive and efficient [30]. The challenge for both Business Intelligence (BI) and Data Warehousing (DW) has been in explaining their mixed success. While many organizations have experienced value from their BI initiatives, many others have fallen short. Some organizations have gained solid benefits, while others have fallen short [34]. Also, recent ComputerWorld and other studies cite BI as a top skill for IS professionals in 2010 [5][25].

The interest in BI/DW, and its corresponding usefulness as a managerial tool, is representative of concerns with all new management approaches: Is it a passing fad, or an enduring fashion? The question is of particular concern to IS researchers and practitioners since IS/IT is driven by technological opportunities [24]. In the 1980’s, there was considerable argument about the value of graphs vs. tables, although in general little difference was found between the two [10], and subsequent interest disappeared. Other technologies, such as CASE and e-commerce, have evolved into widely accepted and valued tools [17]. Bibliographic research has been used to identify IS “fashion waves” based on neo-institutional theory and suggested that a management fashion is a belief that a certain management technique leads to rational progress [2]. Findings recommended that IS research should participate more directly through the use of action and practice research, demonstrate flexibility in dropping research topics that have lost interest among practitioners, provide more practitioner-oriented publication outlets, and create new publication outlets with shorter review cycles [2].

Other research [17] raised the question Are issues fashions or diffused innovations? While generally agreeing with the findings stated above, researchers differed in the manner in which fashion waves should be analyzed. They generally suggest that research should be viewed from the perspective of diffusion, which considers the changing rate of acceptance. Consider the following (entirely hypothetical) example, where one topic (A) represents a fad, and another topic (B) represents a fashion wave (see Table 1 for raw data, and Figure 1 for the cumulative graph). Topic A clearly started strongly, but over time the total (cumulative) number (i.e., diffusion) of articles diminished, while topic B has endured and even accelerated. For the last five time periods, the slope of
the linear regression for topic A is 1.0, while the corresponding slope for topic B is 11.00.

There have been numerous BI/DW articles published from different perspectives, including implementation and available technology [36]. One of the most important and contemporary BI/DW literature reviews [22] published in ten leading journals from 1997 to 2006 in an attempt to establish BI-centric research categories.

<table>
<thead>
<tr>
<th>Time period</th>
<th>Topic A No</th>
<th>Topic A Total</th>
<th>Topic B No</th>
<th>Topic B Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>8</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>16</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>21</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>25</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>29</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>31</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>32</td>
<td>11</td>
<td>45</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>33</td>
<td>11</td>
<td>56</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>33</td>
<td>12</td>
<td>66</td>
</tr>
</tbody>
</table>

Table 1: Illustrative Data

However, there are no current research articles that validate or reject the notion that BI/DW are broadly adopted technologies/techniques. Some researchers [35] claim that DW was a fad “…advanced by many who were impressed by claims but who never tried it…. In the private sector, BI diffusion is generally viewed as an area of growth [36].

We consider the diffusion of BI/DW research (although the approach could be applied to any topic) based on accepted diffusion models using bibliographic research in both published academic journal articles and practitioner papers/articles [17]. There are also several subordinate questions which we address: Are the academic and private sector communities synchronized? Which of the well-referenced diffusion innovation models best predicts the diffusion of BI/DW? What are the differences between the scholarly research and private sector in terms of internal versus external influence? If it is found that influences driving BI/DW diffusion have similar pattern in scholarly research as well as in private sector then it can be argued that BI/DW research and practice have similar information exchange patterns and/or drawbacks. If not, then one can argue that influences driving IS research and IS practice are different which may eventually lead to increasing isolation among the two communities and bridges need to be built for better exchanges and communication. Additionally, if there is evidence of enduring BI/DW diffusion, we can further consider potential opportunities for future research [36].

While previous studies have investigated diffusion of various technologies/concepts based on academic and practitioner’s outlets, none have also considered how source of influences (internal, external, and mixed) make these diffusions similar or different. In this paper, we also compare these influences and discuss how they differ.

2. DW and BI

The early focus of relational databases was on online transactional processing (OLTP), designed to support the daily business operation requirements [6]. Using Codd’s normalization rules, operational systems are typically optimized for fast inserting and updating of data. As companies expanded the applications they supported and consequently, the data stored, attention turned to providing user access for reporting and decision support. However, a different architecture was needed where transactional data from operational systems could be extracted, transformed, integrated, and stored to support the reporting requirements [9]. The DW utility was defined by four necessary components: Subject-oriented, Integrated, Non-volatile, and Time-variant [21]. The DW is designed to collect disparate data from operational systems and uniquely store that data to allow end-users ready access for subsequent analysis.

A DW is intended to provide integrated data from operational systems organized for reporting, whereas BI is intended to provide actionable information. BI can be defined as “…a set of concepts and methodologies to improve decision-making in business through use of facts and fact-based systems.” Simply stated, BI provides the ability to transform data into usable and actionable information for business and organization purposes [18]. BI is an encompassing term that combines data architectures, technical architectures, analytic tools, and methodologies [32].

There is significant synergy and overlap between DW and BI. While a DW is the core repository [32], BI
requires an information infrastructure to provide actionable insight to decision-makers [4]. The synergy between BI and DW has caused them to be viewed as one entity. An industry report [16] notes that “BI/DW is a strategic initiative that has the potential to deliver significant insights unavailable through other means”. The close relationship between BI and DW drives this research to examine both as a single body of published work, reflecting a continuum of applications, technologies, and techniques to support better decision-making.

3. Diffusion

Much attention has recently been given to Information Communication Technologies (ICT) adoption and diffusion, defined as the degree of market penetration. The theory of innovation diffusion considers how a new idea spreads throughout the market over time. The ability to predict accurately new product diffusion is of concern to designers, marketers, managers, and researchers alike.

Innovation diffusion is defined as the process by which the innovation is communicated over external media channels or internal social systems [27]. There are four fundamental components to the traditional diffusion process, including: 1) An innovation such as an idea, product or process, 2) Communication channels such as radio, TV, and newspaper, 3) Time as it pertains to the rate of adoption of the innovation, and 4) A social system that consists of individuals or organizations that represent potential adopters [26].

There are several diffusion models with significant support, including the external, internal, and mixed influence models [12][3]. The external model, assumes that external sources of influence drive the adoption of innovation for an organization. ATMs are an example of external influence where the manufacturers of the equipment aggressively promoted the adoption [12]. The general form of the model is:

$$\frac{dN(t)}{dt} = a(N - N(t))$$

Where: $a$ = External Influence, $t$ = time, $N$ = Potential Total Adopters, and $N(t)$ = population of adopters

The internal model assumes that influence comes from communication or social networks, such as the adoption of seed as a consequence of promotion by farmer associations:

$$\frac{dN(t)}{dt} = bN(t)(N - N(t))$$

Where: $b$ = Internal Influence

The Bass mixed influence model assumes that influence comes from both, internal and external sources. The adoption process starts with mass media influencing early adopters. Individuals are influenced by internal communication early in the diffusion process but the impact declines in later periods. When plotted on a cumulative basis, this adoption rate creates a familiar S-shaped curve [27].

$$\frac{dN(t)}{dt} = (a + bN(t))(N - N(t))$$

This study uses all three models to not only validate or reject hypotheses offered, but also to determine which model best predicts the diffusion of BI/DW.

4. Hypotheses

There is evidence of continued growth in BI/DW and its impact on organizations. The IDC studied 43 leading organizations in North America and Western Europe and found a median Return on Investment (ROI) of 112% for business analytics projects using BI/DW [20]. This relatively high ROI and continued IT spending in BI/DW indicates “The extent of BI pervasiveness is a statistically significant predictor of organizational competitiveness and performance” [33]. A 2008 study [15] concluded that BI was the No. 1 technology priority for three years in a row. Consequently, we believe that data warehousing and business intelligence are experiencing a fashion wave [2].

We propose the following two research questions and related hypotheses.

Research Question 1: How has BI/DW diffused over time? This is an important question to answer since it has significant ramifications for organizations considering the implementation of BI/DW. If BI/DW exhibits an elongated S-curve (fashion wave), one can infer that adopters have benefited from their investment. Conversely, limited diffusion may discourage organizations to further pursue BI/DW.

If BI/DW demonstrates enduring diffusion, the question “Which diffusion model most accurately estimates BI/DW diffusion?” needs to be answered. Previous research has generally favored the Bass mixed-influence model [19], and thus our first hypothesis is:

$H1$ – A Mixed-Influence model will more accurately predict the diffusion of BI/DW in both academic and practitioner’s literature.

Research Question 2: Is there any difference in diffusion between BI/DW academic and practitioner’s literature? Traditionally, there are several ways a
product or an idea is diffused in academic and practitioner's literature:

1. Academic literature generally precedes practitioner literature [7]. There are variants in this phenomenon:
   a) a lag in otherwise similar diffusion curves
   b) academic literature continues, practitioner literature does not take off or slows down earlier resulting in different diffusion curves (for example TAM research)
   c) academic interest shifts, practitioner interest continues—similar as in (b), resulting in different diffusion curves

2. Academic and practitioner literature diffusion corresponds — similar diffusion curves in time, with little or no lag [29].

3. Practitioner literature diffuses earlier than academic research—a lag in diffusion as in 1.a [28], an example could be SOA area.

Although any of these paths could occur, it seems likely that academic research first occurred with relational database design, followed by interest in DW and then BI. Practitioners’ articles also appeared on how to implement these systems, how vendor products are developing, and success/failure stories around the concept. Soon, successful implementations followed in DW and it became a routine affair in industry circles. BI implementation stories proved to be more difficult and esoteric in nature for practitioners' magazines while the academics focused research on this topic with additional mathematical developments. Thus, practitioners’ interest in BI/DW has slowed down whereas academic interest has continued. In addition, academic research in general is significantly influenced by peer review which may not only limit the amount of published work, but delays the ideas or innovations proposed [2].

$H2$ – BI/DW diffusion in academic literature will be slower than in practitioner literature.

A concept in academic environment primarily diffuses through social networks. Academics often discuss ideas with their peers, present their ideas at conferences, receive peer-criticism and review, and then finally prepare the paper for an academic journal. In contrast, practicing managers often receive new ideas, as well as implementation success/failure stories and relevant product information, from vendor ads and articles in magazines, newspapers, and other external outlets (although peer networks also play an important role). Therefore, we propose the following hypothesis:

$H3$ – In the diffusion of BI/DW the role of external influence will be greater for practitioners' literature.

5. Methodology and data

Five top scholarly journals and two practitioner periodicals were searched for papers that reference at least one of the terms, Business Intelligence or Data Warehousing, in the body of the paper. The search covered a 15 year period (1995 – 2009).

5.1. Data

5.1.1. Phase one: select the years of study and periodicals. Focusing on a time span of 15 years, from 1995 to 2009, affords a comprehensive examination of the diffusion of BI/DW. The period also encompasses the most prevalent activity for BI/DW. It was during this period that Bill Inmon, often considered the "father of data warehousing" published his influential book [21], and Ralph Kimball founded Red Brick Systems. There were two overriding constraints used to select scholarly and practice journals for this preliminary study, including: 1) Only journals that were actively publishing the entire 15 year period were candidates, and 2) Only journals that represented a cross-section of IT-centric content and not just BI/DW topics were considered. Given these two constraints, some popular journals were not selected, for example, TDWI BI Journal. This practitioner journal was not publishing the entire 15 year time span and it represents a BI/DW specific content and tone.

We selected five academic journals based on two criteria: first, at least one journal must represent general interest content and the other four specifically MIS, and second, one journal had to be considered a "B" level journal while the other four are rated as "A/A+" journals. The purpose was to get a better cross-section of papers being published as opposed to simply selecting the top five MIS journals. Using the current rankings of journals by Association for Information Systems (AIS) coupled with publication grading from Penn State University and Washington State University (WSU), we selected the following journals: MIS Quarterly, Information Systems Research, Journal of Management Information Systems, Management Science, and Information & Management.

Practitioner periodicals were limited in large part due availability in LexisNexis. We selected two of the most popularly read: Computer World and InfoWorld. While this represents fewer periodicals than used for the academic community, the number of articles...
selected is considerably more (See Table 2 for a summary).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MIS Quarterly</td>
<td>34</td>
<td>Computer World</td>
<td>1329</td>
</tr>
<tr>
<td>Info. &amp; Mgt.</td>
<td>15</td>
<td>InfoWorld</td>
<td>886</td>
</tr>
<tr>
<td>Journal of MIS</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mgt. Science</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISR</td>
<td>59</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>158</strong></td>
<td></td>
<td><strong>2215</strong></td>
</tr>
</tbody>
</table>

Table 2: Periodicals Selected

The approach and process for selecting papers used in this study was based on previous research conducted [1][2][22]. For example, selected the ten top IT journals and used the terms BI, DW, and Data Mining to identify potential papers spanning 1997 to 2006 [22]. While this work identified articles for the purpose of categorizing the type of research being conducted, our intent was to simply identify and count relevant papers as data points for the three diffusion models and descriptive statistics. To that end, we designed the three-phased process outlined below:

5.1.2. Phase two: conduct the initial search for candidate papers. This phase was not a single search, but required 75 specific searches (15 years times 5 journals). For each journal, 15 searches were conducted, one for each year from 1995 to 2009. The three key phrases searched for in each article were: “business intelligence”, “data warehouse”, “data warehousing”. We purposefully did not include phrases such as decision support, data mining, expert systems, or artificial intelligence in our searches in order to mitigate concerns or doubts of the papers included in this study. The specificity of this search process ensured consistency and repeatable results.

5.1.3. Phase three: identify and select papers to include in study. Although the initial search results were specific, there were several other filtering criteria to ensure we selected only those papers that support the research. In order to be selected, the three key phrases must be used in the context of IS support systems and applications. If the article mentions, for example, business intelligence in a literal sense as opposed to reference of an IS based system or application, then the paper was rejected. Also, if the phrases were found only in the reference section, subject terms/key words, or author’s bio, the paper was rejected. Only those papers that used one or more of the three phrases within the body of the paper, its tables or notes, and conformed to an IS reference, were selected.

Final selection of papers for this study was conducted on a paper-by-paper basis. Since the vast majority of papers were stored in a PDF format, the Find option of Adobe Acrobat was used in order to determine if the candidate paper met all the criteria specified above. A test word was chosen at random from the paper in order to test that the Find option worked on the PDF version of the document. If the test search worked, then three separate searches were conducted, one for each of the key phrases.

5.2. The study design

The papers identified, selected, and categorized for this study serve as the data set for the three diffusion models: Internal, External, and Mixed-Influence. The expressions for N(t), the cumulative number of adopters at time t, used for each of the models are outlined below [26][12]:

- **Internal Model**: \( m/(1+ ((m-m0)/m0)*exp(-b*m*t)) \)
- **External Model**: \( (m*(1-exp(-(a)*t))) \)
- **Bass Mixed-Influence Model**: \( (m*(1-exp(-(a+b)*t)))/(1+(b/a)*(exp(-(a+b)*t))) \)

where \( m \) is the number of potential adopters, \( m0 \) is the number of adopters at time \( t=0 \) and other symbols have the same meaning as in section 3.

The Nonlinear Regression model of SPSS was used for each of the three diffusion models of this study.

6. Analysis and results

The results from each of the diffusion models are described below. Accompanying each model is a graph that plots the diffusion model estimates for each quarter and another graph that plots the actual cumulative papers over 15 years. Plotting the cumulative data on articles should result in an S-curve and this is the case in the present paper. Its shape is dictated by the values of internal, external or both [26].

6.1. Internal model (IM): academic research

The IM provides evidence that sustained BI/DW diffusion is occurring. The results suggest that such internal as communication within the organization or social networks do impact diffusion. The sum of squares of errors (SSE) is relatively high and \( R^2 \) value low when compared to the mixed model, thus providing support for H1. Refer to Table 3 for summary statistics. These results suggest that mixed models are more accurate at diffusion estimation than the Internal model.
6.2. External model (EM): academic research

The EM failed to present evidence of diffusion (Figure 3). There is no convergence between the model estimates and the actual cumulative papers over the 15 years of data. This suggests that external influences by themselves, for example, vendors or consulting companies, do not sufficiently promote the diffusion of BI/DW. An SSE of 22383.309 implies that the external model is unable to capture the diffusion of BI/DW. The results do, however, add support to H1 by demonstrating that an EM is less accurate, if not incapable, of measuring diffusion when compared to Mixed-Influence model.

6.3. Bass mixed-influence model (MM): academic research

The results from the Bass MM are given in Figure 4. The findings suggest that both internal and external influences are affecting the diffusion of BI/DW in academic research. The fit of the MM to the cumulative academic articles published is significant with an $R^2$ of 0.998 (and an SSE of 398.061). The results provide additional support for H1 in the case of academic literature.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MM Estimate</th>
<th>IM Estimate</th>
<th>EM Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>198.45***</td>
<td>187.286***</td>
<td>103221.85**</td>
</tr>
<tr>
<td>$a$</td>
<td>0.00186***</td>
<td>---</td>
<td>0.000022**</td>
</tr>
<tr>
<td>$b$</td>
<td>0.08709***</td>
<td>0.001***</td>
<td>---</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.998</td>
<td>0.997</td>
<td>0.868</td>
</tr>
</tbody>
</table>

Table 3. Scholarly Data Set Model Parameter Estimates

$m = \%$ Saturation level = total number of adopters

$a = \text{Coefficient of External influence (e.g., media)}$

$b = \text{Coefficient of Internal influence/imitation (e.g., word of mouth)}$

***: $p < 0.000$, **: $p < 0.05$

6.4. Bass mixed-influence model: practitioner articles

We show only the figure (Figure 5) for the Bass model for practitioner articles. The summary parameter estimates are given in Table 4. Based on the $R^2$ values the Bass model provides the best fit for diffusion of practitioner’s literature. Thus support for H1 can also be found in the case of Practitioner’s literature. This also provides an answer to Research question 1.
Figure 5. Practitioner Data Set Bass Mixed-Influence Model Distribution versus Actual Cumulative

Table 4. Practitioner Data for All Model Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MM Estimate</th>
<th>IM Estimate</th>
<th>EM Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>2610.6***</td>
<td>2193.08***</td>
<td>3444.44***</td>
</tr>
<tr>
<td>a</td>
<td>0.0207***</td>
<td>---</td>
<td>0.0182***</td>
</tr>
<tr>
<td>b</td>
<td>0.0228***</td>
<td>0.000444***</td>
<td>---</td>
</tr>
<tr>
<td>R²</td>
<td>0.9939</td>
<td>0.9765</td>
<td>0.9927</td>
</tr>
</tbody>
</table>

7. Comparison of academic and practitioner diffusions

The inflection point (the maximum penetration rate), where incremental change in cumulative value is maximum, of the MM for academic diffusion occurred at about the 43rd quarter, preceded by a rapid rise in adoption and then reaching a level of sustained maturity. For practitioner diffusion, the MM results in an inflection point much earlier, at about the 3rd quarter. Clearly, the two mixed models differ in structure. This provides support for H2.

Using the results from the MM for both the scholarly and practitioner data sets, we also find support for H3 as well as research question 2. The internal parameter estimate for the practitioner data set is 0.0207, whereas the internal parameter estimate for the scholarly data is 0.08909. The external parameter values for scholarly and practice data sets are .00186 and .0207 respectively. When we consider the ratio of coefficients b and a (i.e., b/a), for practice data set it is 1.101 whereas for the academic data set it is 46.823, a marked difference indeed. Thus external influence has a better impression on practitioners' literature, although both models are mixed in nature.

8. Predictive quality

A related issue is the predictive quality of the Bass Mixed-Influence model. Assessing the predictive validity of a model requires the examination of three aspects: 1) Use of ex ante (out-of-sample) data from data set, 2) Comparison to other simple and established forecast methods, and 3) An adequate sample [8]. We selected the last 10 data points (out of the 60 data points in our data set) as an ex-ante sample from the academic data set.

<table>
<thead>
<tr>
<th>T</th>
<th>Actual</th>
<th>MM</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>136</td>
<td>129.04</td>
<td>119.75</td>
</tr>
<tr>
<td>52</td>
<td>141</td>
<td>132.85</td>
<td>126.00</td>
</tr>
<tr>
<td>53</td>
<td>141</td>
<td>136.54</td>
<td>131.75</td>
</tr>
<tr>
<td>54</td>
<td>145</td>
<td>140.09</td>
<td>136.25</td>
</tr>
<tr>
<td>55</td>
<td>147</td>
<td>143.50</td>
<td>140.75</td>
</tr>
<tr>
<td>56</td>
<td>149</td>
<td>146.77</td>
<td>143.50</td>
</tr>
<tr>
<td>57</td>
<td>151</td>
<td>149.90</td>
<td>145.50</td>
</tr>
<tr>
<td>58</td>
<td>154</td>
<td>152.87</td>
<td>148.00</td>
</tr>
<tr>
<td>59</td>
<td>158</td>
<td>155.70</td>
<td>150.25</td>
</tr>
<tr>
<td>60</td>
<td>158</td>
<td>158.37</td>
<td>153.00</td>
</tr>
</tbody>
</table>

Table 5. Actual vs. Forecasts for Academic Diffusion

The Bass model was again run for a 50-data point series. The parameter values obtained were used to forecast the next 10 periods. As a comparison, we also forecast the ex-ante data using a 4-period simple moving average. Table 5 shows actual cumulative totals as well as the forecasts. The SSE values for the out-of-sample 10 forecasts were 184 (for the MM) and 871 (for the 4-period moving average). Clearly, the MM outperformed the 4-period moving average.

We also repeated the process for the practitioner data set. The Bass MM for 50 samples converged with significant parameter values (m=2360.931, a=.02165, b=.0338). However, the SSE for out-of-sample data from the Bass model was 61263.68 whereas the SSE of a corresponding 4-period simple moving average was 21579.

9. Study limitations

There are three key limitations to this study. First, the publications selected consciously not BI/DW specific. The rationale was twofold: 1) these types of BI/DW journals do not represent a broader cross section of IT papers that we focused on in this paper, and 2) the journals often did not cover the 15 year timespan we studied in this paper. A second limitation was that we
specifically selected articles that materially used the exact terms, “Business Intelligence”, “Data Warehouse”, and “Data Warehousing”. Much of the literature that might be appropriate for BI could include terms such as data mining, statistical analysis, dashboards, and OLAP just to name a few. Thirdly, the publications examined were not considered international journals. Consequently, the results might be overly biased toward the U.S. domestic market.

10. Conclusion and future research

In spite of the limitations cited above, this study contributes in several ways. The analysis demonstrates that BI/DW has achieved a relatively significant level of diffusion over the 15 years from 1995 through 2009. Of the types of diffusion innovation models (Mixed, Internal, and External), the mixed model more accurately estimates the diffusion of BI/DW. The Bass Mixed-Influence model suggests that both, internal and external influences are promoting the adoption of BI/DW technology and techniques for both academic and practitioner’s literature.

Another important finding is that the practitioners’ literature growth seems to be more influenced by external factors than are academic literature growth. Finally, it can be observed that practitioners’ periodical diffusion rate slows earlier than that of academic journals. This may be partly explained by the peer review process of published scholarly work.

The impact of these preliminary findings is particularly significant if they remain consistent as we expand the study. It suggests that academic research in BI/DW is conducted in a closed community of peers. As such, it may make the research results less relevant or disconnected with practitioners. Conversely, the private sector may be overly influenced by external factors such as advertising, vendor influence, and outside consultants. When companies plan to make investments in BI/DW, they must guard against decisions being swayed by vendors and other external sources that may marginalize the internal value of the effort. If this is allowed to happen, these two fields may become disjoint, causing enough future concerns for DW/BI growth in general.

This preliminary research provides a foundation for further study, specifically in four areas: multi-innovation, multi-adoptions, global, and applicability to different technologies.

First, future research can be more specific about the relationship between BI and DW. In our study, the two areas were examined as a single body because it is arguable that significant synergy exists between the two. But doing so masks the underlying, and potentially significant influences between BI and DW. Studying Multi-Innovation Diffusion models may expose important insights. For instance, since BI is a more recent innovation, does it serve as a substitute for DW, is it complementary to DW, or contingent on DW [26]. Understanding the relationship between BI and DW diffusion potentially provides insight for organizations to plan investment strategies among other benefits.

Multi-Adoption Diffusion models examine the possibility of adopters’ repurchases of technical innovations. For the BI/DW community, successive investments in technology to support the BI/DW initiatives as well as the iterative nature of implementation, presents an excellent opportunity for research. From a product innovation perspective, how do first-time buyers and repeat buyers impact the adoption rate [26][11]?

The global research is an important extension to this study. A fundamental research question to ask is whether the diffusion of BI/DW in the U.S. is similar to that in other areas and countries around the globe? Is BI/DW diffusion similar, for instance, between the U.S. and Western Europe? A global variation of this question would be if diffusion for BI/DW is similar between developed and developing countries. Insight into the global diffusion of BI/DW does provide potential guidance for investments, not only from a prospective adopter, but also external influencers, such as product vendors and consulting firms.

We obtained results showing that for DW/BI diffusion, practitioner’s research is being influenced more by external factors and in academic research diffusion internal influence is predominant, progressing differently than practitioner’s literature diffusion. One pertinent question arises: Is this only true for BI/DW diffusion or a generalized enough result that holds true for all technologies such as BPR, CASE, e-commerce, ERP and the like? Empirical studies are needed to explore this question further.

11. References


