Managing the Enterprise Business Intelligence App Store: Sentiment Analysis Supported Requirements Engineering

Michael Goul
Information Systems
W. P. Carey School of Business
Arizona State University
Tempe, AZ USA 85287-4606
Michael.Goul@asu.edu

Olivera Marjanovic
BIS Discipline
University of Sydney
NSW 2006 Australia
Olivera.Marjanovic@sydney.edu.au

Susan Baxley
Director, Comm. Relations
Teradata Corp.

Karen Vizecky
Modeling & Statistical Analysis
Target Corp.

Abstract

This paper posits that an app store delivery model for enterprise business intelligence is inevitable due to compelling business and technology drivers. A model of how employees will adopt BI apps is advanced based on theoretical foundations and recent thinking regarding how IT artifacts are becoming perceived of as “social actors within one’s network.” The theory cited asserts that individual’s resource management instincts impact sentiment. Repeated app usage is clarified as ongoing appropriation. Actual app store review data is used to test the proposition that sentiment analysis can aid in addressing current practice bottlenecks in requirements engineering. A sentiment analysis tool suite is applied to over 5000 reviews of productivity apps as a proof of concept. Findings support that sentiment analysis can address current bottlenecks to requirements engineering, but that certain types of reviews tend to elude algorithmic analysis. Future needs for sentiment analysis algorithms in the space are suggested.

1. Introduction

An interrelated set of business and technology drivers is reshaping the way that Business Intelligence (BI) will be deployed within the enterprise. In general, the consumerization of IT has required that organizational IT departments develop strategies for integrating personal devices like tablet computers and smart phones into future application deployments. Gartner estimates that by 2013, 80% of businesses will support a workforce using tablets, 75% of businesses will plan for end users who will be using personal devices to connect to the enterprise network with or without permission - and that the majority of end user devices will be employee-owned [9]. In addition to IT consumerization, enhanced device capabilities and the expanding availability of service functionalities within a single device are changing work styles. To reflect these trends, BI vendors are moving rapidly to provide fitting solutions. Organizations are at a crossroads with respect to the need to re-evaluate the impact of personal device proliferation on application architectures and platforms, with the concomitant need to pursue some sort of cross-platform application development strategy.

With the growing importance of BI applications in business, the consumption of BI functionality via handheld devices will likely be around 33% in just under two years [11]. For organizations developing a mobile BI strategy, there are a host of adoption impediments. Security will likely be a major concern, with some analysts suggesting that lack thereof could doom approaches that rely on ‘apps’ – an abbreviation for an application that ‘works,’ is often described as being ‘bite size’ and that provides a means to get things done quickly and effectively. App deployment is most often achieved with an app store – a distribution model popularized by Apple. In general, an app store provides a single connection point and centralized directory for software that is easily searchable, has community ratings and feedback, and makes it easy to acquire and install updates. Some proponents point out that app stores for mobile devices are no less secure than web-based counterparts for deploying other types of software, for example, cloud solutions (e.g., www.apps.gov).

Platform builders are evolving enterprise app stores – Jive Apps Market is an example [www.jivesoftware.com]. The company’s next iteration, Jive 5, enables IT departments to manage the visibility of third party vendor apps and in-house apps, although it is envisioned that many organizations may opt to open their app store to any apps being released into an app market. Major BI vendors are following suit with Oracle, MicroStrategy, IBM and SAP, for example, advancing in the space. Generic app store platforms are emerging as well, e.g., MobileIron,
Apperian, Rhomobile, JackBe and appcentral. Many claim that a major advantage of the app store approach is that a market with feedback and comments can open a new dialogue between IT and employee end users. Most agree that major capabilities for an enterprise app store must include: 1) A software distribution model with deployment, installation, updating, deletion or blocking functionalities, 2) Support for policy management, 3) Facilities for inventory management, 4) Security management capabilities (e.g., authentication, encryption, etc.) and 5) Service management functionality (e.g., Figure 1).

![Figure 1: BI app store model](image)

The store’s service management role is consistent with the ideal of connecting the IT Department to employee end users in such a way as to promote app innovation. For example, IBM has adopted a mobile apps store it calls WhirlWind, and the company’s CTO of mobility states, “IBM has embraced a crowdsourcing approach, giving employees the ability to submit, rate and comment on the apps, which has stepped up the pace of development” [13]. Genentech, Kraft, Nygard, Guess and others are finding that the apps approach drives down costs by reducing ‘bloatware’ and the long, traditional and costly structured application development lifecycle - because apps development is typically much faster, and it is done in smaller chunks [13].

Addressing IT consumerization issues, taking a new approach to more closely connecting the IT Department and employee end-users, and reducing the costs of development are among the rationale offered for an app store platform BI deployment approach. But for enterprise BI, in particular, what important advantages can be gained?

In the next section, we discuss current understandings of common enterprise BI development approaches with an eye towards possible enterprise BI app store advantages and disadvantages. From this, we posit that for the BI app store to be an effective deployment and innovation solution, it must address the BI developer/employee end user gap. Since there is little theoretical foundation for exploring this gap in the context of BI apps, we develop a model of BI app appropriation that we validate. That model is anticipated to be an important element for exploring feedback and rating data.

Since a crowdsourcing approach whereby feedback and comments can be collected and analyzed is deemed key to reducing the IT/end user gap, we investigate the potential of ascertaining design innovations from a dataset we scraped from Apple’s iTunes store using two review scraper tools. Our corpus includes a random sample of a total of 5036 reviews of the 23,534 available reviews through iTunes Connect for nine of the top ten best iPhone Productivity apps as recently ranked by BNET [16]. We used Clarabridge’s sentiment analysis tool suite to aid in validating our BI app appropriation model.

Our findings demonstrate that the crowdsourcing approach is valid for identifying potential app improvements, and we discuss how the approach is capable of supporting innovation. We conclude with an assessment of how sentiment analysis can be extended to be even more applicable to innovating and improving enterprise BI solutions.

2. Traditional BI and New Theoretical Motivations

In “Method Support of Information Requirements Analysis for Analytical Information Systems,” Stroh, Winter and Wortmann reviewed the state of the art as based on literature analysis and a survey of practice requirements to posit a research agenda for BI requirements engineering [15]. Requirements engineering is defined as the “sum of all activities (elicitation, documentation, validation and negotiation) which determine the requirements of an information system (component)” [15, p. 33-4]. They point out that by using the term ‘engineering,’ they imply a systematic and repeatable technique can be used to ensure requirements are complete, consistent and relevant. In the context of BI systems, the authors assert that since decisions are being supported, and decision makers are the intended end users, then the quality of requirements analysis has a very high impact on the effectiveness of solutions developed.

The authors’ methods and analysis lead them to conclude that there is an important need for documentation approaches that can more easily be understood by both business and IT. They also found there is substantial practice-based need for elicitation that is tailored to the BI context, i.e., there is too much reliance on methods and techniques that are generic to all systems development efforts. Finally, they conclude that the historic project-based view needs to be complemented with a process-based perspective that
reflects the fact that BI systems are expected to be under continuous modification, expansion and movement, i.e., a state of continuous evolution. The enterprise BI app store approach has the potential to address all three of these needs. However, there are few, if any, formal techniques and methods discussed in the literature directed to analyzing app store feedback and comments for requirements engineering purposes.

2.1. Theoretical Foundations

What would organizations and their employees gain from an enterprise BI app store? To address this, we propose a theoretical lens that has recently been applied in the context of mobile applications. First and foremost, and consistent with consumerization of IT notions, employees may prefer to use their own device to be able to access both personal and work-related information and services. An enterprise app store approach encourages this. There is a reduced learning curve in the case of software solutions that can be consistently navigated with gestures and other common representations that are inherent to devices that have been chosen by the end user for personal uses. But there may be something more.

Richardson’s recent dissertation, titled “Information Lifeline or High-Tech Tether: An Empirical Investigation of Workplace Connectivity Behavior,” posits that conservation of resources theory – a theory that is based in the study of stress – is very relevant to a construct she refers to as ‘connectivity behavior’ (CB) [12]. CB refers to how people use mobile communications technology to engage with work or work-related colleagues during non-work time. Conservation of Resources (COR) theory asserts that individuals strive to obtain, construct and protect what they value – particularly resources that include objects, conditions, energies and personal characteristics [4][10]. COR’s relationship to CB implies that individuals can achieve a perception of better control over their jobs by being virtually connected to their workplace. In the context of the enterprise BI app store, COR predicts high CB if an app provides the ability to scrutinize work performance indicators, drill down to identify issues that may threaten job security or that call for quick action before additional complexities arise - and COR even asserts that an ability to prevent what are referred to as ‘ego-threats’ will increase CB. In COR, stress is defined as a reaction to the environment in which there is a perceived or actual: a) threat of a net loss of resources, b) the net loss of resources, or c) a lack of resource gain following the investment of resources. Resources then become the single unit necessary for understanding stress – they are those objects, personal characteristics, conditions or energies that are valued by the individual or that serve as a means for attainment of objects, personal characteristics conditions or energies (4, p. 516).

COR also predicts what individuals do when they are not confronted with stress – they strive to minimize the net loss of resources. In non-stress periods, individuals may enrich resources by investing in other resources, develop resource surpluses to offset future loss, and if they can develop surpluses, they are likely to experience positive well-being. In the case of the enterprise BI app store, COR predicts high adoption among users who will utilize apps to minimize stress and reduce resource loss. They will also use and explore new apps in periods of non-stress to strive to increase resource surpluses. Further, COR supports the contention that enterprise app store comments and feedback will reflect opinions, sentiments and innovative ideas that reflect an app’s resource cost/surplus implications to individuals.

Overall, with COR as a foundation, the analysis of employee feedback and comments in an enterprise BI app store should go far beyond providing rough-cut crowdsourcing metrics – more complex, information-rich sentiment analysis can likely be leveraged to address the BI requirements elicitation, documentation and continuous improvement gaps identified by Stroh, et al. [15].

2.2. From “Tool” to “Social Actor”

Historical decision support research has framed the ‘system’ or artifact as a ‘tool’ [e.g., 14]. There is increasing evidence that people are viewing their consumer devices (mobile phones and tablets), in combination with the software services they deliver, as more than what might be described of as tools. In Carter, et al., an exploratory study finds that young people are developing emotional connections to their cell phones [2]. This connection became more pronounced as the cell phone was increasingly embedded into the context of an individual’s daily life. Consistent with COR, young people, “… appear to develop deeper emotional ties with the cell phone when usage contributes to a positive sense of self.” [2011, p. 32]. This finding is also consistent with recent thinking about IT adoption whereby factors surrounding the use of IT artifacts are aligned with the utilitarian benefits that users accrue from their interactions with those artifacts. Al-Natour and Benbasat posit that a relationship develops with an IT artifact that influences subsequent usage – with usage at each iteration based on prior beliefs and the status of
the relationship. They assert that an artifact can become a “social actor” within one’s network [1]:

“We adopt a perspective of IT artifacts that goes beyond that of productivity-enhancing tools, and posit that users perceive them as social actors as well. Consequently, users view their interactions with IT artifacts as interpersonal in nature, and react to them as though they are interacting in social situations. In assessing IT artifacts, therefore, users consider both the interpersonal interactions with them and their functioning as productivity tools” [1, p. 663].

The interaction model (Figure 2) they propose identifies appropriation as a key construct, and it includes feedback from both behavioral and relational beliefs to subsequent appropriations. In the context of BI apps and the services they provide, this model implies that over time, beliefs are impacted by appropriations in a manner akin to building some sort of customer loyalty. This dynamic process mirrors what research on services describes as customer loyalty based on positive, repeated service encounters. It also suggests that capturing and filtering beliefs can provide insights into appropriation.

First, an employee may have initial expectations about what an app can or should do based on the app’s description, its average ranking, other reviewers’ comments, etc. This initial, pre-launch perception will be validated or not depending on a set of subsequent usages that stream some value to the employee – i.e., the usages are examined by that employee to determine if the app will become a new resource that an employee can repeatedly leverage. We can examine a value stream for each stage of app appropriation by considering the return it provides. For example, in the simplest case, the initial connection phase, an employee who downloads an app expects that it will install properly, that its initial functionality is understandable and that its interfaces and basic behaviors match basic perceptions from the pre-launch phase. If early expectations are met, then there is a high likelihood that initial uses will test out how well the app delivers basic functionalities. As above, we assume that an app delivers on the basis of COR theory – it either acts to increase surplus resources, or protects resources that are already in the portfolio. In general, we suggest that these initial uses should lead to some form of productive use in what we refer to as structural integration of the app into an employee’s workflows.

If an app delivers a productivity advantage, then it may proceed to a next stage of appropriation. It is at this point that it begins to become more ingrained in how employees do their work. For example, if an app’s functionality can improve productivity to a point where it can inspire an employee to redesign their work processes around its usage, then the value of that app has gone beyond a more simplistic productivity gain to one where a work pattern is altered. It is logical to assume that COR theory would predict that the app has become a more central resource within that employee’s portfolio than it was prior to this phase, and that an employee will do more to protect the status of that resource if there are threats to change or delete it.

Finally, if an app enables the alteration of work patterns, then it may proceed to the highest level of ‘social actor’ interaction that is central to the notions of Al-Natour and Benbasat. We refer to this phase as embedded usage to imply that if the app can cross the boundaries of work/life balance that might have existed prior to the app’s appropriation. It is at this stage that there are potential pitfalls with app usage that can lead to employee burnout – or what some have referred to as an app addiction – consistent with the coining of the term, “crackberry” [3]. While the negative consequences are likely real, there are also potentially important contributions to work/life balance. Being physically free to move about and still manage to complete work tasks can enhance quality of life. In

2.3. Appropriation

In the context of the enterprise BI app store, COR theory predicts that people will gather surplus resources (e.g., ‘good’ apps) when they have time, and they will leverage those resources in ways that enable them to acquire new resources or protect their current portfolio of resources (time, job security, etc.). The appropriation notion of Al-Natour and Benbasat suggests that interactions with those resources will impact beliefs – both behavioral and relational – that subsequently impact repeated appropriation cycles. We posit that this collection of theoretical foundations leads to a model of enterprise BI app appropriation that proceeds through phases, not all of which may be realized.
addition, an app that can address dual resource area needs – i.e., provide resource surpluses to both work and life – then that app have be a very positive overall impact. It will also become a highly valued resource worthy of considerable protection. Figure 3 depicts these phases, and the shape of the curve hypothesizes that attempts to construe meaning from higher stage app reviews is an increasingly complex task. This is discussed in more detail in the next section.

3. Closing the “IT Dept-BI End User” Gap

The gap between IT Department BI developers and BI employee end users was addressed in the beginning of section 2. That gap relates to requirements engineering, specifically issues with elicitation, documentation and continuous improvement. It was posited that sentiment analysis of feedback and reviews posted in an enterprise app store might directly address these issues.

Sentiment analysis is also called opinion mining, c.f., [7][8]. In the context of app store feedback and reviews, sentiment analysis could provide opinion classifications to assert whether a reviewer is positive or negative about an app. More important, it could provide feature-based sentiment analysis. This type of analysis determines opinions on objects and their components, attributes and features. In the context of apps, this means that sentiment analysis can identify those features for which opinions are positive and those features where opinions are negative. This could guide developers in making next version improvements to an app – and the resulting documentation of opinions could address an important omission in existing requirements engineering processes.

There are two main mining tasks in feature-based sentiment analysis: 1) Identify the object features that have been addressed and 2) Determine whether the opinions have been positive, negative or neutral. Since app store reviews are typically free form, there can be multiple opinions and both positive and negative opinions in a single review. It makes most sense to apply sentence-level sentiment analysis to app reviews, and that technique is very complex because in reviews, one must separate sentences that state an opinion from those that assert a fact. An appropriate app store sentiment analysis tool can do feature extraction and then identify opinion orientation to features at the sentence level. In addition, a sentiment analysis tool suite must aggregate opinions by feature. This can be done by clustering what are referred to as opinionated sentences. (Note: A full discussion of sentiment analysis is beyond the scope of this paper, but these foundations are important for subsequent sections.)

In the context of the app store, we must also consider that opinions stated might be associated with a certain appropriation phase. For example, an opinion about a feature may be couched in the context of a level of appropriation – and might not be relevant to other levels. How relevant would a review be that was posted just after an initial download? Would it have meaning to most users? Could one user’s review be at a level of sophisticated appropriation that few users will ever attain? These questions and others require closer analysis of sentiment analysis on actual review data.

3.1. Review Data Set

We selected the iphone/ipad productivity apps category as the surrogate for enterprise BI apps for our study. Productivity apps are designed to help an individual get things done more quickly and efficiently. Other than finance and developer tools, this category was deemed closest to the orientation of the types of apps to be included in an enterprise BI app store. We extracted a random sample of 5036 reviews from the Apple app store using two scrapers – one open source and one proprietary. The first tool suite, provided by appFigures.com, gathers reviews in multiple languages and actually performs translations between languages. It is designed for apps developers and facilitates tracking reviews as they come in to the apps store. We used the public views capability and signed up for six months of service. There is no detailed sentiment analysis capability within the tool suite, but a new API is designed to support developers who want to download reviews for further analysis. Figure 4 shows a screen shot of appFigures review display page with the countries where reviews have been tracked. All reviews have a title, an opinion author, the date of review and a ranking. Rankings can
be sorted by most recent, most helpful, most critical, and most favorable. These latter two categories are applying some sort of sentiment analysis filter, but there are no details on how the filter works. It should be noted that appFigures always downloads reviews in real-time when an app’s reviews are scraped.

Figure 4: appFigures.com review interface

Figure 5 shows comments in French and their English translation. The tool suite supports translations to and from English, Spanish, French, German, Italian, Hebrew, Russian, Chinese and Japanese.

---

### Disappointing!
by prepalx on 11 jul 2011 version 4.0.3

Unable to create files for grading. Limited functionality. Considering the price, one can only be disappointed!

### Shame not to synchronize Dropbox
by alexb on 7 oct 2011 version 4.0.3

Sur documents to go

### Disappointed!
by jankesan on 19 mar 2011 version 4.0.3

perfect operation for one month ... and suddenly, without reason, can not synchronize files. the drainage application does not recognize either iPad or iPhone! Uninstall, restart, nothing happens, very frustrating ...

### Disappointing
by rpee24 on 11 jul 2011 version 4.0.3

Unable to add an image in word processing. At that price, it is unforgivable ...

---

The appFigures.com tool suite also provides app tracking. Figure 6 shows ranking trends for selected countries and by device (ipad, iphone). This information provides app developers with real time data being scraped from the Apple iTunes store through the iTunes Connect facility.

---

**Figure 5: appFigures.com language translation**

The appFigures.com tool suite also provides app tracking. Figure 6 shows ranking trends for selected countries and by device (ipad, iphone). This information provides app developers with real time data being scraped from the Apple iTunes store through the iTunes Connect facility.

---

**Figure 6: appFigures.com tracking**

We also used an opensourse utility to download files that we could then store in .csv format for uploading into a sentiment analysis tool suite. That tool is called AppReviewsFinder at http://www.massycat.co.uk/iphonedev/AppReviewsFinder/.

The reviews in our dataset are among the top 10 rated productivity apps [16]. We abstracted a sample of reviews for the following apps:

<table>
<thead>
<tr>
<th>Docs to Go (Premium)</th>
<th>Quick and dirty document viewing/editing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dragon Dictation</td>
<td>Turns spoken words into text</td>
</tr>
<tr>
<td>Fuze Meeting</td>
<td>Host or join web conferences and share media</td>
</tr>
<tr>
<td>Genius Scan</td>
<td>Scan/send documents, whiteboard contents</td>
</tr>
<tr>
<td>Hey Tell</td>
<td>Send voice messages (like a walkie-talkie)</td>
</tr>
<tr>
<td>Print Magic Business</td>
<td>Wireless printing by beaming documents</td>
</tr>
</tbody>
</table>

---

---

---

---
<table>
<thead>
<tr>
<th>App</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProOnGo Expense</td>
<td>Calculate mileage expense; prepare reports</td>
</tr>
<tr>
<td>Sugar Sync</td>
<td>Cloud storage tool with public link capability</td>
</tr>
<tr>
<td>World Card Mobile</td>
<td>Converts business card scans into contacts database entries</td>
</tr>
</tbody>
</table>

### Table 1: Apps included in analysis

From this sample, we first set out to validate if we could identify review comments that could substantiate the app appropriation phases of Figure 3. For purposes of validation, we analyzed the dataset to ascertain if there were complete reviews that fit each of the appropriation categories identified in the proposed model. We didn’t code the entire data set with a category for each review, but that could be an important part of a follow-up research project. These exemplary complete reviews did become an interesting part of our evaluation of sentence-level, feature-based sentiment analysis for supporting BI app requirements engineering; this is discussed in more detail in Section 3.4.

#### 3.2. The Appropriation Model

For each stage of the appropriation model of Figure 3, we investigated to ensure there was coverage by exemplary reviews in our dataset. This validity test would imply that the model has necessary elements, is practically valid, but it does not allow us to reach the definitive conclusion that it is provably sufficient. Here we present several reviews from our dataset that are representative of each phase of the appropriation model. Docs to Go supports mobile users by providing creation, editing and reviewing support for Microsoft Office documents on iPads and iPhones.

The pre-launch phase of the appropriation model typically has few if any reviews to mine. The initial connection phase, however, has numerous instances of reviews for all of the apps in our dataset. Following are two examples of Docs to Go reviews that fall into the initial connection/usage phase – in brackets are the reviewer’s provided name, the date of the review, the version reviewed and the rating, respectively:

“"I decided to purchase the more expensive version in hope that the email function would be available not only for MS exchange. Unfortunately they released the email function for the lesser expensive app and not for this version.?!?! It is also still not available that you can edit sheets; there are other apps available which let you send email and edit sheets. I cannot recommend this version for the price that it is offered."” [appwiz65, Jul 23, 2009, version 1.1.1, rating = 1]

“"Great app except who has their desktop connected to wifi? It works with my laptop but even though my desktop is through the same router I can't synch because its not wifi. A poorly conceived limitation. Next version should allow you to drag files onto the ipod/iphone or synch via other methods. The purpose of having documents to go is to get them off my desktop so I don't have to carry my laptop. Now I have to ask whtas the point??”” [Simpsoneh, July 23, 2009, version 1.0, rating = 2]

These reviews discuss connection and initial usage issues, they reflect expectations that are carried over from pre-launch perceptions, and they tend to deal with issues that would be discovered from initial app testing by an end-user. Some of the issues may be reconcilable by app support personnel or an end-user’s careful reading of documentation. The next stage is concerned with the structural integration of apps into an employee end user’s workflows. At this stage, the concerns begin to relate to one’s personal productivity and the features that are most significant; following is an example:

“"I have used Dropbox for about a year and use it to keep files in sync across many different computers. Having this in my iPhone will now be invaluable. I can't wait for the application to be updated for iPad as it will be a boon productivity capabilities of the iPad. Apple should take note - the syncing experience is far better than running everything through iTunes.”” [dougie76, April 24, 2010, version 3.2, rating = 4]

In the next phase, reviewers reflect changes to their work processes. They may also comment on specific features in the context of that work pattern shift. This positions feature comments into the context of the workflow changes – it thereby becomes more difficult to assess the broader need for the feature beyond the individual’s workflow. However, an understanding of the need for the feature in that context may well be significant to enable more end users to advance their usage of an app. Consider this review of ProOnGo Expense – an app to aid in tracking expenses - and a review of HeyTell – a walkie-talkie-like app:

“"I am a consultant and used to spend my Fridays/Saturdays punching in my expense report. No more! This product does a great job and its easy to use. I really recommend it!”” [thaltic, January 11, 2010, version .96, rating = 5]

“"This is an app I dl on a whim with one of my buddies while looking for a way to text with voice... Now I use it literally everyday and it's alot safer than driving while texting... Prolly the most used app on my phone after safari”” [Qwerty104, June 16, 2010, version 1.1.0, rating = 5]

These reviews show that a process change has occurred as facilitated by the app, but extracting app
features is more of a challenge. For example, in the earlier reviews, a feature of being able to connect to Dropbox was central. Here, it isn’t clear what specific features enabled the process changes. The final phase of the appropriation model has similar challenges because there is the issue of discovering that the review actually contains work/life balance aspects. Consider the following review for Docs to Go Premium where the issue of having multiple email accounts available implies that both personal and work-related messaging is appreciated:

“I had moved to the iPhone from the pocket pc platform. The only thing I missed was office document editing. Then finally a truly brilliant app came along. I am an IT manager so I need an app that can roll hard all day long with me as I am moving from project to project on the go. Docs to go with Exchange integration has delivered in a big way for me. Works as promised and I can pull docs from attachments or sync files from my laptop. Also I was amazed that I could send emails from all the other email accounts configured on my iPhone. The interface is so much more elegant than any pocket pc mobile office version I have seen. The only thing I wish it could do is pull attachments from folders within my inbox as I have some emails several folded deep. But this limitation can be worked around. If they fix that this app will be perfect. Well worth the investment.” [Steve@, November 3, 2009, version 2.0, rating = 5]

These reviews demonstrate that as appropriation advances, the feature-based aspects of reviews become less discernable, yet it is often the case that these reviews will be very important for developers. However, such a painstaking task of manually examining thousands or millions of reviews with an eye towards identifying app innovations and potential new requirements for a BI app store is not scalable.

3.3. Sentiment Analysis Evaluation

Sentiment analysis tool suites have advanced significantly in their sophistication over the past several years, and the capability to extract important information for marketers, product designers and numerous other business area professionals has become well documented. We sought to assess whether sentence-level, feature-based sentiment analysis could provide support for BI app innovation, needs documentation, and serve to support requirements elicitation. We used Clarabridge’s tool suite to perform comprehensive sensitivity analysis to assess potential. The tool suite has many options; we relied on an unsupervised algorithm for advanced users that performed a categorical analysis and a within category verbatim text clustering. Note that verbatim text is analyzed at the sentence level in the options we selected. As an example of the categories for Docs To go, consider Figure 7:

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>USE -&gt; LOVE -&gt; APP</td>
<td>206</td>
</tr>
<tr>
<td>WORK -&gt; GREAT; WORK -&gt; GOOD</td>
<td>18</td>
</tr>
<tr>
<td>SUPPORT -&gt; EMAIL; SUPPORT -&gt; TECH</td>
<td>10</td>
</tr>
<tr>
<td>IPHONE -&gt; WASTE -&gt; MONEY</td>
<td>7</td>
</tr>
<tr>
<td>DOC -&gt; GOOGLE -&gt; GO</td>
<td>3</td>
</tr>
<tr>
<td>DOCUMENT -&gt; WORD; EDIT -&gt; DOCUMENT</td>
<td>2</td>
</tr>
<tr>
<td>UPDATE -&gt; IPAD; UPDATE -&gt; GREAT</td>
<td>9</td>
</tr>
<tr>
<td>VERSION -&gt; PREMIUM; VERSION</td>
<td>1</td>
</tr>
<tr>
<td>FILE -&gt; DOC</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 7: Categories for Docs to Go

In the context of mining to support requirements engineering, it is important to identify those clusters of verbatims that can guide new feature identification, problem/issue identification, etc. For this reason, the categories for closest scrutiny are those that aren’t typically associated with strong positive sentiment (e.g., less informative categories like: USE -> LOVE; LOVE -> APP). For example, in category 0, (OPEN -> FILE; FILE -> DOC) issues about managing documents were surfaced.

As can be seen in Figure 8, the verbatims are subsets of overall reviews since the sentiment analysis tool suite was set to sentence level feature analysis. Some authors have multiple contributions and are identifiable – an important feature for requirements analysis for scenarios where the BI team may like to follow-up with subsets of employee end users. Category 2 of the analysis included 233 verbatims – they all relate to how well Docs to Go will work with Google Docs. In fact, the verbatims on issues with Google Docs related to editing were separated by the sentiment analysis from comments in Category 7: DOCUMENT -> WORD; EDIT -> DOCUMENT. Interestingly most of the verbatims that were negative and based on more general iPhone or iPad sentiment were clustered together in Category 8: IPHONE -> 4; WASTE -> MONEY. The “4” in that comment refers to the iPhone4, and so the association between the two was strongest for the cluster. We also found the overall sentiment analysis to be helpful in ascertaining how the tool suite was inferring positive and negative connotations to certain words. Figure 9 shows an example.

4175
Figure 8: An insightful cluster

3.4. Discussion

Overall, we found that sentence-level, feature-based sentiment analysis was very informative and can serve to document needed requirements, it can support elicitation processes, and an additional advantage is that it can support continuous analysis of app version releases. These three capabilities clearly address the gap identified by Stroh et al. [15]. However, we also find that the approaches we tested did not perform well on the more subtle phase 3 and 4 reviews of our appropriation model. Those stages are characterized by workflow pattern changes and work/life balance implications. We find this to be disappointing, but it clearly validates the difficulties associated with interpreting reviews that fall into those two phases. But those types of reviews are likely to be among the most important for discovering innovations. In addition, they can support the sharing of best practices throughout an organization.

Figure 9: Negative sentiment summary

We also examined some additional possibilities for sentiment analysis that resulted in interesting conclusions. First, we assessed the relationship between an author’s ranking of an app and the version number and found very little correlation across apps in the dataset. We also examined if sentiments were correlated with date/time. We found that there were no correlations. Many people tend to download an app, but they may not update it, so the comments on a particular version may come at a date/time long after a newer version has been released. This also indicates that many end users do not read other comments as they opt instead to “pile on” their own assessments. This is an important point for sentiment analysis and crowdsourcing assessments of app review content in general. We considered whether a version number was related to ratings. We did not find a correlation, which was somewhat surprising. It seems clear that different end users have different scales that are likely related to how much they use an app, how much they pay for it in comparison to the other apps they make the most use of, etc. Also surprising, there were very few instances of repeated reviews by end users – although there were some instances of a reviewer going back
and updating a review (sentiment analysis techniques did a poor job of assessing these types of reviews).

4. Conclusions

We believe that the ability to do sentiment analysis on reviews in the context of enterprise BI app store deployment will prove to be an invaluable source of organizational knowledge that has been missing in the application of current development paradigms. This bodes well for the organizational value that can be obtained by adopting enterprise BI app store approaches. We find that sentiment analysis tools do an excellent job in clustering sentence-level, feature-based comments that are rich in terms of app evolution needs.

Of course, our study has limitations. We relied on iPhone and iPad reviews, and we selected broad appeal productivity apps as surrogates for BI apps. We validated an appropriation model and a sentiment analysis approach with the same dataset. The model was practically validated, but will need more rigorous examination.

Our findings do demonstrate significant promise in showing how current requirements engineering impediments can be addressed with an enterprise app store and adopting sentiment analysis approaches to investigate employee end user reviews. The inadequacy of current sentiment analysis was shown to relate to the latter phases of our appropriation model where reviews detail an author’s work pattern changes or work/life balance findings that are difficult to extract using regular sentence-level, feature-based analysis. Research related to this inadequacy seems to be advancing. For example, adding domain understanding to current sentiment analysis algorithmic approaches will be an important area of study related to review analysis in the enterprise BI app store context, e.g., [6][7].

References


