Learning from post project reviews

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: CARRILLO, P.M. ... et al., 2010. Learning from post project reviews. IN: Anumba, C.J. ... et al. (eds.) Innovation in Architecture, Engineering and Construction. Proceedings of the 6th International Conference on Innovation in Architecture, Engineering and Construction (AEC), pp. 210 - 219.

Additional Information:

- This is a conference paper. It was presented at the 6th International Conference on Innovation in Architecture, Engineering and Construction, June 9-11, 2010, Pennsylvania State University, USA.: http://www.engr.psu.edu/ae/AEC2010/Proceedings.pdf

Metadata Record: https://dspace.lboro.ac.uk/2134/11302

Version: Accepted for publication

Publisher: © Loughborough University

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
LEARNING FROM POST PROJECT REVIEWS

Patricia Carrillo*1, Jenny Harding**, Alok Choudary**, Paul Oluikpe***
P.M.Carrillo@lboro.ac.uk, J.A.Harding@lboro.ac.uk, A.K.Choudary@lboro.ac.uk, poluikpe@hotmail.com
*Department of Civil and Building Engineering, Loughborough University, UK, **Wolfson School of Mechanical and Mechanical Engineering, Loughborough University, UK, ***Senior Manager – Knowledge Management, Central Bank, Nigeria

ABSTRACT

Post Project Reviews (PPRs) can provide a valuable source of learning for project teams. They are also known by other terminologies such as project closeout, project post mortems, etc, and attempt to document the project experience – both good and bad. In order to reflect their importance, many construction organisations now have policies towards the conduct of PPRs. The reports resulting from these PPRs are done with the best intentions of providing a rich and valuable source of learning. However, because many companies do not have the resources to examine their review reports, either individually or collectively, important insights are missed thereby leading to a missed opportunity to learn from previous projects.

Text mining offers a potential solution to companies that do not have the resources to analyse these reports. Text mining analyses large volumes of text to identify patterns and trends in order to extract information and knowledge that could improve process, and identify both good and bad practice. Text mining is a development of knowledge discovery and data mining; the latter uses numerical data and has been used successfully in a range of industry sectors such as banking, manufacturing and retail to improve customer satisfaction. Text mining is a relatively new approach and uses unstructured text, as found in PPR reports. It is thus ideally suited to overcoming the problem with organisations possessing a large number of PPRs that may provide very useful information and knowledge without the requirement for extra human resources to analyse them.

This paper investigates the potential use of text mining to identify vital sources of knowledge that can lead to learning from Post Project Reviews. Two UK construction contractors provided PPRs reports. The companies adopted radically different approaches to the style and content of their PPRs reports and thus provided an opportunity to investigate the success of text mining for different scenarios. In total 48 PPR reports were analysed. The companies’ reports were first pre-processed to allow then to be used in a text mining tool. The text mining tool also had to be customised, using ontologies, to suit the context of the reports. In addition, both companies were asked to identify key knowledge areas that are important to their businesses; these formed the basis of the key words and phrases that were used for text mining. Two techniques, namely Link Analysis and Dimensional Matrix Analysis were used to identify correlations between key words and phrases that appear across a range of different Post Project Review reports. The initial results are very promising because they help to

1 Corresponding author (P.M.Carrillo@lboro.ac.uk)
identify links and trends that would otherwise be difficult to identify without a substantial amount of manpower. One of the advantages is the graphical representation of the strength of correlations between key words that makes it easy to select areas for further investigation.

**Keywords:** Knowledge Management, Learning, Project Reviews, Text Mining

**POST PROJECT REVIEWS**

Lane (2000) defined a Post Project Review as “a formal review of the project which examines the lessons which may be learnt and used to the benefit of future projects”. There are many benefits that accrue to doing PPR (von Zedtwitz, 2003) but organisations surveyed by Newell et al. (2006) acknowledge not having time to analyse PPR reports to try to learn from these. The post project review sometimes is a huge silo of information which rarely gets analysed critically to reveal patterns of information that could help decision making in the project process. Due to this inability to convert the review contents into useful knowledge quickly, project teams abandon the report on the shelf and move on, thus ignoring knowledge which might be useful and even critical to future projects. There is evidence relating to the fact that most organisations consider post project reviews as an additional constraint (Bowen et al., 1994; Huber, 1996; and Saban, 2000).

Most construction companies carry out post project reviews but there exists some limited research in this area (Sowards, 2005; Carrillo, 2005; Kamara, 2003, CII, 2007). The importance of PPR is underscored by its mention both in knowledge management literature (Tan et al, 2006), construction (Carrillo, 2005), manufacturing (Koners and Keith,2007), information technology (Robertson and Terry, 2006; Disterer, 2002), space project management (Garon, 2006), R & D (von Zedtwitz, 2003), software development (Pyra, 2002), environmental studies (Braniš and Christopoulos, 2005), finance (Terry, 2004) and operations research (Terry, 2003). A lot of the applications of PPRs appear to come from business, information technology, space project management and aerospace.

**Benefits of conducting PPR**

Tan et al. (2006) and Carrillo (2005) indicate that PPRs facilitate collective learning, provide utilisable knowledge, benefit client organisations in a variety of ways, foster better project phase management and prevent knowledge loss. Besides the view that PPRs are viewed as additional constraints and companies often do not have the resources to critically analyse the reports, it is not very apparent why companies appear not to be taking advantage of the benefits of post project reviews.

**Problems in conducting PPR**

Despite the benefits of conducting PPRs, there are some constraints in the PPR process. Tan (2006), Carrillo (2005) and Garon (2006) highlighted the problems of PPR processes.
Staff turnover and redeployment of staff to other projects and the difficulty of remembering issues that happened long ago pose problems for PPR. The lack of an established format for representing knowledge and the time lag between the generation of knowledge and its capture might cause some knowledge loss to the organisation. As a result, Carrillo (2005) recommends that PPRs should be done as soon as the project ends or at intermediate stages. Knowledge should also be captured as the project goes along. This is in consonance with Sowards (2005) view that reviews ought to be done in phases rather than at the end of the project.

**Identified approaches to PPR**

Some typical recommendations on how to conduct PPRs exist in literature. Gibson et al., (2007) start by proposing a three step programme to developing a culture for PPRs. Baird’s (1999) process proposes a phased approach although it stopped short at going into details; Sowards (2005) five stage process focusing on establishing criteria, involving key people, discussing an agenda, documenting key learning points and disseminating these to people who should see them; Freedman and Weinberg (1977) categorisation of project processes and analysis of each constituent part for lessons learned, a viewpoint which Busby (1999) does not share as this could be disjoint and not holistic; Roth and Kleiner (1998) advocated a six stage process which begins with planning and then reflective interviews, distillation, writing, validation and dissemination. Sowards (2005), Roth and Kleiner (1998) and Schindler and Eppler (2003) are more specific in the sense that they categorise the processes into simple and measurable steps. However, these steps are not broken down into constituent units of activity (who, what and how).

A common consensus in the approaches highlighted above is the need to learn from PPRs. This paper explores learning from PPRs from the approach of extracting knowledge using text mining.

**TEXT MINING**

Text mining analyses large volumes of text to identify patterns and trends in order to extract information and knowledge that could improve process, and identify both good and bad practice. It uses such techniques as text analysis, information extraction, information retrieval, visualisation, clustering, categorization, database technology, machine learning, natural language processing and data mining (Harding, et.al, 2006). Karanikas and Theoudoulidis (2002) defined knowledge discovery in text as “the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in unstructured data”. Text mining utilises data mining and natural language processing algorithms, which when subjected to computer processes under certain limitations produce an identifiable pattern from a defined set of unstructured text. Figure 1 illustrates a text mining process from unstructured or semi structured textual data.
Fan et.al., (2006) highlighted the following strengths and benefits of text mining are useful in addressing the limitations and gaps found in PPR processes.

- TM can be used to extract relevant information of different types from one or more documents.
- TM can be used to gain insight about trends, relationships between people/places/organizations etc., by automatically aggregating and comparing information extracted from the documents of a certain type.
- TM can be used to classify and organize documents according to their content. This may be useful in classifying PPR documents at the start of a new project. The most relevant documents are pre-selected into groups on a specific topic, analysed for lessons learned and then disseminated to the most appropriate person at the start of new projects.
- TM tools and techniques such as link terms and clustering can be used to retrieve documents based on various sorts of information about the document content.

To explore the application of text mining in discovering knowledge from PPRs, a research methodology was needed which would process the PPPR reports and submit it to a text mining tool of which there are several commercial products on the market, each providing different functionality (Choudary et al, 2009).

**RESEARCH METHODOLOGY**

This research involved two construction sector collaborators. Company A is a services, building and maintenance group which provides services across the whole life of many types of buildings and infrastructure such as hospitals, schools, offices, industrial plant, bridges, waterworks or roads. Company B is an architectural and construction company which works with financial, property and retailing companies. Each company was asked to provide a number of their PPR reports and to identify approximately six key knowledge areas their company would be interested in for the purposes of text mining. Each of these key knowledge areas were then subdivided into more detailed topics e.g. “Safety” as the high level knowledge area and “Accidents” as a sub-set of Safety. In total, 27 reports were obtained from Company A and 21 reports from Company B. Figure 2 shows the stages of the process used to text mine the reports.
During the Preparation stage the PPR reports were pre-formatted and prepared for text mining in order to maximise the text mining output. These involved tasks such as removing unwanted text, the tabular layout, colour coding schemes and photographs as well as converting the .doc files into .txt files. The PPR reports were also manually examined to check the occurrence of keywords and phrases to provide a datum so that after the experiments the research team could determine if the automatic text mining process had successfully identified the available knowledge.

The Run Experiments stage involved pre-processing which is particularly important to improve the relevancy of the results achieved, and two types of pre-processing were carried out. Firstly, frequently occurring words which were unlikely to contribute to the identification of any useful knowledge were identified and subsequently “ignored” in the analysis. Secondly, synonyms or different representations of the same word were identified and marked in the text mining dictionary so that they were treated as being the same word. Various algorithms such as such as rules application, text analysis, ontology based approach and Link Analysis and Dimensional Matrix were then used to text mine the PPR reports. This was done using the PolyAnalyst text mining software after a review of possible text mining tools had been undertaken.

In the Analyse Results stage the results of the text mining were checked for relevancy, consistency and completeness using an iterative process of review within the project team.

The final stage consisted of Dissemination/Evaluation. Individual consultations were carried out with both companies. This led to a series of refinements and actions on the part of either the companies or the research team that would help the companies to use the results to the text mining exercise.

RESULTS
This section will focus primarily on the results of the Run Experiments and Analyse Results stages in order to demonstrate the outcome achieved from the text mining. The Run Experiment stage consists of a number of sub-stages. These are text analysis, rules application, Link Analysis and Dimensional Matrix.
Text Analysis
Text Analysis extracts and counts the most important words and word combinations from the PPR reports (see Figure 2). The left side of Figure 2 shows the keywords and phrases in the form of rules. A rule may be treated as a command to the software to retrieve data from the reports that agrees with specified keywords or a phrase. A combination of these words and phrases can be used to generate a new rule using Boolean operators (“AND”, “OR”). For example, these “rules” can be applied to retrieve information related to both “quality” AND “audit”. The results will include all reports related to both keywords.

![Image of software interface](image)

Figure 2: Example results of text analysis

Rules Application
Based on the key knowledge areas, a set of rules were created using a set of keywords and phrases. These rules were further applied to the Companies’ reports to identify a subset of reports dealing with similar issues. Rule application is an area in which domain experience is very important as an expert’s input is needed to identify a set of phrases which should be used to create a useful rule and the expert is subsequently also needed to identify the relevancy of the results and issues of importance. Furthermore, these rules can be used to as an input to other visualisation techniques e.g. Link Analysis and Dimensional Matrices.

Link Analysis
Link Analysis was applied to visualise the correlation between a set of keywords and phrases identified during the text analysis stage. The identified keywords/phrases under
were stored in the form of rules, which are then applied to the dataset. Examples of Link Analysis findings are shown in Figure 3.

![Figure 3: Correlation of keywords and phrases](image)

Each key knowledge area is identified by a colour coding. The strength of a correlation is indicated by the thickness of the lines linking the words or phrases. For example, there is a strong correlation between “safety” and “loss”, and between “delay” and “success”. These correlations could be of interest to the companies and may merit further examination to discover useful knowledge.

The major constraints observed during the application of Link Analysis were:

- The links identified could be either a positive correlation (e.g. “maintaining a good relationship” and “success of project”) or a negative correlation (e.g. “safety” and “loss”)
- Repetition of keywords/phrases. Some keyword appeared under multiple high-level headings and were therefore duplicated;
- High number of keywords/phrases. 30+ keywords made the Link Analysis very complicated and it was not easy to pick out the most important correlations; and
- The threshold values set during the text mining process. The higher the threshold value, the fewer the correlations found because there were insufficient links to be flagged up.

Dimensional Matrix Analysis

Dimensional Matrix Analysis compares a number of keywords in the reports and investigates their influence on each other. A dimensional matrix was created using six columns representing each knowledge area with several key words and rules. Each column consisted of different cells where each cell represented the keyword(s) to be searched for within the PPR reports. The user defines the values of one or more cells and then browses the subset of records, belonging to the selected cell which represents a
keyword or combination of keywords. For example, keywords/phrases that come under the “Finance” column might have occurred with other combination of keywords in different columns (such as time, quality, health and safety, etc) in the matrix. In Figure 4 the phrase “Additional Cost” occurred in six reports. When combined with “Extension of time” in the column Time, both terms appeared in four reports. The knowledge derived can be interpreted as due to the extension of time, four projects incurred additional cost. Furthermore, these two can be combined with Quality to identify how extension of time and additional cost affected quality of product.

LIMITATIONS

The research found that text mining is valuable to companies and could enhance the decision making process of companies based on lessons learned from mining the PPR reports. However, a number of issues were identified which could improve the results obtained as follows:

Keywords Identified
In the Preparation stage companies were asked to identify high level keywords (e.g. Time, Finance, Safety, etc.) that were important to their businesses. They were also asked to identify a number of sub-keywords that would fit under the selected headings (e.g. “Contract Programme” under “Time”). The results showed that in a number of cases the PPR reports did not include the sub-keywords identified. This lead to an anomaly of

Figure 4: Dimensional Matrix representing key knowledge areas.
the company not recording what was considered important or, these keywords were discussed in meetings other than at the PPR.

**Distance between keywords**
Caution should be exercised in making decisions using the relationships between keywords because the distance between keywords and their relevance to each other needs further exploration. The final decisions as to the value and worth of the identified knowledge must be made based on human judgement and experience.

**Threshold Values**
The research found that Link Analysis is sensitive to threshold values set during the text mining process. When a threshold value is set, certain keywords are excluded because they fall below the defined threshold. Since the threshold value is subjective, if it is set too low, it will include several links that are not really valid. Conversely, if the threshold values are set too high, it will ignore some of the important correlations. For example, when a threshold of 3 is set, the number of reports containing the keyword “quality” is reduced to 12 from 20.

**CONCLUSIONS AND WAY FORWARD**
The research found that text mining is indeed valuable to companies and could enhance the decision making process of companies based on lessons learned from mining the PPR reports. The text mining experiments identified some useful and relevant knowledge within the reports that would not have been previously picked up manually. For example, one of the companies was able to identify a “loss” with projects of a certain type. It is also possible that text mining will identify some irrelevant knowledge. However, this can be minimised by careful use of the set up parameters for the text mining techniques. The final judgements as to the value and use of identified knowledge must always lie with the human user. However, the text mining approach reduces the effort required by users when making these decisions, as it highlights potentially relevant reports or sections of text, hence substantially reducing the quantities of documentation that need to be read before useful knowledge is found.

The research also exposed some weaknesses in companies approach to PPRs if text mining was to be successful. Some of these are as follows:
- Although a company may have a corporate policy for the conduct and content of a PPRs, these were not universally followed with different regions adopting their own procedures.
- Companies cannot readily locate their PPR reports because these are stored in an ad hoc manner in several different locations and formats;
- The content and context of the report is important if non-attendees are to learn from the experience. In other words, it must be sufficiently detailed to explain the situation and the outcome but not too lengthy;
• The length of report was also one of the limitations as long reports can contain more irrelevant information. Techniques such as Link Analysis and Dimensional Matrix Analysis are more useful for comparatively small reports varying from 3-8 pages.
• The process of extracting knowledge from PPRs is iterative. Keywords and key knowledge areas needed refinement by the companies’. This was necessary to ensure that results are relevant to the business context of companies.

In terms of the actual text mining, the Text Analysis, Link Analysis and Dimensional Matrix Analysis worked best in analysing the PPR reports. The reason was that the companies structured their reports consistently across various headings. Text analysis identified the keywords and key knowledge areas in the form of rules which were important for analysis purposes. Application of these rules enabled Link Analysis to consistently identify the linkages between the chosen knowledge areas. It was also possible to compare results with the manual extraction of keywords and link key knowledge areas across different reports, because of this consistency of approach.

REFERENCES

Baird (1999)
Bowen et al., 1994
Braniš and Christopoulos, 2005
Busby (1999
Carrillo, 2005
CII, 2007
Disterer, 2002
Freedman and Weinberg (1977
Garon, 2006
Gibson et al., (2007
Huber, 1996
Kamara, 2003
Koners and Keith, 2007
Lane (2000)
Newell et al. (2006)
Pyra, 2002
Robertson and Terry, 2006
Roth and Kleiner (1998
Saban, 2000).
Schindler and Eppler (2003
Sowards, 2005
Tan (2006)
Tan et al, 2006)
Terry, 2003
Terry, 2004
von Zedtwitz, 2003
Fan et al., (2006)
Choudary et al, 2009