Text Mining of Post Project Reviews

Paul Oluikpe,
Loughborough University, UK (email: p.i.oluikpe@lboro.ac.uk)
Patricia Carrillo,
Loughborough University, UK (email: p.m.carrillo@lboro.ac.uk)
Jenny Harding
Loughborough University, UK (email:j.a.harding@lboro.ac.uk)
Alok Choudhary
Loughborough University, UK (email: a.k.choudhary@lboro.ac.uk)

Abstract

Post Project Reviews (PPR) are a rich source of knowledge and information for organisations - if they have the time and resources to analyse them. Too often such reports are stored, unread by many who can benefit from them. PPRs attempt to document the project experience – both good and bad. If these reports were analysed collectively, they may expose important detail, perhaps repeated between projects. However, because most companies do not have the resources to examine these PPR, either individually or collectively, important insights are missed thereby leading to a missed opportunity to learn from previous projects. Hidden knowledge and experiences can be captured by using knowledge discovery and text mining to uncover patterns, associations, and trends in data. The results might then be used to enhance processes, improve customer relationships, and identify specific problem areas to address.

This paper outlines an ongoing research project that investigates the use of knowledge discovery and text mining on Post Project Reviews. An illustrative example will be presented using case studies from the construction sector. The PPR processes of two construction companies were mapped with the aim of understanding the context, format, terminologies used and key knowledge areas suitable for text mining. The textual examination of the PPR reports was complemented by semi-structured interviews and workshops to understand the production and content of the reports. Preliminary results highlight that although organisations have publicised, standard processes for PPR, there is a variance in how these are conducted and produced on a regional basis. These variances provide a number of challenges for organisations from a corporate perspective. Also, there is an over-reliance on key individuals with little attempt to make some of their knowledge more explicit and therefore easier to disseminate between project team members. This paper summarises the challenges in identifying the type of knowledge to be text mined, the format of PPR reports and the process of conducting PPR. It will also highlights the development of suitable

ontologies for text mining PPR reports and provides recommendations on how to improve the PPR process of companies.

Keywords: Post Project Reviews, knowledge management, text mining, construction, learning.

1. Introduction

Post Project Reviews (PPR) have been said to represent a rich source of information, knowledge and data for organisations, that is, if organisations have the time and resources to exploit them. The recent interest in literature on PPR is not unconnected to the perception that within these reports could lie hidden knowledge in the form of previously unknown patterns and associations that could lead to improved business decisions, effective trouble-shooting and organisational learning. Analysing these reports collectively could expose important details probably repeated between projects. However, the fact remains that most companies have little or no resources to analyse the reports and take advantage of the business insight they may provide. Against this backdrop, this paper aims to address the issue of extracting knowledge from PPR in the construction sector by using text mining tools and techniques. Illustrative case studies of two construction companies have been used to provide opportunities for applying theoretical and research concepts into practice.

2. Review of Post Project Review Literature

A Post Project Review (PPR) is defined as "a formal review of the project which examines the lessons which may be learnt and used to the benefit of future projects"[1]. The basic motivation for doing PPR is to learn from successes and failures. A review of literature however suggests that in some organisations, the conduct of PPR is in danger of becoming a mere formality rather than an opportunity to learn. A survey in research and development (R&D) indicates that organisations are aware of the benefits of conducting PPR but do not utilise the full opportunity to learn from these [2]. The reasons for failing to learn from results of PPR were explained by Newell *et al.,* [3] as relating to the lack of awareness that critical knowledge vital to process improvement resides in the reports. Williams [4] concludes that in theory, although organisations do have PPR processes in place, in practice, PPRs frequently do not take place and lessons learned from failed projects are often abandoned for a variety of reasons which include fear of management consequences. Organisational apathy towards the PPR process and failure to use the results of PPR is highlighted in Bowen et al., [5], Huber [6], and Saban et al., [7]. There is evidence in literature that PPR

appear to be considered as an additional constraint to adhere to management stipulations rather than as an opportunity to learn. However, there are benefits accruing to the conduct of PPR.

2.1 Benefits of Post Project Reviews

Benefits accruing to organisations from conducting PPR are highlighted in Tan et al., [8], and Carrillo [9]. Post Project Review meetings can provide opportunity for facilitating collective knowledge and can also yield knowledge that can be utilised. In addition, they benefit the client organisations by improving processes and relationships. PPR have potential for enhancing better project phase management and also prevent the loss of useful knowledge.

2.2 Post Project Review in Construction

There is little research on PPR in construction. Published research includes Sowards [10], Carrillo [9], and Kamara et al., [11]. The importance of PPR is underscored by its frequent mention in knowledge management literature [8], construction [9], manufacturing [12], information technology [13,14], space project management [15], R & D [2], software development [16], environmental studies [17], finance [18] and operations research [19]. Therefore, there is a need for further research the area in construction.

2.3 Post Project Review Approaches

A number of approaches to conducting PPR were reviewed and categorised from literature. A breakdown of these approaches is presented in Table 1.

Table 1: PPR approaches in literature

PPR as KM technique	Process
• Tan et al., [8]	• Sowards, [10]
• Robinson et al., [20]	• Roth and Kleiner, [21]
• Newell et al., [3]	• Branis and Christopoulos, [17]
	• Von Zedtwitz, [22]
Systemic Approach	**Collective Learning**
• Garon, [15]	• Carrillo, [9]
• Williams, [4,19]	• Grobelnik and Mladenic, [18]

The Knowledge Management approach is common in Tan et al., [8], Robinson et al., [20] and Newell et al., [3], while the Process Approach is advocated by Sowards [10], Roth and Kleiner [21], Branis and Christopoulos [17], and von Zedtwitz [22]. The Systemic Approach is seen in

literature such as Garon [15] and Williams [19], suggesting the use of cognitive mapping and system dynamics in carrying out PPRs. The Collective Learning Approach is advocated by Carrillo [9] and Grobelnik and Mladenic [18]. This perspective is not significantly different from the knowledge management perspective, however it could be said it is a specifically focused and targeted approach of conducting PPR which links key people within the project.

3. Methodology

The research reported in this paper is part of an ongoing research project which aims to extract knowledge from PPR of construction and manufacturing companies. The aim of the project is to improve project performance by providing access to relevant trends, patterns and observations from previously completed projects. More specifically, the project is exploring the use of knowledge discovery and text mining on construction and manufacturing PPR reports to extract potentially vital knowledge and information. This paper looks at the structure, content, format and process of conducting PPR, and will also describe the process of developing ontologies for text mining PPR reports of collaborators.

The methodology adopted for reviewing the PPR processes of companies is as follows:

- Review of documentation with regard to content, format, structure and identification of key knowledge areas;
- The use of semi-structured interviews to identify and confirm the PPR process; and
- The use of workshops and interviews to clarify the PPR process, key knowledge areas and other issues identified from documentation provided.

Case study Selection: This project has two case studies associated with two collaborators of the project. Company No 1 is a services, maintenance and building 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 No 2 is an architectural and construction company which works with financial, property and retailing companies. They also work with the entertainment and leisure industry as well as manufacturers in engineering, pharmaceuticals and food. Company No. 2 has a fairly good spread of services across the UK and also has a system for conducting PPR. The sections that follow will present the findings of the review of PPR reports with regard to structure, content, format and the process of conducting reviews.

4. Findings

This section presents the findings from the review of PPR reports with regard to the expectations of companies from doing PPR, structure and format of reports, key knowledge areas, PPR process and the challenges of the PPR process for companies.

4.1 Expectations and key issues in companies' PPR

Both of the companies participating in this study expect the results of PPR to lead to improved process, client retention, increased turnover and competitiveness in the market. These forces drive the conduct of PPR within the companies. Earlier indications from interviews and review of documentation during the first phase of the study showed that the PPR processes of the companies are managed. However, it is doubtful whether the results of PPR are adequately collected and disseminated to people who should see them.

4.2 Structure and format of Post Project Reviews

The PPR reports of the companies were structured based on important headings critical to the companies' processes.

Company No 1: Two types of PPR reports exist within this company. One is called a "Site Debriefing" and utilises a 16 heading structure, minimal text, and consistency across board. The other is called "Project Post Completion Review" and utilises a 7 heading structure, ample text and description but yet there is some inconsistency in length, format and subheadings. This report incorporates as much text as possible and is particularly narrative. Its advantage over the Site Debriefing is the use of ample text while the Site Debriefing has the advantage of being better structured.

Company No 2: The PPR report of Company No 2 is called a Project Closeout. A Project Closeout is typically between 16 and 20 pages long. It has a 16 heading structure and is consistent across the board. The closeout report contains very short sentences or phrases which are direct to the point. However, to someone who did not attend the meeting, these might not convey sufficient information and therefore limit the understanding of the context behind the reports.

4.3 Key Knowledge Areas

The review of documentation resulted in the identification of words and phrases which constitute important areas for both companies. The process of reviewing company documentation was iterative and the result was that companies were asked to confirm, delete or include additional words and phrases from the reports which reflected the important areas of their business. These words and phrases were considered the key knowledge areas of the companies. The key knowledge areas are very useful in the process of developing ontologies for text mining of the reports. It was considered important to identify those words and groups of words which are relevant to the companies in order to enhance the text mining process on the reports.

4.4 Post Project Review Processes

The PPR processes of both companies were mapped based on documentation provided. The PPR process maps were presented to the companies during workshops and meetings for clarification, remapping, adjustments and agreement on what should be reflected on the process map. Three project stages were identified as relevant to both companies. These are: Pre-construction, Construction and Post-construction stages. Table 2 below illustrates the PPR processes of the companies.

Table 2: Post Project Review Processes

	Preconstruction	Construction	Post Construction
Company 1	**Pre-commencement Meeting** • Project team • Risk Register • Contracts/Requirements • Text mining Reports • Client Requirements **Outcomes** • Critical Success Factors	*Project Execution*	**Agenda** **PPR meeting** • Site Debriefing report • PPR report **Outcomes** Lessons Learned • Store-BPL(Intranet) • Disseminate • Reuse Knowledge (PCM) • Improve Process
Company 2	**Project Launch Workshop** • Risk Register • Client Requirements • Project Team • Contract/Requirements **Outcomes** • Success Factors • Roles/Responsibilities • Desired Outcomes	*Project Execution*	**Agenda** **Project Closeout Meeting** • Closeout Report **Outcomes** Highlights/Lowlights Key Learning Points • Disseminate-P2P • Store on intranet • Improve process

Findings suggest that the Pre-construction stages of companies' PPR process have inputs and outcomes. The inputs and outcomes are bulleted under Pre-commencement meetings for Company 1 and Project Launch Workshops for Company 2. For Company 1, the Critical Success Factors feed into the Construction stage of the project while the Success Factors, Roles and Responsibilities and the Desired Outcomes feed into the Construction stage for Company 2. The Post-construction stages of both companies are characterised by the formulation of an agenda leading to the PPR meeting or Project Close Out meeting. The outcome of the review meetings are the PPR/Site Debriefing report and the Project Closeout Report respectively. These reports have outcomes in the form of Lessons Learned and Key Learning points. The outcomes of the pre-construction stage are different for both companies and the method of utilising lessons learned from the review meetings are also different. In Company 1, lessons learned are fed back to pre-commencement meetings and dissemination of the lessons learned is also done using the Best Practice Library on the intranet. Interviews and meetings however reveal that this method is not working efficiently. For Company 2, dissemination of the key learning points is done on people to people basis. There is not the tendency to refer staff to the intranet to retrieve key learning points of particular projects. Again, it has been discovered that these methods of dissemination and utilisation are not efficient in gaining the attention of people who should see the results of PPR.

4.5 Potential and criteria for text mining

The PPR reports of the companies have potential for being text mined, however, in their current format, it is doubtful if text mining processes could be maximised. The reports therefore need restructuring to enable the companies to get the best results. This conclusion arises from preliminary text mining investigations carried out on the reports. The review of PPR reports and processes of collaborators identified a number of challenges faced by the companies in regard to PPR.

4.6 Challenges and critique of PPR processes

A number of challenges are faced by the companies in regard to Post Project Reviews.

Reliance on key individuals: There is an over-reliance on a few individuals in designing, conducting and presenting the reports of PPR. While this might have the advantage of consistency in the format of the reports, the drawbacks include loss of knowledge if the member of staff leaves the organisation and also a lack of input from other staff who might have ideas and contributions on how to improve the process.

Format of reports: These staff also format the reports based on the headings they consider useful and important, hence it is not certain that other people's perspective and approach are taken into consideration. The structure of some of the reports may present difficulties for text mining processes. For consistency in results of text mining, there needs to be uniformity in presentation of the reports. The reports from Company 2 have very short sentences which are disadvantageous for text mining purposes. The reports also require major restructuring to be able to give rich insight into the projects reviewed.

Continuity of staff at meetings: Sometimes the staff who attend the start-up meetings are not the same people who attend the review meetings so there is an issue with continuity. It might be difficult for people who were not in the review meetings to understand the contents of the reports and the context in which they are presented.

Dissemination of results: The companies rely on a few staff to disseminate the results of the reviews and this is completely informal. Staff responsible might forget to share such knowledge and others will miss out on useful information. Unless this is made a requirement at the pre-commencement stage of the project, there will be a lost opportunity to re-use knowledge gained from previous projects.

Reusing knowledge from the reports: Finally, there is no systematic mechanism for storing, retrieving, utilising and disseminating lessons learned. Interviews and workshops provide a clue that companies do not adequately feedback the results of the review meeting to future projects. There is the need to implement feedback mechanisms that will ensure that the relevant people will see the results of the review. The companies need to do more to ensure that results of PPR are systematically analysed and utilised. Text mining the PPR reports may provide opportunities to learn lessons from all the reports over a period of time. Following the identification of challenges of PPR of collaborators, the PPR reports were further examined and analysed with the aim of developing ontologies to aid the understanding and mining of text from the reports. This will enhance the extraction of knowledge from these reports collectively or individually to facilitate learning from the project process.

5. Ontology Development

In order to search the PPR reports for trends, the use of certain words and phrases as well as their relationships need to be classified. Thus ontologies need to be developed. The methodology adopted for developing ontologies for the PPR reports of the companies followed a two stage process. As mentioned earlier, key knowledge areas were identified based on interviews and examination of the PPR reports. These key knowledge areas were used as the basis for analysing and classifying ontologies for the reports. Two groups worked independently on the manual examination of the PPR reports. The first group approached the development of ontologies by matching key knowledge areas with relevant words and phrases in the body of the reports. These words and phrases were in turn collated, classified and tabulated under the broad key knowledge areas earlier established with references to the pages where they occurred and the frequency of occurrence. The second group approached the development of ontologies from word classification using the tree structure whereby parent words were broken down into various sub-classes in a hierarchical structure. It was however observed that some issues in the reports which were clearly important to the companies were not highlighted in the key knowledge areas identified. On the other hand, the reports did not reflect consistently some of the key knowledge areas highlighted by collaborators. The manually collated and proposed ontologies were discussed in team meetings and agreements were reached on the inclusion or exclusion of some words or phrases from the list of proposed ontologies. The two different approaches to ontology development were discussed and agreements reached on combining these for the purposes of generating a definitive list of ontologies for the PPR reports of collaborators. The next stage of the research will entail using these ontologies to experiment with available text mining software. The following section will summarise and conclude the findings of this research.

6. Findings and Conclusions

This section summarises the findings of the research and concludes the paper.

6.1 Findings

A number of findings were made from the review of PPR reports and consultation with collaborators as have been presented and described in earlier sections.

a) Project pre-commencement meetings and activities influence the PPR process as documentation and decisions taken during this period are considered during the Post Project Review meeting.

b) There is a reliance on key people to disseminate the lessons learned from the Post Project Review meeting. This could create problems for the organisations if such staff leave the company.

c) There is scope for improvement in the contents and structure of PPR reports and the dissemination of PPR results.

d) It was discovered that pre-commencement meeting reports, the PPR reports and the agenda for both reports have a relationship. This is because what was discussed in the pre-commencement meetings are often reviewed during the PPRs. In essence, one can track how they are implemented or otherwise. This creates the opportunity for any text mining process to link results within these documents in the project process and to compare these with the project outcomes or lessons learned.

e) The collaborators have identified potential knowledge areas that could be usefully text mined. These knowledge areas have been ranked according to priorities set by the collaborators.

f) The kinds of information needed by new projects as identified by collaborators are within the categories that could be identified from text mining of PPR reports.

g) The process of development of ontologies identified that certain key knowledge areas identified by companies' staff during the review of PPR reports were not reflected adequately in the reports. On the other hand, there were key areas which stood out in the reports but were not reflected as key knowledge areas.

6.2 Conclusions

This paper investigated the scope for applying text mining on PPR reports of construction companies using two industrial case studies. This led to the mapping and validation of the process using semi-structured interviews and workshops.

A number of recommendations were made for improving the PPR processes of collaborators. First, with the PPR reports of Company No.1, there needs to be consistency across the units of the company. This will help text mining results to be consistent. Company No.2 on the other hand needs to incorporate more text in the body of the reports to help readers to understand the context of the reports. Secondly, more people need to be involved in the PPR process. This is to ensure

that knowledge is not lost when staff responsible for PPR meetings leave the company. Third, the dissemination of the results of PPR meetings needs to be taken more seriously by the companies. There is not enough being done at the moment to ensure that the right people get to know about the results and utilise them. A policy of dissemination of PPR results needs to be put in place by the companies. There needs to be a periodic review of the structure, format and content of the reports to ensure that they are meeting the objectives of doing PPRs.

The next stage of the project will develop ontologies for classifying and analysing Post Project Review reports of collaborators. This process will be iterative involving meetings with project team members to review and agree on proposed ontologies. Following this will be the experimental stage where the ontologies are input into text mining software to help determine the relationships and trends found in the PPR reports.

References

[1] Lane, K. 2000. Project Management Today. URL: <http://www.projectnet.com> [accessed 25.03.07].

[2] von Zedtwitz, M. 2003. Post-project reviews in R& D. Research-Technology Management 46(5): 43-49.

[3] Newell, S, Bresnen, M ,Edelman, L, Scarbrough, H and Swan, J. 2006. Sharing knowledge across projects: limits to ICT-led project review practices. Management Learning 37(2): 167-185.

[4] Williams, T. 2004. Identifying the hard lessons from projects easily. International Journal of Project Management 22(4): 273-279.

[5] Bowen, H, Clark, K, Holloway, C and Wheelwright, S. 1994. The perceptual enterprise machine—seven keys to corporate renewal through successful product and process development. New York: Oxford University Press.

[6] Huber, G. 1996. Organizational learning: a guide for executives in technology-critical organizations. International Journal of Technology Management 11(7/8): 821-832.

[7] Saban, K, Lanasa, J, Lackman, C and Peace, G. 2000. Organizational learning: a critical component to new product development. Journal of Product & Brand Management 9(2): 99-119.

[8] Tan, H, Carrillo, P, Anumba, C, Kamara, J, Bouchlaghem, D and Udeaja, C. 2006. Live capture and reuse of project knowledge in construction organisations. Knowledge Management Research and Practice 4: 149-161.

[9] Carrillo, P. 2005. Lessons learned practices in the engineering, procurement and construction sector. Engineering, Construction and Architectural Management 12(3): 236-250.

[10] Sowards, D. 2005. The value of post-project reviews. Contractor 52(8): 35-36.

[11] Kamara, J, Anumba ,C, Carrillo, P and Bouchlaghem,N. 2003. Conceptual framework for live capture and reuse of project knowledge in Amor, R. (ed.) Construction IT: Bridging the Distance, *Proceedings of CIB W78 International Conference on Information Technology for Construction*, Waiheke Island, New Zealand, 23-35 April, 1999: 178-185.

[12] Koners, U and Keith, G. 2007. Managers perception of learning in new product development. International Journal of Operations and Production Management UK 27(1): 49-68.

[13] Robertson, S. and Williams, T. 2006. Understanding project failure using cognitive mapping in an insurance project. Project Management Journal 37(4): 55-71.

[14] Disterer, G. 2002. Management of Project Knowledge and Experiences. Journal of Knowledge Management 5: 512-520.

[15] Garon, S. 2006. Space project management lessons learned: a powerful tool for success. Journal of Knowledge Management 10(2):103-112.

[16] Pyra, J. 2002. Risk management post analysis: gauging the success of a simple strategy in a complex project. Project Management Journal 33(2): 41-48.

[17] Braniš, M and Christopoulos, S. 2005. Mandated monitoring of post-project impacts in the Czech EIA. Environmental Impact Assessment Review 25(3): 227–238.

[18] Grobelnik, M and Mladenic, D (2005). Automated knowledge discovery in advanced knowledge management. Journal of Knowledge Management 9(5): 132-149.

[19] Williams, T. 2003. Learning from projects. The Journal of the Operational Research Society 54(5): 443.

[20] Robinson, H, Carrillo, P, Anumba, C and Al-Ghassani, A. 2005. Knowledge management practices in large construction organisations. Engineering, Construction and Architectural Management 12(5): 431 – 445.

[21] Roth, G and Kleiner, A. 1998. Developing organisational memory through learning histories. Organisational Dynamics 27(2): 43-60.

[22] von Zedtwitz, M. 2002. Organisational learning through post-project reviews in R&D. R&D Management UK 32(3): 255-68.