



# A Text Mining Approach for Extracting Lessons Learned from Project Documentation: An Illustrative Case Study

**Benjamin Matthies**

*South Westphalia University of Applied Sciences, Germany*

Lessons learned are important building blocks for continuous learning in project-based organisations. Nonetheless, the practical reality is that lessons learned are often not consistently reused for organisational learning. Two problems are commonly described in this context: the information overload and the lack of procedures and methods for the assessment and implementation of lessons learned. This paper addresses these problems, and appropriate solutions are combined in a systematic lesson learned process. Latent Dirichlet Allocation is presented to solve the first problem. Regarding the second problem, established risk management methods are adapted. The entire lessons learned process will be demonstrated in a practical case study.

*Keywords:* project knowledge management, project documentation, lessons learned, text mining

## Introduction

Lessons Learned (LLs) can be described as key experiences (e.g., failures or success factors) that have a specific business relevance for correcting future behaviour in a positive way (Jugdev, 2012; Kotnour & Kurstedt, 2000; Weber, Aha, & Becerra-Fernandez, 2000; von Zedtwitz, 2002). By their very nature, LLs are therefore particularly important building blocks for continuous learning in constantly evolving project-based organisations (Boh, 2007; Disterer, 2002; Jugdev, 2012; Kotnour, 2000; Navimipour & Charband, 2016; Schindler & Eppler, 2003; Williams, 2007). A corresponding LL process has the specific purpose to support the acquisition, archiving and the targeted reuse of the key experiences gained during completed projects in planning future projects (Disterer, 2002; Keegan & Turner, 2001; Middleton, 1967; Parnell, Von Bergen, & Soper, 2005). The aim is to avoid repeating already made mistakes or solving problems that have already been solved in previous projects. In this context, several studies show that the consistent reuse of this wealth of experience can have a positive influence on performance in projects (see, e.g., Hong, Kim, Kim, & Leem, 2008; Kululanga & Kuotcha, 2008).

Nonetheless, the practical reality is that LLs are often not consistently reused for organisational learning and for the necessary modification of existing policies, procedures or tasks in project-based organisations. Several research studies have proven that historical key lessons often remain disregarded in future projects, even if such evidence is codified in LL documentation and has been made available in public knowledge databases for project managers (see, e.g., Almeida & Soares, 2014; Carrillo, Ruikar, & Fuller, 2013; Duffield & Whitty, 2015; Newell, Bresnen, Edelman, Scarbrough, & Swan 2006; Williams, 2007). In this context, two problem areas are commonly described in studies on the reuse of LLs. First, the archives of mostly text-based LL documentation are too extensive to manually identify and summarise the relevant knowledge by project employees with reasonable effort (Barclay & Osei-Bryson, 2010; Choudhary, Oluikpe, Harding, & Carrillo, 2009; Haksever, 2001; Newell et al., 2006; Menon, Tong, & Sathiyakeerthi, 2005). This problem of information overload (see Haksever, 2001) becomes even more obvious when one looks at large project-based organisations where several hundred projects are often being carried out in parallel and at least as much project documentation is being archived (see Prencipe & Tell, 2001). Several researchers therefore emphasise the need for computer-aided techniques that can process the comprehensive, mostly textual inventories of documents and can identify, query and also summarise the knowledge that is relevant for the respective project task (Al Qady & Kandil, 2013; Choudhary et al., 2009; Menon et al., 2005). The second problem with the consistent reuse of LLs is that even when LLs have been identified and the project manager is aware of them, they are rarely implemented effectively into an organisational learning process (see Almeida & Soares, 2014; Barclay & Osei-Bryson, 2010; Newell et al., 2006). In this regard, Newell et al. (2006) were able to determine a general lack of awareness of the existence and the actual value of the knowledge contained in project documentation. Furthermore, Barclay and Osei-Bryson (2010) also make the critique here that appropriate assessment and subsequent implementation of LLs into project planning practices often lie with the intuitive judgements made by the individual project manager, influenced by his or her subjective opinions, experience, diligence or care. However, in order to learn consistently from the collective experience of project-based organisations, project management methods and appropriate LL processes are required to ensure that LLs are not only recorded, but are also properly assessed and consistently utilised in future projects (see Fosshage, 2013; Weber, Aha, & Becerra-Fernandez, 2001; Wellman, 2007; Williams, 2007).

The two problem areas described above are evident in the project management practice and research. Conventional project management guidelines, such as the PMBOK® guide or the PRINCE2 framework (see Project

Management Institute, 2013; Office of Government Commerce, 2009), reinforce the importance of reusing historical LLs in future projects, but they do not provide any concrete recommendations for extracting them from the extensive knowledge repositories for assessment (see also Duffield & Whitty, 2015). The present work addresses these problem areas, and appropriate solutions are combined in a LL process containing the semi-automatic extraction of LLs, their assessment and implementation. The ultimate goal is to contribute to the further development of dealing with LLs in project-based organisations. For this purpose, Latent Dirichlet Allocation (LDA) is presented in relation to the first problem described above. Regarding the second problem, risk management processes and methods are adapted and subsequently implemented in the assessment and implementation of the previously extracted LLs. For a better understanding, the entire LL process will be demonstrated and evaluated in an illustrative, practical case study. This case study demonstrates how 13 meaningful LLs are extracted from a textual database consisting of a total of 68 sets of project documentation. The extracted LLs will then be assessed and implemented by means of established risk management tools.

This article is structured as follows: the second section describes the essence of LL processes in project-based organisations. The third section presents the design of the proposed LL process. The fourth section contains the demonstration of the LL process in a practical case study. The fifth section provides a final discussion of the contributions, implications and limitations of the proposed approach.

### Lessons Learned Processes in Project Environments

Project-based organisations are characterised by the fact that essential business functions of a company are not actually carried out in a rigid functional organisation, but rather executed in unique, temporary and interdisciplinary projects (see Middleton, 1967). In such a constantly evolving project environment, LLs play a central role in organisational learning (Disterer, 2002; Jugdev, 2012; Kotnour, 2000; Schindler & Eppler, 2003; Williams, 2007). To be able to understand LL processes in project-based organisations, it is helpful first to understand the fundamental nature of a LL as such. A common definition of LLs was formulated by the American, European, and Japanese space agencies (Secchi, Ciaschi, & Spence, 1999, as cited in Weber et al., 2001):

A lesson learned is a knowledge or understanding gained by experience. The experience may be positive, as in a successful test or mission, or negative, as in a mishap or failure. Successes are also considered sources of lessons learned. A lesson must be significant

in that it has a real or assumed impact on operations; valid in that it is factually and technically correct; and applicable in that it identifies a specific design, process, or decision that reduces or eliminates the potential for failures and mishaps, or reinforces a positive result.

Two central understandings of LLs can be derived from this definition. First, LLs deal with valid, correctly represented experiences (positive as well as negative) that have a significant influence on existing business operations. Consequently, LLs must be assessed in project management practice for validity, correctness and significance. Second, LLs are designed for consistent reuse in future endeavours. Therefore, LLs must be immediately usable and applicable within their own specific practical environment. This requires first a systematic identification of relevant LLs and then also a consistent implementation in their respective business functions. In general, these tasks are performed in the discipline of project knowledge management (see Handzic & Bassi, 2017; Hanisch, Lindner, Mueller, & Wald, 2009; Oun, Blackburn, Olson, & Blessner, 2017). Corresponding knowledge management processes support the project managers during the identification and creation of relevant knowledge, the transfer and sharing of knowledge and, finally, the acquisition and implementation of helpful knowledge in the planning of new projects (Gasik, 2011; Navimipour & Charband, 2016).

To ensure consistent reuse of LLs in accordance with the understanding presented here, formalised, i.e. standardized, processes are necessary for the purposes of project knowledge management (see Fosshage, 2013; Herbst, 2017; Weber et al., 2001; Wellman, 2007; Williams, 2007). A comprehensive definition of such a LL process is provided by Weber, Aha, Muñoz-Ávila, & Breslow (2010, p. 39): 'Lessons learned (LL) processes [...] are knowledge management (KM) solutions for sharing and reusing knowledge gained through experience (i.e., lessons) among an organisation's members.' A more specific definition is given by Fosshage et al. (2016, p. 25): 'A lessons learned process describes the tools and practices whereby information about experiences (lessons or good practices) is collected, verified, stored, disseminated, retrieved for reuse, and assessed for its ability to positively affect organisational goals.' The tools and practices used in such a process can take various forms (see Weber et al., 2010), with differences being recognised in principle between organisational components (e.g., responsibilities, work instructions, organisational culture) and technological components (e.g., information and communications technology, knowledge databases, intelligent expert systems). Fosshage et al. (2016) describe a LL process at Sandia National Laboratories, where an electronic LL database forms the basis of the LL process. In this database, LL documentation is prepared from the LL owner and is saved according to a prede-

financed taxonomy. An organisational control board checks the corresponding entries and ensures that the LLs meet a certain standard of quality and then distributes them within the organisation. Project team members and special LL analysts should then access the stored documents, search for and interpret the relevant LLs and, if valuable, direct them into the planning of new projects. Weber et al. (2010), as another example, propose an intelligent lessons learned process that facilitates LL reuse through a specially developed representation that highlights reuse conditions for decision-making tasks.

The typical LL processes, like the example here demonstrates, usually focus on the first steps of the process, i.e. the acquisition, storage, distribution and retrieval of LL contents. The steps entailing the summary of relevant LL content (i.e., reading, interpreting, and synthesizing comparable findings into meaningful LL), as well as their assessment and final implementation into project planning practices, are mainly the responsibility of the project manager and are left up to his or her individual discretion. The approach described in this paper addresses these steps by first proposing a solution for the computer-aided, semi-automatic extraction of LLs from large collections of LL documentation, and then recommending a tool set for subsequent assessment and implementation of the extracted LLs.

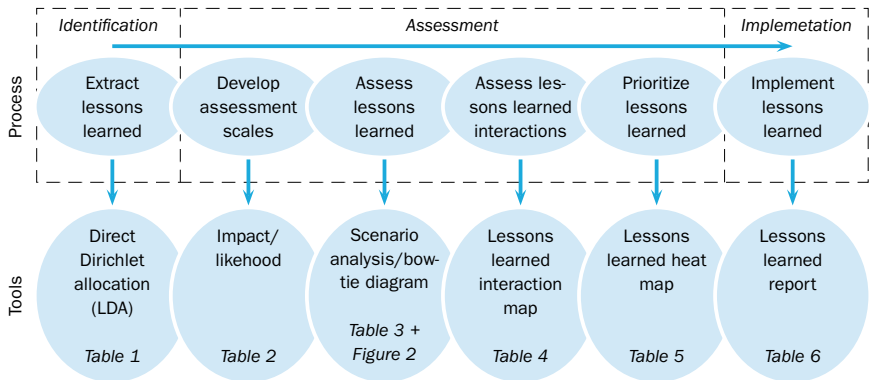
## Methodology

### Lessons Learned Process

The goal of this article is to contribute to the further development of dealing with LLs in project-based organisations. A solution will be presented for the semi-automatic extraction of LLs and their subsequent assessment and implementation. Other typical tasks in LL processes, such as the steps of codifying and storing the LLs in databases, are not being considered here.

The conceptual framework for the proposed LL process is provided by the established COSO (Committee of Sponsoring Organizations of the Treadway Commission) framework for enterprise risk management (see Moeller, 2011), which also has parallels to typical project risk management cycles (see, e.g., Chapman & Ward, 2003) but is less complex and therefore more suitable for the purposes of pragmatic LL handling. A central assumption is that LLs (i.e., positive or negative experiences) can basically be handled the same as risks (or opportunities). The process (including the tools and methods used) is illustrated in Figure 1. It includes the following sub steps:

- Semi-automatic extraction of LLs from a collection of textual project documentation using Latent Dirichlet Allocation (LDA).
- Assessment of LLs (assessment scale development, LL assessment, LL interaction assessment, LL prioritization) using risk management



**Figure 1** Process for LL Identification, Assessment, and Implementation

tools (i.e., scenario analysis, bow-tie-diagram, LL interaction map, LL heat map).

- Implementation of LLs into the project design using a LL report.

The application and practicability of the proposed approach is illustrated and evaluated below in a practical case study. The case study demonstrates the LL process in the context of an e-business project, which includes the conception, development and implementation of an online shop. The goal is to screen a collection of LL documentation for relevant LLs and then to assess and incorporate them into the conception phase of the project.

### Latent Dirichlet Allocation

A key contribution of this article is to demonstrate a semi-automatic solution for the extraction of LLs from large textual databases. In order to efficiently handle the large volume of documentation that can be found in normal project-based organisations, computer-aided text analysis is indispensable (Al Qady & Kandil, 2013; Choudhary et al., 2009; Menon et al., 2005). Against this backdrop, this article follows the idea of incorporating text mining approaches into the LL process. ‘Text mining’ can be understood as an umbrella term for a whole range of techniques for the computer-aided discovery of useful knowledge in texts (see Fan, Wallace, Rich, & Zhang, 2006). In this context, the Latent Dirichlet Allocation (LDA) is a specific text mining technique that allows the exploratory summarisation of large textual databases (see Blei, Ng, & Jordan, 2003; Blei, 2012). The LDA algorithm applies statistical analysis to extract the probability of co-occurring word patterns (i.e., correlating words regularly occurring together), which can be interpreted as latent topics hidden in the underlying corpus of documents. The underlying assumption of this probabilistic topic modelling

method is that each document is made up of a particular composition of a few topics and that a characteristic set of specific words (vocabulary) can be assigned to each of these topics. Based on this assumption, it is the statistical task of the LDA algorithm to estimate the probability distribution with which certain topics, each with specific word bundles, are to be found in the underlying collection of documents (0% = the discussion of a topic is missing from a document entirely; 100% = a document discusses only one specific topic exclusively). Using this method, such topics can be used for the exploratory summarisation of large textual databases by identifying the statistically most probable co-occurring words and interpreting these word bundles in combination. A simplified example: The co-occurring terms 'data,' 'exchange,' 'security,' 'client,' 'employee,' 'access,' 'digital,' and 'computer' can be interpreted collectively as the topic 'data security.' Likewise, each document is assumed to be composed by a specific set of such topics and, therefore, specific documents with the thematic content of 'data security' can be grouped together.

LDA is particularly suitable for the purposes of this article, because this exploratory technique is able to solve the problem of information overload in project-based organisations by mostly automating the knowledge discovery and summarisation of textual project documentation, which is very difficult or impossible to deal with manually. The practicability of this technique for extracting semantically meaningful topics from a large corpus of texts has already been proven in several studies (see, e.g., Chang, Boyd-Graber, Gerish, Wang, & Blei, 2009). In addition to the demonstration of a LDA study in the following case study, the comprehensive tutorial of this technique in Debortoli, Junglas, Müller, and vom Brocke (2016) can be recommended as a supplement.

## Demonstration

### Lessons Learned Extraction

In the first step of the LL process presented here, meaningful LLs should be extracted from a larger text database using LDA. The typical implementation of the LDA technique involves the following steps (see also Debortoli et al., 2016): (1) compiling and preparing the textual database; (2) defining the number of topics to be extracted; (3) preparing the text data in the context of pre-processing; and (4) interpreting and titling the extracted topics. In the following, the individual steps are described in more detail.

1. *Compilation and preparation of the textual database.* In total, a collection of 68 post-project reports from previous e-business projects (i.e., conception, development, and implementation of online shops) was available for this case study. The authenticity, credibility, represen-

tativeness and meaningfulness of the documentation was checked prior to being used. Since the results of the LDA essentially depend on the contents of the database to be summarised, the database must be thematically specified, which ensures the exclusive extraction of LLs and thus also guarantees the significance and validity of the results. For the purposes of this analysis, only the LL sections of the documents were exclusively selected (i.e., the discussion of positive and negative experiences; each between 0.25–1.50 pages long). Other contents contained in the project reports (e.g., general project descriptions, cost statements or technical details) were excluded from the analysis. To put the text data into a format more efficiently analysable by a computer, the text content was transferred to a relational database consisting of the document number and LL description.

2. *Definition of the number of topics to be extracted.* The database was then read into a text mining tool and prepared for analysis (the cloud-based tool MineMyText – [minemytext.com](http://minemytext.com) – was used for the LDA application). In the first step, an adequate number of topics to be extracted must be determined. One recommendation here is to evaluate different numbers of topics in iterative test runs (see Debortoli et al., 2016). To evaluate the significance of the topics generated during the test runs, the general recommendation is to assess the individual topics on whether they provide a meaningful statement that is useful for the issue to be analysed and that is also coherent. Based on the available database, a variant with 13 topics to be extracted gave the most promising results. Relevant topics would be lost in a variant with fewer topics, and a variant with more topics would not lead to meaningful topics or would lead to duplicates. The corresponding evaluations were carried out by two analysts (one of the authors of this article and a project manager). Differences were discussed, and final decisions were agreed upon by consensus.
3. *Preparation of text data in the context of pre-processing.* In the next step, the database must be processed, i.e. cleaned, to ensure proper accuracy of the results. The goal of this pre-processing is to remove as much ‘statistical noise’ from the text data as possible. This process includes several steps. A tokenising process is carried out first. In this process, the database of the text descriptions is disassembled into separate, individual words, and these are assigned to the corresponding documents. Next, general stop words were removed from the database, i.e. words that do not carry meaning and thus do not contribute anything to interpretable topics (e.g., ‘and,’ ‘or,’ ‘this’). The



analysts also defined special stop words that would not provide any added value to the analysis (e.g., the words 'lesson' or 'learned'). The technique of lemmatisation was also used to reduce to a basic form different variants of a word that were nonetheless identical content-wise (e.g., 'plans,' 'planning,' or 'planned' to their dictionary form 'plan'). This process reduces the complexity of the database for the subsequent statistical calculations.

4. *Interpretation and titling of the extracted topics.* Thirteen topics were extracted using the LDA algorithm, each topic consisting of bundles of most likely co-occurring words (see Table 1). The combined interpretation of these commonly associated word bundles makes it possible to title concrete LL issues. This interpretation was again carried out independently by two analysts and appropriate titles were then agreed upon by consensus. A specific example of this kind of interpretation and titling: Topic #4 contains the correlated words 'supplier,' 'order,' 'procurement,' 'online,' 'purchase,' 'company,' 'eprocurement,' and 'process,' which in combination can be interpreted and titled as the topic 'e-procurement.' Because this subject was extracted from a database that consists exclusively of text discussions of LLs of earlier e-business projects, this means that the importance or difficulty of e-procurement regularly seems to play a central role in such projects, and consideration of this fact should therefore be significant and applicable in comparable future endeavours. According to the definition of a LL in the second section, the necessary requirements in this example are thus fulfilled (= a lesson should be significant, valid, technically correct, and applicable). Other LLs extracted from the database include, for example, topics on system integration (topic #3), data security (topic #5), and supplier collaboration (topic #12). The last column contains the probabilities of the topics (the higher the probability, the more the documentation discusses these facts), which may be helpful in the next phase of the LL assessment, when these LL facts are specified, reflected, and prioritised.

### **Assessment Scale Development**

To be able to assess and compare the extracted LLs consistently, a form of quantitative assessment is needed. Without the appropriate measurement scales, LLs can only be assessed and prioritised according to qualitative criteria and, finally, according to the subjective judgement of the individual project manager.

Typical assessment scales used in risk management include 'Impact' and 'Likelihood' (Moeller, 2011). 'Impact' represents the potential conse-

**Table 1** Extracted LL from the LDA

#	Topic Title	Keywords (excerpt)	Probability
1	Project scope & planning	Solution, system, implementation, project, process, cost, time, change, customer, data, key, business, result	3.1%
2	Customer demand management	Customer, solution, system, offer, service, order, range, option, delivery, business, case	8.0%
3	System integration	Solution, system, software, integration, online, shop, eshop, erp, catalogue, product, launch, standard	2.9%
4	E-procurement	Supplier, order, eprocurement, purchase, online, company, procurement, process, support	7.7%
5	Data security	Client, access, exchange, security, employee, knowledge, digital, obtain, data, computer, stock, product	2.9%
6	Business chain & logistics	Business, chain, creation, supply, process, logistics, company, product, tool	7.5%
7	Customer relationship management	Service, crm, mobile, communication, application, experience, exchange, support, data, time	4.4%
8	Digitalisation potential	Process, cost, saving, document, electronic, interface, reduce, potential, supplier, invoice	1.3%
9	E-business strategy alignment	Company, internet, ebusiness, case, technology, trade, market, position, success, management, industry, cost, sale	25.2%
10	Software and platform usability	Employee, launch, function, user, software, platform, connect, requirement, advantage, benefit, clear, simple, difficult	3.9%
11	Electronic payment	Ebusiness, introduction, payment, card, online, pay, activity, distribution, transfer, channel, cost, impact	4.3%
12	Supplier collaboration	supplier, collaboration, erp, portal, trust, full, communication, define, workflow, advantage	11.8%
13	Stakeholder management & communication	project, partner, stakeholder, success, factor, process, network, involve, information, management, team, party, important	5.5%

quences of an identified LL not being implemented (i.e. the non-occurrence of a positive effect) or of its occurrence not being prevented (e.g., the recurrence of a problem). ‘Likelihood’ represents the probability that the consequences of a LL can be expected. A scale of 1–5 provides the basis for quantitatively comparative assessments. Sample scales are shown in Table 2. The assessment variables (qualitative/quantitative) for the concrete definition of the scale levels can vary depending on the company and project, and must be defined individually. An impact level of 5 would be reached in this case study if the targets of the project (e.g., time, quality or outcome) were fatally exceeded and the project was certain to be can-

**Table 2** Assessment Scales

	Rating	Descriptor	Definition (examples)
Impact scale	5	Extreme	Project targets cannot be achieved (> 30%); certain project cancellation; game-changing loss of market share; extreme impact on stakeholders; multiple senior leaders leave and high turnover.
	4	Major	Major deviation from targets (20% up to 30%); possible project cancellation; major loss of market share; significant impact on stakeholders; key employees leave and high turnover.
	3	Moderate	Moderate deviation from targets (15 up to 20%); project delay; moderate loss of market share; moderate impact stakeholder; widespread staff morale problems and increase in turnover.
	2	Minor	Minor deviation from targets (10% up to 15%); minor loss of market share; minor impact on stakeholders; minor staff morale problems.
	1	Incidental	Calculated deviation from targets (5% up to 10%); no relevant loss of market share; no or minor impact on stakeholders; isolated staff dissatisfaction.
Likelihood scale	5	Almost certain	Almost 90% or greater likelihood of certain occurrence in a project
	4	Likely	Likely 65% up to 90% likelihood of occurrence in a project
	3	Possible	Possible 35% up to 65% likelihood of occurrence in a project
	2	Unlikely	Unlikely 10% up to 35% likelihood of occurrence in a project
	1	Rare	Rare <10% likelihood in a project

celled. In this context, a ‘Likelihood’ of 5 describes a high probability of 90% or more for the occurrence of this impact.

**Lessons Learned Assessment**

In risk management, one of the main tasks is assessing potential risk impacts in the context of scenario analysis (Raz & Michael, 2001). The same understanding should also apply to LLs. In the present case study, the 13 extracted LLs were independently assessed and discussed along the ‘Impact’ and ‘Likelihood’ scales by two experts (see Table 3). The assumptions from the respective scenarios were based on the evaluation of past values from comparable projects and estimates. The probability distributions of the extracted topics also supported the assessment of the distribution of the individual LL facts. Furthermore, a closer examination of individual project reports assigned to the different LLs with a high probability supported this assessment as well. For example, LL #3 stands for the fact that the integration of the various systems into the online shop platform (e.g., content

**Table 3** Scenario Analysis

#	Scenario Description	Detailed Assumptions	IR	LR
1	Project scope (time, cost, capacity, ...) is underestimated and not met	Deviation from time schedule (up to 20%) Cost overruns (up to 15%)	2.5	2.0
2	Customer demand (offer, product range, ...) and needs (services, ...) are not met	Potential decrease in market share of 10% Lack of customer orientation and damage in reputation	4.5	3.0
3	System integration difficulties and disruptions	Extended implementation and testing phase (+2 weeks) Need for experts n and increase in capacity (+8.500\$)	2.0	1.5
4	e-Procurement process inconsistencies and disruptions	Delivery problems and disruptions Loss of sales due to delivery problems (-x\$)	3.5	2.0
5	Data security breaches	Reputational damage and loss of trust Loss of customers and loss in market share of 15%	4.5	2.5
6	Problems in the business chain and logistic problems	Delivery problems and disruptions Loss of sales due to delivery problems (-x\$)	3.0	2.5
7	Customer relationship management (CRM) is inappropriate	Reputational damage Loss of customers (10%) and loss in market share of 5%	2.5	1.5
8	Digitalisation potential (e.g., cost savings) is not fully exploited	Data inconsistencies and data transfer problems Manual reworking and increased process costs (+15%)	2.5	1.0
9	e-Business strategy alignment is inappropriate	Inappropriate market penetration strategy Loss of customer (20%) and loss in market share of 15%	4.0	2.0
10	Software and platform usability is inappropriate	User and employee dissatisfaction Loss in market share of 12.5 %	2.0	3.0
11	Electronic payment is disrupted	Disrupted trade and payment functions Reputational damage and loss of customers (10%)	3.5	1.5
12	Supplier collaboration is ineffective	Increased level of problem management Extended implementation and testing phase (+2 month)	2.5	4.0
13	Stakeholder management and communication is inappropriate	Dissatisfaction of stakeholders Reputational damage	1.5	2.5

**Notes** IR – impact rating, LR – likelihood rating.

management system, warehouse management system) caused difficulties and delays (Impact = 2.0; Likelihood = 1.5). The assumption here is that this circumstance is unlikely (10%–20% likelihood of occurrence). However,

if it does occur, it will involve an extended implementation and testing phase (+ 2 weeks) requiring increased capacity (+ 8.500\$). The examination of the project reports with the highest probability to contain this topic also supported this definition. Corresponding scenarios were also simulated for the other 12 LLs.

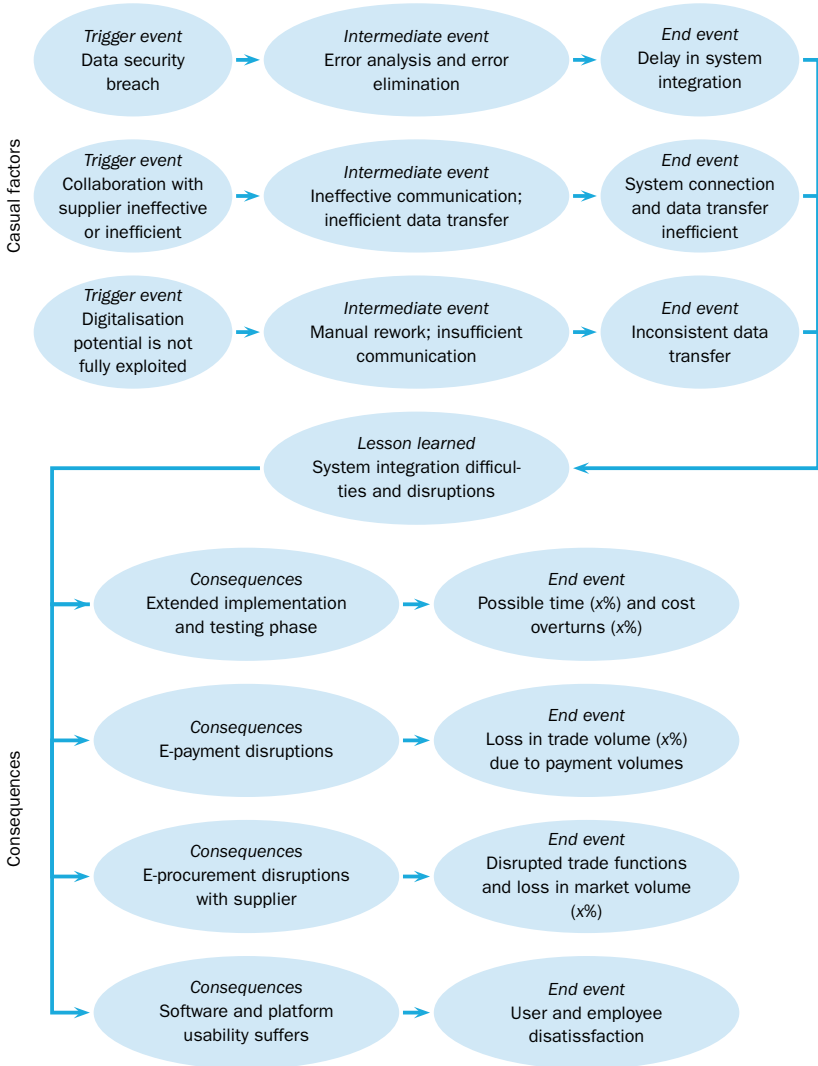
The bow-tie diagram is another potential tool that supports the analyses of the causes and consequences of LLs. The bow-tie method divides the LLs into their individual components (i.e., causal factors and consequences), making it possible to better analyse and understand the complex interrelationships of the LLs. Figure 2 demonstrates this type of analysis on LL #3 (system integration). A trigger event for the occurrence of this issue could be a data security breach, for example, which involves error analysis and error elimination (intermediate event), and ultimately leads to a delay in system integration (end event). The eventual consequences would be that required repairs delay the implementation and testing phase, and create additional costs associated with the expanded capacity.

### **Lessons Learned Interaction Assessment**

LLs usually have interrelations with other LLs. A LL interaction map (see Table 4) covers corresponding networks of relationships with other LLs by placing them opposite each other in a matrix (× indicates a relationship between two different LLs). The knowledge of such relationships, which also includes any possible dependencies or influences, can be relevant when a decision must be made as to whether a LL should be implemented or not. This assessment was performed by analysing the previously constructed bow-tie diagrams and by discussing potential relationships together with experts. For this purpose, workshops with different domain experts can potentially improve the understanding of a LL by combining various perspectives (Moeller, 2011). Table 4 demonstrates that, for example, active consideration of LL #4 (data security) has a direct (positive) effect on LL #3 (system integration). On the other hand, neglect of LL #12 (supplier collaboration) would imply a most likely negative influence on e-procurement activities (LL #4).

### **Lessons Learned Prioritisation**

A LL heat map (see Table 5) can be useful in providing an overview of the current LL portfolio and in precisely prioritising individual LLs. This prioritisation can help in the decision-making process if not all of the LLs can be implemented considering the usually scarce capacities in projects. This heat map combines the previously created assessments of 'Impact' and 'Likelihood' into a matrix and divides the LL portfolio into three visual areas with three different levels of strategic prioritisation. In the present case,



**Figure 2** Bow-Tie-Diagram

this will define three groups, with the group that has both LL #2 (customer demand and needs) and LL #5 (data security) having the highest priority for implementation in accordance with the high likelihood of impact.

**Lessons Learned Implementation**

Reports are usually recommended in the PM guidelines for the implementation of the LLs that have been identified. The PRINCE2 framework pro-

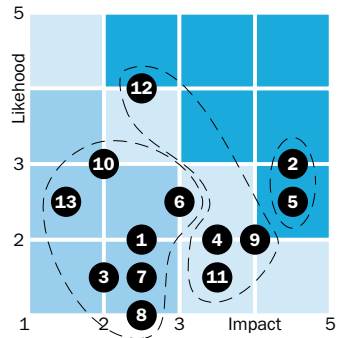
**Table 4** Lessons Learned Interaction Map

	1	2	3	4	5	6	7	8	9	10	11	12	13
1		x							x				x
2	x						x		x	x			
3				x	x	x		x		x	x	x	
4			x			x		x					
5			x					x			x		
6			x	x				x	x			x	x
7		x											x
8			x	x	x	x			x		x		
9	x	x				x		x					x
10		x	x										
11			x		x			x					
12			x	x		x							x
13	x					x	x		x			x	

**Notes** Column/row headings are as follows: (1) project scope, (2) customer demand and needs, (3) system integration, (4) eprocurement, (5) data security, (6) business chain & logistics, (7) customer relationship management (CRM), (8) digitalisation potential, (9) ebusiness strategy alignment, (10) software and platform usability, (11) electronic payment, (12) supplier collaboration, (13) stakeholder management and communication.

**Table 5** Lessons Learned Heat Map

#	Lesson learned	IR	LR
1	Project scope	2.5	2.0
2	Customer demand and needs	4.5	3.0
3	System integration	2.0	1.5
4	Eprocurement	3.5	2.0
5	Data security	4.5	2.5
6	Business chain & logistics	3.0	2.5
7	Customer rel. man. (CRM)	2.5	1.5
8	Digitalisation potential	2.5	1.0
9	Ebusiness strategy alignment	4.0	2.0
10	Software and platform usability	2.0	3.0
11	Electronic payment	3.5	1.5
12	Supplier collaboration	2.5	4.0
13	Stakeholder man. and comm.	1.5	2.5



poses, for example, LL implementation in the project planning phase using a so-called ‘lessons log,’ which collects all the LLs relevant for the current project (Office of Government Commerce, 2009). A similar approach will be pursued in this case with the LL report, which summarises essential information from the extracted LLs and contains the classification in the relevant project phase, as well as recommendations for dealing with each LL (see

**Table 6** Lessons Learned Report

#	Stage of project	Subject	IR	LR	Recommendation
1	Project conception & planning	Project scope (time, cost, capacity, ...) is underestimated and not met	2.5	2.0	Business case, project plan and requirement catalog should be regularly confirmed by project committee and product owner. LL #2 and #10 should be closely reviewed.
2	Project conception & planning	Customer demand (offer, product range, ...) and needs (services, ...) are not met	4.5	2.5	Conduct a profound market research. The concept for the online shop should be confirmed by product owner and customer representatives. Key customers should be part of the steering committee. Close alignment with LL #9 is necessary.
3	Project execution: implementation and testing	System integration difficulties and disruptions	2.0	1.5	Plan more capacity for testing of system connectivity and implementation. LL #4 and #5 should be closely reviewed.
4	Project execution: system implementation	e-Procurement process inconsistencies and disruptions	3.5	2.0	Close collaboration with supplier and logistics service provider is necessary. Conduct a project start-up workshop together with suppliers and service providers. Connections to LL #12 should be reviewed.
5	Project execution: system development	Data security breaches	4.5	2.5	Consider extended security functions. External audit of data security aspects should be considered. Review connections to LL #8 and #11.
6	Project execution: system implementation	Problems in the business chain and logistic problems	3.0	2.5	Logistic solutions should be discussed and evaluated together with suppliers and logistics service provider. LL #12 should be taken into account.

*Continued on the next page*

Table 6). The LL #3 documentation, for example, describes its relevance in the project execution phase (implementation and testing), the defined assessment scales, as well as the recommendations for dealing with this LL (including consideration of the network of relationships).

## Conclusions

'Learning [...] has to be managed together with the project and must be integrated into project management as standard practice. It has to become a natural experience with projects' (Ayas, 1996, p. 132). Following this recommendation of Ayas (1996), the goal of this article is to contribute to further developments in dealing with LLs in project-based organisations. For this



**Table 6** *Continued from the previous page*

#	Stage of project	Subject	IR	LR	Recommendation
7	Project closing: change management	Customer relationship management (CRM) is inappropriate	2.5	1.5	Plan more capacity for customer service during roll-out of the online shop. Additional marketing campaigns. LL #2 and #10 should be closely reviewed.
8	Project conception & planning	Digitalization potential (e.g., cost savings) is not fully exploited	2.5	1.0	Business processes should be evaluated in terms of lean management principles (e.g., waste analysis). Lean project team should be implemented. LL #1 should be appropriately adapted.
9	Project conception & planning	e-Business strategy alignment is inappropriate	4.0	2.0	Carry out a strategy workshop with all necessary stakeholders, guided and moderated by the external consultants. LL #1 and #2 should be adapted accordingly.
10	Project execution: system development	Software and platform usability is inappropriate	2.0	3.0	Include employees and customers into the design phase. Continuous testing of the functionalities through customer representatives. LL #1 should be closely reviewed.
11	Project Execution: system implementation	Electronic payment is disrupted	3.5	1.5	Hire technical experts for designing and testing payment solutions. External audit of data security aspects is necessary. LL #8 and #5 should be closely reviewed.
12	Project planning & execution	Supplier collaboration is ineffective	2.5	4.0	Conduct a project start-up workshop together with suppliers and logistics service provider. A key supplier should be part of the steering committee. LL #4 and #6 are affected.
13	Project planning & execution	Stakeholder man. and communication is inappropriate	1.5	2.5	Stakeholder analysis should be conducted and an appropriate communication established. LL #2, #10 and #12 should be closely reviewed.

purpose, a LL process (including a tool set) is recommended for the semi-automatic extraction of LLs from large collections of project documentation and their formal assessment and implementation in the conception of new projects. Such formalised LL processes are necessary in project-based organisations so that, as expressed by Williams (2007), ‘lessons observed’ can become ‘lessons learned.’

The proposed LL process provides a contribution to project knowledge management practice first by proposing a solution for the computer-assisted exploration of large textual document collections and for efficient extrac-

tion of LLs with the application of LDA. This solution provides the project manager with a tool for overcoming information overload in project-based organisations and the associated time-consuming research in text-based knowledge archives. Another contribution is made with the proposal of a systematic procedure for the assessment and implementation of those LLs that were previously extracted. In practice, this kind of formalised process can help reduce the individual discretion of the project manager and instead promote a standardised evaluation and consistent implementation of relevant LLs. It should be noted here that the individual sub-steps and the application of the individual tools are not sequences that must be strictly adhered to. Individual process steps can be varied as needed or omitted depending on the project or capacity. For example, assessment of the LLs could also be limited to scenario analysis and LL reporting.

The research is provided with development paths. Firstly, a text mining technique is demonstrated for handling extensive text-based collections of knowledge in project management. Computer-aided text analysis reveals generally interesting potentials for (project) knowledge management (see, e.g., Choudhary et al., 2009; King, 2009). Future research efforts should focus on the evaluation of these techniques in project management. Secondly, a further contribution is provided by transferring and applying the principles and methods of (project) risk management to the handling of LLs. The basic assumption here is that LLs represent positive experiences (opportunities) and negative experiences (risks), which also require formalised handling similar to that of project risks. The application of the broad range of risk management procedures and tools (see Raz & Michael, 2001) could drive consistent confrontation with LLs and their implementation and inspire the development of future LL processes.

The proposed approach is of course not without limitations. In particular, the methodology of computer-assisted extraction of LLs from inventories of text documents requires further in-depth evaluations. In this context, aspects of validity, objectivity and reliability must be examined. The following influencing factors to be examined should also be listed here. On the one hand, the influence of the nature and extent of the textual database, consisting of project documentation texts, should be evaluated. A different composition or amount of LL documentation could easily produce different results. On the other hand, the effect of subjective influences in the interpretation of the topics extracted by LDA is a limitation. Although the results were interpreted and discussed by two experts, a subjective bias in the definition of the LL cannot be ruled out, which could particularly affect the objectivity and reliability of the results. For text mining studies used for research purposes, there is a series of even deeper-reaching evaluation methods (see Debortoli et al., 2016). However, these are too comprehen-

sive for the practical purposes of efficient project management and would, therefore, not be used in the context of this practical case study.

## References

- Al Qady, M., & Kandil, A. (2013). Document management in construction: Practices and opinions. *Journal of Construction Engineering and Management*, 139(10). Retrieved from [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000741](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000741)
- Almeida, M. V., & Soares, A. L. (2014). Knowledge sharing in project-based organizations: Overcoming the informational limbo. *International Journal of Information Management*, 34(6), 770–779.
- Ayas, K. (1996). Professional project management: A shift towards learning and a knowledge creating structure. *International Journal of Project Management*, 14(3), 131–136.
- Barclay, C., & Osei-Bryson, K.-M. (2010, 12–15 August). *An exploration of knowledge management practices in IT projects: A case study approach*. Paper presented at the Americas Conference on Information Systems, Lima, Peru.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine learning research*, 3(January), 993–1022.
- Blei, D. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84.
- Boh, W. F. (2007). Mechanisms for sharing knowledge in project-based organizations. *Information and Organization*, 17(1), 27–58.
- Carrillo, P., Ruikar, K., & Fuller, P. (2013). When will we learn? Improving lessons learned practice in construction. *International Journal of Project Management*, 31(4), 567–578.
- Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. (2009, 6–8 December). *Reading tea leaves: How humans interpret topic models*. Paper presented at the Neural Information Processing Systems Conference, Vancouver, Canada.
- Chapman, C., & Ward, S. (2003). *Project risk management: Processes, techniques, and insights*. Chichester, England: Wiley.
- Choudhary, A. K., Oluikpe, P. I., Harding, J. A., & Carrillo, P. M. (2009). The needs and benefits of text mining applications on post-project reviews. *Computers in Industry*, 60(9), 728–740.
- Debortoli, S., Junglas, I., Müller, O., & vom Brocke, J. (2016). Text mining for information systems researchers: An annotated topic modeling tutorial. *Communications of the Association for Information Systems*, 39(7), 110–135.
- Disterer, G. (2002). Management of project knowledge and experiences. *Journal of Knowledge Management*, 6(5), 512–520.
- Duffield, S., & Whitty, S. J. (2015). Developing a systemic lessons learned knowledge model for organisational learning through projects. *International Journal of Project Management*, 33(2), 311–324.

- Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Communications of the ACM*, 49(9), 76–82.
- Fosshage, E. (2013). *Considerations for implementing an organizational lessons learned process* (Sandia Report SAND2013-3671). Sandia National Laboratories (SNL-NM), Albuquerque, NM.
- Fosshage, E. D., Drewien, C. A., Eras, K., Hartwig, R., Post, D. S., & Stoecker, N. K. (2016). *Implementing a lessons learned process at Sandia National Laboratories* (Sandia Report SAND2016-0275). Sandia National Laboratories (SNL-NM), Albuquerque, NM.
- Gasik, S. (2011). A model of project knowledge management. *Project Management Journal*, 42(3), 23–44.
- Haksever, A. M. (2001, 2–6 April). *A model to predict the occurrence of information overload of project*. Paper presented at the 15th CIB World Building Congress, Wellington, New Zealand.
- Handzic, M., & Bassi, A. (2017). *Knowledge and project management: A shared approach to improve performance* (5th ed.). New York, NY: Springer.
- Hanisch, B., Lindner, F., Mueller, A., & Wald, A. (2009). Knowledge management in project environments. *Journal of Knowledge Management*, 13(4), 148–160.
- Herbst, A. S. (2017). Capturing knowledge from lessons learned at the work package level in project engineering teams. *Journal of Knowledge Management*, forthcoming.
- Hong, H. K., Kim, J. S., Kim, T., & Leem, B. H., 2008. The effect of knowledge on system integration project performance. *Industrial Management & Data Systems*, 108(3), 385–404.
- Jugdev, K. (2012). Learning from lessons learned: Project management research program. *American Journal of Economics and Business Administration*, 4(1), 13.
- Keegan, A., & Turner, J. R. (2001). Quantity versus quality in project-based learning practices. *Management Learning*, 32(1), 77–98.
- King, W. R. (2009). Text analytics: Boon to knowledge management? *Information Systems Management*, 26(1), 87–93.
- Kotnour, T. (2000). Organizational learning practices in the project management environment. *International Journal of Quality & Reliability Management*, 17(4/5), 393–406.
- Kotnour, T. G., & Kurstedt, H. A. (2000). Understanding the lessons-learned process. *International Journal of Cognitive Ergonomics*, 4(4), 311–330.
- Kululanga, G. K., & Kuotcha, W. S. (2008). Measuring organisational learning through project reviews. *Engineering, Construction and Architectural Management*, 15(6), 580–595.
- Menon, R., Tong, L. H., & Sathiyakeerthi, S. (2005). Analyzing textual databases using data mining to enable fast product development processes. *Reliability Engineering & System Safety*, 88(2), 171–180.
- Middleton, C. J. (1967). How to set up a project organization. *Harvard Business Review*, 45(2), 73–82.

- Moeller, R. R. (2011). *COSO Enterprise Risk Management: Establishing Effective Governance, Risk, and Compliance Processes* (2nd ed.). Chichester, England: Wiley.
- Navimipour, N. J., & Charband, Y. (2016). Knowledge sharing mechanisms and techniques in project teams: Literature review, classification, and current trends. *Computers in Human Behavior*, 62, 730–742.
- Newell, S., Bresnen, M., Edelman, L., Scarbrough, H., & Swan, J. (2006). Sharing knowledge across projects: Limits to ICT-led project review practices. *Management Learning*, 37(2), 167–185.
- Office of Government Commerce. (2009). *Managing and directing successful projects with PRINCE2TM* (5th ed.). London, NY: Office of Government Commerce. Retrieved from <http://www.bpug.nl/downloads/finish/30-1-5-2010-prince2/1318-brochure-ogc>
- Oun, T. A., Blackburn, T. D., Olson, B. A., & Blessner, P. (2016). An enterprise-wide knowledge management approach to project management. *Engineering Management Journal*, 28(3), 179–192.
- Parnell, J. A., Von Bergen, C. W., & Soper, B. (2005). Profiting from past triumphs and failures: Harnessing history for future success. *SAM Advanced Management Journal*, 70(2), 36–47.
- Prencipe, A., & Tell, F. (2001). Inter-project learning: Processes and outcomes of knowledge codification in project-based firms. *Research policy*, 30(9), 1373–1394.
- Project Management Institute. (2013). *A guide to the project management body of knowledge (PMBOK® Guide)* (5th ed.). Newtown Square, PA: Project Management Institute.
- Raz, T., & Michael, E. (2001). Use and benefits of tools for project risk management. *International Journal of Project Management*, 19(1), 9–17.
- Schindler, M., & Eppler, M. J. (2003). Harvesting project knowledge: A review of project learning methods and success factors. *International Journal of Project Management*, 21(3), 219–228.
- Secchi, P., Ciaschi, R., & Spence, D. (1999). A concept for an ESA lessons learned system. In P. Secchiu (Ed.), *Proceedings of alerts and LL: An Effective way to prevent failures and problems* (Tech. Rep. WPP-167) (pp. 57–61). Noordwijk, The Netherlands: ESTEC.
- von Zedtwitz, M. (2002). Organizational learning through post-project reviews in R&D. *R&D Management*, 32(3), 255–268.
- Weber, R., Aha, D. W., & Becerra-Fernandez, I. (2000). Categorizing intelligent lessons learned systems. In D. W. Aha & R. Weber, *Intelligent lessons learned systems: Papers from the AAAI 2000 workshop* (pp. 63–67). Menlo Park, CA: AAAI Press.
- Weber, R., Aha, D. W., & Becerra-Fernandez, I. (2001). Intelligent lessons learned systems. *Expert Systems with Applications*, 20(1), 17–34.
- Weber, R., Aha, D., Muñoz-Ávila, H., & Breslow, L. (2010). An intelligent lessons learned process. In *Foundations of Intelligent Systems* (pp. 39–41). Berlin, Germany: Springer.

Wellman, J. (2007). Lessons learned about lessons learned. *Organizational Development Journal*, 25(3), 65–72.

Williams, T. (2007). *Post-project reviews to gain effective lessons learned*. Newtown Square, PA: Project Management Institute.

**Benjamin Matthies** Benjamin Matthies is a research assistant at the South Westphalia University of Applied Sciences, Germany. Before joining the team, he worked as business analyst for a multinational retail company and was involved in a variety of IT and business projects. His research interests include project and knowledge management, management accounting, and business analytics. Within the field of knowledge management, he particularly focuses on the application of text mining techniques for organisational learning purposes. *matthies.benjamin@fh-swf.de*



This paper is published under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).