DESIGN
2018

CHANGES AND SENTIMENT: A LONGITUDINAL EMAIL ANALYSIS OF A LARGE DESIGN PROJECT

S. A. Piccolo, J. Wilberg, U. Lindemann and A. Maier

Abstract

Changes are part of any project. Although previous research provides methods to deal with changes, understanding of changes in relation to sentiment is still unclear. This is important as people's mood can affect performance and decisions. We implement an approach to quantify "change language" in emails and study its relation to sentiment. We find that sentiment decreases when problems or changes emerge, and increases when changes are implemented successfully. We discuss the implications of our findings for research and project engineering practice, providing avenues for further work.

Keywords: design process, process analysis, data mining, changes, sentiment

1. Introduction and literature background

Understanding and supporting engineering design processes is critical to creating good designs with market and operational success. However, design processes can be challenging to handle and to execute, and there is a need for practical approaches for design process understanding and management - practical approaches for data-driven design process diagnostics and monitoring.

Design processes are inherently iterative (Wynn and Eckert, 2016) and changes, whether intended or emergent, are a part of every design and development project (Fricke et al., 2000; Eckert et al., 2004; Jarratt et al., 2011) in a number of industry sectors, including construction (Wilberg et al., 2015). We characterise change as a subset of iterations as all changes involve iterations, yet not all iterations involve changes or rework; and whilst having the potential for leading to product improvements in particular or system innovations more generally, changes and rework may negatively affect the design process. Changes have been shown to propagate (Clarkson et al., 2004), increasing process completion time and, in extreme cases, leading to design churns (Yassine et al., 2003) or even to a lack of convergence (Braha and Bar-Yam, 2007). Changes and rework have been reported as potentially very costly and taxing for an engineering organisation, both financially (Jarratt et al., 2011; Langer et al., 2012) and emotionally (Pirola-Merlo et al., 2002; Love, 2002; Love and Edwards, 2004; Love et al., 2005).

In summary, changes and rework do not only affect the artefact, it has been shown that rework and changes are negatively correlated to people's mood, which in turn affects decision-making and performance (Pirola-Merlo et al., 2002)., i.e. the execution and handling of the design process. Although many methods to handle or anticipate changes exist (Clarkson et al., 2004; Hamraz et al., 2013), methods to learn from previous changes and to monitor the design process in real time are still scarce.

In order to learn from the past and monitor the health status of a process in real time, emails exchanged during a project are a good data source. Wasiak et al. (2010, 2011) used emails to perform a qualitative content analysis, classifying emails for content and purpose, showing that such an approach can highlight project phases and problems. Snider et al. (2017) analysed emails to develop a quantitative technique to monitor the design process through the emails' topics. They derived patterns and

characteristics of communication that are associated with issues, focused working, and background discussion.

What has so far still been missing in literature, is the understanding of people's sentiment toward changes. By analysing the emails exchanged during a large engineering project of designing a biomass power plant, we join the above conversation to shed light on the relation between changes and sentiment in the design process. In other words, this paper focuses on the relation between changes and sentiment. We use sentiment analysis and a custom approach for detecting "change language" in emails, plus a topic analysis over time to provide context and to evaluate the unfolding of the process (see §2.4). As a result, we quantify sentiment and change scores over time and analyse the trends they describe. The behaviour of sentiment and change scores maps onto process stages and highlights the emergence of problems in the design process. The succession of topics over time describes the evolution of the process and can be used to reason about the order followed to complete the design activities; thus, providing content-based ecological validity to the numerical scores (see §3). By reading and re-phrasing the content of the emails, we provide a narrative account or story of the design process under investigation that supports the trends in changes and sentiments detected and characterised by the scores we computed. Thus, we advance knowledge on the relation between changes and sentiment, by showing that the project's sentiment appears to overall decrease when changes emerge or when their impact increases; at the same time, by quantitatively showing that sentiment increases while changes are successful and lead to project progression, we provide empirical evidence that changes are not always negative. Furthermore, we qualitatively show that sentiment and changes are closely related and influence each other (see §4). These results support the practice of managing engineering projects to reason on the importance of sentiment in connection with change and to develop appropriate strategies successfully to handle changes. Finally, the practical approach developed in this paper and used to analyse ~54000 emails from a large engineering project, shows that is possible to investigate and understand past changes from email text. In fact, we are able to detect periods with impactful changes and detect the associated topics, thus evidencing the suitability of our approach for design process profiling and process diagnostics. This can be leveraged, in future research, as a basis for defining and implementing process diagnostics or "health status" metrics to monitor a project in real time or to understand and learn from previous ones.

2. Methods

To gain understanding of the evolution of the design process over time with respect to sentiments and changes, we aim to detect periods in which changes occurred using real-world project data. To detect the occurrence of people using words related to change in their emails, we start out with topic analysis over time (see Figure 1). Following on, we need some way of finding out if an email relates to changes or not. For that, we rely on sentiment analysis and develop a custom approach to score the presence of "change language"; that is, emails containing words related to changes. We use both sentiment analysis and "change language" analysis to assign a numerical score to each email. We plot the normalised z-scores over time to detect potential relations between changes and sentiment (see Figure 2). Upon visual inspection of the change and sentiment curves, we can form expectations concerning the relation of change and sentiment in the email text (see §3.2). Guided by the relative topics, we then proceed to read the emails associated to the periods of interest to provide context and ecological validation to our expectations.

2.1. E-mail data

Investigating changes and sentiment together, we analyse email exchanges from a Danish organisation involved in the design of power plants. Our data refer to the design stage of a large project (£160,000,000+) of a renewable energy power plant for electrical energy generation (Parraguez et al., 2015). The dataset contains ~54000 emails spanning a total of five years and covering all the stages of the design process and the erection of the power plant. The dataset covers intra-organisational communication (i.e. communication between the company's employees) and inter-organisational communication (i.e. communication between the company's employees and employees from other companies). The company that designed this power plant is Danish and we found in an almost equal

proportion emails in both Danish and English. The email exchanges involve more than 2300 unique email addresses, ~170 of which are associated with the company's employees.

2.2. Topic analysis over time

To understand the topics of conversation over time, we use the burst detection algorithm developed by Kleinberg (2003). This algorithm models a stream of text using an infinite-state automaton, in which bursts appear as state transitions. In this way, the algorithm can detect the appearance and disappearance of a word in a given period (a "spike"). The analysis of the keywords and relative periods generated by the algorithm reveals the evolution of the conversation over time. The keywords emitted by the algorithm are different from the words used most often. Indeed, the words used most often and with similar frequency over time - and are thus not bursty - are articles, pronouns, propositions, and the like. Kleinberg showed the goodness of this approach by analysing a set of paper titles from two conferences on databases and two conferences on computing theory, revealing the evolution of the principal topics over several decades in the conferences.

2.3. Sentiment analysis

Sentiment analysis aims to understand the mood of a text. Many approaches exist (for a review see Ravi and Ravi, 2015). A common approach is to use a lexicon of words annotated with the sentiment polarity (usually from -5 to +5) from the most negative to the most positive. The text is analysed matching each word in the text with the words in the lexicon and the sentiment score is then a combination of the scores for each word (e.g. the sum or the mean). As our dataset includes emails in both English and Danish, we use the AFINN sentiment analyser (Nielsen, 2011), a lexicon based approach with a wordlist developed to detect the sentiment as a numerical score (on a scale from -5 to +5) of Danish as well as English texts. The final sentiment score for each email is computed as the average of the scores associated with each word.

2.4. Change language analysis

As we are explicitly looking for changes, we complement the sentiment analysis using a similar dictionary-based approach to score emails by the amount of "change language" they contain. To make this possible, we need to create a custom wordlist containing words that are likely to appear if the email is related to changes. We created an initial wordlist analysing the bursty words in the emails' text using Kleinberg's algorithm (§2.2). We used our domain knowledge to select the words that could be related to changes, such as "change", "modification", "to alter", "to correct", "re-change", "revision", "to adapt" etc. as well as business and engineering acronyms such as "ncr" (non-conformity report), and "cco" (change contract order). We complemented this wordlist with synonyms using The Oxford English Dictionary. Lastly, we asked two experts with pluriannual research and industry practice experience in engineering change management to independently complement this list.

Our final wordlist comprises 381 words in both Danish and English that are likely to relate to changes. Then, we asked two experts to evaluate the wordlist by assigning to each word a value between 0 and 10 to describe the likelihood that a word could appear in an email related to changes. The two experts scored a Krippendorff alpha (Krippendorff, 1970) of 0.7 (95% confidence interval [0.6 - 0.76] computed on 1000 bootstrap replicates), showing a substantial agreement. We finally average and normalise the scores of the two experts in [0, 1] to obtain a single score for each word. We use the obtained wordlist to assign a "change language" score to each email with a similar procedure used by lexicon-based sentiment analysis: we match each word in the text with the words in the lexicon and we average over the scores.

3. Results

3.1. Topic analysis over time

To gain background understanding of the overall design process through associated topics of email conversations, we began by analysing the topics of conversation over time.

Figure 1. Topics over time elicited from emails exchanged during the development of a biomass power plant

For this, we applied the burst detection algorithm (see §2.2) to our email corpus and analysed the subject of each email. We refined the results of the burst analysis by visual inspection of related emails and aggregated some words that appeared in the same period. This was the case for "external piping", "internal piping", "wood chips", "steel structures", "pressure parts", and "spare parts". This was also the case for acronyms, e.g. request for quotations (rfq), selective non-catalytic reduction (sncr), flue gas converter (fgc), air preheater (aph), quality control plans (qcp), and non-conformity report (ncr).

Additionally, we applied the algorithm to the body of each email but the results were hard to interpret. Emails contained both English and Danish language, the text was noisy with project, companies and people names, technical measures, and included a high number of short messages (~30000) with fewer than 20 words. As people in the company appeared to have used the email subject line to identify the topic or the issue at hand, we focus on the email subjects.

The topic analysis over time revealed the evolution of the different phases of the design process with a particular focus on the components of the biomass power plant. The results of the burst analysis are presented in Figure 1, where the period of occurrence for each topic of conversation is highlighted.

Applying topic analysis, we find the following: topic analysis over time reveals three stages in this design process, indicated by bursts. A stage of conceptual design and negotiations centred on the boiler, the drum, and the selective non-catalytic reduction (sncr). In this stage, the company sent many requests for quotations to evaluate possible suppliers to employ during the erection of the power plant. The big burst "contract" highlights the negotiations that took place in this period: contractual agreements with suppliers, financial investors, and the client.

A second stage starts at the end of 2011, when the company secures the contract for the erection and commissioning of the power plant. In this stage, the detailed design begins and the topics of conversation shift toward detailed terms, mainly referred to components.

The "steel structure" is the first topic to be addressed, then it comes the turn of components such as "fan", "membrane", "grate", "superheater", "valves", "fgc", and "aph". We note also the appearance of terms related to structures, i.e. "walls", "external piping", and "internal piping". The burst topic "comments" indicates that from April 2012 to January 2013 people reviewed drawings and designs. We know that the company implements design reviews and that some documents were revised. Thus, this burst highlights the iterative nature of this stage. In this stage, we also observe the development of

quality control plans for the tests to conduct in the next stage. The period from the middle of 2012 to the middle of 2013 involves the testing of the components delivered to the site and the quality of the design. Words such as "inspection" and "report" mark the two principal activities during the testing stage. The acronym "ncr" instead, marks the non-conformity reports issued in this stage, signalling the presence of problems and successive corrective actions, thus iterations.

From approximately mid-2013 to mid-2014 and after the second burst of requests for quotations, we find the commissioning stage where the power plant is developed and tuned for the operating conditions. In the commissioning stage, we find the following bursts: "insulation", "spare parts", "burner", "fuel", "oil", "stoker", "cleaning", "site", and "straw". These words signal site work and work on the feeding system of the power plant.

As it will be discussed in more detail in the next section, these words turned out to be associated with the occurrence of problems and rework: a turbine which was leaking oil, straw lumps getting stuck in the stoker, missing spare parts on site, and problems with a supplier who did not perform cleaning operations after insulation of some pipes.

3.2. The analysis of "change language" and sentiment

To understand in what way changes and sentiment may be connected, we compute change and sentiment z-scores for each email and we plot their trends over time, smoothing the data every 15 days. For this analysis, we consider the email body without attachments. Figure 2 shows the relation between the "change language" z-score and the sentiment z-score. We observe that from January 2012 the sentiment decreases until January 2014 from when the sentiment increases again. For the change score, we note that for most of the project duration it is higher than the average level of change language (z-score higher than zero).

Again, we find three periods. The first period spans from the beginning of the design project (November 2009) to December 2011, encompassing the conceptual design stage followed by a period in which the project was on hold for negotiations. Due to the lack of a sufficient number of emails, we did not compute the sentiment and "change language" scores for this period; thus, it is not shown in Figure 2. However, we describe this period in the following to show the uncertainty and the kind of changes and iterations that happened. For the other two periods, considering the negative effects that changes can have on people's mood (Love, 2002; Pirola-Merlo et al., 2002; Love and Edwards, 2004; Love et al., 2005) and observing the behaviour of sentiment and "change language" scores, we can set expectations about the nature of the events underlying the period of interest. We phrase these expectations in the more general framework of iteration stereotypes (Wynn and Eckert, 2016): progressive, corrective, and coordinative iterations. The second period spans from mid-2012 to mid-2013. We observe that both scores are decreasing, thus expecting churns and other coordinative iterations. The third period spans from mid-2013 until the end of the project. We observe a sudden increase of change score followed first by a decrease of the sentiment, thus expecting problems and corrective iterations and then by an increase in the sentiment, thus expecting progressive iterations.

In what follows, we are describing and analysing the design process of this large engineering project over time. For reasons of confidentiality, the content of the emails is summarised in prose and woven into the text without direct quotations. As such, the next sections provide results in the form of a narrative account or story of the design process under investigation and as a qualitative validation for the numerical scores presented in Figure 2.

3.2.1. Project's kick-off: Concepts, requirements negotiations, financing, planning and re-planning

While waiting for the notice to proceed (NTP) from the client, the project started and a plan for the workload with a time-schedule of 25 months for the project execution was agreed. As the NTP was expected for the 1st of February 2010, the company tried to obtain the most out of the period preceding the NTP by sketching some conceptual drawings and producing a 3D model of the power plant. This period was surrounded by high uncertainty as requirements were underspecified and the exact components to use were unknown. To cope, people relied on their experience, similarities with previous projects, and negotiated technical and functional specifications in different meetings and developed a parametric version of the previous 3D model that allowed to experiment with various settings. This

model was also useful to anticipate some problems with the layout and other practical problems that would have taken place on site during the erection stage. Process and instrumentation diagrams (PID) were iterated many times to accommodate requirements and to provide important details for future suppliers and calculations of temperature and pressure for valves have been performed.

The NTP did not arrive because one or more investors retired and, consequently, the project was put on hold. While the client was searching for new investors, some suppliers who had already received orders and manufactured some mechanical parts (mainly drivers and axles) required the payment of the job regardless the status of the project. The company had to pay, otherwise it would have been subjected to additional cost for late payment. This anecdote shows the financial uncertainty and risk associated with this stage.

The new investors were found and the client requested more efficient turbines and a boiler able to output 1 MW more than planned. These changes increased global project risk. The company evaluated the possibility of receiving some kind of monetary compensation for the increased risk and, at the end, agreed a contract with a fast schedule, bonuses for early delivery and for each MW more in output. The project was re-planned and the official delivery time was agreed to be 30 months, although as an internal strategy the company decided to try to deliver in 25 months. This was motivated as a risk management strategy because, as stated in an internal email between two project managers, "the risk of more than 5 months delay is very unlikely". Comforted by the promising news and the fact that the financial problem was solved, the company put good effort in detailed calculations to be ready to start the "detailed design" stage as soon as the NTP was received. The final signature to the contract, along with the NTP arrived late in December 2011 and the event was marked with a post on the company's website to celebrate the awarded project.

3.2.2. The good, the bad, and the ugly: Progression, re-work and churn, and time wasted

In the period spanning May 2012 to August 2013, we observe a progressive decrease in both sentiment and change scores, with some peaks (see Figure 2). Thus, we expect to observe some progression, a substantial period of design churn and coordination problems, and corrective actions.

The fast schedule agreed during the previous period was implemented as a concurrent engineering process, where many documents were unspecified with a high amount of missing information.

The overall idea was to be fast and to figure out the details on site. Thus, documents describing the air pre-heater (APH) and the layout were iterated to detail missing information. The flue gas condenser (FGC) was completed despite some deliveries were late and the company was forced to change materials. The 3D model of the plant was updated to include the wood chips feeding system. Pressure test and other quality control plans (QCP) were discussed and arranged. The design of grates, transportation units, fans, internal and external piping progressed as well with some iterations to fill the missing information. As a general management strategy during this stage, the company implemented weekly internal progress meetings to evaluate the status of the project and to decide on the next steps. In addition, the company asked each supplier who performed site-work to compile a monthly progress report with pictures included that describe the status of the job.

We found churn iterations and problems of coordination in the realisation of the welds for the piping system. The company in charge of the erection of the piping system produced many wrong welds that did not respect the quality standards agreed and expected for this project. Consequently, non-conformity reports (NCR) were issued and changes were conducted accordingly. However, the changes generated other errors and, thus, other NCRs and changes. This happened because the supplier tried to correct some errors using procedure that were not approved by the company and the procedures were revealed to be wrong and not respect the quality standards.

The situation was worsened by the lack of information in the documents related to the design of the piping system, which did not report the exact welding points and the requirements of the welds. Furthermore, the company shared with the supplier documents that were not yet finalised, with the consequence of adding more confusion. The supplier had already worked with this company in other projects, thus they had a relation of trust. However, the failures described, the corrective actions that the supplier took without the company approval, and the use of documents that were not yet finalised

produced scepticism regarding the ability of the supplier to complete the production in time. Thus, the company asked the supplier to confirm in writing that the deadlines would be respected.
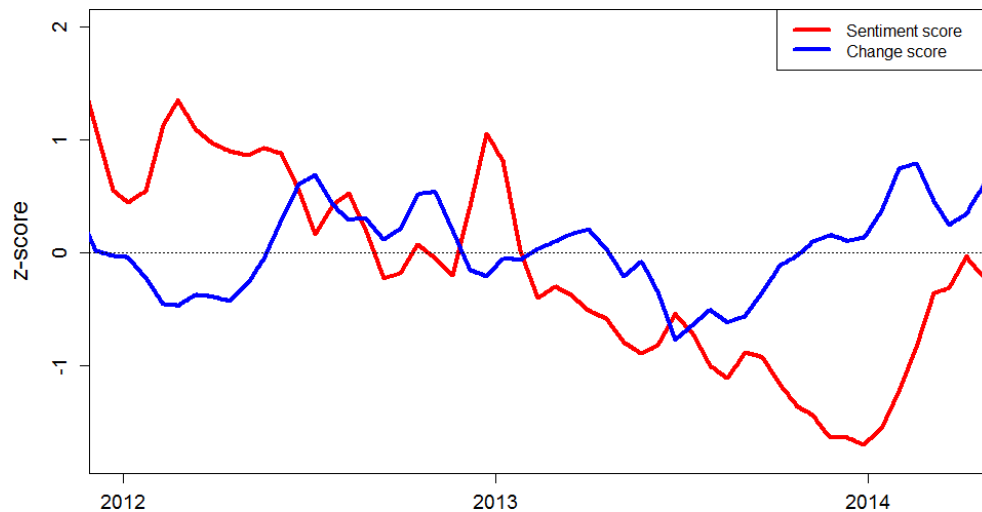
Figure 2. "Change language" and sentiment scores (z-score) over time. The trend lines are obtained by smoothing the scores every 15 days

This whole series of iterations was so critical, took a lot of time (globally around nine months to agree on the correct welding procedures), and increased the risk of not delivering the power plant on time. This drove the company to protect itself from similar future failures by issuing "yellow cards" followed by "red cards" to suppliers that would not follow rules, procedures, and risk assessment methods (RAMs).

The major source of waste was represented by problems in procurement and logistics. Indeed, the company experienced many missing deliveries and also deliveries of wrong or non-conformal parts.

The following anecdote is instructive to understand the waste due to missing/wrong deliveries. The material for the steel structure was supposed to be delivered on site around mid-September 2012 but the supplier did not send it on time. The supplier was contacted many times and, besides being unable to give a reasonable motivation for the missing delivery, sent also some non-conformal material. Within the internal discussion on how to solve this issue, some people proposed to change supplier and stop the current deliveries as they did not add any value but only congestion. In fact, the missing material produced a time and monetary loss as the workers on site were unable to erect the steel structure. At the end, they agreed to have a delivery on site for mid-October, which instead happened at the end of October. To ensure that the right material would be sent on time, the company needed to send a person to the supplier for coordination.

A major challenge and an important milestone for the project was the "main pressure test", which consists of a comprehensive test of all the main connection pipes, boiler, drum, super-heaters, vents, drains, soot blowers, etc. Because of all the problems and waste described before, the timely delivery of the power plant was questioned and successfully passing the main pressure test would have been difficult. In an email to the project leader, one team leader expressed the scepticism and the severity of the situation with peremptory terms. The team leader argued that any further failures due to missing delivery, incorrect design, or missing information would have led to fail the test, making the project late. The team leader warned the project leader that the due rework, to correct all the mistakes, was not "a matter of person-hours" and that the discovery of new mistakes, during the erection stage, was expected. Consequently, the time schedule was revised in the attempt to meet the requirement to successfully pass the pressure test. Further delays were termed unacceptable and the schedule became tight. There was some more rework on site to address the described issues and more detailing of documents where information was unavailable or ambiguous. The corrective actions were successful and the pressure test went "very well" with only minor but doable glitches in the form of three leaking

pipes. The successful passing of the test was celebrated and the good news spread among the project participants.

3.2.3. Roller-coaster ride: Problems, losing the bonus, and successful project completion

From September 2013 until the end of the project, we observe a steady increase of the change score with a sentiment score that decreases until January 2014 and then increases (see Figure 2). As such, we expect a certain amount of rework and then convergence toward the completion of the project.

We found some minor rework in the electrical drawings to obtain approval from the health and safety certifier plus rework on the finishes of the boiler chamber. The steel structure received some corrective iterations due to wrong sizes. The ash conveyor required rework due to a wrongly designed measurement system which did not detect an increasing deposit of ash in the conveyor's box. The company tried many patches but had to resign to change the actual measurement system to a new one.

The air-preheater (aph) showed some vibrations during the tests for commissioning. This problem lead to the additional installation of two sets of guide vanes. In general, rework led to other rework, and other problems discovered during the testing phase for commissioning led to other modifications and additional interventions to the steel structure and some pressure parts.

A fan was found to leak oil and the supplier was called to arrange an investigation to solve the problem on site as soon as possible. However, the supplier was not as responsive and started the investigation around one month after having been notified. The supplier repaired the leaking fan in December 2013, but after 14 days of operations it was leaking again, thus leading to ulterior rework. The problem appeared to be solved after another month of rework. However, it reappeared in April 2014 and then it was finally solved.

Another problem was found while operating the stoker for the commissioning. Some straw bales prevented the correct functioning of the stoker, because lumps of straw got stuck into its clocks. This caused the interruption of the normal functioning of the power plant for around 30 minutes per incident - meaning up to a total of twelve hours per day in the worst-case scenario. This problem could depend on a number of factors, such as the homogeneity and humidity levels of the straw bales, the characteristics of the stoker, and the length of the duct. Many possible changes were suggested, including the installation of a synchronising unit and radical modifications to the position of the bale opener to change the flow of the lumps of straw. After a meeting with commissioning people and the operators, and a test of the stoker's functioning, it was observed that the problem was in most cases only present in one stoker's screw and it was solvable by reversing its rotation. As such, the solution proposed was to change the automatic rotation mechanism to work on the single screw rather than on the whole stoker. The solution proposed for the stoker blockage due to lumps of straw was not fully working and some blockages happened again. Thus, the company started to investigate whether the problem was caused by the bales' humidity level. This generated a long discussion about which method to use to measure the humidity level and in the end, it was found that bales with more than 25% of humidity and bales made of rape seed had to be rejected.

The next milestone was the installation and commissioning of the generator's turbine that had to be performed in collaboration with a partner company (in a consortium). Around mid-November 2013, after its installation, the turbine was found to not work properly and the partner worked to solve the problem. However, people in the company lamented difficulties of getting information from the partner about the specific problem and the status of the turbine. Due to the damage in the turbine, its test and commissioning were delayed to the beginning of 2014. From the emails of some project and technical managers and team leaders, a strong lack of communication between the company and the partner emerge, which underlie a fundamental lack of trust. In fact, people complained that the partner was ignoring many of their technical observations. In December 2013, the malfunctioning turbine rotor was sent back to its supplier, repaired and re-installed. However, it became clear during the vibration tests that the cause for the turbine damage was a problem with the generator shaft and the damage to the turbine was caused by the vibrations. Furthermore, no cooling system to the generator was installed. It was then found that the partner company knew about this problem, as they already detected the vibrations back in November, but did not take any actions nor did they share this piece of information with the company. This negligence had the effects of delaying the project of two months, losing a bonus

of 1.6 million £, generating additional costs to keep the site open for at least two more months. The company had a meeting with the partner to agree on a "realistic action plan" for the future. The company proposed to dismantle the generator and send it back to the supplier, the partner wanted to understand the problem and see if it was solvable on site. As a result of this dispute, no action was taken on the generator until the end of March 2014. At the end, it was decided to send the generator back to the supplier and have it repaired. The repaired generator was then re-installed on site around the end of May 2014. After that and during the test for the final commissioning, a fire in the stoker happened. After investigation, it was found to have been caused by a non-tested change in the programming system made by the partner when the generator was dismantled. Eventually, all problems were solved, the last changes and tuning were performed, and the project was successfully completed and deployed. The power plant was officially opened in November 2014.

4. Discussion, conclusions, and further work

In summary, this paper contributes to understanding the relation between changes and sentiment in design processes and it provides a working methodology to monitor and diagnose a process in real time as well as to learn from previous processes. This paper makes methodological contributions to analysing changes and sentiment expressed in emails (§4.1), contributions to literature on iterations in design processes (§4.2), and contributions to project engineering practice (§4.3).

4.1. Methodological contributions for engineering email analysis

From a methodological point of view, results in this paper show that the "change language" score we developed lets us capture periods with high number of changes and/or periods with high impact changes. Complementing the change score with sentiment analysis lets us capture aspects of the design process that would otherwise be obscured. Thus, we contribute to heightened process understanding.

The qualitative results exposed through the narrative account of the design process provide convincing evidence that our numerical scores capture the behaviour and interrelatedness of sentiment and changes. To illustrate with an example found through the email corpus of the large real-world design project analysed: During coordination for setting up the right welding procedures, we notice that both sentiment and change scores decrease in their trend. We show that even without access to specific change records, that is, even in the absence of specific data for errors, changes, and rework, it is still possible with the help of email analysis, to understand previous design processes and to diagnose them while being executed. Our specific contributions lie in observation of combination of change language and sentiment scores to detect and potentially predict points needing further inspection and attention. Thus, we add our voice to previous studies that leveraged emails to analyse the design process (Wasiak et al., 2011) or to diagnose the "health status" of a project (Gopsill et al., 2014).

Further, such an analysis provides also an opportunity to complement design process analyses using other data traces, e.g. documents. With respect to the analysis of the documents produced during this particular design process (Piccolo et al., 2017), our email analysis confirms previous findings about the negotiations taking place during the conceptual design stage and the iterative nature of this design process. Moreover, it adds findings on the nature of the commissioning stage that was not covered in the document log because the design stage was completed. In other words, the combination of change language and sentiment analysis enables additional insights into a real-world design process, complementing the picture painted by other data-driven insights such as the analysis of document logs.

4.2. Contributions to research on design process understanding

In literature on the impact of change on sentiment, rework and changes have been shown to affect peoples' mood, leading to stress, de-motivation, and having other indirect consequences such as making mistakes (Love, 2002; Love and Edwards, 2004; Love et al., 2005). Further, researchers have found that obstacles such as technical failures that as such may involve change or rework are correlated with negative sentiment (Pirola-Merlo et al., 2002). Our results provide empirical support for correlations between changes and sentiment in general, and between a negative relation between sentiment and changes in particular. Indeed, we found periods in which the change language score increases while the

sentiment score decreases. Contributing to the design research literature (Wynn and Eckert, 2016) on iterations, this pattern may signal the presence of changes or rework and, in general, corrective iterations as we have shown in our results. Researchers have also found that a team's affective climate correlates positively with performance (Pirola-Merlo et al., 2002), implying that a negative team's affective climate correlates with decrease in performance. Thus, negative sentiment may trigger even more mistakes, activating an undesired spiral. An example of this is given by disagreements between the two companies working on addressing the generator failure. This appears to have extended project completion time and increased expenditure. We also found periods in which both change and sentiment scores increase. These periods could indicate convergence. Converging toward the objective may involve changes but may also increase the mood of people involved. Similarly, periods in which sentiment increases while the change index decreases might indicate progression in the design, while periods in which both sentiment and change scores decrease might indicate periods in which people are waiting for data or information, periods in which problems not involving changes happened such as missing deliveries or periods with churns.

These patterns show that while on a general level changes and sentiment are negatively associated, changes can also have positive effects and relate positively to sentiment. Overall, this implies that in order to understand changes, it is important to understand the sentiment toward them. Thus, as changes can affect the sentiment and sentiment can affect performance and decision-making, we call for a change management process that explicitly includes sentiment understanding and - managing.

Coincidentally, when analysing engineering email relating change and sentiment to obtain points of interest for further process investigation and potential intervention into design process practice, when reading the emails associated with the inflection point in question, we noticed incidences of miscommunication. These instances appeared to have been leading to changes with financial consequences for the company and impacted the mood of the project employees (as per sentiment score). Consistently with previous research which termed communication as significant during changes (Fricke et al., 2000; Tavčar and Duhovnik, 2005), we also found miscommunication in our real-world case of the design and development of a power plant. For example, the change in the software that produced the fire in the stoker, the problems with vibrations in the turbine and generator, and the iterations with suppliers to agree on welding procedures may all be interpreted as instances of some form of miscommunication. The change in the software was neither communicated nor tested; for the problems with the turbine and generator, the company commented that the partner did not clearly communicate the problem and the status of the subsequent repairing work; and regarding the iterations for the welding procedures, many important (possibly crucial) pieces of information were communicated late. For instance, an email explaining to the supplier that a single crack in a horizontal weld can expand along the pipe causing ulterior leaks. Only after months of iterations was this email sent to request the needed materials for the welds. We can reasonably assume that such information should have been sent as a requirement to the supplier in the first place and not after so many iterations. This may represent an error in the communication; however, it is unclear if this is the problem or the effect. In this specific case, one could also reasonably argue that the problem and the importance of welds were overlooked. We found instances in email communication where similar reasoning applies. What is evident is that communication is important and its analysis can help to understand and spot problems, as many problems are reflected in the communication (Maier et al., 2008).

In summary, such an email analysis combining change language and sentiment scores also functions as a detector for instances of process anomalies and communication mishaps and can be leveraged to see in real-time what is happening in the project.

4.3. Contributions to project engineering practice

In this research, we show that using email topic and sentiment analysis of a real-world design project for diagnosis and monitoring produces actionable insights for future design processes. To illustrate with a concrete example on how the design process in the large engineering project under analysis here unfolded and how using email analysis as a diagnostic instrument gives insights into alternative ways: The temporal order of topics (Figure 1) shows that the steel structure was discussed before other components that needed to be fitted in. This means that the structure was addressed before addressing

other components that can influence it. Given that the process is based on the integration of many components provided by the suppliers, it may have been worth considering addressing component selection before discussing the creation of ad hoc structures to integrate the various components. This order could be preferential as the characteristics of the components can influence the characteristics of the structure of the biomass power plant. Such order would also be suggested by the application of a sequencing technique based on DSMs (Eppinger and Browning, 2012) that schedule activities influencing other activities to be executed prior to activities that are merely influenced. This is to minimise the possibility of feedback loops and thus, iterations. Ironically, the company had to make changes to the steel structures during the commissioning stage as some dimensions turned out to be incorrect. Topic analysis of engineering email allows for detecting topics of conversations that may or may not be expected at point of occurrence and thus indicates potential deviations from the project plan.

4.4. Avenues for further work

An interesting perspective for enhancing design process learning from past processes to explore in further work is the combined analysis of topics, sentiment and change scores. In that way, it would be possible to understand what are the topics associated with most of the changes and what kind of sentiment there is toward them. We can see this by looking, for example, at the commissioning stage in this particular real-world case. Results from our email topic analysis illustrated in Figure 1 give us the following topics: "insulation", "spare parts", "burner", "fuel", "oil", "stoker", "cleaning", "site", and "straw". Visual inspection of the results from the change language and sentiment score analyses illustrated in Figure 2 point us to a pattern of decreasing sentiment while the use of change language in the emails increases in occurrence. When this happens, as we have seen in our case, such topics appear to be associated with some severe changes and mistakes. It would be interesting to explore the potentials of such an analysis and further test to which level this would be automatable. We intend to do this in future research by formally assessing the ability and precision of our score to detect changes through emails. Furthermore, we will research into algorithms that classify emails related to changes and that detect the type of change or the object of changes.

Acknowledgements

References

Braha, D. and Bar-Yam, Y. (2007), "The Statistical Mechanics of Complex Product Development: Empirical and Analytical Results", *Management Science*, Vol. 53 No. 7, pp. 1127–1145. https://doi.org/10.1287/mnsc.1060.0617

Clarkson, P.J., Simons, C. and Eckert, C. (2004), "Predicting Change Propagation in Complex Design", *Journal of Mechanical Design*, Vol. 126 No. 5, pp. 788-797. https://doi.org/10.1115/1.1765117

Eckert, C., Clarkson, P.J. and Zanker, W., (2004), "Change and customisation in complex engineering domains", *Research in Engineering Design*, Vol. 15 No. 1, pp. 1–21. https://doi.org/10.1007/s00163-003-0031-7

Eppinger, S.D. and Browning, T.R. (2012), *Design Structure Matrix Methods and Applications*, MIT Press.

Fricke, E., Gebhard, B., Negele, H. and Igenbergs, E. (2000), "Coping with Changes: Causes, Findings, and Strategies", *Systems Engineering*, Vol. 3 No. 4, pp. 169–179. https://doi.org/10.1002/1520-6858(2000)3:4<169::AID-SYS1>3.0.CO;2-W

Gopsill, J., Jones, S., Snider, C., Shi, L., McMahon, C.A. and Hicks, B.J. (2014), "Understanding the engineering design process through the evolution of engineering digital objects", *Proceedings of the DESIGN 2014 / 13th International Design Conference, Dubrovnik, Croatia, May 19-22, 2014*, The Design Society, pp. 1773-1784.

Hamraz, B., Caldwell, N.H.M. and John Clarkson, P. (2013), "A holistic categorization framework for literature on engineering change management", *Systems Engineering*, Vol. 16 No. 4, pp. 473–505. https://doi.org/10.1002/sys.21244

Jarratt, T.A.W., Eckert, C.M., Caldwell, N.H.M. and Clarkson, P.J. (2011), "Engineering change: An overview and perspective on the literature", *Research in Engineering Design*, Vol. 22 No. 2, pp. 103–124. https://doi.org/10.1007/s00163-010-0097-y

Kleinberg, J. (2003), "Bursty and Hierarchichal structure in streams", *Data Mining and Knowledge Discovery*, Vol. 7 No. 4, pp. 373–397. https://doi.org/10.1023/A:1024940629314

Krippendorff, K. (1970), "Estimating the Reliability, Systematic Error and Random Error of Interval Data", *Educational and Psychological Measurement*, Vol. 30 No. 1, pp. 61–70. https://doi.org/10.1177/001316447003000105

Langer, S., Maier, A.M., Wilberg, J., Münch, T.J. and Lindemann, U. (2012), "Exploring differences between average and critical engineering changes: Survey results from Denmark", *Proceedings of the DESIGN 2012 / 12th International Design Conference, Dubrovnik, Croatia, May 21-24, 2012*, The Design Society, Glasgow, pp. 1–11.

Love, P.E.D. (2002), "Influence of Project Type and Procurement Method on Rework Costs in Building Construction Projects", *Journal of Construction Engineering and Management*, Vol. 128 No. 1, pp. 18–29. https://doi.org/10.1061/(ASCE)0733-9364(2002)128:1(18)

Love, P.E.D. and Edwards, D.J. (2004), "Determinants of rework in building construction projects". Engineering, Construction and Architectural Management, Vol. 11 No. 4, pp. 259–274. https://doi.org/10.1108/09699980410547612

Love, P.E.D., Edwards, D.J. and Smith, J. (2005), "A forensic examination of the causal mechanisms of rework in a structural steel supply chain", *Managerial Auditing Journal*, Vol. 20 No. 2, pp. 187–197. https://doi.org/10.1108/02686900510574593

Maier, A.M., Kreimeyer, M., Hepperle, C., Eckert, C.M., Lindemann, U. and Clarkson, P.J. (2008), "Exploration of Correlations between Factors Influencing Communication in Complex Product Development", *Concurrent Engineering*, Vol. 16 No. 1, pp. 37–59. https://doi.org/10.1177/1063293X07084638

Nielsen, F.Å. (2011), "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages*, pp. 93–98.

Parraguez, P., Eppinger, S.D. and Maier, A.M. (2015), "Information Flow Through Stages of Complex Engineering Design Projects: A Dynamic Network Analysis Approach", *IEEE Transactions on Engineering Management*, Vol. 62 No. 4, pp. 604–617. https://doi.org/10.1109/TEM.2015.2469680

Piccolo, S., Lehmann, S. and Maier, A. (2017), "Using data- and network science to reveal iterations and phase-transitions in the design process", *Proceedings of the 21st International Conference on Engineering Design (ICED17), Vol. 2: Design Processes, Design Organisation and Management*, The Design Society, pp. 11–21.

Pirola-Merlo, A., Härtel, C., Mann, L. and Hirst, G. (2002), "How leaders influence the impact of affective events on team climate and performance in R&D teams", *The Leadership Quarterly*, Vol. 13 No. 5, pp. 561–581. https://doi.org/10.1016/S1048-9843(02)00144-3

Ravi, K. and Ravi, V. (2015), "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications", *Knowledge-Based Systems*, Vol. 89, pp. 14–46. https://doi.org/10.1016/j.knosys.2015.06.015

Snider, C., Gopsill, J., Škec, S. and Hicks, B. (2017), "The characterisation of engineering activity through email communication and content dynamics, for support of engineering project management", *Design Science*, Vol. 3, pp. 1–31. https://doi.org/10.1017/dsj.2017.16

Tavčar, J. and Duhovnik, J. (2005), "Engineering change management in individual and mass production", *Robotics and Computer-Integrated Manufacturing*, Vol. 21 No. 3, pp. 205–215. https://doi.org/10.1016/j.rcim.2004.07.017

Wasiak, J., Hicks, B., Newnes, L., Dong, A. and Burrow, L. (2010), "Understanding engineering email: The development of a taxonomy for identifying and classifying engineering work", *Research in Engineering Design*, Vol. 21 No. 43, pp. 43–64. https://doi.org/10.1007/s00163-009-0075-4

Wasiak, J., Hicks, B., Newnes, L., Loftus, C., Dong, A. and Burrow, L. (2011), "Managing by E-Mail: What E-mail Can Do for Engineering Project Management", *IEEE Transactions on Engineering Management*, Vol. 58 No. 3, pp. 445–456. https://doi.org/10.1109/TEM.2010.2090160

Wilberg, J., Elezi, F., Tommelein, I.D. and Lindemann, U. (2015), "Using a Systemic Perspective to Support Engineering Change Management", *Procedia Computer Science*, Vol. 61, pp. 287–292. https://doi.org/10.1016/j.procs.2015.09.217

Wynn, D.C. and Eckert, C.M. (2016), "Perspectives on iteration in design and development", *Research in Engineering Design*, Vol. 28 No. 2, pp. 1–32. https://doi.org/10.1007/s00163-016-0226-3

Yassine, A., Joglekar, N., Braha, D., Eppinger, S. and Whitney, D. (2003), "Information hiding in product development: The design churn effect", *Research in Engineering Design*, Vol. 14 No. 3, pp. 145–161. https://doi.org/10.1007/s00163-003-0036-2

Sebastiano Antonio Piccolo, PhD Student
Technical University of Denmark, Management Engineering
Produktionstorvet, 2800 Kongens Lyngby, Denmark
Email: sebpi@dtu.dk