

---

## Leveraging process mining techniques for up-to-date resource profiles

Mathijs Creemers<sup>1</sup> and Mieke Jans<sup>1</sup>

**Abstract:** Keeping an up-to-date overview of the knowledge and skills available in an organization is a difficult task. Looking at this problem we propose the use of data, that already exists in the organization's information systems, as a means to keep user knowledge profiles up to date. To extract the required information from these information systems, we suggest the use of process mining. We illustrate the current state of our research and the end goals we aim to reach with it. We state a general research question. We then apply this to a software development setting, applying knowledge gained during an ongoing case study at a software development company.

**Keywords:** Process Mining, Knowledge Management, Resource Metrics

### 1 Introduction

The real world problem we address is the following: does an organization have an up-to-date view of the knowledge and skills available in their employees? [LSC16]. We first provide some background in the two related themes of our research problem: knowledge management and process mining.

According to Davenport, knowledge management is the process of capturing, distributing, and effectively using knowledge[Da94]. A more formal definition is given by Duhon[Du98]:

"Knowledge management is a discipline that promotes an integrated approach to identifying, capturing, evaluating, retrieving, and sharing all of an enterprise's information assets. These assets may include databases, documents, policies, procedures, and previously un-captured expertise and experience in individual workers."

We focus on the second part of this definition, meaning we consider the expertise and experience in individual workers. In many companies information about the activities performed by employees is logged in information systems. When looking for information about a person, we look at the activities this person performs in the organization. Almost all of the tasks performed by someone will show up in the logs. We aim to leverage this information to obtain a better view of the skills and knowledge the person has to offer. The way to mine information from these event logs is process mining.

Process mining is the bridge between data mining and process modelling. Starting from an event log we use data mining on this log to uncover process information [VdAWM04]. The

---

<sup>1</sup> Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium mathijs.creemers@uhasselt.be

main focus of process mining is process discovery, the discovery of the work-flow of the process, leading to an increasing amount of discovery algorithms. We can however also use process mining to find information about the people (resources) that are part of the process. A social network can be established based on event log information [SvdA08]. Information about the performance of the resources can also be extracted [Pi15]. We can even determine the work prioritisation patterns of resources [Su17].

The structure of this paper is as follows, in the next section we illustrate how we got the subject of our research and state our main research question. In the following section we expand on our research design. After that, we state some challenges or possible concerns. Finally we show the expected output of the research in a paragraph outlining the expected contributions.

## 2 Research Question

The initial focus of research was on bridging the fields of process mining and knowledge management. The motivation behind this was to extend the reach of process mining, moving the focus from work-flow and performance to resource information.

To map out this bridge we performed a literature study of both knowledge management challenges and resource metrics available in process mining. This helped narrowing down the research to three main challenges available in knowledge management, that appear suited to be enhanced by process mining techniques:

1. The creation of a knowledge map
2. Team selection
3. Brain drain

A knowledge map allows a company to have an overview of the knowledge resources available in a company. This can include both physical or digital repositories, but also the knowledge contained in the heads of the employees. The map can be used by employees when they need to find an expert to help them answer a particular knowledge question. The map can also be used by companies to keep an overview of their knowledge.

Team selection is a second common problem in the field of knowledge management. To create the perfect team, we need to know the skills that each employee has. If this is known, the best combination of people can be found for a certain team or project. When it comes to picking a leader for a team, insight in the social network can help. The best leader is one who has a central position in the social network.

Finally, brain drain is a risk faced by most companies. This is the name for the risk that someone leaving results in knowledge being lost from the company. For example if there is only one person who knows how to operate a certain machine, and he leaves the company, the knowledge of how to operate this machine could be lost.

A more detailed description of these problems and an indication of which process mining metrics might be applicable to them can be found in our earlier work [CJ16].

Building on this foundation we kept on increasing the focus of our research. One of the main issues with every solution for knowledge management is keeping it up to date. Consider for example a company where the employees need to update their CV every 6 months. However, most of the information they will add, is already known by the company. The information systems know what projects an employee worked on. We would like to eliminate the manual effort required when updating knowledge management information. Especially when it comes to team selection, an up-to-date profile of the current skills and knowledge of possible team members is needed. Such a profile should contain an overview of the skills and knowledge of the employee. In an ideal situation this profile would also add a maturity level to each of the skills.

This leads us to the following research question:

Can we keep employee knowledge profiles up to date based on data obtained from information systems, in organizations where all of the work is logged in these information systems.

The goal of this research is to be able to support the team selection process with up to date information on each employee while taking the manual updating away. This allows the employees to focus on their work instead of update tasks.

### 3 Research Design

The main goal of our research is to offer a solution to an existing problem. Because we are designing a solution from scratch, this research can be conducted as design science research. We aim to create an artefact that provides a solution to the research problem. We follow the Design Science framework of Peffers et al.[Pe06]

We have already completed the problem identification phase. To complete this phase, we have conducted a literature study into both knowledge management and process mining.

Looking at process mining literature, it can be seen that the focus is currently on performance or compliance measures and work flow discovery. To our knowledge, none of the metrics, that address resource behaviour in a process, is related to the knowledge of the resource. This shows that there is an opportunity to broaden the field of process mining, by adding new metrics that focus on the knowledge of the resources.

A second part of our problem identification was a case study. For our case study we selected a company with a large focus on software development. This corresponds to our research question since the process of software development is extensively logged. When making use of the SCRUM methodology for agile work, as is the case for our case study, there exists a SCRUM board on which all of the activities performed by the team are logged. When these SCRUM boards are electronic, they are an information system we can extract data from. The company we do the case study at is Cegeka, a large Belgian company, situated in the IT business. They offer broad IT-support, ranging from software development to back-end services or infrastructure. At this company we interviewed members of the software development group. Mainly employee in charge of team selection or project management.

These interviews confirmed that the research problem is very relevant in their organization. They currently struggle with team selection and see merit in an automated system that aids them with this. They also see a clear business value for a solution to the problem.

Our research has allowed us to establish a general context for the research problem. We assume an environment in which we have information systems that give detailed information about the activities performed by employees. This includes time registration and human resource information. There should also be a way to map these activities onto skills that are relevant to the team selection or brain drain problem.

We can give a concrete example of this context when looking at the environment of our case study. Here we can identify the following information systems that correspond to the software development context:

- Time registration software
- Electronic scrum boards
- software commits in a version control system

We are currently finishing the next phase in which we define the objectives of our solution. We plan on getting started on the development phase soon. We will then follow the rest of the design science research steps as laid out by the framework.

We will build on our case study during the development phase. The company where the case study was performed has agreed to collaborate during the following design science phases. They will allow us access to their data and support us when handling it.

## **4 Challenges**

Given our context and the lack of previous work surrounding knowledge management using process mining techniques, there are some challenges we can anticipate in the coming phases of the research.

One of the risks is the granularity level of the available data. Data extracted from existing information systems might be too fine grained. This will introduce difficulties when trying to establish a connection between events in the log and knowledge or skills this corresponds to.

A second concern is the source of the data. When working with version control systems, there is no clear notion of activities. We just have a log of commits submitted to the code, but the activity of writing this code may take place before the commit is done. A possible solution for this concern lies in the work of Bala [Ba15]. This shows that it is possible to mine activity information from version control systems.

The final challenge regarding this research is validation. Knowledge and skills are often a subjective measure. We might think someone is good or bad at a certain thing, but there is not always an objective way to measure it. Especially when trying to add a certain level of capability to each skill. A medium skill level in one interpretation could be a high or low skill level in another interpretation. We have to carefully consider our validation methods

ahead of time, to ensure we have an objective way of verifying the profiles that result from our artefact.

## 5 Expected Contributions

We see two main contributions. One is fundamental, showing that it is possible to create an artefact that allows the use of event logs to assist in keeping the profiles of employees up to date. This will also show that process mining techniques can be used when addressing challenges in the field of knowledge management.

A second contribution we see is the validation of the economical value of this extension. Since we are continuing our case study, we will be able to implement the artefact in a business setting. This will allow us to judge the economical impact of the artefact. The results of this contribution can be indicative for further research into the crossover between knowledge management and process mining.

## References

- [Ba15] Bala, Saimir; Cabanillas, Cristina; Mendling, Jan; Rogge-Solti, Andreas; Polleres, Axel: Mining Project-Oriented Business Processes. In: *Business Process Management: 13th International Conference, BPM 2015, Innsbruck, Austria, August 31–September 3, 2015, Proceedings*. volume 9253. Springer, pp. 425–440, 2015.
- [CJ16] Creemers, Mathijs; Jans, Mieke: Social mining as a knowledge management solution. *CEUR Workshop Proceedings*, 2016.
- [Da94] Davenport, Thomas H.: Saving IT's Soul: Human-Centered Information Management. *Harvard business review*, 72(2):119–31, 1994.
- [Du98] Duhon, Bryant: It's all in our heads. *Inform*, 12(8):8–13, 1998.
- [LSC16] Leyer, Michael; Schneider, Christian; Claus, Nina: Would you like to know who knows? Connecting employees based on process-oriented knowledge mapping. *Decision Support Systems*, 87:94–104, July 2016.
- [Pe06] Peffers, Ken; Tuunanen, Tuure; Gengler, Charles E.; Rossi, Matti; Hui, Wendy; Virtanen, Ville; Bragge, Johanna: The design science research process: a model for producing and presenting information systems research. In: *Proceedings of the first international conference on design science research in information systems and technology (DESRIST 2006)*. ME Sharpe, Inc., pp. 83–106, 2006.
- [Pi15] Pika, Anastasiia: Mining process risks and resource profiles. Thesis, Queensland University of Technology, 2015.
- [Su17] Suriadi, Suriadi; Wynn, Moe T.; Xu, Jingxin; vander Aalst, Wil M. P.; ter Hofstede, Arthur H. M.: Discovering Work Prioritisation Patterns from Event Logs. *Decision Support Systems*, February 2017.
- [SvdA08] Song, Minseok; van der Aalst, Wil M. P.: Towards comprehensive support for organizational mining. *Decision Support Systems*, 46(1):300–317, December 2008.
- [VdAWM04] Van der Aalst, Wil; Weijters, Ton; Maruster, Laura: Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, 16(9):1128–1142, 2004.