Brainstorm seminar: The domain-expert-in-the-loop in process mining

Program

15-16 September 2023

This initiative is funded by the Scientific Research Community on Process Mining *(www.srcprocessmining.com)*

UHASSELT
KNOWLEDGE IN ACTION

KU LEUVEN

GHENT UNIVERSITY

Table of contents

About the seminar

The importance of domain expertise is often highlighted in the process mining field, for instance to parameterize process mining algorithms or to give meaning to certain patterns in process mining output. While there is strong consensus on the pivotal role of domain expertise, fairly little research has been done on how to efficiently and effectively involve domain experts in process mining. This has recently been demonstrated by Koorn et al. (2021)[1] with a particular focus on the evaluation of process mining findings: while domain experts were often involved in this stage of a process mining project, a lack of structure in how this happens was often observed.

Against this background, this brainstorm seminar will focus on the **efficient and effective involvement of domain experts in process mining**. This broad topic opens a plethora of specific questions, including: How to identify the right research questions by interacting with domain experts? How to involve domain experts during data preprocessing and event log building? Which domain expert input is needed to create effective process mining output? How should domain expertise be represented? How to evaluate process mining outcomes with domain experts? … Input for further questions will be collected amongst the participants in preparation of the seminar.

The aforementioned questions are of general interest to the process mining field and relevant in a wide range of process contexts, especially in complex processes. Consider, for instance, healthcare processes. When using process mining in healthcare, interaction with domain experts such as medical doctors and nurses is critical given the complexity and variability of healthcare processes. At the same time, healthcare professionals often have no background in process mining (or even data analysis), which has implications on how input needs to be collected. Moreover, medical doctors often have limited availability, implying that the feasibility of approaches to involve them should also be critically assessed.

The two-day seminar will encompass several thematic sessions during which different facets of this topic are discussed. In this way, interesting directions for future research can be identified. Based on the discussions during the brainstorm seminar, a position paper on the seminar's topic is envisioned. Besides the position paper, the brainstorm seminar also aims to initiate research collaborations on specific topics between the attendees.

Scientific Research Community on Process Mining

This brainstorm seminar is funded by the Scientific Research Community on Process Mining. The **Scientific Research Community on Process Mining** aims to interchange research ideas and aspires synergetic research collaborations all over the world. In the second term of the research community, running from 2022 until 2026, a specific focus will be attributed to the following three challenges:

- Challenge #1 – Evaluation of process mining algorithms
- Challenge #2 – Predictive and prescriptive process monitoring
- Challenge #3 – Event data and behavioural analytics

More information on the research community can be found on www.srcprocessmining.com.

[1] Koorn, J. J., Beerepoot, I., Dani, V. S., Lu, X., van de Weerd, I., Leopold, H., & Reijers, H.A. (2021). Bringing rigor to the qualitative evaluation of process mining findings: an analysis and a proposal. *Proceedings of the 2021 International Conference on Process Mining*, pp. 120-127.

Seminar participants

Name	Institution
Iris Beerepoot	Utrecht University (the Netherlands)
Elisabetta Benevento	Università di Pisa (Italy)
Carlos Fernandez-Llatas	Universitat Politècnica de Valencia (Spain)
Thomas Grisold	University of St. Gallen (Switzerland)
Xixi Lu	Utrecht University (the Netherlands)
Mieke Jans	Hasselt University (Belgium)
Owen Johnson	Leeds University (United Kingdom)
Jelmer Koorn	MLC Customer Excellence (the Netherlands)
Felix Mannhardt	Eindhoven University of Technology (the Netherlands)
Niels Martin	Hasselt University (Belgium)
Renata Medeiros de Carvalho	Eindhoven University of Technology (the Netherlands)
Jan Mendling	Humboldt-Universität zu Berlin (Germany)
Luise Pufahl	TU München (Germany)
Hajo Reijers	Utrecht University (the Netherlands)
Alessandro Stefanini	Università di Pisa (Italy)
Marcos Sepúlveda	Pontificia Universidad Católica de Chile (Chili)
Moe Wynn	Queensland University of Technology (Australia)
Francesca Zerbato	University of St. Gallen (Switzerland)

Intended deliverable

Based on the discussions during the brainstorm seminar, a position paper on the seminar's topic is envisioned. This paper, which will be co-authored by all participants, will be submitted to a scientific journal. The set-up and outlet will be discussed during the second day of the seminar.

Contact persons

In case any issue would occur during the seminar, don't hesitate to contact one of the seminar's organizers:

Niels Martin
Hasselt University – Research group Business Informatics
E-mail: niels.martin@uhasselt.be
Mobile phone: +00 000 00 00 00

Iris Beerepoot
Utrecht University – Research group Business Process Management and Analytics
E-mail: i.m.beerepoot@uu.nl
Mobile phone: +00 000 00 00 00

Program overview

Friday, September 15th 2023

	Session description	Chair(s)
09.00 09.20	Arrival with coffee at *Stadskasteel Oudaen* (location: Oudegracht 99, Utrecht - room Linteloozaal)	-
09.40	09.30 (sharp): Welcome and getting acquainted	Niels Martin Iris Beerepoot
10.00 10.20 10.40	10.00: Setting the stage	Niels Martin Iris Beerepoot
11.00	10.50: Coffee break	-
11.20 11.40	11.10: Brainstorm session 1 How to define the appropriate research questions for a process mining project with domain experts?	Francesca Zerbato
12.00 12.20 12.40	12.00: Lunch break	-
13.00 13.20	13.00: Brainstorm session 2 How to leverage domain expertise during the extraction of an event log?	Mieke Jans
13.40 14.00	13.40: Brainstorm session 3 How to leverage domain expertise to ensure the data quality of an event log?	Xixi Lu
14.20	14.20: Coffee break	-
14.40 15.00 15.20	14.40: Brainstorm session 4 How to incorporate domain expertise as an additional input for process mining methods (besides the event log)?	Renata Medeiros de Carvalho
15.40 16.00	15.30: Brainstorm session 5 How to let process mining methods interact with domain experts?	Elisabetta Benevento
16.20	16.20: Coffee break	-
16.40 17.00 17.20	16.40: Brainstorm session 6 How to use visualizations to enhance understanding and interaction with domain experts?	Jan Mendling
17.40		
18.00		

18.20		
18.40		
19.00 19.20 19.40 20.00	19.00: Dinner at restaurant *Hemel & Aarde* (location: Keistraat 8, Utrecht)	-
…		

Saturday, September 16th 2023

	Session description	**Chair(s)**
09.00 09.20	Arrival with coffee at *Stadskasteel Oudaen* (location: Oudegracht 99, Utrecht - room Linteloozaal)	-
09.40 10.00	09.30 (sharp): Brainstorm session 7 How to evaluate process mining outcomes with domain experts?	Jelmer Koorn
10.20 10.40	10.20: Brainstorm session 8 What is needed to enable domain experts to translate process mining results into process improvement actions?	Niels Martin
11.00	11.00: Coffee break	-
11.20 11.40	11.20: Synthesis session and closing remarks	Niels Martin Iris Beerepoot
12.00 12.20 12.40	12.00: Goodbye lunch and end of the seminar	-
13.00		
…		

Detailed program

Composition of the program

The program consists of eight brainstorm sessions. Each brainstorm session focuses on a particular question related to the seminar's topic. Each brainstorm session will consist of two parts:

- The chair will open the session by giving a **short introductory presentation** on the session's leading question. The presentation aims to set the stage for an interesting discussion. Besides covering some pointers from the state-of-the-art in literature, the chair's point of view on the topic can also be incorporated.
- After the introductory presentation, an **open discussion** will take place in which all participants can actively participate to exchange views and ideas. The chair will moderate the discussion.

Friday, September 15th 2023

09.00 – 09.30: Arrival with coffee

09.30 – 10.00: Welcome and getting acquainted

- Welcome by the organizers: Niels Martin and Iris Beerepoot
- Before starting the discussions, it is nice that we get to know each other a little. Consequently, each participant will receive one minute to introduce themselves and to mention why they feel that the seminar's topic is important to discuss.

10.00 – 10.50: Setting the stage

Chairs:	**Niels Martin, Iris Beerepoot**

In this session, the organizers will introduce the brainstorm seminar's topic and its program. Moreover, they will propose a conceptual model regarding the involvement of domain experts in a process mining context. Feedback will be gathered to further refine this conceptual model.

10.50 – 11.10: Coffee break

11.10 – 12.00: How to define the appropriate research questions for a process mining project with domain experts? [Brainstorm session 1]

Chair and introduction:	**Francesca Zerbato**

Examples of subquestions covered by this topic:

- Who to involve in the definition of research questions (background, role within the organization,...)? Should certain roles/backgrounds definitely not be involved in this process?
- How to operationalise this involvement?
- How to leverage domain expertise for process mining prompt engineering?
- …

Reflections/statements:

Name	*Reflection/statement with some background explanation*
Niels Martin	**Good research questions require some basic process mining understanding on the part of domain experts** This statement builds upon the assumption that an organization is confronted with problems where process mining can be helpful (which already requires a first exploration phase)). To define good research questions in this setting, domain experts should have a basic level of understanding of what process mining is and the versatility of its applications. Otherwise, there is a risk that the full potential is not leveraged or that inflated expectations will lead to disillusion. When starting a process mining project in a new environment, it is important to assess the level of knowledge that is available and, if needed, to provide the required basics for them. This can also stimulate thinking in directions that have not been explored before.
Jelmer Koorn	**Balancing scientific and business value** For me the pain point in these projects is finding the middle ground between a question that is scientifically relevant (it is a research question after all) while simultaneously being a question to which the answer generates value to the business. Finding the profile of domain experts that understand and can work within these boundaries is extremely valuable. Another reflection has more to do with how domain experts see a process, this builds on the comment of Niels. The domain expert not only needs to understand process mining as a technique, the researcher and domain expert need a shared understanding of what a process and its contents are (what is a process step, when does a process start and end, etc.).
Jan Mendling	I am a bit confused here. Am I correct to assume that we are talking about what happens in real-world process mining projects? If so, we have to talk about business objectives of such a project, not research questions. Also, speaking of a process mining project is misleading, as much as writing a book is not a Microsoft Word project. A question that I find useful to ask is: which type of business projects make use of process mining technology?
Francesca Zerbato	**Externalizing domain knowledge for (collaborative) question generation** Domain experts can bring into a PM project their knowledge about the domain, the specific process and, sometimes, the data. Often this knowledge is exchanged with process analysts, who might use it together with the domain experts to set the focus of the PM project and derive questions for the analysis. However, I find that there is a lack of support for domain experts in externalizing this knowledge in a way that is usable by process analysts, i.e., that can be linked to the data and used to formulate hypotheses on it. Similar to building a common understanding of what a process is between analysts and domain experts, we could reflect on how to support knowledge sharing and externalization for collaborative question and hypothesis generation. Disclaimer: I have taken a narrower view on "questions" here compared to the broader objectives of a PM project. As such, this view can also fit the Mining&Analysis Phase.
Carlos Fernandez-Llatas	**Define Research Questions that solve REAL interesting (for experts) problems** The questions to solve should attract experts in order to catch their real attention. Doctors usually are working by interest and can expend a high quantity of time for research when the problem is interesting for them. Sometimes, we, as researchers/process ingeniers, we see very interesting problems/ineficiencies in the clinical process and we can try to reduce the time by patient, or even improve offline cares, but doctors are thinking on simpler cases like lack of communication with patients and thay want more time with patients and reject offline consultations (for example in primary care). Also, continuing with the comment of Niels, Not only is necessary to teach que basic of process mining (even the concept of process) but also is necessary to teach in how Process Mining can be used for medical research.

Carlos Fernandez-Llatas	**Define "Doable" research Questions** Sometimes, even in questions very interesting to doctors, the data is not available, has not the needed quality or it is in free text; the questions can be interesting for managers, but not by doctors or viceversa; what is the effects of the research on nursing… . When defining a Research Question it is crucial to take into account all the stakeholders in order to assure that the research is possible in real conditions
Thomas Grisold	**Clarify and align expectations of all stakeholders** It seems to me that process mining initiatives are often implemented without or only partial awareness about the expectations. Partial, here, means that some stakeholders are involved but not others. Process mining initiatives involve stakeholders on multiple levels (e.g. management level with little connection to the actual process, as well as process participants who are involved in the actual enactment of the process). From this angle, questions are particularly useful in the beginning to clarify and align perspectives and expectations of project members (even before the business goals are being defined, see Jan Mendling's comment). The following questions have proven useful in my experience: What expectations do you have about process mining implementation in the next X years? How will that improve X aspect of the organization? What can be challenges to get there? Answers to these questions, in turn, provide the grounds to shed light on blind spots in terms of required knowledge structures and further infer required domain knowledge from all key stakeholders.
Mieke Jans	It is not clear to me whether we are talking about academic research questions or organizational research questions. For academic research questions, I think a question should be formulated by the academic researcher. Of course, this could be inspired by the difficulties that exist when domain experts and process mining experts communicate in their own language, and probably have different responsibilities in the firm. If we are talking about an organisational research question, are we then referring to the challenges of connecting domain experts with process analysts again? (as you can see, the question puzzles me a bit, sorry :)).

12.00 – 13.00: Lunch break

13.00 – 13.40: How to leverage domain expertise during the extraction of an event log? [Brainstorm session 2]

Chair and introduction: Mieke Jans

Examples of subquestions covered by this topic:

- How to efficiently identify the relevant process data in a database?
- How to determine the appropriate level of granularity of an event log with domain experts?
- How do you ensure you extract the right data for a multi-perspective process mining project?
- …

Reflections/statements:

Name	*Reflection/statement with some background explanation*
Jelmer Koorn	**Multi-perspective process mining** Domain experts often think in silos about their processes, in some projects it is interesting to look at a process from multiple perspectives (IT view, task view, role view, etc.). The challenge here is to find an efficient way to extract the right knowledge about the systems in which this data is stored and how it can be extracted and subsequently connected (but the latter is pre-processing I suppose).

Elisabetta Benevento / Alessandro Stefanini	**Inputs from domain expert are crucial** It is frequently hard to understand the data to be collected without the indirect or direct help of the experts. Experts can suggest which data are useful and how to interpret them. Specifically, in complex business environments, the "meaning" of the data is not obvious. In addition, if some data is missing, it could be difficult to extract them later in particular in public service sectors (e.g. healthcare).
Jan Mendling	This task is a challenge because the usual business people driving process improvement projects do not know the data structures. In the worst case, there is nobody in the company who knows the data structures, such that this expertise has to be bought in from technical consultants.
Luise Pufahl	**Business and technical experts need to colloborate** Usually in this step, business experts and a technical expert who is aware which and how data is stored in the databases/dataware houses. The business expert bring into plate what is relevant and the technical experts know what is available. Enterprise Architecture Managment is a field that like to bridge this gap with different methods and techniques. However, I am also not fully an expert there. My questions would be: • Does it makes sense to involve methods, techniques, models from Enterprise Architecture Mangement to allow to involve domain experts in the extraction? Can LLM on top of the databases play a role for supporting the explanation of the available data and constructing the queries?
Francesca Zerbato	**Less is more: Can we support targeted data collection?** My experience with this phase, which is mostly with collecting data from systems that are not process-aware (such as Moodle), is that there is a tendency to first collect as much data as possible and, only afterward, think what their function and meaning could be in the context of the process under study. The line between having rich data and data that are too noisy or complex to prepare and analyze with PM is very thin and, often, there is a lot of effort spent in trying to understand and prepare the data. In which circumstances would PM projects benefit from targeted data collection?
Carlos Fernandez-Llatas	**Real Clinical protocols should be provided by medical experts** Decission events are not in Databases, are in the clinical protocol that usually is in each Doctor mind. It is not usual that the decision that doctors take are categorized properly in the databases. Even, there are many problems in the categorization on Diagnosys in standards like CIE-10. The standard is so huge that usually real doctors does not codify properly the patient diagnosis so this data is not always trustable or incomplete or is written in free text. The decision made by doctors can be inferred from the decision taken after a visit. If doctors can communicate the protocol that the usually follow we can define which data we need from hospital data. This process should be done in collaboration of IT department that is the only that really know what is collected and their possibilities of use it(is trustable? is usable?).
Xixi Lu	**It depends?** For data collection, the challenges of this step highly depend on the nature of the project, in my experience. I have experienced different situations : (1) a single research/consulting project with a company. (2) A large consulting project (by well-established consulting companies). And (3) large projects directly with PM vendors. (1) single research/consultation projects, then we have different roles (a) PM experts, (b) data engineers, and (c) business users who need to work together to extract the right set of data. Here, some challenges are (i) privacy issues, (ii) knowledge exchange between the three types of experts, …

	(2) a project with large consulting companies. These companies usually already have made a huge data dump from their clients' ERP and stored the data dump in their servers. They have BI consultation projects with their customers. Thus, the consulting team tends to act both as the data expert and the business expert. The PM expert in such a team has some freedom to collect data from this server but has no access to the original data. I think the challenge here may be less related to data collection and more related to data scoping. (3) large vendor projects. The data collection tends to be a very huge data connector or ETL project coming from the clients' data lake or BI systems directly. Such a data collection project should be sustainable, easy to maintain, and continuous data collection. The challenges here are more related to the data model and ETL pipeline design and the use of domain knowledge to improve this process.
Thomas Grisold	**Assess data relevance from domain expert's POV** Oftentimes, technical analysis plays out on levels and with goals that are detached from what the domain experts perceive to be important. Early relevance-checks (in terms of what can we see? What does that tell us?) can drastically increase the relevance of process mining-related insights in the long run.

13.40 – 14.20: How to leverage domain expertise to ensure the data quality of an event log? [Brainstorm session 3]

Chair and introduction: Xixi Lu

Examples of subquestions covered by this topic:

- How to assess process data quality with domain experts?
- How to improve process data quality with domain experts?
- ...

Reflections/statements:

Name	*Reflection/statement with some background explanation*
Niels Martin	**Interactive data quality assessment/improvement is the future** In order to detect process data quality problems and to determine appropriate actions to tackle them, the input of domain experts is crucial. Without their expertise, it is very hard for process mining experts to make judgments on these matters, especially for more subtle data quality issues. Asynchronous interaction between domain experts and process mining experts makes data quality assessment/improvement a very time-consuming step. Interactive approaches, in which the domain expert is presented with the data and potential data quality issues is the way to go as they can immediately indicate which issues represent genuine problems and how to handle them. This will speed up this crucial part of a process mining analysis.
Niels Martin	**Context is everything** If we want to involve domain experts during process data quality assessment/improvement, considering the context of process execution is everything. Context can be defined in a multitude of ways: it can relate to the time of day, the day of the week, the busyness of the process,... It is crucial to also incorporate this context to enable a domain expert to judge whether a potential data quality problem actually is one because the same pattern might be a data quality problem in context A, while it is completely normal in context B. Hence, there is a need for context-aware data quality assessment and improvement.

Niels Martin	**Not only remedy, also prevent** At the moment, the focus in research is often on identifying and "solving" process data quality problems. However, it is also important to devote attention to mechanisms to prevent these issues from occurring in the future. This is always preferred to having to handle them afterwards. A first step towards preventing data quality problems is understanding why they have occurred. Hence, there is a need for structured and user-friendly approaches to identify the root causes of process data quality problems together with domain experts are needed. The recently developed Odigos framework is a nice example moving in this direction. One could take it a step further and also develop instruments to support finding solutions to take away these causes, taking into account the trade-off between costs and benefits.
Jelmer Koorn	**Predicting data quality issues with domain experts** Often we start with thinking about data quality after we get the data (that makes sense as we do not know our baseline before we have data). However, perhaps it would be interesting to think of methods that allow for domain experts to indicate where they would expect or know that data quality issues exist in the data that is to be collected. This way, rather than stumbling upon issues, you have some idea about where to look for them and can start thinking about potential fixes before the data is there.
Jelmer Koorn	**Using domain experts to differentiate between types of data quality issues** Another thought, domain experts can be used to differentiate between the gravity of data quality issues, on a high level they could help indicate mission-critical versus nice-to-have-solved quality issues. This can help speed up the process of getting to results in a project without losing too much time on fixing (irrelevant) smaller data quality issues.
Jan Mendling	This problem could be reformulated as "How to check the plausibility of the data available for process mining."
Francesca Zerbato	**Leveraging meta-data for process mining** Reading the thoughts above about the context led me to think about meta-data. I did some search before, and I am not aware that meta-data is explicitly reported for event logs. Meta-data could be a viable way to make knowledge about the context accessible for the Mining & Analysis phase as well as a way to document data quality (and its assessment over time for specific purposes) for future analyses.
Carlos Fernandez-Llatas	**Interactive Data Correction increase doctors trustability** Supporting absolutely the points proposed by Niels, in my experience, the Interactive data analysis, incorporating the clinical expert in the process of cleaning data as well as IT experts of the hospital, not only support in the creation of better data for the process mining work, but also, makes expert aware of what are the real quality of the data that they have, and what are the ineficiencies/problems/needs of culture change needed to improve that data quality that finally will improve the quality of their analysis. With that we achieve a double objective, on one hand, the expert can adapt their level of trustability on the processes discovered, and, on the other hand, makes professional aware of the changes needed on the collection data process to improve the data quality.
Thomas Grisold	**Create shared frames of reference** As it was said before, those who analyze event log data are often disconnected from those who leave traces behind. Based on my own experience, this can lead to flawed inferences/interpretations. In the beginning, it is thus crucial that all project teams jointly create shared frames of reference, that is, knowledge structures through which all stakeholders look at the data/outcomes. One approach that I find useful in that regard is that technical terms analyze some chunks of the data and present their interpretations/conclusions to e.g. those who are involved in the process. This often leads to clarification on both sides in terms of what data is useful and represents what aspects of work.

Xixi Lu	**Decomposing data quality issues and data preprocessing tasks** From a practical point of view, not all data quality issues are relevant for the analysis. Not all data quality issues are worth fixing. How to use domain knowledge optimally (minimally) to achieve the most valuable results? I agree with Niels that context is important and interactive approaches are needed. It would be nice to use the inputs we received from the experts during the interactive approaches to automate the preprocessing pipeline.
Mieke Jans	**Organizational perspective (optimalisation) of gathering input** Even with an intensive knowledge sharing phase between the domain expert and the PM expert at the start of data collection, the domain expert (and perhaps also the system expert) need to be available when checking the data quality. An interesting aspect is when and how to organize these interactions. Are there formats of information exchange that work more efficiently than others? I can imagine that sending e-mails every couple of hours with detailed questions (typically without context!), is less efficient than presenting a broader setting once a week. But is there an optimal stage in the event log building phase to gather this feedback?

14.20 – 14.40: Coffee break

14.40 – 15.30: How to incorporate domain expertise as an additional input for process mining methods (besides the event log)? [Brainstorm session 4]

Chair and introduction: Renata Medeiros de Carvalho

Examples of subquestions covered by this topic:

- Which type of domain expertise would be useful?
- How to elicit domain expertise?
- How to represent domain expertise?
- ...

Reflections/statements:

Name	*Reflection/statement with some background explanation*
Jelmer Koorn	**The varying role of the domain expert** In my experience, in process mining projects we tend to vary in what role we ascribe to a domain expert. In the core of process mining, we state that information systems are a better source of truth. In that sense, we see the role of domain experts as a source of *additional information* - the data is leading and the domain expert complements this information. However, as we see in the statement above about data quality, domain experts can be put in a position to overrule the source of truth (the event log), thereby, switching their role to *primary source of information* - the domain expert can now overrule the data. Beyond these two examples, we vary a lot in what role we ascribe to the domain expert. Mapping this, and better understanding when and how, and especially WHY we incorporate domain experts in various stages in various roles is really interesting to help improve process mining methods.
Elisabetta Benevento / Alessandro Stefanini	**Expert suggestions are usually "rules"** From our perspective, the expert knowledge is expressed by "forbidden sequences", for example A must not happen before B, or by "compulsory sequence", like if/when B happens C should follow it. Such a list of "rules" may be an interesting input for the process mining methods. Maybe, Declarative language can help in this goal.

Jan Mendling	Research on visual analytics has developed techniques for experts to efficiently edit data that is used as input for analysis. I have not seen this in process mining so far.
Francesca Zerbato	**Supporting Knowledge Generation throughout the Analysis** In my experience, analysts tend to structure their analysis based on the domain problems (and the questions defined together with the domain experts in the first phase). However, there is often a step in which analysts need to translate questions to concrete hypotheses and analysis steps and link them to the data in the log. This translation step is often not done at once but is achieved incrementally throughout the analysis by learning from the testing of "not-so-precise'' hypotheses. *Intermediate findings* might be shown to domain experts, who might confirm or disconfirm a hypothesis or even, as Niels writes, overrule the data. However, these findings can also be new for both analysts and domain experts, generating new knowledge that informs future hypotheses and analysis steps. How can we support analysts and domain experts in this knowledge generation process?
Carlos Fernandez-Llatas	**Processes in Health Care Domain are represented as Clinical Guidelines** In case of the analisis of clinical Process, from the 90's Medical Doctors have been working in the creation of "guidelines" for supporting the standardization of the care in the most formal way as possible. Most of these guidelines are based on recommendations, Rules (as Elisabetta and Alessandro said), and DFD representing algorithms. Those guidelines are published and accepted by the medical community . There was some attempts in literature for formalize these guidelines in formal computerized languages, but the manual creation of those are arduous. I think Process MIning is a great data-driven way to fuse both worlds and I'm working actively on that. here are position paper on this issue, that cwe publish with doctors, process Miners and Clinical guidelines experts: (https://www.mdpi.com/1660-4601/17/18/6616). In those guidelines, as well as the way of medical doctors apply those guidelines in real practice is the key to acquire the knowledge about the process. These guidelines can support us in the preprocessing of event log for providing models in the same way that doctors want. Even, Interactive Process Discovery techniques can take advantage of those for creating better models
Thomas Grisold	**How to Find and Articulate *Implicit* Domain Knowledge** It is often hard to "just ask" domain experts about their expertise because much of what they know/do is implicitly represented. For me, the biggest challenge here is to make such knowledge explicit, especially from the side of those who are quite unaware about the domain (e.g. the technical support team of a process mining vendor). A great thing to do would be to develop some systematic procedure (e.g. in the form of typical questions to ask) that can reveal those things that are important to know. (There has been quite some research around this in the knowledge management field, see e.g. works by Nonaka, von Krogh and others; perhaps it would be cool to map findings from this field to process mining research).

15.30 – 16.20: How to let process mining methods interact with domain experts? [Brainstorm session 5]

Chair and introduction: Elisabetta Benevento

Examples of subquestions covered by this topic:

- How to operationalize the interaction between a process mining method and domain experts?
- How to avoid confirmation bias?
- Which knowledge/skills are required to enable the interaction between process mining methods and domain experts?
- …

Reflections/statements:

Name	*Reflection/statement with some background explanation*
Niels Martin	**Beware of the slippery slope of confirmation bias** Process mining has the key advantage of reflecting the real-life behavior of business processes, instead of reflecting how actors hope/think the process is executed. As a consequence, process mining can bring inconvenient truths to the surface. If the domain expert is actively steering the analysis, this might lead to confirmation bias, i.e. the data is consciously or unconsciously used to prove his/her gut feeling, which might only be a selective view on reality. Hence, sufficient safeguards need to be in place to avoid that giving control of the analysis to the domain expert leads to confirmation bias.
Elisabetta Benevento / Alessandro Stefanini	**Expert with PM knowledge Vs Expert-friendly PM methods** Process Mining (PM) methods/tools usually require PM and process modelling skills. Unfortunately, experts often do not have such skills. A first solution is the help of a PM analyst. This may limit the actual application of such techniques in real business contexts. Potential solutions may be the enhancement of expert knowledge through brief training courses or similar initiatives, or the development of more expert-friendly methods/tools. Many times this does not require necessarily the development of new algorithms/methodology but the simplification of the interface with the expert.
Elisabetta Benevento / Alessandro Stefanini	**Confirming Vs Modelling** The emerging interactive techniques mainly follow two different approaches: 1) the PM method creates one/more process models and the expert has to confirm, decline, or modify the proposed model; 2) the expert creates the process model starting from the inputs of the PM method/tool. Which approach is better? Is it dependent on the context?
Jan Mendling	I recommend Du, F., Shneiderman, B., Plaisant, C., Malik, S., & Perer, A. (2016). Coping with volume and variety in temporal event sequences: Strategies for sharpening analytic focus. *IEEE transactions on visualization and computer graphics*, *23*(6), 1636-1649.
Jelmer Koorn	I think it would be interesting to consider the goal of having a domain expert model their own processes. I can imagine that there are a variety of reasons to do this, each with its own pros and cons in terms of method / technique used. For example, having domain experts model a process because of their in-depth knowledge of a process might be better served through the support of a confirming technique as the domain experts will easily and quickly identify the appropriate process model and its variants. Whereas I can imagine that if the goal is to create ownership of a model it might be better served with a modeling technique as domain experts will feel they have built the model themselves.
Carlos Fernandez-Llatas	**Co-creating Interactive Process Indicators in interactive Data Rodeos** Medical Experts usually have not knowledge about PM tools. Even, the majority of Medical doctors have not enough knowledge about processes language to design or sometimes understand the workflow language. For that, one of the best ways of creating fully understandable/trustable PM systems for doctors is the co-creation of them. We have designed a methodology incorporating the medical expert and the IT crowd of the hospital in a series of collaborative sessions (Interactive Data Rodeos) with the aim to create enhanced process models that can serve as an alternative to classical KPIs as indicators (Interactve Process Indicators) for the analysis of processes performance. We have tested this methodology in several hospitals around europe with good results. In the book (Fernandez-Llatas, C. (Ed.). (2021). Interactive process mining in healthcare. Cham: Springer) we depicts our experiences with Interactive Process Indicators and Data Rodeos .

Thomas Grisold	**Ensure sustainable momentum** I have observed in several organizations how process mining fell into oblivion after a short "hype" (e.g. 3-4 weeks). A key challenge is to frame the methods in a way that domain experts find useful means to explore and analyze aspects of the process in the long term. For me, this often depends on teaching the domain experts to ask the right questions and define clear goals to be followed up in the long term (which can only be answered through PM). This could also be prompted (and refined) by the program itself. E.g. depending on progress, the user is invited to perform a little more complex analysis in iterative fashion. PS. Niels' comment made me think: Indeed, one reason for domain experts to "forget" about process mining is that they see what they expect to see (see Niel's comment) which gets boring after some time.
Mieke Jans	**Avoid finger-pointing** Oftentimes, PM projects provide insights in differences between groups. Aside the risks that are already mentioned above, there is the risk of creating the impression of finger-pointing towards 'less performing' groups. Instead, communication skills are very important to convey the message of 'lets's have a look at the context of the different groups, in order to understand the differences in process executions'. If this isn't done correctly, you risk loosing support of important stakeholders.

16.20 – 16.40: Coffee break

16.40 – 17.30: How to use visualizations to enhance understanding and interaction with domain experts? [Brainstorm session 6]

Chair and introduction: Jan Mendling

Examples of subquestions covered by this topic:

- What are the key requirements that process mining visualisations should fulfill?
- How to assess the appropriateness of visualizations?
- …

Reflections/statements:

Name	*Reflection/statement with some background explanation*
Niels Martin	**Visuals need to be made robust for misinterpretation** There is definitely significant research potential in further exploring how the visual representation of process mining outcomes can be extended and improved. Making things visual is highly beneficial when interacting with domain experts. However, according to my knowledge, little is known about how domain experts interpret process mining outcomes. As this interpretation step is key for the translation from insights to actions (which relates to the topic of brainstorm session 8), it is important that visuals are made robust for misinterpretations. Good practices can be put in place (inspired by other domains) to guide the community in the right direction.
Niels Martin	**Less is more** There is definitely a need for further research on the visualisation of process mining outcomes. However, we should be careful not to overengineer process mining visualisations. 'Less is more' might also hold for process mining visualisations. As a consequence, it is important to always carefully assess the understandability of a visualisation with users whenever introducing something new.

Jelmer Koorn	**Dependencies with other processes/products/services visualised** In practice I see a lot of domain experts requesting the visualisation of dependencies between processes. This might be more interesting to study especially if a process mining project is not taking place in a vacuum of a single process, but has a place in a larger process architecture. My experience thus far is that visualising dependencies often limits the readability of a process, but not including them leaves the domain expert unsatisfied.
Elisabetta Benevento / Alessandro Stefanini	**How to visualize the variation of model quality during interactive mining?** The PM tools should permit the visualization of the changes in process model quality (F-score, fitness, simplicity, etc.) due to the expert interventions on the model. This may increase the awareness of experts about the changes they are proposing and the related effects on the model.
Francesca Zerbato	**Understanding PM tasks and purposes** In my view, PM needs fit-for-purpose visualizations that are created to support specific tasks. I have the feeling that, despite their limitations, most users rely on DFGs for many tasks (understand the control-flow, visually identify bottlenecks, identify/show rework, compare processes) because they are "fast to read" and "interactive". To identify key requirements for visualizations I see the need to identify common process mining tasks and their requirements.
Francesca Zerbato	**Augment Control Flow Visualizations** PM tools tend to put control-flow analysis in the focus, "forgetting" that users might want to have interactive multi-perspective visualizations that show other process perspectives besides the control-flow or on top of it.
Francesca Zerbato	**Consider Visualization Literacy** Besides being fit-for-purpose, visualizations should also consider the target audience they are designed for. Domain experts might not be able to read and configure visualizations that are meant to be used by analysts. Still they are often exposed to them, e.g., in the Evaluation phase. It might be important to complement visualizations with guidance on how to read them, especially when interactive visualizations that are sensitive to the data and the user configuration setting (e.g., dotted chart) are given to domain experts or used to present results.
Carlos Fernandez-Llatas	**Take advantage of PM characteristics for creating interactive control-flow visual Navigable Models coexisting with classical/new Visualizations** I agree with Francesca points. in my opinion the Control-flow views is the crucial differential aspect of PM over all the Bussiness Intelligence perspectives. PM control flow systems allow an easier and navigable way to investigate into the process looking for specific issues that are not shown in other techniques. Making those models coexisting in the screen with other classical visualizations or the KPIs that medical experts usually manage makes not only that users understand better the process, but also, the Clinicians can understand better the control flow. However, the creation of visualization tools with classical KPIs can have a problem. Experts could obviate the PM Control flow and center in that they really know (KPIs). For that is very important to make efforts in make health professionals in undertanding the power of control flow visualization over others.
Xixi Lu	**How to start such a subfield?** This topic is very interesting and seems to me like an entire subfield. There have been earlier efforts to involve Information Visualization experts. However, this subfield seems not receiving much attention? To me, the question seems to be how can we get senior and young researchers being interested in this topic and push this research topic forward.
Mieke Jans	**Bridging the 2 research fields** Following up on Xixi's question, I think the first step is to understand the mechanics of the different fields. We think in terms of process instances, activities, control-

	flow… The other field has their own building blocks. A research project to understand how these elements could be brought together is in my opinion a first step.

19.00 - …: Dinner

In the evening, we can enjoy dinner at the fine-dining restaurant *Hemel & Aarde* in the city center of Utrecht (Keistraat 8, Utrecht). More details on the restaurant can be found on the last page of this booklet.

Saturday, September 16th 2023

09.00 – 09.30: Arrival with coffee

09.30 – 10.20: How to evaluate process mining outcomes with domain experts? [Brainstorm session 7]

Chair and introduction: Jelmer Koorn

Examples of subquestions covered by this topic:

- What should be the subject of the evaluation?
- Which approach to use in order to evaluate process mining outcomes with domain experts?
- What are the key pitfalls regarding the evaluation of process mining outcomes with domain experts?
- How to assess whether domain experts understand process mining outcomes?
- …

Reflections/statements:

Name	*Reflection/statement with some background explanation*
Niels Martin	**More structure is needed** In process mining research, the evaluation of outcomes with domain experts often lacks structure. In many papers, it is merely reported that outcomes have been presented and/or validated to domain experts. From this, it follows that the findings "make sense". As process mining is maturing as a research field, more structured and rigorous approaches should also be put in place to evaluate process mining outcomes with domain experts. As a community, we do not have to reinvent the wheel and we could learn from other fields to define guidelines to appropriately evaluate process mining outcomes with domain experts.
Francesca Zerbato	**What makes a process mining outcome a "good" one?** I know that this question might open a Pandora's box but I like to think that there are ways to assess, together with domain experts, whether the result of an analysis is good or not. While, again, evaluating a result is much narrower than assessing whether a business objective has been achieved or not, I think students (and process analysts themselves) might need more objective criteria to assess the "quality" of PM outcomes. Here quality can be measured in many ways, considering the analysis process, the results, the way the results have been reported, etc.
Carlos Fernandez-Llatas	**Integrate PM tools in Lean Healthcare, Value-Based HealthCare, High Value Care, and Evidence Based Medicine perspectives.** In last years, from the Clinical perspective, there are some of works dealing with the concept of value in the sense of what are the value by resources compromised that the patiens/professional receive/perceive. There are initiatives, such as ICHOM, that are looking for indicators that are able to measure the value chain of each process action, in terms of cost, patient health, patients experience, professional experience, and, even, ethics. In my opinion, PM has a huge advantage in this area because is able to create models that could represent all tha value chain and discover inefficiencies, wastes, adverse effects, bottlenecks, etc. that are related to one or several dimensions of the value defined. In my opinion, a Good model is the one that make professionals understand de value chain and supports the analysis and the iterative optimization of the process analized. Also, taking into account Evidence Based Medicine paradigm, PM techniques can support in the creation formal Data driven Clinical guidelines. This could be very

	interesting for creating formal gold standards that could drive to a way to quantitatively and qualitatively measure the process variability in healthcare. As Niels said, we do not need to reinvent the wheel, Healthcare domain has several works that are trying to deal with this problem and I think Process Mining can offer some lights in this line.
Thomas Grisold	**Evaluate initial goals and keep developing new goals** See topic 1 -> if the goals are clearly defined and the expectations are aligned in the beginning, then it is important to reflect if goals have been met (e.g. after 6 months). This, in turn, may now lead to more refined/granular goals. This is often enabled through better knowledge about process mining, which can be combined with existing domain knowledge.

10.20 – 11.00: What is needed to enable domain experts to translate process mining results into process improvement actions? [Brainstorm session 8]

Chair and introduction: Niels Martin

Examples of subquestions covered by this topic:

- What are the current barriers that might inhibit domain experts to translate process mining results into actions?
- How can those barriers be overcome?
- …

Reflections/statements:

Name	*Reflection/statement with some background explanation*
Jelmer Koorn	**Process Mining as part of a continuous improvement culture or as a stand-alone project?** It is interesting that we often refer to process mining projects as projects, inherently this means that it has a clear start and end and is non-recurring. In my experience, process mining is best situated as part of a continuous improvement cycle. This would also ensure that the outcomes of a process mining exercise are by standard of the continuous improvement method incorporated in a next step where their feasibility is assessed and (perhaps) implementation will follow. The question is, should process mining be presented as a project style effort when the only outcome is a list of fixes of which the implementation is (almost) never a part of the initial project proposal (timeline and budget)? Should process mining not stay true to its own promise of generating value by working in a process style manner?
Elisabetta Benevento / Alessandro Stefanini	**Process Mining as a standard evaluation method** Process Mining methods should be implemented in the performance management systems. This may be possible if the methods, through appropriate tools, are easy to use for business managers. In this way, managers can exploit the potential of PM methods for anayzing process performances and investigating the potential root cause of "problems".
Elisabetta Benevento / Alessandro Stefanini	**Process improvement actions are usually difficult to carry out** The main barriers to this issue are process mining related or are process-improvement related? Usually, taking process improvement action is a non-trivial task, independently from the analysis methods. We are wondering if the main difficulties in translating process mining results into process improvement actions are related to the flexibility of the (management) systems.

Carlos Fernandez-Llatas	**Process Mining as a tool for Continuous Improvement of healthcare process** Continuing the point proposed by Jelmer, in my experience, the use of Process mining tools as Interactive Process Indicators (IPIs) a way to provide the real status of the process and to measure the effects of change catch the interest of both, hospitals managers and doctors. Continuous improvement technologies such as Lean Healthcare, use classical KPIs that have partial views of the reality of the process and does not offer a way to investigate the causes of the processes changes. Process Mining based indicators, such as IPIs, can offer not only a way to show and measure, creating enhanced modules, but also compare the differences among processes implementation thanks to conformance checking techniques. These indicators can offer a data driven, objective, navigable way to analyze and measure the healthcare process to allow a better assessment in the Continuous improventent of Healthcare process. As Elisabetta and Alessandro said, these indicators, can serve as a standard way to measure the best practices in the implementation of a new policy in a health center.
Thomas Grisold	**Define points of leverage** Process mining enables a "new way of seeing" re: process performance. My impression is that domain experts are often searching for systematic guidance in terms of what they can actually do now with these insights. What could be relevant: A systematic overview (taxonomy?) that maps insights with potential points of leverage, i.e. specific points/aspects that can be improved through a given set of actions.

11.00 – 11.20: Coffee break

11.20 – 12.00: Synthesis session and closing remarks

Chairs: **Niels Martin, Iris Beerepoot**

The goal of the synthesis session is to bring together some key points from the several brainstorm sessions. Moreover, we will revisit the conceptual model that was introduced in the introductory session on the first day to see whether insights have evolved. At the end of the synthesis session, we will also discuss the position paper, which is the intended deliverable of the seminar.

12.00 – 13.00: Goodbye lunch and end of the seminar

Practical information

Seminar location

The brainstorm seminar will take place in ***Stadskasteel Oudaen*** in the city center of Utrecht. In particular, the seminar will take place in room ***Linteloozaal***.

Stadskasteel Oudaen
Address: Oudegracht 99, 3511 AE Utrecht, the Netherlands
Website: https://www.oudaen.nl

Dinner

In the evening of September 15th, we will have dinner at the restaurant ***Hemel & Aarde*** (literally translated as: '*Heaven & Earth')* in the city center of Utrecht.

Restaurant Hemel & Aarde
Address: Keistraat 8, 3512 HV Utrecht, the Netherlands
Website: https://www.hemel-aarde.nl