UNIVERSITY OF TARTU
Institute of Computer Science
Software Engineering Curriculum

Janna-Liina Leemets

Persuasive Visual Presentation of Prescriptive Business Processes

Master's Thesis (30 ECTS)

Supervisors: Fredrik Payman Milani
Alexander Udo Nolte

Tartu 2022

Persuasive Visual Presentation of Prescriptive Business Processes

Abstract:

Organizations create data through business processes, which can be utilised to build prescriptive systems. A prescriptive business process prescribes the process worker a recommendation of either actions to take or resources to use. However, a worker may choose to follow their knowledge or intuition instead of trusting the given prescription. Therefore, to make the prescriptions acceptable for the worker, they must be visually presented in a persuasive way. This thesis addresses the research question of how to visually present prescriptive business processes persuasively in order to optimise process outcomes and efficiency. The contribution of this thesis is a set of persuasive principles to guide the visual presentation design of prescriptive business processes. These principles combine existing persuasion and visualisation literature. The results of this thesis can be useful to those who design and develop prescriptive business processes as they can make prescriptive systems more useful, easy to use and trustworthy for process workers.

Keywords:

Prescriptive business process monitoring, PrBPM, recommendation, prescription, persuasion, visualisation, persuasive principles

CERCS: P170

Soovitussüsteemidega äriprotsesside veenvalt kujutamine

Lühikokkuvõte:

Organisatsioonide äriprotsesside käigus tekivad andmed, mida saab rakendada automaatsete soovitussüsteemide loomiseks. Soovitussüsteemiga äriprotsess kirjeldab töötajale kas järgmist parimat tegevust või ressurssi, mida kasutada. Töötaja võib aga soovituse asemel järgida enda teadmisi või intuitsiooni ning põhjustada organisatsiooni protsesside ja nende tulemuste tõhususe vähenemist, mistõttu on vajalik soovitussüsteeme kuvada veenval viisil. Antud lõputöö käsitleb uurimisküsimust, kuidas visuaalselt kujutada soovitussüsteemidega äriprotsesse veenvalt, et parandada protsesside efektiivsust. Lõputöö panuseks on põhimõtete kogum, mis kirjeldab, kuidas visuaalselt esitleda selliseid soovitussüsteeme. Põhimõtted loodi visualiseerimis- ja veenvusteooria kombineerimisel. Antud lõputöö tulem on kasulik neile, kes disainivad ja arendavad soovitussüsteemidega äriprotsesse, sest see aitab luua töötajate jaoks lihtsasti kasutatavat, kasulikku ja usaldusväärset süsteemi.

Võtmesõnad:

Soovitussüsteemidega äriprotsess, soovitus, veenvus, joonis, veenvuspõhimõtted, kujundus

CERCS: P170

Table of Contents

1. Introduction

Recent decades have witnessed an outstanding growth of artificial intelligence (AI) implementations. Much of life is connected to information technology and thus, a substantial amount of data is created. This data can be used to make descriptive, predictive and prescriptive AI models. Where descriptive models answer the question "what happened" and predictive answer questions such as "what will happen", the most complex of the three, prescriptive models, answer the question "how to make something happen" [1]. Such AI-based systems of recommendations are becoming more popular in many fields, including data-driven management of business processes.

A business process is a chain of interrelated operations performed in an organisation to achieve business goals [2]. Event data gathered during various stages of a business process execution is written in event logs. These event logs can be used for evidence-based management of business processes. Methods that use event logs to manage different aspects of business processes are collectively called process mining [3].

Although process mining was initially concerned with descriptive and diagnostic analysis, it has recently extended to predictive business process monitoring techniques. Predictive monitoring constructs a predictive model from event log data and can then be used to predict the next action to take, process outcomes etc. [4]. Predictive process monitoring has adopted advanced machine learning methods that allow it to create consistent and precise predictions. However, a prediction alone is not enough to make an informed decision and should be paired with a recommendation [5]. A collection of techniques for recommending or triggering actions (i.e. interventions) during the execution of a business process is called prescriptive process monitoring (PrBPM [6]). It is being researched to help process workers monitor and make decisions of key importance on the fly. PrBPM aims to optimise business processes with regard to one or more performance indicators. PrBPM methods propose interventions that improve the probability of a desired outcome by recommending actions to take, which resources should perform a task or alarm process workers about a possible negative outcome [5].

Prescriptive business processes can only be fully leveraged if the prescriptions it offers are followed [7]. A prescriptive system, also called a recommender system [8], takes advantage of complex machine learning techniques and by nature these systems work as a black box. A black box system only outputs the final result, such as a recommendation, and lacks transparency with regards to how a prescription was made [9]. Consequently, the user cannot directly understand the inner mechanisms of the prescriptive system and might not consider a recommendation generated by AI. Instead, a worker may base their decision on intuition or previous experience. A recommender system cannot make the user abide by the recommendations it makes despite the precise results it offers. Therefore, these systems and the way they are presented must be understandable and trustworthy for process workers to accept them. This is a concern that must be addressed during business process improvement and execution.

Prescriptive systems inherently aim to be persuasive, as their goal is to influence a decision a user makes. As defined by Fogg [9], persuasive technology is "an interactive technology that changes a person's attitudes or behaviours". Persuasion has been extensively studied in psychology to help solve, for example, addiction or dieting issues [10]. These persuasive principles

have also been successfully applied to human-computer interaction [11], [12] and new computer-specific behaviour change theories have been developed [12]. As persuasiveness of software solutions has been proven, the same theories can be applied to applications used in business processes.

Little research exists on the topic of visual presentation of persuasion. The most notable examples compare the persuasive power of tables and charts [13] and the effect of interactivity on persuasiveness [14]. Even further, as far as the author of this thesis is aware, presenting recommender systems with visualisations or persuasive textual messages within business processes has yet to be researched. This work seeks to reduce this gap by examining how to present recommender systems within business processes to make them more acceptable for the user. In this thesis, acceptance is defined as willingly making decisions based on a recommendation with or without agreeing with the recommendation. In other words, the user trusts the system.

This thesis aims to research existing prescriptive business processes and persuasion theories, identify aspects for improving the acceptance of PrBPM recommendations and condense the findings into principles for visualising such a persuasive system in business processes. On the basis of these principles, a prototype will be created for evaluation. The research goal is to ultimately create a guide for making the visual presentation of prescriptive business processes persuasive for process workers. The following research question has been formulated to guide the thesis' research:

RQ: How can prescriptive business process monitoring be visualised in a persuasive manner?

To answer this question, the Design Science framework is followed [15]. According to the framework, the first step is to explore the problem's relevance. For this, a review of literature is conducted to collect information about persuasion and visualisation principles, in addition to the persuasiveness of recommender systems. Presentation of prescriptive business processes is a novel field and thus, all papers will be collected and then the findings will be contextualised to fit PrBPM techniques. Based on the gathered data persuasive principles and requirements are elicitated and further utilised to create an artefact – the prototype. The prototype will be addressed with the acronym PERSEVERE (PERsuasive System for prEscriptiVe businEss pRocEss). Finally, PERSEVERE will be evaluated with potential users from the industry.

The rest of this paper is structured as follows. Section 2 gives an overview of the background of prescriptive business monitoring, visualisation and persuasion research. Section 3 details related works in persuasion, PrBPM and AI explainability. Section 4 outlines the methodology of this thesis. In section 5, persuasive principles are described. Section 6 describes requirements and their elicitation. Section 7 provides an overview of the development of the prototype. In section 8, the evaluation of the prototype is outlined. Section 9 contains the discussion of the findings and a discussion. Section 10 concludes the thesis with a summary.

2. Background

This section introduces concepts and research about prescriptive business process monitoring (section 2.1), visualisation (section 2.2), and persuasion (section 2.3).

2.1 Prescriptive Business Monitoring

Prescriptive business process monitoring currently aims to achieve two goals: optimise process (herein also called a case) outcomes and process efficiency [5]. An example of the first goal is a missed deadline or risk such as a customer turnover increase, the rate of which the prescription aims to decrease. An example of the second goal is improving key performance indicators (KPIs) by doing case optimization, reducing cycle time or increasing service quality. Some PrBPM methods use alarms that trigger interventions to avoid or reduce defect rates. Others, as mentioned before, recommend actions to take next and which resources should perform a task, e.g. which clerk should carry out an activity.

PrBPM has adopted different complex machine learning models, which in turn base their interventions on either predictions or similarity scores calculated by machine learning models in predictive business monitoring [5]. Interventions are triggered by policies given to the model, for instance, an attribute exceeding metric limits, a set of resource rules or the probability of an event being above a certain threshold. PrBPM is a novel field and thus, not many specialised algorithms currently exist. The most common techniques include decision trees, long short-term memory network (LSTM) models, random forests, neural networks and nearest neighbour algorithms.

Following are some recently proposed examples of PrBPM frameworks and the underlying algorithms they use. Weinzierl et al. [6] used a LSTM model to recommend the next best action out of predicted next actions based on an optimal KPI. The KPI, in this case, is the remaining time of a case. Metzger et al. [16] implemented deep recurrent neural network (RNN) models to create a predictive model and used the predictions to create adaptive interventions. By leveraging online reinforcement learning, it is possible to determine when a change should be made during process run time. Fahrenkrog-Petersen et al. [17] presented a framework for a random forest and gradient boosted tree classifier-based prescriptive alarm system that would notify of a predicted negative outcome to prevent or mitigate it. Gröger et al. [8] used real-time prediction-based recommendations that would prevent deviation from metrics during a case. Classification techniques were used for predictions and decision trees for recommendations.

A recommender system presentation consists of input, process and output characteristics [18]. Input characteristics include all inputs the user can provide for the system, such as search bars, input fields and filters. The data needed for input is either control flow, resource, temporal or domain-specific [5]. Methods that recommend items impacting control flow mostly use control flow data, for example, task duration or frequency. Resource data can be used for different prescriptions, i.e. for optimising a resource's performance or allocating a resource to a task. Temporal data (or time data) is often co-implemented with other data types and can be used, for instance, to recommend the most suitable resource for a task per time of day. Domain-specific data are specific features, such as manufacturing materials or medical data.

Process characteristics are displayed when the prescription is being calculated [18]. Examples of these cues are loading screens, visualisations of the search process and the system's response time. These visual indicators give a clear overview for the user about the effort the system is making to give a prescription and should be present to avoid confusion during loading times.

Output characteristics are related to the given prescription. The output data can be, regardless of whether the recommendation is a resource or activity, either [19]:

- one-dimensional, e.g. item cost, time taken by process, process step count or probability of a risk
- multi-dimensional, e.g. a process with cost, average execution time or steps taken,
- text, e.g. a textual description.

2.2 Visualisation

Data visualisation is a common tool for presenting information, as it enables to visually encode complex statistical data. The goal of visualisation is to help us to understand data by leveraging our innate skill to spot patterns, trends and outliers [20]. An effective visual representation of data is a shorthand for cognitive reasoning. According to Chen et al. [21] "Graphic displays are often very effective at communicating information. They are also very not effective at communicating information", which, in other words, means that an effective visualisation is not concerned with the amount of information delivered, but rather how easy it is for viewers to comprehend it. As this paper's purpose is to increase the persuasiveness of a recommendation presentation, it is crucial to select visualisation tactics that aid in making the prescription acceptable. The existing body of work on general principles must be followed, such as Munzer's rules of thumb [22], KISS (an acronym for "keep it simple, stupid"), or Shneiderman's Visual Information Seeking Mantra [23], as well as principles of persuasion such as the four dimensions defined by Fang et al [24].

Munzner presented seven rules of thumb or best practices when visualising data [22]:

- The rule of "No unjustified 3D" states that 3D visualisations are often misused and are only necessary when visualising spatial data. When implementing 3D visualisations, occlusion occurs when many data points are hidden behind each other or overlap in-depth and even though interacting with the visualisation (by zooming in or filtering) can resolve the issue, this costs time and puts a cognitive load on the viewer. 2D is almost always a better choice for charts, graphs and text. Some spatial data can be transformed to 2D and other views for better data abstraction.
- The rule of "No unjustified 2D" is similar to the first rule, but for 2D visualisations. Spatial data and topological structures and contexts are preferred in 3D.
- The rule "Eyes beat memory" states that external cognition, for example giving a comparison with side-by-side views, is easier to comprehend than comparing a visible item to memory.
- Giving visual feedback during processes, for instance, showing a progress bar for long operations and an hourglass for multi-second operations is suggested in the rule "Responsiveness is required".

- "Function first, form next" suggests starting with a focus on functionality, rather than aesthetics, which can be perfected later.
- Munzner also points out some basic graphic design principles, such as contrast, proximity, and repetition. If two items are not identical, they should have very contrasting colours (e.g. black and white). Items related to one another should be grouped and any equal whitespace between unrelated visualisations should be avoided, as it creates a sense of connectivity. If any data is repeated, this consistency should be stressed and unified.
- Another best practice to follow is labelling a visualisation to make it as self-documenting as possible. Axes should have labels and all plotted elements a legend. Titles, labels and legends should be meaningful.

The KISS principle brings focus to simplicity. It originates from a mission the U.S. Navy carried out in 1960, named "Project KISS" [25]. According to the principle, most systems work best when simplicity is the key aim. Simple interfaces increase ease of use and thus, increase user acceptance.

Shneiderman's Visual Information Seeking Mantra, a widely practised principle for creating effective user displays, is expressed as "Overview first, zoom and filter, then details-on-demand" [23]. The idea is to construct a visualisation out of layers, where the first is the general overview (or summary) layer for context and presenting major components and their interrelationships [26]. The second layer is for zooming in and filtering to reduce the complexity and organise the visualisation. The last layer, details-on-demand, is for inquiring about more specific data with a simple interaction such as a mouse click or mouseover.

As the topic of this paper is concerned with persuasiveness, the following is an overview of a study about making visualisations persuasive. Fang et al. [24] stated that persuasive strategy visualisations must fit into four dimensions for better persuadability. The four dimensions can be applied separately, but they can also overlap. The dimensions are as follows:

- **Ambient.** The degree to which the persuasive system calls the attention of a user. Most behaviour change systems display information in a way that does not need much attention on the user's part and hence, is not interruptive.
- **Aesthetic. T**he degree to which the system provides visual comfort and is attractive. For sustained system use, aesthetic appearance is important.
- **Emotionally engaging**. The degree to which the system elicits an emotional connection to the user. Emotionally engaged users are more likely to change behaviour according to the system.
- **Metaphorical.** A metaphor is the understanding of an idea in terms of another. It is applied in user interface design and can be applied to behaviour change systems for better persuasiveness. An example of this could be the UbiFit mobile system, in which a user's physical activity progress is displayed as a digital garden, which encourages the user to work out more to make the garden bloom [27].

2.3 Persuasion Research

Human behaviour and attitudes have been studied extensively [10] and the theories and models developed as a result of that research have been used in persuasion studies. Coercion and betrayal are not included in persuasion research for ethical reasons.

Fogg identified three roles persuasive technology can play: it can be a tool, medium or social actor [11]. As a tool, it makes actions easier or more efficient. As media, technology can offer believable simulated experiences and thus, change behaviour and attitudes. As a social actor, technology implements social cues to persuade the user. For each role, Fogg defined certain design principles.

When technology is filling the social actor role, it can leverage five types of cues to persuade: physical, psychological, language, social dynamics and social roles. Physical attractiveness via simple and aesthetic interfaces and likeness to humans can help the system seem more persuasive. Psychological cues make it seem the computer has a consciousness: emotions, preferences, motivations and personality. The effect is even greater if the computer seems like-minded or similar to the user in some way. Expressing a personality and praising the user are ways to persuade the user through language. Following along with social dynamics, the unwritten rules for interacting with others, can help make the system seem more human-like and trustworthy. A special social dynamic that could be leveraged is the principle of reciprocity, wherein if a system has done something for the user, they will feel the need to reciprocate the favour. Finally, the system can present itself in a social role as a teacher, judge or expert to give it the power of authority.

The Elaboration Likelihood Model (ELM), as described by Petty and Cacioppo [28], is one of the most established theoretical models of persuasion. It outlines two routes to persuasion: a thoroughly constructed central route, which encourages thoughtful consideration of the argument by the user, and the peripheral route, which uses cognitive shorthands to persuade more simply. Either route can be effective to the same degree, although the central route can bring about a more persistent behaviour change.

In ELM, an attitude is defined as the evaluations a person has regarding themselves, others, objects and issues. The model measures attitude change in several ways, which are all based on seven postulates. These postulates describe the natural inclinations in behaviour, e.g. we are motivated to hold correct attitudes, and every person has a different level of willingness to elaborate on a persuasive message. Elaboration is defined as “the process of relating the to-be-evaluated recommendation and arguments to other issue-relevant information in memory” and usually results in self-generated attitudes towards the given information. It is influenced by many different factors like motivation, ability and opportunity. When the likelihood of elaboration is low, a person will not be motivated to give thorough thought to an argument and therefore persuasion occurs via the peripheral route and vice versa.

According to ELM, elaboration is also influenced by variables. A variable can either increase or decrease the persuasiveness of an argument. Some examples of variables are expertise, mood and attractiveness. With a high elaboration likelihood, a variable can become an argument itself, and with a low likelihood, it can serve as a peripheral cue or as a cognitive shortcut.

In addition to explaining attitudes and persuasion, Petty and Cacioppo also outline how to assess them. A scaled system, for example the Likert scale, is often used, which helps gauge how much a person likes or dislikes a certain stimulus. Scales can be semantically differential, i.e. a series of adjectives are presented with their antonyms, or single-item. In some cases, to avoid social desirability from interfering with test results, a cover story might be used, such as telling a subject the test is for another goal entirely and as their own beliefs might influence the way they answer, their opinions must be recorded. This allows the actual persuasive test to be veiled as a simple opinion quiz.

According to ELM, influencing attitudes can be achieved by varying the argument's quality, employing peripheral cues and by determining the direction or extent of argument processing. The strength of an argument determines whether it can provide a consistent behaviour change. A weak argument will fundamentally generate unwished results, no matter the elaboration likelihood. A peripheral cue, as mentioned before, is a simple variable that can affect attitudes because it triggers primitive states that become associated with the object under argumentation. Variables can direct message processing in a relatively objective or biased manner. To test these three constructs of ELM, Petty and Cacioppo have used five different procedures:

- self-reports of effort, in which the subject is asked how much effort they made to understand the message, e.g. "To what extent were you trying to evaluate the message?"
- argument recall, in which the subject is asked to recall the persuasive arguments
- thought listing, in which the subject is asked to think out loud during the processing of the persuasive message
- electrophysiological activity monitoring, where EMG activity is recorded
- argument quality manipulation, in which the argument's quality is either strong or weak and the effect it has on processing is recorded

Oinas-Kukkonen and Harjumaa developed a method called Persuasive System Design (PSD) for designing behaviour change support systems (BCSSs) [29], [30]. A BCSS is defined as "a socio-technical information system with psychological and behavioural outcomes designed to form, alter or reinforce attitudes, behaviours or an act of complying without using coercion or deception". The PSD method consists of three steps, where the first step is understanding seven postulates concerning persuasive systems. The postulates are derived from the seven postulates defined in the ELM model. They are as follows:

- Information technology is never neutral.
- People like their views about the world to be organised and consistent.
- Direct and indirect routes are key persuasion strategies.
- Persuasion is often incremental.
- Persuasion through persuasive systems should always be open.
- Persuasive systems should aim at unobtrusiveness.
- Persuasive systems should aim at being both useful and easy to use.

The second step is to analyse the context within which the persuasion takes place: who is the persuader and the persuadee, what is the message conveyed and through which channel, as well as what is the larger context. More generally, this analysis is categorised into three parts:

- The intent. One must determine who is the persuader, as a computer does not have an intention of its own. Recommender systems covered in this paper have an endogenous or exogenous intention, respectively meaning the intention of those who create or produce the technology or those who distribute the technology. One must also analyse the change type aimed for, particularly whether it's an attitude and/or behaviour change that the behaviour change intervention should create. A one-time behaviour change is the easiest, a permanent behaviour change more challenging and an attitude change may be the most difficult to achieve.
- The event. One must consider the problem domain, the user context and technology. An important part of the user context is the user's goals.
- The strategy. One must choose and/or recognize suitable persuasion strategies for the current system. Another question is what route the strategy will be delivered through: a direct or indirect one. A direct route will require a more in-depth evaluation of the strategy on the user's part, while the indirect route means the user will use stereotypes and simple cues to quickly take in the information.

The third step in the PSD method implements Fogg's design principles as well as new principles which aim to improve persuasiveness. In total, 28 persuasive design principles are listed. A persuasive strategy can either be for primary task support (features that support the target behaviour), dialogue support (feedback the system guides the user with), increasing perceived system credibility (features that help perceive the system's expertise) or for social influence (features that create motivation through social influence). The principles are:

- Reduction. Reducing complex behaviour into simpler tasks or steps, which help the user achieve the target behaviour
- Tunnelling.
- Tunneling. Guiding the user through a process and persuading them along the way
- Tailoring. Tailoring the system's information given according to the user, usage context, or other factors
- Personalisation. Personalising the system's content or services provided for the user
- Self-monitoring. The system tracks the user's performance for them, helping them achieve their goals
- Simulation. Simulating possible before and after situations
- Rehearsal. Providing a way for the user to rehearse a behaviour
- Praise. Offering praise and therefore making the user more open to persuasion
- Rewards. Rewarding target behaviours
- Reminders. Reminding the user of a target behaviour
- Suggestion. Offering fitting instructions, actions, ideas etc.
- Similarity. Similarity of the system to the user or something else meaningful
- Liking. Visual attractiveness of the system
- Social role. The system adopts a social role
- Trustworthiness. A seemingly trustworthy system can be more persuasive
- Expertise. A seemingly expert system can be more persuasive

- Surface credibility. The appearance of competency and credibility can be more persuasive
- Authority. A seemingly authoritative system can be more persuasive
- Real-world feel. Naming the organisation or people behind the system
- Third-party endorsements. Displaying endorsements, especially from reputable sources
- Verifiability. Allowing users to verify the system's content through outside sources
- Social learning. Allowing users to observe others performing a target behaviour
- Social comparison. Allowing users to compare their performance with others
- Normative influence. Making use of peer pressure to increase the likelihood of a target behaviour being performed
- Social facilitation. Allowing users to see that other users are performing the behaviour with them
- Cooperation. Motivating users by allowing cooperation
- Competition. Motivating users by allowing competition
- Recognition. Offering public recognition for a user or group

Al-Ramahi et al. [31] researched persuasive design principles by doing a systematic analysis of user feedback, differentiating them from other persuasive frameworks, which followed experts' opinions. They compared the gathered principles to commonly used principles described in Oinas-Kukkonen and Harjumaa's Persuasive Systems Design framework [30]. They defined three persuasive design dimensions. The first was *structural*, which involves connecting supporting elements in the user's context, for example connecting the user's app with external devices such as insulin pumps and external supporting actors like doctors. The second dimension, *technical*, consists of the primary functionalities of a behaviour change system, such as self-monitoring, informative presentation, effort expectancy and persuasive messages. The last dimension, *social*, involves a community of peers that can offer social support to the user via a forum.

Marcu et al. [32] reviewed the literature on bipolar disorder and developed a framework called the Patient-Clinician-Designer (PCD) Framework. The framework can be utilised for designing persuasive illness management and monitoring devices. The design is based on user-centric values and is addressed to, as the name suggests, the patients, clinicians and designers of the illness monitoring system. The framework combines user-centred design principles with persuasion research. For every phase of the user-centric design process, they proposed questions that all three shareholders of the system should ask themselves. In addition to questions focused on bipolar disorder and its treatment, they also gave many examples of general questions for the designers of behaviour change systems.

The questions were concerned with three goals: adoption, acceptance and sustained use of the system. For designers, the main questions to address from the user's perspective to achieve the three goals are the following:

- What do users and others gain from using the system?
- What will motivate them to use the system?

- What can users learn about themselves, and what can they and others do with this knowledge?
- Will the system fit into the daily lives of users?
- Is data collection nonintrusive and convenient enough to sustain regular use of the system?
- Can the system provide evidence for a suggestion? The system must adhere to established business practices to be incorporated into the daily practice.
- Does it potentially impose additional work on the users' part? It is important to consider how the system integrates and works with the existing systems infrastructure.

The authors also put forward questions to be addressed from the point of view of the business. The questions help designers guarantee motivation, benefit and the cost of the system. The questions are:

- What is the motivation for building the system?
- What happens to the data, who owns it, what it can be used for, and in what format it is in?
- What is its cost-benefit trade-off? The motivation for the designers of the system is related to the existence of a clear and evidence-based "business case" for the system; i.e., is there clear evidence that a solution potentially will help patients in their treatment, and what other related and competing solutions already exist?

As an evaluation method, they developed the MONARCA monitoring system using the PCD framework. They implemented metaphors, feedback, data sharing and visualisation as persuasive elements. Although the system is mental illness oriented, the questions proposed in the framework can be applied in other persuasive fields.

The health recommender framework is a persuasive system design framework focusing on health promotion [33]. This framework follows the guidelines set by Fogg in his behaviour model and borrows from multiple different behaviour theories such as the Goal Setting Theory, theory of reasoned action / planned behaviour etc. They analysed existing principles and divided them according to if they applied for the user or the system. The user aspects are about their beliefs, status, goals, perceptions and intentions and the system aspects are about the content of recommendations and the persuasive elements it contains. For a persuasive health system to be successful, all these elements must be addressed.

The health recommender framework defines requirements for both the user and system and their aspects. For example, for the user, some requirements were "Users should receive information about the health risks and the seriousness of the threats", "Users should receive the information about the health benefits", "The difficulty of the recommendation items should increase automatically and step by step so that it enhances self-efficacy and users' ability" and "Recommendations should be socially appropriate for the user". For the system, some examples include "The relevant recommendation items (i.e. interventions) should be repeated and practiced in order to make it a natural habit for stressful situations", "One should therefore minimize the barriers of following the recommendation" and "the system should enable users to track their own behaviours and observe the outcomes." The health recommender framework

authors created a mobile prototype based on the principles they defined. The prototype's goal was to reduce stress in the user by recommending stress-coping activities (e.g. physical activity and positive thinking).

In conclusion, the field of PrBMP is still relatively new. PrBPM aims to improve business processes by providing prescriptions generated by complex AI models. To visually present these prescriptions, many visualisation techniques and principles can be applied. Some of these describe which elements to use in what context, some how to approach the designing of these elements and others cover the colour and placement of them. Visualisations can also be designed with persuasion in mind by utilising ambient, aesthetic, emotionally-engaging or metaphorical qualities. Similar to visualisation, many frameworks and techniques exist in persuasion research, which can be applied to a business process context. These works describe how attitudes and behaviours form and how to best influence them.

3. Related Work

Persuasion has been the subject of many studies in psychology [10], but not much has been done in the field of persuasive visualisation. The complete research on persuasive psychology is beyond the scope of this paper, but in the following section, an overview of research on persuasion relevant to visualisation is given (section 3.1). Also, research on persuasive recommender systems within business processes is described (section 3.2).

3.1 Research on Persuasive Visualisation

Pandey et al. [13] tested the visualisation of different topics with tables and charts and found that presentation type may influence persuasion. With a pre-questionnaire, they categorised participants per topic according to the Likert scale: they were either negatively polarised (NP), neutral/weakly polarised (NWP) or positively polarised (PP) towards a given topic. Each participant was presented a web page with a persuasive message and a supporting graph or chart and table per topic. The persuasive message contained a claim, contextual information and the evidence supporting the claim (e.g. "during this time X increased and Y decreased"), which was based on the data being visualised. After showing the message to the participant, they asked each to what degree they agreed with the claim. They used the ELM model to calculate the persuasive power of each presentation.

They discovered that using tables to display information had an increased probability of creating attitude change for participants who were NP towards a topic. For NWP and PP participants, charts had the highest persuasion likelihood. This research, however, was just limited to the persuasiveness of a single general message and its visualisation not an entire persuasive system with recommendations. The focus of this thesis is on persuasive prescriptive business processes, within which tables, graphs and charts can be used to encourage trust in the system.

Oh et al. [14] studied why data visualisation influences user engagement and proposed interactivity as an underlying mechanism. In order to empirically test this theory, the authors visualised information about obesity and physical inactivity and combined interactivity literature and narrative processing theories[1]. The interactivity of visualisations was defined into two separate levels: modality and message interactivity. Modality interactivity is the interactivity degree of the medium, such as a website, where features enabling manipulation could be, for example, sliders, drags, or mouse-overs. Message interactivity refers to the interactivity degree of the data visualisation, where features could be, for example, click, mouseover, slide, zoom-in, or swipe. In the study, modality interactivity was operationalized as a variety of interaction techniques, such as a slider bar for controlling a heat map, a clickable heat map, and a clickable line graph. As interactive visualisation alone only carries descriptive and collective information, a textual concrete message was added with the use of narrative persuasion. As a test control, a low interactivity webpage was shown.

The study found that high modality interactivity in data visualisation enhanced user engagement and changed the users' perception and attitudes pertaining to the obesity content. The authors suggest that the high interactivity may have distracted users from fully concentrating

[1] The reference for the test webpage can be found at https://stateofchildhoodobesity.org/adult-obesity/

on the persuasive message shown, in turn reducing their defensiveness, and instead pulled their focus to the interface. The results also indicated that data visualisation has more impact when paired with a narrative, but only so if both simultaneously have high interactivity.

The authors of this paper studied the persuasiveness of visualisation and narrative in a general sense. The persuasiveness of modality and message interactivity can be a relevant input for this thesis, however, this paper is less concerned with the modality and more with the messaging of persuasive interventions.

3.2 PrBPM and Explainability

Although the persuasiveness of PrBPM methods has not been explicitly studied, trustworthiness of the system has been a point of interest in a number of papers. Namely, these works focus on solving the issue of AI black boxes as the black-box nature of PrBPM applications suppresses their operationalization in the industry. The opaqueness of deep learning approaches has already been a larger focus in other fields of research and leveraging this, a recent trend of explainable AI (XAI) has emerged in PrBPM research. The following is a brief overview of such papers.

Galanti et al. [34] tackled the issue of explainable predictive business process monitoring. The created framework is independent of the underlying machine- or deep-learning technique and implements the state of the art of explainable AI. The explanation is presented with regards to a target KPI either at run-time or offline to create an understanding of how the trained model makes predictions. For each case, an explanation would include the set of attributes that influence a prediction the most and how each one separately increases or decreases a KPI. The explanations are presented in the form of a heatmap and a table of attributes with respective values that influence an event the most (e.g. “CLOSURE TYPE=Porting” and “CLOSURE REASON!=1 - Client lost”). The explanations were generated with Shapley Values, a game theory approach that generates local explanations.

Farrugia et al. [35] leveraged different machine learning techniques to find fraudulent users in the iGaming industry. They implemented the latest XAI research advancements to present which attributes contributed most to determining whether a user was a fraud or not. The explanation would help proactively find frauds in the industry and take action against them.

Mehdiyev and Fettke [9] described a framework for explainable process predictions. Despite many XAI solutions being proposed over the last decades, the user’s needs for understandability have not been fulfilled. Many circumstances such as the decision-making environment and the user’s personality can affect the explainability of a system. In order to enable the user to comprehend the recommendation generation process, the authors posit five elements that comprise the framework for explainable process predictions: subject, objectives, techniques, generation time, and outcomes. An explanation has to cater to subjects (or stakeholders), fulfil different objectives, examples of which are transparency, effectiveness, persuasiveness, satisfaction etc., employ model-specific or model-agnostic techniques for either global (e.g. PDP, global surrogate models, SHAP dependent plots, Accumulated Local Effects or ALE Plots) or local (for instance Expectation or ICE Plots, Shapley Additive Explanations or SHAP, LIME) explanation approaches, be generated either before, during or after the model building and,

finally, it can be used to provide information in terms of different outcomes and model robustness. In addition to a framework, the authors proposed a post-hoc local explanation approach. The predictions the interventions depend on are generated with a deep learning model. For explanations, they implemented a novel algorithm similar to K-LIME.

Mehdiyev and Fettke, together with Rehse [36], studied the opportunities and challenges of process predictions and explainable AI. For this, they used a fully automated LEGO brick factory and predicted possible outcomes of the factory's processes. The predictions, which are made by a deep LSTM model, are made understandable for the workers with model, local and textual explanations. The model explanations present the most important features that can influence a prediction and the model's metrics (for example accuracy, precision, recall etc.). The local explanations show rules which most determined the current prediction and which feature contributed the most to it. A textual explanation is added for a more detailed look into the rules and feature contributions. Textual explanations are generated semi-automatically with pre-defined templates.

All four aforementioned papers took a technical approach to explainability in descriptive systems and aimed for explaining a system to, among other things, make the system more persuasive. However, this thesis aims at making a prescription more understandable for the process worker, explanations being just one of the applied techniques. The explanation's semantic content and persuasiveness are also not studied by any of the papers, consequently requiring a considerable amount of effort to be understood. We build upon the existing research by applying persuasive principles and visualisation techniques in addition to explanations.

4. Methodology

This section describes the research process of the thesis and details how data was collected and analysed in order to answer the research question. First, the chosen methodology is described. Then an overview is given of how the persuasive principles were extracted from literature (section 4.1), how requirements elicitation was done (section 4.2) and how creating the PERSEVERE prototype was approached (section 4.3). Finally, evaluation is explained (section 4.4).

This research aims to understand how to present prescriptive business processes for user acceptability and persuadability. To address the research question posited in this paper, *How can prescriptive business process monitoring be visualised in a persuasive manner?*, we followed the Design science methodology as described by Geerts, L [15]. This methodology fits our purposes as the goal is to solve an issue (visualising prescriptive business processes) by creating and evaluating an artefact (a prototype based on generated principles). As stated in the Design Science guidelines, first the problem is identified and its solution motivated, after which the objectives of the solution are defined. Then the solution is designed and developed, which is followed by a demonstration and evaluation. Finally, the result is communicated.

As stated before, the first step of the Design Science methodology is to identify the research problem and justify the solution's value. The field of prescriptive business processes is novel and research on the topic of persuasive system visualisation is underdeveloped. In addition, this topic can be used in business process improvement to maximise the value of recommender systems in business processes. The developed solution can guide the team in carrying out process improvement and can be an addition to the studies of persuasive visualisation.

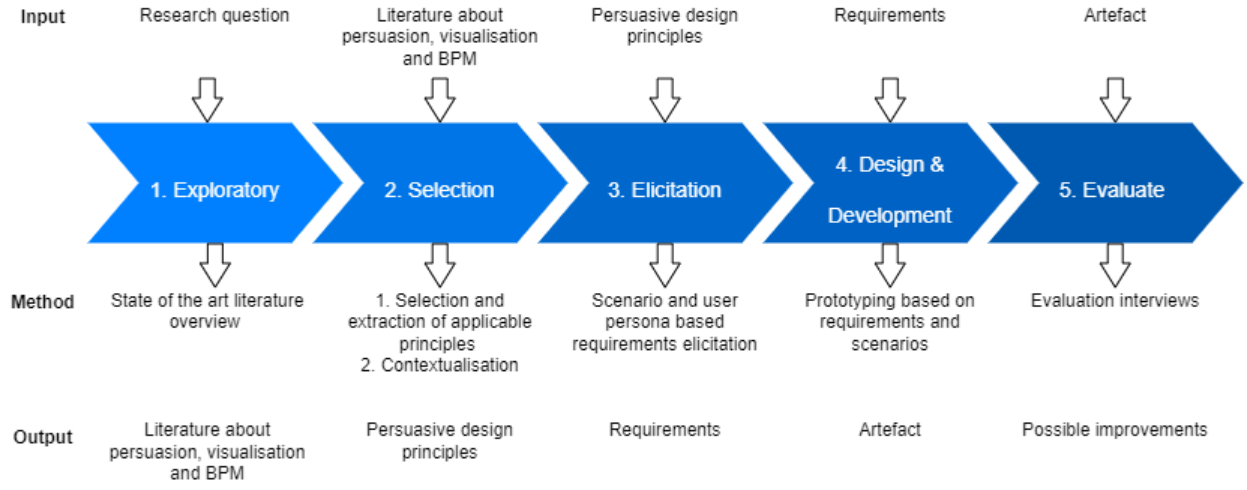

Figure 1. Research process

The rest of the steps in the Design Science framework were divided into five phases for this thesis and were numbered from 1 to 5. These steps are presented in Figure 1 and are explained in further detail below.

4.1 Persuasive Principles

We based the requirements and development of PERSEVERE on visualisation and persuasive technology strategies and principles gathered from literature. This research was done in the first, *exploratory* phase. Guided by the research question of the thesis we searched the state of the art literature in business process management (BPM) (section 2.1), visualisation (section 2.2) and persuasion (section 2.3) and gathered any suitable works from IEEExplore[2] and the ACM digital library[3]. More specifically we focused on works containing the keywords "recommender" or "prescriptive systems", "acceptability", "persuasion", "persuasive visualisation" and "trust". During searching the publication date was set to be later than the year 2009 to gather the most up to date literature. When a suitable paper was found, we searched its references for other similar works with no date limit.

As a result of the exploratory phase, 40 papers were gathered. We removed papers that were not written in English and any that did not contain sufficient information about the methodology or evaluation. Papers that were included had to answer the research question of the thesis in some capacity, e.g. discuss persuasive visualisation, prescriptive recommender systems in businesses or persuasion in human-computer interaction. Subsequently, not all papers were suitable for the context of this thesis or covered prescriptive systems or business processes. Therefore, in the following phase analysis would be required to select any suitable principles or frameworks from them.

We selected three frameworks to serve as a larger basis for the logic of the principles (section 2.3). The PSD framework is the state-of-the-art framework for designing behaviour change systems and has been applied in many different fields and thus, here we utilised it for PrBPM as well [12]. Marcu et al. [32], in their Patient-Clinician-Designer (PCD) framework, proposed many user-centric questions from the user's perspective be taken into account by designers of persuasive mental illness applications. These questions, if generalised, help generate active user involvement in a behaviour change system when used during the design phase and hence provide useful input for the main research question of this thesis. The last framework we selected was the health recommender framework [33]. This framework describes requirements for a persuasive health recommender system. Although many of the requirements are derived from Fogg's model, there are some that are unique compared to the rest covered in this thesis. These requirements can be selected and contextualised.

What followed was generalisation and contextualisation of the principles to fit the purpose of the thesis, i.e. the domain-specific principles were changed to apply to PrBPM. For instance, we transformed a question from the PCD framework "Can the system provide evidence for a suggestion?" into the principle PE1 *The system should provide textual or numerical explanations for an intervention*. Another example is the 1st postulate of the PSD framework, which states that persuasion is a process, rather than a single act, and this was utilised to create the principle PE29 *Persuasion should be incremental and adaptable to the business process and its workers and goals.* All this was done in the *selection* phase. This process and its resulting

[2] https://ieeexplore.ieee.org/

[3] https://dl.acm.org/

set of persuasive design principles are outlined in section 5. The principles were further grouped into categories based on if they describe how to display an intervention (section 5.1), what content an intervention should contain (section 5.2) or when an intervention should be presented (section 5.3).

4.2 Requirements Elicitation

According to the Design Science framework steps, the persuasive principles needed to be evaluated. For this, a prototype is necessary. We collected requirements for the prototype in the *elicitation* phase. During this phase we chose a setting and created a user persona, on which we based two scenarios. The setting we chose was a bank loan business process, as we had the aid of a domain expert with 30 years of experience in this field, had contacts in an Estonian bank for finding interviewees and had access to an event log consisting of loan application data. The persona was used as the actor in requirements. The scenarios were utilised in requirement elicitation and prioritisation. Both are also used in the next phase, the evaluation. In the following paragraphs we will discuss the creation of user personas, scenarios and requirements in more detail.

4.2.1 User Persona and Scenarios

First, we created a user persona, as it aids in keeping the focus of the requirements on the end-user and their needs [37]. The persona is a generalisation of middle-level loan process workers, allowing us to apply it to as many users as possible. The creation of the user persona was informed by the domain expert. The persona can be found in section 6.1.

Some of the data used in this phase was sourced from the Dhana bank loan application handling process log[4]. The log details all offers for accepted loan applications of a Dutch bank from 2016 till February 1st of 2017. For each loan application, details such as the requested amount, application type, reason and ID are available. The log also describes all offers with details: an ID, offered amount, first withdrawal amount, monthly repayment cost, number of payback terms, the customer's credit score, the employee's info who created the offer, whether the offer was selected and if it was accepted. One application can have multiple offers, but only one can be accepted. An application can have one of many states:

- application created – an application has been created via the bank's website
- submitted – a customer has submitted an application via the bank's website
- complete – one or multiple offers have been sent to the customer and the bank is waiting for a returned signed offer with other documents
- concept – an application has just been submitted and first assessment has been done automatically
- accepted – application is completed, assessed and an offer can be made
- complete – offers have been sent and the bank is waiting customer's response
- validating – an offer and documents have been received and are being validated
- incomplete – documents are not correct or are missing

[4] The process log is available here

- pending – all documents have been received and the assessment is positive
- denied – a loan cannot be offered
- cancelled – the loan is cancelled

The persona is accompanied by two scenarios that outline possible PrBPM system uses during a normal business process flow (section 6.2). The possible prescription outputs (the best next action and the best resource for achieving a goal) are given in section 2.1. We formed the scenarios with the aid of the domain expert mentioned before. Both scenarios include loan applications selected semi-randomly from the Dhana bank process log. The imaginary bank the scenario takes place in was named after the process log.

In the first scenario we chose the application's goal to be a car lease, since middle-level loan officers work with smaller mass products such as car loans most commonly. For the second scenario, we chose a mortgage as these types of bigger loans usually are handled by multiple loan officers and are less common for middle-level workers to have in their portfolio. Once a loan goal was selected, we searched the process log for a random instance of such an application, which was then used as the example application in the scenario.

4.2.2 Requirements

We developed a set of requirements based on the persuasive principles (section 6.3). The user persona was utilised as the user's perspective for the requirements. First, we transferred the persuasive principles to affinity notes. Then, combined with situations described in the scenarios, we analysed every note and then added new notes as requirements. Some principles were broken apart into multiple requirements and sub-requirements, some retained a one-to-one relationship, while others were combined into one. For example, the principle PE4. *Visualisations should follow the principles of good visualisation so they can enable different levels of understanding of the data and avoid information overload* was broken down into 10 requirements, one of which is R.D.3.4 *The worker can toggle accordions for explanations and visualisations*, while the requirement PE2 *For a higher likelihood of persuasion, in addition to messages and explanations, the system should provide visual evidence for an intervention* was directly transformed into one requirement only: R.E.2 *Explanations contain visualisations*. Next, we organised the requirements with the affinity diagram method [38]. By using affinity notes, this method allows for quick organization of data as the notes can be easily grouped by similarity. As a result, categorisations of requirements emerged: Design of the System, Flow of the System, Explanation of the Prescription and Text.

The affinity note method allowed for requirements to be easily prioritised as well (section 6.3.1). We grouped the notes and analysed them according to the MoSCoW method, in which requirements are prioritised in the order of importance of implementation. The categories of priority are the *must-haves*, *should-haves*, *could-haves* and *won't-haves*. The aforementioned scenarios were utilised as a basis for prioritisation.

For must-haves, we selected the supportive requirements for the core functionality of PERSEVERE, for instance presenting evidence in support of a prescription. In addition, the degree of

how much the requirement would alleviate the black box problem was another factor for selection, since much of the trustworthiness of an AI system is lost in its opaque nature. An example must-have requirement is R.E.2. *Explanations contain visualisations.*

When choosing requirements for the should-haves, we included requirements that are a smaller support to the main activity of the process workers (for instance R.T.8 *The system displays the total number of possible prescriptions or analysed cases*). As some of these requirements can be more complex to implement compared to others, they cannot be part of the must-have selection, but they can be a large persuasive influence and therefore should be implemented if possible.

We selected other requirements by comparing their suitability in either the should- or could-have categories, as the most trivial and smaller tasks naturally fell into the latter. For example, the requirement R.T.4 *The worker receives rewards, e.g. tags, badges or titles, for interactions with the system* is one that doesn't have a strong guarantee of benefit because in some fields it isn't allowed or it may not be suitable in a business environment and hence was moved to the could-have category. Those requirements that have no value for PERSEVERE or were not possible to implement were moved under the won't-have category, for instance R.D.8. *The system employs visualisations utilised in other applications in use in the business.*

4.3 PERSEVERE

During the *design and development* phase, we relied on the requirements to develop the PERSEVERE prototype in Figma[5], a free popular prototyping tool. This is detailed in section 7. This thesis is only concerned with the visual representation of a persuasive recommender system and as such, PERSEVERE is a mock front end of a loan application system and the recommendations themselves are static, i.e. they are not a result of a prescriptive model. The prescriptive model is a combination of a mock decision tree, since according to a domain expert this is the most common model behind loan application engines, as well as a mock neural network model accompanied by XAI methods, which can output complex analytics about a prescription.

The user interface of PERSEVERE had to have a certain structure to be persuasive. We designed a generally applicable persuasive layout based on the prioritised requirements, which we then applied to all modals (section 7.1). A common layout allows for the two recommendations to be compared. The layout also ensures the most crucial requirements are present and in the most beneficial order for the system to be persuasive. The most important requirements were placed so that they can be seen in the first glance after the system has opened. The less important an element is or the more negative impact it can have on the recommendation being accepted, the less visible it is. The layout includes conditional content so that it can be modified according to the needs of any system.

PERSEVERE was divided into screens and overlays as that is how Figma is designed to be used. The screens are a variation of one desktop-sized view: the loan overview. Different states of information are contained within every screen that change according to the user's actions.

[5] https://www.figma.com/

The screens (section 7.2), however, were mostly added for immersion and therefore contain the most non-persuasive elements. We placed the persuasive elements within overlays such as modals, alerts and tooltips (section 7.3). Overlays can be triggered on the screens via buttons or by hovering over elements. As they are interactive, they fulfilled multiple requirements that pertain to the flow of the prototype that a regular screen wouldn't be able to.

We designed PERSEVERE's user flows and setting such that in comparison with a real loan process without a recommendation component, the prototype would not be more complex, since that would deter users from using the system. In addition, the persuasion should not be strongly polarizing and hence should be realistic and believable. The baseline assumption is that the prototype is useful and the recommendations shown are acceptable, since a recommendation would not be shown otherwise.

4.4 Evaluation

This section gives an overview of how the evaluation user study was designed and how data was collected and analysed to evaluate the prototype. This is the *evaluation and refining* phase, in which practitioners in the field of loan processes were interviewed and surveyed to confirm PERSEVERE fulfilled its purpose.

4.4.1 Data Collection

As the literature on the impact of evidence presentation in prescriptive business processes is limited, a strategy is necessary for evaluating the prototype outlined in section 7. Some key questions we were confronted with when approaching a testing strategy were "How does a persuasive tactic impact the user's decision? What element is the most persuasive? How to measure persuasion?". In order to answer these questions a user study can be used [39]. Study goals were derived from the research question of this thesis:

1. How did the participants use the interface?
2. How did the persuasive elements impact the decision process?
3. What improvements can be made to the interface?

When selecting the participants, we were guided by the user persona, as it describes the most general potential users of the system. The target users were domain experts who have experience with leasing and loan applications. All participants had to be employees of a selected Estonian bank, since different banks have different rules and analytics for evaluating a loan application. This restriction was set to ensure the experience of the practitioners would be as similar to one another as possible, so the interview results would be easier to compare. Due to these limits, the subject pool was quite small and we added no further restrictions. Both junior and senior employees were selected to participate so different levels of experience would be represented. In total, eight field practitioners participated in the study and their information can be seen below (Table 1).

Table 1. PERSEVERE evaluation study participants

Code	Most common loan type	Experience	Age	Gender	Position
I1	Consumer loans, credit cards, mortgage loans	1 year	21	Female	Loan specialist
I2	Consumer loans, home renovations, car loans, credit card loans, travel loans	1.5 months	21	Female	Loan specialist
I3	Consumer loans, small loans, credit cards	1 year	24	Male	Loan specialist
I4	All loan products the bank offers	10 years	51	Female	Solicitor
I5	Unsecured loans, leasing, mortgage loans	8 years	54	Female	Deputy head of retail banking
I6	Small business loans, mortgages	21 years	43	Male	Credit manager
I7	Mortgages	20+ years	50	Female	Team lead
I8	Consumer and business loans, mortgages, small loans	20+ years	50	Male	Credit risk senior expert

We conducted online interviews individually with the participants using the Microsoft Teams software. The duration of the individual studies was between 41 and 60 minutes. The user study was structured as follows:

1. Topic description. The participant was provided with a neutral and short description of the topic of the study without mentioning the persuasive aspect
2. User persona and setting description. The participant was asked to imagine themselves as the user persona and their work place, Dhana bank, was described to them
3. User interaction with Figma. The participant was presented with PERSEVERE and asked to complete the two recommendation flows to make a decision based on the presented information
4. Interview. The participant was interviewed for their thoughts on PERSEVERE
5. Questionnaire. The subject's attitudes and thoughts towards PERSEVERE were recorded after the persuasive treatment had been done.

Methods used in steps 3, 4 and 5 were all informed by the ELM model [28]. It has been used in numerous behaviour change studies, which is why we selected and structured the study methods with it in mind [13], [40]–[42]. ELM, like other attitude change theoretical frameworks, aims to document changes in attitude when a person is exposed to a persuasive message.

During the second step of the study, the following introduction to the persona was given to the participant: "*Imagine you are a loan officer called Anne who works for Dhana. Dhana is a large international bank and you work in a team of loan officers. You have mainly worked on*

small and medium standard credit cases. You have multiple cases this week to work on. Currently, you are logged into the internal web application of the bank and you are viewing the loan overview screen, where you manage your assigned loans. You can see loan application cases with their IDs and additional information can be seen when you interact with the prototype." During the third step we conveyed the details of the first scenario to the user. After the user had finished the first scenario, the second scenario was introduced in the same way. The user was allowed to freely interact with the prototype with limited guidance. Due to the study goals, the usability of PERSEVERE was of interest and therefore having as little guidance as possible was another way of measuring it. We directed any participants to the prescriptions, if necessary, but that was the limit of the aid given.

During the decision process the participants were encouraged to take their time, explore the overlays and when ready, accept or reject the recommendation. During the demonstration of PERSEVERE, we recorded the participant's screen for more context for the answers they would give during the following interview and questionnaire. We also asked them to think aloud. Thinking aloud or thought listing is an ELM model method of assessing persuasion. This approach is beneficial for answering the first two study questions *How did the participants use the interface?* and *How did the persuasive elements impact the decision process?*

We chose interviews to be the main method of the study as the study goals were to thoroughly understand the subject's reasoning and expert knowledge, and any possible improvements that could be made to the interface [43]. The interviews would also help assess the subject's argument recall, which is one of the persuasion assessment methods described in the ELM model [28]. The duration of the interviews was between 23 and 52 minutes.

A semi-structured interview method was chosen due to the qualitative and exploratory nature of the study goals [44]. A semi-structured interview allows for questions to be predetermined, as well as spontaneous ones to be included during the interview according to need. We transformed each study question into multiple interview questions, which we then organized by the visual blocks of the recommendation modals. These were then divided into sub-questions and probes that help collect detailed information. A probe is a situational question, while a sub-question is a question that should be asked in any situation. To avoid participants forgetting what they had seen, we organised the interview questions in the order of the user flow and showed the user interface to them during questioning.

Since the aim is to understand the user's decisions, we first asked each subject to reiterate their decision process. If this method or their previous thought-listing had not covered the most important factor for the decision, then we inquired about it explicitly: "*Was there an element that was the deciding point for you?*" Then each visual block of the prototype was discussed from top to bottom with the goal of determining the useful, not useful and confusing part of each block, e.g. "*How did you use this block for your decision? Was there something unclear about this block?*" We encouraged the participants to offer suggestions for improving each block. For the block containing an XAI explanation, the user was also asked if it made the recommendation more trustworthy or less and why. If the subject did not notice a part of the modal (for instance the collapsible elements, tooltips), then we would show this during the interview and ask questions such as "*Now that you have seen this element, how would this have affected your*

decision?". After covering the visual blocks individually, we then asked the participant about their overall decision confidence: "*How confident are you in your decision? Why / why not? What would make you more confident?*", as well about how well they felt they understood the recommendation: "*How well did you understand the recommendation engine? Do its recommendations make sense to you?*". This line of questioning was repeated for both recommendation modals.

Following the specific questions, we presented a set of general questions applying to both recommendations. First, the participant was asked to explain how the layout of the recommendations helped or hindered their decision-making. We also encouraged the participants to comment on aesthetic choices. They were given a chance to mention any issues that had not been discussed yet. Finally, we asked each of them about their personal experience with loans: "*How familiar are you with loan processes? What loans have you worked the most on?*"

The ELM model states that attitudes can be measured with scales. Self-reports of effort are an assessment method also covered in ELM model. A questionnaire with scales allows for a more quantitative analysis of the user's effort and ease of use. Therefore, the fifth step in the study was a questionnaire, which contained four parts:

1. Demographic questions about the subject's age, job experience with loans and gender
2. Perceived usefulness multi-item 5-point scale [45]
3. Perceived ease of use multi-item 5-point scale [45]
4. Satisfaction single-item 5-point scales

The questionnaire can be found in Appendix III. The perceived usefulness and ease of use scales, as defined by Davis and Davis [45], are widely in use and have been validated for software evaluation. Both usefulness and ease of use play a part in the acceptability of a system. Both scales contain six items and allow for simple questions to be asked and answered. The satisfaction section contains four single-scale items that assess the user's satisfaction with PERSEVERE. By asking the participants the same questions in the questionnaire, their experience could be easily assessed and compared. The data gathered could be used as additional information for the qualitative part of the user study.

4.4.2 Data Analysis

In order to answer the study questions and main research question, we started by transcribing the interviews with the tool Otter.ai[6]. The interviews served as our primary data source. After the automatic transcription, we reviewed each one and manually edited some since some words were interpreted incorrectly. In addition to the transcriptions, we viewed the screen recordings for more context for their answers (e.g. if they noticed an interactive feature) and noted down any important timestamps.

Similar to requirements, we analysed the transcriptions with the affinity note method [38]. We read through the data first and then noted down all explanations, ideas and critiques that answered the study questions. Each participant was assigned a colour and the notes were colour-

[6] https://otter.ai/

coded accordingly so they could easily be differentiated. The notes were then grouped into three large categories:

1. Specifics about the loan application recommendation modal
2. Specifics about the assign-to modal
3. General comments about and issues with the user interface of PERSEVERE.

Each grouping was then further sub-grouped by common topics. For example, both the first and second scenario groups contained smaller ones describing missing information, the most important factor for a decision, notes connected to risks etc. The result of this process is detailed in section 8.1. Any possible improvements derived from or proposed during interviews can be seen in section 8.2. Any possible improvements that couldn't be applied in this thesis are outlined in section 8.2.1.

The questionnaire was analysed for average scores of usefulness, ease of use and satisfaction. Per every question, the average score was calculated for all answers by giving values to textual answers from 5 (best) to 1 (worst).

5. Conceptual Foundation

The research question of this thesis is *How can prescriptive business process monitoring be visualised in a persuasive manner?* When creating a visual presentation of a persuasive recommender system in a business process, guiding principles are necessary. The principles describe how to integrate behaviour change and visualisation strategies in order to create a system acceptable for the process worker. Due to a universally applicable model or framework for business processes and even for persuasive visualisation not existing, we combined different commonplace theories that have been successfully used in other fields. In addition to singular principles, we selected three frameworks as a larger basis for the logic of the principles. Those three frameworks are PSD (Persuasive System Design) from section 2.3, PCD (Patient-Clinician-Designer Framework) from section 2.3 and the health recommender framework [33]. All principles are denoted as PE-s or *persuasive principles*.

The principles defined in this paper cover only the *technical* persuasion dimension, as defined by Al-Ramahi et al [31] (section 2.3). The technical dimension of persuasion consists of primary functionalities of a behaviour change system, such as techniques, informative presentation, effort expectancy, also known as ease of use, and persuasive messages. Behaviour change theories and techniques provide direct input for the research question and will encompass the entirety of the technical dimension.

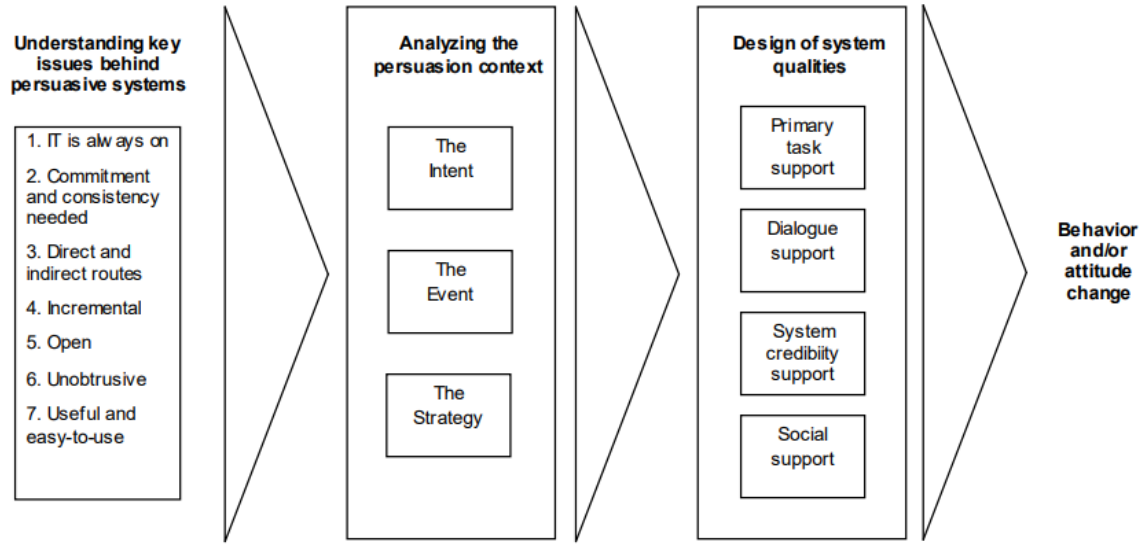

Figure 2. Phases of the PSD framework [30]

At the beginning of system creation, the designer of a system should consider the adoption and acceptance of a behaviour change system. The first step in the PSD framework, as seen in Figure 2, is comprehending seven postulates regarding behaviour change systems. The second step in the PSD framework is to analyse the persuasion context, which has been done already as the thesis is guided by the boundaries set by the research question. The third step is to select persuasive strategies. A persuasive strategy is a principle as defined by Fogg [46] and Oinas-Kukkonen et al. [30]. Both the postulates and strategies will be the basis for the

principles described, as they can be well contextualised for business processes.

The Patient-Clinician-Designer (PCD) framework proposed many design questions for different perspectives during system development. Out of these, we selected 4 questions which apply to the technical dimension and contextualised them as principles for prescriptive business processes.

The health recommender system framework, which combines multiple persuasion theories, was the last of the three frameworks applied. The framework describes the requirements for a persuasive health recommender system and as this thesis is concerned with a business recommender system, the requirements can be filtered and contextualised for that purpose.

All of these frameworks answer the following three questions: *how to show a message*, *when to show a message* and *what message should be shown.* As this thesis is concerned with process visualisation, the same three questions should be answered to fulfil its goal. Therefore, the following principles will be categorized according to *how*, *when* and *what*.

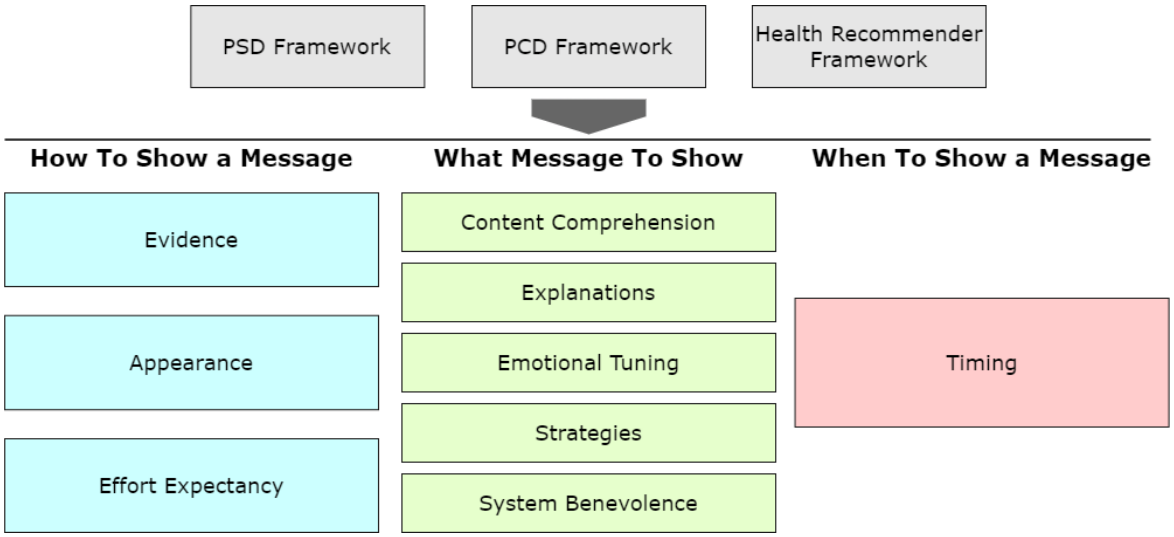

Figure 3. The origin of persuasive principles and their categorisation

To conclude, this section will cover persuasive principles for PrBPM visual presentation, which are mostly derived from three frameworks (Figure 3). The three categories of principles (*how*, *when* and *what*) are further divided into subcategories, which will be expanded upon in their respective sections. In total, 29 persuasive principles are described.

5.1 How to Show a Message

This thesis pertains to the persuasiveness of interventions. A persuasive intervention can have one or multiple components, which are all entailed in the technical dimension of persuasion and increase the intervention's impact. According to Pandey et al [13] the components are as follows:

1. a claim or suggestion.
2. the evidence or explanation itself.
3. the data provided in support of the prescription visualised as a table, chart or graph.

The claim is the main component of a recommendation. It can take many forms, e.g. “The next recommended course of action is X” or “This item has a trait Y, due to which the system recommends Z” [47]. It must be present no matter what the purpose of the system is and as such, it doesn’t need to be covered here. However, evidence, explanations and visualisations are flexible in their use. They increase persuasiveness, as the more data the user knows about the recommendation, the more trustworthy it may seem [18]. These components answer the question of this chapter.

The technical dimension of persuasion entails not only visualisations, persuasive strategies, messages and explanations, but also effort expectancy. Effort expectancy is also an answer to the “how” question of this chapter as it defines a characteristic of the persuasive message carriers. Presenting visualisations and evidence can quickly falter at appearance creation. The appearance of a prescription does not change its content or timing but can increase persuasiveness if used to its full effect. Thus, appearance is another answer to the “how” question.

In summary, this section will cover principles about evidence, effort expectancy and appearance. All three describe the components of an intervention, or the so-called container or carrier of a persuasive message, not the content or timing of them.

5.1.1 Evidence

Evidence for a prescription is a numerical or textual (or a combination of both) explanation of the logic behind it. It is often made of numbers and trends, for example “the item was recommended because during this time X increased and Y decreased”, and “You have been recommended the alternative X because you have chosen filters such as…” [18]. The principle for persuasive messages and explanations is based on the question posited in the PCD framework “Can the system provide evidence for a suggestion?”. The principle is:

PE1. The system should provide textual or numerical explanations for an intervention.

As mentioned before, the evidence can also have supportive visualisations. Visualisations should be used as data for explaining and supporting a recommendation. There are a large number of visual representations available, such as index charts, stacked graphs, scatter plots, node-link graphs, etc. [20]. Only certain visualisation techniques apply for a given purpose and data type and therefore, selecting the correct visualisation is crucial. There also exists the question of when to prefer a graph, table or chart. Overall, a simple key to selecting the correct chart doesn’t exist, but there are guidelines to avoid common pitfalls [48].

Anagnostopoulou et al. [49] found that combining persuasive visualisations and messages affected users’ decisions more than just textual messages alone. Hence, they should be co-implemented. For visualising charts, graphs and tables, the question “Can the system provide evidence for a suggestion?” was, in combination with available persuasive visualisation theory, condensed into a principle. The resulting principle is:

PE2. For a higher likelihood of persuasion, in addition to messages and explanations, the system should provide visual evidence for an intervention.

The persuasive principle of comparison can be combined with visualisation very well. Weber's law [50], which applies to all our senses, posits that "a noticeable change in stimulus is a constant ratio of the original stimulus"; in other words and in terms of visualisation, the bigger the change from the starting visual, the more we notice it. This can be leveraged with comparisons when the difference in data points crosses a certain threshold, for example for a before and after visualisation [48]. Juxtapositioning (or side-by-side comparison), superpositioning (or overlaying two items) or transforming two visualisations with explicit encoding (or calculating the difference between the two and representing that in a visualisation) can be persuasive and works well with charts such as scatter plot and bar charts.

PE3. As evidence for a prescription, comparative visuals should be applied.

5.1.2 Appearance

In addition to the message container itself, certain appearance characteristics must be present for a high likelihood of persuasion. Nunes et al. [51] similarly to the PCD framework [32] referenced user-centric design principles for persuasive self-care applications and proposed questions for visualisation that can be generalised and applied to other fields, including PrBPM.

Due to the large size of datasets recommender systems usually use to make recommendations, any graphs, charts and tables can grow quite large. The questions "Is the type of visualisation adapted to the characteristics of the process? If not, can users configure it to their needs?", "Is the visualisation suited for different levels of understanding of the data?", and "Does the visualisation strategy take a possible information overload into account?" by Nunes et al. [51] also point the attention of the designer toward the information overload problem.

Figure 4. Details-on-demand interactivity features (left) vs. low interactivity (right) [14]

For solving the issue, Richthammer et al. [19] recommend using interactive graphs where the user can zoom in-out and add or remove constraints and conditions (Figure 4). This also

aids with effort expectancy, as it reduces information overload and enables users to explore the relationships between items, increasing the likelihood of them finding a suitable item [52].

Not coincidentally, this details-on-demand principle is also recommended by Munzer's rules of thumb [22] and Shneiderman's Visual Information Seeking Mantra [23]. For most cases, however, the KISS principle applies well; simplicity can be very convincing. Adding emphasis points to large visualisations makes for effective communication. Critical information, e.g. the most compelling part of evidence or explanation, should be the first thing the user notices [53]. This can be, for example, the feature contributing the most to the intervention, which is highlighted in a bar chart as a floating, separate sector from the rest of the visualisation. The zoom level of a visualisation can be set so that items of importance are highlighted.

PE4. Visualisations should follow the principles of good visualisation (section 2.2) so they can enable different levels of understanding of the data and avoid information overload.

An important detail in visualisation, in addition to sizing, is choosing the correct colour coding [18] (Figure 8). Colour coding and applying contrast strategically are also covered in section 2.2. Using colours instead of black and white can generate interest and certain colours (for instance red, green and yellow) can have a bigger impact than others [53]. An interactive visualisation, where items can be filtered by colour, can provide clarity and therefore, enable the user to trust the system more. Basic graphic design principles regarding contrast apply to visualisations in recommender systems. Grouped items should have similar colours and opposite data must have contrasting colours [22]. Although aesthetic colour combinations can give cues of expertise, it is often best to keep it simple.

PE5. Aesthetics and colour groupings of prescription evidence should be strategically used to create a higher probability of persuasion.

In addition to the aforementioned principles, a designer of a system should also focus on the system's appearance from the business point of view. The 2nd postulate of the PSD framework states that users prefer their views about the world to be organised and consistent. This cognitive (in-)consistency can be used to create dissonance and thus, motivate behaviour change or to make the system seem familiar by imitating known cognitive shorthands, therefore inspiring trust. Imitating the user's way of thinking or acting has proven to be successful in multiple tests with recommender systems [18]. This helps the system look more like a social actor to be trusted rather than a black box machine. The common beliefs the process workers have towards possible recommendations should be leveraged in appearance design. The usual way of operating within a case should also be mimicked. Consequently, the principle is:

PE6. The recommendation system should follow, at least visually, the usual decision-making logic of a process worker.

5.1.3 Effort Expectancy

The 6th and 7th postulates, as described in the PSD framework, state that a persuasive system must be unobtrusive and easy to use. When all technologies (applications and other infrastructure solutions) support each other and can be used together seamlessly, no additional work is imposed on the user by the system and can hence be seen as a helpful tool instead of additional work. Every user has a different degree of effort expectancy and thus stressing the flexibility of the system helps persuade people to use it. This applies not only to messages proclaiming flexibility but also to the flexibility of visualisations as well. For effort expectancy management, the PCD framework questions "Will the system fit into the daily lives of users?" and "Does it potentially impose additional work on the users' part?" were refined into the following principles:

PE7. The persuasive system should fit into the daily business process activities and should be integrated into the existing business infrastructure.

PE8. The persuasive system should require a minimal amount of input.

5.2 What Message to Show

When a suitable message carrier is selected, the content needs to be generated or selected. The content can have characteristics such as comprehensiveness (or thoroughness of text) and emotional tuning (i.e. what emotion the content is tuned into). These characteristics will be covered separately.

The technical dimension of persuasion entails explanations and persuasive strategies, which all describe the content of a message. Explanations are necessary to mitigate the black box problem of AI systems. Strategies, while sometimes also possibly describing the "how" of a message, mainly convey what persuasive content or algorithms can be added to the system for better persuasiveness. Therefore, both topics will be covered here.

An important part of persuasion is the trust the user has in the system. A seemingly good-willed or, in other words benevolent, system can be easier to trust. Benevolence can be conveyed through messaging, which is why system benevolence is also one of the topics covered in this section.

In summary, this section will cover principles about content comprehensiveness, explanations, emotional tuning, strategies and system benevolence. All five describe the content of an intervention.

5.2.1 Content Comprehensiveness

The 3rd postulate in the PSD framework states that persuasion can be carried out via a direct or indirect route. The direct route is the more comprehensive one, where deep analysis of the shown content is encouraged. The indirect route relies on visual and cognitive shorthands and cues for persuasiveness in a little amount of time. Guidelines set by Gedikli et al. [54] support comprehensiveness, because users would rather take their time to read in order to make good decisions. As this thesis is concerned with business processes, in which risks

must be mitigated and important decisions made, the clear preference is for a direct route to be taken. Thus, the principle is:

PE9. Content within business processes should be more comprehensive rather than not for a higher likelihood of positive impact on the worker as well as the business.

5.2.2 Explanations

Recommender systems are based on complex algorithms and often work as so-called black boxes, where only the input and output are apparent to the user. This, however, reduces trust in the system. To solve this issue, recommender system explanations can be added [18], [19], [52]. As discussed in section 5.1.1, evidence can be shown either as visualisations or as textual or numerical explanations. This chapter focuses on what textual or numerical evidence should be presented for a higher likelihood of persuasion.

Explanations can vary based on the algorithm and architecture of the prescriptive and descriptive models. The explanation can be about an input feature of the recommender system or generated connections between items. As mentioned before, they can be in textual or numerical form, examples of which are given in the chapter on related works (section 3.2)

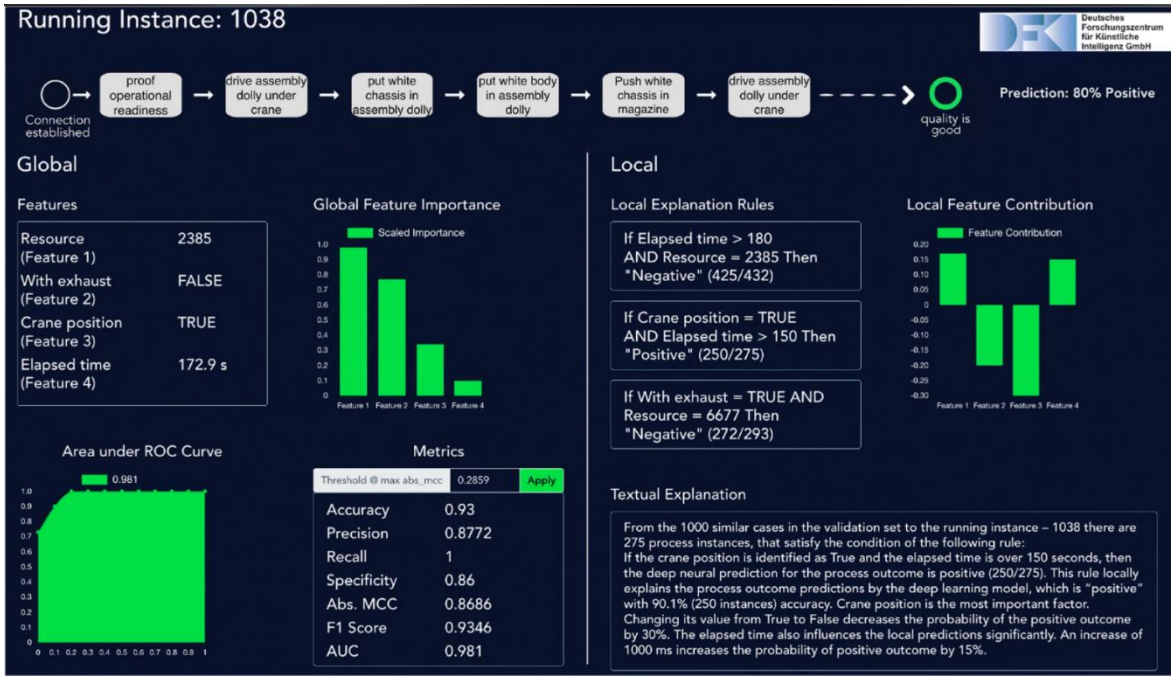

Figure 5. Model, local and textual explanations [36]

In research regarding explainable prescriptive and descriptive business processes so far, the common methods explain how one or multiple features impacted a prescription through numerical values [9]. This explanation style is similar to knowledge-based algorithms, which recommend items based on domain knowledge about how an item's attributes meet user needs [18]. Therefore, these simplistic numerical explanations could be brought to a more easily comprehensible level with similar explanations to knowledge-based recommender systems, for example, "This item has a trait that's better than in alternative recommendations, which will make it suitable for you".

As seen in Figure 5, feature importance explanations can also be made textual, e.g. “X is the most important feature for this recommendation” and the rules the system makes its predictions upon can be turned into an explanation, for example, “If the feature A is X, then the prediction is Y” [36]. System metrics can be used in explanations for increasing trust, an example being “This model’s accuracy is 95%” [36]. Explanations of this kind can give the impression that the system has a personality, as a result allowing the computer to fulfil the role of a social actor [11].

PE10. Textual or hybrid explanations are preferred to purely numerical comparisons.
PE11. The model behind the system should be explained visually or with textual messages.

In addition to explanations for recommendations, a system should also clarify why no recommendation was made [18]. This type of explanation can, for example, be in the form of a sentence such as “There are no events with an X KPI higher than Y and a Z KPI lower than Q.” The explanation can also give pointers about what the user should do next. System transparency in the form of both explanation types can help the users more confidently judge the reliability and trustworthiness of a system.

PE12. Explanations should be shown in the case a recommendation cannot be made.

CASE ID	REMAINING TIME	Explanations for increasing remaining_time	Explanations for decreasing remaining_time
201810011258	5d 6h 7m	ACTIVITY=Evaluating Request (NO registered letter)	CLOSURE_TYPE!=Inheritance
201810000206	5d 2h 12m	ROLE=DIRECTOR	CLOSURE_TYPE=Bank Recess
201811010829	2d 2h 31m	ROLE!=BACK-OFFICE (-1) AND ACTIVITY!=Service closure Request with BO responsibility (-1)	-
...	...	...	...

Figure 6. Prescription explanations based on Shapley values by Galanti et al. [34]

According to Pandey et al. [13], if a person disagrees with a topic, a table will have a high probability of changing their attitude, whereas neutrally or positively opinionated people will be more influenced by a chart. Therefore, a designer should use their knowledge of the business process and process workers’ attitudes and beliefs to choose tables or charts/graphs for a recommendation’s evidence.

If the visualisation counters a user’s held beliefs strongly, simply showing more can backfire [48]. Instead, in such a situation affirmation (e.g. praising the user) might work better in lowering a user’s scepticism.

PE13. Tabulated visualisations of a prescription’s supporting evidence should be used if the user is likely to disagree with it.

PE14. Graphs and charts are more persuasive if the user is more likely to agree or be neutral.

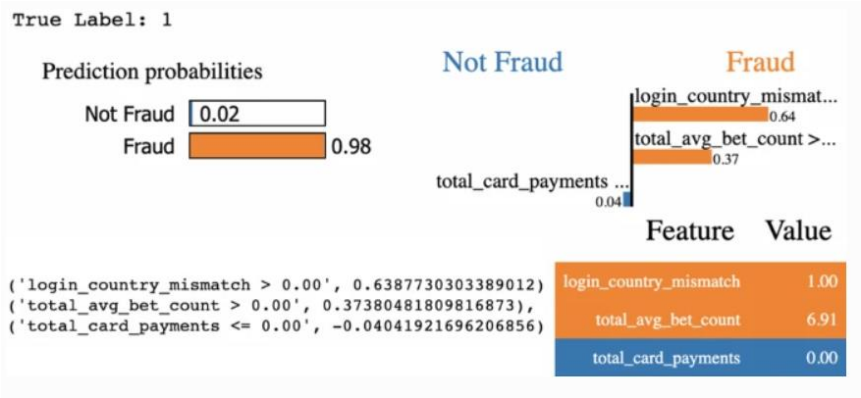

Figure 7. Explanations based on XAI methods [35]

XAI methods used for explanations of prescriptive monitoring (section 3.2) often have a suitable plotting functionality included. Examples of this are ICE plots, LIME plots, PDP plots and ALE plots [9] (Figure 6 and Figure 7 [34], [35]). However, if this is not the case, Richthammer et al. [19] proposed some visualisation techniques to show the user how the algorithm makes the recommendation:

- For one-dimensional data simple charts such as pie or bar charts should be used. One-dimensional data explanations are used with opinion-based recommender systems, where an opinion is, for example, a numerical rating. The same can be done in a business context if, for instance, the model uses a one-dimensional feature to prescribe an intervention.
- For multi-dimensional data approaches such as matrices, scatter plots, heat maps, node-link representations, parallel coordinates, cluster representations [52] and hybrids can be used. The item itself is multi-dimensional but consists of multiple one-dimensional data types. Matrices are better for dense networks of data, node-link cluster representations for small graphs and coordinates for showing any conditions and constraints that have been set on the system.
- For textual data word clouds can be applied.

PE15. Prescription evidence can be supplied via XAI method visualisations or, depending on the data dimension, pie and bar charts, scatter plots, heat maps, word clouds etc. can be applied

5.2.3 Emotional Tuning

Persuasive messages are influenced by the framing effect, which is the idea that a person's decision can vary due to cognitive irrationalities and the framing of alternate choices [55]. The framing effect can be used to the system's benefit to create a positive or negative connection with a recommendation. Framing effect management also coincides with principles in PSD.

PE16. Persuasive messages should be managed with the framing effect to create negative or positive emotions and thus reinforce the target behaviour.

Nyström et al. [55] proposed framing effect principles for sustainability, which are borrowed and contextualised for PrBPM. Some framing effect management principles are already covered by other persuasive principles, such as risk management, system benevolence and ease of use. Others that can be applied to persuasive messages in business processes are:

- stress mainstream social practice of some activity. A positive framing effect will tell the user that using the system will make them a part of and a valuable asset to the social community of the business. An addition to this is stressing group inclusion importance, which tells the user that by interacting with the system they aren't alone and can have an impact on the collective.
- frame the activity as doing something good or noble. A negative framing effect will make the user see not participating as a risk or as if they are missing out on something positive.
- demonstrate the advantage of the system. Persuasiveness could be increased if the long term positive effect the system has on a process is presented.

5.2.4 Strategies

In the PSD framework, the third step when creating a system is selecting suitable strategies for carrying out persuasion. Out of the total 28 strategies (section 2.3), not all can be applied to recommender systems in business processes. Some techniques are inherently already applied to an extent and some cover aspects of the "how" and "when" question, and thus will not be included under the "what" persuasive strategies.

We have selected some potential strategies to be used based on the most commonly used ones that have seen success in other fields of research [32], [49], [56] in addition to possible situational suitability. It is important to note that for brevity some have been combined. The combined principles can be seen in Table 8 [29].

5.2.4.1 Primary Task Support

Primary task support is one of the four categories of strategies in the PSD framework. These strategies help in the carrying out of the user's main task in the system. Some primary task support strategies are not included here, for instance tunnelling (the system already is a step-by-step process from input to output), and reduction (the user's effort during a business process is lessened by recommending a solution).

Included under these strategies, in addition to ones from the PSD framework, is goal-setting. Goal-setting theory can help persuade a user to act towards a goal and therefore, if given business goals of the company, can motivate them to meet those goals [57]. As the goals of the business can be fulfilled with the help of the persuasive system, goal-setting fits well under primary task support.

PE17. Although *simulation* is inherently already applied to an extent (recommendations are made based on predictions most often created with simulated cases), it can be included as visualisation content. A *simulation* of a decision during a process persuades by presenting clear consequences of every action. A process worker might usually trust their instincts or knowledge, but a simulation might persuade them not

to. Business process simulation (BPS) methods can be used [2].

PE18. During a business process, a worker does not always have the resources to monitor their performance. By implementing *self-monitoring* the system does that for the user and through it, informs, reminds and persuades the user to do a target behaviour.

PE19. People respond differently to different persuasion strategies and messages [18], [49]. Studies have found that users perceive a recommender system as a social actor and thus this facet should also be considered to make the system's decision-making process similar to the users [18]. *Tailoring* and *personalization* are widely applied [30] and have big impacts on system persuadability, and that is why they should be applied to business processes as well. Persuasive visualisations and messages should be adapted to the characteristics of the business process as well as to the process workers. If such personalizations cannot be made, the visualisation should be configurable to a user's needs.

PE20. As business processes are goal-oriented [2], leveraging those goals in persuasion is a simple yet effective strategy. *Goal-setting* can hence be persuasive for process workers.

5.2.4.2 Social Strategies

Social support strategies help connect the user and persuasive system. As a business process is tied less with social facilitation within the company and more with achieving business goals, only one strategy was chosen to be applied to a business environment.

PE21. The inner-business social dynamics can be leveraged by employing *social comparison*, which is one of the more often applied strategies [30] and can be a good motivator [58], and encouraging either *competition* or *cooperation*.

5.2.4.3 Feedback

Another category of strategies in the original framework is dialogue support. The strategies in this category provide feedback to the user. Two strategies of this category, liking and social role, are covered in other chapters, as liking does not influence the content of the message and social role is a part of benevolence, which is a separate topic in its own right. The strategy of suggestion is inherently already applied by a persuasive recommender system, as the recommendation is a suggestion. Thus, the principle for feedback is:

PE22. *Praise* and *reward* can be great motivators [58]. Praise especially can provide a chance for comparison. *Praise* and *reward* may be persuasive in business processes as well.

5.2.5 System Benevolence

Users see recommender systems not only as support tools but as social actors. Wang and Benbasat [59] found that recommender system users perceive traits like benevolence and integrity in the systems, both of which are dimensions of trustworthiness. Benevolence is defined as the system's interest in and acting in the user's interest and overlaps with system verifiability, whereas integrity is described as the system's adherence to a set of social principles the user deems acceptable, such as honesty. We have grouped these two definitions

under benevolence, as integrity can be perceived similarly to benevolence.

Textual cues can help create perceptions of benevolence and integrity. An example of this is presenting the total amount of possible recommendations that can be made, out of which a few were chosen. Another is asking the user for feedback on the recommendations. This is the basis for the principle:

PE23. Informing the user about the amount of work the system has done during the search and recommendation process can increase persuasiveness.

In the PSD framework, a separate category of principles is for improving the system's credibility. As these strategies describe characteristics that are tied to benevolence and integrity, they are outlined in this section. A business process takes place in a professional environment, which is why from system credibility support characteristics, verifiability, trustworthiness, expertise and credibility are the most suitable for prescriptive system credibility support. All four of those characteristics can aid in giving the impression of a benevolent persuasive system.

Figure 8. Feedback prompt [49]

For increasing verifiability, the user must be allowed to critique the system. This can be implemented by, for example, prompting for feedback at the end of a recommendation process [60]. This can be done as a popup notification as in Figure 8 [49]. Not only does this help improve the algorithm behind the system, but it may give the users a clearer understanding of the features needing clarification and may make them more inclined to try to use the system again.

PE24. Allowing users to critique the system can increase its persuasiveness.

Torkamaan and Ziegler, in their health recommender system framework [33], described the requirement of expertise as such: "The system should provide the source of each recommendation and explain the reason for its effectiveness. It should also provide supporting and source information, such as an accessible link to additional information concerning the credibility and trustworthiness of the items. Users should be able to access the source easily." As credible sources are key to business processes, the principle is:

PE25. Prescriptions should point toward credible sources to give the system an image of expertise.

Presenting risks can seem benevolent to the user. In addition to system credibility, Torkamaan and Ziegler focused on risk management. They described the requirement of "Users should receive information about the health risks and the seriousness of the threats. Users should receive information about the health benefits." As a business process entails certain risks, this requirement should be applied as a persuasive principle. The framing effect also creates transparency regarding risk. The risk the user's actions can create or mitigate must be communicated. The system should provide feedback and explanations about these risks.

PE26. Risks concerning possible disuse or ignorance of the persuasive system should be communicated.

The 5th postulate of the PSD framework states that a persuasive system should always be open, i.e. reveal the designer's bias and only base the content on truthful information. A question proposed by the PCD framework [32], "What will motivate them to use the system?", is also tied to system openness. The user should be able to perceive the full impacts of the system. Honesty is a part of benevolence and therefore, a principle is formed:

PE27. The system should communicate why it exists and what the user can gain from using it.

5.3 When to Show a Message

A persuasive message needs to be presented at a certain time to have a bigger impact. Persuasion does not happen at once and often takes many iterations to create an attitude or behaviour change. This section outlines principles related to the timing of persuasive messaging.

The amount and frequency of textual messages are important. Repetitive messages can become an annoyance and deter users. Restricting the number of persuasive messages and adding multiple messages per persuasive strategy helps persuade users more [47], [49], [61]. Anagnostopoulou et al. [47], [49] used 98 persuasive messages for eight different contexts (e.g. nice weather, the destination is within biking distance etc.), which the users found non-repetitive and persuasive. On average, they had 12 messages for one situation. That is a possible upper bound to strive for per persuasive strategy, with the lower bound being at least 2, consequently:

PE28. The amount of repeating persuasive messages should be restricted.

The 1st postulate of the PSD framework states that persuasion is a process, rather than a single act. Different factors, such as the user's needs and stances, can change the likelihood of persuasion during every attempt and thus, a persuasive system should be able to adapt. The 4th postulate pertains to the incremental nature of persuasion. A simpler goal to achieve is to influence someone in steps rather than with one suggestion. Therefore, a principle is:

PE29. Persuasion should be incremental and adaptable to the business process and its workers and goals.

6. Solution Design

This section describes the requirements elicitation process for PERSEVERE to be developed. To elicit requirements for the visual presentation of a prescriptive business process, the perspective of potential users should be considered. To this end, a user persona was created for examining how the system would affect them, which is described in section 6.1. Applying a scenario to the persona helps even further demonstrate the potential benefits of the proposed persuasive principles for the user. The scenarios are outlined in section 6.2. Finally, using the scenario and the persuasive principles (section 5) as a basis, requirements are elicitated. Section 6.3 gives an overview of the requirements, their categories and prioritisation.

6.1 User Persona

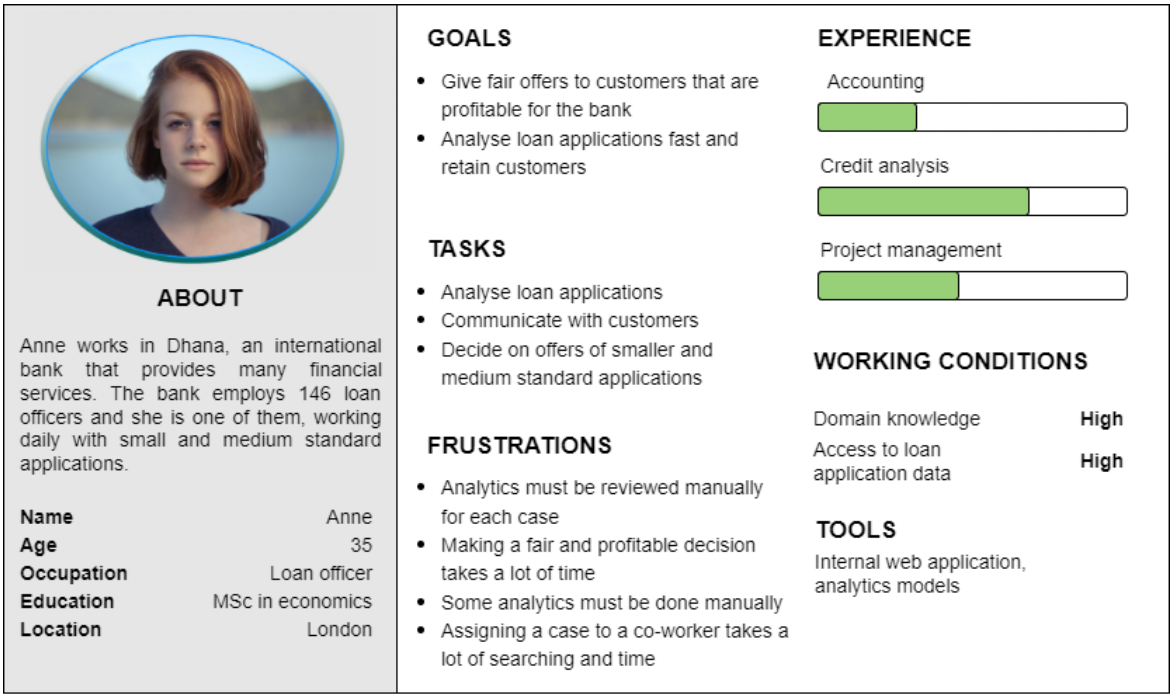

Figure 9. User persona

As the research question of this thesis is *How can prescriptive business process monitoring be visualised in a persuasive manner?* we chose the user persona to be a process worker that has to make impactful decisions every day, while still belonging to the biggest user group of middle-level employees. For the persona, the following points are covered: a general description, after which a detailed account of their background and work environment is given. The persona can be seen in Figure 9.

Description. Anne is a 35-years-old loan officer in Dhana, an international bank based in London. Anne is one of 146 employees who handle loan applications.

Background. Anne has 10 years of experience in banking and out of that 6 in credit analysis. She has an MSc degree in economics. Anne studied macro and micro economics and finance in school. She spent the first three years after university in a private business as a project leader for a payment default registry. Soon after, she found the job in a private company to

be too time-consuming and single-focused on a narrow topic with limited career options and switched jobs to an international bank. Although Anne started studying economics to be a project manager or finance officer, after 4 years of working as a junior specialist in the finance department managing daily accounting tasks such as controlling and verifying financial transactions she found herself more interested in the process of credit analysis and review. She moved to the corporate department as an assistant supporting account manager and prepared credit cases belonging to small and medium-sized enterprises. After one year she was attracted by a job offer by Dhana bank, hence taking the next step in her career. She has been working for Dhana as a credit officer for 5 years.

Environment. Dhana is a large bank and provides many financial services, loans being one of them. As lending is one of the main sources of income for Dhana, many credit professionals are employed including Anne. She has many tasks to do daily and as a credit officer, she is facilitating lending to the bank customers, assessing their creditworthiness and processing relevant paperwork. She is also the sole decider for the small standard credit cases given to her.

Every loan application is different and requires focusing for a fair and profitable decision to be made. Last week Anne worked on one mortgage loan application and 8 small standard loan applications, including car leasing. This week altogether, Anne analysed 9 loan applications. Next week she will need, in addition to daily tasks, to write a monthly report about recent applications and take part in a meeting regarding one large business loan application, as the ultimate beneficial owner of the company is a private person in her portfolio and thus she already has good background knowledge of this person. Recently, the number of loan applications and therefore, the wait times for customers, have increased. This has led to customer complaints and a non-negligible churn rate. Anne has multiple loan applications waiting in the pipeline to be analysed. Her superior keeps demanding faster results despite there not being enough time to correctly assess every case given.

6.2 Scenarios

This section covers the scenarios PERSEVERE and evaluation of the thesis will be based upon. A scenario functions as a basis for identifying a persona's goals, needs and pains. The scenarios are diversified against different possible prescription outputs: the best next action and the best resource for achieving a goal (section 2.1). The application details were semi-randomly chosen from the Dhana bank loan application handling process log.

Scenario 1. Anne is working on a car leasing application. She has worked with car leasing cases many times before. The customer has submitted an application via the bank's website, but the application is missing the required amount. The customer has now emailed Anne to let them know the sum they need, which is 5700 euros. Anne first has to update the application's details and then decide whether or not to make an offer. Normally Anne would analyse the application, the customer's income and cost structure, repayment ability and credit habits to ascertain if an offer is feasible. Sometimes she would miss an important detail or under- or overestimate a customer's solvency, in turn either inadvertently increas-

ing the churn rate or costing the business money in other ways. A more advanced recommendation tool she could trust and abide by to decide would lessen the risks of deciding by herself, and also shorten the process's cycle time. The second improvement of the recommendation tool is a more automated assessment of historical credit behaviour by combining internal overdue payments history with the external credit information bureau data.

Scenario 2. An application has been submitted for purchasing a home with the requested amount of 53 000 euros. Anne has looked through the application but needs to delegate the case to another worker who has more experience with mortgages. For this, she will have to disturb the 145 other loan professionals in the group chat or ask her manager, who would in turn have to ask other managers. When she has done so before, it would take at least a day to find someone with previous experience. In addition, Dhana is a large bank and Anne does not know many of her co-workers and so she doesn't have anyone specific in mind. Thus, Anne could complete the task quicker if she had a tool that gave her a resource recommendation, i.e. a co-worker for her to assign the case to.

6.3 Requirements Specification

As the aim of this thesis is to create a guide based on persuasive principles and evaluated prototypes, requirements are necessary for further development. This section covers the requirements for the visual presentation of a prescriptive loan handling process, as well as the prioritisation of these requirements. A requirements specification [37] is developed based on the persuasive principles defined in section 5. Requirements, compared to user stories, are more comprehensive and focus less on the reasoning behind a feature and more on what the feature should be. The requirements were created with the perspective of the user persona in mind (section 6.1).

In total, 40 requirements were defined. An overview of all requirements can be seen in Table 9 (Appendix II). When adapting the persuasive principles into requirements, groups emerged. Altogether, the requirements are grouped into four categories: Design of the System, Flow of the System, Explanation of the Prescription and Text. The groups are as follows:

- **Design of the System**. Requirements related to the visual, interactive and aesthetic design of the prototype. This group also includes requirements that describe how the system should look according to the parent system. For example, R.D.1 *3D visualisations are used with spatial or topological data and 2D visualisations for all other types of data*.
- **Flow of the System**. Requirements that pertain to the user flow of the prototype. This group contains requirements that describe the structure and layout of the system, along with its integration within the parent system. This includes the before and after of using the system. For example, R.F.1 *Within the system, structure of visualising content should be the same as the decision-making process of the worker.*
- **Explanation of the Prescription**. A group of requirements that relate to explanations of a prescription. These can either describe the message carrier type or the

content of the explanations. For instance, R.E.1 *Textual, numerical or hybrid explanations are shown for a prescription.*

- **Text**. Requirements related to what text should be presented to the user. This group of requirements is mostly derived from principles concerned with persuasive strategies and system benevolence. R.T.6 *The system's overall success rate is displayed* is an example.

Some persuasive principles were excluded from the prototype. PE13 and PE14 (section 5.2.2) are only applicable in the case of repeated use, where the user's behaviours can be learned over time. PE19 (section 5.2.4.1) can also only be applied over repeated use of the system, during which the most suitable content for a user is revealed. Similar to these three, one whole category of principles, Adaptiveness, was discarded during requirements elicitation, as the prototype created in this thesis does not allow for incremental development, but rather a one-time evaluation.

To efficiently develop PERSEVERE, the requirements were prioritised. Requirements prioritisation mainly helps decide on the core requirements of a system and exclude trivial ones, in addition to many other benefits [62]. For this, the MoSCoW method of prioritisation was implemented, wherein requirements are divided into four categories [63]. Requirements are prioritised according to the importance of implementation. Requirements in a single category are of similar priority. The categories are:

1. M or **must-haves**, i.e. the core requirements the system cannot function without
2. S or **should-haves**, i.e. high value requirements that the system can function without
3. C or **could-haves**, i.e. desirable requirements, but not necessary for the system
4. W or **won't-haves**, i.e. requirements that will be excluded from the system.

In the **must-have** category, requirements R.E.1. *Textual, numerical or hybrid explanations are shown for a prescription*, R.E.2. *Explanations contain visualisations,* and R.E.7. *The explanations fit the prescriptive model by showing evidence through a tree visualisation* alleviate the AI black box problem the most. From a business standpoint, goals and risks play a large role within a process, and thus requirements that focus on those aspects were included in the must-have category (R.T.7 and R.T.9). The must-haves of the persuasive prototype are presented in Table 2.

Table 2. Must-have requirements

#	Requirement	Based on
R.E.1	Textual, numerical or hybrid explanations are shown for a prescription.	PE1
R.E.2	Explanations contain visualisations.	PE2
R.E.7	The explanations fit the prescriptive model by showing evidence through a tree visualisation.	PE15

R.D.1	3D visualisations are used with spatial or topological data and 2D visualisations for all other types of data.	PE4
R.D.2	The most important details of a visualisation should be emphasised.	PE4
R.D.3	Complex visualisations are interactive for details-on-demand.	PE4
R.D.3.4	The worker can toggle accordions for explanations and visualisations.	PE4
R.D.4	The system presents visualisations with colour codings and a legend.	PE4
R.D.5	Contrast, proximity and repetitions of visual elements should be utilised for indicating identicality, similarity and patterns.	PE4, PE5
R.D.6	Visualisations should be ambient, aesthetic, emotionally-engaging or metaphorical.	PE4, PE5
R.F.2	The system should be easily openable and dismissible.	PE7
R.F.3	Within the system, the user is asked for a minimal amount of input, i.e. accept and reject.	PE8
R.F.4	When opening the system, information should be automatically queried from the parent system.	PE8
R.F.5	When having reached a decision, the system automatically conveys information to the parent system.	PE8
R.F.6	The system displays a progress bar during prescription creation.	PE23
R.T.7	System should present clear business goals to be attained through the system.	PE20
R.T.9	The system displays risks regarding a prescription.	PE26

As this thesis focuses on visual presentation, the most basic visualisation principles must be followed and were included as well (R.D.1, R.D.2, R.D.3, R.D.4 and R.D.5). The requirement R.D.6 applies to all visualisations. A visualisation that aims to be persuasive should be, in any case, at least ambient and not disturb the viewer. The other persuasive dimensions can be applied as suitable for a situation. As information overload is easy to happen, R.D.3.4 must be applied to allow the user to select what they want to see.

In order to give the system an edge over alternatives, requirements R.F.4 *When opening the system, information should be automatically queried from the parent system* and R.F.5.

When having reached a decision, the system automatically conveys information to the parent system were included. A vital part of enabling these requirements are R.F.3 and R.F.2, which ensure the system is a part of the business infrastructure and were therefore included.

Lastly, the must-haves contain a smaller but very important requirement that ensures the user's trust in the system. R.F.6. *The system displays a progress bar during prescription creation* lets the process worker know how the system is doing and can help avoid creating frustration during system loading times. Altogether, there are 17 must have requirements.

The **should-have** category contains 14 requirements. Self-monitoring, as described in R.T.6, can aid in keeping the worker interested in using the system, but is not a necessity. The requirement R.D.9 describes the usage of simulation and comparisons, and although properly done simulations can be persuasive, implementing them can be difficult and costly and thus cannot be a must-have. Comparisons can also be very persuasive, as covered in R.E.3, but only if there is something to be compared. In the loaning business process as described in the scenarios, not enough suitable data exists for comprehensive visualisation comparisons to be made.

Social dynamics can play a role in a business setting (requirement R.T.2). The selection between encouraging cooperation or competition, however, depends on the workers and business. In a bank setting, where professionalism should be maintained, more positive tone is preferred in messaging (R.T.2.1). A negative tone (R.T.2.2) is a riskier option in the current scenario and is categorised under the could-haves. Another social requirement is that of giving praise (R.T.5), which also encourages trust and continued use of the system, but is not necessary for persuasion to take place.

For complex visualisations such as decision trees, searching for an element can help save the user's time and cognitive load. Thus, R.D.3.2 *The loan officer can search visualisations* should be included in the prototype.

Requirements R.D.7 *The system's visual design elements follow the business' common user interface design* and R.F.1 *Within the system, structure of visualising content should be the same as the decision-making process of the worker* aid in making the system visually a part of the business' infrastructure and workflow. They are not necessary for the system but should be applied for better persuasiveness. The requirement R.T.8 *The system displays the total number of possible prescriptions or analysed cases* shows the amount of work the system has done, but is a small feature and thus not a must.

The requirement R.E.4 describes a preference for text over numeric text. As loan officers must be accustomed to numbers due to analysis, numeric explanations may suffice already. Therefore, this requirement is not a must. However, since text is much easier to comprehend, it should still be implemented.

The requirement R.T.1, which states that thorough information about a prescription should be available on-demand, can in a smaller system, such as the one in the prototype, create additional cognitive load. Hence it is not a must, but as it does help with the black box

problem, it should be implemented, if possible. This is especially true when a recommendation model is very opaque, but as most loan applications use simple models, this is not such a case.

The requirement R.T.3. *The system's overall success rate is displayed* can be quite persuasive if the system is successful. However, most machine learning models do not have a success rate high enough to create trust and hence this requirement must be used sparingly when possible. Although in this thesis there is no actual model behind the recommendations, the success rate of the system is still within believable constraints.

The **could-have** category has 3 requirements. The requirement R.E.5, which explains the prescriptive model's characteristics is more trivial as users don't always need to be explicitly informed of it. According to a domain expert, the model most commonly used for loan application prescriptions are decision trees and process workers already have a good understanding of them, which is why any additional explanations can create noise. These requirements could still be implemented, as not all employees will have a thorough understanding and ambient model explanations can create more trust for them.

Using negative user statistics as a persuasive element in a working environment can be detrimental to the good spirit of the workplace, due to which the requirement R.T.2.2 could be implemented only if the designer of the system knows the work force and is not a should or must category requirement.

In a workplace scenario, gamifying a system can be unprofessional. This is why the requirement R.T.4, which describes gifting badges, achievements etc. in case a recommendation is followed, should be sparingly applied. Some users can find this strategy fun and persuasive, others not. Consequently, this requirement is in the could category.

The requirements elicited were a result of a selection and filtering process, which is why there are only 6 **won't-haves**. Most so trivial requirements were already excluded when the conceptual foundation (section 5) was formulated. However, during development, the requirement R.E.6. *The system displays an explanation for no prescription* aids was moved under the won't-haves. Although it can enormously help the user if the system returns no recommendation and the user wishes to remedy the situation making it a must in other cases, it is outside the two scenarios described earlier and thus is excluded.

The requirement R.E.8, which describes presenting a source of model data, was also categorised as a won't-have during development, as in this case the source is the company's process logs and thus users don't need to be explicitly informed of it.

During development, the limited interactivity of the prototype software meant that requirements R.D.3.1 and R.D.3.3 were assigned to the won't-have group. As bank systems are vast and complex, the requirement R.D.8. *The system employs visualisations utilised in other applications in use in the business* can aid the user, but does not ensure visualisations of use. Commonly used visualisations within the bank infrastructure can be too complex, simple, or just present a different data type, so forcing their implementation would carry no benefit. Therefore, this requirement is usually a could-have and should be applied sparingly.

However, as in this prototype no visualisations were included outside the persuasive elements, this requirement was set as a won't-have.

The self-monitoring requirement R.T.7 can aid in keeping the worker interested in using the system but as this would require repeated use of the system, it was set as a won't-have. Due to the limitations of the prototype's interactivity, the requirement R.F.7 was also excluded from development.

After prioritisation, 17 must-have, 13 should-have, 4 could-have and 6 won't-have requirements were outlined.

7. Development

This section gives an overview of the development of PERSEVERE. It is based on the user persona, scenarios and requirements outlined in section 6. A prototype with limited interactivity was created using Figma. In Figma every screen is static, but hotspots can be defined that provide functionality, for example clicking on a button navigates the user to another screen or opens an overlay. Interactive elements (e.g. alerts, tooltips and modals) can be implemented with overlay elements.

All data and analytics used in PERSEVERE are fictional or taken from the Dhana process log. The analytics used in both persuasive modals include some that are currently in use in the industry, some non-finance specific KPIs and some, which an industry expert would need to know in order to make a decision. The generic KPIs were chosen to ensure they would apply to any loan process. In the activity recommendation, the analytics consist of the loan applicant's past behaviour, for instance credit score, overdues and some calculations based on income and loan details. In the resource recommendation, important loan officer portfolio details are included such as cases past due date, most common case type and average portfolio profits.

Section 6.2 presents two scenarios – one about an activity recommendation and the other about a resource recommendation. Therefore, there are two user flows in the prototype. The activity recommendation flow has the user either approve or reject a loan application. The resource recommendation suggests a co-worker the user can assign their case to. Overall PERSEVERE consists of eight screens and 16 overlays.

The visual presentation of prescriptive systems in business processes is presented as a part of an imitation bank internal web application. Hence, elements outside the scope of the requirements were added, for example, a mock Dhana bank logo and buttons for adding a new loan application or filtering existing ones. The external views were based on a free template[7]. Every element inside and outside of the requirements' scope is developed in such a way that it would be suitable for a desktop experience.

The prototype's visual design is informed by both the requirements, as well as existing bank systems in Estonia. The colour palette was chosen purposefully muted: the main colours grey and white allow for a simple user interface while creating a large contrast between any other emphasising colours. The white, grey and dark blue theme was inspired by a bank currently operating in Estonia.

7.1 Persuasive Layout

A persuasive layout is necessary for generalising the findings of this thesis. The layout was designed based on gathered persuasive principles and prioritised requirements and should be applied when creating a visual representation for a recommendation. It ensures the most important elements in persuasion are seen first, then the mandatory elements tied to the

[7] https://uikitfree.com/insurance-dashboard-figma-template/

decision and finally additional details on demand. The layout presumes a modal or something similar is being utilised to present information, but it can be expanded to a larger screen size.

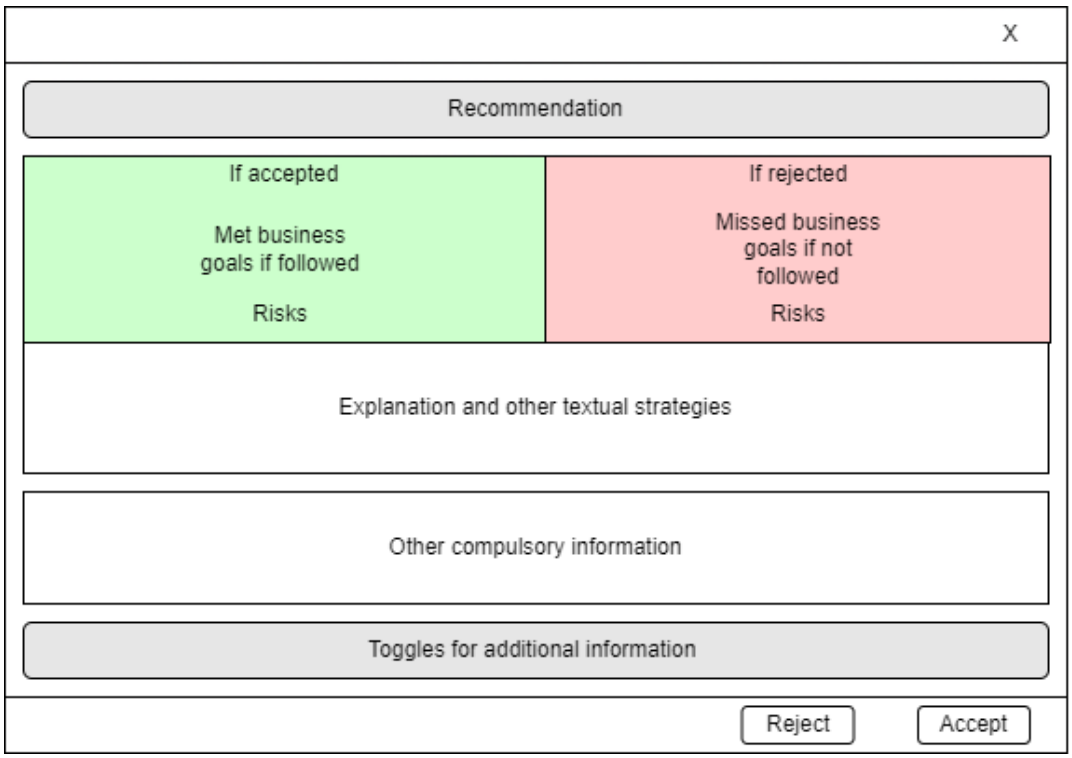

Figure 10. Persuasive layout

The information that is shown during a recommendation should be segmented, with must-haves having priority over other requirements. This satisfies the requirements R.T.1 *Thorough information and details can be viewed about a prescription upon demand* and R.D.3.4 *The worker can toggle accordions for explanations and visualisations*. In this prototype, this was achieved with collapsible accordion elements. The layout consists of the following components (Figure 10):

1. When a modal is opened, the first thing the user should see is the recommendation or suggestion, as it is the most important piece of information. If accordions are in use, this could be a collapsible header element. Font size and boldness can be experimented with for a bigger eye-catching effect.
2. The second elements that should be visible at first look are the possible positive and negative aftereffects that accepting or rejecting the recommendation can have. This is referred to as the **If Accepted/Rejected** block. A juxtaposition comparison should be presented for a bigger impact. This realises the requirement R.D.9 *The system should provide comparisons of simulations of the possible after-effects of following and not following the decision.*
    a. Both sides of the comparison should contain the biggest probable changes to business goals that could happen if the suggestion is accepted or denied. This feature covers the must-have requirement R.T.7 *System should present clear*

business goals to be attained through the system. The business goals should be emphasised with large font size and boldness in addition to icons.

   b. Risks, which can be realised in case of either decision, are the topic of the must-have requirement R.T.9 and therefore must be included. However, these risks must be less impactful than the business goals. In this prototype, risks are hidden from first view by moving them into a separate tooltip. The tooltip only is visible while the cursor is hovering above a red alert icon next to the "risks" keyword. Thus, an additional requirement R.T.1 *Thorough information and details can be viewed about a prescription upon demand* is implemented.

3. Textual and hybrid explanations and strategies should be presented next and must also be visible at first view. Any content here fulfils the requirement R.E.1 *Textual, numerical or hybrid explanations are shown for a prescription.* This is called the **Why Should You Accept this Recommendation?** block.
4. If the decision to be made requires some analytics and data without a prescriptive component, then this information has to be present in the layout. However, as it is compulsory, it can be last in the layout's order to direct more attention toward persuasive elements. Persuasive design can be applied to these elements to further strengthen the argument of the recommendation. This block is named **Statistics**.
5. Big visualisations and graphs should be in separate collapsible elements, closed by default when first opening, or be scrolled down to. One or multiple collapsible elements can be included. All these additional elements, subsequently, will have the **Additional Information** name.

7.2 Prototype Screens

All eight screens of the prototype are variations of one view that presents an overview of current loan applications the worker has assigned to them (Figure 11). This view contains a card per loan application. As the imaginary bank Dhana used in the scenarios of this thesis provides a range of different financial services, the sidebar is necessary for visualising that aspect and communicating that the loans category is currently open in the user interface.

The sidebar also includes a user section, where the user is presented a "Log out" button to convey they are currently logged in, and, for the same purpose, furthermore displays the user persona's name. The sidebar additionally has a section at the bottom for adding business clients. For immersion purposes, the screens include a toolbar at the top of the layout for adding loan applications and filtering them.

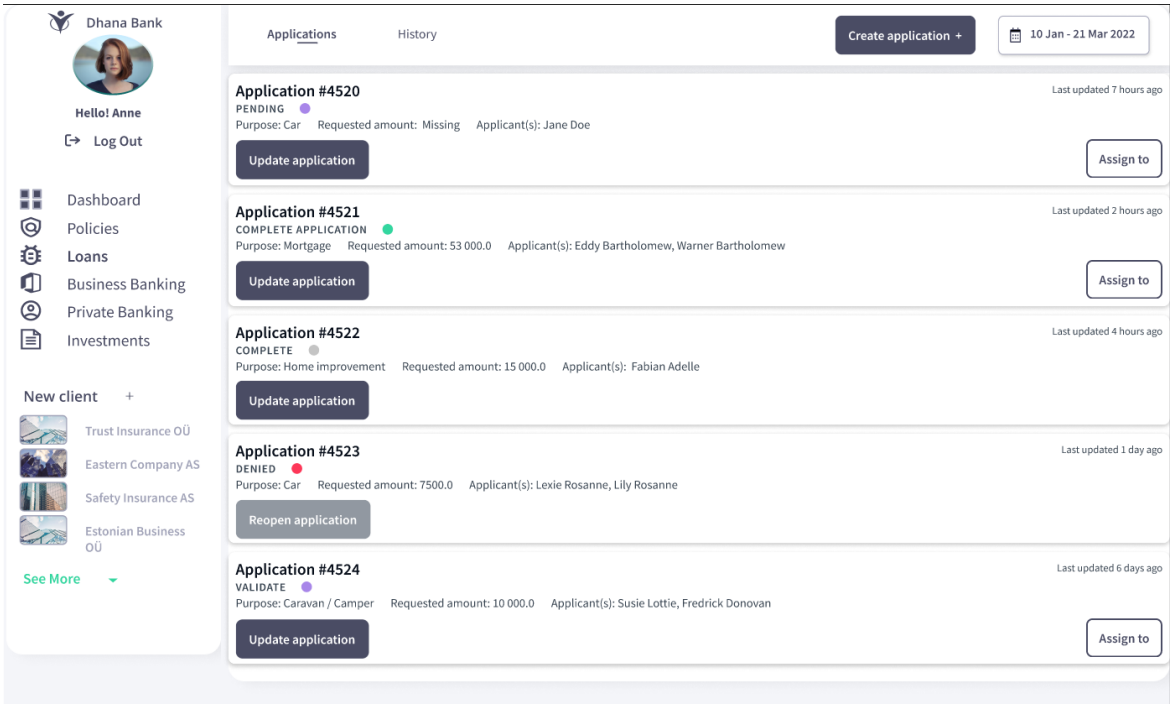

Figure 11. Loan overview screen

The eight variations of the loan overview screen all offer different information. Every loan application card contains an application ID, a status (one of the possible ten statuses in the process log and a corresponding colour we assigned to it), when it was last updated and minimal information about the application contents, e.g. its purpose, requested amount and applicant(s). These details change when the interactive flows are followed, for example, the status and its colour change after an application has been denied. Each application in the prototype was taken from the Dhana bank process log.

7.3 Prototype Overlays

Overlays can be opened and closed on top of a screen and they contain most the requirements of section 6.3. Overlays are presented either in the form of a modal, tooltip or alert. Overlays were chosen to implement the requirements, as they fulfil the requirement R.F.2 *The system should be easily openable and dismissible*.

All overlays either open automatically by themselves after a certain event has taken place or can be opened via a button click on the screen or by hovering on a triggering element, thus ensuring that they can be easily accessed and opened. The same applies to closing the overlays: they can be dismissed via an “X” or “Back” button or they disappear automatically when the cursor is no longer hovering on a triggering element. There are three different modals with different states:

- **loan application details modal**, where the details of an application can be viewed and updated
- **assign-to selection modal**, where the user can select whether they want to search for a worker or be recommended someone to assign a case to

- **loan application recommendation modal** or the activity recommendation modal, where the recommendation, persuasive elements and analytics are presented
- **assign-to recommendation modal** or the resource recommendation modal, where the user can see which co-worker was recommended to assign a case to and other persuasive elements

All four modals have a similar design. The persuasive layout (section 7.1) was applied to both persuasive modals of this prototype: the loan application and assign-to recommendation overlays. According to the requirement R.D.7 *The system's visual design elements follow the business' common user interface design*, the colour palette of all four modals is white, grey and dark blue, as those are the main colours of the loan overview screen (section 7). Other colours, such as red, green, and yellow are utilised for emphasis in visualisations and for highlighting elements. The useful shorthand of those colours implicates how positive or negative a certain element is. Similar elements are placed symmetrically; for example, business goals, icons and buttons, so that the user can connect the elements subconsciously. All the aforementioned design features, among others, implement the requirement R.D.5 *Contrast, proximity and repetitions of visual elements should be utilised for indicating identicality, similarity and patterns.*

In loan applications as well as when recommending another co-worker, no spatial data was in use. Therefore, all visualisations were 2D and fulfil the requirement R.D.1, which describes using 2D visualisations for all other types of data except spatial and topological. All visualisations have a legend and colour coding, hence realising the requirement R.D.4.

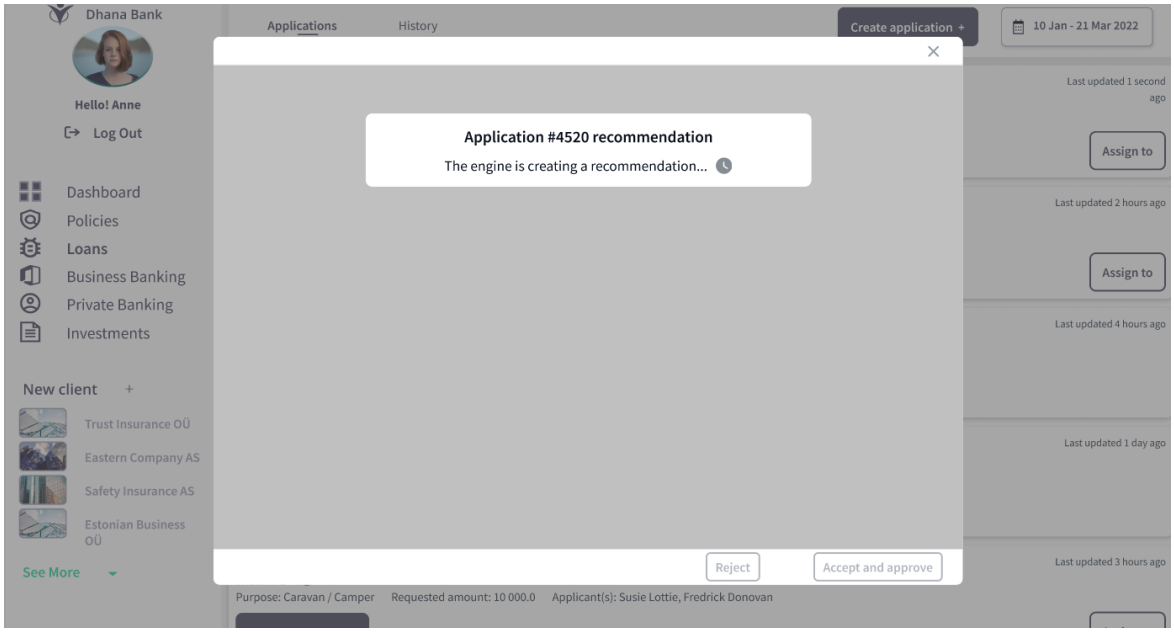

Figure 12. Activity recommencation modal loading screen

All visualisations in the modals implement the requirement R.D.6 *Visualisations should be ambient, aesthetic, emotionally-engaging or metaphorical*:

- Pie charts have grey and brown colour coding, allowing them to be colour-compatible with the rest of the user interface. In other words, the charts are ambient. They are also minimal and aesthetic.

- The bar chart used in the assign-to modal has striking green and purple colours to possibly create an emotional reaction in the user, as well as an aesthetic effect due to the colours contrasting.
- A decision tree visualisation has nodes in grey, green and red. The grey is alike the rest of the user interface, therefore being ambient, and the red and green invoke an emotional connection.

Both pie charts in the modals also emphasise important details by using a bolder and larger font where necessary. This is related to the requirement R.D.2 *The most important details of a visualisation should be emphasised.*

When either the loan application or assign-to modals are first triggered, they show a loading status (Figure 12). As machine learning models usually take some time to make a recommendation, an arbitrary number of seconds was chosen to delay showing the result, in turn making the mock engines feel more realistic. This loading screen is an implementation of the requirement R.F.6 *The system displays a progress bar during prescription creation*. All names used in PERSEVERE are random.

7.3.1 Additional Modals

The details and assign-to selection modals are overlays that were outside of the scope of the requirements, much like the screens outlined in section 7.2, but were added to showcase requirements R.F.3, R.F.4 and R.F.6. Both have limited interactivity and were only added for immersion purposes.

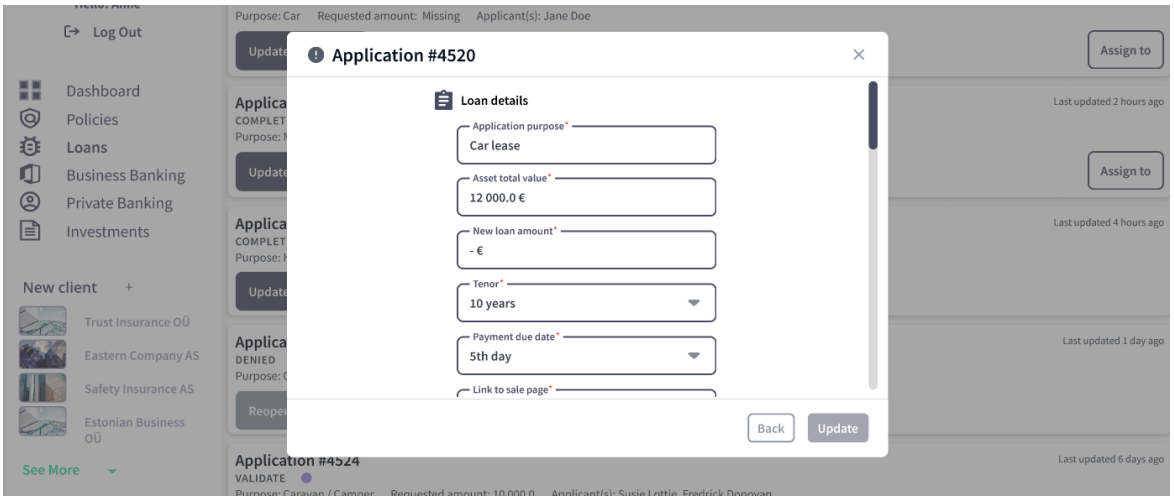

Figure 13. Loan application details modal

The details modal is only a part of the activity recommendation scenario flow, where the requested sum is missing in the application. The loan officer has received info about the correct sum and therefore has to update the application via the details modal, which can be triggered with the “Update application” button of application #4520 in the loan overview screen. This modal can be seen in Figure 13.

The details modal is limited to only six visible inputs, out of which only one can be interacted with. An imitation scrollbar has been added to give the modal a more in-depth feel. In addition, a “Loan details” label has been added to give the impression that other categories of information, which are usually in a loan application (e.g. the applicant’s or asset’s details) could be found if scrolled lower.

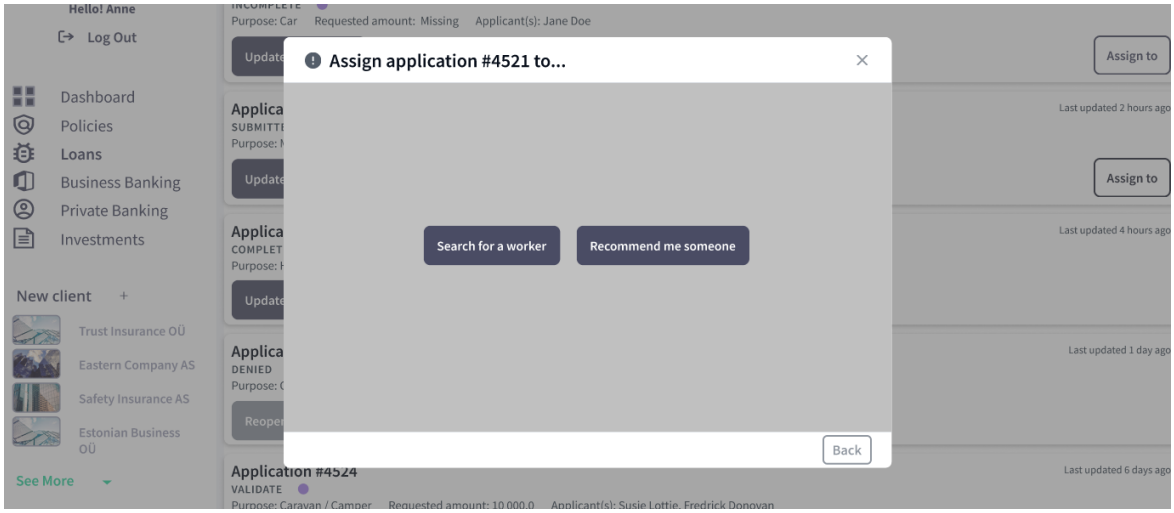

Figure 14. Assign-to selection modal

The assign-to selection modal is a part of the resource recommendation scenario flow (Figure 14). According to the scenario, the user wishes to assign the loan application #4521 to another worker. The assign-to selection can be triggered first by clicking the button “Assign to” in the card of application #4521. In this modal, the user is given an option between searching for a worker, if they already have someone in mind, or having a co-worker recommended to them. The first option for searching does not have any interactivity.

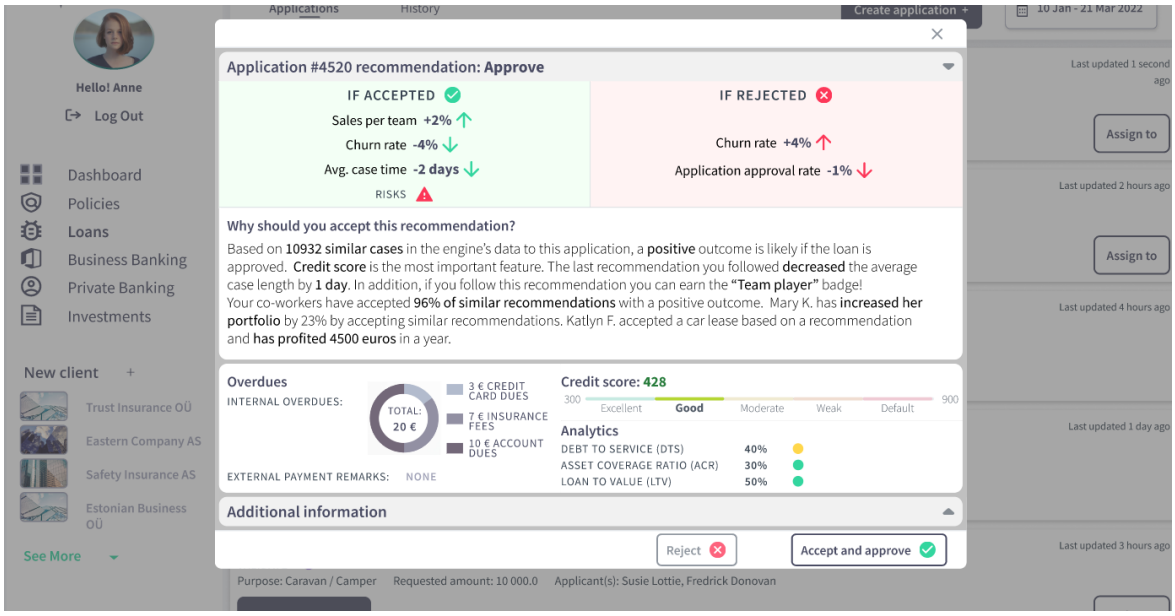

Figure 15. Activity recommendation modal

Both modals showcase the requirement R.F.3 *Within the system, the user is asked for a minimal amount of input, i.e. accept and reject*. In the details modal, the mandatory fields

of the application were already automatically filled and no other input or separate command to create a recommendation was required. As the user updates the "New loan amount" input and clicks the button "Update", the system automatically opens the activity recommendation modal. The assign-to modal can only be opened from a loan application's card and therefore the application's details are available to it. If the user clicks "Recommend me someone", the recommendation system opens with no additional input needed. This also fulfils the requirement R.F.4 *When opening the system, information should be automatically queried from the parent system*, as the user does not need to enter information into either recommendation engine again.

7.3.2 Loan Application Recommendation Modal

After a loading screen, the user is navigated to a recommendation modal, which presents a recommendation to approve case #4520. (Figure 15). To realise the requirement R.T.7, which describes the usage of business goals, some KPIs were chosen that would be realistic in a banking environment [64]. These KPIs, such as sales per team, churn rate and average case time, are presented in the "If accepted" area and others, for instance application approval rate, in the "If rejected" area.

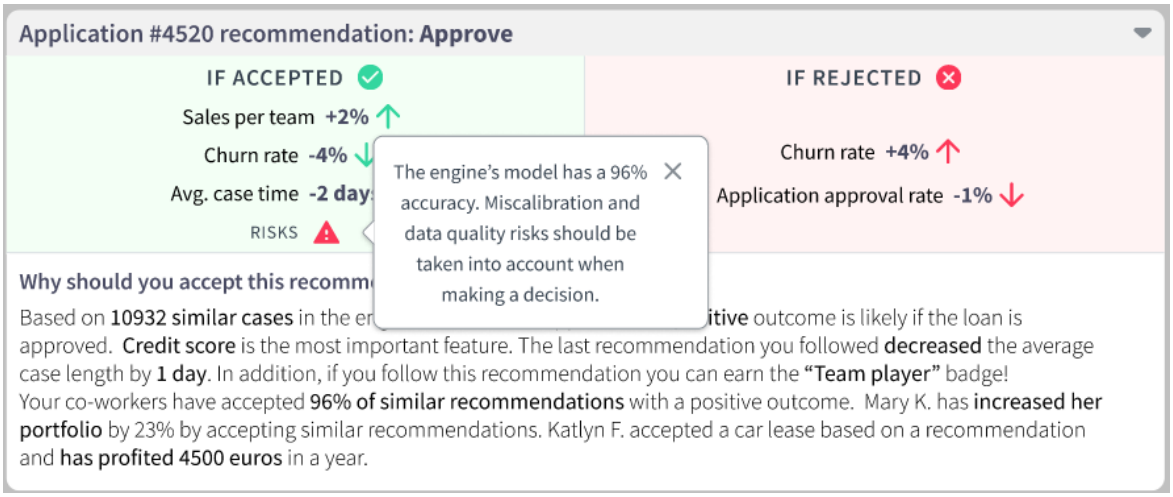

Figure 16. Loan application recommendation risks tooltip

The must-have risks covered in requirement R.T.9 are present as well. For this scenario, risks such as model miscalibration and possible low quality of the model's input data were chosen, since those are common with machine learning models [65] (Figure 16). To alleviate risks in the mind of the user and assure them of the model's trustworthiness, the model's high accuracy (96%) is shown as well. This feature covers the requirement R.T.3 *The system's overall success rate is displayed*.

The explanation area information was composed with the aid of the domain expert, for instance they proposed a range for the similar case count, the most important feature and possible profits in comparisons. The explanation includes multiple requirements:

- The requirement R.T.8 *The system displays the total number of possible prescriptions or analysed cases* is covered by mentioning the total number of similar cases the system analysed.

- To bring transparency to the decision tree model, the requirement R.E.5 *The feature weights of the prescriptive model are described* is implemented by describing the most important feature the model bases its recommendation on, which is the applicant's credit score.
- The self-monitoring requirement R.T.6 has been implemented by letting the user know their previous accepted recommendation improved the KPI average case length.
- The requirement R.T.2, which covers showing other users' data to motivate the user, and its' sub-requirement R.T.2.1, are implemented with a comparison of other co-workers' system usage and the successes of imaginary co-workers Mary K and Kaitlyn F.
- Requirements R.T.4 and R.T.5, which describe rewarding and praising the user, have been combined into a badge called "Team player", which is rewarded in case the recommendation is followed. This badge was chosen due to the team business goals the user is helping achieve if they follow the recommendation.
- As most numerical data, such as the number of similar cases or user statistics, could be presented in a tabular form with only a label and number, but instead, all information is given as proper sentences, the requirement R.E.4 *The system displays textual or hybrid explanations in favour of purely numerical ones* is fulfilled.

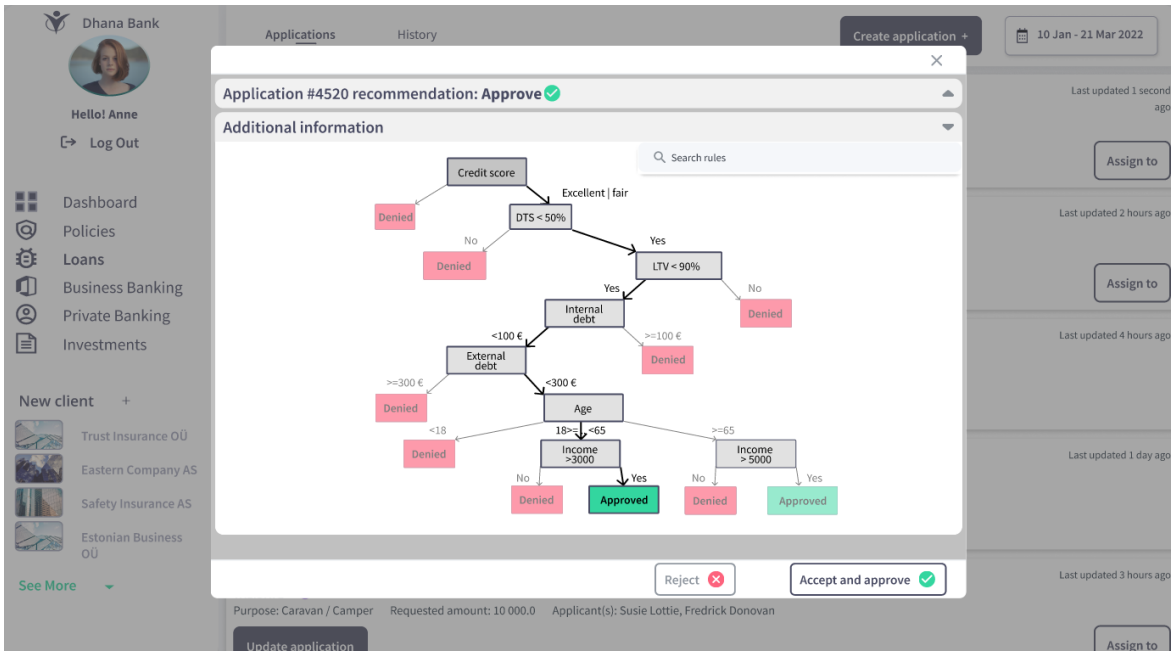

Figure 17. Activity recommendation modal decision tree visualisation

According to a domain expert, when considering a loan application, the applicant's overdues, credit score and other analytics such as debt-to-service ratio (DTS), asset coverage ratio (ACR) and loan-to-value (LTV) ratio are considered. The order of analysing these ratios might change from bank to bank, but as far as the authors of this thesis are aware, at least one bank in Estonia has usually inspected these three data groups in the particular order presented in the modal. Therefore, the requirement R.F.1 *Within the system, structure of visualising content should be the same as the decision-making process of the worker* is covered by the modal.

The bottom of the persuasive layout is the location for an accordion or multiple for additional information. In this modal, there is one accordion and upon clicking on it, the recommendation and analytics accordion closes and a decision tree visualisation can be seen instead (Figure 17). This fulfils the requirements R.E.2 *Explanations contain visualisations* and R.E.7 *The explanations fit the prescriptive model by showing evidence through a tree visualisation*. As this visualisation is usually quite large, keeping it in a separate collapsible element allows for more information to be shown. The size of the visualisation also may bring about a need for interactivity, which is covered by the requirement R.D.3. The tree can be searched for a node, which then brings the result to the centre of the modal. Thus, R.D.3.2 *The loan officer can search visualisations* is covered by this feature. The path the decision tree progressed by when creating this decision is highlighted and all the other nodes are set at half transparency. This design feature is related to the requirement R.D.2 *The most important details of a visualisation should be emphasised.* The elements in the tree and their ordering were suggested by the domain expert.

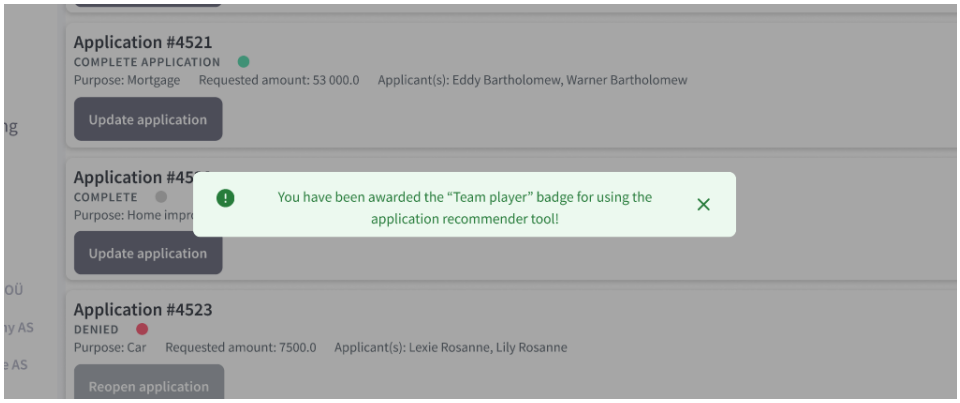

Figure 18. Loan application recommmendation reward alert

At the bottom of the overlay are buttons for accepting the recommendation (and approving the application) and rejecting it. Clicking either one closes the recommendation and updates the application's data automatically according to the user's decision. If accepted, the loan application has the status "Approved", otherwise the status is "Denied". This feature covers the requirement R.F.5 *When having reached a decision, the system automatically conveys information to the parent system*.

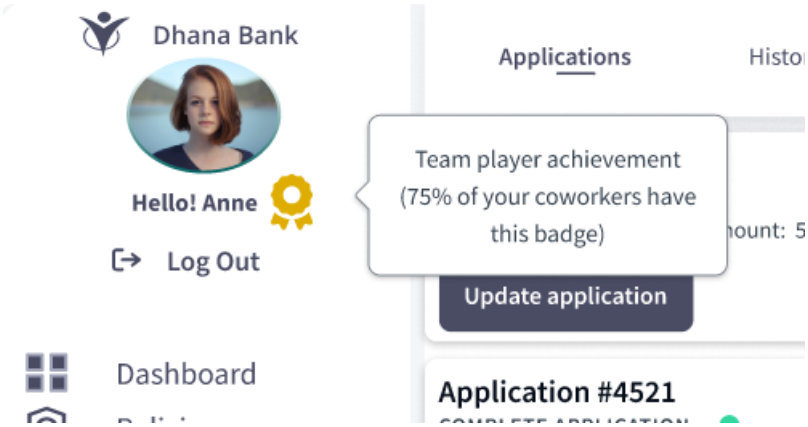

Figure 19. Loan application recommendation reward badge and explanation

If the recommendation is accepted, the user is alerted about the award they have received – the "Team player" badge (Figure 18). Accordingly, the loan overview screen has also been updated to show the user an award icon next to their profile (Figure 19). When hovering above the icon, the user is told how many other people have this badge (75% in this case), which, if the number is low enough, can be persuasive. The scenario ends there.

7.3.3 Assign-To Modal

When the user has clicked "Recommend me someone" in the assign-to selection modal and a loading screen has been displayed for some seconds, the user is navigated to the assign-to modal. A recommendation to assign case #4521 to Ron K. is presented (Figure 20).

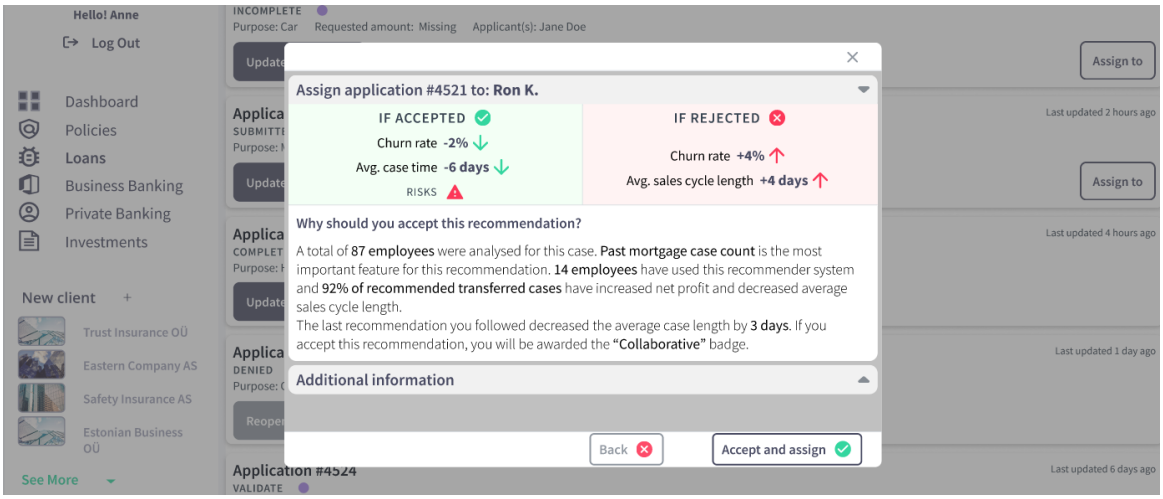

Figure 20. Assign-to modal

The must-have business goals for this recommendation are tied to the time a case can take (e.g. average case time). The must-have risk for this scenario is a case time increase if the application is misassigned (Figure 21).

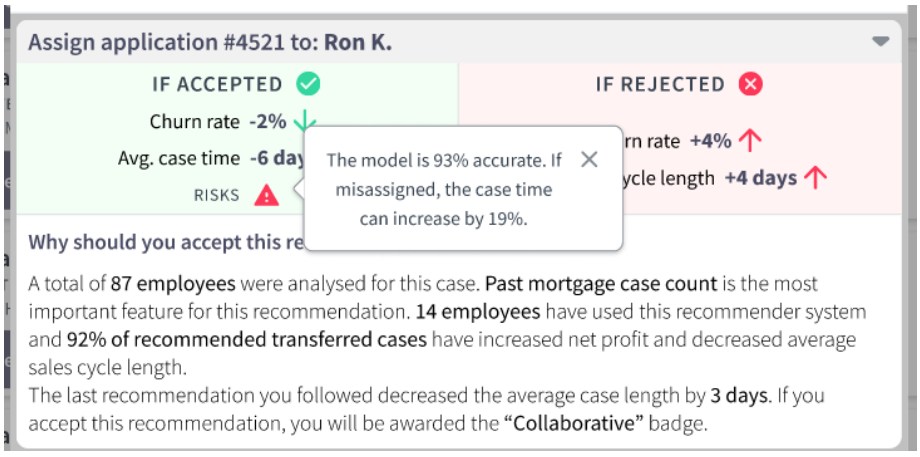

Figure 21. Risks of the assign-to recommendation

Again, as was done with the loan application recommendation modal, to lessen the risks in the mind of the user and to increase the model's trustworthiness, its accuracy of 93% is presented.

Evaluating the prototype is made easier by having a similar ordering of explanations in this modal as in the loan application modal. Some numbers and most important feature arguments were suggested by the domain expert. The persuasive messages are as follows:

- The requirement R.T.8 *The system displays the total number of possible prescriptions or analysed cases* is covered by mentioning the total number of employees considered for recommendation.
- As this recommender model is not familiar to casual users and thus may be opaque, the user is informed of the model's most important feature weight, which was chosen to be the worker's past mortgage case count. This implements the requirement R.E.5.
- The self-monitoring requirement R.T.6 has been realised by letting the user know their previous accepted recommendation decreased average sales cycle length.
- Users statistics, covered in the requirement R.T.2 and its sub-requirement R.T.2.1, are realised with a comparison of how successfully others have used the system.
- The rewarding and praising requirements R.T.4 and R.T.5 have been combined into a badge called "Collaborative", which is rewarded in case the recommendation is followed. This badge was chosen, as it indicates the user has used this system enough times and is hence collaborating with other co-workers often.

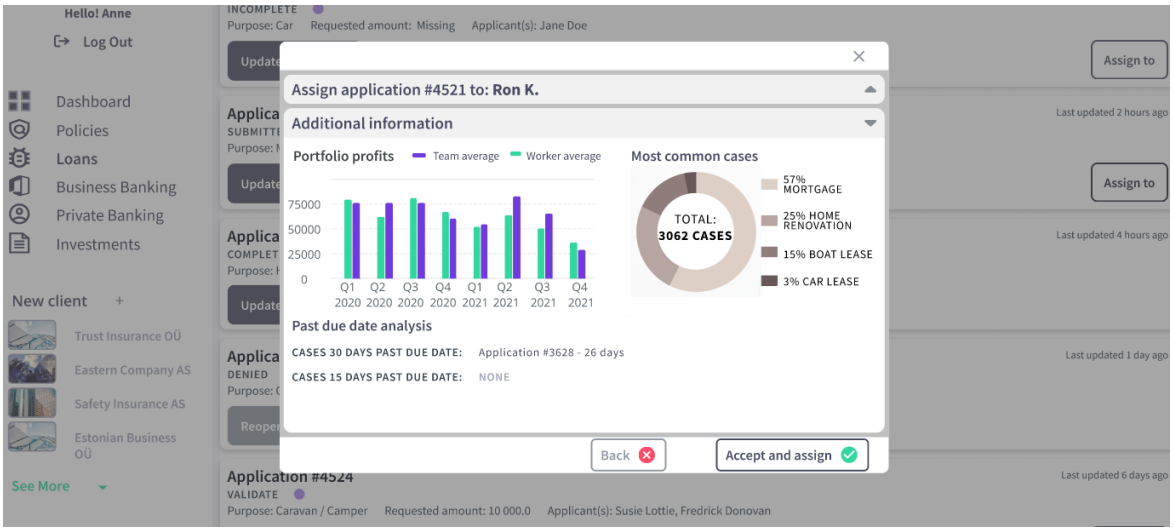

Figure 22. Bar chart in the assign-to modal

The different types of data presented in the additional information were proposed by the domain expert. In this modal, the requirement R.E.2 *Explanations contain visualisations* has been implemented with a bar and pie chart (Figure 22). The bar chart compares the suggested worker's portfolio profits per quarter to the team's average. The pie chart gives an overview of the worker's most common cases. Below the visualisations is the worker's tendency to let cases go over the due date presented as text.

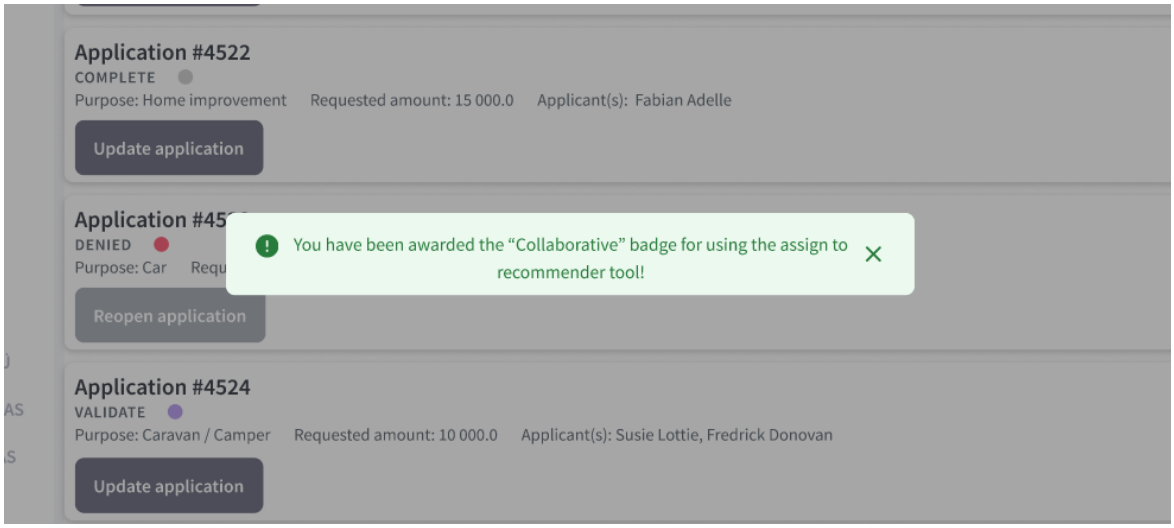

Figure 23. Assign-to modal reward alert

At the bottom of the overlay are the buttons for accepting the recommendation and approving the application, and going back to the loan overview screen. Going back does not change anything and the same recommendation can be generated again. Clicking "Accept and assign" updates the application's data automatically and the loan application card is removed from the loan overview screen, indicating it is now Ron's task.

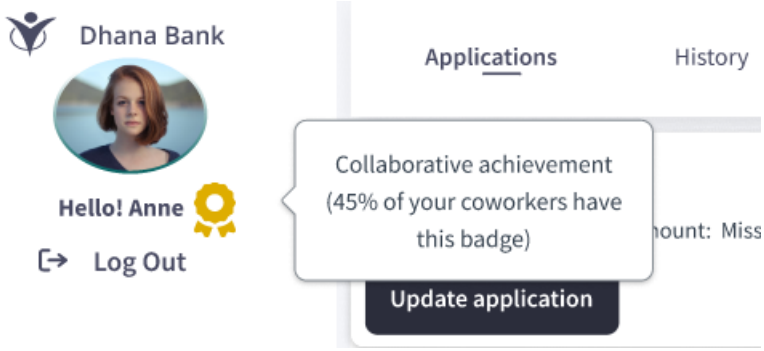

Figure 24. Assign-to modal reward badge and explanation

This feature covers the requirement R.F.5 *When having reached a decision, the system automatically conveys information to the parent system*. As it was with the loan application modal, if the recommendation is accepted and the modal has closed, the user is shown an alert about the reward they have received. This can be seen in Figure 23. Assuming workers don't use this recommender system daily like they do the other, this badge is much rarer (Figure 24). Scarcity can make the badge seem more rewarding. The scenario ends there.

8. Evaluation

This section describes the evaluation of the prototype developed in the previous section. Section 8.1 outlines the user study findings. Section 8.2 lists possible improvements to persuasive principles and requirements.

8.1 User Study Findings

In this section we will first describe findings related to the loan application recommendation modal (section 8.1.1), then the assign-to recommendation modal (section 8.1.2) and finally general findings that apply to both modals (section 8.1.3). The findings are ordered and grouped according to the persuasive layout (section 7.1) with general feedback at the end. The If Accepted/Rejected block has been split into metrics and risks for better understandability. The grouped findings are assigned codes for further referencing, e.g. L-IAR (loan-if-accepted-rejected), L-R (loan-risks), A-MI (assign-missing-info), A-WSYA (assign-why-should-you-accept) and GF (general-feedback).

8.1.1 Loan Application Recommendation Modal

This recommendation was accepted by all participants except one (I2). The reasoning for the decisions is outlined below.

The **If Accepted/Rejected** block had a largely negative reaction (**L-IAR**). Most participants (I3-I8) did not find it useful in their decision-process, as the KPIs do not influence a loan decision ("*For decision-maker it doesn't give me any information*", I5). A common criticism was that the KPIs were too team- or sales-centric and should be more focused on the customer ("*It is not relevant to me how the team is doing*", I7; "*We're not awarded for higher sales… I'm used to disregarding the sort of sets of benefits*", I6). However, had they worked on a sales team, they would have found this data very useful, since "*It's important to know how much this case is impacting the sales volumes*" (I8). Overall, this block was perceived to be the least important by the participants.

Although the block reportedly had little useful information for decision-making, one participant still found that the "*churn rate might be useful*" (I3). Two participants, I4 and I8, appreciated that the block existed, as it shows data has been analysed for them ("*It shows me the data has probably been analysed for me and showed me the result*", I4) and helps keep a perspective on the case time ("*accepting would speed the process up*", I2). Some participants had trouble understanding the KPIs, especially the term churn rate (I1, I3, I6, I7). Participants I1 and I2 found the positive and negative number design (for example "*churn rate -4%*" under positive changes) confusing, as for them a minus implies a negative impact and vice versa.

The **risk popup** inside the If Accepted/Rejected block was initially seen only by three participants (I2-I4). These findings belong to the **L-R** grouping. Although the other five reported seeing the red icon and wondering what the risks entail, they did not conclude that a popover can be activated. I5 remarked that the icon pulled their attention right after the recommendation modal was opened even though she/he did not understand it ("*…it's some kind of big mark, it's always high importance information, a red flag…*", I5).

The risks were received with mixed feelings. For four participants (I1, I3, I4, I8), the accuracy displayed in the risk popover felt useful and increased their confidence and trust in PERSEVERE. I6 thought the risks did not influence his decision, as they decide purely based on case-specific statistics. The rest of the participants stated the risks decreased their trust in the system, since the model's accuracy must be 100% ("*If an engine is built, then it should be working 100% of time*", I7) or because of the red icon design, which without triggering the popover made one participant feel less confident ("*there might be some data the automatic program can't find*", I2). One participant mistook the recommendation's risks as the car's risks ("*I understand it's about the car and engine*", I4). The participants I5 and I6 wanted more information about how the accuracy is calculated ("*Without knowledge of the model, it's difficult to trust it [accuracy]*", I6).

The **Why Should You Accept this Recommendation?** block received contradictory comments (**L-WSYA**). All participants found it to be useful to a varying degree, for example I1 and I4 said it to be overall the most useful part of the recommendation, while other participants found specific parts of the text to be of use. One participant would remove the explanation altogether, as it did not contain enough case-specific info ("*This statistic is relevant if you're analysing how your system is working*", I7). Instead I7 suggested it should contain more statistics about their work and decision quality, e.g. how quickly they are dealing with cases and with whom. The participant I2 made a similar remark, in which the interviewee said the whole text is not suitable for a work setting due to it reminding them of scam websites ("*It's similar to fake investment web sites… And on trustworthy sites like in a workforce it seems out of place*", I2). I5 did not read the text when first making the decision, as they said "*numbers always spoke to me more than text*" (I5), but later remarked it was nice to have, although not necessary for decision-making.

All the comments participants had per persuasive arguments can be seen in Table 3. The similar case argument was reported as the most well-liked. One argument, the badge, received only negative comments.

Table 3. Feedback regarding persuasive arguments in the loan application recommendation modal

Persuasive argument	**Comments**
"Based on 10 932 similar cases in the engine's data to this application, a positive outcome is likely if the loan is approved"	• The participants I3, I5 and I8 were assured by the system using a lot of data and statistics to make a recommendation, which increased their trust in it ("*More than 10 000 cases, it's quite okay to me*", I8; "*10 000 similar cases is quite a big sample set*", I5)
"Credit score is the most important feature"	• The participant I8 felt assured by the model's most important feature, because "*the text is confirming my historical experience*" (I8)

"In addition, if you follow this recommendation you can earn the "Team player" badge!"	• I3 thought the badge's benefits are not clear ("*I'm [supposed to be] more motivated?*", I3) • Participants I6 and I7 said the badge is not relevant for decision and is therefore useless ("*Depends on the organisation. Where I work, these badges… are not useful*", I6) • I8 argued that the badge's effect depends on the person using the system, but for them it had no purpose • I8 stated that had the badge been more focused on the user's portfolio quality and less about the team, it could have been of more use. Otherwise it may impact the user's decision in an undesirable way ("*If it's a borderline case then maybe the employee would still approve to get the badge…*", "*…it should be more about the personal portfolio quality revenues*", I8)
"Your co-workers have accepted 96% of similar recommendations with a positive outcome"	• I2 and I4 both thought the 96% of successful recommendations was assuring and trustworthy • I2 added this argument could undermine trust in the system, as "*in sketchy sites there are similar things like '99% people recommend this or that'*" (I2)
"Mary K. has increased her portfolio by 23% by accepting similar recommendations. Katlyn F. accepted a car lease based on a recommendation and has profited 4500 euros in a year."	• zzI2 claimed the comparison with other co-workers is another tactic used by scam websites and therefore decreased their trust in the system ("*Previously all these types of sentences have seemed untrustworthy to me*", I2) • I2 thought the comparison with Mary K. put unnecessary pressure on them to accept the recommendation • The comparisons to Mary K. and Kaitlyn F. seemed useless for participant I3 as these arguments did not help decide better ("*It doesn't mean I can analyse more cases or make more money*", I3) • I1 felt the comparisons had an impact on their decision to accept the recommendation. In a similar vein, I5 felt it helped them feel less alone ("*it's quite valuable for a decision-maker that he's not alone and is thinking in the right way*", I5)

Despite the arguments being found useful or trustworthy, a few participants still did not trust the explanation nor the recommendation. For example, I2 felt the "*Your co-workers have accepted 96% of similar recommendations with a positive outcome.*" (I2) argument helped them feel more confident in their decision, but also found that it undermined that confidence. To some participants the arguments also seemed too sales-focused, too long or too narrative

("*the sales data and badges were a little distracting*", I4; "*I would like to see detailed statistics of similar cases… It's still too narrative*", I6; "*It's a lot of text*", I7). One participant, I6, described the arguments as too "*personal*" for making a credit decision. I7 said the word "*should*" in the title "*Why should you accept this recommendation*" was too strong and made her felt pressured to accept the recommendation.

There was some confusion among participants about the terms used in the text. For example, two participants wished the term "*positive outcome*" was more explicit (I8, I6). I1 misinterpreted the most important feature argument as a sort of guidance, not as a feature the prescriptive model used the most. The participant I6 asked for a comparison between the current and model's sample case average statistics, for instance the risk metrics and overdues. I6 also wished the term "*similar cases*" to be expanded upon as it was too vague currently ("*What is the default rate for those? Which have substantial overdues?*", I6).

The block **Statistics** was received well (**L-S**). Seven participants, I1 and I3-I8, felt it was the most useful and one thought it was useful, but not persuasive enough ("*there should be something that makes the client less weak in their profile… If the credit score was excellent, I would accept it automatically*", I2). Although two participants did not specify the most important feature for their decision (I2, I3), four listed credit score (I1, I4, I5, I7), I8 credit score in combination with the debt-to-service ratio and I6 the loan-to-value ratio as their most influencing feature. Thus, this block was perceived as the most important for the decision.

Although I3 and I6 did not know some of the statistical ratios included, for 87.5% of the participants the statistics shown in this block were good enough to accept the recommendation. The same applies to three participants who wished to know how the credit score is calculated (I3, I6, I7). One of them, I3, presumed the credit score included all the data they usually use, such as "*the client's obligations in other banks*" (I3). In addition to a strong credit score and analytical ratios, another reason for accepting the recommendation were the small overdues. The only participant to reject the recommendation, I2, decided so due to a subjectively weak credit score.

Participant I4 suggested a different ordering of elements in this block. They said it is "*confusing that the credit score is after this [If accepted/rejected] part*" (I4) and therefore the credit score and the loan-to-value ratio should be prioritised more. The interviewee also wanted the loan sum to be displayed in the modal. Two people, I2 and I7, commented on the design of the overdues, which they felt was too unnoticeable. This applies more to the external overdues, which some people seemed to miss based on their answers, e.g. I3 was worried about external overdues not being displayed. I2, I4 and I7 found the design to be clear and easy to understand, and especially commented on the colour-usage and credit score visualisation ("*It's easy to see the credit score is good, excellent or whatever*", I4). The yellow dot used to signify the debt-to-service ratio mediocre level was confusing for I2, as well as the credit score range of 300 to 900 ("*Usually I would think that the highest score is the best*", I2).

The **Additional Information** block received positive feedback (**L-AI**). Due to the additional information containing more information about the customer and how the recommendation was made, all participants thought it was useful to a degree and two of them, I2 and I5, even preferred it over the information presented in the first collapsible accordion ("*it helps exclude good and bad criteria… if it shows the data automatically, like age and income, it is really useful*", I2). This block allegedly increased the trust and confidence of four participants I1, I4, I6, I8 ("*it gives additional confidence… it is considering all the key parameters for decision-making in a reasonable way*", I8; "*it would be useful in real life work*", I4). I3 thought the visualisation was well designed and easy to understand ("*If it [application] doesn't get automatic approval, you can go and look at it where it went wrong… So you can ask more questions from the customer, like 'why did you have these problems?'*", I3).

On the negative side, only participant I2 noticed and opened this block during decision-making. Although I2 stated the information shown is beneficial, especially to a beginner worker ("*in banking there are so many terms… it's a lot to remember, but this will be really convenient*", I2), another participant, I7, thought the visualisation was too complex to read for the first time, because "*everything is new and confusing*" (I7). They mentioned repeated use of PERSEVERE would aid with such an issue. Two participants (I7 and I6) would have preferred a tabled sheet of data instead of the visualisation, with I7 calling themselves an "*Excel person*" (I7).

In general, seven participants (I1-I3, I5-I8) felt there was some **missing information** after seeing all the available data and visualisations (**L-MI**). All seven of them wanted more data about the customer, for instance their income, job position, liabilities and education. I8 also wished to know more about the asset, the car. I2 specifically also wanted to see the loan period and a comparison of their past decisions and the current recommendation ("*compare the recommended and my decisions, how they match or mismatch*", I2).

Table 4. Confidence, understandability of and trust in PERSEVERE during loan recommendation

Code	Confidence	Understandability	Trust
I1	Confident about decision ("*I made the right decision*", I1)	Didn't understand the system due to missing data	Trusted the recommendation
I2	Felt their confidence was dependent on "*if I had a bunch of work I would more easily accept*" (I2), but was more or less confident in current decision	Understood the recommendation engine and why such a recommendation was made ("*I understand it is based on previous cases… and how likely it is to be a good decision if accepted*", I2)	Didn't trust PERSEVERE. They explained that "*from previous work experience the automatic engines usually miss a lot of important things*" (I2). Repeated use of PERSEVERE and their future decisions matching

			the recommendations would increase trust
I3	Wasn't confident as they felt information was missing ("*let's say she just took a huge obligation for another loan from another bank… that would change everything*", I3)	Understood the recommendation engine and why such a recommendation was made ("*it took some time… it's basically automated system to make the approval by itself*", I3)	Trusted the recommendation
I4	Confident about decision	Didn't understand PERSEVERE as the explanations were too sales-oriented	More or less trusted the system, since it is comfortable and does the analysis for them
I5	Confident due to a high credit score	Did not understand the recommendation engine	Trusted the recommendation, as they said "*if it's shown to me, I need to believe it*" (I5). The recommendation was nice to have, but it should only be supportive to the final decision. External audits would help trust it even further ("*some third party, official institution needs to approve this scoring model, then I can believe*")
I6	Wasn't confident when making the decision, but was more so during the interview after seeing the additional information, because "*the most important aspects are covered in this visualisation*" (I6)	Understood the recommendation engine and why such a recommendation was made	He will inherently try to trust an internal system ("*If I'm told to trust the model, then there's very little for me to do*", I6). Suspects PERSEVERE has not looked at all available data and needs to know the model has been tested and approved to trust it, otherwise the human decider wouldn't be necessary at all ("*if I know that [the data it used] or who did this recommendation system… I believe that if the model is tested and approved and one can rely on it*", I6). They also wanted

			to know the age of the system and wished for a more in-depth explanation of the model and its flaws
I7	Was not confident due to missing data, such as credit score formula	Did not understand the engine ("*why is credit score not excellent? What's missing? What are the analytics?*", I7)	If they are told to trust PERSEVERE, they will do so. Trusted the system inherently ("*I have to trust it if the bank made it*", I7)
I8	Confident in decision ("*besides the fact it's a standard transaction with a conservative loan-to-value ratio… then looking at the income level, credit score, no external payment remarks… overall picture looks good to me*", I8)	Understood the recommendation and engine	Trusted the system

Overall, the **confidence**, **understanding** and **trust** of the participants in this recommendation were mixed (**L-CUT**). Details can be seen in Table 4. Half of them were confident in their decision, although some participants remarked external audits would make them more confident. I2 said they were more or less confident, depending on the situation. The remaining two were not confident (I3, I7).

In summary, the loan application recommendation was accepted by 87.5% of participants. Its elements received varied feedback. The If Accepted/Rejected block received criticism by many for not containing useful information for the decision. Some participants struggled to understand the KPIs. The risks displayed within this block created confusion for some and was given contradictory feedback regarding trustworthiness. Participants were similarly divided on the usefulness of the Why Should You Accept this Recommendation? block. The compulsory information contained in Statistics was well-received by almost all participants and the same applies to the Additional Information part of the modal, which many preferred over the first two. Some participants proposed suggestions for improving the content (e.g. to include certain missing information) or to the design of the modal. Most participants felt confident in their decisions and trusted the recommendation, but only half of them understood how the engine worked or why such a recommendation was made.

8.1.2 Assign-to Modal

This recommendation modal was, similarly to the previous one, accepted by all participants except I5. In this modal, I5 experienced deep confusion regarding the recommendation. Further details are below.

The **If Accepted/Rejected** was found to be useful by three participants mainly due to the average case time KPI (I2, I3, I4). These findings are categorised under the **A-IAR** group. Of those participants who mentioned having a most important element influencing their decision, I5 stated the average case time in this block to be theirs as they were primarily interested in the loan application's quick resolution. As was the case with the first recommendation, this time there was also confusion around the term "*churn rate*" (I5). I2 found that the churn rate KPI was not useful and in a similar vein I8 thought it was too obvious. One participant, I7, was confused who the KPIs were about ("*if we are talking about this concrete mortgage case, if accepted and if rejected, so does it mean that similar cases were accepted in six days and similar cases in four days?*", "*is it about Ron?*", and "*if this is about this concrete mortgage case then again for me it's unclear how it can be rejected and accepted answers for one case*", I7).

Risk associated findings of the assign-to modal belong to the **A-R** grouping. The risks popup was opened by three participants – I3, I6 and I8. For I3, the case time increase was worrying, but for I6 and I8 the accuracy was good and risks were low ("*The model is 93% accurate, not bad at all. If misassigned the case time can increase. Well that's obvious*", I8). I2 explained that "*a mortgage loan takes a bunch of time anyway, then the time increase by 19% doesn't seem that high*" (I2). In this modal the participant I5, despite not opening the popover, said the risks' red icon was large and worrying, due to which their confidence decreased ("*I see under if accepted there is a quite high risk*", I5).

Table 5. Feedback regarding persuasive arguments in the assign-to recommendation modal

Persuasive argument	**Comments**
"A total of 87 employees were analysed for this case"	• I4 found the total number of analysed employees to be useful, as it proved the model was objective • I2 appreciated the number of analysed co-workers, since it made them feel "*it seems like more time has been put into it [the recommendation]*" (I2) • I6 wanted to know how many employees there are in total to better understand it • The word "*analyse*" seemed like a smarter and more trustworthy word to I2 when comparing it with the similar case argument in the previous modal
"Past mortgage case count is the most important feature for this recommendation"	• I8 felt assured by the most important feature, as they thought it was the most important as well ("*if it would have been something about sales increase, for instance, or maybe even average case time… I wouldn't feel so confident.*", I8). I8 explained that "*then I wouldn't know if the recommended person would make really high quality decision or there may be theoretically the possibility that he or she just makes fast decisions without a thorough analysis or considerations*" (I8) • I8 and I4 regarded this argument as the most important for their decision

	• The feature seemed reasonable to I6, which in turn made this argument useful
"14 employees have used this recommender system and 92% of recommended transferred cases have increased net profit and decreased average sales cycle length"	• I3 said the amount of other employees who have used PERSEVERE was "*a bit worrying*" (I3) • Data about other employees made it more trustworthy for I2
"The last recommendation you followed decreased the average case length by 3 days"	• I7 stated that a decrease in case time is important for the customer, which is why this argument was the most useful • Personal data felt useful to I3 and I2 ("*I have previously used the system and decreased [average case length]. Oh, that's great. I'm gonna use it again*", I3;" *I think this leaves a better impression, because this part of this text is based on my previous work, but at that car loan, it was like based on someone else. So I think it's more personal, that it shows how your work has been more effective or better with automatic decision or like automatic recommendations*", I2) • I1 referred to this argument as the most important influence to their decision
"If you accept this recommendation, you will be awarded the "Collaborative" badge"	• Although receiving a badge made I1 feel "*happy*" (I1), it is useless to them ("*it isn't saying [anything] to me and I think it doesn't help me to make the decision*", I1). They would not make a decision based on the badge award • I2 thought the badge is too gamified and unfair, as they explained "*you want to look at the applications yourself and then you find something that's important that automatic system missed. So because you're doing like the right thing, trusting your gut you miss a badge that actually doesn't give you anything*" (I2) • I5 felt confused about the necessity of the badge

The **Why Should You Accept this Recommendation?** block received positive feedback from 87.5% of participants (**A-WSYA**). Only participant I7 found it to be semi-useful due to it not containing statistics about Ron and it having numbers that are too generic ("*if we are talking about customer consultants, and then if we are in one team, then I will not start to think how often he's getting a positive answer. I'm thinking how quickly he is taking this on the table, and also how quickly he's responding and dealing with it*", I7). All specific comments about the text are shown in Table 5. I6 said the block was useful in general. I1 found this explanation easier to understand compared to the previous one ("*I noticed the important text better*", I1). Two participants, I5 and I4, were initially confused about the subject of the explanation, mistaking it to being about Ron, not the recommendation ("*so I understood that he's experienced in mortgage loans*", I4). Out of the five mentioned most

important features, four were from the explanation text, so this was overall perceived as an important block.

The **Additional Information** related findings are grouped under **A-AI**. It was clicked on only by two participants during the initial decision, I2 and I7. The block got mixed reactions. Five participants found it to be useful (I2, I3, I4, I7, I8). More specifically, I7 appreciated the portfolio comparison, because "*if there was one line, then it wouldn't say nothing to me because I cannot compare it to anything else, but the average [team profit] is lower then it means that he's [Ron] a good worker and he gets more profits*" (I7). I2 reported liking Ron's due date analysis and case count visualisation. However, for interviewee I5 Ron's data had no value, since they were more concerned about the speed of the case being resolved and did not care who would get it done ("*it's important the client gets result, positive or negative… it's important the case is worked on quicker than if I don't assign it to another person*", I5). I1 did not find anything useful to support their decision-making from the portfolio visualisation ("*I don't know what I'm doing with that information*", I1). Three participants were initially confused who the additional data was about (I2, I3, I7).

I2 suggested naming the portfolio profits visualisation header to "*Ron's portfolio profits*", because they found that Ron's info was not explicit enough. The interviewee felt the portfolio profits visualisation can be misleading if Ron only works with smaller loans compared to his team ("*he can just have customers with smaller loans, and then the profit will be smaller*", I2) or if the worker is new in the company. I2 suggested not showing the current quarter as that can be misleading because it's "*it's going to be smaller anyway*" (I2), despite the current quarter not being shown in the visualisation.

The participant I6 wanted more information about Ron's cases, for example how many active and closed applications he has, and what quality his decisions have, e.g. are his approved loans paid back, are they on time and etc. I6 also wanted to know how the portfolio profits were calculated.

A few participants felt some **information was missing** from this modal (**A-MI**). I7 wished to know more about Ron's decisions, how quickly he works on cases and more information about the mortgage loan itself. In order to avoid creating trouble for Ron, I2 wanted to know if he was currently available and "*how much time would it take for him to start with the project [case]*" (I2). I6 wanted to know more about Ron's past mistakes.

One user, I5, rejected the recommendation initially, since they thought it was a loan application approve/reject decision. This misunderstanding was due to their poor English skills. When I5 realised the goal of the modal the interviewee stated that they would have accepted it instead. This influenced their responses during the interaction and interview. For example, in the beginning I5 felt there could have been more information about the case, similarly to the previous modal (e.g. analytics, customer's income). This prevented I5 from making a confident decision.

Table 6. Confidence, understandability of and trust in PERSEVERE during assign-to recommendation

Code	Confidence	Understandability	Trust
I1	Confident about decision	Understood the recommendation and PERSEVERE	Trusted the recommendation
I2	Confident due to there being less risk with this decision	Did not comment on understandability	Trusted the system to recommend a specialist
I3	Confident about decision, although initially wanted a second recommendation. Accepted easily as there were no possible negative effects	Said the recommendation is new to them, so understandability will come with time. The recommendation "*made sense to me after like, reading about it a few times*" (I3)	Trusted the recommendation
I4	Confident about decision	Understood the engine and recommendation because "*it was really about colleague*" (I4)	Trusted the recommendation and presumed that Ron would handle the case quickly. Appreciated that it gives "*objective information to make a decision*" (I4)
I5	Not confident during the misunderstanding, confident after I5 understood the goal of the recommendation. They did not care if the recommendation was correct due to it being an in-bank low risk decision ("*maybe it's Ron, maybe it's David, it's not important to me*", I5)	Did not understand the recommendation	If a model uses statistics, I5 trusts it due to its objectivity ("*if it shows me the volume how many applications it will go through… then I can believe*", I5)
I6	They felt they didn't have much to rely on for the decision, but what they did have were "*positive arguments*" (I6). Reported themselves as both not and semi-confident about the decision,	Has not used such a system and did not understand it well	They had no alternative and therefore accepted the recommendation

	but approved since there was no reason to reject		
I7	Was not confident due to missing data, but accepted due to low risk decision	Did not comment on understandability	Inherently trusts the recommendation ("*I trusted it because it was there*", I7)
I8	Confident in decision due to low risk decision for bank and client ("*there is no risk to the to the bank, no risk to the customer*", I8)	Understood the recommendation surprisingly well for first time usage ("*surprisingly, because I do not have any personal experience with any that kind of recommendation tool*", I8)	Trusted the system as it's arguments aligned with his

Overall, the recommendation seemed like a useful tool to many participants. The participants' confidence in their decision, their trust in the recommendation and level of understanding are presented in Table 6. These findings are coded as **A-CUT**.

In conclusion, the assign-to recommendation was accepted by 87.5% of participants. Its elements received mostly positive feedback. One participant demonstrated major confusion regarding the purpose of the recommendation. The If Accepted/Rejected block and the risks it contains had both positive and negative reactions. Again, some participants struggled to understand the KPIs. Almost all participants found the Why Should You Accept this Recommendation? block useful. The Additional Information part of the modal created varied reactions. A few suggested including certain information. Over half of the participants felt confident in their decisions, most trusted the recommendation, however only half of them understood how the engine worked or why such a recommendation was made.

8.1.3 General Feedback

All eight participants answered the questionnaire. Our interviewees gave generally positive feedback in the questionnaire. The total average score of perceived usefulness, ease of use and satisfaction of using PERSEVERE can be seen in Figure 25. Out of a total of 128 answers, only around 7% were negative and around 9% neutral. Participants agreed most with the statements "*I would find PERSEVERE easy to use*" and "*It would be easy for me to become skilful at using PERSEVERE*" and least with the statements "*I would find it easy to get PERSEVERE to do what I want it to do*" and "*I would find PERSEVERE to be flexible to interact with*".

PERSEVERE received an average score of 4.2 out of 5 in usefulness. This means most participants believe to a degree that PERSEVERE would enable them to work quicker, improve their performance, productivity and effectiveness on the job. They also reported a degree of belief that it would make their job easier to perform.

The highest average score of 4.4 was given to ease of use of PERSEVERE. All participants thought to some degree that learning and mastering PERSEVERE would be easy for them,

as well as that they would find it easy to use. They also reported to agreeing that using PERSEVERE was clear and understandable. Over half (57.5%) thought it easy to get PERSEVERE to achieve what they want and believed it to be flexible. A few participants (25%) remained neutral regarding some of the questionnaire points and one participant (12.5%) disagreed with most.

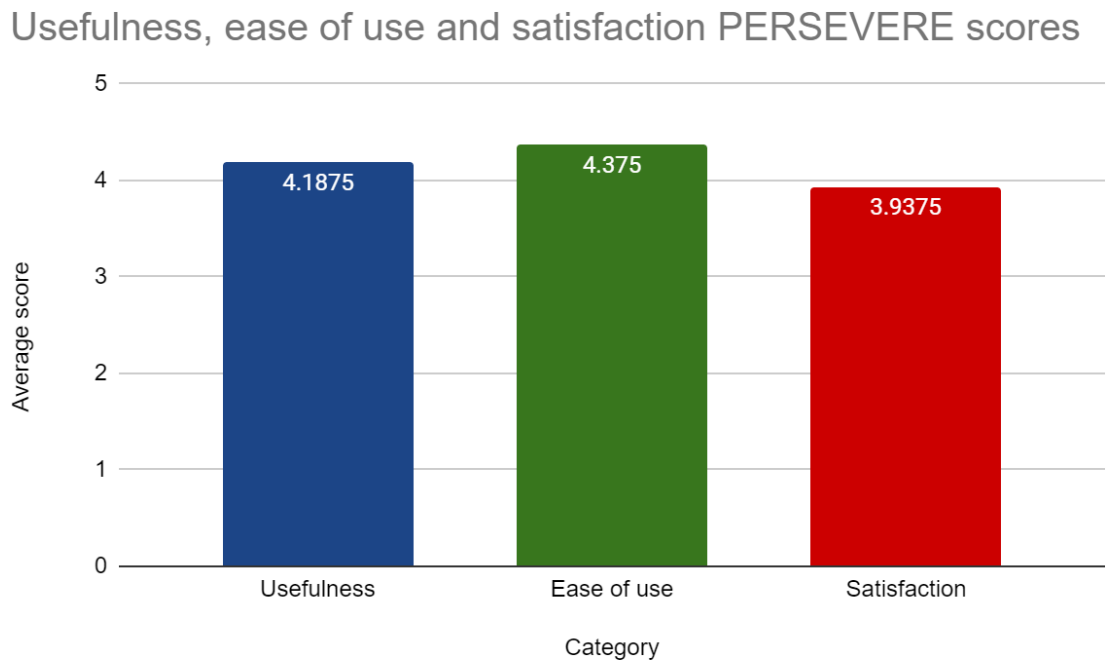

Figure 25. Usefulness, ease of use and satisfaction of using PERSEVERE

The interviewees were also generally satisfied with using PERSEVERE, as in this category, the average score was around 3.9. Most interviewees felt somewhat satisfied with PERSEVERE. A few (25%) felt very pleased and one participant (12.5%) very contented and delighted with PERSEVERE. The feeling of being somewhat pleased, contented and delighted was reported by 62.5%. The rest were neutral.

All participants were asked to comment on the design and structure of the persuasive layout and elements. The resulting findings are grouped under the code **GF**. All participants liked the colour-design of the elements, specifically the use of green, yellow and red in the modals ("*I like the yellow, green and red dots. It indicates red is always important and I need to look through it more, green is always positive and if the system uses it, I can believe that… Yellow should maybe be looked through to make final decision*", I5; "*it's easier to catch the things that are coloured*", I2). Participants I5 and I7 appreciated the "If accept/reject" colour-coded areas ("*how it [the block] looks is clear*", I7). I2 mentioned the icons and arrows drew their attention well to the necessary information. The credit score visualisation got praise from a few participants. I8 thought the design in general is simple and modern, unlike many current systems in use in Estonian banks at the moment, but not "*over-visualised [with] adding lots of special effects and lots of colours*" (I8) so it is suitable for everyday use. I8 commented that these sorts of tools should be designed with daily use and the bank's design in mind.

For the explanation blocks, one criticism was that the thicker font did not summarise the text enough, because it was used too often ("*some parts I think catch my eye but I still read the whole text anyway*", I2). I7 commented about the explanation that "*I have to read through very, very many fields*" (I7) which was problematic. However, I3 appreciated that the explanation allows the user to see further details. Similarly, I8 said that having access to more details is always a positive thing.

The Additional Information block went unnoticed for both modals and many participants. Some of them suggested removing the collapsible design and replacing it with a scrollbar (I1, I2), but I6 also remarked that this is only a first-time use issue ("*During a second time already, this wouldn't be a problem*", I6). I4 felt similarly towards the risk popover, which many participants missed. The participant also suggested that "*the risks part is kind of bigger when there are really big risks. It can just flash*" (I4).

I1 wished the modals were larger, so more information could be contained within. I2 suggested making the collapsible header fonts to be made bigger so they would stand out more, because "*it's grey and matching with the [explanation] part that has a lot of text so that [header] can definitely go on unrecognized*" (I2). They also wished that the most important information of the recommendation for them, the "*loan amount and the loan term*" (I2), were emphasised more. The system should have an introduction, thought participant I4. I7 felt both modals would have benefitted from better chosen KPIs ("*if there would be better numbers [KPIs] and there were more relevant things then it's good to read those numbers that go up and down*", I7). I4 suggested moving the Why Should You Accept this Recommendation? block to be the first and the If Accepted/Rejected block to be the last, as I4 prefers to first read and gather information and then make a decision. Additional information, which was useful for many in both recommendations, should be prioritised according to I6.

The feedback for visualisations was mostly positive. Participants appreciated the visualisations and I8 stated that for a tool that many people will use, well-designed visualisations will always be useful ("*if it properly highlights the key things then you can always first grasp the key, the real core of that case*", I8). The exceptions were I5 and I7. I5 felt strongly that "*numbers always spoke to me more than text*" (I5) and visualisations. I7 had similarly commented about preferring numbers earlier in section 8.1.1. Visualisations are acceptable when it allows the user to see "*all information in one picture*" (I5) and, for I8, only if they complement the important data for decision-making, i.e. when "*the visualisation will not take over the content*" (I8). No users used the details-on-demand feature of the visualisations.

In summary, the visual design of elements received praise from all participants and were given many small suggestions for improvement both content- and layout-wise. Many specifically commended the colour usage and credit score and decision tree visualisations. Both explanations were criticized due to the amount of text and overuse of bold text emphasis. The interactive blocks (risk popover and collapsible accordion) were noticed by very few participants, but with repeated use this would not be an issue. The ordering of elements was

another topic some proposed changes to. Participants found PERSEVERE to be mostly useful and easy to use and felt satisfied with using it.

8.2 Improvements

Based on the findings, we amended and developed an additional set of requirements. The requirements were formulated in the same way as in section 6.3. Some possible improvements were not made due to various reasons, which are listed under section 8.2.1.

Keeping the findings in L-WSYA, L-CUT AND L-S in mind, the content and prioritisation of the requirement R.E.8 was updated. During the first times PERSEVERE is being used, any explanations containing numbers may seem too similar to scam websites. The participants did not implicitly trust statistics shown to them in the Why Should You Accept this Recommendation? block and wanted more clear definitions to terms such as "*positive outcome*" and "*similar cases*". Our presumption that the users would know the data's source and meaning was false in some cases. Therefore, the requirement R.E.8 was defined as *Explanation mentions sources of prescription and explanation data (e.g. data warehouse, external credit bureaus, formulas)* and belongs to the should have priority group. With repeated use of the system, users will understand the sources and terms more and this problem should lessen. This requirement also helps alleviate some issues covered in L-WSYA and L-R, for example by adding links to the other users who are referenced in R.T.2, as well as to their decisions or to a summary of analysed cases in R.T.8. Any formulas that were used to calculate ratios, accuracy etc. can be added as popovers.

The requirement R.E.9 was added to the R.E category. It is a could-have requirement as the system may not be externally audited yet and therefore it may not be possible to implement. If it can be added, however, it could help users feel more trusting towards PERSEVERE.

For the system design requirement updates, we referred to findings in L-IAR and L-S. The requirement R.D.9 *The system should provide comparisons of simulations of the possible after-effects of following and not following the decision* was divided further into a sub-requirement R.D.9.1 for specifying the design of the KPIs. It has the same priority as its parent requirement. In L-S, the grey overdues visualisation received criticism, therefore an additional requirement specifying colour usage was added with the must priority, since it helps avoid information going unnoticed.

The design of risks was not covered by the previous set of requirements. To accommodate the different levels of risks and decrease mistrust the previous design created, the requirement R.D.11 was added with the must-have priority. Alternatively, if the system accuracy is low, presenting them can be omitted.

The explanation text requirements were changed with the findings from L-WSYA A-IAR, GF and A-WSYA in mind. The requirement R.T.2.3 was added to help remove the social pressure a participant reported feeling due to comparisons. As it may also lessen the persuasiveness of the argument for some people, the requirement is a could-have. The requirement R.T.4 *The worker receives rewards, e.g. tags, badges or titles, for interactions with the system* was set as a won't have due to the large amount of negative responses. An alternative approach is to remove any text mentioning the badge from the explanations, but keep it in

the parent system of PERSEVERE and add more information to the hoverable text, for instance "You received this badge for accepting a loan recommendation. 75% of your co-workers have this." To give users more incentives to get a badge, text such as "Accept X more recommendations to get the Y badge" can be added to the hoverable text. A more visual option is including a progress bar.

The requirement R.T.6 was prioritised as a must-have requirement, because personal arguments were more effective according to L-WYSA and A-WYSA findings. The requirement R.T.7 *System should present clear business goals to be attained through the system* was refined into *System should present clear business goals connected to the current business process to be attained through the system* to resolve the issue of the chosen KPIs not being important for the current decision.

The requirement R.T.8 previously stated that the total number of prescriptions or analysed cases must be displayed. However, according to that requirement, if a subsection of the total cases (or resources) is analysed because of some constraint, then that amount is shown and the actual total is not displayed. To include this missing information, R.T.8.1 was added.

Table 7. Amended and new requirements

#	Requirement	Based on	Prioritisation
R.E. Explanations for Prescriptions			
R.E.8	Explanation mentions sources of prescription and explanation data (e.g. data warehouse, external credit bureaus, formulas)	L-WSYA, L-S, L-CUT	Should
R.E.9	Explanation includes proof of external system audits	L-CUT	Could
R.D. Design of the System			
R.D.9.1	Up and down arrows, minus and plus operators and indicative words (e.g. "decrease" / "increase") should be used to communicate possible changes	L-IAR	Should
R.D.10	Less impactful visualisations have non-greyscale pastel colour variations, more impactful ones stronger variations	L-S	Must
R.D.11	The system displays risks regarding a prescription in a size and/or colour according to the risk level	L-R, A-R, GF	Must
R.T. Text			
R.T.2.3	Names of other users are anonymised	L-WSYA	Could

R.T.4	The worker receives rewards, e.g. tags, badges or titles, for interactions with the system	L-WSYA, A-WSYA	Won't
R.T.6	The system self-monitors previous cases and customer satisfaction and gives feedback	L-WSYA, A-WSYA	Must
R.T.7	System should present clear business goals connected to the current business process to be attained through the system	L-IAR, A-IAR, GF	Must
R.T.8.1	The total number of resources that could be analysed is presented.	A-WYSA	Should

Table 7 gives an overview of all amended and new requirements formulated based on the findings. Four requirements were improved upon and six new were added.

Some improvements were not substantial enough to be a separate requirement. These are ideas that can improve the persuasiveness of single arguments in explanations. According to A-WSYA, the argument "14 employees have used this recommender system" can mention the amount of transferred cases, for example "14 employees have used this recommender system to transfer 102 cases". According to L-R, model accuracy lower than 100% can worry users. Technically achieving 100% model accuracy is near impossible, so instead the descriptive text of the risk can say the accuracy is "high". As covered in A-AI, the collapsible additional information block was not descriptive enough. Thus, its header could instead be "Additional information about Ron K." or the portfolio visualisation's title could be "Ron's portfolio profits" to make the subject clear. According to GF, the header can also be bolder. L-S included an issue regarding the design of no external overdues. Such a positive argument can be in a bolder font.

8.2.1 Excluded Improvements

We derived many improvements based on the findings. However, some of these possible improvements were excluded as they require repeated use of the system to take effect. Repeated use of the system was not accounted for when designing and evaluating PERSEVERE. In addition, when the user's effort expectancy is increased by a potential requirement, it cannot be included, as ease of use affects persuasion.

According to L-WSYA and L-AI, some people prefer numbers and tables over text. In extreme cases explanations could be in a tabulated form. As explained in section 5.2.2, tabulation is better for users likely to disagree with the recommendation. In the beginning of system use this is likely to be true as the user is sceptical of the recommendations, and therefore this could be an additional requirement with repeated use in mind. When users have enough experience and trust in the system, text or a hybrid approach are better options and hence R.E.1 would gain a higher priority.

Another point of concern in L-WSYA was the argumentation, which seemed scam-like to one participant. An optional resolution to this would be removing those arguments. However, in a real world scenario this issue is alleviated as the people referenced might know each other or can search each other up.

Under the findings of GF, a scrollbar was proposed instead of the collapsible design covered in R.D.3.4 *The worker can toggle accordions for explanations and visualisations.* This proposal wasn't applied since the criticism is influenced by the participants' experience with the system. Another reason is that scrolling can take many more movements with a mouse than one click collapsing or opening of an accordion, therefore ease of use could decrease.

Another finding from GF was the explanation bold text design issue. When the wrong words or too many words are stressed, the user cannot quickly glean information anymore and the design serves no purpose. The first issue can be resolved with iterative development, where the system designer can over time choose better words to make bold for optimal effect. The second issue is solved with repeated use, as the arguments will stay the same but the numbers will change and the user will naturally begin to associate the bold words with its argument, which in turn lessens their cognitive load.

The last findings from GF group were regarding the ordering of elements in the persuasive layout. However, the layout wasn't changed, since it would decrease the persuasiveness. This issue can also lessen when workers get used to the ordering over time. Similarly, under L-R, A-R and GF the design of the risk popover was a point of concern. With repeated use workers can learn to expect a popover and thus no changes were made.

In L-AI the tree visualisation received very polarising comments. For one participant reading it was too complicated, while seven others thought it was easy to read. Therefore, no altercations were made to it. Under A-AI findings some of the information presented was of little to no value to the participants. These comments were discarded as it's additional information, which by definition doesn't have to always be of use.

Findings about any missing information (L-MI, A-MI) were disregarded as the scenario used was a simplified one and not all case data could be given. For a designer of a real system, the necessary data is known and can be added accordingly. Hence, no requirements were created from that feedback.

In the A-WYSA group, one criticism was that there was no information about Ron. However, the explanation or Why Should You Accept this Recommendation? block is by design not meant for that information and rather is connected to the recommendation and engine. Ron's info was in the additional information section instead. The confusion about the explanation's subject can be cleared with sustained use of the system. Hence, no additional changes were made based on these findings.

9. Discussion

In consideration of the research question *How can prescriptive business process monitoring be visualised in a persuasive manner?* the findings indicate that utilising existing persuasive and visualisation strategies from different research areas in the context of prescriptive business process visualisation can allow workers to follow recommendations with more confidence and increase trust towards them. Persuasion is multi-faceted, situational and requires a deep understanding of the business process, user, domain and context it is embedded in.

Some business processes are such that process workers require specific data in order to make decisions. The findings suggest that this data can be displayed with persuasive strategies in mind to further strengthen a recommendation. In addition, process workers can be influenced with different types of supplementary persuasive messages, such as visualisations and explanations about the prescription. These results are consistent with persuasive research, such as Fogg's design principles [11], the PCD [32], PSD framework [29], [30] and Thomson et al. [56].

The findings provide indication that visualisations can be an effective shorthand for persuasion, especially so for beginners or first-time users. Well-done visualisations and comparisons can convey the essence of a persuasive argument in a simple way. This is supported by visualisation theories, such as the work by Berinato [48]. The findings also suggest that explaining a machine learning model via model-specific visualisations, such as a decision tree, can make a system more transparent and therefore, trustworthy. In accordance, Gedikli et al. [54] emphasise that transparency can increase satisfaction and Verbert et al. [52], in the same vein, confirmed that model-specific visualisations are useful.

The findings also suggest that aesthetics are an important factor for persuasive visual presentation. Aesthetics were an important factor not only in visualisations, but in the whole system in general. These results were in accordance with ELM [28] and the work of Fang et al. [24], in which they proposed a persuasive visualisation can have four dimensions, one of which is aesthetic. One other dimension they outlined is ambience. While a user interface in a business environment cannot have the most up-to-date visual flourishes, as it would not be suitable there, simplicity and beauty may make the system more trustworthy and clear for the user, as is evident by the findings. Moreover, there is evidence that colours play an important role in aesthetics and effective persuasion, as they can serve as an emotional cue. This finding further confirms the association between the work of Fang et al. [24], in which the third dimension of persuasion in visualisation is emotional.

The results showed a high degree of ease of use when working with PERSEVERE. This indicates that workers appreciate the clarity of the user interface, the minimal amount of input the system requires and the work it does for them. These results match those observed in earlier studies by Oinas-Kukkonen and Harjumaa, as well as Marcu et al. [29], [32].

The findings include evidence in support of displaying comprehensive information for a recommendation. However, workers may have concerns about trustworthiness of the persuasive messaging if it's not supported by credible sources. While some workers found the

more comprehensive form of messaging to be superfluous, displaying more information allows workers the freedom of investigating the data by themselves and therefore can increase the clarity of the recommendation. This finding is corroborated by the PSD framework and ELM, which state that comprehensive data, although not inherently better than less data, can cause a more permanent behaviour change [28], [29]. Guidelines set by Gedikli et al. [54] also found users would rather take their time to read to make good decisions. There is also evidence that workers appreciate the inherent objectivity of an automatic prescription. Providing credible sources can emphasise this objectivity. All this, in turn, can increase persuasiveness. These findings were mirrored by Torkamaan and Ziegler in their health recommender system framework [33].

There is evidence that many different explanation strategies can be employed with success. This finding is corroborated by the PSD framework and Fogg's work [11], [29]. However, our interviewees gave answers which show that each strategy's success is relative to their understanding and personality. For most, self-monitoring has an affirming effect, which is in agreement with Thomson et al. [56], who found this strategy to be effective with physical activity behaviour change apps. A technically high prescriptive model accuracy can decrease persuasiveness for some users who are not aware of the limitations of machine learning models. Therefore, the system should either be designed by someone who knows their target user base or be personalised to users. In accordance with these results, previous study by Marcu et al. [32] have demonstrated that a system's persuasion is highly dependent on the designer's knowledge. Personalisation and tailoring of persuasive content is one of the topics that requires further research.

The findings contain evidence that although different persuasive strategies can help the worker understand a prescriptive system better, it is not necessary for trust or confidence in decision-making. In other words, users can trust a recommendation guided by persuasive principles regardless of if they understand it or not. In theory, a good understanding of the prescriptive system and its prescriptions can perhaps create a longer-lasting, more in-depth persuasive effect [28], [29], [54]. Although many persuasive principles of this thesis focus on understandability and the black box problem, the findings suggest that understandability can be circumvented with other persuasive methods. Thus, successful persuasion can be achievable without focusing on understandability. However, XAI evidence is still an important supporting factor for a higher likelihood of persuasion.

The findings suggest that having persuasive explanations and comparisons ordered before mandatory information ensures that the users see them better. Although it may seem unfamiliar to workers first, it ensures that they will see the arguments in favour of the recommendation. Further research could be done regarding repeated use of the system to see how users perceive the ordering and the persuasive strategies. Another point of interest that could be further looked into is the limiting of the persuasive strategies used over time. Either switching out strategies or removing and adding them over time could change the persuasive effect.

Findings point out that the chosen metrics and KPIs used in simulation and goal-setting strategies play an important role in persuasion. This seems to be in accordance with research

by Consolvo et al. [57], who successfully utilised goal-setting theory in a persuasive physical activity application. However, when the metrics aren't directly connected to the decision and process, they retract from system usefulness. This finding agrees with guidelines set by Gedikli et al. [54], one of which states that domain-specific data should be used in explanations. This is especially important when the workers know the process well. The designer must have an in-depth understanding of the business process in order to maximise the persuasive power of metrics and KPIs.

Although some of the feedback given by interview participants was mixed, in general PERSEVERE was perceived as useful and satisfying. This suggests that even when parts of the persuasive messages aren't useful or have a detrimental effect, the overall persuasion can still be successful.

We could not confirm that purely textual and hybrid explanation styles are better than just numerical. Some participants did not like visualisations and reported a preference for tabulation, which may be connected to the work of Pandey et al. [13], who proposed that displaying tables may be more persuasive for users who are likely to disagree with a topic (the recommendation). These participants reported feeling not confident and not trusting of the system unless told to do so.

We also could not confirm the findings of Oh et al. [14], which propose that interactivity in visualisations can increase persuasiveness. The interviewees had a chance of interacting with a visualisation by searching it, but nevertheless, they did not use this feature. This could be explained by the minimal design of the search box or by the small size of the visualisation. Thus, further research could be done to corroborate this theory.

Although rewards can be applied successfully, an example being Thomson et al. [56], our study has been unable to demonstrate that. Our findings suggest that gamifying the business process too much can create confusion among workers. Instead of rewards, further research could be done to study the effect of praise, for example congratulating the user in case of target KPIs being met or telling them they've done a good job after following a recommendation.

Due to the limited scope of this thesis, sustained use of persuasive elements could not be included in PERSEVERE nor evaluated. Moreover, some persuasive principles were not included in the evaluation that would require repeated use of the system. Thus, the principles defined in section 5.3 need further scrutiny. Giving the user a chance to criticise the system may have a positive effect on persuasion, but was also not included in the evaluation. This needs further looking into in the future.

9.1 Limitations

In the name of answering the research question of this thesis, a design science approach was utilised. However, this has certain limitations [66]. In order to keep our perspective user-centric, we gathered input from industry practitioners. Eight practitioners with loan experience took part of the evaluation. Although the user study was theoretically founded and the participants selected based on a user persona, the varying experience of the participants and their subjective perspectives can affect the results. Therefore, to gain deeper insights, more

participants should be involved. The internal threat of bias and subjectivity of the interviewees, which can influence the interpretation of data gathered, was mitigated by analysing multiple sources (screen recordings, interview, post-questionnaire). In addition to interviewee subjectivity, the system was designed and evaluated only by the author of this thesis, which means any findings should be subject to scrutiny.

The setting of this thesis was a bank loan business process with no repeated use. We took this limitation though to have less interference from other factors such as the domain. It was presumed that the recommendations given in the evaluation are acceptable and that PERSEVERE is not more complex than a regular process without the prescriptive element. This setting was chosen due to existing resources, such as an event log and domain expert. Hence, the development and evaluation of PERSEVERE was purely done in the context of that setting. A different setting may require a different approach.

Evaluation was limited to one version of PERSEVERE. This study is the first step in a longer research process that needs to be continued in the future. The data used in it was such that the user was more inclined to accept the recommendation. This was due to the presumption that PERSEVERE is highly accurate or otherwise it would not have been included in the workflow. Whether the user would trust the system in a more adverse situation was not evaluated and this calls for more thorough testing.

10. **Conclusion**

The aim of this thesis was to research the persuasiveness of the visual presentation of prescriptive business processes. Although a prescriptive system can automatically give recommendations regarding activities to do or resources to use, the worker may not follow this recommendation. Thus, due to human error the business may lose profits. To make the prescriptions more acceptable, a set of 29 persuasive principles was formulated by combining persuasion and visualisation research from different areas. For evaluation, an interactive user study was carried out with a prototype called PERSEVERE and practitioners from finance.

To understand a process worker and how to make a persuasive prescriptive system for them, we created a user persona that describes a general middle level loan officer. Moreover, we outlined two scenarios – a loan application recommendation and co-worker to assign a case to recommendation scenario. 40 requirements were specified and prioritised based on the persuasive principles, user persona and scenarios using the MosCow method. All requirements except for those with the "won't have" priority were implemented in PERSEVERE, in which static screens are linked through clickable zones and triggerable overlays. PERSEVERE was evaluated in a round of interviews with eight bank workers with different amounts of loan experience. As a result of the user study, several possible improvements to PERSEVERE were discovered and formulated as additional requirements or updates to existing ones. The findings also suggested that the proposed persuasive principles can have a persuasive effect, i.e. it can create trust in the system and confidence in the decision-maker no matter if they understand the recommendation or not. Interviewees found PERSEVERE to be useful and easy to use. However, most of the persuasive principles must be applied in a very specific manner to the business process to be useful for the worker.

There are many opportunities for future research of PrBPM persuasiveness. First, repeated use of a prescriptive system can be studied to see how it affects the persuasiveness of the principles. Second, personalisation and tailoring of persuasive messages, among other omitted persuasive strategies, can further be researched to determine the influence these strategies can have. Third, a different setting than a loan business process should be utilised to evaluate the principles against. Finally, a user test could be done with both a pre- and post-questionnaire and a less competent recommendation to see the level of persuasion the principles can create.

References

[1] D. Frazzetto, T. D. Nielsen, T. B. Pedersen, and L. Šikšnys, 'Prescriptive analytics: a survey of emerging trends and technologies', *VLDB J.*, vol. 28, no. 4, pp. 575–595, Aug. 2019, doi: 10.1007/s00778-019-00539-y.

[2] M. Dumas, M. L. Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management*. Springer Publishing Company, Incorporated, 2013.

[3] W. Aalst, *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, vol. 136. 2011. doi: 10.1007/978-3-642-19345-3.

[4] C. Di Francescomarino, C. Ghidini, F. M. Maggi, and F. Milani, 'Predictive Process Monitoring Methods: Which One Suits Me Best?', in *Business Process Management*, Cham, 2018, pp. 462–479.

[5] K. Kubrak, F. Milani, A. U. Nolte, and M. Dumas, *Prescriptive Process Monitoring: Quo Vadis?* 2021. [Online]. Available: https://arxiv.org/pdf/2112.01769.pdf

[6] S. Weinzierl, S. Dunzer, S. Zilker, and M. Matzner, 'Prescriptive Business Process Monitoring for Recommending Next Best Actions', Sep. 2020.

[7] M. Dees, M. de Leoni, W. Aalst, and H. Reijers, *What if Process Predictions are not-followed by Good Recommendations?* 2019.

[8] C. Gröger, H. Schwarz, and B. Mitschang, 'Prescriptive Analytics for Recommendation-Based Business Process Optimization', in *Business Information Systems*, Cham, 2014, pp. 25–37.

[9] N. Mehdiyev and P. Fettke, 'Explainable Artificial Intelligence for Process Mining: A General Overview and Application of a Novel Local Explanation Approach for Predictive Process Monitoring', *ArXiv*, vol. abs/2009.02098, 2020.

[10] D. J. O'Keefe, 'Persuasion: Theory and Research', 1990.

[11] B. J. Fogg, 'Persuasive Technology: Using Computers to Change What We Think and Do', *Ubiquity*, vol. 2002, no. December, Dec. 2002, doi: 10.1145/764008.763957.

[12] H. Oinas-Kukkonen, 'A Foundation for the Study of Behavior Change Support Systems', *Pers. Ubiquitous Comput.*, vol. 17, no. 6, pp. 1223–1235, Aug. 2013, doi: 10.1007/s00779-012-0591-5.

[13] A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini, 'The Persuasive Power of Data Visualization', *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, pp. 2211–2220, 2014, doi: 10.1109/TVCG.2014.2346419.

[14] J. Oh, H. S. Lim, J. G. Copple, and E. K. Chadraba, 'Harnessing the persuasive potential of data: The combinatory effects of data visualization and interactive narratives on obesity perceptions and policy attitudes', *Telemat. Inform.*, vol. 35, no. 6, pp. 1755–1769, 2018, doi: https://doi.org/10.1016/j.tele.2018.05.004.

[15] G. L. Geerts, 'A design science research methodology and its application to accounting information systems research', *Int. J. Account. Inf. Syst.*, vol. 12, no. 2, pp. 142–151, 2011, doi: https://doi.org/10.1016/j.accinf.2011.02.004.

[16] A. Metzger, T. Kley, and A. Palm, 'Triggering Proactive Business Process Adaptations via Online Reinforcement Learning', Jun. 2020.

[17] S. A. Fahrenkrog-Petersen *et al.*, 'Fire now, fire later: alarm-based systems for prescriptive process monitoring', *Knowl. Inf. Syst.*, Dec. 2021, doi: 10.1007/s10115-021-01633-w.

[18] F. Ricci, L. Rokach, and B. Shapira, 'Recommender Systems Handbook', in *Recommender Systems Handbook*, vol. 1–35, 2010, pp. 1–35. doi: 10.1007/978-0-387-85820-3_1.

[19] C. Richthammer, J. Sänger, and G. Pernul, ‘Interactive Visualization of Recommender Systems Data’, in *Proceedings of the 4th Workshop on Security in Highly Connected IT Systems*, New York, NY, USA, 2017, pp. 19–24. doi: 10.1145/3099012.3099014.
[20] J. Heer, M. Bostock, and V. Ogievetsky, ‘A Tour through the Visualization Zoo’, *Commun ACM*, vol. 53, no. 6, pp. 59–67, Jun. 2010, doi: 10.1145/1743546.1743567.
[21] C.-H. Chen, W. Härdrdle, and A. Unwin, ‘Handbook of Data Visualization’, 2016.
[22] T. Munzner, ‘Information Visualization Rules of Thumb’, University of British Columbia, Mar. 2020. [Online]. Available: https://www.cs.ubc.ca/~tmm/courses/436V-20/slides/rulesofthumb.pdf
[23] B. Shneiderman, ‘The eyes have it: a task by data type taxonomy for information visualizations’, in *Proceedings 1996 IEEE Symposium on Visual Languages*, 1996, pp. 336–343. doi: 10.1109/VL.1996.545307.
[24] W.-C. Fang, J. Yung, and J. Hsu, ‘Design Concerns of Persuasive Feedback System’, Jan. 2010.
[25] T. Dalzell and E. Partridge, *The Routledge Dictionary of Modern American Slang and Unconventional English*. Routledge, 2009. [Online]. Available: https://books.google.ee/books?id=5F-YNZRv-VMC
[26] B. Craft and P. Cairns, ‘Beyond guidelines: what can we learn from the visual information seeking mantra?’, in *Ninth International Conference on Information Visualisation (IV’05)*, 2005, pp. 110–118. doi: 10.1109/IV.2005.28.
[27] S. Consolvo *et al.*, ‘Flowers or a Robot Army? Encouraging Awareness & Activity with Personal, Mobile Displays’, in *Proceedings of the 10th International Conference on Ubiquitous Computing*, New York, NY, USA, 2008, pp. 54–63. doi: 10.1145/1409635.1409644.
[28] F. Marquart and B. Naderer, ‘Communication and Persuasion: Central and Peripheral Routes to Attitude Change’, in *Schlüsselwerke der Medienwirkungsforschung*, M. Potthoff, Ed. Wiesbaden: Springer Fachmedien Wiesbaden, 2016, pp. 231–242. doi: 10.1007/978-3-658-09923-7_20.
[29] H. Oinas-Kukkonen and M. Harjumaa, ‘Persuasive Systems Design: Key Issues, Process Model, and System Features’, *Commun. Assoc. Inf. Syst.*, vol. 24, Mar. 2009, doi: 10.17705/1CAIS.02428.
[30] K. Torning and H. Oinas-Kukkonen, ‘Persuasive System Design: State of the Art and Future Directions’, New York, NY, USA, 2009. doi: 10.1145/1541948.1541989.
[31] M. Al-Ramahi, O. El-Gayar, and J. Liu, ‘Discovering Design Principles for Persuasive Systems: A Grounded Theory and Text Mining Approach’, in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 2016, pp. 3074–3083. doi: 10.1109/HICSS.2016.387.
[32] G. Marcu, J. E. Bardram, and S. Gabrielli, ‘A framework for overcoming challenges in designing persuasive monitoring and feedback systems for mental illness’, in *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, 2011, pp. 1–8. doi: 10.4108/icst.pervasivehealth.2011.246097.
[33] H. Torkamaan and J. Ziegler, ‘Integrating Behavior Change and Persuasive Design Theories into an Example Mobile Health Recommender System’, in *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, New York, NY, USA: Association for Computing Machinery, 2021, pp. 218–225. [Online]. Available: https://doi.org/10.1145/3460418.3479330

[34] R. Galanti, B. Coma-Puig, M. de Leoni, J. Carmona, and N. Navarin, ‘Explainable Predictive Process Monitoring’, in *2020 2nd International Conference on Process Mining (ICPM)*, 2020, pp. 1–8. doi: 10.1109/ICPM49681.2020.00012.
[35] D. Farrugia, C. Zerafa, T. Cini, B. Kuasney, and K. Livori, ‘A Real-Time Prescriptive Solution for Explainable Cyber-Fraud Detection Within the iGaming Industry’, *SN Comput. Sci.*, vol. 2, no. 3, p. 215, Apr. 2021, doi: 10.1007/s42979-021-00623-7.
[36] J.-R. Rehse, N. Mehdiyev, and P. Fettke, ‘Towards Explainable Process Predictions for Industry 4.0 in the DFKI-Smart-Lego-Factory’, *KI - Künstl. Intell.*, vol. 33, no. 2, pp. 181–187, Jun. 2019, doi: 10.1007/s13218-019-00586-1.
[37] M. Maguire and N. Bevan, ‘User Requirements Analysis’, in *Usability: Gaining a Competitive Edge*, J. Hammond, T. Gross, and J. Wesson, Eds. Boston, MA: Springer US, 2002, pp. 133–148. doi: 10.1007/978-0-387-35610-5_9.
[38] C. Plain, ‘Build an affinity for KJ method’, *Qual. Prog.*, vol. 40, p. 88, Mar. 2007.
[39] R. Kosara, C. G. Healey, V. Interrante, D. H. Laidlaw, and C. Ware, ‘User Studies: Why, How, and When?’, *IEEE Comput. Graph. Appl.*, vol. 23, pp. 20–25, 2003.
[40] C. Zanbaka, P. Goolkasian, and L. Hodges, ‘Can a Virtual Cat Persuade You? The Role of Gender and Realism in Speaker Persuasiveness’, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2006, pp. 1153–1162. doi: 10.1145/1124772.1124945.
[41] D. Schulman and T. Bickmore, ‘Persuading users through counseling dialogue with a conversational agent’, in *ACM International Conference Proceeding Series*, Jan. 2009, vol. 350, p. 25. doi: 10.1145/1541948.1541983.
[42] S. Stern, J. Mullennix, C. Dyson, and S. Wilson, ‘The Persuasiveness of Synthetic Speech versus Human Speech’, *Hum. Factors*, vol. 41, pp. 588–95, Jan. 2000, doi: 10.1518/001872099779656680.
[43] M. C. Harrell and M. A. Bradley, *Data Collection Methods: Semi-Structured Interviews and Focus Groups*. Santa Monica, CA: RAND Corporation, 2009.
[44] P. Runeson and M. Höst, ‘Guidelines for conducting and reporting case study research in software engineering’, *Empir. Softw. Eng.*, vol. 14, no. 2, p. 131, Dec. 2008, doi: 10.1007/s10664-008-9102-8.
[45] F. Davis and F. Davis, ‘Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology’, *MIS Q.*, vol. 13, p. 319, Sep. 1989, doi: 10.2307/249008.
[46] B. Fogg, ‘Persuasive Computers: Perspectives and Research Directions’, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, USA, 1998, pp. 225–232. doi: 10.1145/274644.274677.
[47] E. Anagnostopoulou *et al.*, ‘From mobility patterns to behavioural change: leveraging travel behaviour and personality profiles to nudge for sustainable transportation’, *J. Intell. Inf. Syst.*, vol. 54, no. 1, pp. 157–178, Feb. 2020, doi: 10.1007/s10844-018-0528-1.
[48] S. Berinato, *Good Charts: The HBR Guide to Making Smarter, More Persuasive Data Visualizations*. Harvard Business Review Press, 2016. [Online]. Available: https://books.google.ee/books?id=_kJBCgAAQBAJ
[49] E. Anagnostopoulou, B. Magoutas, E. Bothos, and G. Mentzas, ‘Persuasive Technologies for Sustainable Smart Cities: The Case of Urban Mobility’, in *Companion Proceedings of The 2019 World Wide Web Conference*, New York, NY, USA, 2019, pp. 73–82. doi: 10.1145/3308560.3317058.
[50] ‘Weber’s law’, *Encyclopedia Britannica*. Jan. 31, 2020. [Online]. Available: https://www.britannica.com/science/Webers-law

[51] F. Nunes, N. Verdezoto, G. Fitzpatrick, M. Kyng, E. Grönvall, and C. Storni, 'Self-Care Technologies in HCI: Trends, Tensions, and Opportunities', *ACM Trans Comput-Hum Interact*, vol. 22, no. 6, Dec. 2015, doi: 10.1145/2803173.
[52] K. Verbert, D. Parra, P. Brusilovsky, and E. Duval, 'Visualizing Recommendations to Support Exploration, Transparency and Controllability', in *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, New York, NY, USA, 2013, pp. 351–362. doi: 10.1145/2449396.2449442.
[53] C. H. Chih and D. S. Parker, 'The Persuasive Phase of Visualization', in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2008, pp. 884–892. doi: 10.1145/1401890.1401996.
[54] F. Gedikli, D. Jannach, and M. Ge, 'How should I explain? A comparison of different explanation types for recommender systems', *Int. J. Hum.-Comput. Stud.*, vol. 72, no. 4, pp. 367–382, Apr. 2014, doi: 10.1016/j.ijhcs.2013.12.007.
[55] T. Nyström and M. M. Mustaquim, 'Managing Framing Effects in Persuasive Design for Sustainability', in *Proceedings of the 19th International Academic Mindtrek Conference*, New York, NY, USA, 2015, pp. 122–129. doi: 10.1145/2818187.2818277.
[56] C. Thomson, J. Nash, and A. Maeder, 'Persuasive Design for Behaviour Change Apps: Issues for Designers', New York, NY, USA, 2016. doi: 10.1145/2987491.2987535.
[57] S. Consolvo, P. Klasnja, D. W. McDonald, and J. A. Landay, 'Goal-Setting Considerations for Persuasive Technologies That Encourage Physical Activity', New York, NY, USA, 2009. doi: 10.1145/1541948.1541960.
[58] J. Mumm and B. Mutlu, 'Designing motivational agents: The role of praise, social comparison, and embodiment in computer feedback', *Comput. Hum. Behav.*, vol. 27, no. 5, pp. 1643–1650, 2011, doi: https://doi.org/10.1016/j.chb.2011.02.002.
[59] I. Benbasat and W. Wang, 'Trust In and Adoption of Online Recommendation Agents', *J AIS*, vol. 6, Mar. 2005, doi: 10.17705/1jais.00065.
[60] O. Artemenko, V. Pasichnyk, N. Kunanec, and D. Tabachyshyn, 'Using context analysys for providing real time recommendations in e-tourism mobile location-based recommender systems', in *2019 IEEE 14th International Conference on Computer Sciences and Information Technologies (CSIT)*, 2019, vol. 3, pp. 166–169. doi: 10.1109/STC-CSIT.2019.8929822.
[61] M. de Leoni, M. Dees, and L. Reulink, 'Design and Evaluation of a Process-aware Recommender System based on Prescriptive Analytics', in *2020 2nd International Conference on Process Mining (ICPM)*, 2020, pp. 9–16. doi: 10.1109/ICPM49681.2020.00013.
[62] P. Berander and A. Andrews, 'Requirements Prioritization', in *Engineering and Managing Software Requirements*, A. Aurum and C. Wohlin, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 69–94. doi: 10.1007/3-540-28244-0_4.
[63] A. Hudaib, R. Masadeh, M. Qasem, and A. Alzaqebah, 'Requirements Prioritization Techniques Comparison', *Mod. Appl. Sci.*, vol. 12, Jan. 2018, doi: 10.5539/mas.v12n2p62.
[64] Indeed Editorial Team, 'Key Performance Indicators (KPIs): Definition and Examples', Oct. 2019. https://www.indeed.com/career-advice/career-development/key-performance-indicators
[65] A. Artasanchez, '9 Reasons why your machine learning project will fail', Jul. 2018. http://thedatascience.ninja/2018/07/12/why-your-machine-learning-project-will-fail/
[66] S. Gregor and A. Hevner, 'Positioning and Presenting Design Science Research for Maximum Impact', *MIS Q.*, vol. 37, pp. 337–356, Jun. 2013, doi: 10.25300/MISQ/2013/37.2.01.

Appendix

I. Persuasive Strategies

Table 8. PSD strategies used in this thesis

Principle	Description	Motivation
Self-monitoring	The system helps the user achieve their goals by tracking their performance and visualising it or giving feedback in textual form.	During a business process a worker does not always have the resources to monitor their performance. The system does that for the user and through it, informs and persuades the user to do better.
Simulation	The system persuades by simulating their decision and presenting the cause and effect relationship of it.	A simulation of a decision during a process persuades by presenting clear consequences of every action. A process worker might usually trust their instincts or knowledge, but a simulation might persuade them not to. Business process simulation (BPS) methods can be used.
Tailoring and personalization	The system presents persuasive content most suitable per user. It can be tailored to potential needs, interests, personality, usage context, and other factors.	People respond differently to different persuasion strategies and messages [18], [49]. In most cases, tailoring and personalization have big impacts on system acceptability, and that is why they should be applied to business processes as well.
Comparison, competition and cooperation	Socially comparing user performances can persuade them to perform a target behaviour so that they can be included in the social group. This can be done to a larger degree	The inner-business social dynamics can be leveraged by employing social comparison

	by motivating users to cooperate or compete.	and encouraging either competition or cooperation.
Praise and rewards	The user is praised via text, image, symbols, or sounds or digitally rewarded in some form after a behaviour is performed.	Praise and reward are universal motivators and may be persuasive in business processes as well.
Goal-setting	System should present clear, attainable goals to be achieved through the system that help towards performing a target behaviour.	As business processes are goal-oriented [2], leveraging those goals in persuasion is a simple yet effective strategy.
Trustworthiness, expertise, system surface credibility and verifiability	System should provide truthful, fair, accurate and unbiased information, and seem competent and knowledgeable. The system can be verified, i.e. the content it provides can be verified via outside sources.	A business process takes place in a professional environment, which is why trustworthiness, expertise, credibility and verifiability are the most suitable system credibility support.

II. Requirements

Table 9. Requirements

#	Requirement	Based on	Prioritisation
R.E. Explanations for Prescriptions			
R.E.1	Textual, numerical or hybrid explanations are shown for a prescription.	PE1	Must
R.E.2	Explanations contain visualisations.	PE2	Must
R.E.3	Explanations contain comparisons of visualisations.	PE3	Should
R.E.4	The system displays textual or hybrid explanations in favour of purely numerical ones.	PE10	Should
R.E.5	The feature weights of the prescriptive model are described.	PE11	Could
R.E.6	The system displays an explanation for no prescription.	PE12	Won't
R.E.7	The explanations fit the prescriptive model by showing evidence through a tree visualisation.	PE15	Must
R.E.8	Explanation mentions sources of prescription data (e.g. data warehouse, external credit bureaus).	PE25	Won't
R.D. Design of the System			
R.D.1	3D visualisations are used with spatial or topological data and 2D visualisations for all other types of data.	PE4	Must
R.D.2	The most important details of a visualisation should be emphasised.	PE4	Must
R.D.3	Complex visualisations are interactive for details-on-demand.	PE4	Must
R.D.3.1	The loan officer can filter visualisations.	PE4	Won't
R.D.3.2	The loan officer can search visualisations.	PE4	Should
R.D.3.3	The loan officer can zoom in and out of visualisations.	PE4	Won't

R.D.3.4	The worker can toggle accordions for explanations and visualisations.	PE4	Must
R.D.4	The system presents visualisations with colour codings and a legend.	PE4	Must
R.D.5	Contrast, proximity and repetitions of visual elements should be utilised for indicating identicality, similarity and patterns.	PE4, PE5	Must
R.D.6	Visualisations should be ambient, aesthetic, emotionally-engaging or metaphorical.	PE4, PE5	Must
R.D.7	The system's visual design elements follow the business' common user interface design.	PE6, PE13, PE14, PE15	Should
R.D.8	The system employs visualisations utilised in other applications in use in the business.	PE6	Won't
R.D.9	The system should provide comparisons of simulations of the possible after-effects of following and not following the decision.	PE3, PE17	Should
R.F. Flow of the System			
R.F.1	Within the system, structure of visualising content should be the same as the decision-making process of the worker.	PE6, PE7	Should
R.F.2	The system should be easily openable and dismissible.	PE7	Must
R.F.3	Within the system, the user is asked for a minimal amount of input, i.e. accept and reject.	PE8	Must
R.F.4	When opening the system, information should be automatically queried from the parent system.	PE8	Must
R.F.5	When having reached a decision, the system automatically conveys information to the parent system.	PE8	Must
R.F.6	The system displays a progress bar during prescription creation.	PE23	Must
R.F.7	The system prompts the user for critique after a decision has been made.	PE24	Won't
R.T. Text			

R.T.1	Thorough information and details can be viewed about a prescription upon demand.	PE9	Should
R.T.2	Other users' statistics of system use are shown.	PE16, PE21	Should
R.T.2.1	Other users' successful uses of the system are displayed with side impacts (e.g. promotion of co-workers, happy customers, increase of KPIs).	PE16, PE21, PE27	Should
R.T.2.2	Other users' negative uses of the system are displayed with side impacts (e.g. decrease of KPIs).	PE16, PE21, PE27	Could
R.T.3	The system's overall success rate is displayed.	PE16	Should
R.T.4	The worker receives rewards, e.g. tags, badges or titles, for interactions with the system.	PE16, PE22	Could
R.T.5	The user is praised after a recommendation is followed.	PE16, PE22	Should
R.T.6	The system self-monitors previous cases and customer satisfaction and gives feedback.	PE18	Should
R.T.7	System should present clear business goals to be attained through the system.	PE20	Must
R.T.8	The system displays the total number of possible prescriptions or analysed cases.	PE23	Should
R.T.9	The system displays risks regarding a prescription.	PE26	Must

III. Questionnaire

PERSEVERE Evaluation Post-Questionnaire

Finally, as the last step of the study we ask you to fill out this questionnaire. You will need to answer questions regarding the usefulness and ease of use of the PERSEVERE system. This helps us evaluate the system on a quantitative basis.
Fill the questionnaire as yourself.

* Required

Information about you

1. Age *

2. Gender *

Mark only one oval.

- ◯ Male
- ◯ Female
- ◯ Non-binary

3. Job tenure *

Usefulness of PERSEVERE

4. Please indicate your level of agreement with the following statements related to the USEFULNESS of PERSEVERE. *

Mark only one oval per row.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
Using PERSEVERE in my job would enable me to accomplish tasks more quickly	◯	◯	◯	◯	◯
Using PERSEVERE would improve my job performance	◯	◯	◯	◯	◯
Using PERSEVERE in my job would increase my productivity	◯	◯	◯	◯	◯
Using PERSEVERE would enchance my effectiveness of the job	◯	◯	◯	◯	◯
Using PERSEVERE would make it easier to do my job	◯	◯	◯	◯	◯
I would find PERSEVERE useful in my job	◯	◯	◯	◯	◯

Ease of use of PERSEVERE

5. Please indicate your level of agreement with the following statements related to the EASE OF USE of PERSEVERE. *

Mark only one oval per row.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
Learning to operate PERSEVERE would be easy for me	◯	◯	◯	◯	◯
I would find it easy to get PERSEVERE to do what I want it to do	◯	◯	◯	◯	◯
My interaction with PERSEVERE would be clear and understandable	◯	◯	◯	◯	◯
I would find PERSEVERE to be flexible to interact with	◯	◯	◯	◯	◯
I would find PERSEVERE easy to use	◯	◯	◯	◯	◯
It would be easy for me to become skillful at using PERSEVERE	◯	◯	◯	◯	◯

Satisfaction of PERSEVERE

6. How do you feel about your OVERALL EXPERIENCE of system use *

Mark only one oval per row.

	1	2	3	4	5
(1) Very dissatisfied to (5) Very satisfied	⬭	⬭	⬭	⬭	⬭
(1) Very displeased to (5) Very pleased	⬭	⬭	⬭	⬭	⬭
(1) Very frustrated to (5) Very contented	⬭	⬭	⬭	⬭	⬭
(1) Absolutely terrible to (5) Very delighted	⬭	⬭	⬭	⬭	⬭

IV. License

Non-exclusive licence to reproduce thesis and make thesis public

I, Janna-Liina Leemets,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,
   Persuasive Visual Presentation of Prescriptive Business Processes,
   supervised by Fredrik Payman Milani and Alexander Udo Nolte.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Janna-Liina Leemets
08/08/2022