

Promoting Timely Completion of Multi-Step Processes - A Visual Approach to Retrospective Analysis

Kostas Pantazos^{#1}
Software Development Group
IT University of Copenhagen
Copenhagen, Denmark

Sureyya Tarkan^{‡1,2}
Catherine Plaisant^{§1}
Ben Shneiderman^{¶1,2}
Human-Computer Interaction Lab ¹
& Dept. of Computer Science ²
University of Maryland
College Park, MD, USA

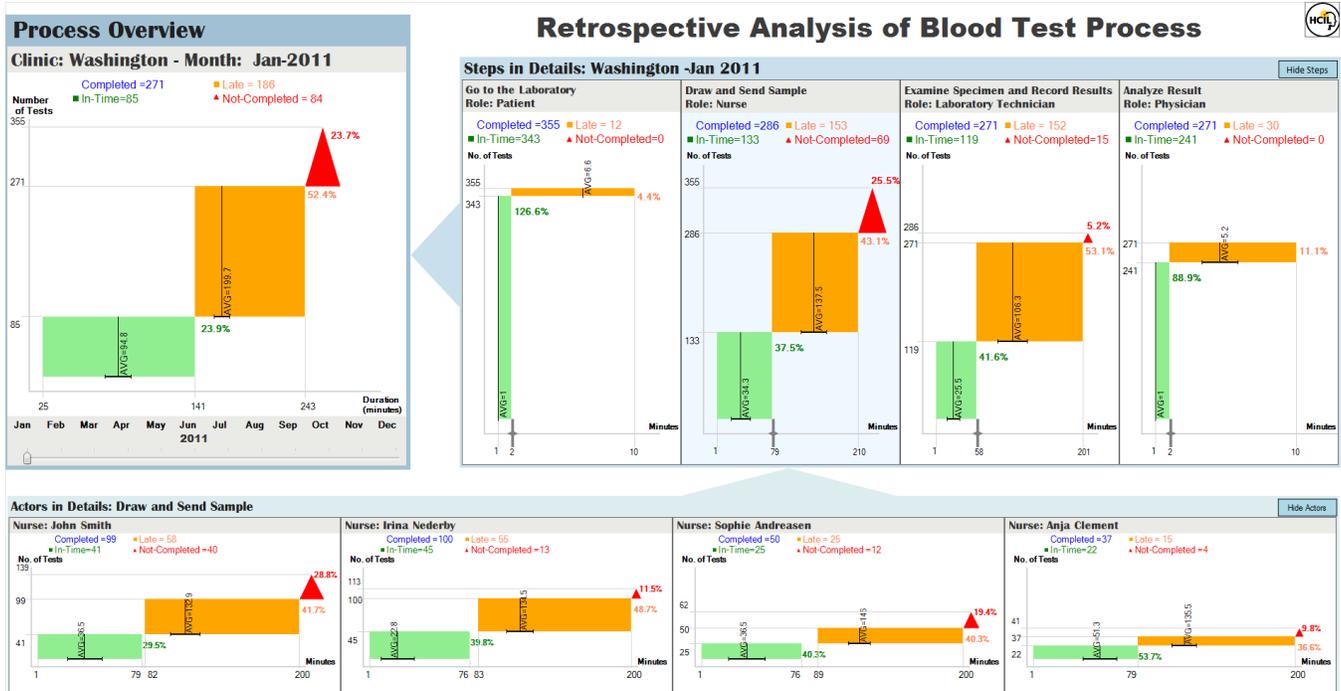


Figure 1: The MSProVis combines 3 views in a single presentation, allowing managers to explore process, steps, and actors' performances at a glance. Each view uses one or more PCDs to visualize in-time, late and not-completed completions. This interactive visualization allows users to view historical data using the slider in the Process Overview, and adjust threshold of lateness in the Steps in Details view. Changes in one view reflect immediately in the other two views. Users can view details-on-demand inside tooltips by clicking on shapes.

ABSTRACT

Multi-step processes are common in healthcare, software development, and other dynamic environments. Managers who monitor processes are overwhelmed by information, and often limited in the time to investigate data from different viewpoints. We propose a visual approach to facilitate retrospective analysis, to promote

timely completion of multi-step processes, and improve performance. It uses a novel visualization called the Process Completion Diagram (PCD) that aggregates event-logs into in-time, late and not-completed completions, and visualizes those using shapes and colors. The interactive Multi-Step Process Visualization (MSProVis) combines a number of PCDs and allows users to review and compare series of PCDs at three levels of detail, allowing comparisons between steps or between actors executing those steps. MSProVis computes default thresholds that define in-time and late completions. These might trigger alarms, and allow managers to adjust those thresholds interactively.

A pilot usability test with five participants showed that all participants could learn how to read almost all the information presented in the PCD without training in less than five minutes, and use MSProVis effectively. They carried out nine tasks successfully, which lasted on average seven minutes. Suggestions for improvement include color usage, and label and shape positioning.

* e-mail: kopa@itu.dk

[†] This work was done while visiting the University of Maryland, College Park, Maryland, USA

[‡] e-mail: sureyya@cs.umd.edu

[§] e-mail: pleasant@cs.umd.edu

[¶] e-mail: ben@cs.umd.edu

Index Terms: Visual Analytics, Process Completion Diagram, Multi-Step Process Visualization.

1 INTRODUCTION

Advancing technology has accelerated the growth of stored data, but extracting meaningful insights remains a substantial challenge. The challenge is amplified in dynamic environments, where multi-step processes are present. Examples include the lab tests in healthcare, the paper-review process in academic journals, the return merchandise authorization process in the electronics industry, the package shipping process in the logistics industry, the bug fix process in software maintenance, etc. A multi-step process is composed of a number of steps, where each step is executed by one or more actors from an organization. The performance of an actor can be *in-time* (when completion's duration is lower or equal to an expected duration time), *late* (when completion's duration is higher than the expected duration time), and *not-completed* (when completion's duration is undefined). Retrospective analysis can assist in improving these processes, locating the critical steps, understanding the root-cause and supporting managers in making better decisions. However, investigating completions (the raw data of event-logs) is a time-consuming and tedious process.

In multi-step processes, managers who monitor the process are overwhelmed by information, and often limited in the time to investigate data from different viewpoints. According to the "Bounded Awareness" theory, the limited cognition of humans and large volume of available information can harm the decision-making process [5]. Information visualization enhances human cognition by visually presenting abstract data [4] and revealing patterns, trends and outliers [20]. Visual analytics is the science of combining interactive visualizations with analytical reasoning techniques to enable users to understand their data, reflect more effectively, and make better decisions [10]. The bounded awareness challenge in multi-step processes can be addressed by introducing visual analytics. Exploring the data using a visual analytics approach can support managers in obtaining an overview, identifying the bottlenecks, and highlighting the root causes and contributing factors. In addition, it can deal with fundamental questions such as: what is the average time for the whole process and each step; how many steps were performed in time and late; are there consistent patterns for not-completed completions; which actor is causing the delay; how are processes and steps performed over different points in time etc. Therefore, visual analytics has the potential to promote timely completion of multi-step processes, and assist managers in making better decisions.

The contributions of this paper are:

- A novel information visualization technique that categorizes and summarizes event-logs using duration time, called the **Process Completion Diagram (PCD)**,
- The interactive **Multi-Step Process Visualization (MSProVis)** that combines several PCDs, which aims to promote timely completion of multi-step processes.

. This paper also reports the results from a pilot usability test with 5 participants. Although this approach was evaluated using a medical domain example, it can be applied for multi-step processes in other domains as well.

The remainder of this paper is organized as follows. First, the motivation section describes several examples from the medical domain and the software development. Then, the paper provides a summary of the related work, describes the PCD and MSProVis, and presents results of our evaluation. At the end, this paper presents the conclusions and future works.

2 MOTIVATION

The study of Hickner et al. [8] reports that patients' lives were threatened when missing test results caused delay in care (24%), time and financial costs (22%), patients' pain (11%), unpleasant clinical consequence (2%). In 18% of these cases patients were harmed, while in 28% harm status was not defined. Previous research [30], [28], [21], [24], [29] in the healthcare environment has shown that there is room for improvement in the process of ordering tests. A test, such as blood chemistry, can be performed in-time, late or not-completed. The process starts when a physician orders a test, and finishes when the result is reported to the physician who acts upon the result. Once the order is placed the patient has to come to the appointment, where another medical professional draws the blood and sends it to the laboratory. The lab technician analyzes the blood then sends back the results to the ordering physician who take actions (i.e. orders more tests or a medication/procedure). Each step in this process is triggered when the previous is completed. A delay in one step may cause a delay in the whole process. The performance of an actor affects the duration of the step and the entire process.

Steindel and Howanitz [23] investigated the test turn-around-time in the emergency department (ED). Their results showed that delayed lab test results lengthened the ED treatment time and the ED stay length. Although often laboratories may cause the delay, there are cases [6] when they are not the root cause of the delay. Therefore, there is a need for identifying the bottlenecks in different situations [23] and support decision-making, which can improve healthcare delivery [28].

Software development is another dynamic environment, where project's success is a challenge. On 2009, the Standish Group published "Chaos Summary 2009"[25]. This report showed that only 32% of projects were delivered in time, on budget and supporting the required functionalities, 44% were late, over budget and supporting less functionalities, and 24% were never completed. Therefore, companies use measurement systems to manage and monitor their projects/products. Measurement systems can be developed according to the International Standard ISO/IEC 15939: 2007, in a standardized way [9]. The software measurement process involves data collection, storage, and analysis and results interpretation. Information needs (e.g. estimate productivity of future project using historical data) is the outcome of the measurement process [9]. This information is used by managers to estimate the process and ensure the product is developed in time, on budget, and good quality is guaranteed. They use software metrics indicators (variables derived from basic and derived measures in the measurement system to support data analysis and decision-making), but still project managers are overwhelmed by the generated information [17]. For example, a department manager wants to analyze retrospective data of a project. The project followed the waterfall development model that consisted of several steps (e.g. requirements analysis, design, coding, etc.). Each step involved several actors (e.g. analysts, developers, testers, etc.). One actor was responsible for executing several tasks. Some of those were performed in-time and late, and others never completed. An exploratory case study was conducted in a development unit of Ericsson, on how to improve the presentation of software metrics indicators in industrial measurement systems [17]. Their results showed that managers need space-efficient and interactive overviews to monitor the projects, and details-on-demand to identify bottlenecks, which in response can support decision-making.

The complexity of analyzing a multi-step process and deriving insights from retrospective data produces a need for a simple and intuitive presentation integrated in a visual analytics approach. The purpose of a visual analytics approach is to allow managers to gain an overview of the process, interact with the raw data, identify the critical areas, highlight the root causes, and make better decisions.

Finally, we present some example questions that are applicable in both cases presented above: what is the overall process performance; which step is performed in-time, late or not-completed; how does the overall process performance change when the value of the threshold of lateness increases or decreases; who is performing in-time or late; how does an actor's performance change if the threshold of lateness increases or decreases; what is the average execution time for the process, step and actor respectively; how does process performance differ over time?

3 RELATED WORK

In order to gain an overview of the previous work related with multi-step process visualizations, we reviewed existing literature on: Process Management, Process Mining, Information Visualization and Visual Analytics.

3.1 Process Management

A multi-step process can be modeled using one of the several business process management notation, for example using Petri Nets[6], Business Process Modeling Notation [33], Yet Another Workflow Language [27], etc. These models assist in defining the workflow of the process in an ideal environment, which may fall far from reality [26].

Process management models do not provide information about what is actually happening or happened in practice, because these models are not bound to data [15]. Therefore, they do not provide a summary of event-log data or ensure timely completion of multi-step processes. Poorly designed models can affect the process and then lead to incorrect decisions [26].

3.2 Process Mining

Considering the limitations of process modeling, a new research discipline, named process mining, has emerged. The foundations of this discipline are machine learning, data mining, process modeling and analysis [26]. Process mining techniques are applied to integrate business processes and event-logs aiming at discovering, monitoring and improving real processes [26].

Mans et al. [15] conducted a case study where they applied process mining techniques to analyze the healthcare process. In that study, the authors used a dot-chart, which can be considered as a visual analytics technique, to gain a better understanding of the event-logs in the process. Users were able to obtain new insights and identify patterns in the data. Their study showed that process mining techniques assist in improving the healthcare process, and visual analytics techniques enhance users' capability in exploring the data.

3.3 Information Visualization and Visual Analytics

Several previous studies [2], [7], [22], [19],[14] have shown the effectiveness of visual approach in analyzing data from different domains. A visual representation assists users in making better decisions for different types of tasks [22],[19]. Savikhin et al. [19] applied a visual analytics approach for economic decision-making. Their evaluation showed that a visual analytics approach was more successful in decision-making compared to a simple visual and tabular representation. Furthermore, their results suggest that a visual analytics approach can address the bounded awareness challenge. Therefore, this project aims at providing a visual analytics approach for multi-step processes, which allows users to review and compare performances at process, step and actor levels.

Several interactive visualizations have been developed to present event-log data. LifeLines [18] is an interactive visualization that presents an overview of a patient's medical record. This visualization was the inspiration for LifeLines2 [31], but it visualizes a collection of medical records, where users can explore the event-logs for temporal patterns. LifeLines2 is limited in showing an

overview of event sequences. This limitation was addressed in LifeFlow [34]. LifeFlow provides an overview of event sequences by summarizing all possible sequences and visualizing the temporal space of events within sequences [34]. ProcessLine visualizes time-series data of a process combining several visualization techniques in one [13]. Their evaluation using time-series data from a beer industry showed the effectiveness of the ProcessLine. McNames et al.[16] developed a graphical display of semiconductor manufacturing processes to monitor tools performance in the process. This approach uses color-encoding to represent tools and transition usage. However, all the aforementioned visualizations share the same limitation: *they do not categorize and provide a summary of event-log data*. This means aggregating event-logs durations into in-time, late, and not-completed completions, and then providing an overview to users who can identify new insights regarding the process, step, and actors' performances. A recent study [11] proposes visualization requirements for business processes, where users can gain a better understanding of the changes in a process. These requirements can be used as guidelines for investigating further how to visualize changes in business processes. Although considerable work [18], [31], [34] has been done in visualizing event-log data, when it comes to visualizing multi-step processes, there is room for improvement.

To the best of our knowledge, there is no previous study which addresses the issue of summarizing event-log data, and promoting timely completion of multi-step processes.

4 DESIGN

This section presents the PCD and MSProVis that combines several PCDs. MSProVis shows different levels of detail so as to support decision-making by managers to continuously improve overall organization and individual actors' performances.

4.1 Process Completion Diagram (PCD)

In a process, completions (e.g. lab-tests, tasks in software development projects) can be executed in-time, late or not-completed. Figure 2 shows how event-logs are aggregated, and presented in the PCD (Figure 2.c). The PCD visualizes completions using colors and shapes. The PCD is composed of two rectangles, one for the in-time and one for the late completions. Furthermore, it has a triangle for not-completed completions. These three shapes are placed in a time series plot, where X-axis shows the completion duration and Y-axis shows the number of completions. The PCD design was guided by the desire to support managers to make rapid recognition of tall-narrow green rectangles that indicate good performance or large orange rectangles and big red triangles that point out the problems. More explicitly, the PCD shows the number of in-time (green rectangle), late (orange rectangle), and not-completed (red triangle) completions (Y-axis). Figure 2 shows how 454 blood tests are summarized and presented using the PCD. The PCD shows that there are 224 in-time, 185 late, and 45 not-completed blood tests.

The PCD also shows the minimum, average, and maximum duration for the in-time and late groups of completions (X-axis). The in-time completion durations range from 1 to 50 minutes and the late completion durations range from 50 to 200 minutes. The PCD also presents a large amount of detailed information such as the threshold value, the standard deviation around each average (a small horizontal mark at the bottom of each rectangle), and the percentages for in-time, late, and not-completed tests.

At the top of the PCD there is a textual list of the absolute numbers of total completed, which is made up of the in-time and late completions. The textual list also presents the number of not-completed completions. Small colored shapes are placed close to each label, to link the text with the shapes in the plot.

To encode effectively the information regarding the in-time, late and not-completed completions, the PCD uses colors and shapes

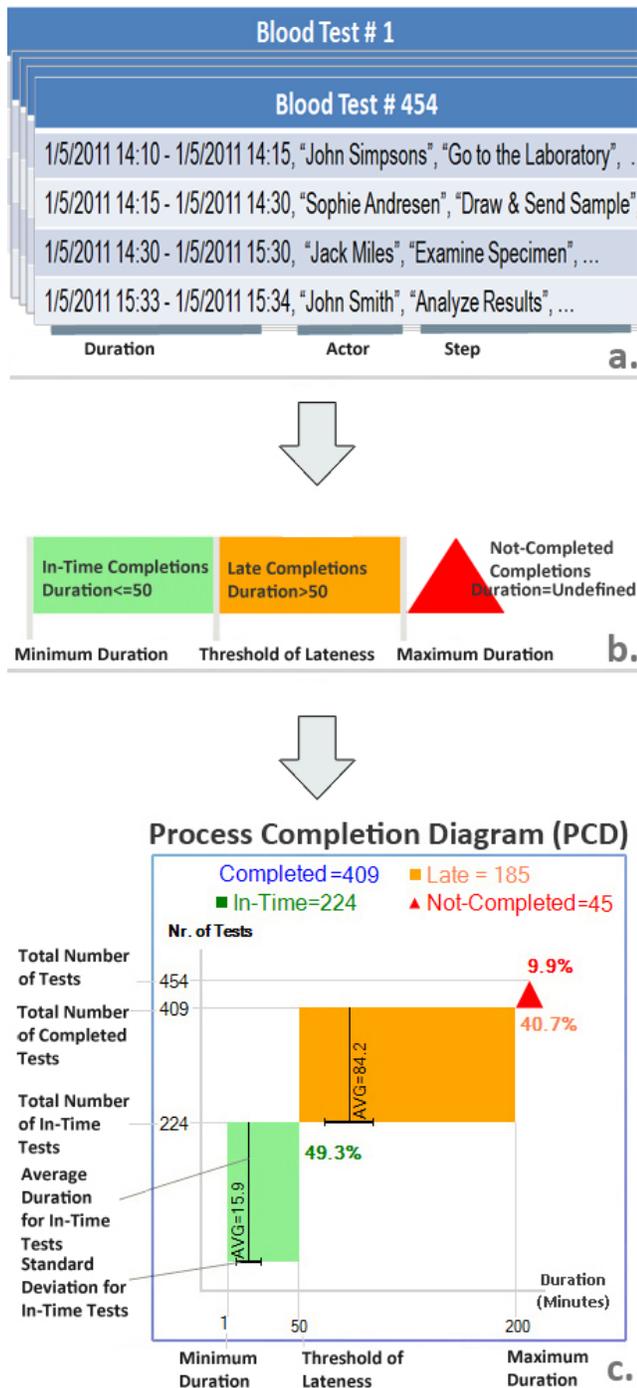


Figure 2: Creating the Process Completion Diagram (PCD): a. It presents an example of an event-logs dataset for a blood test process. The PCD uses the *Duration*, *Actor* and *Step* fields in the table. b. The event-logs are summarized in three groups using the duration and the threshold of lateness. c. Data are visualized using shapes and colors in a time series plot, where X-axis presents the duration and Y-Axis the number of tests. The PCD shows the average completion time, the standard deviation, minimum and maximum duration time, and the number for the in-time, late and not-completed completions.

(information encoding techniques suggested by Bertin [3]). Ware [32] suggested a 12-color palette for presenting data, and we decided to use green, orange and red. These colors are consistently used in the shape and the label coloring. For example, the late rectangle and labels related to it are colored in orange. Initially, we considered using yellow, rather than the orange, but on a white background the yellow color is not very readable. The red color usually denotes critical situations, such as the red sign in traffic lights, and in our case red highlights the critical aspect of the not-completed completions.

To distinguish further the completed from not-completed completions, the PCD uses a different shape, the triangle. In this design, the size of the shape denotes the process, step or actors' performances. Ideally, managers would see only tall and narrow green rectangles. This means that the duration time is low and the number of completions is high. In the PCD, the size of the shapes is calculated by aggregating all the event-log data from the database, and grouping them in: in-time, late and not-completed completions. The height of the shapes is bound to the number of event-logs in the database, and the width of the rectangles represents the duration. As the not-completed completions do not have a duration time, the triangle's width is not bound to the X-axis, and can expand up to predefined pixels number (based on available space). When the number of the not-completed completions is smaller than defined pixels number, the triangle's width refers to the triangle's height. This restriction is applied to obtain a space-efficient representation, and because duration is not reflected in the triangles width.

Classifying completions into in-time and late is realized using the threshold of lateness. As shown in Figure 2, the threshold of lateness corresponds to the maximum duration for the in-time, and minimum duration for the late completions. However, this might not be always the case. The maximum duration of the in-time can be lower, or the minimum duration of the late completions can be higher. As a result, there will be a gap between the green and the orange rectangle. The presence of the gap provides more detailed information, especially when managers compare individual actors' performances. An example is presented in section 4.2.3 and shown in Figure 5. Furthermore, the PCD positions the shapes using the completion duration and the number of completions. The orange rectangle and red triangle are placed at the top-right corner of the green rectangle (unless there is a gap) and the orange rectangle respectively. In this way, the PCD presents hierarchically the grouped completions, visualizes the total number of completions using the shapes' heights, and is a space-efficient presentation.

The PCD provides managers with a visual representation of the average duration and the standard deviation for the in-time and late completions. It uses a vertical line inside the rectangle to show the average duration, and a horizontal line, placed at the bottom of each rectangle, for the standard deviation. The numeric value of the average is aligned vertically and positioned close to the average line. Showing the average and the standard deviation indicates to managers, that a change in a threshold may have a high or low impact on the in-time or late completions. For example, let's assume the threshold of lateness is set to 79 minutes, the average value for lateness is 137, and the standard deviation is 20. Adjusting the threshold from 79 to 90 will not affect dramatically the size of green rectangle (i.e. the number of completions considered late in the process). While moving the threshold bar close to the average value (i.e. 136), will significantly change the rectangles' sizes (see Figure 6 and 7) In the PCD, the colored percentage labels for each performance type are positioned on the right side of the shape, aiming at confirming assumptions derived from shapes' size.

The final version of the PCD is a result of an iterative design, where different representation aspects were considered. For instance, initially the axes were inverted, showing the number of completions in X-axis, and duration in Y-Axis. Taking into con-

sideration that time is usually presented horizontally (see examples in [1]), directed us into changing the design.

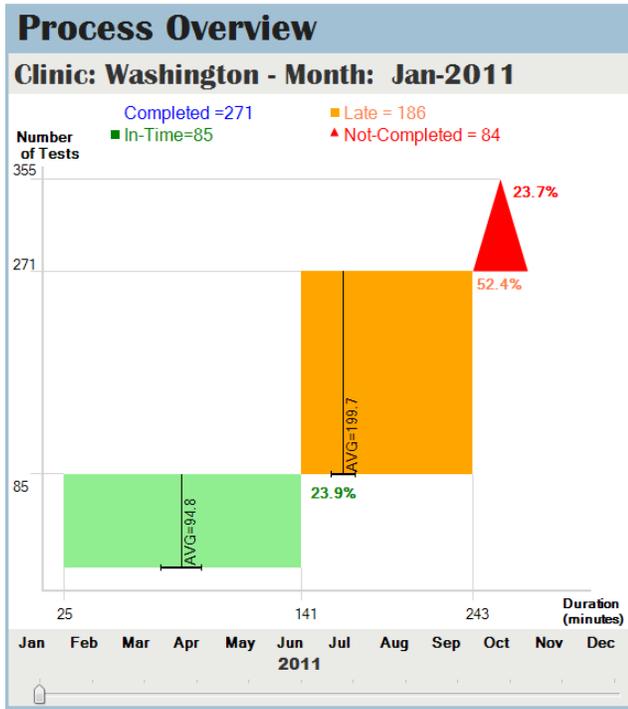


Figure 3: Process Overview presents the overall process performance. Users can use the slider to view data for another month.

4.2 Multi-Step Process Visualization (MSProVis)

The MSProVis is an interactive visualization composed of three main views: *Process Overview*, *Steps in Details*, and *Actors in Details*. MSProVis allows managers to review and compare series of PCDs at different levels of detail, allowing comparisons between steps or between actors executing those steps. MSProVis computes default thresholds that define lateness, and allows managers to adjust those thresholds interactively.

4.2.1 Process Overview

The Process Overview uses only one PCD that summarizes the overall process performance. The process duration is computed using all the process steps' durations. The threshold of lateness for the process is calculated using the threshold of lateness defined for each step in the process (described in section 4.2.2).

As an example, Figure 3 shows the overview of the blood test process using data from January 2011. The manager can click on rectangles or the triangle, which opens a tooltip that shows details-on-demand such as number of tests, minimum value, maximum value, average and standard deviation. Furthermore, the manager can interact with the slider at the bottom of the area, and view data of another month. As a result, the manager can gain insights on a specific month, but also compare the overall process performance for different months. Next, the manager double-clicks on the Process Overview and the Steps in Details view (Figure 4) is shown.

4.2.2 Steps in Details

A process can have one or more steps and the Steps in Details uses PCDs to show the performance of each step. In our example, the blood test process consists of four steps: "Go to Laboratory", "Draw and Send Sample", "Examine Specimen and Report

Results", and "Analyze Report". Therefore, the Steps in Details is composed of four PCDs (Figure 4).

The in-time and late completions do not affect the performance of the proceeding steps, although clearly impacts the performance of the overall process. Not-completed completions of a step has an impact on the next one, as the total number of performed tests in the next step is decreased. For example, 355 were performed in step two, while in step three only 286 tests, as 69 tests were not-completed from step two. Positioning PCDs in order and using the same normalized scale for the number of completions, allows managers to view quickly how the number of completions changes through steps, compare their performances, and identify bottlenecks and good performances.

Different steps may have different lengths of execution time (10 minutes for step one, 210 minutes for step two). This means that timescales used for the X-axis may differ from one step to another. Time-oriented visualizations face the challenge of presenting data in different timescales and using different granularity of time [1],[10]. PCDs use consistently the same granularity of time (i.e. minutes), but as space is limited and time span may differ significantly, we decided to evenly divide the space in the Steps in Details view. As a result, we use different timescales for the X-axis.

The manager can interact with the thresholds of lateness in the PCDs of the Steps in Details. This means that the manager can reconsider what should be treated as in-time and late completions for a step. When the visualization is opened for the first time, the threshold of lateness in each step has a default value. The default values usually correspond to estimations made from process management, but as we did not had this information we decided to set values to the average of step's duration instead of using random values. In our example (Figure 4), in the second step (*Draw and Send Sample*) the default threshold is 79. As a result, the PCD shows 133 in-time (where duration time is lower or equal than 94 minutes), 153 late (where duration time is higher than 79) and 69 not-completed completions. The manager can adjust the thresholds interactively by moving the threshold line, shown as a thicker gray line with two small arrows over the X-axis. When the threshold value changes, the Process Overview and Actors in Details are updated. This feature allows the manager to investigate several scenarios and extract insights from the data. The thresholds of lateness can also be used to adjust the triggering of alarms. Next, the manager double-clicks inside the second step (*Draw and Send Sample*) in Steps in Details, and the Actors in Details is shown (Figure 5).

4.2.3 Actors in Details

A step can be executed by more than one actor. The Actors in Details presents each actor's performance involved in a step, using a PCD per actor. It applies a descending ordering of the PCDs using the total number of completions, and can contain an unlimited number of PCDs. However, as space is limited, users have to use the scroll bar when there are more than 7 PCDs. In our example, there are four PCDs, one for each nurse involved in step two (Figure 5).

The same principles apply for the minimum and maximum value of the late completions. The minimum can be equal or higher than the step's threshold of lateness, and the maximum correspond to the highest duration performed by the actor (210 minutes). A gap between the green and orange rectangle is possible, whenever the actors maximum in-time performance and minimum late performance does not match the step threshold of lateness. For example, the actor *Sophie Andreasen* has executed in-time completions in 76 minutes, and her minimum time for late completions correspond to 89 minutes. While, the actor *Anja Clement* drew and sent samples equal to the threshold of lateness. As a result, managers can rapidly identify who has caused the lengthy step.

The PCDs use the same scale for X-axis (duration) and Y-axis

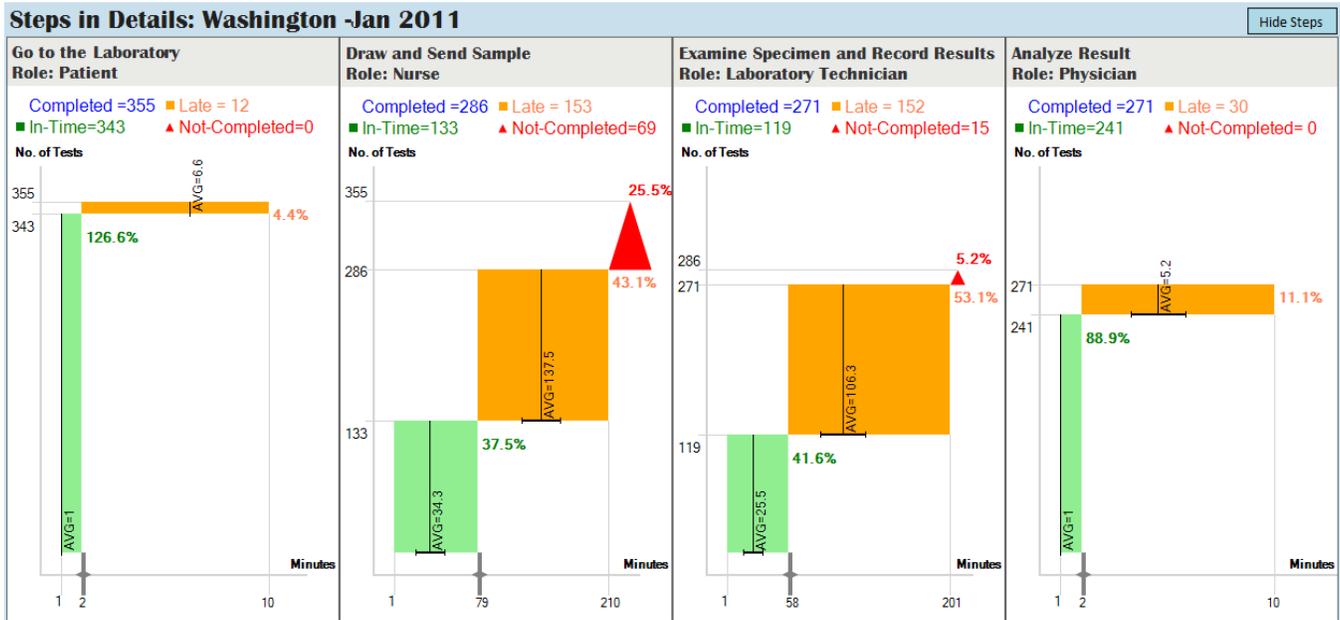


Figure 4: Steps in Details contains four PCDs. Each PCD correspond to a step in the process. Users can interact with the threshold of lateness (a thicker gray line with two small arrows over the X-axis), and investigate different scenarios. In this case, step two has the highest number of not-completed and step one has the best performance. Managers can refer to the percentage values of each shape to confirm their visual understanding.

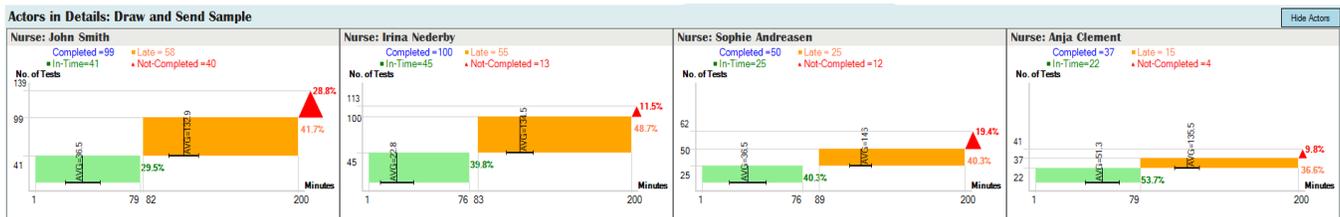


Figure 5: Actors in Details shows the actors' performances involved in step two. In this case there are 5 actors represented by 5 PCDs. This figure shows that *John Smith* and *Anja Clement* has the highest and lowest number of not-completed completions respectively. *Irina Nederby* has the best performance for the in-time completions. The gap between the green and orange rectangle shows that *Irina Nederby* has executed in-time and late completions in less and more time respectively than the defined threshold of lateness in the step in details.

(number of completions). More precisely, the timescale in the X-axis correspond to the timescale of the step. While, the Y-axis maximum value corresponds to the maximum number of completions performed by one of the actors. Therefore, comparing durations between actors is easier than between steps, where different timescales were used. Managers can access details-on-demand (i.e. average duration, standard deviation, etc.) by clicking on each shape.

The number of in-time, late and not-completed completions for each actor is computed by selecting those completions performed by the actor in the selected step. For example, the actor, *John Smith*, in Figure 5 has 41 in-time, 58 late and 40 not-completed completions. From this dataset, the PCDs show the minimum value and the maximum value based on the threshold value of the selected step. The minimum value (1 minute) is equal to the minimum duration performed by this actor, which means that there is a test executed in 1 minute. The maximum duration (79 minutes) for the in-time completions can be smaller than or equal to the threshold of lateness (94 minutes) of the step.

4.2.4 Multi-View Visualization

Figure 6 presents the MSProVis that combines three views (the Process Overview, the Steps In Details, and the Actors in Details) in a single presentation (Figure 1).

The design of MSProVis was driven by the need to efficiently present an overview of the multi-step process, and details-on-demand for steps and actors. We attempt to provide a space-efficient presentation by arranging the views as shown in Figure (Figure 1). The height, header color and position of each view represent the task analysis hierarchy: obtain an overview of the multi-step process, investigate steps, and analyze actors' performances. This multi-view visualization allows managers to review and compare series of PCD at different levels of detail, allowing comparisons between steps or between actors executing those steps. This visual approach aims to facilitate retrospective analysis, and to promote timely completion of the multi-step process and improve performance.

The Process Overview and Steps in Details are the most interactive parts. In the Process Overview, the managers can interact

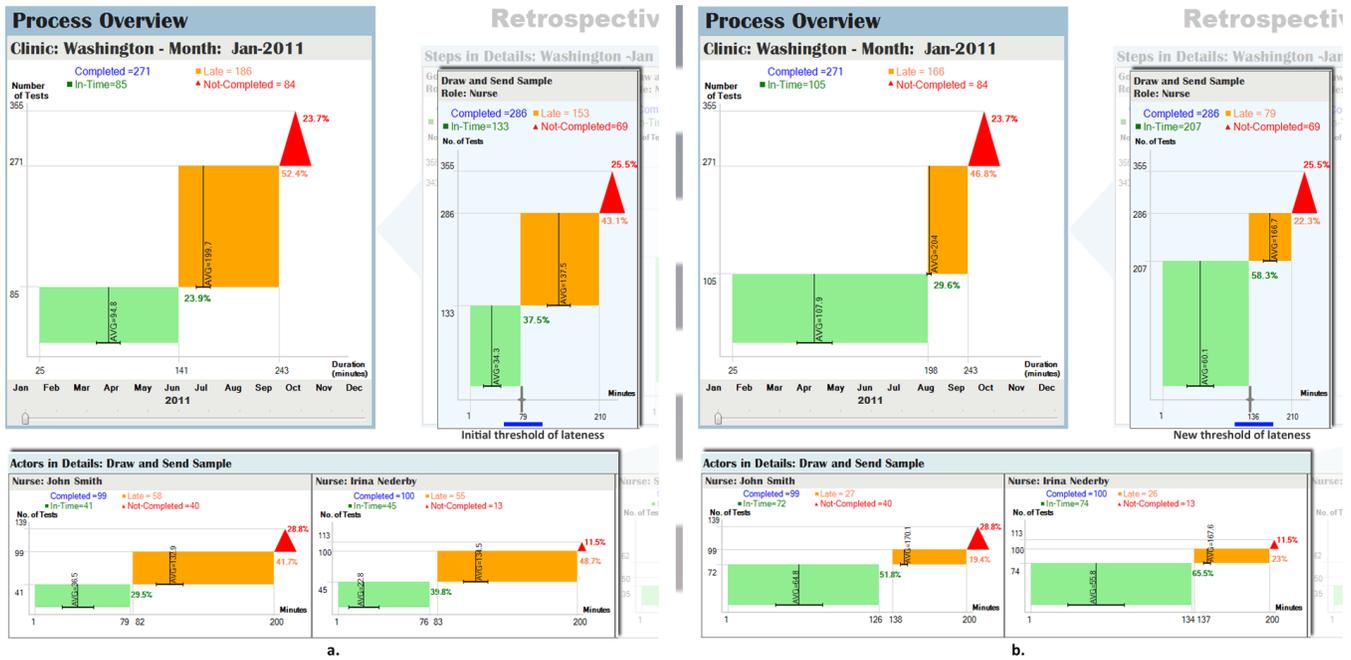


Figure 6: a. For space reasons, this figure shows how some parts of the MSProVis changes when the manager adjusts the threshold of lateness from 79 minutes to 136 minutes (close to initial average value of late completions). Actors' performance has improved almost 50 % of their initial one. From the gap, managers can see that *Irina* was executing completions in less time initially, but as the threshold's value increased, *John* is doing better. The triangles still remain the same, as the changes in the duration do not affect the not-completed completions. If we look at the process overview, the number of late completions has not dropped much. Looking at the new process' average, managers can notice that this happens because most of the late completions have an average equal 204. So, increasing the threshold by 6 minutes will considerably change the number of late completions, see Figure 7.

with the slider and view historical data for different months. Automatically, the Steps and Actors in Details views show steps and actors' performances for the selected month. MSProVis calculates initial predefined thresholds of lateness, and allows managers to adjust those thresholds interactively in Steps in Details. Threshold changes are also reflected in the Process Overview and Actors in Details view.

Figures 6 and 7 show how process, steps, and actors' performances changes when the manager changes the threshold of lateness from 94 to 136, and to 143. For space reasons, the figures show only some parts of MSProVis.

5 IMPLEMENTATION

MSProVis is constructed in uVis [12], a formula based visualization tool for relational databases developed at the IT University of Copenhagen. uVis also supports user interface development, and allows users to extend their applications with visualizations.

uVis has an integrated development environment, where designers can create visualizations by dragging and dropping controls (e.g. triangle, rectangle, etc.) in the design panel, and specifying spreadsheet-like formulas in the property grid. By means of formulas, controls can be bound to data, refer to other controls, define the appearance, and the behavior of each control. The formula language and the integrated development environment allowed us to rapidly create several visualization prototypes and led us to MSProVis.

6 EVALUATION

The purpose of the usability study was to investigate if users are able to understand the information presented in the PCD and efficiently use MSProVis to perform representative tasks. This study

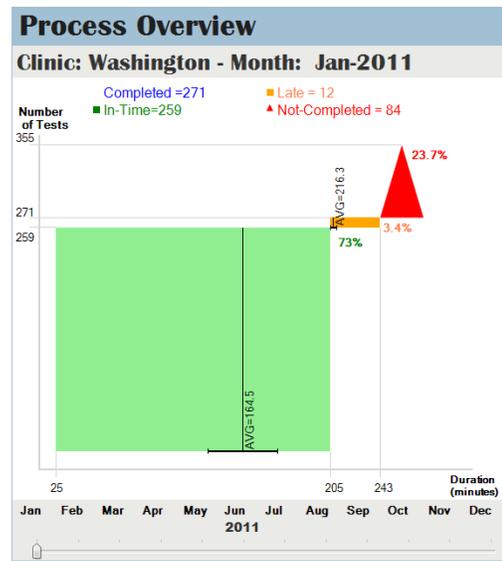


Figure 7: It presents how the number of late completions changes when the threshold of lateness in step 2 increases by 64 minutes, from 79 (Figure 6.a) to 143 minutes. This means that the process overview threshold will be adjusted to 205 (141+64). Note the initial average is smaller than 205 for late completions. That is why the late completions dropped noticeably.

also focused in identifying potential usability issues, and providing suggestions for further improvement.

We conducted a pilot usability test with five graduate students (3 female and 2 male) from the University of Maryland. Although they were good at computer usage in general, they were not familiar with process analysis and did not have a medical background. All participants were introduced for the first time to the PCD and MSProVis during this usability test.

In this study, we used a de-identified medical dataset of 1606 records, extracted from a four step medical process in a hospital. However, as this dataset did not include information about not-completed tests, or information regarding the actors involved in each step, we processed this dataset and added some synthetic data for our usability study.

6.1 Procedure and Tasks

Each usability test consisted of 4 parts and was conducted in 30 minutes. First, participants were shown the PCD using the Process Overview screen, asked to interpret the PCD without any training, and whenever there was a misinterpretation the conductors of the study assisted them before they moved on the next step. The second part took 10 minutes and consisted of an exploration of MSProVis. After getting an understanding of the PCD and MSProVis, participants used 10 minutes to carry out these tasks in a think-aloud manner:

For April 2011:

1. How many blood tests were in-time, late, and not-completed?
2. What is the minimum, maximum and average duration for performing late blood tests?
3. Which step has the highest percentage of in-time, late, and not-completed?
4. If you increase the threshold of lateness in step 3 “Examine Specimen and Record Results”, how many late tests are there?
5. Why do you think the number of late tests changed in that way?
6. Which actor in Step 2 “Draw and Send Sample” has the highest percentage of not-completed tests?
7. If you decrease the threshold of lateness, how does it affect the performance of this actor?
8. If you compare March, April and May, which month has the highest percentage of late and not-completed tests for Step 2 “Draw and Send Sample”?
9. Which actor for each month is responsible for the not-completed tests?

These tasks tested if participants were able to quickly obtain an overview of the process, view details for steps and actors, adjust threshold interactively, and compare performances for different months.

In the last 5 minutes, participants were asked if they had any suggestion for improvements, and to respond to these subjective evaluation questions using a 10-point scale:

1. Give your evaluation of how comprehensible is the PCD.
2. Give your evaluation of how comprehensible is the MSProVis.
3. Give your evaluation of the interactive features in the MSProVis.

6.2 Results

6.2.1 Understanding data in PCD

The evaluation showed that all users were able to read most of the information presented in the PCD. They were able to identify easily the number of in-time, late and not-completed completions in the Y-Axis and associate them with colors. Information presented in the X-axis was more difficult to understand at a glance.

Positioning the triangle next to the orange rectangle made two of the participants assume that not-completed were considered those that have a duration time higher than the maximum time in late completions. One of them suggested that using a circle, instead of a triangle may have been less confusing. Also, two participants found the average confusing because the text was aligned vertically and the line was not extended till the X-axis. Standard deviation notation was not guessed by two participants, while the others related the horizontal bar with the average and recognized it. Positioning the in-time percentage performance below the orange rectangle caused a misunderstanding to one of the users, who associated the percentage with the late completions. The color-encoding did not help in this case. While others found the consistent color encoding helpful and easy to associate with shapes.

Participants interacted with the interfaces and accessed through tooltips details-on-demand. Two participants would prefer to obtain details on mouse hover instead of mouse click.

6.2.2 Exploring data using MSProVis

After the first step, where participants were introduced to the PCD, they started exploring the MSProVis. They started exploring the data presented in the Process Overview by interacting with the slider. They were able to compare the performances between months. Participants were told that by making a double-click on the Process Overview, the Steps in Details would open. One of the participants clicked over an orange rectangle and expected to see only information related to the late completions.

All participants found the Steps in Details view intuitive and easy to understand. They could recognize the steps in the process and compare the number of tests that were completed and not-completed in a step. However, the different timescales in the steps caused some confusion when participants compared the duration time. They appreciated the fact that they could adjust the threshold of lateness. All participants were able to understand what happened when the threshold changes. However, threshold of lateness bar were not very noticeable as interactive affordances. Participants were able to compare different steps identifying which step caused the lengthening of the process.

Finally, the participants explored the last view of the MSProVis, the Actors in Details, by double-clicking in one of the steps. They were able to rapidly gain insight on actors’ performances through the PCDs. Participants were able to more easily compare actors’ performances than steps’ performances, due to the consistent scales used in X-axis and Y-axis between actors. Initially, participants found it strange to see the gap between the green and orange rectangle. However, they were able to figure out what it meant on their own. One of the participants pointed out that a marker for the threshold of lateness might have assisted them more in understanding the gap.

6.2.3 Executing tasks

The participants used the data from April 2011 for the first seven tasks. In task 8 and 9, participants were asked to compare performance for March, April and May 2011.

All participants answered all tasks correctly. The first seven tasks were performed easily and quickly. Average execution time for all 7 tasks was 5 minutes. The last two tasks were more complex, and took approximately 2 minutes for both. Users had to switch from one month to another, and memorize the late and not-performed

percentages for 3 months. Although, the comparison was successful, the results indicated the need for a better approach in comparing processes, actors and steps' performances different points in time.

The usability test showed that users familiar with the PCD and MSProVis were able to rapidly perform representative task, and gain an overview of the process, view details-on-demand for steps and actors, interact and create scenarios to extract insights, and compare performances across different levels.

6.2.4 Summary and Suggestions for Improvement

Overall, participants could interpret the PCD and use MSProVis. On average, they rated them 8.5 out of 10 for being comprehensible. Interactive features in the visualization were rated 7 out of 10. Showing tooltip on mouse hover, and having a thicker adjustable threshold, should improve interactivity. Accessing Steps and Actors in Details was easy and intuitive once they knew how to interact with MSProVis.

Our evaluation showed that the PCD allows users to quickly identify good and critical performances by means of color, shape, and size. Comparing steps and actors PCDs was easier than comparing PCDs in different months, as MSProVis is limited in showing data for a single month. During the study we observed that users initially were seeking information using the size and color of shapes, and once they identified large colored shapes they referred to the labels to ensure their assumptions were correct (list of details on the top of each PCD, percentage, etc.).

Several easy improvements were identified and will be included in the future versions of the PCDs and MSProVis. For example, the percentage labels should be placed at the right-bottom corner of each rectangle. The current position was misleading. Introducing a marker for the default threshold of lateness, would indicate the initial value of the threshold. Currently, the average label is aligned vertically. Aligning horizontally the average label, and connecting the average line with X-axis, would be easier to read. Introducing mouse-hover tooltips, and making the threshold line easier to select, would improve interactivity.

More challenging suggestions related to timescales used in the X-axis of Steps in Details, and the size of the triangle, need further investigation. One of the participants suggested that the triangle should have a proportional width and height, and the steps should not be distributed evenly. Regarding the triangle position, a suggestion was to explicitly distinguish the triangle from the duration on the X-axis. Also, using a circle might solve the ambiguity introduced by the triangle and the duration X-axis. However, these changes need further investigation as it may affect the space-efficiency characteristic of the PCD, and cause confusion.

7 CONCLUSIONS

This paper presents a visual analytics approach to promote timely completion of multi-step processes. A novel visualization called PCD was designed to summarize event-log data and present them in a space-efficient and comprehensible way. We developed MSProVis, which combines several PCDs, to visualize process, steps and actors' performances. MSProVis is a multi-view presentation that allows managers to analyze and compare several scenarios interactively using thresholds of lateness and different points in time.

We evaluated our approach using an example from the medical domain. Users were able to understand most of the data visualized in the PCD without any prior training, and effectively use MSProVis after 10 minutes of exploring the tool. The PCD and MSProVis can be applied in other domains, where multi-step processes are present.

This research is a first attempt to address the complexity of summarizing event-logs of multi-step processes in an efficient and effective way.

8 FUTURE WORK

In the future, we are planning to conduct usability studies with managers in the medical domain, using realistic data. Also we will apply the PCD and the MSProVis in other domains.

Our approach is limited in dealing only with linear processes, but the PCD's principles could be applied in parallel processes as well. Further, the Actors in Details can contain a large number of PCDs, but it shows only seven without interacting with the scroll bar. Future work will address these challenging issues.

The feedback from our usability study will be used to produce a better version of the PCD and MSProVis. Implementations include: implementing mouse hover for more information, designing more visible threshold bars, re-positioning some of labels, etc. More challenging improvements such as comparing more than one month in a single screen will be considered as well.

9 ACKNOWLEDGMENTS

The authors would like to thank Professor Soren Lauesen and the members of the uVis project at the IT University of Copenhagen. Authors also thank Zach Hettinger for providing valuable comments on drafts, Hank Rappaport for providing sample of data, and all the participants in the study. This work was partially supported by Grant No. 10510592 for Patient-Centered Cognitive Support under the Strategic Health IT Advanced Research Projects Program (SHARP) from the Office of the National Coordinator for Health Information Technology.

REFERENCES

- [1] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of Time-Oriented Data*. Human-Computer Interaction. Springer Verlag, 1st edition, 2011.
- [2] N. Ayoub, R. Martins, K. Wang, H. Seki, and Y. Naka. Two levels decision system for efficient planning and implementation of bioenergy production. *Energy Conversion and Management*, 48(3):709 – 723, 2007.
- [3] J. Bertin. *Semiology of graphics: diagrams, networks, maps*. University of Wisconsin Press, Madison, WI, 1983.
- [4] S. K. Card, J. D. Mackinlay, and B. Shneiderman, editors. *Readings in information visualization: using vision to think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.
- [5] D. Chugh and M. H. Bazerman. Bounded awareness: what you fail to see can hurt you. *Mind Society*, 6(1):1–18, 2007.
- [6] J. Desel, W. Reisig, and G. Rozenberg, editors. *Lectures on Concurrency and Petri Nets, Advances in Petri Nets*, volume 3098 of *Lecture Notes in Computer Science*. Springer, 2004.
- [7] G. W. Dickson, G. DeSanctis, and D. J. McBride. Understanding the effectiveness of computer graphics for decision support: a cumulative experimental approach. *Commun. ACM*, 29(1):40–47, Jan. 1986.
- [8] J. Hickner, D. Graham, N. Elder, E. Brandt, C. Emsermann, S. Dovey, and R. Phillips. Testing process errors and their harms and consequences reported from family medicine practices: a study of the american academy of family physicians national research network. *Qual Saf Health Care*, 17(3):194–200, 2008.
- [9] International Standard Organization and International Electrotechnical Commission. Ieee standard adoption of iso/iec 15939:2007 systems and software engineering measurement process. *IEEE Std 15939-2008*, pages C1 –40, 30 2009.
- [10] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, editors. *Mastering The Information Age - Solving Problems with Visual Analytics*. Eurographics, November 2010.
- [11] S. Kriglstein and S. Rinderle-Ma. Change visualizations in business processes. In *Proceedings of the International Conference on Information Visualization Theory and Applications*, 2012.
- [12] S. Lauesen, M. Kuhail, K. Pantazos, S. Xu, and M. B. Andersen. Extending applications with visualization. Submitted to INFOVIS 2012, 2012.
- [13] X. Luo, F. Tian, W. Liu, D. Teng, G. Dai, and H. Wang. Visualizing time-series data in processlines: design and evaluation of a process

- enterprise application. In *Proceedings of the 2010 ACM Symposium on Applied Computing, SAC '10*, pages 1165–1172, New York, NY, USA, 2010. ACM.
- [14] A. Malik, R. Maciejewski, B. Maule, and D. Ebert. A visual analytics process for maritime resource allocation and risk assessment. In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on*, pages 221–230, 2011.
- [15] R. S. Mans, M. H. Schonenberg, M. Song, W. M. P. van der Aalst, and P. J. M. Bakker. Process mining in healthcare - a case study. In *HEALTHINF (1)*, pages 425–438, 2008.
- [16] J. McNames, D. Abercrombie, D. Turner, and T. Shannon. Semiconductor manufacturing process visualization. In *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, volume 5, pages 4940–4944 vol.5, oct. 2003.
- [17] K. Pandazo, A. Shollo, M. Staron, and W. Meding. Presenting software metrics indicators - a case study. In *Proceedings of the 20th International Conference on Software Product and Process Measurement*, 2010.
- [18] C. Plaisant, D. Heller, J. Li, B. Shneiderman, R. Mushlin, and J. Karat. Visualizing medical records with lifelines. In *CHI 98 conference summary on Human factors in computing systems, CHI '98*, pages 28–29, New York, NY, USA, 1998. ACM.
- [19] A. Savikhin, R. Maciejewski, and D. Ebert. Applied visual analytics for economic decision-making. In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 107–114, oct. 2008.
- [20] B. Shneiderman and C. Plaisant. *Designing the user interface: strategies for effective human-computer interaction*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2010.
- [21] D. Solomon, H. Hashimoto, L. Daltroy, and M. Liang. Techniques to improve physicians' use of diagnostic tests: a new conceptual framework. *The Journal Of The American Medical Association (JAMA)*, 280(23):2020–7, 1998.
- [22] C. Speier and M. G. Morris. The influence of query interface design on decision-making performance. *MIS Quarterly*, 27(3):397–423, Sept. 2003.
- [23] S. J. Steindel and P. J. Howanitz. Physician satisfaction and emergency department laboratory test turnaround time observations based on college of american pathologists q-probes studies. *Archives of Pathology*, 125(July):863–871, 2001.
- [24] S. Tarkan, C. Plaisant, B. Shneiderman, and A. Hettinger. Reducing missed laboratory results: defining temporal responsibility, generating user interfaces for test process tracking, and retrospective analyses to identify problems. In *AMIA Annu Symp Proc*, volume 2011, pages 1382–1391, 2011.
- [25] The Standish Group. Chaos summary 2009. <http://www.standishgroup.com/newsroom/chaos.2009.php>, 2009.
- [26] W. M. P. Van Der Aalst. *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, volume 136. Springer, 2011.
- [27] W. M. P. van der Aalst and A. H. M. ter Hofstede. Yawl: yet another workflow language. *Inf. Syst.*, 30(4):245–275, June 2005.
- [28] C. Van Walraven and C. D. Naylor. Do we know what inappropriate laboratory utilization is? a systematic review of laboratory clinical audits. *The Journal Of The American Medical Association (JAMA)*, 280(6):550–558, 1998.
- [29] M. van Wijk, M. Mosseveld, and J. van der Lei. Design of a decision support system for test ordering in general practice: choices and decisions to make. *Methods of Information in Medicine*, 38(4-5):355–61, 1999.
- [30] C. V. Walraven, V. Goel, and B. Chan. Effect of population-based interventions on laboratory utilization: a time-series analysis. *The Journal Of The American Medical Association (JAMA)*, 280(23):2028–2033, 1998.
- [31] T. D. Wang, C. Plaisant, A. J. Quinn, R. Stanchak, S. Murphy, and B. Shneiderman. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, CHI '08*, pages 457–466, New York, NY, USA, 2008. ACM.
- [32] C. Ware. *Information Visualization, Second Edition: Perception for Design (Interactive Technologies)*. Morgan Kaufmann, 2 edition, Apr. 2004.
- [33] M. Weske. *Business Process Management: Concepts, Languages, Architectures*, volume 54. Springer-Verlag New York Inc, 2007.
- [34] K. Wongsuphasawat, J. A. Guerra Gómez, C. Plaisant, T. Wang, M. Taieb-Maimon, and B. Shneiderman. Lifeflow: visualizing an overview of event sequences (video preview). In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems, CHI EA '11*, pages 1747–1756, New York, NY, USA, 2011. ACM.