Intelligent Environments 2018
I. Chatzigiannakis et al. (Eds.)

doi:10.3233/978-1-61499-874-7-346

Empowering Ergonomy in Workplaces by Individual Behavior Modeling Using Interactive Process Mining Paradigm

Carlos FERNANDEZ-LLATAS [a,1], Gema IBANEZ-SANCHEZ [a] Vicente TRAVER [a] and Fernando SEOANE [b,c,d]

[a] *ITACA-SABIEN, Universitat Politècnica de València, Spain*
[b] *Department for Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden.*
[c] *Department of Biomedical Engineering, Karolinska University Hospital, Stockholm, Sweden*
[d] *Swedish School of Textiles, University of Boras, Boras, Sweden*

Abstract. Work-related disorders account for a significant part of total healthcare expenditure. Traditionally muscle-skeletal disorders were predominant as source of work absenteeism but in last years work activity-related disorders have increased remarkably. Too little activity at work, sedentarism, or too much work activity leads to stress. The individualized behavioural analysis of patients could support ergonomy experts in the optimization of workplaces in a Healthier way. Process Mining Technologies can offer a human understandable view of what is actually occurring in workplaces in an individualized way. In this paper, we present a proof of concept of how Process Mining technologies can be used for discovering the worker flow in order to support the ergonomy experts in the selection of more accurate interventions for improving occupational health.

Keywords. Behavior Modeling, Process Mining, IoT, Smart Environments

1. Introduction

With the arrival of new mobile personal technologies and wearable sensors, the quantity of data available for monitoring the behavior of people is dramatically growing [6]. The rapid digitization of society leads to an exponential growth data from Internet of Things (IoT) devices [3]. According CISCO Visual Networking Index Prediction[1], the number of connected things on the Internet will arise to 26.3 billion by 2020. All this information, added to information already stored in Electronic Health Records (EHR), social media or patient portals, among others, suppose a great opportunity to extract a valuable knowledge that will help to improve the quality of life of citizens[19]. Telemedicine and telehealth is being part of this IoT revolution[8]. It will be a critical piece of the digital

[1] Corresponding Author: Carlos Fernandez-Llatas Universitat Politècnica de València, Camino de Vera S/N 46022 Valencia Spain; E-mail:cfllatas@itaca.upv.es

transformation of healthcare. Currently, the market is already plenty of wearables and mobile apps for medical providers, disease-specific apps, such as diabetes [2], medical education and teaching, apps for patients and general public, including health and fitness apps, diet and nutrition [7]. Also apps for providing EHR access and patient information and the ability to share it with caregivers, family, or clinicians, or telemedicine and tele-healthcare stakeholders. This is useful in a large quantity o cases such as in stroke or acute trauma[4]. As seeing, the IoT healthcare market is growing at breakneck speed. It is due to the increase of chronic diseases associated with lifestyle and the fact that healthcare applications are capable of providing cost-effective solutions, improving communication between patients and healthcare providers.

This opens the door towards a new generation of sensors, lab equipment, employee wearables for working monitoring, where IoT will massively increase the amount of data available for the analysis of the ergonomy at work places, bringing a new complexity level. This makes think that precision medicine[17] can be a reality before expected although hardly exists today. Initiatives as Obamas 4P[20] (personalized, predictive, preventive and participatory) are pioneer where, in his words, *a new model of patient-powered research that promises to accelerate biomedical discoveries and provide clinicians with new tools, knowledge, and therapies to select which treatments will work best for which patients*. This can lead to use more advanced analytics, visualizations and decision support tools to improve accuracy in the diagnostics, allowing more effective and precise treatments.

In this line, Work-related disorders account for a significant part of total healthcare expenditure. Traditionally muscle-skeletal disorders were predominant as source of work absenteeism but in last years work activity-related disorders have increased remarkably. Too little activity at work, sedentarism, or too much work activity leads to stress. Work interventions should be executed to ensure a healthy distribution of work among employees and other stakeholders to avoid sedentarism that contributes to obesity or excessive workload that leads to occupational stress and depression.

Using all the information gathered from the plain monitoring of working environment it is possible to extract patterns that will allow to track, analyze and optimize the behavioral patterns of people at workplace. In this line, Process Mining [21,15], can be a powerful solution for supporting ergonomy experts in the process of understanding the patients. This technology, that traditionally was used for improving and optimizing the business processes in enterprises, is based on the application of syntactic machine learning technologies for inferring processes in an human understandable way. Process Mining techniques can be used for discover and understand processes by processing Internet of Things (IoT) data available data from monitoring environments [12]. Using these techniques it is possible to build different patterns that shows the general behaviour model of workers in an human understandable way.

Also, using Process Mining Conformance Technologies and Clustering algorithms it is possible to stratify the workers depending on their health behavior [11]. This allow to ergonomy experts to detect the workers behavioral status in an understandable way and propose actions to correct unhealthy behaviors. In addition, by comparing the personal models in times it is possible to detect behavioral changes that will support ergonomy experts in the measuring of the accuracy of their proposed actuation [11] and evaluating the distance between the current health status of the worker and their desired status.

To track the movements of the employees the facilities will provide with an accurate indication of the work activities and duration, that will be inferred, using Process Mining tools, as formal work behavioral models. Such models will allow the occupational health expert to assess work environments working life behaviors, evaluating healthy risk and detecting hazardous individual working behavior. The occupational health expert using the solution will have the opportunity to propose work interventions aiming to reduce risk and promote a healthy working lifestyle with a better and accurate knowledge of the working environment.

In this paper, we present a proof of concept on the application of Interactive Process Mining technologies for supporting Ergonomy experts in the discovery of unhealthy and unsafe patterns in monitored working scenarios. For that, we simulate individual behaviors of workers for building a Internet of Things (IoT) event logs and, applying Process Mining techniques. Then, we will show how ergonomy experts will be enabled to understand the actual behavior of workers in those scenarios.

The results shows how Process discovery Techniques are able to infer the general behavior of workers in a general way, as well as, using clustering Process Conformance algorithms, to stratify the behavior of different workers in different kind of flows depending on its evolution. This will support ergonomy experts in the understanding of the behavioral aspects of individual workers as well as detecting optimization possibilities for improving the ergonomy in workplace in a easier way.

In order to test it, we have designed a scenario based on the actions of janitors at an university. The idea is to simulate the different behaviors of janitors and on one hand present the general behavior of janitors over working days, and, on the other hand, show how the different behaviors can be discovered and presented to ergonomy experts, using Process Mining techniques.

This paper is structured as follows, the next section the simulation process is stated and the Process Mining technology used in this work is presented. In Results section, the flows inferred using Process Mining technology was explained. Finally, a discussion part concludes the paper.

2. Materials and Methods

Process Mining technology is a relatively new paradigm based on syntactical data mining framework that is though to support process experts in the understanding of the processes. Process Mining provides, algorithms, tools and methodologies to show what is really occuring inside the actual process that, usually, not correspond with the perceived one [10]. Process Mining technologies sacrifice accuracy in the learning process in order to provide more human understable processes. This easy understandability allows experts to add its own kwowledge to the learning process by correcting the actions in an iterative way. In this way, the application of Process Mining technologies can be used for supporting health experts in the management in occupational health via providing more understandable models, evaluating his actions and making them conscientious of the specific characteristics of workers in a general and individualized way.

There are three main kind of techniques associated to Process Mining:

- *Process Discovery* that are algorithms that produce graphic human understandable flows from event logs.

- *Process Conformance* that are algorithms that are able to compare logs and models in order to decide if the log is according the model or even measure the difference between two models.
- *Process Enhancement* that are tools that provide an augmented view of the process that allow highlight their specific characteristics in order to make easier their understandability by experts. For example, heat maps showing the most common paths in a flow

The general flow can provide a common view of the status, however due to the different personalities of the stakeholders, not all the actors contibute the actions in the same way to the general flow. In order to detect the differences among the individual users it is possible to use Clustering techniques [9] for creating partitions that maximize the similarity among the elements of each cluster. There are several clustering algorithms in the literature. K-Means [9] is one of the most known Clustering algorithm. This method is able to split a set of samples in a given number of partitions, maximizing the similarity among the group members. However, in our problem, the number of partitions is not known. Other algorithm available in literature is Quality Threshold (QT) Clustering[18]. This algorithm requires a threshold distance that define the maximum distance among the members of each cluster. In that case, we can define a distance and the algorithm will build the groups that maximize the similarity among their members taking the given distance as the maximum distance.

Nevertheless, clustering algorithms require a distance function between two samples in order to construct the groups. Usually, the samples are data vectors that use classical distances like euclidean. However, in our problem, the samples are synctactical ordered traces and the euclidean, and other geometrical distances that does not take into account the order, are not a good distance measures in our case. For calculate the distance over syntactically ordered data there are classical specific algorithms like Levenshtein [5] , and in case of Process Mining techniques Edition Distance Workflow Algorithm (EDWA) [11] takes into account the topology of the samples for computing

3. Results

In this paper, we will test the possibilities for detecting different human behaviors at workplaces using Process Mining techniques. For testing Process Mining in the behavior discovery in workers, we have designed an experiment via simulating different possible behaviors of Janitors at work. In this way, we have simulated 4 types of behaviors: a) Janitors that stay all day at the office, b) janitors that worked in the morning in some active tasks (Open and Closing class rooms, and Mailing delivery) and stay in the office after lunch, c) janitors that stay in the office in the morning, and perform the active tasks in the evening, and d) Janitors that performs active task during all day. We create these event logs using an Ambient Assisted Living simulator [16], that present the different activities performed by the users. We simulated a total of 140 traces 50 of type a), 40 of type b) 30 of type c) and 20 of type d). The available activities performed by Janitors are presented in Table 1.

In this table, all the possible activities performed by Janitors is presented. We have selected activities that could be collected by common domotic systems available currently in universities, like for example, card control for rooms entry, and the associated

Table 1. Set of simulated actions performed by janitors

Action	Average (in minutes)	Standard Deviation
Breakfast Room	60	15
Concierge Office	240	15
Mail Delivery	120	5
Open Classroom	120	15
Close Classroom	120	15
Lunch Room	60	15

semantic information. For example, Breakfast and lunch rooms activities are in the same location, but depends in the hour the activity has a different semantic meaning.

ID	Name	Start	End
065180c91efc4d89bae36bfa5f816b21	BreakFast Room	13/03/2018 8:21	13/03/2018 9:32
065180c91efc4d89bae36bfa5f816b21	Mail Delivery	13/03/2018 9:32	13/03/2018 11:30
065180c91efc4d89bae36bfa5f816b21	Close Class Room	13/03/2018 11:30	13/03/2018 13:28
065180c91efc4d89bae36bfa5f816b21	Lunch Room	13/03/2018 13:28	13/03/2018 14:30
065180c91efc4d89bae36bfa5f816b21	Concierge Office	13/03/2018 14:30	13/03/2018 18:52
065180c91efc4d89bae36bfa5f816b21	Home	13/03/2018 18:52	13/03/2018 19:47
068882948a1e43bb8288f539d9efcc41	BreakFast Room	13/03/2018 8:29	13/03/2018 9:28
068882948a1e43bb8288f539d9efcc41	Concierge Office	13/03/2018 9:28	13/03/2018 13:35
068882948a1e43bb8288f539d9efcc41	Lunch Room	13/03/2018 13:35	13/03/2018 14:51
068882948a1e43bb8288f539d9efcc41	Concierge Office	13/03/2018 14:51	13/03/2018 18:38
068882948a1e43bb8288f539d9efcc41	Home	13/03/2018 18:38	13/03/2018 19:46
0822d40cd7624abbb4bfe3a22d7a245f	BreakFast Room	13/03/2018 8:58	13/03/2018 10:08
0822d40cd7624abbb4bfe3a22d7a245f	Concierge Office	13/03/2018 10:08	13/03/2018 14:30
0822d40cd7624abbb4bfe3a22d7a245f	Lunch Room	13/03/2018 14:30	13/03/2018 15:21
0822d40cd7624abbb4bfe3a22d7a245f	Concierge Office	13/03/2018 15:21	13/03/2018 19:11
0822d40cd7624abbb4bfe3a22d7a245f	Home	13/03/2018 19:11	13/03/2018 20:22

Figure 1. Event Log simulated

In Figure 1 a part of the simulation resultant event logs is shown. In this log each janitor has a different ID. The name represent the activity performed, and Start and End represent the time when the action take place.

Using the logs simulated, we will use a Process Discovery system for inferring their flow associated. In order to do that we will use PALIA Suite Tool [12]. Also we apply heat maps over he flow to show the time spent by stakeholders in each one of the activities of the flow.

Figure 2 represents the flow inferred by PALIA Suite using the simulated logs. As a Heat Map we use a color gradient for representing the time spent in each one of the activities (nodes) and in the transitions the color represents the number of events occurred, that means the number of janitors that follow this path. In the flow we can see that the janitors seems follow a correct general behavior expending time in active actions (Open and Close Class Rooms, and Mail Delivery) in the same degree that static one (stay in concierge office).

Seeing this view, it can be though that, in general, janitors have a healthy behavior alternating active and static actions. However, according the simulation we know that

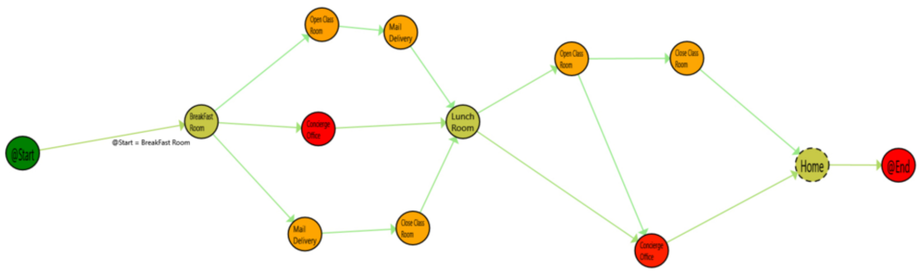

Figure 2. All Janitors behavior inferred

there are janitors that have different behaviors. In order to present how Process Mining technologies are able to detect and show the type of behaviors of janitors we will use Clustering algorithms. PALIA Suite Tool has currently implemented a Quality Threshold Clustering algorithm using a Edition Distance Workflow Algorithm. So for this proof of concept we will use this combination of algorithms in order to provide the clusters. PALIA suite allows the selection of thresholds between 0.0 and 1.0 representing the percent distance between two samples. A maximum distance (1.0) represents traces that have not common activities, while a minumin distance (0.0) represents equal flows.

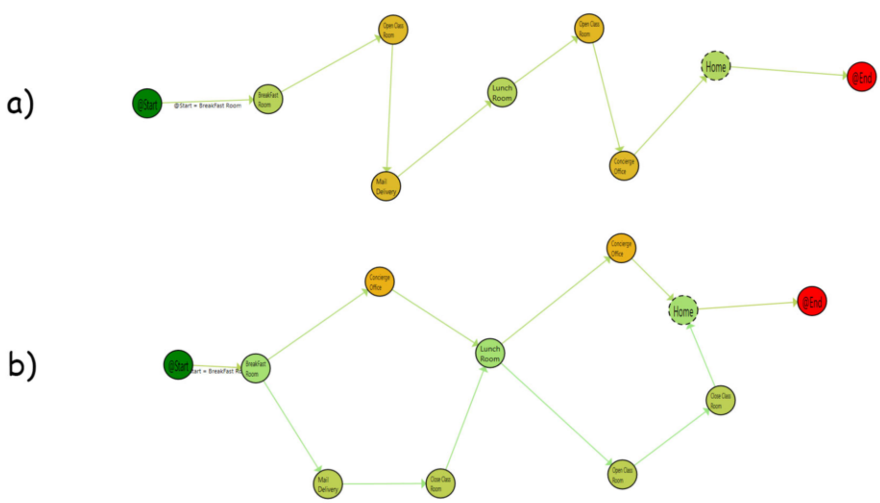

Figure 3. Janitors groups after clustering with a 0.3 threshold

Figures 3 and 4 and shows the different behaviors groups detected by the clustering algorithm. With a threshold upper than 0.4, PALIA suite consider all the traces in only one group. Figure 3 shows the results of applying clustering algorithm with a threshold of 0.3. In the Figure is possible to see people that are active almost all day a) (Stay less time in the Concierge Office) separated from those that are more static b).

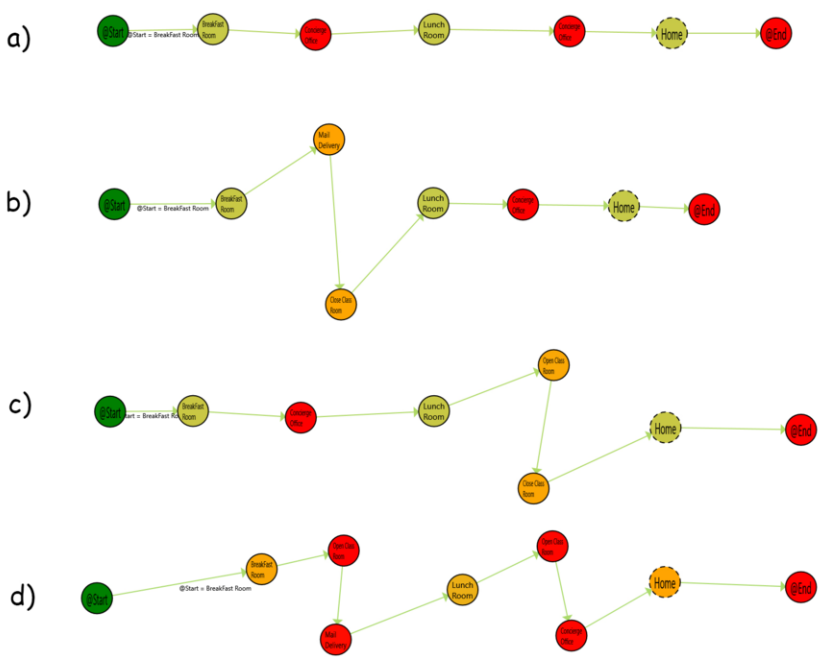

Figure 4. Janitors groups after clustering with a 0.2 threshold

On the other hand, in we apply a threshold below 0.2 (Figure 4) we can detect all the different groups that we simulated, a) those that are more inactive; those that b) are active at morning; c) those that are active at evening; and d) those that are active all day.

4. Discussion and Conclusions

The application of Process Mining over the information available in activity logs can be used for identify behavioral models that can be used for experts to optimize processes and correct anomalous behaviours. Classical data mining technologies and machine learning are able to provide better accurate models for classify actions, infer accurate models for predicting undesired situations, or even discover models that classify the behavior of the worker.

However, although the Machine Learning tools will support in the general understanding of the behavioral status of the worker, these techniques are not able to show the behavioral process in an easy an understandable way to ergonomy experts. So, although, there is possible to characterize the status of the general aspects of the workplace and compare it with individualized behavior of personal workers it is not possible to show the reasons of these differences in order to support ergonomy experts in the selection of best actions for improving the safety and health of people at workplaces.

Interactive Pattern Recognition paradigm [14] was born for providing a machine learning framework not only for understanding better what occurs inside the inferred models, but also to allow the incorporation of the experts experience within the inferred models. However, Interactive paradigm requires that the result of Machine Learning algorithms was human understandable. The application of this paradigm with Process Mining technologies can be the solution for include the expert in the middle of the learning process in order to make him conscientious of the characteristics of the problem to solve and providing heuristic clues to the automatic machine learning system for improve the accuracy and the efficiency of the inference.

In this paper we have tested a simple model with a limited granularity. The more events we add to the algorithm the more complete information flow we can achieve. However, the more complete was the workflow the more complex could be to understand, due to the higher quantity of edged and nodes in the flow. This effect is commonly known as *Spaghetti Effect* [13]. In order to avoid this undesired effect is necessary to select the adequate granularity for a better understanding of the process by human experts. Also, the selection of adequate threshold factors will allow the experts to show different situations and behaviors beyond the general process allowing them provide more accurate behavioral correction actions for individul workers.

Process Mining technologies not only can help in the discovery of the processes, but also, can support experts in the evaluation of the correction actions proposed. For example, in our case, the expert can propose to more static janitors to exchange tasks with more active ones. The experts can compare the previous flows with the new traces in order to evaluate the degree of adherence of the janitors to the intervention proposed by ergonomy experts.

In addition, as Process Mining technologies infers formal models it is possible to create simulation models [16] that can be used by experts to evaluate the intervention before propose it. This will support ergonomy experts in the selection of the most adequate interventions for optimizing the system.

References

[1] Cisco vni predicts near-tripling of ip traffic [https://newsroom.cisco.com/press-release-content?type=press-release&articleid=1771211].

[2] Madlen Arnhold, Mandy Quade, and Wilhelm Kirch. Mobile Applications for Diabetics: A Systematic Review and Expert-Based Usability Evaluation Considering the Special Requirements of Diabetes Patients Age 50 Years or Older. *Journal of Medical Internet Research*, 16(4), April 2014.

[3] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The Internet of Things: A survey. *Computer Networks*, 54(15):2787–2805, October 2010.

[4] Maged N. Kamel Boulos, Ann C. Brewer, Chante Karimkhani, David B. Buller, and Robert P. Dellavalle. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online Journal of Public Health Informatics*, 5(3):229, February 2014.

[5] Tilo Buschmann and Leonid V Bystrykh. Levenshtein error-correcting barcodes for multiplexed dna sequencing. *BMC bioinformatics*, 14(1):272, 2013.

[6] Min Chen, Shiwen Mao, and Yunhao Liu. Big Data: A Survey. *Mobile Networks and Applications*, 19(2):171–209, April 2014.

[7] Steven S. Coughlin, Mary Whitehead, Joyce Q. Sheats, Jeff Mastromonico, Dale Hardy, and Selina A. Smith. Smartphone Applications for Promoting Healthy Diet and Nutrition: A Literature Review. *Jacobs journal of food and nutrition*, 2(3):021, 2015.

[8] Dimiter V. Dimitrov. Medical Internet of Things and Big Data in Healthcare. *Healthcare Informatics Research*, 22(3):156–163, July 2016.

[9] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. Wiley-Interscience, New York, NY, 2 edition edition, November 2000.

[10] Carlos Fernandez-Llatas, Jose Luis Bayo, Alvaro Martinez-Romero, Jose Miguel Benedi, and Vicente Traver. Interactive pattern recognition in cardiovascular diseases management. a process mining approach. In *Proceedings of the IEEE International Conference on Biomedical and Health Informatics 2016*, las Vegas, EEUU, 2016.

[11] Carlos Fernandez-Llatas, Jose-Miguel Benedi, Juan M. Garcia-Gomez, and Vicente Traver. Process mining for individualized behavior modeling using wireless tracking in nursing homes. *Sensors*, 13(11):15434–15451, 2013.

[12] Carlos Fernandez-Llatas, Aroa Lizondo, Eduardo Monton, Jose-Miguel Benedi, and Vicente Traver. Process mining methodology for health process tracking using real-time indoor location systems. *Sensors*, 15(12):29821–29840, November 2015.

[13] Carlos Fernandez-Llatas, Antonio Martinez-Millana, Alvaro Martinez-Romero, Jose Miguel Benedi, and Vicente Traver. Diabetes care related process modelling using Process Miningtechniques. Lessons learned in the application of InteractivePattern Recognition: coping with the Spaghetti Effect. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2127–2130, 2015.

[14] Carlos Fernandez-Llatas, Teresa Meneu, Vicente Traver, and Jose-Miguel Benedi. Applying evidence-based medicine in telehealth: An interactive pattern recognition approximation. *International Journal of Environmental Research and Public Health*, 10(11):5671–5682, 2013.

[15] Carlos Fernandez-Llatas, Bernardo Valdivieso, Vicente Traver, and Jose Miguel Benedi. Using process mining for automatic support of clinical pathways design. In Carlos Fernández-Llatas and Juan Miguel García-Gómez, editors, *Data Mining in Clinical Medicine*, number 1246 in Methods in Molecular Biology, pages 79–88. Springer New York, January 2015.

[16] Carlos Fernndez-Llatas, Juan Bautista Mochol, Pilar Sala, Juan Carlos Naranjo, Salvatore F Pileggi, Sergio Guilln, and Vicente Traver. Ambient Assisted Living spaces validation by services and devices simulation. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society/*, 2011:1785–1788, 2011.

[17] J. Larry Jameson and Dan L. Longo. Precision medicine-personalized, problematic, and promising. *The New England Journal of Medicine*, 372(23):2229–2234, June 2015.

[18] Xin Jin and Jiawei Han. Quality Threshold Clustering. In *Encyclopedia of Machine Learning*, pages 820–820. Springer, Boston, MA, 2011.

[19] Burke W. Mamlin and William M. Tierney. The Promise of Information and Communication Technology in Healthcare: Extracting Value From the Chaos. *The American Journal of the Medical Sciences*, 351(1):59–68, January 2016.

[20] William T. Riley, Wendy J. Nilsen, Teri A. Manolio, Daniel R. Masys, and Michael Lauer. News from the NIH: potential contributions of the behavioral and social sciences to the precision medicine initiative. *Translational Behavioral Medicine*, 5(3):243–246, September 2015.

[21] Wil M. P. van der Aalst. *Process Mining: Data Science in Action*. Springer, Heidelberg, 2 edition, 2016.