Linking Diagnostic-Related Groups (DRGs) to their Processes by Process Mining

Alessandro Stefanini[1], Davide Aloini[2], Riccardo Dulmin[2] and Valeria Mininno[2]
[1]Department of Enterprise Engineering, University of Rome Tor Vergata, Via Orazio Raimondo, Rome, Italy
[2]Department of Energy, Systems, Territory and Construction Engineering (DESTEC), University of Pisa, Pisa, Italy

Keywords: Healthcare, Process Discovery, Process Mining, Diagnostic-Related Group (DRG), Patient-flow.

Abstract: The knowledge of patient-flow is very important for healthcare organizations, because strongly connected to effectiveness and efficiency of resource allocation. Unfortunately, traditional approaches to process analysis are scarcely effective and low efficient: they are very time-consuming and they may not provide an accurate picture of healthcare processes. Process mining techniques help to overcome these problems. This paper proposes a methodology for building a DRG related patient-flow using process mining. Findings show that it is possible to discover the different sequences of activities associated with a DRG related process. Managerial implications concern both process identification, analysis and improvement. A case study, based on a real open data set, is reported.

1 INTRODUCTION

Healthcare is one of the most relevant field in every modern society, for two main reasons: the great interest of people for health and the meaningful impact on the world economy. Healthcare is one of the most important economic sector in the developed countries: "OECD Health Statistics 2014" shows that the Health spending accounted for 9.3% of GDP in the OECD countries. Due to the impact of Healthcare on Economy and mostly on Public expenditure, the efficient use of resources is a fundamental point particularly during and after the recent global economic crisis.

Process organization and resources allocation on activities play a key role in achieving the required service level and efficiency. For this purpose, the knowledge of patient-flow is essential in order to understand the current organization of processes and to identify how resources could be allocated more efficiently on organizational units (Vissers, 1998; Haraden and Resar, 2004; Brailsford et al., 2004).

In this context, Diagnosis-Related Group (DRG), as an empirical classification of the final products in a hospital (Fetter and Freeman, 1986), represents a topic of concern for research and a critical success factor for healthcare systems worldwide. In fact, most of the developed countries have already introduced DRG-systems (Busse et al., 2011) for driving resources allocation.

Despite DRGs classify patient groups with similar expected patterns of resource use depending on their diagnosis and treatments (Fetter and Freeman, 1986), it is not still clear the association between DRGs and related care processes (Mans et al., 2009). Therefore, managers are not able to relate the resources received (and theoretically consumed) with the process flow.

Difficulties in process discovery and analysis are mainly due to the peculiar complexity of healthcare processes: highly interconnected, patient dependant and multi-disciplinary in nature (Anyanwu et. al, 2003; Mans et al., 2009; Rebuge and Ferreira, 2012). Given these characteristics, the process mining techniques are potentially useful in identifying process workflow in healthcare environments (Rebuge and Ferreira, 2012). This paper proposes a methodology for creating DRG patient-flow using process mining techniques. Through the proposed approach, it may be possible to understand the various sequences of activities associated with a DRG and, thus, to identify and model the main unit of analysis of a healthcare system. Expected contributions and implications aim to support:

1. *Process identification* in order to better comprehend the actual way of working (Weerdt et al., 2013).
2. *Process analysis* in order to calculate

Stefanini, A., Aloini, D., Dulmin, R. and Mininno, V.
Linking Diagnostic-Related Groups (DRGs) to their Processes by Process Mining.
DOI: 10.5220/0005817804380443
In *Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2016) - Volume 5: HEALTHINF*, pages 438-443
ISBN: 978-989-758-170-0
Copyright © 2016 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

performance indicators, assess resource consumption, detect bottlenecks, (Mans, 2008; Rebuge and Ferreira, 2012; Rovani et al. 2015) and facilitate cost accounting, particularly in an Activity Based Costing perspective.

3. *Process improvement* in order to streamline business processes and optimize resource planning. Moreover, the expected results may also support regional or national authorities in demand aggregation and resource planning at district level.

Furthermore, the paper presents a case study based on a real open data set of the AMC (Academic Medical Center) hospital in Amsterdam, Netherlands.

2 THEORETICAL BACKGROUND

The healthcare industry represents an important and growing research context. It can be characterized on providing individualized cares and, at the same time, on efficiency (Dobrzykowski et al., 2013). In the last years, scholars have strongly focused their attention on many healthcare related sub-streams. In Operation Management and Supply Chain Management literature, the most prominent are: ICT and Technology Assessment in healthcare (e.g. Chau and Hu, 2001; Tzeng et al. 2008); Quality and Lean Thinking (e.g. Stock et al., 2007; Aronsson 2011); SCM and Strategy (e.g. Li et al. 2002; Lee et al., 2011); Resource Planning, Capacity Management and Scheduling (e.g. Bretthauer et al., 1998; Jebali et al., 2006); Business Process Management and Patient Flow in healthcare (e.g. Haraden and Resar, 2004; Brailsford et al., 2004).

2.1 Patient-flow and Healthcare Process

Patient-flow is the movement of patient, with related information and equipment, between departments, staff groups or organizations as part of a patient's care pathway. It is characterized by a sequence of processes/activities that a patient follows from the first contact with the hospital to the discharge. Usually activities in patient flow are highly interconnected, heterogeneous and numerous, but all of them are necessary for the achievement of final outcome (Lenz and Reichert, 2007; Rebuge and Ferreira, 2012; Partington et al., 2015).

The knowledge of patient-flow is very important for healthcare organizations, because strongly connected to effectiveness of care and efficiency of resource allocation (Vissers, 1998; Haraden and Resar, 2004; Brailsford et al., 2004). As a consequence, some authors started studying patient-flows for optimizing resource allocation both in the short-term (e.g. scheduling of treatment) and medium-long term (e.g. planning of resource inside the hospital) (e.g. Vissers, 1998; Haraden and Resar, 2004).

Although business process analysis is recognized as an extremely important activity for healthcare organizations, traditional approaches are low effective and scarcely efficient in such a context: they are very time-consuming and may not provide an accurate picture of the healthcare processes (dynamic, multi-disciplinary, interconnected, numerous and ad hoc) (Anyanwu et. al, 2003; Mans et al., 2009; Rebuge and Ferreira, 2012; Rovani et al., 2015).

The recent introduction of process mining techniques has partially changed the scenario helping to overcome problems in finding and mapping patient-flows. Until today, the studies on patient-flow through process mining have mainly focused on medical and methodological aspects (e.g. Bose and Van der Aalst, 2011; Huang et al., 2012; De Weerdt et al., 2013; Mans et al. 2013; Delias et al., 2015; Rovani et al., 2015).

In particular, at our best knowledge, literature does not present works that link the patient-flows with the DRGs, i.e. the "final products" of hospitals. The lack of a clear association between the healthcare service and its process and, thus, cost structure, is an important issue that needs to be faced in order to improve the efficiency of healthcare systems.

2.2 Process Mining in Healthcare

Due to the complexity of healthcare processes, process mining could be successfully applied to support process (patient flow) discovery in healthcare sector (Van der Aalst et al., 2007). Process mining, in fact, aims to derive meaningful insights from the complex temporal relationships existing between activities and resources involved in processes (Partington et al., 2015).

Although various authors have already proved its appropriateness (e.g. Mans et al. 2008; Mans et al., 2009; Huang et al., 2012; Rebuge and Ferreira, 2012; De Weerdt, 2013; Rovani et al., 2015), the application of process mining technique in healthcare is a relatively new and unexplored field. In particular, different authors have mined patient control-flow using process discovery techniques.

In process discovery context, authors have oriented their attention on different issues. Some scholars have proved and discussed the suitability of process discovery technique to healthcare services (e.g. Mans et al. 2008; Mans et al., 2009; Kaymak et al., 2012; Mans et al., 2013), typically presenting case studies. Other researchers have focused their studies on proposing innovative methodologies and/or new algorithms (e.g. Huang et al., 2012; Rebuge and Ferreira, 2012; Delias et al., 2015). Others have directed their attention on comparing the discovered process workflow with the expected pathways, the medical guidelines or the analogous patient flows in other hospitals (e.g. Montani et al., 2014; Caron et al., 2014; Rovani et al., 2015; Partington et al., 2015).

On the other hand, the use of "conformance checking" and "enhancement" process-mining techniques seems to be still currently understudied in the healthcare field (Partington et al., 2015).

Discovering the patient-flow is strongly dependent on the adopted patient classification. Some authors use new clustering techniques or clustering tools to categorize patients (e.g. Mans et al., 2009; Rebuge and Ferreira, 2012). Others categorize patient basing on main treatment (e.g. De Weerdt, 2013; Caron et al., 2014) or on a preliminary classification inside the emergency centre (Partington et al., 2015).

Nobody has classified patient basing on DRG.

3 RESEARCH DESIGN

3.1 Research Objectives

This paper aims to suggest a methodology for identifying and mapping patient-flow by process mining, following the lens of DRGs.

Process mining is the combination between data mining and traditional model-driven Business Process Management (Van der Aalst, 2011). The main purpose is to discover, analyze and improve processes by extracting knowledge from event logs readily available in enterprise information systems (Van der Aalst, 2011; Van der Aalst et al., 2012).

Process mining in fact exploits information in the event logs to understand how processes are actually performed, rather then what is prescribed or supposed to happen (Mans et al., 2009; Huang et al., 2012).

3.2 Research Methodology

The proposed methodology (Figure 1) refers to process mining methods suggested by previous authors (e.g. Bozkaya et al., 2009; Rebuge and Ferreira, 2012; Mans et al., 2009; De Weerdt, 2013) and consists of four phases: log preparation, log inspection, process discovery, validation of process model.

1. *Log preparation* aims to set up the event log by pre-processing the event data gathered from one or more information systems in order to select the adequate unit of analysis for analysing the patient-flows.

2. *Log inspection* provides first insights about the investigated process. It includes statistical information on the number of cases, the total number of events, the number of different sequences, etc.. In addition, these activities help to filter incomplete and outlier cases.

3. *Process discovery* phase extracts the process model from the event log. In order to point out a clear process workflow from highly complex processes, it could be also necessary to refine the final model ruling out less frequent paths or excluding less relevant activities. Therefore, the log may need to be re-filtered looking at retaining events covering at least 80% of the cases (Bozkaya et al., 2009). Among many available techniques to act process discovery like α-Algorithm, heuristic miner, fuzzy miner, genetic miner and region-based miner, most suitable approaches for complex environment seems to be heuristic mining, genetic mining and fuzzy mining (Van der Aalst, 2011), which are specialized in noise filtering. While highly performant in dealing with noise, genetic process mining is not very efficient for larger models and logs, due to excessive computation times (Van der Aalst, 2011). Heuristic mining algorithm enables users to focus on the main process flow effectively, but it needs to be supported by the application of clustering technique to the event log.

The fuzzy mining addresses the issue of mining unstructured processes using a mixture of abstraction and clustering techniques and attempts to make a more suitable representation for analysts. Fuzzy mining automatically provides a high-level view on the process by abstracting and aggregating details (Mans et al., 2009). Finally, some authors (e.g.

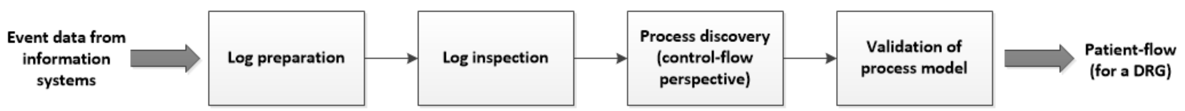

Figure 1: The proposed methodology.

De Weerdt et al., 2013; Mans et al., 2009; Partington et al., 2015) adopt heuristic and fuzzy mining conjointly in order to exploit the advantage of both of them.

In the case study, we have decided to use the fuzzy mining for the motivations previously cited.

4. *Validation* is the final step. Results obtained, i.e. patient-flow model related to the DRG, has to be reviewed by experts in order to check for completeness and correctness. More in-depth validation analysis are also available, for example it is possible to apply the conformance checking technique on the achieved workflow model (Bozkaya et al., 2009; Van der Aalst, 2011; Rovani et al., 2015).

Due to the complexity and the client dependency, which are typical in healthcare processes, it is expected that the patient-flow model presents more than one possible path for the related DRG.

The method could be extended and applied to different DRGs inside the hospital, to obtain the overall process model of the organization.

3.3 Case Study

The case study tests the applicability of the proposed methodology on an open dataset from a gynaecological oncology process at the AMC hospital in Amsterdam (doi:10.4121/uuid:d9769f3d-0ab0-4fb8-803b-0d1120ffcf54). The repository contains about 150.000 events of more than 1100 patient cases. Each case in the event log corresponds to a single patient, so that data present a wide variety of care activities (De Weerdt et al., 2013).

It should be noted that the case study is limited to the first three stage of the methodology.

The first phase of log preparation was very limited, because the datasets had already been integrated and pre-processed by scholars at "Technische Universiteit Eindhoven". We just selected the patient instances with the same group diagnosis ("maligniteit ovarium|tuba"), recognizable in The Netherlands as DBCs (equivalent to DRGs)

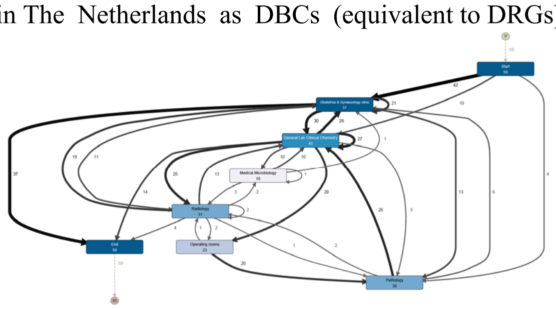

Figure 2: The discovered process model mined by Fluxicon Disco® (the more infrequent activities are excluded).

although not yet introduced at that time. The number of selected case was 60.

In this step, we also decided to designate the hospital operational unit as the adequate level of abstraction for our study and, thus, for aggregating more specific activities. For example, blood test is usually recorded like a set of different events with a distinct name depending on the specific analysis. In this case, we substituted this set with a single event reporting the involved organization unit, the General Lab Clinical Chemistry department.

The second phase of log inspection provided us with a first impression about the processes through event log data and some statistics such as the total number of events, the number of different sequences etc.. Outlier cases were removed. After this operation, the dataset contained 59 cases.

Process discovery was conducted using the Fuzzy Miner implemented by Fluxicon Disco®. We analysed process behaviours and mapped the workflow for the process under investigation. In Figure 2, the less frequent activities have been removed. Please note that just cutting out less than 10% of the events, the simplification in the process map is very relevant. This is significant and let us more confident about the suitability of the methodology and the appropriateness of the selected process mining technique.

Exploring the discovered process models allows us to point out some interesting insight about the patient-flow. For example, as shown in Figure 2, the majority of patients start with a Gynaecological visit; they deepen the diagnosis through a Laboratory analysis or/and Radiological examination (sometimes repeated more than ones); if serious diseases occur the process goes on with a surgery operation and post-operation care, otherwise patient usually end her path after a final Gynaecological visit.

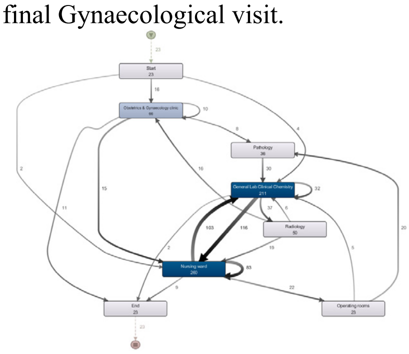

Figure 3: The discovered process model for a classify set of cases (patients who need a surgical operation).

As an addition, the tool allows us to filter cases using specific attributes and sketches patient-flow for a specific case or for a set of patients, e.g. to understand specific practices and to find out possible anomalies.

As shown in Figure 3, we were finally able to determine the main process flow for patients who need a surgical operation: the process generally starts with a Gynaecological visit, then a Laboratory analysis and a radiological assessment are executed, hence the patient is hospitalized and other Laboratory analysis can be done in the meanwhile, after that the patient undergoes a surgical operation, an examination at Pathology depart, again Gynaecological visit and, if needed, additional Laboratory analysis. Finally, the patient is discharged from hospital.

Note that the process discovery phase, while depicting the process flows, also reveals many interesting information about how many times one activity occurred, for example we found out that Gynaecological visit takes place 144 times and Laboratory analysis 304 times. This is useful in order to define resource consumption and considerations about process efficiency. This activity is useful to support resource planning both in the short and medium-long term. Looking at the DRG “maligniteit ovarium|tuba”, we can state that a patient needs an average of 2.5 Gynaecological visits, 5 Laboratory analysis and 5.7 hospitalization days.

Besides, we can find some patterns of resource consumption. For example, only in 38% of cases the patient undergoes a surgical operation but these cases implicate the 62% of diagnosis and treatments and the 77% of nursing activities. This evidence could be useful in order to support process verification and process analysis.

4 CONCLUSIONS

The case study supports the potential applicability of the proposed methodology for process discovery in the healthcare context. Process map and other interesting insights about the patient flow for the “maligniteit ovarium|tuba” cluster are obtained: most relevant process path, number of cases, recurrence of activities (for example Gynaecological visit or Laboratory analysis), total hospitalization days and resource consumption.

The comprehension of the patient flow (and subsequent comparison of the actual way of working within the hospital unit) enables to check and verify business processes, to support process analysis and improvement and to better define resource consumption/allocation (Vissers, 1998). Besides, the tool also provides simultaneous functionalities of clustering, filtering and modifying the abstraction level of the model, which allows decision makers to understand processes for a specific set of patients.

The extension of this methodology to a hospital system (i.e. mapping and integrating processes for all the DRGs) might help managers to:

- Estimate the levels of activities and plan hospital resources in the medium-long term.
- Support cost accounting, particularly in an Activity Based Costing perspective.
- Take decisions accordingly to the different “profitability levels” of various subsets of patients.

Moreover, the approach can be also expanded to wider care organizations, at district or regional level. This would be helpful for regional or national authorities, since it might support demand aggregation and integrated resource planning, then system efficiency.

This work also shows some limitations for the proposed methodology. The main critical issue is the possible concurrent coexistence of two or more diseases (DRGs). These cases of comorbidity could alter the final analysis and require more accurate/ad hoc/time consuming pre-processing activities during the log preparation or log inspection phases. Furthermore, the case study concerns a limited number of cases and a single DRG; this condition can clearly affect the meaningfulness and generalizability of the investigation.

However, results gained in this preliminary study are mostly positive and encourage us to extend the research to a wider set of DRGs to support hospital process management. In future, we plan to assess a new case study in collaboration with an Italian hospital. This could allow gaining a more in-depth test of the method and integrating the process structure of n-DRGs in order to really support a hospital resource planning system.

REFERENCES

Anyanwu, K., Sheth, A. P., Cardoso, J., Miller, J. A., & Kochut, K. J. (2003). Healthcare enkmterprise process development and integration.

Aronsson, H., Abrahamsson, M., & Spens, K. (2011). Developing lean and agile health care supply chains. Supply Chain Management: *An International Journal*, 16(3), 176-183.

Bose, R. J. C., & van der Aalst, W. M. (2011). Analysis of Patient Treatment Procedures. In *Business Process Management Workshops* (1) (pp. 165-166).

Bozkaya, M., Gabriels, J., & Werf, J. M. E. M. (2009). Process diagnostics: a method based on process mining. In Information, Process, and Knowledge Management, 2009. eKNOW'09. International Conference on (pp. 22-27). *IEEE*.

Brailsford, S. C., Lattimer, V. A., Tarnaras, P., & Turnbull, J. C. (2004). Emergency and on-demand health care: modelling a large complex system. Journal of the Operational Research Society, 55(1), 34-42.

Bretthauer, K. M., & Cote, M. J. (1998). A model for planning resource requirements in health care organizations. Decision Sciences, 29(1), 243.

Busse, R., Geissler, A., & Quentin, W. (2011). Diagnosis-Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals. McGraw-Hill Education (UK).

Caron, F., Vanthienen, J., Vanhaecht, K., Van Limbergen, E., De Weerdt, J., & Baesens, B. (2014). Monitoring care processes in the gynecologic oncology department. Computers in biology and medicine, 44, 88-96.

Chau, P. Y., & Hu, P. J. H. (2001). Information technology acceptance by individual professionals: A model comparison approach. Decision Sciences, 32, 699-719.

De Weerdt, J., Caron, F., Vanthienen, J., & Baesens, B. (2013). Getting a grasp on clinical pathway data: an approach based on process mining. In Emerging Trends in Knowledge Discovery and Data Mining (pp. 22-35). Springer Berlin Heidelberg.

Delias, P., Doumpos, M., Grigoroudis, E., Manolitzas, P., & Matsatsinis, N. (2015). Supporting healthcare management decisions via robust clustering of event logs. Knowledge-Based Systems, 84, 203-213.

Dobrzykowski, D., Deilami, V. S., Hong, P., & Kim, S. C. (2014). A structured analysis of operations and supply chain management research in healthcare (1982–2011). *International Journal of Production Economics*, 147, 514-530.

Fetter, R. B., & Freeman, J. L. (1986). Diagnosis related groups: product line management within hospitals. Academy of Management Review, 11(1), 41-54.

Haraden, C., & Resar, R. (2004). Patient flow in hospitals: understanding and controlling it better. Frontiers of health services management, 20(4), 3.

Huang, Z., Lu, X., & Duan, H. (2012). On mining clinical pathway patterns from medical behaviors. Artificial intelligence in medicine, 56(1), 35-50

Jebali, A., Alouane, A. B. H., & Ladet, P. (2006). Operating rooms scheduling. *International Journal of Production Economics*, 99(1), 52-62.

Kaymak, U., Mans, R., van de Steeg, T., & Dierks, M. (2012, October). On process mining in health care. In Systems, Man, and Cybernetics (SMC), 2012 *IEEE International Conference on (pp. 1859-1864). IEEE*.

Lee, S. M., Lee, D., & Schniederjans, M. J. (2011). Supply chain innovation and organizational performance in the healthcare industry. International Journal of Operations & Production Management, 31(11), 1193-1214.

Lenz, R., & Reichert, M. (2007). IT support for healthcare processes–premises, challenges, perspectives. Data & Knowledge Engineering, 61(1), 39-58.

Li, L. X., Benton, W. C., & Leong, G. K. (2002). The impact of strategic operations management decisions on community hospital performance. *Journal of Operations Management,* 20(4), 389-408.

Mans, R., Schonenberg, H., Leonardi, G., Panzarasa, S., Cavallini, A., Quaglini, S., & van der Aalst, W. (2008). Process mining techniques: an application to stroke care. Studies in health technology and informatics, 136, 573.

Mans, R. S., Schonenberg, M. H., Song, M., van der Aalst, W. M., & Bakker, P. J. (2009). Application of process mining in healthcare–a case study in a dutch hospital (pp. 425-438). Springer Berlin Heidelberg.

Mans, R. S., van der Aalst, W. M., Vanwersch, R. J., & Moleman, A. J. (2013). Process mining in healthcare: Data challenges when answering frequently posed questions. In Process Support and Knowledge Representation in Health Care (pp. 140-153). *Springer Berlin Heidelberg*

Montani, S., Leonardi, G., Quaglini, S., Cavallini, A., & Micieli, G. (2014). Improving structural medical process comparison by exploiting domain knowledge and mined information. Artificial intelligence in medicine, 62(1), 33-45.

Partington, A., Wynn, M., Suriadi, S., Ouyang, C., & Karnon, J. (2015). Process mining for clinical processes: a comparative analysis of four Australian hospitals. ACM Transactions on Management Information Systems (TMIS), 5(4), 19

Rebuge, Á., & Ferreira, D. R. (2012). Business process analysis in healthcare environments: A methodology based on process mining. Information Systems, 37(2), 99-116.

Rovani, M., Maggi, F.M., De Leoni, M., Van Der Aalst, W. (2015). Declarative process mining in healthcare. Expert Systems with Applications, 42 (23), 9236-9251.

Stock, G. N., McFadden, K. L., & Gowen, C. R. (2007). Organizational culture, critical success factors, and the reduction of hospital errors. International Journal of Production Economics, 106(2), 368-392.

Tzeng, S. F., Chen, W. H., & Pai, F. Y. (2008). Evaluating the business value of RFID: Evidence from five case studies. *International Journal of Production Economics*, 112(2), 601-613.

Vissers, J. M. (1998). Patient flow-based allocation of inpatient resources: a case study. European Journal of Operational Research, 105(2), 356-370.

Van der Aalst, W. M., Reijers, H. A., Weijters, A. J., van Dongen, B. F., De Medeiros, A. A., Song, M., & Verbeek, H. M. W. (2007). Business process mining: An industrial application. Information Systems, 32(5), 713-732.

Van Der Aalst, W. (2011). Process mining: discovery, conformance and enhancement of business processes. Springer Science & Business Media.

Van Der Aalst, W., et Al. (2012). Process mining manifesto. In Business process management workshops (pp. 169-194). *Springer Berlin Heidelberg*.