Extracting Customer Traces from CRMS: From Software to Process Models

Article · January 2019

DOI: 10.1016/j.promfg.2019.02.261

2 authors:

Cristina-Claudia Osman
Babeş-Bolyai University

Ana-Maria Ghiran

Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 32 (2019) 619–626

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

The 12th International Conference Interdisciplinarity in Engineering

Extracting Customer Traces from CRMS: From Software to Process Models

Cristina-Claudia Osman[a,]*, Ana-Maria Ghiran[a]

[a]*Business Informatics Research Center, 58-60 T. Mihali Street, Cluj-Napoca, 400591,Romania*

Abstract

Digitization has subjected nowadays business models to dramatic changes. This commitment to innovation forced even the giants of technology to shift and adapt their own business models (e.g. Google developed a rival social network for Facebook, and a cloud file-storage in order to propose an inexpensive version of Dropbox or One Drive, Microsoft developed an online version of Microsoft Office). Organization's agility must be embedded in companies' strategies to survive in a market where everything changes with a tremendous quickness. Customer journey mapping, embedded with business process management systems, can assure an improved digital customer experience. Design thinking, a methodology used by several high-tech companies like Apple and IBM, helps the designers to better understand customers' needs in order to provide desirable solutions. This paper proposes a new approach of generating customer traces from information systems like CRMS to improve customer experience. A germane fact is that companies should use state of the art methods and techniques to better understand customers' needs and Process Mining algorithms are among the most suitable solutions for this type of analysis.

Keywords: Customer Relationship Management Systems; Customer Journey Map; Process Mining, Design Thinking; Business Process Management

1. Introduction

The tremendous development of digital experience strategy from the last decade brought several changes in companies' need to understand customers' expectations. Some of the authors consider digital experience as not

* Corresponding author. Tel.: +040264405300.
E-mail address: cristina.osman@econ.ubbcluj.ro

2351-9789
10.1016/j.promfg.2019.02.261

being an IT-driven initiative, but a customer-needs driven one [1]. However, this strategy is based on IT tools, methods and methodologies (e.g. Design Thinking, Business Process Management (BPM), Big Data Analytics, Machine Learning, Artificial Intelligence, Internet of Things (IoT) etc.) used in an optimal manner.

The integration of several channels as web, mobile or social in designing Customer Journey Map (CJM) is a really challenge for nowadays companies as the major issue is to employ disparate information systems in company-clients interaction. On the other hand, the companies' departments have known a process-centric approach of the business processes within the company. Moreover, employees from marketing departments have to include the channels mentioned above in Customer Relationship Management Systems (CRMSs). Thus, receiving inconsistent answers for the same issue from different employees that work in the same company (for example call center officer and physical store employee) may create customer frustrations or, even worse, losing them. The step of identifying new trends in companies' portfolio is a mandatory task in maximizing opportunities for cross-selling and up-selling. Furthermore, a recent study from 2017 [2] realized an analysis of more than 3100 companies (100 of them being leading companies in their domain) from all over the world regarding their position on digital transformation. These companies belong to different industries like manufacturing, banking, consumer goods or public sector. The study shows that 96% of the leading companies consider digital transformation as a main business goal, compared with 61% of other companies. The same study indicated increased investments in Big Data Analytics technologies within leading companies (94%) compared with 60% of other companies.

Next, we will summarize the ground covered in this paper. Second section gives a brief overview of related work and problem statement. In this section data mining approaches applied in Customer Relationship Management System are presented. Then, a short introduction into Design Thinking is made, followed by a brief description of Business Process Management (BPM) and Process Mining (PM) fields. Second section also presents the general approach to extract customer behavior. In the third section a case study is introduced and the event log generated by the information system used by an Italian software company is presented. The discussion continues in the fourth section where the generated process models are explained and analyzed. Both, control-flow and resource perspectives are investigated to provide a more comprehensive picture of the reality caught by the event log. Our conclusions regarding the current research are drawn in the final section.

2. Background and problem statement

The approach described in this paper proposes the use of Process Mining algorithms in detection of customer behavior aiming to extract customer traces from the interaction of clients with Customer Support department. Next we will describe the existing data mining approaches in Customer Relationship Management Systems analysis. In order to introduce Process mining field, we briefly present Design Thinking and Business Process Management domains.

2.1. Customer Relationship Management Systems (CRMSs)– data mining approaches

CRMS plays an important role in organizational agility. This concept measures the ability of companies to respond to market fluctuations. Ever since 1998 three success factors were introduced to depict organizational agility: reading the market, mobilizing rapid response and embedding organizational learning [3]. On the other hand, operational agility refers to ability of companies' business processes to run as fast and as accurate as possible to provide optimal solutions on a competitive market by analyzing opportunities [4]. IT is considered as being an enabler of operational agility [5].

The literature provides a plethora of approaches of data mining techniques applied on CRMS including associations [6,7], classification [8,9], clustering [7,10,11], regression [12], predictive analysis [9,10,12,13], sequence discovery [12] or visualization [9]. Although all these approaches use data mining techniques, none of them provide a process centric approach.

2.2. Design Thinking (DT)

Design Thinking (DT) has known an increased popularity in the last ten years. This can be considered as an agile method that helps companies to easily adapt to the new changes imposed by digitization or other factors. Initially, it was defined as the cognitive process of designers [14, 15]. Currently, DT is identified as an exciting new paradigm for dealing with problems in many professions, most notably Information Technology (IT) [16], Engineering [17], Business [18], Education [19, 20], or Medicine [21]. DT proposes a more visual overview of company's processes, integrates emotions and feelings into company processes, allocates time for failure part of a process and not at least, is based on collaboration within company's actors and interactors. Next, we will focus on DT involvement into business processes. This emerging multidisciplinary agile field is human-centered and it was depicted as having several steps depending on the model used.

DT model 3I [22] model depicted by IDEO in 2001 consists of 3 elementary steps: Inspiration, Ideation, respectively Implementation. One year later, IDEA developed another model: HCD Model [23]. Its initials stand from the basic steps: Hear, Create and Deliver. In 2015, a Field Guide for this model was launched (Ideo.org, The Field Guide To Human-Centred Design, http://www.designkit.org//resources/1). The drawback of these models is that they do not depict the whole design process (for example testing step is omitted). Another model came from Hasso-Plater-Institute [24]. It consists of 6 phases: Understand, Observe, Define, Ideate, Prototype and Test. Double Diamond Model (Discover, Define, Develop, Deliver) was proposed in 2005 by Design Council [25]. This model has been implemented in giant companies like Microsoft, Yahoo or Starbucks. An extended case study is presented in [25]. Both last two mentioned models before incorporated the testing phase. An iterative process of DT is proposed by [26]: Service Design Thinking Model. It consists of 4 steps: Exploration, Creation, Reflection and Implementation. The last step also embeds testing and improvement activities.

Regardless of the model used, the focus of DT is represented by users. Brainstorming is a DT method. This method is a very common method used even from 1939 for creative problem-solving. Another DT method that has the basis in cinematography is storyboard [27]. Basically, use cases are depicted via a series of visual representations which are put together in natural language. Post It notes are used in order to represent the steps depicted in the analysed process. Storytelling [28] is based on natural language and illustrations. It has a sequence and it also has a narrator who is telling the story. Managers may use this tool in resolving conflicts or facing challenges. Moreover, this tool may be used in creating stronger relationships with the customers.

2.3. Business Process Management (BPM)

Business Process Management (BPM) refers to the methods, techniques and tools that allow the design, enactment and management of business processes [29]. It also includes Business Process Modelling, an approach that offers conceptual representations of business processes. Davenport was one of the first authors that defined a business process as being a set of activities that interact to produce a business outcome [30].

Solutions of business approaches like Business Process Management (BPM) offer a holistic process-centric overview on how an organization works. One of the aims of BPM is to organize, improve business processes and to automatize them.

2.4. Process Mining (PM)

Process Mining is a blended approach using concepts from both Process Modelling and Machine Learning [31]. It is quite different than Business Intelligence (BI) and Data Mining as it is process oriented. The main open source tool that supports hundreds of plugins is ProM Framework [32]. One commercial Process Mining tool is Disco [33]. The latest one has a friendlier UI than ProM and supports the main ProM functionalities. Both on them require XES (eXtensible Event Stream) files [34] as inputs for process models extraction. Although CSV files are also accepted if they are converted into XES files a priori their use.

There is a link between Process Mining and DT. The first one is also used in *Story Mining* [35]. It embeds Process Mining in Storytelling. Initially, a tool called *TellStories* was developed in order to transform tacit knowledge from stories into explicit knowledge [35, 36]. Then, this was converted into *Story Mining* for process model extraction using business people stories as input. This is a text mining approach where a business process is built and, finally, generated models are reviewed by participants.

2.5. General approach

Operational agility is given preponderantly by IT technology. Using appropriate information systems and technologies can monitor and improve companies' KPIs (Key Performance Indicators). Moreover, every information system generates enormous quantities of data that are used in reaching different company's goals. For example, real time customer behavior models may help to target the customers on the right time, therefore, providing a customer tailored experience.

Process mining algorithms provide one of the best solutions to extract customer traces from event logs generated by Customer Relationship Management Systems. This paper aims to provide a new approach in business process analysis and improvement by using Process Mining methods and techniques.

Starting with a CSV/XES file we propose a modern approach using several Process Mining algorithms in order to generate business process models. We preponderantly focus on the order of activities (control-flow perspective) as the most part of ProM plugins propose diagrammatic models in terms of Petri Nets, Causal Nets, Fuzzy Nets, Heuristic Nets or BPMN diagrams (see Fig. 1). Our analysis went further by involving other process perspectives, namely resource perspective. On this matter, we generated a Social Network depicting work handover between employees. Each resulted process model highlights customer behavior extracted from the event log.

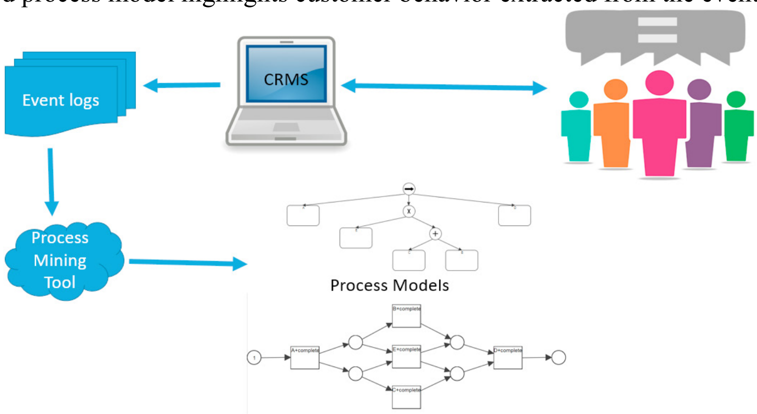

Fig. 1. General approach

3. Log analysis

This paper describes a process mining use-case on the event logs generated by the information system used by an Italian software company [37]. The data set consists of 21348 events belonging to 4580 cases spread from 13th of January 2010 until 3rd of January 2014. Although the event log depicts 4580 cases, the number of customers who interacted with help desk is 397. At a first sight the process seems to be a simple one, but by applying several process mining algorithms the process complexity is revealed.

The analysis performed offers a comprehensive understanding of the process model hidden behind the event log. Our findings are based on two main process model perspectives: control-flow perspective and resource perspective. Mapping these perspectives to process mining terms we discuss about process *("How?")* and organizational (*"Who?"*) perspectives [31]. For all events the state is considered to be *complete*. The main activities extracted from event log are: *Insert ticket, Assign seriousness, Take in charge ticket, Resolve ticket, Schedule intervention, Require upgrade, Create SW anomaly, Resolve SW anomaly* (see table below).

The excerpt consists of 3 traces belonging to the event log and each trace has between 4 and 5 events, while the number of traces belonging to the entire event log is 4580 and a trace has at least 2 events and at most 15 events. The shortest case lasted approximatively 30 days, while the longest took around 60 days.

Table 1. Excerpt from event log

CaseID	Activity	Resource	Complete Timestamp	customer	product
Case 1	Assign seriousness	Value 1	2012/10/09 14:50:17.000	Value 1	Value 1
Case 1	Take in charge ticket	Value 1	2012/10/09 14:51:01.000	Value 1	Value 1
Case 1	Take in charge ticket	Value 2	2012/10/12 15:02:56.000	Value 1	Value 1
Case 1	Resolve ticket	Value 1	2012/10/25 11:54:26.000	Value 1	Value 1
Case 1	Closed	Value 3	2012/11/09 12:54:39.000	Value 1	Value 1
Case 2	Assign seriousness	Value 4	2012/04/03 08:55:38.000	Value 2	Value 2
Case 2	Take in charge ticket	Value 4	2012/04/03 08:55:53.000	Value 2	Value 2
Case 2	Resolve ticket	Value 4	2012/04/05 09:15:52.000	Value 2	Value 2
Case 2	Closed	Value 5	2012/05/19 09:00:28.000	Value 2	Value 2
Case 3	Assign seriousness	Value 6	2010/10/29 10:14:06.000	Value 3	Value 3
Case 3	Take in charge ticket	Value 7	2010/11/03 16:16:11.000	Value 3	Value 3
Case 3	Resolve ticket	Value 7	2010/11/03 16:21:17.000	Value 3	Value 3
Case 3	Closed	Value 5	2010/11/30 11:20:18.000	Value 3	Value 3

4. Process analysis

The overview of our approach was depicted in sub-section 2.5. Basically, Help desk employees record all customer interactions with the CRMS. Then, the event log is extracted and by using process mining algorithms different process models are generated and analyzed.

4.1. Control-flow perspective

The type of Process Mining used in our analysis is play-in, in other words process discovery. In order to depict control-flow perspective we employed several ProM algorithms having different outputs. As the event log has multiple start and end events we former customized the event log by adding artificial start and end events [38]. This step was mandatory as several Process Mining algorithms ask for unique start/end events. Even though the event log extends to 30508 events (the number of events of each trace increased by 2), this artifice does not affect the results of our analysis. First, we applied an improved version of first Process Mining algorithm, Alpha Miner, namely Alpha++ [39] whose output is a Petri Net.

The Petri Net shows a series of repeating activities and the flow of the process is difficult to be followed. In order to depict the most frequent behavior observed in the event log we applied *Filter in High-Frequency Traces (Single log)* plug-in and the number of traces decreased to 3875, while the number of events decreased to 25192. During the analyzed period of time, from 13th of January 2010 until 3rd of January 2014, 395 clients asked for software anomalies solutions for 21 products provided by the company. Therefore, the model changed into a more comprehensible and simpler one (see Fig. 2). The analysis performed from now on refers to this filtered event log.

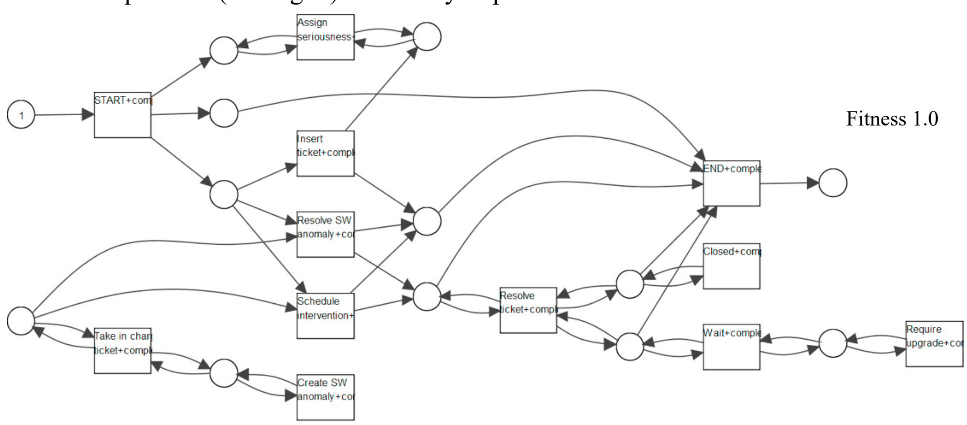

Fig. 2. Petri Net resulted after filtering most frequent behavior

Next, we applied other discovery process mining algorithms to generate the control-flow perspective. The BPMN diagram discovered by using *BPMN Miner* plugin and the Process tree extracted using *Mine a process Tree with ETMd* [40] can be seen in figure below.

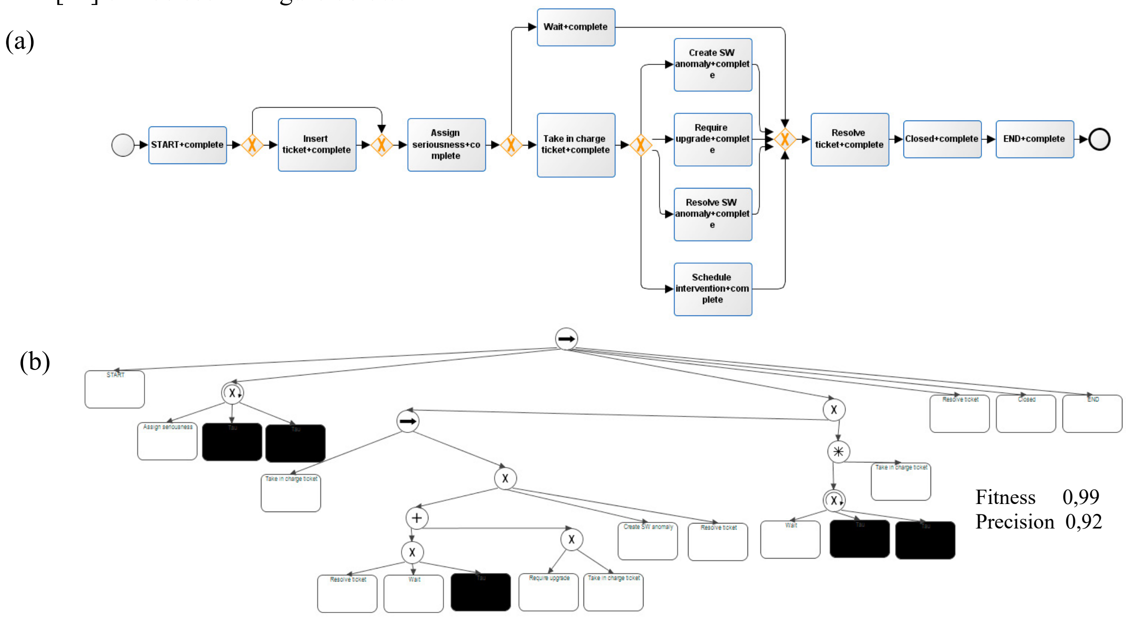

Fig. 3. (a) BPMN diagram; (b) Process tree

Other process models may be discovered using *Interactive Data-aware Heuristic Miner* [41]: Directly-follows graph, Dependency graph and Causal Net. The Directly-follows graph shows that most of cases starts with *Assign seriousness* activity (3791), but there also are 84 cases that start with *Insert ticket* activity. According to the event log, there are 4 degrees of seriousness and a case, during its lifecycle may take multiple degrees of seriousness. *Assign seriousness* activity is executed 4166 times, out of which 291 times is executed in the same case. Then, in the majority of cases *Take in charge ticket* activity follows (4219, out of which 65 times is repeated in the same case). The next activity to be executed is *Resolve ticket* (4067 times), but there are also few cases when this activity is executed without executing *Take in charge ticket* (4 times). Finally, the case is closed and the process ends. The flow described previously is also depicted by BPMN diagram and Process Tree from above.

The correctness of the resulted process models is validated using indicators like fitness and precision (where available). For example, for Process Tree evaluation, we used *Compute projected fitness and precision (log and process tree)* plugin.

4.2. Resource perspective

Although most of Process Mining algorithms focus on control-flow perspective, there also exist plug-ins that emphasize resource or data perspectives. One of the plugins focused on resource perspective is *Mine for a Handover-of-Work Social Network* [42]. The aim of this plugin is to provide and evaluate the associations between departments and/or employees. In other words, the social network provides the communication flow between organization's resources. The figure below shows work transfer from a resource to other and resources' interaction. There are 21 resources involved into the 3875 analyzed cases, plus the artificial resources assigned for start and end events. There is not a certain flow between work handover as customer support officers propose solutions to software anomalies described by clients and send the ticket to the employees in charge with the specified anomaly. The network shows that the most engaged resource in the activities depicted by the event log is *Value 2* (3385 activities), while the resource executing less activities is *Value 20* (14 activities).

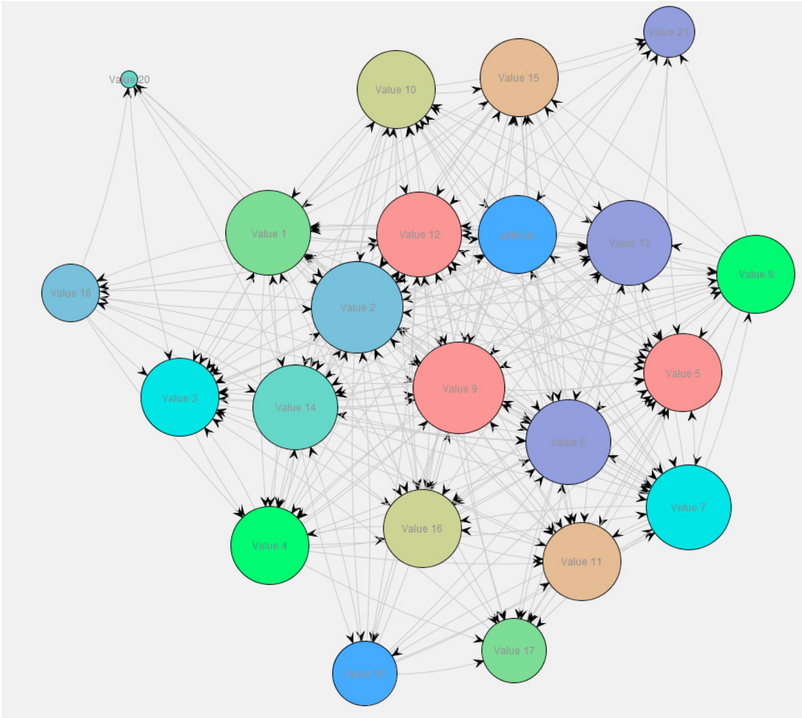

Fig. 4. Social network – resource perspective

5. Conclusions and future work

Considering the increasing amounts of existing digital information within a company, this paper emphasizes the value of using new methods and techniques for the purpose of supporting business processes without having any information about the analyzed process, only the event log (or a CSV file containing certain attributes). Then, different process mining algorithms have been applied in order to extract graphical visualizations. On this line, we evaluated diagrammatic models using different modelling languages (for example: Petri Nets, Causal Nets or BPMN diagrams). Fuzzy and Heuristics nets were not analyzed as they emphasize most frequent behavior and that kind of behavior was already embedded into the discovered Petri Net or BPMN diagram.

The approach proposed may be used together with Data Mining techniques reminded in sub-section 2.1 or other machine learning, artificial or business intelligence methods/techniques. A combination of process mining and classic business process reengineering also offers a promising solution in customer behavior analysis. The current research resumes to the use of Process Mining in customers traces extraction. These findings suggest the opportunities for future research to generate more diagrammatic models like Customer Journey Maps by involving several communication channels like e-mail, instant messages, company's website, etc. in order to better respond to clients' needs.

References

[1] M. Hinshaw, CMO by Adobe, 18 December 2012, http://www.cmo.com/opinion/articles/2012/12/18/digital_experience_t.html#gs.AdpuBY4, Accessed 10 March 2018.

[2] SAP and Oxford Economics, https://www.sap.com/dmc/exp/4-ways-leaders-set-themselves-apart/index.html?url_id=ctabutton-blog_DTES_4traits-Digital_Transformation_Executive_Study-sap.com-Study_Landing_Page-DTES-DTES, (2017), Accessed 12 March 2018.

[3] L. Dyer and R. A. Shafer, From human resource strategy to organizational effectiveness: Lessons from research on organizational agility, *CAHRS Working Paper Series (1998)* 1-34.

[4] V. Sambamurthy, A. Bharadwaj and V. Groover, Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms, *MIS Quarterly,* vol. 27, (2003) 237-263.

[5] P.-Y. Huang, T. H. Ouyang, S. L. Pan and T.-C. Chou, The role of IT in achieving operational agility: A case study of Haier, China, *International Journal of Information Management,* vol. 32, no. 3, (2012) 294-298.
[6] S. R. Ahmed, Applications of data mining in retail business, *Proceedings of Information Technology: Coding and Computing*, vol. 2, Las Vegas, (2004).
[7] W.-Y. Chiang, Applying data mining for online CRM marketing strategy: An empirical case of coffee shop industry in Taiwan, *British Food Journal,* vol. 120, no. 3, (2018) 665-675.
[8] G. S. Linoff and J. B. Berry, Data mining techniques: for marketing, sales, and customer relationship management, Indianapolis: John Wiley & Sons, (2011).
[9] G. Shmueli, P. Bruce, I. Yahan, N. R. Patel and K. J. J. Lichtendahl, Data Mining for Business Analytics: Concepts, Techniques, and Applications, John Wiley & Sons, (2017).
[10] J.-T. Wei, M.-C. Lee, H.-K. Chen and H.-H. Wu, Customer relationship management in the hairdressing industry: An application of data mining techniques, *Expert Systems with Applications,* vol. 40, no. 18, (2013) 7513-7518.
[11] C. Giraud-Carrier, O. Povel, Characterising data mining software, *Intelligent Data Analysis,* vol. 7, no. 3, (2003) 181-192.
[12] C. Rygielski, J.-C. Wang and D. C. Yen, Data mining techniques for customer relationship management, *Technology in society,* vol. 24, no. 4, (2002) 483-502.
[13] F. T. Bahari and M. S. Elayidom, An efficient CRM-Data Mining Framework for the prediction of Customer Behaviour, *Procedia computer science, 46,* (2015) 725-731.
[14] N. Cross, K. Dorst, N. Roozenburg. Research in Design Thinking: Proceedings of a Workshop Meeting Held at the Faculty of Industrial Design Engineering*, Delft University of Technology, the Netherlands, May 29-31, 1991*. Delft University Press, (1992).
[15] C. Eastman, W. Newstetter, M. McCracken, (Eds.). Design knowing and learning: Cognition in design education. Elsevier, (2001).
[16] F.P. Brooks Jr., The design of design: Essays from a computer scientist. Pearson Education, (2010).
[17] H.A. Simon, The Sciences of the Artificial, MIT Press, Cambridge, MA (1969).
[18] R.L. Martin,*The design of business: Why design thinking is the next competitive advantage*. Harvard Business Press. (2009).
[19] Koh J.H.L., Chai C.S., Wong B., Hong HY, Design Thinking and Education, *Design Thinking for Education*. Springer, Singapore, (2015), 1-15.
[20] A. Scheer, C. Noweski, C. Meinel, Transforming Constructivist Learning into Action: Design Thinking in education. *Design and Technology Education: an International Journal*, vol.17, no. 3, ISSN 1360-1431 (2012), 8-19.
[21] T.C. van de Grift, R. Kroeze, Design thinking as a tool for Interdisciplinary education in health care. *Academic Medicine*, *91*(9), (2016), 1234-1238.
[22] T. Brown, Change by Design. How Design Thinking transforms Organizations and inspires Innovation. New York: Harper Collins Publishers (2009).
[23] T. Brown, J. Wyatt, Design Thinking for Social Innovation. Stanford Social Innovation Review, (2010) 31-35.
[24] K. Thoring, R.M. Müller, Understanding the Creative Mechanisms of Design Thinking: An Evolutionary Approach, *Proceedings of the DESIRE'11 Conference Creativity and Innovation in Design.* ACM, Eindhoven, (2011) 137-144.
[25] Design Council, Eleven lessons: managing design in eleven global brands A study of the design process, http://www.designcouncil.org.uk/sites/default/files/asset/document/ElevenLessons_Design_Council%20(2).pdf, Accessed 14 March 2018.
[26] M. Strickdorn, J. Schneider, This is Service Design Thinking: Basics-Tools-Cases, Amsterdam: BIS Publishers, (2010).
[27] G. Curtis, Storyboards and sketch prototypes for rapid interface visualization, *ACM SIGCHI'90 Tutorial,* (1990).
[28] D.M. Boje, The storytelling organization: A study of story performance in an office-supply firm, *Administrative science quarterly*, (1991) 106-126.
[29] W.M. Van Der Aalst, A.H. Ter Hofstede, M. Weske, Business process management: A survey. In *International conference on business process management*, Springer, Berlin, Heidelberg, (2003)1-12.
[30] T.H. Davenport, Process Innovation, Harvard Business School Press, 1993.
[31] W.M.P. van der Aalst, Process Mining: Data Science in Action, Springer Verlag (2016).
[32] H.M.W., Verbeek, J.C.A.M. Buijs, B.F. van Dongen, W.M.P.van der Aalst, ProM 6: The process mining toolkit, *Proceedings of BPM Demonstration Track 2010,* vol. 615, (2010) 34–39.
[33] C.W. Günther, A. Rozinat, Disco: Discover Your Processes. *BPM (Demos)*, 940 (2012) 40-44.
[34] C. Gűnther, E. Verbeek, XES Standard Definition, Technische Universiteit Eindhoven. (2014).
[35] R. Perret, M.R. Borges, F.M. Santoro, "Applying group storytelling in knowledge management", in Proc. of International Conference on Collaboration and Technology, (2004), 34-41
[36] J.C. de AR Goncalves, F.M. Santoro, F.A. Baiao, Business process mining from group stories, I*EEE of The 13th International Conference on Computer Supported Cooperative Work in Design (CSCWD 2009)*, (2009) 161-166.
[37] M. Polato, Dataset belonging to the help desk log of an Italian Company. University of Padova, Padova (2017), Dataset: https://doi.org/10.4121/uuid:0c60edf1-6f83-4e75-9367-4c63b3e9d5bb, Accessed 12 April 2018.
[38] J. Claes, Add artificial events plugin, http://www.janclaes.info/post.php?post=addartificialevents, Accessed 12 April 2018.
[39] L. Wen, W.M. van der Aalst, J. Wang, J. Sun, Mining process models with non-free-choice constructs, *Data Mining and Knowledge Discovery*, 15(2), (2007) 145-180.
[40] J.C.A.M. Buijs, Flexible evolutionary algorithms for mining structured process models Eindhoven, Technische Universiteit Eindhoven, PhD thesis, (2014).
[41] F. Mannhardt, M.de Leoni, H.A. Reijers, W.M.P. van der Aalst, Data-Driven Process Discovery - Revealing Conditional Infrequent Behavior from Event Logs. *Dubois E., Pohl K. (eds) Advanced Information Systems Engineering. CAiSE 2017. Lecture Notes in Computer Science*, vol. 10253. Springer, Cham (2017) 545-560.
[42] W. van der Aalst, M. Song, Mining social networks: uncovering interaction patterns in business processes, *Proceedings of International Conference on Business Process Management (BPM 2004)*, Berlin, Germany, (2004) 244-260.