Process Mining Analytics For Industry 4.0 With Graph Signal Processing

Georgios Drakopoulos[1] [a], Eleanna Kafeza[2] [b], Phivos Mylonas[1] [c] and Spyros Sioutas[3]

[1] *Department of Informatics, Ionian University, Greece*

[2] *College of Technology and Innovation, Zayed University, UAE*

[3] *Computer Engineering and Informatics Department, University of Patras, Greece*

{*c16drak, fmylonas*}*@ionio.gr, eleana.kafeza@zu.ac.ae, sioutas@ceid.upatras.gr*

Keywords: Process Mining, Industry 4.0, Graph Signal Processing, Graph Mining, Multilayer Graphs, PM4Py, Neo4j

Abstract: Process mining is the art and science of (semi)automatically generating business processes from a large number of logs coming from potentially heterogeneous systems. With the recent advent of Industry 4.0 analog enterprise environments such as floor shops and long supply chains are bound to full digitization. In this context interest in process mining has been invigorated. Multilayer graphs constitute a broad class of combinatorial objects for representing, among others, business processes in a natural and intuitive way. Specifically the concepts of state and transition, central to the majority of existing approaches, are inherent in these graphs and coupled with both semantics and graph signal processing. In this work a model for representing business processes with multilayer graphs along with related analytics based on information theory are proposed. As a proof of concept, the latter have been applied to large synthetic datasets of increasing complexity and with real world properties, as determined by the recent process mining scientific literature, with encouraging results.

1 INTRODUCTION

Recently the theory and practice of manufacturing underwent a series of radical evolutionary transformations after a long period covering Antiquity and the Middle Ages where humans, whether slaves or highly paid technicians and professionals, animals, and simple machines such as Heron's steam engine or *Aeolipile* were the primary means of production. The roots of each major milestone can be respectively traced in the following historical periods:

- The Victorian era[1] in the wake of a major scientific wave saw the massive transition to hydraulic power for a broad spectrum of applictions. The uncontested colophon of that era was the development of steam engine.
- Between the French-Prussian war of 1871 up to the start of First World War in 1914 heavy emphasis was placed on developing extensive networks, whether physical, such as railroads and post offices, or telecommunication ones, like the telegraph and local telephone systems. These networks prompted the construction of massive assembly lines and supply chains.
- Finally, after the end of the Second World War in 1945 and until the beginning of the 21st century focus shifted on digitization and miniaturization, eventually giving rise to microelectronics and digital computers. The main paradigm shift here was the reinforcement not only of the human body but of the brain as well.

Currently Industry 4.0, originally a set of specifications compiled in 2011 by the *Bundesregierung*, namely the federal German government, aims to transform manufacturing landscape by introducing the use of sensors, artificial intelligence (AI), and Internet of Things (IoT) technology in order to increase productivity, cybersecurity, and personnel safety. In this way diverse operational objectives from various scopes can be achieved even under quite adverse circumstances. At the same time human-to-machine and machine-to-machine will become seamless and more efficient through wearable electronics for humans and reconfigurable sensor arrays for machines.

In this digital enterprise setting the role of process mining is becoming increasingly more important as large event logs are created by a multitude of commer-

[a] https://orcid.org/0000-0002-0975-1877
[b] https://orcid.org/0000-0001-9565-2375
[c] https://orcid.org/0000-0002-6916-3129

[1] The technology of that era and the promises it brought about human life led to the *steampunk* subculture and literary genre.

cial business applications and big process graphs are generated for various production purposes. Given that data volume and its high generation rate, errors are almost bound to happen. They are frequently manifested in the absence or addition of spurious vertices or edges at the process graphs. However, a more insidious result is the changes to process graph semantics as errors are more subtle and can be thus propagated undetected in the process graph.

The primary research objective of this conference paper is the development of edge, path, and triangle similarity metrics for evaluating the difference between any template process graph and a corresponding variant one. Said difference is evaluated with a metric enriched with semantics represented as edge labels which is derived from information theory. This work differentiates from previous approaches in two ways, namely the use of multilayer graphs in order to represent long Industry 4.0 processes and the use of the emerging field of graph signal processing (GSP).

The remaning of this work is structured as follows. In Section 2 the recent scientific literature pertaining to process mining and multilayher graphs is briefly reviewed. Section 3 contains the formal definition of as well as some intuituion about mulilayer graphs. The proposed methodology is described in detail in section 4. The results of applying it to synthetic process benchmark graphs of increasing complexity are given in section 5. Section 6 recapitulates the main results and outlines future research directions. Technical acronyms are defined the first time they are encountered in the text. In definitions parameters are given after formal arguments following a semicolon. Finally, table 1 summarizes the notation of this work.

Table 1: Notation of this conference paper.

Symbol	Meaning	First
$\triangleq$	Equality by definition	Eq.(1)
$\{s_1,\ldots,s_n\}$	Set consisting of $s_1,\ldots,s_n$	Eq.(2)
$(t_1,\ldots,t_n)$	Tuple with $t_1,\ldots,t_n$	Eq.(1)
$\lvert S\rvert$	Set or tuple cardinality	Eq.(3)
$S_1 \setminus S_2$	Asymmetric set difference	Eq.(3)
$[e_1,\ldots,e_p]$	Path of edges $e_1,\ldots,e_p$	Eq.(9)
$\mathcal{H}(\cdot)$	Harmonic mean	Eq.(5)
$\mathrm{diag}[\cdot]$	Diagonal matrix	Eq.(12)

2 PREVIOUS WORK

Industry 4.0 is a major milestone in the history of industrial organization and production (da Rosa Righi et al., 2020). It aims to the full digitization of industrial production through a wide array of sensors installed in machinery and in wearable electronics for human operators as well as through delegation of minor, mundane, or dangerous tasks to computer-operated equipment (Bigliardi et al., 2020). Various sensor architectures based on the Industry 4.0 requirements have been proposed and compared in (Bajic et al., 2020). Operational criteria and considerations for the industrial equipment are examined in (Culot et al., 2020). The connections between Industry 4.0 and circular economy are explored in (Rajput and Singh, 2019). The principal question of sustainability is put in (Bai et al., 2020). An extensive review of the relevant bibliography about Industry 4.0 is given in (Souza et al., 2020).

Process mining relies heavily on the parsing of automatically generated process logs in order to discover patterns, latent dependencies, and persistent anomalies (Mitsyuk et al., 2017; Reinkemeyer, 2020). The IEEE extensive event stream (XES) or IEEE standard 1849-2016 is a standard log file format designed for the explicit purpose of process mining proposed in (Acampora et al., 2017). Automated log mining is explained in (Egger et al., 2020). PM4py is a Python package for process mining complete with methods for pattern discovery and miners such as A and A^+ (Berti et al., 2019). Dealing with malformed or otherwise imperfect process logs is examined in (Suriadi et al., 2017). Context-aware process mining with the introduction of advanced graph mining is the topic of (Becker and Intoyoad, 2017). The role of process mining to auditing information systems is described in (Zerbino et al., 2018). Finally, among the various surveys covering the topic are (Lopes and Ferreira, 2019) and (Verenich et al., 2019).

Multilayer or multiplex graphs allow parallel edges between the same pairs of vertices (Caimo and Gollini, 2020; Halnaut et al., 2020). As with ordinary graphs massive graph mining for this class can take place with the help of graph analytics (Zhou and Cheung, 2019) including attribute engineering (Drakopoulos and Mylonas, 2020). Also multilayer graphs have been proposed as a scalable IoT model (Xie et al., 2020). Functional and structural aspects of brain circuits are combined to form multilayer graphs in (Mandke et al., 2018). Visualization techniques for multilayer graphs are explored in (McGee et al., 2019). Semi-supervised learning methods for this class of graphs are proposed in (Mercado et al., 2019). Multilayer graphs have been used for image segmentation (Wang et al., 2016), spectral graph clustering (Chen and Hero, 2017), fast graph transform mining (Drakopoulos et al., 2021). A versatile and space efficient data structure which can be used among others for process storage and additionally supports persistency is proposed in (Kontopoulos and Drakopoulos, 2014).

3 MULTILAYER GRAPHS

Informally speaking, the class of multilayer graphs represents graphs with multiple edge labels. The name comes from the fact when considering only a single given label, then an ordinary graph termed a *layer* results. Thus, a multilayer graph can be decomposed to various layers. The total activity in such a graph comes from the following interacting factors:

- Activity in each separate layer. This happens at the vertices and edges of the specific layer.
- Activity across layers. Typically this takes place at the vertices belonging to at least two layers.

The above imply that any extension of Metcalfe's law (Metcalfe, 2013) to multilayer graphs should take into account both these factors if the true graph value is to be determined. Possibly this entails a composite power law which will be a function of the overall average degree or the average degree of each layer.

Formally, the combinatorial structure of a multilayer graph is given by definition 1.

Definition 1 (Mutilayer graph)**.** *A multilayer graph is the ordered quadruple of equation (1).*

$$G \stackrel{\triangle}{=} (V, E, L, h) \tag{1}$$

In equation (1) the tuple elements are the following:

- *The vertex set V contains the vertices of the graph. In this context vertices represent special states, namely the beginning or the end of a process or important intermediate steps.*
- *The edge set $E \subseteq V \times V \times L$ contains the labeled edges of the graph. They indicate dependencies or the various connections between either process states or entire processes.*

4 PROPOSED METHODOLOGY

4.1 General Notes

In this section the proposed methodology based on the class of multilayer graphs will be described. First the way edge similarity is computed will be presented followed by applications to paths and triangles, two of the most common structural pattterns encountered in process mining graphs. Then the edge signal to noise ratio, a concept borrowed and adapted from the field of information theory, will be also presented.

At this point it is important to highlight that the theory developed here is based on the following underlying fundamental assumption.

Assumption 1 (Alignment assumption)**.** *The template and the variant process graphs are aligned.*

This is not a trivial observation since alignment is a major research topic in graph mining, ontology discovery, and in related fields.

Moreover, emphasis should be placed that the comparison metrics described in this section were explicitly designed for evaluating distances between the original process graph and the variant graph, explained respectively in definitions 2 and 3.

Definition 2 (Process graph)**.** *The process graph is the template describing in detail the desired process mining assumptions, approach, and operational characteristics of an organization.*

Definition 3 (Variant graph)**.** *The variant graph is the process mining graph constructed (semi)automatically from parsing process logs, equipment sensors, personnel reports, and any other technical means deployed in the field.*

Since the original process graph and any variant one deriving from it are aligned, each edge e in the latter has a unique counterpart e_0 in the former. Hence it makes perfect sense to refer in the text to the counterpart of e without any further clarification.

4.2 Label Noise

Since multilayer graphs allow multiple edges between the same pair of vertices, for comparison purposes as well as for notation simplification a group of labeled edges can be replaced with a single edge with a set, the *edge set*, containing the labels of the respective individual edges. In figure 1 is shown how various parallel labeled edges can be substituted with an equivalent label set. This step is crucial for developing the analytics presented in later sections.

Therefore, in a process graph for a given vertex pair a group of connecting edges $e_1, \ldots, e_n$ with corresponding labels $l_1, \ldots, l_n$ L is replaced by a single edge e with the edge set of equation (2):

$$L \stackrel{\triangle}{=} \{l_1, \ldots, l_n\} \tag{2}$$

The basic building block for assessing the similarity between process patterns is edge similarity. In order to evaluate the similarity between two edges, one from the process graph and one from the template graph, it suffices to compare the respective label sets. To this end the asymmetric Tversky index will be employed. The latter evaluates the divergence between

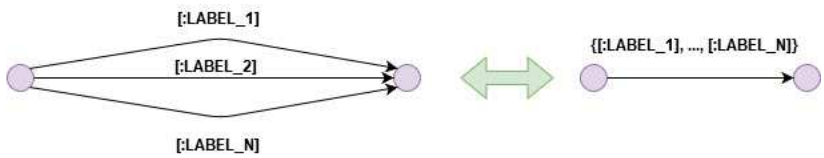

Figure 1: Construction of the edge label set. Source: Authors.

two sets T and V where the former is considered to be a template and the latter a variance thereof. Thus these two sets are by construction not interchangeable. This fundamental property is reflected in the index mathematical definition (Tversky, 1977):

$$\tau(T,V;\alpha_0,\beta_0) \triangleq \frac{|T \cap V|}{|T \cup V| + \alpha_0 |T \setminus V| + \beta_0 |V \setminus T|} \tag{3}$$

In equation (3) the parameters α_0 and β_0 denote respectively the weights for the number of elements present in T but absent in V and vice versa. Although their only real constraint is that they are non-negative, frequently their sum is normalized to one such that α_0 and β_0 become relative weights. This is further illustratred by typically selecting their values such that their ratio takes a predetermined and application-dependent value γ_0 as shown in equation (4):

$$\frac{\alpha_0}{\beta_0} = \gamma_0 \tag{4}$$

These changes to process graphs labels can be thought of as noise similar to that present in digital electronics-based wired (DEBW) telecommunications systems. However, the properties of the label noise are fundamenally different because of the following reasons:

- In contrast to DEBW systems where the primary source of noise is continuous, any changes to edge labels are discrete.
- In DEBW systems the noise is numerical in nature, where in process graphs the noise results in semantic errors.
- In DEBW systems noise comes from the electronics components located in the transmitter and the receiver or from the propagation medium, whereas changes to labels stem primarily from design or communication errors.

Thus, given the above it is clear that the additive white Gaussian noise (AWGN) model is not appropriate in this context and by extension neither is the Gaussian distribution a proper model for the label noise.

The signal to noise ratio (SNR) is a fundamental concept in information theory which serves in the development for metrics of signal distorion over telecommunication channels.

Definition 4 (Edge SNR)**.** *For a single edge of the variant process graph the SNR is defined as the logarithm of the ratio of to as shown in equation (5):*

$$s(e) \triangleq \ln\left(\frac{\tau(L,L_0)}{1-\tau(L,L_0)}\right) \tag{5}$$

In equation (5) $0 \leq \tau(L,L_0) \leq 1$

$$\log_a b = \frac{\log_c a}{\log_c b}, \qquad a,b,c \neq 0 \tag{6}$$

The numerical behavior of s with respect to $\tau(L,L_0)$ in equation (5) is degrading as label noise vanishes as shown in equation (7):

$$\frac{\partial s}{\partial \tau} = \frac{1}{\tau(1-\tau)} \tag{7}$$

$$\frac{\partial^2 s}{\partial \tau^2} = \frac{2\tau - 1}{\tau^2 (1-\tau)^2} \tag{8}$$

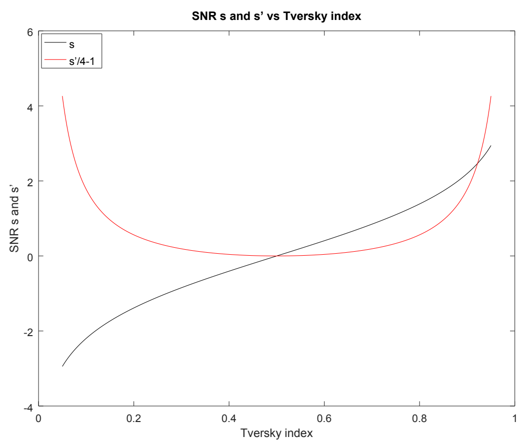

Figure 2: Label SNR vs Tversky index. Source: Authors.

4.3 Path Noise

Let p be a directed path in a process graph consisting of n of labeled edges as shown in equation (9):

$$p \triangleq [e_1, \ldots, e_n] \tag{9}$$

Definition 5 (Path SNR)**.**

$$s(p) \triangleq \frac{n}{\sum_{k=1}^{n} \frac{1}{s(e_k)}} \triangleq \mathcal{H}(s(e_1), \ldots, s(e_n)) \tag{10}$$

The harmonic mean of equation (10) has many algorithmic and numerical properties which make it ideal for this conference paper.

$$\nabla s(p) \triangleq \left[\frac{\partial s(p)}{\partial s(e_1)} \quad \frac{\partial s(p)}{\partial s(e_2)} \quad \ldots \quad \frac{\partial s(p)}{\partial s(e_n)} \right]^T$$
$$= n \left[s(e_1)^2 \quad \ldots \quad s(e_n)^2 \right]^T \tag{11}$$

$$\nabla^2 s(p) \triangleq \begin{bmatrix} \frac{\partial^2 s(p)}{\partial s(e_1)^2} & \cdots & \frac{\partial^2 s(p)}{\partial s(e_1)\partial s(e_n)} \\ \frac{\partial^2 s(p)}{\partial s(e_2)\partial s(e_1)} & \cdots & \frac{\partial^2 s(p)}{\partial s(e_2)\partial s(s_n)} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 s(p)}{\partial s(e_n)\partial s(e_1)} & \cdots & \frac{\partial^2 s(p)}{\partial s(e_n)^2} \end{bmatrix}$$
$$= 2n \operatorname{diag}\left[s(e_1), \ldots, s(e_n) \right] \tag{12}$$

4.4 Triangle Noise

Triangles are the simplest yet most fundamental community blocks in graphs as well as the first closed graph structural pattern. By extending the path SNR metric to any given triangle yields equation (13):

$$s(T) \triangleq \mathcal{H}(s(e_1), s(s_2), s(s_3)) \tag{13}$$

5 RESULTS

In this section the similarity metrics presented earlier are put to test. Synthetic datasets based on the following real world Industry 4.0 requirements were constructed. Specifically, the benchmarks will be graph datasets generated to have many of the process graph properties reported in the recent process mining scientific literature in works such as (Verenich et al., 2019) and (Acampora et al., 2017). These properties include:

- The number of vertices and edges as well as the number of labels.
- The average graph diameter as well as the effective 70
- The expected number of triangles.
- The expected path length and the associated variance.

Table 2 contains the synopses of template graphs used in this work. Each is a Kronecker graph coming from a generator graph of lower size. In order to create the variant graphs labels where either added or removed at random from edges of the template graph until an average SNR was met. For each template graph and for each SNR a thousand instances were created. The average values and the respective variances for each metric were recorded. Coding was done in Python 3.8 with the numpy and the scipy packages for analysis. Graphs were created and handled with the NetworkX package.

From the dataset synopses presented in table 2 it follows that they have an increasing level of complexity, implying that more complex datasets pose a bigger challenge for analytics designers.

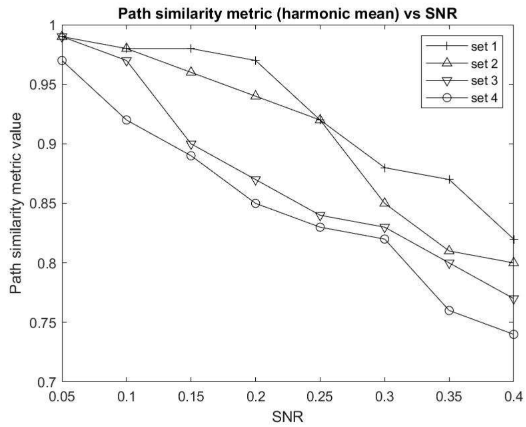

Figure 3: . Source: Authors.

6 CONCLUSIONS

This conference paper focuses on a process mining model for Industry 4.0 based on the class of multilayer graphs as well as on associated analytics. This class of graphs extends the ordinary ones by adding edge labels, essentially semantics based on the underlying process logs. This is appealing since edges can have properties depending on their role in the overall process and, moreover, edges denoting tasks executed in parallel along the same check points can be combined to a single one with a label set. As high degree task parallelism, typically due to multiple sensor readings, is a very common characteristic of an Industry 4.0 setting, edges with label sets of even a moderate size arise frequently. In turn, these sets can be the building blocks for a number of analytics for the distance between the process graph, namely the actual graph as mined from the various system

Table 2: Dataset properties.

Set 1	
Generator vertices	5
Generator edges	7
Template vertices	3125
Template edges	16807
Label set size	16
Labels per edge	6.53
Diameter	11
80% effective	7
90% effective	8
Number of triangles	625
Set 2	
Generator vertices	5
Generator edges	8
Template vertices	15625
Template edges	262144
Label set size	32
Labels per edge	11.67
Diameter	13
80% effective	9
90% effective	11
Number of triangles	33125
Set 3	
Generator vertices	7
Generator edges	13
Template vertices	16807
Template edges	371293
Label set size	48
Labels per edge	28.44
Diameter	15
80% effective	11
90% effective	13
Number of triangles	67617
Set 4	
Generator vertices	7
Generator edges	17
Template vertices	16807
Template edges	1419857
Label set size	64
Labels per edge	32.33
Diameter	16
80% effective	12
90% effective	15
Number of triangles	212881

and process logs, and the template graph, namely the blueprint process graph as derived by system designers. Analytics based on this distance metric include path and vertex similarity metrics as well as a modified clustering coefficient. Experiments conducted with synthetic datasets indicate that these analytics can discover errors in multilayer graphs while at the same time being algorithmically robust and numerically stable, given the large number of floating points operations required to derive the final result.

ACKNOWLEDGEMENTS

This conference paper was supported by the Research Incentive Fund (RIF) Grant R18087 provided by Zayed University, UAE.

REFERENCES

Acampora, G., Vitiello, A., Di Stefano, B., van der Aalst, W., Günther, C., and Verbeek, E. (2017). IEEE 1849: The XES standard. *IEEE Computational Intelligence Magazine*, 12(2):4–8.

Bai, C., Dallasega, P., Orzes, G., and Sarkis, J. (2020). Industry 4.0 technologies assessment: A sustainability perspective. *International Journal of Production Economics*, 229.

Bajic, B., Rikalovic, A., Suzic, N., and Piuri, V. (2020). Industry 4.0 implementation challenges and opportunities: A managerial perspective. *IEEE Systems Journal*.

Becker, T. and Intoyoad, W. (2017). Context aware process mining in logistics. *Procedia Cirp*, 63:557–562.

Berti, A., van Zelst, S. J., and van der Aalst, W. M. (2019). PM4Py Web Services: Easy development, integration and deployment of process mining features in any application stack. In *BPM (PhD/Demos)*, pages 174–178.

Bigliardi, B., Bottani, E., and Casella, G. (2020). Enabling technologies, application areas and impact of Industry 4.0: A bibliographic analysis. *Procedia Manufacturing*, 42:322–326.

Caimo, A. and Gollini, I. (2020). A multilayer exponential random graph modelling approach for weighted networks. *Computational Statistics & Data Analysis*, 142.

Chen, P.-Y. and Hero, A. O. (2017). Multilayer spectral graph clustering via convex layer aggregation: Theory and algorithms. *IEEE Transactions on Signal and Information Processing over Networks*, 3(3):553–567.

Culot, G., Nassimbeni, G., Orzes, G., and Sartor, M. (2020). Behind the definition of Industry 4.0: Analysis and open questions. *International Journal of Production Economics*, 226.

da Rosa Righi, R., Alberti, A. M., and Singh, M. (2020). *Blockchain Technology for Industry 4.0.* Springer.

Drakopoulos, G., Kafeza, E., Mylonas, P., and Iliadis, L. (2021). Transform-based graph topology similarity metrics. *NCAA*, 1(1).

Drakopoulos, G. and Mylonas, P. (2020). Evaluating graph resilience with tensor stack networks: A keras implementation. *NCAA*, 32(9):4161–4176.

Egger, A., ter Hofstede, A. H., Kratsch, W., Leemans, S. J., Röglinger, M., and Wynn, M. T. (2020). Bot log mining: Using logs from robotic process automation for process mining. In *International Conference on Conceptual Modeling*, pages 51–61. Springer.

Halnaut, A., Giot, R., Bourqui, R., and Auber, D. (2020). Deep dive into deep neural networks with flows. In *VISIGRAPP*, volume 3, pages 231–239.

Kontopoulos, S. and Drakopoulos, G. (2014). A space efficient scheme for graph representation. In *ICTAI*, pages 299–303. IEEE.

Lopes, I. F. and Ferreira, D. R. (2019). A survey of process mining competitions: The BPI challenges 2011–2018. In *International Conference on Business Process Management*, pages 263–274. Springer.

Mandke, K., Meier, J., Brookes, M. J., O'Dea, R. D., Van Mieghem, P., Stam, C. J., Hillebrand, A., and Tewarie, P. (2018). Comparing multilayer brain networks between groups: Introducing graph metrics and recommendations. *NeuroImage*, 166:371–384.

McGee, F., Ghoniem, M., Melançon, G., Otjacques, B., and Pinaud, B. (2019). The state of the art in multilayer network visualization. In *Computer Graphics Forum*, pages 125–149. Wiley Online Library.

Mercado, P., Tudisco, F., and Hein, M. (2019). Generalized matrix means for semi-supervised learning with multilayer graphs. *arXiv preprint arXiv:1910.13951*.

Metcalfe, B. (2013). Metcalfe's law after 40 years of Ethernet. *Computer*, 46(12):26–31.

Mitsyuk, A. A., Shugurov, I. S., Kalenkova, A. A., and van der Aalst, W. M. (2017). Generating event logs for high-level process models. *Simulation Modelling Practice and Theory*, 74:1–16.

Rajput, S. and Singh, S. P. (2019). Connecting circular economy and Industry 4.0. *International Journal of Information Management*, 49:98–113.

Reinkemeyer, L. (2020). *Process Mining in Action: Principles, Use Cases and Outlook*. Springer Nature.

Souza, M. L. H., da Costa, C. A., de Oliveira Ramos, G., and da Rosa Righi, R. (2020). A survey on decision-making based on system reliability in the context of Industry 4.0. *Journal of Manufacturing Systems*, 56:133–156.

Suriadi, S., Andrews, R., ter Hofstede, A. H., and Wynn, M. T. (2017). Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs. *Information Systems*, 64:132–150.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4):327.

Verenich, I., Dumas, M., Rosa, M. L., Maggi, F. M., and Teinemaa, I. (2019). Survey and cross-benchmark comparison of remaining time prediction methods in business process monitoring. *ACM TIST*, 10(4):1–34.

Wang, T., Ji, Z., Sun, Q., Chen, Q., and Jing, X.-Y. (2016). Interactive multilabel image segmentation via robust multilayer graph constraints. *IEEE Transactions on Multimedia*, 18(12):2358–2371.

Xie, C., Yu, B., Zeng, Z., Yang, Y., and Liu, Q. (2020). Multilayer Internet-of-Things middleware based on knowledge graph. *IEEE Internet of Things Journal*, 8(4):2635–2648.

Zerbino, P., Aloini, D., Dulmin, R., and Mininno, V. (2018). Process-mining-enabled audit of information systems: Methodology and an application. *Expert Systems with Applications*, 110:80–92.

Zhou, Y. and Cheung, Y.-m. (2019). Bayesian low-tubal-rank robust tensor factorization with multi-rank determination. *IEEE TPAMI*.