Problem Characterization for Interactive Knowledge-Assisted Causal Analysis in Injection Molding

Jan Vrablicz[1,*], Alexander Rind[1] and Christian Mittermayr[2]

[1]Josef Ressel Centre for Knowledge-Assisted Visual Analytics for Industrial Manufacturing Data, St. Pölten University of Applied Sciences, Austria
[2]Greiner Packaging International, Gewerbestraße 15, 4642 Sattledt, Austria

Abstract

Manufacturing companies have to continuously optimize their processes and quickly respond to incidents. Therefore, they need a causal understanding of their domain to intervene. Advances in causal discovery and the advent of knowledge graphs as a generic data storage allow to merge human expertise and data-driven causal algorithms. This short paper characterizes the domain problem of causal analysis in injection molding based on a series of participatory design sessions with domain experts. It describes and abstracts data, users, and tasks. Finally, a knowledge-assisted visual analytics framework is proposed that allows engineers to perform causal discovery and subsequently test interventions on the causal model.

Keywords

Causal Knowledge, Industrial Systems, Human-in-the-Loop, Problem Characterization

1. Introduction

Industrial manufacturing companies have to set themselves apart constantly to stay relevant. As the growth due to globalization declines, manufacturing companies are put in a position where they have to operate on thin profit margins while also constantly showing improvements to their performance. However, finding further improvements becomes increasingly difficult. The modern understanding of efficient manufacturing is largely shaped by the Toyota Production System (TPS) [1]. TPS focuses on creating processes with little waste, which highlight issues as soon as they occur. This allows manufacturing companies to produce high quality products while reducing the working capital to a minimum. As a side-effect this method leads to issues being found more closely to the location and time where they originated. Complex environments benefit especially from this since less time is spent on understanding and searching and more time is spent on fixing the actual problems. The success of the TPS lead to cause-and-effect-based frameworks being incorporated as a central part in modern manufacturing. Methods such as the 5-Why approach and Failure Mode and Effect Analysis (FMEA) are common tools to steer efforts of manufacturing companies [1] to understand the root cause. While the TPS improves visibility for issues in the manufacturing process, the system relies on workers' experience to develop causal understanding. Knowledge-based root cause analysis methods such as the 5-Why approach [2] only function if the correct cause and effect chain can be identified by the workers applying the method. There are no verification systems that stop wrong conclusions and neither does the depth of five questions necessary correspond to the real depth of the root cause [3]. In contrast, modern data-driven causality approaches such as described by Peters et al. [4] try to lessen the effect of individuals by constraining their causal interpretations to the actual data rather than relying on experience and assumptions. While identifying causal relations is key to properly intervene in manufacturing processes, Hasan et al. [5] identified the lack in usability of causal discovery methods for real-world tasks as their runtime scales poorly with many nodes. Recent gradient-based approaches

1st Workshop on Leveraging SEmaNtics for Transparency in Industrial Systems (SENTIS), co-located with SEMANTiCS'25: International Conference on Semantic Systems, September 3–5, 2025, Vienna, Austria
*Corresponding author.
jan.vrablicz@fhstp.ac.at (J. Vrablicz); alexander.rind@fhstp.ac.at (A. Rind)
0009-0004-3694-9469 (J. Vrablicz); 0000-0001-8788-4600 (A. Rind)

[6, 7] aim to close this gap by providing approximate solutions scaling almost linearly with the amount of nodes. However, since these methods do not identify the correct causal relations but rather a plausible causal graph that fulfills the acyclicity requirement given the data they still require supervision by domain experts to validate and adjust the proposed causal relationships. Industrial injection molding generates data at a volume, velocity, and variety typical for big data environments [8]. Leveraging observational data to deepen causal understanding into the injection molding process can drive further improvements, even in highly optimized production environments.

This work aims to integrate human expertise with data-driven approaches to understand causal relationships in industrial injection molding. For this purpose, this short paper characterizes this domain problem and proposes a knowledge-assisted human-in-the-loop framework for causal analysis based on knowledge-assisted visual analytics [9]. Section 2 introduces the relevant background regarding causality. Section 3 surveys relevant works related to human-in-the-loop causal analysis especially in manufacturing and summarizes recent developments. Section 4 provides the methodological background in problem-driven design study research. Section 5 characterizes the industrial domain by characterizing data, users, and tasks. The knowledge-assisted framework for interactive causal analysis is outlined in Section 6. Finally, Section 7 summarizes and gives an outlook to future work.

2. Background

Rubin [10] elevates statistical analysis by introducing the idea of typical causal effect which can be estimated by taking the difference in effect between a control group and a treatment group in a randomized controlled trial. Rubin emphasizes the importance of using randomization to obtain an unbiased estimator while observational studies in the light of uncontrolled confounders only produce biased estimators. This idea takes a step forward from studies of association. This idea is later formalized [11] as the Rubin Causal Model (RCM) using the language of potential outcomes as a way to think about causal effects.

Pearl [12] introduces structural causal models (SCM). SCMs are directed acyclic graphs (DAG) that use directed edges to convey causal influence instead of just mathematical equations. Pearl also identifies the back-door/front-door criteria evolving the RCM by providing criteria to find suitable adjustment sets for observational studies. Adjusting by these sets, yields an unbiased estimation of the causal effect. Pearl [13] also provides the notion of d-separation, a graphical test to identify the potential information flow between nodes and the impact on conditioning on data via a graphical model. Pearl [14] further provides the notation of the do-Calculus arguing it alleviates overcomplicated mathematical expressions that derive from the potential outcome framework.

Schölkopf and von Kügelgen [15] compare causal models to statistical and mechanistic models on a spectrum of predictive power. Highlighting the short-comings of purely statistical modeling, they identify causal models as the key to provide robust out of distribution predictions and argue that enhancing traditional machine learning paradigms (such as surrogate models or autoencoder) with causality-based constraints will be the way forward.

Spirtes et al. [16] introduces the PC algorithm for causal discovery which recursively thins out edges of a fully connected graph. This reduces the amount of d-separation testing one has to do. Zheng et al. [7] re-imagine score-based causal discovery as a continuous constrained optimization task using the NOTEARS algorithm. They provide a novel differentiable continuous loss-function that implies “DAGness”. This allows the use of common gradient descent methods to fit a DAG for a given dataset which scales even for big-data regime problems. The GOLEM algorithm by Ng et al. [17] is an iterative upgrade to the NOTEARS algorithm using a likelihood-based objective instead of a regression-based objective. This transforms the continuous constrained optimization to continuous unconstrained optimization which is easier to solve. Bello et al. [6] introduces the DAGMA algorithm which improves the NOTEARS algorithm by constraining the search space to M-Matrices and using a log-determinate-based loss function. DAGMA provides lower run-times and higher performance in terms of structural Hamming distance (SHD) than PC, NOTEARS and GOLEM.

3. Related Work

Human-in-the-loop approaches for causal analysis have been under research in visual analytics (VA) as well as in industrial systems.

The Visual Causality Analyst [18] is a VA approach to involve domain experts in causal discovery allowing them to verify and edit causal relations. The Causal Structure Investigator [19] identifies subdivisions in the data and discovers causal models for these. DOMINO [20] focuses on discovering causal relations from time-series data. D-BIAS [21] is a VA approach that builds a causal model as a medium to visually audit social biases in tabular data set and weaken them based on domain knowledge. CausalChat [22] integrates large language models in a VA approach for causal discovery. VAINE [23] allows validation of causal relations but only on one outcome variable. Causalvis [24] is a python package providing visualization modules for causal structure modeling, cohort construction/refinement, and treatment effect exploration. SeqCausal [25] is a VA approach to explore and refine causal relations from event sequence data.

Pfaff-Kastner et al. [26] compares methods to create causal Bayesian networks based on ontologies in industrial manufacturing. Furthermore, they synthesize an ontology based on the basic formal ontology and in accord with the Industrial Ontology Foundry. They then only regard entities in their ontology to test for causality. Pfaff-Kastner et al. [26] identifies human experts as cornerstones to apply knowledge in the real world and formalize knowledge via ontologies and causal graphs. As ontologies are not common in manufacturing, their creation requires an additional effort before causal discovery can begin.

Fujiwara et al. [27] utilize JIT-LiNGAM [28] to understand the possibly time-dependent causal relationship in non-stationary time-series data of simulated vinyl acetate plants. They frame this as a continual learning task which they term continual causality. JIT-LiNGAM provides snapshots at different points in time of the current causal relationship while being amendable via new information. Wehner et al. [29] describe an interactive system for generating a causal graph for root-cause analysis based on observational data in electric vehicle manufacturing. They use a knowledge graph to explicitly store expert knowledge which is used to reduce the search space for causal discovery. Using expert knowledge significantly speeds up the learning process for cause-effect relationships. Vuković et al. [30] offers a different approach to aid the causal discovery process. Instead of investigating causal relationships for one large dataset they propose to apply the PC-algorithm to a slice of the data filtered by the machine for every machine. This results in many DAGs to be analyzed jointly via k-means clustering.

None of the VA approaches directly addresses industrial manufacturing as application domain and also the approaches found in industrial systems research either lack integration with domain experts or tackle different industries with diverging practices, e.g., tighter quality monitoring through full traceability of produced parts in automobile industry. Furthermore, the demonstrated data is often of smaller scale than expected for injection molding.

4. Methods

This work follows the paradigm of problem-driven research as outlined in the design study methodology [31]: starting from a real-world problem, researchers engage with domain experts, design an interactive system, validate it, and reflect the outcomes, thereby building on and extending the scientific body of knowledge. The first core stage of a design study is problem characterization and abstraction, which implies learning about the domain problem and abstracting its requirements. Sedlmair et al. [31] establish the problem characterization as a research contribution on its own.

The problem characterization for an interactive knowledge-assisted causal analysis approach is the scope of this work, which is conducted as participatory design (PD) [32] by the three co-authors with complementary expertise in data science, visual analytics, and industrial manufacturing using injection molding. Methodologically, it is based on literature research and a series of PD group meetings, which were scheduled on average every three weeks. Additional collaborators were added to some of the PD group sessions. The results of each meeting were documented by written notes.

5. Results

After providing contextual information on injection molding, the problem characterization is structured along the three aspects of data, users, and tasks as proposed by Miksch and Aigner [33]. Furthermore, information on the industrial application context is provided.

5.1. Industrial Application Context

Injection molding is a process that transforms polymer granulate into a 3D-object. The process is cyclic and has 4 phases: (1) clamping, (2) injection, (3) cooling, and (4) ejection.

(1) The form is clamped with predetermined pressure, so the mold remains closed throughout the injection. (2) Pellets come from a hopper into a barrel and are pushed forward by a screw. The friction of the rotating screw melts the pellets and the molten plastic accumulates in front of the screw (to be called metering). The screw moves forward and when it reaches the transfer position it goes to the second stage of injection which is "pack and hold". This is the switch point where it changes from velocity control to pressure control. At this point the cavities have been filled with plastic by more than 95 percent and during the "pack and hold" time the parts fill the rest of the way out. The amount of plastic left in front of the nozzle at the end of the fill is known as cushion. (3) The cooling of plastic starts when it gets in contact with the inside the mold. The mold has channels with a cooling liquid that allows the heat to dissipate. As the plastic cools it also hardens and it will take the desired shape. The part may shrink slightly during cooling. (4) Only when the cooling period has elapsed can the mold be opened. When the mold opens the part is pushed out. Force must be used because the part shrinks and may stick to the mold. After ejection, the mold can be closed and another shot can be injected for the process to begin again.

Parameters that determine the quality are associated with different components and parts of the cycle. There are material properties (e.g., melting temperature and viscosity), the geometric form and construction of the mold (e.g., aspect ratio, number of cavities), the machine and its settings. Besides the machine settings also measurements of some machine parameters are available. Additionally, there might be environmental parameters, like the cooling liquid (e.g., temperature, pressure) and ambient conditions like temperature and humidity. Quality control of the product provides information on the weight, various dimensions and surface quality.

Although it should be possible to completely model the injection molding process by thermodynamics, the process is complex and has too many unmeasured parameters such that perfect modeling has not been achieved yet. Furthermore, there are unpredictable disturbances (e.g., fluctuations in the material properties) that are not measured.

5.2. Data

Given the complicated nature of the industrial context, large scale data collection systems are vital parts to govern such systems. The collected data can be categorized by several attributes. First, the data can be divided by what it describes into *metadata* and *transactional data*. It can be further divided into the time scale and granularity on which data is reported and into the primary purpose why the data is collected. Finally, the reported values can be categorized by their scale and type of data that is reported. Metadata contains mostly categorical data such as material and machine identifiers which are used to group sets of transactional data together in a homogeneous manner but also contains preexisting knowledge about the industrial processes such as suitable temperature ranges for a given product-machine-tool combination. Transactional data documents transactions of business processes such as the creation of a particular part and the relevant attributes describing this creation of mostly continuous or ordinal values. Transactional data can be further split into groups of purpose such as *settings data*, *measurements* *data* and *product quality data* where settings data contains machine settings provided by an operator to the machine, measurement data contains the measurements of sensors and product quality data contains results of quality tests. The time scale and granularity of transactional data can differ and is loosely correlated to the groups of purpose. Primarily, the data collection of measurements data and settings data is orchestrated around the event of a shot. A shot describes the process of melting and injecting a polymer, and cooling and ejection of a finished part. During this shot sensors located inside the tool measure specific physical quantities of the injection molding process. Furthermore, the industrial control system is responsible for gathering process-relevant information such as cycle times or keeping track of the total shot count as well as settings provided to the machine by an operator. Settings data only sometimes changes. Thus, only change events are captured to avoid redundancy. This can be expanded on the time dimension without loss of information (disregarding the initial setting which is only available if all the data is viewed). Measurements data that is not orchestrated around the event of a shot contains information regarding the state of the environment, such as temperature and humidity of the shop floor. Since this information can be captured at higher frequencies, it can always be artificially aligned with shot events.

Furthermore, product quality data are collected on different granularities and intervals. Visual or incidental defects are identified in downstream processes, and thus, it is not always possible to attribute these defects to a specific shot. Product quality data can also be generated during offline quality check. This process takes samples of produced parts and checks for deviations of customer/process specifications. Furthermore, there are incident reports that cannot be directly attributed to a single shot event, but rather to an aggregation of shots or to a time interval. This data refers to semi-structured reports and documentation regarding incidents and their resolution.

Abstraction/Takeaways The gathered data can be viewed in a tabular relation where rows are shots and columns are settings, measurements of machine sensors, or environmental sensors. These attributes are either of continuous or categorical nature. Rows, i.e. shots, are time-stamped and they can be joined with other data at a coarser temporal granularity. Missing data can occur at any time either on a row or column level due to differences in granularity or due to technical issues during gathering, transporting or storing of the data. Differences in the provided soft- or hardware of tooling or machines can also lead to mismatches and thus missing data.

5.3. Users

In the industrial context, three groups of *domain experts* will need to perform causal analysis. *Process engineers* oversee the production process from procurement to production. They ensure process reliability and are responsible for checking the sensors used. In the event of process deviations, the process engineers carry out a cause analysis and rectify problems. They act as pivotal link between quality engineers, machine engineers and suppliers. *Quality engineers* ensure that the produced product satisfies all quality criteria and clear it for the customers. If quality violations are detected, they have to identify the root cause of the problem and take appropriate actions in close cooperation with the process engineer and machine engineer. Therefore, the quality engineer possesses extensive knowledge of the product itself as well as knowledge about the production processes and quality standards involved. *Machine engineers* are responsible for choosing and setting the correct machine settings. During the life cycle of the product, they will optimize the settings to minimize costs. In the case of problems during production they are instrumental in finding the root cause of the problem and fixing it. The machine engineers have expertise in machine-related topics.

In addition, *data scientists* provide data extraction and analysis pipelines, check data quality and performance of models. With their expertise in analytical methods, they can manage data-driven processes yet need to cooperate with domain experts for tasks requiring domain knowledge.

Abstraction/Takeaways Causal analysis involves both engineers as domain experts and data scientists with data analysis expertise, whereby the focus should be on enabling direct involvement of domain experts via self-explaining visual interfaces.

5.4. Tasks

Domain expert users are primarily tasked with keeping the current production running as is. In case of an incident, they need to be able to understand what happened and how to intervene to resolve the issue aiming to minimize the time the manufacturing process is impacted. Such incidents could lead to suboptimal quality, to reduced output, or even to stoppages. Partly users can utilize sensor data of the manufacturing process to gain insight into the current state of the system. Furthermore, some sensors have predefined corridors based on domain knowledge which indicate if a sensor is currently in an acceptable state. Resolving an incident requires the user to be able to interpret the incident, find the root cause why it occurred and return the system to an operational state. At best the user is able to anticipate an incident before it happens and prevents the incident by acting proactively. To achieve this, the user needs to understand not only the direct relationships between sensors, but also any temporal relationships.

The secondary task is to improve the current manufacturing process. For that the user has to first define the outcome they want to achieve. Users can either try to achieve this outcome via trial-and-error testing or in a systematic way via design of experiment (DoE). Regardless of what they do, they end up investigating the causal effect of their interventions. If the causal effect is already known due to previous experiments or through expert knowledge this can be applied directly and changes can be made accordingly. Otherwise, extensive testing in conjunction with data analysis is needed.

Abstraction/Takeaways On a more abstract level these two tasks share the same steps of (S1) identifying a need to intervene, (S2) define the objective of the intervention, (S3) identify how to intervene, (S4) apply the intervention, and (S5) evaluate if the objective was achieved. A causal model of the injection molding process can be utilized to solve steps (S2) and (S3) and provide the basis for step (S5).

5.5. Summary

A KAVA framework for causal analysis should enable engineers as domain experts to build and work with causal models of the injection molding process from tabular, time-stamped data in order to resolve incidents or improve performance. The framework should abstract away many of the complexities associated with causal models for domain experts to be able to iteratively add knowledge to a persistent knowledge store. Furthermore actionable items should be the outcome of applying such a framework instead of only predictions as it is usual from machine learning methods.

6. Knowledge-Assisted Visual Analytics Framework

To tackle the domain problem of causal analysis in injection molding, we propose a knowledge-assisted visual analytics framework consisting of automated analysis processes (Section 2) and interactive visual interfaces for the user groups involved (Section 5.3). A scalable solution to store domain knowledge about any entities and their relationships are knowledge graphs (KG) [34]. The building blocks for KG-based machine-human collaboration in industrial manufacturing by Nagy et al. [35] provides guidance on how to properly implement and harness the strengths of KGs in such a framework.

The proposed framework has three phases as shown in Figure 1 since a causal model for each machine needs to be discovered and fitted to the data before optimization or incident response tasks can be addressed. First, a knowledge-constrained causal discovery algorithm dissects the transactional data providing a SCM, which is a DAG of causal relationships. As data volume is large, efficient discovery algorithms are preferred such as described by Bello et al. [6] and Zheng et al. [7]. Since these discovery algorithms do not discover the real underlying causal structure but instead propose a DAG given the constraints, human supervision and refinement are needed. Domain experts such as process engineers

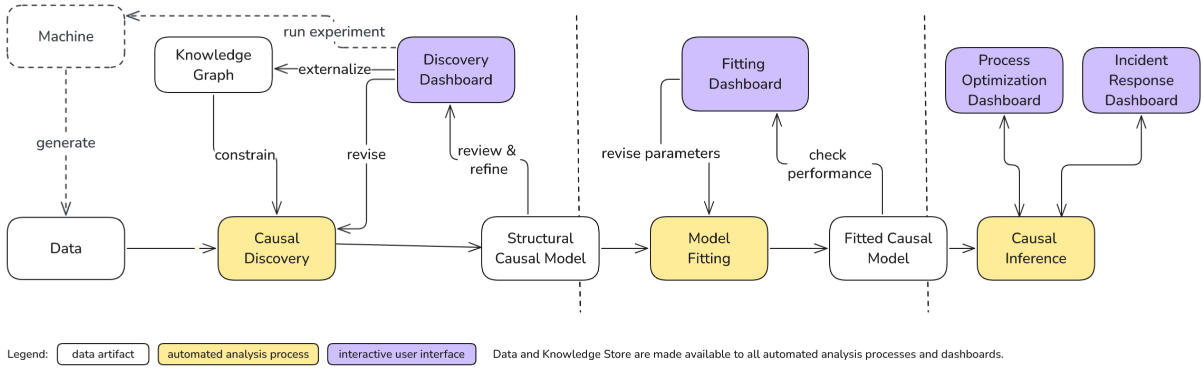

Figure 1: Components and processes of a knowledge-assisted framework for causal analysis. The diagram is split into three phases named after the automated analysis processes (yellow) causal discovery (left), model fitting (middle), and causal inference (right)

can do this via the provided discovery dashboard. Furthermore, approaches, such as Hoyer et al. [36], can help verify the direction of causal effects if the corresponding assumptions are met. Findings from these processes can be externalized into the knowledge store in order to reduce future refinement effort. If insufficient data is present the user can decide to actively pursue an experiment to gather new data. Then, this process can be repeated. In the second phase, once the domain experts are satisfied with the SCM, data scientist can fit models in classical machine learning fashions for each causal relationship using only the immediate parents to predict the effect variable. A dedicated dashboard will support them in this phase. Finally, once a sufficiently accurate model is fitted, domain experts can utilize the provided dashboards to investigate their optimization inquiries and incident responses. Even though both tasks follow the same steps (S1–S5), separate dashboards allow focusing on relevant aspects of the data (e.g., temporal trends before the incident vs. a holistic view for optimization purposes).

7. Conclusion

In this paper we characterized the problem domain of causal analysis for industrial injection molding. For that, the data, users, and tasks were elaborated and abstracted. The characterization was supported by domain experts in industrial injection molding through a series of interviews. As a partial solution for the identified tasks, we propose a knowledge-assisted framework that allows engineers to explicitly store their causal understanding of the domain and automatically discover unknown relationships between domain entities while also using operational data to constrain said discovery. Furthermore, the proposed framework enables engineers to formulate and test their interventions before applying them to the real machine.

Future work towards the realization of this framework includes

- defining a suitable schema for storing engineering knowledge in a knowledge graph [34],
- designing effective dashboards that support domain expert engineers in working with causal models and provide sufficient onboarding [37] to correctly interpret and operate these algorithms,
- integrating domain knowledge from the KG into causal discovery, and
- implementing and evaluating the proposed framework.

Acknowledgments

The financial support by the Austrian Federal Ministry of Labour and Economy, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] T. Ohno, N. Bodek, Toyota Production System: Beyond Large-Scale Production, 1 ed., Productivity Press, 2019. doi:10.4324/9780429273018.

[2] D. Pietsch, M. Matthes, U. Wieland, S. Ihlenfeldt, T. Munkelt, Root Cause Analysis in Industrial Manufacturing: A Scoping Review of Current Research, Challenges and the Promises of AI-Driven Approaches, Journal of Manufacturing and Materials Processing 8 (2024) 277:1–277:19. doi:10.3390/jmmp8060277.

[3] A. J. Card, The problem with '5 whys', BMJ Quality & Safety 26 (2017) 671–677. doi:10.1136/bmjqs-2016-005849.

[4] J. Peters, D. Janzing, B. Schlkopf, Elements of Causal Inference: Foundations and Learning Algorithms, The MIT Press, 2017.

[5] U. Hasan, E. Hossain, M. O. Gani, A survey on causal discovery methods for I.I.D. and time series data, Transactions on Machine Learning Research (2023). URL: https://openreview.net/forum?id=YdMrdhGx9y.

[6] K. Bello, B. Aragam, P. Ravikumar, DAGMA: Learning dags via m-matrices and a log-determinant acyclicity characterization, in: S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, A. Oh (Eds.), Advances in Neural Information Processing Systems, volume 35, Curran Associates, Inc., 2022, pp. 8226–8239.

[7] X. Zheng, B. Aragam, P. Ravikumar, E. P. Xing, DAGs with NO TEARS: Continuous optimization for structure learning, in: Proc. 32nd Int. Conf. on Neural Information Processing Systems, NIPS'18, Curran Associates Inc., 2018, pp. 9492–9503. doi:doi/10.5555/3327546.3327618.

[8] I. Emmanuel, C. Stanier, Defining Big Data, in: Proc. Int. Conf. Big Data and Advanced Wireless Technologies, BDAW '16, ACM, 2016, pp. 5:1–5:6. doi:10.1145/3010089.3010090.

[9] P. Federico, M. Wagner, A. Rind, A. Amor-Amorós, S. Miksch, W. Aigner, The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics, in: Proc. IEEE Conf. Visual Analytics Science and Technology (VAST), 2017, pp. 92–103. doi:10.1109/VAST.2017.8585498.

[10] D. B. Rubin, Estimating causal effects of treatments in randomized and nonrandomized studies., Journal of Educational Psychology 66 (1974) 688–701. doi:10.1037/h0037350.

[11] J. D. Angrist, G. W. Imbens, D. B. Rubin, Identification of Causal Effects Using Instrumental Variables, Journal of the American Statistical Association 91 (1996) 444–455. doi:10.1080/01621459.1996.10476902.

[12] J. Pearl, Causal Diagrams for Empirical Research, Biometrika 82 (1995) 669–688. doi:10.2307/2337329. arXiv:2337329.

[13] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, San Francisco, CA, USA, 1988.

[14] J. Pearl, Causal inference in statistics: An overview, Statistics Surveys 3 (2009). doi:10.1214/09-SS057.

[15] B. Schölkopf, J. von Kügelgen, From statistical to causal learning, in: Proc. Int. Cong. Mathematicians (ICM), volume VII, EMS Press, 2022, pp. 5540–5593. doi:10.4171/ICM2022/173.

[16] P. Spirtes, C. Glymour, R. Scheines, Causation, Prediction, and Search, 2 ed., The MIT Press, 2001. doi:10.7551/mitpress/1754.001.0001.

[17] I. Ng, A. Ghassami, K. Zhang, On the role of sparsity and DAG constraints for learning linear dags, in: H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin (Eds.), Advances in Neural Information Processing Systems, volume 33, Curran Associates, Inc., 2020, pp. 17943–17954.

[18] J. Wang, K. Mueller, The visual causality analyst: An interactive interface for causal reasoning, IEEE Trans. Visualization and Computer Graphics 22 (2016) 230–239. doi:10/ps3v.

[19] J. Wang, K. Mueller, Visual causality analysis made practical, in: 2017 IEEE Conference on Visual Analytics Science and Technology (VAST), 2017, pp. 151–161. doi:10/ggkjbc.

[20] J. Wang, K. Mueller, DOMINO: Visual causal reasoning with time-dependent phenomena, IEEE Trans. Visualization and Computer Graphics 29 (2023) 5342–5356. doi:10/ps3w.

[21] B. Ghai, K. Mueller, D-BIAS: A causality-based human-in-the-loop system for tackling algorithmic bias, IEEE Trans. Visualization and Computer Graphics 29 (2023) 473–482. doi:10/pz3r.

[22] Y. Zhang, A. Kota, E. Papenhausen, K. Mueller, CausalChat: Interactive causal model development and refinement using large language models, arXiv:2410.14146, 2024. doi:10/g8pv82.

[23] G. Guo, M. Glenski, Z. Shaw, E. Saldanha, A. Endert, S. Volkova, D. Arendt, VAINE: Visualization and ai for natural experiments, in: 2021 IEEE Visualization Conference (VIS), 2021, pp. 21–25. doi:10.1109/VIS49827.2021.9623285.

[24] G. Guo, E. Karavani, A. Endert, B. C. Kwon, Causalvis: Visualizations for causal inference, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, ACM, 2023, p. 462:1–462:20. doi:10.1145/3544548.3581236.

[25] Z. Jin, S. Guo, N. Chen, D. Weiskopf, D. Gotz, N. Cao, Visual causality analysis of event sequence data, IEEE Trans. Visualization and Computer Graphics 27 (2021) 1343–1352. doi:10/ghgt6f.

[26] M. M.-L. Pfaff-Kastner, K. Wenzel, S. Ihlenfeldt, Concept Paper for a Digital Expert: Systematic Derivation of (Causal) Bayesian Networks Based on Ontologies for Knowledge-Based Production Steps, Machine Learning and Knowledge Extraction 6 (2024) 898–916. doi:10/pz3s.

[27] D. Fujiwara, K. Koyama, K. Kiritoshi, T. Okawachi, T. Izumitani, S. Shimizu, Prospects of Continual Causality for Industrial Applications, in: M. Mundt, K. W. Cooper, D. S. Dhami, A. H. Ribeiro, J. S. Smith, A. Bellot, T. L. Hayes (Eds.), AAAI Bridge Program on Continual Causality, volume 208 of *Proceedings of Machine Learning Research*, PMLR, 2023, pp. 18–24.

[28] D. Fujiwara, K. Koyama, K. Kiritoshi, T. Okawachi, T. Izumitani, S. Shimizu, Causal Discovery for Non-stationary Non-linear Time Series Data Using Just-In-Time Modeling, in: Proceedings of the Second Conference on Causal Learning and Reasoning, PMLR, 2023, pp. 880–894.

[29] C. Wehner, M. Kertel, J. Wewerka, Interactive and intelligent root cause analysis in manufacturing with causal bayesian networks and knowledge graphs, in: Proc. IEEE 97th Vehicular Technology Conference (VTC2023-Spring), 2023, pp. 1–7. doi:10/ps2g.

[30] M. Vuković, P. Haferlbauer, B. Mutlu, S. Thalmann, Unveiling operational dynamics: Causally interpretable clustering for understanding the usage of plastic injection molding machines, in: Proc. 26th Int. Conf. Business Informatics (CBI), 2024, pp. 218–227. doi:10/pzv6.

[31] M. Sedlmair, M. Meyer, T. Munzner, Design study methodology: Reflections from the trenches and the stacks, IEEE Trans. Visualization and Computer Graphics 18 (2012) 2431–2440. doi:10/f4fv7x.

[32] S. Jänicke, P. Kaur, P. Kuzmicki, J. Schmidt, Participatory visualization design as an approach to minimize the gap between research and application, in: Proc. VisGap – the Gap between Visualization Research and Visualization Software, 2020, pp. 35–42. doi:10.2312/visgap.20201108.

[33] S. Miksch, W. Aigner, A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data, Computers & Graphics 38 (2014) 286–290. doi:10/f3szvk.

[34] A. Hogan, E. Blomqvist, M. Cochez, C. D'amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier, A.-C. N. Ngomo, A. Polleres, S. M. Rashid, A. Rula, L. Schmelzeisen, J. Sequeda, S. Staab, A. Zimmermann, Knowledge Graphs, ACM Computing Surveys 54 (2021) 71:1–71:37. doi:10.1145/3447772.

[35] L. Nagy, J. Abonyi, T. Ruppert, Knowledge Graph-Based Framework to Support Human-Centered Collaborative Manufacturing in Industry 5.0, Applied Sciences 14 (2024) 3398. doi:10.3390/app14083398.

[36] P. Hoyer, D. Janzing, J. M. Mooij, J. Peters, B. Schölkopf, Nonlinear causal discovery with additive noise models, in: Advances in Neural Information Processing Systems, volume 21, Curran Associates, Inc., 2008, pp. 689–696.

[37] C. Stoiber, D. Ceneda, M. Wagner, V. Schetinger, T. Gschwandtner, M. Streit, S. Miksch, W. Aigner, Perspectives of visualization onboarding and guidance in VA, Visual Informatics 6 (2022) 68–83. doi:10.1016/j.visinf.2022.02.005.