AI SUPPORTED ROOT CAUSE ANALYSIS AT TRUMPF

Co-Innovation project by Trumpf and Xplain Data

CASE STUDY

Shortened English version

[1]

CREATED BY

Emanuel Sideras, Dr. Michael Haft

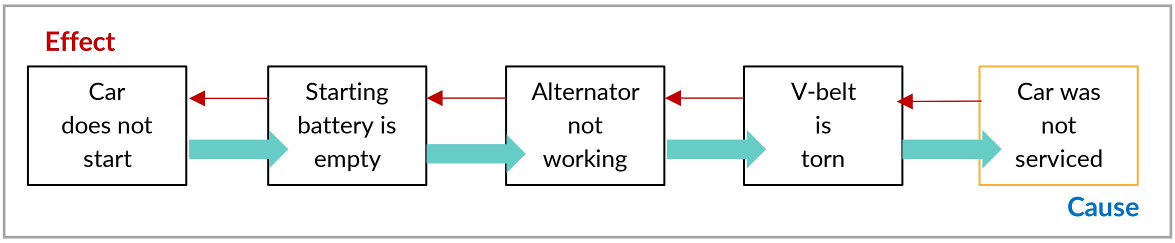
DATE ISSUED

10.09.2022

"AI helps us to identify cause-and-effect relationships in complex data," says **Dr.-Ing. Mathias Kammüller**, Chief Digital Officer and Trumpf Executive Board member. "The trick is to combine AI with existing domain knowledge." Those responsible at Trumpf are meeting this challenge in a current project.

An essential task in quality management is the monitoring of key metrics. If a key metric slips, the question arises - why?

But: Typically, there is not "a single cause" - often a chain of events leads to a problem. Which element in this chain is relevant (as a cause) depends on your goal - see Figure1.



[2]

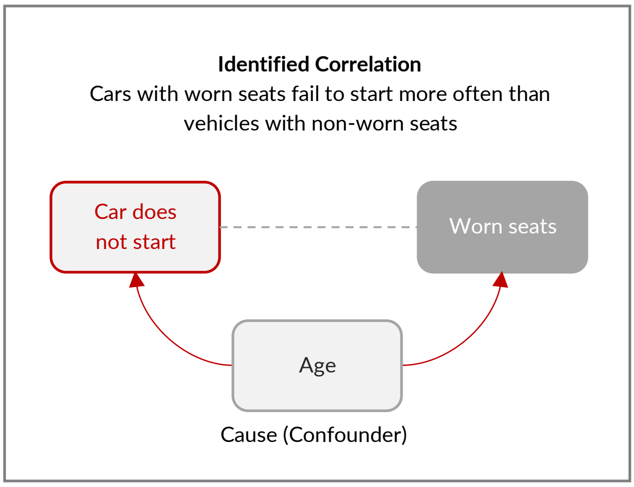
Figure 1: Sequence of events (illustrative) that ultimately lead to the problem "vehicle does not start". The repeated question of "Why?" leads up the chain (5-Why method). Which level is relevant? For the automotive workshop foreman, it is level 3 (broken V-belt). For the reliability engineer who wants to optimize maintenance intervals, it is level 4 (vehicle was not maintained) ... and so on ...

The answer to the root cause question, therefore, depends on the role of the person and its mission - consequently, a meaningful root cause analysis process requires the close involvement of the domain expert.

In each individual case, a chain of events may have led to the observed problem. In the statistical analysis of many failure events with different patterns of failures, it becomes more difficult: a wide range of possible causes - a multitude of variables in a network of dependencies - must be understood. There is a huge number of observed correlations within this network as well as to the error variable - and correlation does not mean causality!

Otherwise, the pharmaceutical industry would not have to conduct expensive randomized trials (RCTs) to prove the efficacy of a therapy. In "Real World Data" (observational data – as opposed to data collected in a controlled experiment) there are endless trivial correlations – without causal relationships – which can mostly be explained by confounders (a common cause – see Figure 2).

[2] Image Source: ©Trumpf GmbH + Co KG



[3]

Figure 2: Example of a correlation without causality (dashed line): If a vehicle does not start, it is statistically more likely to observe worn seats. However, the wearing of the seats does not cause starting problems. The correlation arises from a common cause - the confounder "age". In a real-world example, the condition of a machine is described by 1000s of variables - confounders, such as age or hours of operation, create correlations between almost all variables. Identifying the few potential causalities in this "forest of correlations" is the task of Causal Discovery.

AI methods are able to sift through these correlations in complex data (Big Data), analyze them for possible confounders, and exclude irrelevant correlations. Thus, the multitude of correlations is boiled down to a small number of potential cause-and-effect relationships.

With that, the task is still not done. What if relevant confounders are missing from the data? This is another reason why experts have to evaluate the results. Once again, we arrive at the same conclusion: in a cause-and-effect research process, AI algorithms must work hand in hand with experts, who use their domain knowledge to make a reasoned assessment of the results.

No sooner said than done: Trumpf decided to embark on a co-innovation project with the startup Xplain Data to overcome the above challenge with their AI algorithms. As part of a proof of concept, methods that were originally designed for the pharmaceutical sector were successfully adapted to Trump's requirements and applied to current issues.

[3] Image Source: ©Trumpf GmbH + Co KG

The Power of AI: Identify and quantify potential

What exactly do Xplain Data AI methods offer to help experts uncover causal relationships in observational data?

1. **The reduction of an unmanageable number of correlations in a complex data schema to a few hypotheses for potential causes.**
   Missing confounders are a critical factor in a root cause analysis. Therefore, comprehensive information is required, which necessarily implies a complex data model. Xplain Data´s AI algorithms dig through large amounts of information in a complex model of interrelated tables – thereby searching for confounders and segregating away meaningless correlations to the failure variable. What remains is a small number of well-founded hypotheses for potential causes. These possible causes are presented to the expert for assessment.
   In reverse, the expert may express his assumptions for root causes – which are then evaluated against data.

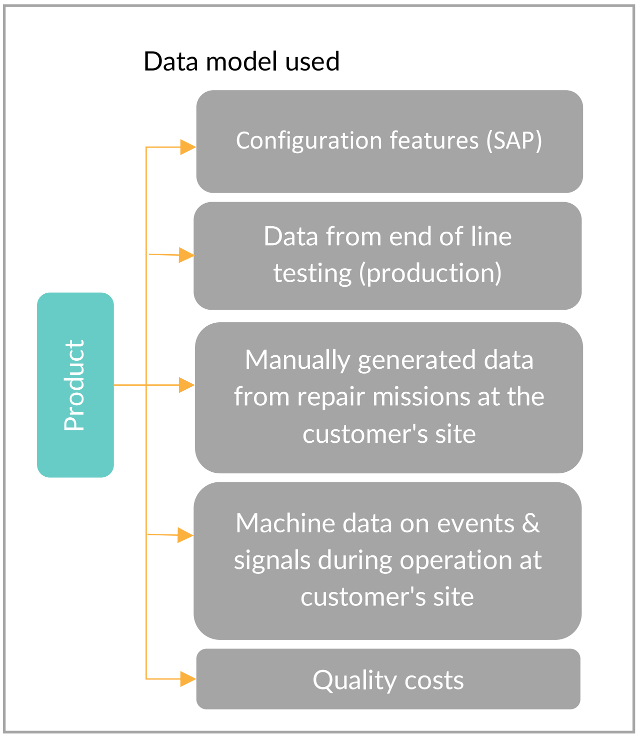
2. **Results that are intuitively comprehensible for the domain expert – the basis for an iterative analysis process.**
   The successful interaction of experts and AI requires intuitively understandable analysis results (or intuitively understandable model parameters): So-called ICI models (*Independence of Causal Influences*) enable this. An overall effect is decomposed into independent contributions: This enables the expert to interpret and evaluate factors independently. At the same time, interactive interfaces support a process of question and follow-up questions (5Ws).

3. **Quantification of the effects of causes.**
   The decomposition into independent contributions (see 2.) also facilitates a quantitative evaluation of individual factors and provides information on whether these are statistically significant (p-value).

What is the goal of a root cause analysis? To intervene in the development, production, or operation of a machine in a way that **prevents future causes for failures**. Only with a quantified factor, it can be assessed whether a measure is adequate.

Encouraging results from the Proof of Concept

The purpose of the POC project was to understand causes that lead to failures of a machine (e.g., a laser cutter) in the field. The following data was available:



[4] *Figure 3: Complex data model at Trumpf in the PoC project. A wide variety of data is tied to the product (the machine), such as configuration and inspection characteristics at delivery or time series data from signals in the field.*

For Causal Discovery, it is essential to have data which is as comprehensive as possible. As explained above: the more comprehensive the information available, the lower the risk of overlooking important confounding factors and drawing incorrect conclusions about possible causes. Fig. 3 sketches the different streams of data – typically different time series of events – which are attached to each product (e.g., each laser cutter operated in the field). Algorithms for Causal Discovery must be able to process such complex data models.

[4] Image Source: ©Trumpf GmbH + Co KG

Example results (surprise included)

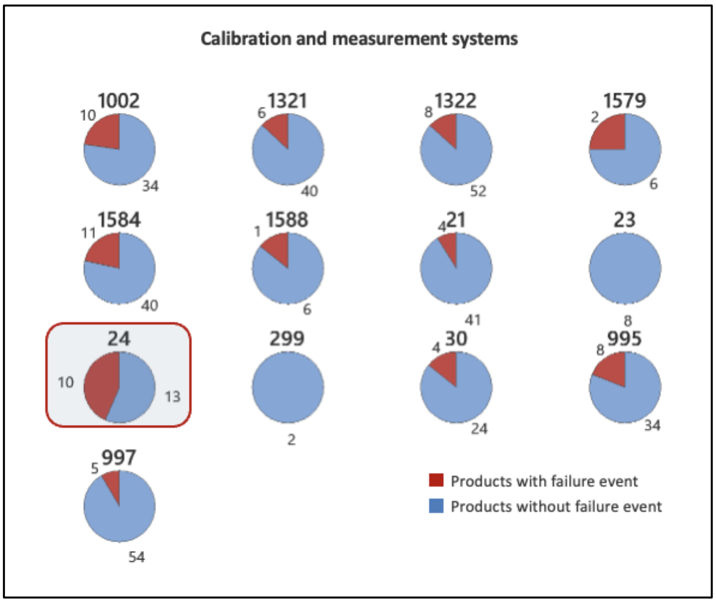
Example 1: Failures of machine operation are related to test values measured at the end of the assembly line, and surprisingly also to the device used for end-of-line calibration.

Findings:

- Certain test values correlated with subsequent failure rates in the field.
- Some of these correlations could not be explained by confounders.
- The data therefore suggested a potential causal relationship.

For factual clarification, these findings were presented to the experts responsible for the production process. These experts were able to judge whether the introduction of an acceptance limit for test parameters was sensible (and profitable). This provided a reasonable trade-off between required rework after production and reduced likelihood of operational failures.

Surprisingly, amongst the potential causes identified, there also was a link to the testing and calibration system used at end of the assembly line (see Figure 4). Note that there might be trivial reasons why failure rates are associated with used test systems: for example, simply because specific testing units are used for specific classes of machines that inherently have different failure rates. However, no such explanation (no confounder) could be found that explains the observed association. The suspicion that there might indeed be a causal relationship here was confirmed in the subsequent factual analysis which revealed problems with specific testing units.

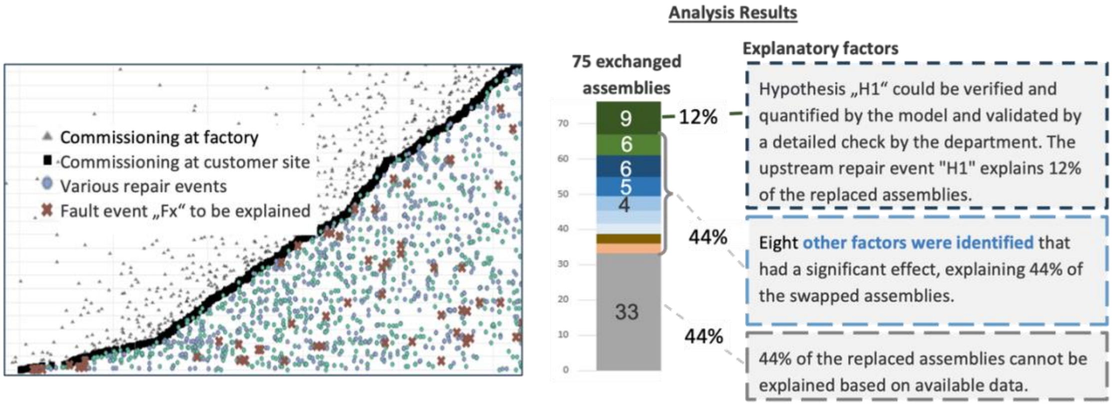


[5]

Figure 4: Noticeable increase in operational failures for calibration system no. 24

[5] Image Source: ©Trumpf GmbH + Co KG

Example 2: Operational failures are related to previous repairs



[6]

Figure 5: Machines installed at the customer's site plotted one above the other (Y-axis) and the chronological progression of their repair and failure events (X-axis). The failure event Fx, which was observed 75 times, requires explanation (exchange of an assembled component).

The expert hypothesis: A previous repair event H1 leads to later trouble in terms of the required exchange of assembly F1. Those events correlate, but how many exchanged assemblies F1 can actually be causally traced back to previous repair event H1?

The challenge in answering this question:

- Alongside repair event H1, often various other events (other repairs) took place before the assembly exchange F1, which could also have potentially caused the required exchange.
- Inherent product characteristics at delivery (and many other things) are also possible explanatory factors.

The model assessed and quantified the H1 repair event as a significant potential cause. The responsible department undertook a case-by-case review, thereby validated both, the causal relationship between H1 and F1 and its quantification (Figure 5, Explanatory Factors).

Example 3: Forecasts for individual products

A good explanatory model can be used to predict the probability of product failures. If a defined limit value is exceeded, an early warning signal is sent to the service department or the customer, and appropriate measures can be initiated.

In practice, for example, an alarm is triggered for all machines with a calculated probability of failure that exceeds 70%. Out of 10 such alarms, only one is wrongly triggered; in 9 cases, early intervention can take place and the failure can be prevented.

[6] Image Source: ©Trumpf GmbH + Co KG

Résumé & Operationalization

"Despite initial concerns - as important data has not been collected so far and existing data originates from a manual collection process with questionable data quality - the results far exceeded expectations. It was therefore decided to implement this successful Xplain Data methodology as standard in the Trumpf QM analysis processes."
Dr. Volker Hettich, Head of Central Quality Management.

The intensive collaboration of many departments (including IT, information security, development, field analysis, and production) made the successful implementation of the AI methodology possible. It is used as follows:

1. **Manual**: Individual root cause analyses can be configured (by authorized users) via an interactive interface, executed and saved as a template.
2. **Automated:** Standardized analyses for various cases can run overnight and are available to departments in the morning, with daily updates.

Use Cases at a glance:

- **Field Analysis & Development**: Clarification of causes of high failure rates
- **Quality management**: Clarification of causes of high-quality costs
- **Production**: Identification of problematic inspection features or adjustment and measurement systems
- **Condition monitoring**: Early detection of products with high failure probability

Through the Xplain Data AI methodology, Trumpf is now able to automatically search not only individual tables but entire relational data models during root cause analyses. A scalable methodology for generating system knowledge is the pleasing - and profitable - result.

A detailed report is available in the German quality management magazine "QZ-online":
https://www.qz-online.de/a/fachartikel/ki-hilft-bei-der-ursachenanalyse-2836691

AUTHORS

Emanuel Sideras is Six Sigma Master Black Belt in central quality management at the company TRUMPF SE + Co. KG.

Dr. Michael Haft is CEO of Xplain Data GmbH.

ABOUT TRUMPF

The Trumpf Group - still family-owned with their headquarters in Ditzingen, Germany - was founded in 1923 as a mechanical workshop and has become a technology and market leader in machine tools for flexible sheet metal processing and industrial lasers. For example, the laser specialist supplies the EUV lasers for exposing the latest generation of chips for the latest Apple devices. In the fiscal year 2020/21, the company generated sales of 3.5 billion euros with 14,767 employees and 80 global subsidiaries.

ABOUT XPLAIN DATA

Xplain Data GmbH, founded 2015, focuses on the development of innovative technologies in the field of Machine Learning and Artificial Intelligence. Xplain Data algorithms enable companies in all industries to identify the few, potentially causal relationships in their "real world data" that are hidden behind a plethora of trivial correlations. Users can leverage this cause-and-effect knowledge to intervene in their business processes to eliminate the causes of errors or achieve a desired effect. Xplain Data customers include leading enterprises in the mechanical engineering, manufacturing, and healthcare sectors, which use the technology not only for sophisticated data analyses but also for predictive maintenance and yield optimization.

CONTACT

TRUMPF SE & Co. KG
emanuel.sideras@trumpf.com | www.trumpf.com

XPLAIN Data GmbH
michael.haft@xplain-data.com | www.xplain-data.com