

preHAZOP: Graph-Based Safety Analysis for Early Integration into Automated Engineering Workflows

Jonas Oeing*, Tim Holtermann, Wolfgang Welscher, Christian Severins, Marius Vogel, and Norbert Kockmann

DOI: 10.1002/cite.202200222

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



Supporting Information
available online

The increasing digitalization and standardization within the process industry lead to a high availability of digital, machine-readable processes and plant descriptions. In particular, the publication of the DEXPI standard provides a digital representation of plant topologies including a complete description of all specifications. In early planning phases, this can be used as the basis for an automated safety assessment since digital availability significantly simplifies accessibility for smart search algorithms. This paper presents the preHAZOP search algorithm, which was developed to analyze P&IDs in DEXPI format and to detect safety-critical deviations regarding their risk according to a classical HAZOP analysis. The preHAZOP is of particular interest in early process development stages and can be easily integrated into modern, digital engineering workflows.

Keywords: Detail engineering, DEXPI, HAZOP, Safety analysis, Smart engineering

Received: December 15, 2022; *revised:* March 14, 2023; *accepted:* April 25, 2023

1 Introduction

Safety engineering is an indispensable element in the process industry [1]. The fact that safety is non-negotiable and safety analyses therefore can take a large part of the engineering phase means that safety-related time and cost savings are of particular interest [2]. The following part highlights past attempts to automate safety assessments and the issues that have hindered the widespread application of these methods.

For over 30 years, there have been attempts to perform safety analyses in the process industry. This results from the enormous potential in terms of computer-aided execution by deterministic algorithms. One example is the RAST tool [3], which can perform the safety assessment of equipment. The tool accesses a scenario database and calculates the worst-case consequences, as well as their impact. Before this happens, however, the tool relies on the user to provide all the input, such as the actual equipment, operating conditions, material data, and more [3]. In addition to the RAST tool, commercial packages exist that offer frequently occurring scenarios for specific equipment or processes. However, automated detection is not included [3]. The development of a HAZOP (hazard and operability) tool, which

aims at far-reaching automation of the overall concept, requires a digitally accessible plant model that includes all necessary information for a subsequent safety assessment. In the past, several approaches have been developed, in which the plant model must first be elaborately created. Usually, the created models are based on graph theory due to the obvious transferability, whereas different model languages are used depending on the method. In some approaches, graphical editors support the time-consuming creation, but the time required for model creation contradicts the goal of saving time [4].

¹Jonas Oeing <https://orcid.org/0000-0002-8410-7651>

(jonas.oeing@tu-dortmund.de), ¹Tim Holtermann,

²Wolfgang Welscher, ³Dr. Christian Severins, ³Marius Vogel,

¹Prof. Dr.-Ing. Norbert Kockmann

¹Department of Biochemical and Chemical Engineering, Laboratory of Equipment Design, TU Dortmund University, Emil-Figge-Straße 68, 44227 Dortmund, Germany.

²X-Visual Technologies GmbH, James-Franck-Straße 15, 12489 Berlin, Germany.

³Bayer AG – Engineering & Technology, 51368 Leverkusen, Germany.

As data is increasingly being converted into digital formats in the context of digitization, the model of the production plant, which takes the form of the P&ID flow diagram, is already available in machine-readable form. In addition, it is possible to store process data when using correspondingly powerful CAD software. The integration of already existing plant models into a HAZOP tool being developed is therefore of great importance to circumvent the problem of model creation.

A relevant example is the method LDGHAZOP (layered digraph HAZOP), which accesses a P&ID flowsheet by using a digital interface [5]. Hazard identification is performed on an automatically assembled graph. Using a complex model-based approach, an attempt is made to automatically generate the potential hazard scenarios via an algorithm, but the degree of generated scenarios that turn out to be irrelevant when viewed manually is too high. In addition, the interface used limits the applicability to the CAD software SmartPlant P&ID [4, 5].

In another method, a CAEX model was used to perform the safety analysis. The CAEX format was generally developed as a data exchange format for CAE systems and can also be used as a machine-readable description of the equipment. In addition to using the CAEX model, the method has a knowledge base that has equipment-type deviations with associated, possible causes and consequences. An appropriately designed evaluation algorithm then outputs possible combinations of deviation impact and cause for each component in the equipment model. A limitation that arises here, but also in relation to other methods, is the manual input of process data, which only ensures a complete data basis. An integrated retrieval of this information from existing simulation models would enormously reduce the effort of manual data maintenance [6].

Despite the numerous attempts shown, the widespread practical application of automation tools for safety assessment in the process industry does not exist until now. The main reasons for this are the data structures and the missing standardization of information. For this reason, this work uses a graph-based information model, which includes both plant data (topology, specifications, etc.) and process information (materials, pressures, temperatures, concentrations, etc.).

In this paper, an engineering-assistance tool based on AI-supported HAZOP-studies is presented. Standardized plant topologies provide a good basis for the application of AI methods and deterministic algorithms [7]. Wiedau et al. pointed out that DEXPI in particular is a good way to describe plant topologies in a machine-readable and standardized way and thus represents a useful baseline for the entire asset lifecycle [7]. Modularization of chemical equipment [8] and modular plants [9] assist the standardization using the DEXPI format and the related safety analysis.

The present work can be subdivided into three sections. First, the relevant theoretical background is presented. Subsequently, the suggested workflow is presented. In this

section, first, the data preprocessing and then the structure of the preHAZOP are explained. Finally, a detailed validation of the preHAZOP takes place on the basis of an example process.

2 Related Models and Methods

2.1 HAZOP Analysis

The HAZOP process is an important safety engineering tool [10]. The workflow of an HAZOP analysis is defined in the standard IEC 61882 [11]. HAZOP stands for Hazard and Operability and can be divided into the following steps: predicting deviations, finding the cause(s), evaluating the impact(s), and taking countermeasures. These are the main steps performed, when conducting a HAZOP procedure. It is the standard procedure for hazard identification in the context of a safety assessment in process industries and is characterized by its systematic procedure, transparency, and holistic approach. It is also flexible and can be used for all types of plants but has often some typical aspects for the individual companies or products. The procedure is carried out by an interdisciplinary team to be able to access expertise from many specialist areas such as production, safety, and technology [12]. In addition to systematic analysis, the main aim is to learn from existing mistakes [13–15].

The identification of deviations is followed by the detection of the associated cause based on experience and process knowledge. Typical causes are the failure of technical equipment, power failure or human error. At this point, the complexity of the study depends heavily on the comprehensiveness of the root cause investigation. The next step in the HAZOP process is an assessment of the impact of which deviations from the intended operation entail. These can be hazards to people and the environment, damage to property, operational downtimes, or other factors. Suitable measures are formulated to counteract the possible effects. Safety precautions are taken that are tailored to the respective scenarios. The precautions taken can limit damage and prevent occurrence of an incident. Any safety precautions that may already be in place have not been considered for the time being. Reason is the questioning of the necessity and determination of the quality requirement. The results are safety-relevant hazard scenarios, including suitable countermeasures (see Tab. 1). The listing in a standardized table enables transparent and uniform documentation. If necessary, additional columns are added to describe the scenario or to subdivide the measures into those that already exist and those that need to be taken. [10, 12]

2.2 DEXPI

DEXPI is a machine-readable P&ID exchange format under development by the DEXPI Initiative [16]. The initiative

Table 1. Example of HAZOP scenarios.

Guide word	Parameter	Cause	Consequence	Countermeasure
High	Pressure	Fire	Exploding vessel	Relief valve
Low	Pressure	Leakage	Product loss	Safety valve
Other	Component	Leakage	Corrosion	Stainless steel
...

consists of owner/operator companies as well as engineering, procurement & construction (EPC) companies, software vendors, and research institutions. The latest data model and the associated DEXPI specification 1.3 [16] were published in 2021. Within the specification, different international standards for the description of engineering relevant data for P&IDs are combined (e.g., ISO 15926 [17], ISO 10628[18], IEC 62424 [19], ISO 10209 [20]). In particular, these include plant breakout structures, instrumentation, properties of equipment and components, and piping topology. The DEXPI information model is already offered by some manufacturers and is exchangeable via a Proteus XML schema [21]. Due to the good accessibility and storage of all plant-specific data, DEXPI is particularly well suited as a data basis for performing automated preHAZOPs within the scope of this work.

2.3 Process Simulation

In addition to plant topology and specification, the process carried out in the plant is a crucial factor for the safety assessment. Pressures, temperatures, and concentrations occurring during operation have a direct impact on safety and are meanwhile calculated with the aid of process simulations. In this publication, the open-source process simulator DWSIM [22] is used. Besides the advantage of being open source it also has an XML export function of the simulation results. This allows for easy extraction of process data into the information model relevant for the preHAZOP. Another advantage is the free availability of the process simulator DWSIM, which means that the preHAZOP tool can also be used by anyone using freely accessible input formats.

2.4 Graphs

A graph represents a data structure consisting of nodes and edges. Nodes represent objects or properties, whereas edges are interactions or connections between the individual objects. When using graphs, a distinction is made between undirected and directed graphs, the latter also being called digraph. If it is a digraph, a direction is assigned to the edges [23]. For the representation or exchange of graphs, the XML-based file format GraphML (see Fig. 1) [24] will

be used. The nodes and edges are listed as the only elements of the file, with the option to store more detailed information about nodes and edges in the form of attributes under the corresponding elements. In this way it is possible to store relevant data specifically in the graph in addition to the topology of the graph [25].

3 Suggested Workflow

The automated safety assessment described in this paper can be divided into two parts: data collection and the following preHAZOP. The workflow is shown in detail in Fig. 2. Starting points are a P&ID in the standardized DEXPI [16] format, which contains a machine-readable and well-documented plant topology. In addition, the simulation concerning the process is used. The required results are documented as XML export and are generated with the open-source process simulator DWSIM. Within the data collection a graph is generated from the DEXPI P&ID with the help of a function DEXPI2graph. The graphs contain the plant topology and specifications and are supplemented with relevant process data at the relevant points by mapping from the simulation. The generated graphs are then processed in preHAZOP, where previously defined scenarios are detected by a search algorithm and evaluated by a downstream risk assessment. The automatically generated HAZOP results are output in the form of a table.

In the following, both the data collection and the preHAZOP are described in more detail. This includes an explanation and detailed analysis of the data basis and the algorithms developed.

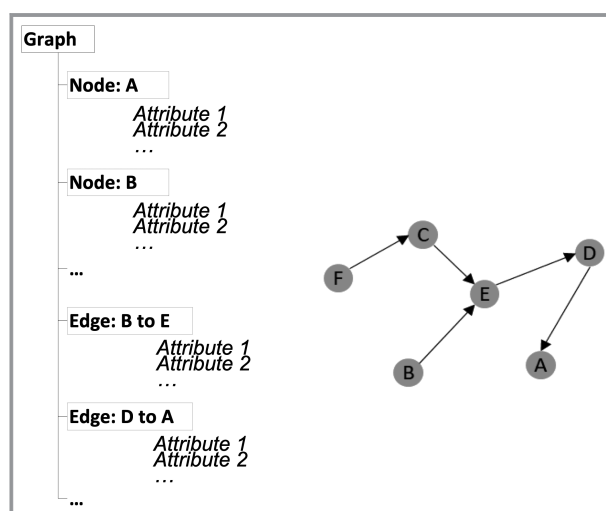


Figure 1. General structure of a GraphML file representing a directed graph.

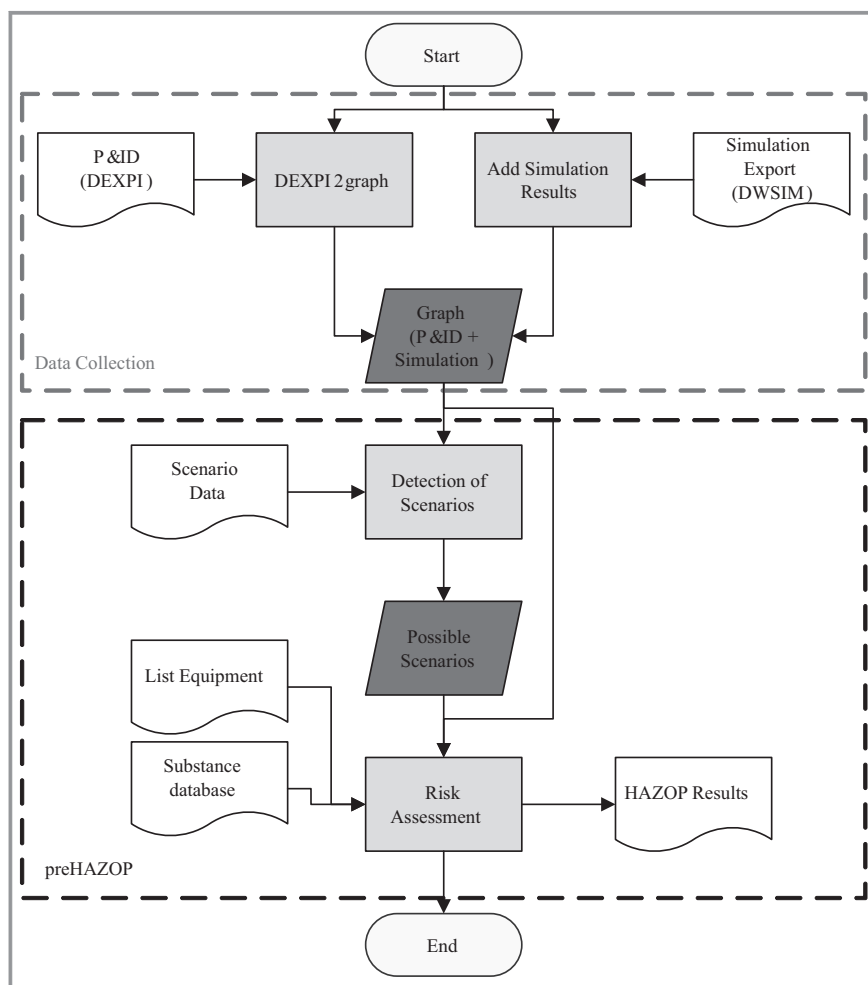


Figure 2. Workflow of the automated, graph-based preHAZOP based on standardized DEXPI P&IDs and DWSIM process simulation exports.

3.1 Data Collection

To enable a reliable safety assessment, it is important that both plant data from P&IDs and process data from process simulations are included in the evaluation. An automated safety assessment therefore requires a data basis that contains both P&ID and process data in a machine-readable form. Within the scope of this publication, a graph-based information model is generated, which is based on a machine-readable DEXPI [16] P&ID, which is enriched with important process information from a simulation using DWSIM [26]. The detailed creation of the information model is described in more detail below.

3.1.1 P&ID Data

The data collection takes place using a DEXPI2graphML converter, which was already shown in a previous publication [27]. The P&ID stored in the DEXPI format is described by three information levels, which are used to

extract the topology. *Equipment*, which contains the information of all components of the plant. The *Piping-NetworkSystem* describing the piping and piping components, as well as the *InstrumentationFunction* describing the signal connections and process control. All relevant plant information is extracted according to [27] and serves as a baseline for the preHAZOP presented in this publication. The P&ID graphs generated by the DEXPI2graphML converter [27] are subjected to further processing to obtain the most efficient generalization adapted to safety assessments. In particular, the processing of equipment groups, the detection of safety-relevant functional assemblies and the processing of measuring points are important.

Functional Groups

To be able to perform an automated safety analysis based on graphs, it is important to group equipment, valves and connections with regard to their safety-relevant function (see Fig. 3). In this way, safety-critical scenarios can usually be transferred to components with the same technical process function. For this reason, equipment with the same function is grouped together to enable particularly good accessibility for the deterministic algorithms of the preHAZOP. In case of valves, a more

differentiated classification is required. For example, it is important to classify whether the valve is used to operate the system or if it has a specific safety-relevant function. If this is the case, the existing valve is additionally grouped with regard to its function into the classes *safety valve*, *check valve*, or *breather valve*. Connections within the graph can be divided into *piping*, *signal connection* and *process connection*. In the case of a *piping*, attributes are assigned to *main pipe*, *heat transfer pipe* and *secondary pipe* regarding the respective function.

Detection of Process Control Loops

The following describes how the existing process control technology is processed within the graph. The function marked according to DIN 19227-1 is taken from each measuring device and stored temporarily. Subsequently, the measuring location of a measuring device must be identified. For a better understanding Fig. 4 shows an example.

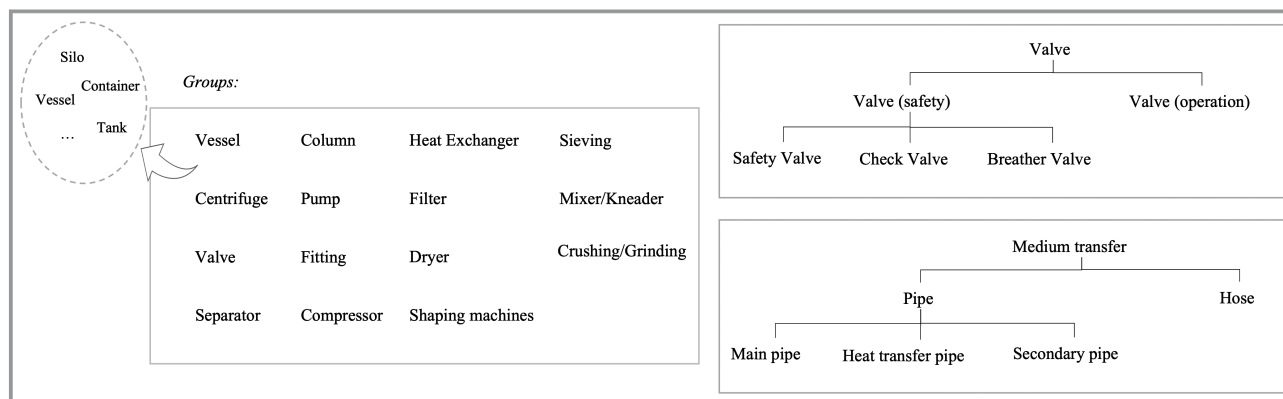


Figure 3. List of equipment groups (left), valve groups (top right) and piping groups (bottom right).

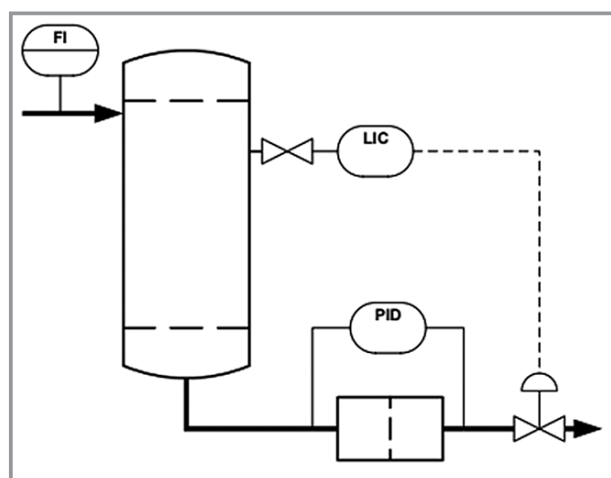


Figure 4. Example of the detection of process control equipment.

For identification, the sensor is always considered as the potential measurement location, while intermediate valves, which are used for installation and removal, are skipped. In addition, in the case of double process connection, the component between both connection nodes is selected (see pressure drop measurement in Fig. 4). A special case is the measurement at a pipeline or at a T-piece, flange or similar. To keep the information of the measurement for upstream and downstream components, the storing of the information takes place for both components, where additionally the attributes *Inlet* and *Outlet* are added. The flow indicator in Fig. 4 measures the flow in the inlet of the column, while the node of the columns is assigned with the information "Inlet Flow Indicator". All functions of a sensor are processed analogously and passed to the identified nodes as attribute *Measurements*. To consider signal lines, the first step is to identify common actuators such as pumps and valves with incoming signal lines. Subsequently, all functions of the corresponding sensor are read out. This is limited to the signal-specific functions *shut off* and *control*,

because possible other functions exclusively affect just the measurement location. For example, Fig. 4 shows a valve, which controls the level of the column. This results in an entry *Level Control* for the attribute *Signals* in the graph. Since the process control technology can play a central role in the safe operation of a plant (functional safety), the attribute *Safeguards* additionally stores all entries of the attributes *Signals* and *Measurements* for the respective nodes as a safety precaution. Only the indication as well as the recording of measurement parameters are not considered as *Safeguards*, because these functions only monitor the process.

Detection of Existing Safety Devices

The protection of equipment against overpressure is an important part of safety engineering. Common safety devices such as rupture discs and overpressure valves are used for this purpose. The automated data collection identifies devices that are protected against overpressure due to a corresponding fitting in the graph and stores the information in the attribute *Safeguards*. Equipment is identified, which is connected to overpressure fittings via pipes. It is important that only T-pieces or rupture discs are located between the safety fitting and the equipment since other components such as valves lead to a deactivation of the safety mechanism in a closed position.

3.1.2 Process Data

For a reliable risk assessment, it is important to use information of used substances, which is not part of a P&ID by default. In order to provide machine-readable process data in addition to the plant data, the graph is therefore extended to include process data from a DWSIM simulation. The results of the simulation are available as an XML file. It is important that the components of the simulation flow-sheet receive the same tags as the components in the P&ID in order to be able to map the process values to the respective nodes in the graph. In addition, the connections in the simulation must be defined by a name containing the start

and end equipment. By matching labels, it is explicit which node represents which component in the graph. By linking P&ID and simulation, process values can be specifically assigned to individual nodes and edges. First, properties of the material streams are queried and entered the graph in corresponding attributes in order to generate as much information as possible about the process medium (temperature, pressure, substances, density, liquid fraction, vapor fraction, solid fraction). Only data from substances that make up more than 1 mol% of the process medium are extracted, since in the subsequent risk assessment it is assumed that highly diluted substances do not influence the hazard level due to their low impact.

The data collection provides a graph adapted for the safety analysis via the preHAZOP application. In addition to the topology of the plant as well as equipment specifications, this graph contains information on operation conditions and substance data that is relevant for subsequent risk analysis. A detailed description of the graph and its attributes are shown in the Supporting Information (SI, Fig. S1, Tabs. S1, S2).

3.2 preHAZOP

The preHAZOP consists of two parts. In the first part, possible hazardous scenarios will be detected. In the second part a risk assessment is carried out for the detected scenarios. The detailed workflow of the preHAZOP is shown in the following chapters.

3.2.1 Scenario Detection

To perform automated scenario detection, it is important to use a machine-readable and uniform database. This database is applied to the graph-based P&IDs using a deterministic search algorithm. All hierarchically structured information, which are essential for the detection, are shown in Fig. 5 and consist of the three sections *HAZOP-scenario*, *Detection* and *Evaluation*. The individual sections are explained in more detail below.

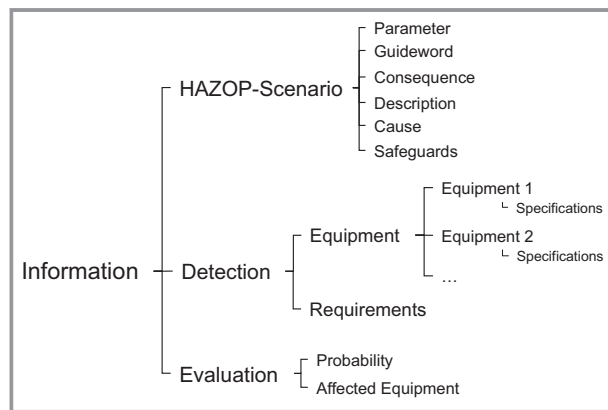


Figure 5. Structure of the standardized scenario database.

HAZOP Scenario

Part of the information is the preconceived scenario with all standard entries, as they are carried out in a HAZOP analysis. These are used for the general description of deviations but also to specify safety measures that can be specifically searched for later on in the graph, thus enabling the automated detection of safety measures that are already in place. The relations between HAZOP scenario, equipment and evaluation are shown in the SI (Tab. S3).

Detection

It is also necessary to have information that can be used to judge whether the scenario under consideration will occur or whether the P&ID graph under investigation can meet the requirements. It is crucial that the necessary equipment for the occurrence of the scenario is available. The simple example in Fig. 6 illustrates the scenario detection mechanism. The sheer presence or absence of manual valves upstream and downstream of a centrifugal pump completely changes the risk profile of this equipment. With existing closing valves, scenarios such as high pressure or low/no flow become possible at all. The algorithm detects particularities in the P&ID graph and assigns the appropriate preconceived scenarios.

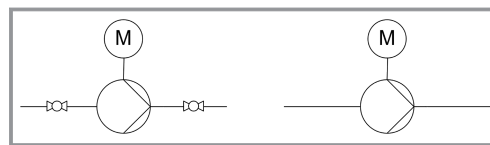


Figure 6. Example of a safety-relevant deviation in P&IDs.

In case more than one component is necessary, the topology has to be considered as well. The equipment must be available in a certain arrangement, which can be defined by numbering. In this way it is possible to search specifically for a linear structure with a defined sequence. The restriction to a linear structure is usually not a problem, because in this work common scenarios, which are rarely bound to complex structures, are considered. By giving specifications, it is possible to presuppose a certain property that the equipment must exhibit. If one or more conditions that the scenario entails are not testable, these must also be included. These can be conditions that are not apparent from the data or conditions whose query is not integrated in the expert tool. By listing those under *Requirements* (see Fig. 5), the validity of the scenario can later be checked manually by the user.

Scenarios are detected by automatically comparing the equipment or equipment sequences stored in the scenario database with the graph-based P&ID. The combinations of components relevant to the scenarios are automatically identified in the P&ID and assigned to the predefined scenarios.

Evaluation

All other information is used for the risk assessment of a scenario, where the consequence, which represents the central information, can already be taken from the specifications of the preHAZOP scenario section. This information does not necessarily indicate which equipment is actually affected by the consequence; hence, this information must be provided additionally (*Affected Equipment*). In addition, the probability of occurrence is one of the two central parameters that directly influence the risk. Based on common scenarios, it is possible to estimate the probability of occurrence, provided that the deviation can be traced back to a common cause. The prediction of the extent, however, is not possible across plants, as it depends on many factors that vary from process to process.

3.2.2 Scenario Database

All previously listed information required for the preHAZOP are stored in a machine-readable table (.xlsx, .csv), which represents the scenario database (see SI, Tab. S3).

The good accessibility of the spreadsheet enables the user to add own scenarios and to extend the functionality of the scenario detection in this way. For a better understanding of the table, an introductory example is given in Fig. 7. The table provides all essential information for a scenario. In addition, Fig. 7 shows a P&ID that represents the example preHAZOP scenario of the table.

The index ensures the explicit traceability of the scenarios. The description in the second column in combination with the entries *Guideword*, *Parameter*, *Cause*, and *Consequence* defined according to HAZOP standard [11] sufficiently explain the present scenario. In the example, the inert system of a vessel B1 fails, causing a vacuum when the contents are pumped out by the pump P1, which in turn can lead to denting of the vessel. To be able to perform an automated risk assessment later, the entries in the *Consequence* field are limited to *Damage*, *Seal Leakage*, *Leakage* and *Rupture*. All possible entries within the *Scenarios* database are shown in the SI (Tab. S6). Another column is also provided to specify common safety measures. To ensure

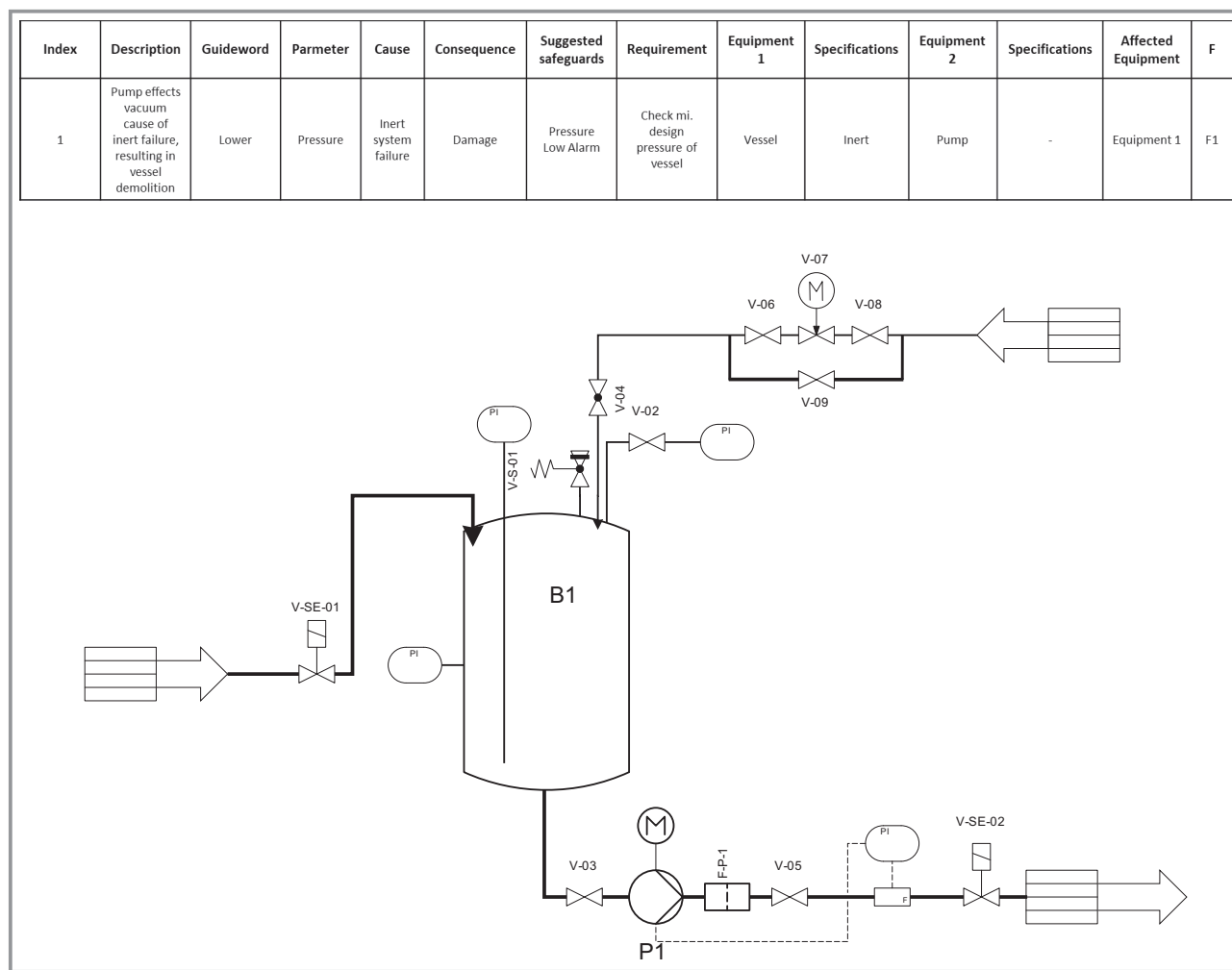


Figure 7. Example of the standardized preHAZOP scenario description (top) and P&ID of the example preHAZOP scenario (bottom).

recognition in the graph, the entries are similar to those listed in the graph under the *Safeguards* attribute. The entry under *Requirements* indicates that it is necessary to check whether the vessel is sufficiently designed for the possible vacuum. Subsequently, the required equipment is specified, which enables the preHAZOP algorithm to detect the scenarios in the P&ID. Here, *Equipment_1*, *Equipment_2* and *Equipment_3* represent sequentially consecutive components, with the *Affected_Equipment* indicating which component experiences the consequence arising from the scenario. In this example, a pump (*Equipment_2*) follows a vessel (*Equipment_1*), which has the specification of an inert system. Finally, each scenario has the information of the probability of occurrence, which is relevant for the subsequent risk assessment.

3.2.3 Risk Assessment

As part of the safety analysis executed by the preHAZOP tool, the detected hazard scenarios are additionally evaluated in an initial risk assessment. The baseline for this is the results table from the previous scenario detection, in which the respective probability of occurrence is already defined. To estimate the risk, the exact extent of damage for the respective scenario must be determined in the following. Since the risk assessment is an automated process, the classification of the extent of damage is based on quantitative criteria. Based on the previously defined consequences, a distinction is made between hazards with substance leakage (*leakage*, *seal leakage*, *rupture*) and hazards with additional damage to the component (*damage*).

Leakage, Seal Leakage, Rupture

In the case of a substance leakage, the way in which the leakage takes place is of crucial importance for the extent. This information is already provided by the formulation of the consequence. In addition to the form of leakage, it is important which substances leak out or which substances are contained in the affected equipment. This information is derived from the stored simulation data and can be easily accessed from the corresponding attributes of the extracted P&ID graph. To enable the preHAZOP tool to assess the hazards resulting from the substances, it contains a substance database. This database is shown in the SI (Tab. S4) and provides information about the hazard classes as well as a corresponding signal word (danger, warning, -). If required, the database can be extended by the user. Since process streams are mostly a mixture of substances, the substance that shows the greatest danger is used for risk assessment.

Furthermore, the quantity of the hazardous substance leaking has a crucial influence on the extent of the consequence. In this context, it is important to distinguish between a continuous and a batch plant. In the case of a batch process, the leaking quantity is to be set equal to the content within the equipment. It should be noted that in some cases the quantity in the upstream equipment is more important

than the quantity in the equipment itself. An example of this is a valve or pump that has a leak. If a vessel is connected upstream, the entire volume of the vessel may leak. For this reason, the quantity of the upstream equipment is considered in such cases. On the other hand, if the plant is being operated continuously, the potential quantity leaking is determined by the mass flow rate. In this case, a substance leakage of 20 min is assumed, which allows for calculating the leakage quantity [28, 29]. Based on this information, the leaking volume (S0–S4) can be classified automatically and justifiably with the help of the matrices shown in Tab. 2. The extent of the damage can then be determined using Tab. 3. If the determination of the damage extent is not unambiguous, the larger damage extent is used for the risk assessment.

Table 2. Matrix of the preHAZOP tool to categorize the extent of damage (signal words for different leakage amounts in kilogram, regarding to [10]).

	0–4.5	4.5–45	45–450	450– 4500	4500– 45 000	> 45 000
Danger	S3	S2	S1	S1	S0	S0
Warning	S4	S3	S2	S2	S1	S0
–	S4	S4	S4	S3	S3	S2

Table 3. Matrix of the preHAZOP to categorize the extent of damage (signal words for different leakage types, regarding to [10]).

	Seal Leakage	Leakage	Rupture
Danger	S2	S1	S0
Warning	S3	S2	S1
–	S4	S3	S2

Damage

For the evaluation of a damage of an equipment, the incurring costs are used in the preHAZOP since these can be determined particularly reliably. Since the development of the plant is usually advanced during the preHAZOP, it is assumed that the equipment costs are known and available to the preHAZOP tool as a predefined list. In combination with the limit values shown in Tab. 4, an assessment of the extent of damage is made. To simplify the estimation, the worst-case scenario is assumed that the entire equipment will be damaged irreparably. Since damage to the equipment is always accompanied by substance leakage, the resulting leakage is also considered according to the procedure described above.

Together with the specified probability of occurrence, the risk is finally classified using a risk matrix (see Tab. 5) that distinguishes between high and tolerable. If the risk assessment result cannot be determined automatically, it is possible to make the decision manually. The reason for this can

Table 4. Extend of damage regarding cost limits.

Loss [€]	Severity
0–1000	S4
1000–10 000	S3
10 000–100 000	S2
100 000–1 000 000	S1
> 1 000 000	S0

Table 5. Risk matrix to evaluate the detected scenario results.

	S4	S3	S2	S1	S0
F0	Tolerable	High	High	High	High
F1	Tolerable	Tolerable	High	High	High
F2	Tolerable	Tolerable	High	High	High
F3	Tolerable	Tolerable	Tolerable	High	High
F4	Tolerable	Tolerable	Tolerable	Tolerable	High
F5	Tolerable	Tolerable	Tolerable	Tolerable	Tolerable

be unknown consequences for the preHAZOP tool or missing information about the probability of occurrence. The evaluation of the risk provides information about the necessity and the required reliability of the previously developed safety measures.

3.3 Graphical User Interface

For a user-friendly application, a graphical user interface shown in Fig. 8 is available. During the application, only the required files (P&ID in DEXPI format, DWSIM export of the simulation results) have to be selected. After the start, the progress of the automated preHAZOP can be tracked by loading bar. After completion, a corresponding message

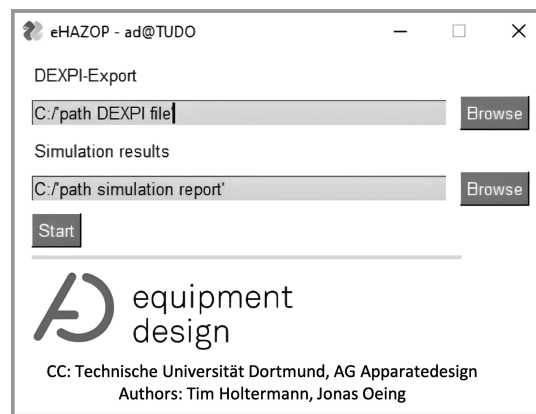


Figure 8. GUI of the preHAZOP tool.

appears. The results can be accessed in the folder structure of the expert tool under *Output*. There the user will find the file *HAZOP-results.xlsx*, which contains the results and an additional error log. Furthermore, a GraphML file of the processed P&ID can be accessed.

4 Validation

In the following, the preHAZOP will be validated. Therefore, the preHAZOP tool is applied to a distillation plant, which separates a mixture of ethanol and water. The P&ID of the distillation plant is shown in Fig. 9 and contains 242 components and 269 connections. In the plant used for validation, a feed stream of 5000 kg h⁻¹ with an ethanol content of 78 wt % is used. The separation is performed at normal pressure (1.013 bar). In the distillation plant, temperatures between 25 and 100 °C occur.

The process is simulated with DWSIM [31] using the unit operation *Distillation Column*. An overview of the simulated process flow diagram (PFD) is shown in Fig. 10. The used property package is *NRTL* [32].

To validate the preHAZOP tool, the user loads the P&ID (Fig. 9) in DEXPI format as well as the DWSIM simulation export (Fig. 10) using the developed preHAZOP GUI. Fig. 11 shows the automatically created results. In the following different areas of the results will be investigated and analyzed regarding their quality. Thus, the results of the scenario detection, the risk assessment as well as the influence of the substances used are analyzed.

Scenario Detection – Validation

As the preceding chapters have shown, the safety assessment and thus its results can be divided into scenario detection and the subsequent risk assessment. Due to different equipment configurations, some scenarios are detected more often than others. Thus, the results show 21 individual scenarios, which can be divided into ten different scenario types. A closer look shows that nonsensical constellations are avoided. Each hazard scenario indicated by the preHAZOP can initially be confirmed by examining the underlying P&ID in Fig. 9. The detailed description of the scenario in combination with the equipment involved helps to ensure clear recognition. In addition, the relevant equipment is highlighted in red in the P&ID. It is important to note the information under *Requirements* in the results table. For example, the scenarios listed with the index 8.1–8.3 require a check of the maximum permitted operating temperature with regard to the increased inlet temperature. In the case of cooling power failure, as described in the scenarios, the increased temperature amounts to 78 °C in the distillate stream and to almost 100 °C in the sump stream. Since the vessels each have a maximum permitted operating temperature of 150 °C, leakage of the substances is not to be expected. The user can adjust the results table accordingly. If the results are compared with the scenario database (see SI, Tab. S3), scenarios that are not listed can be identified. The

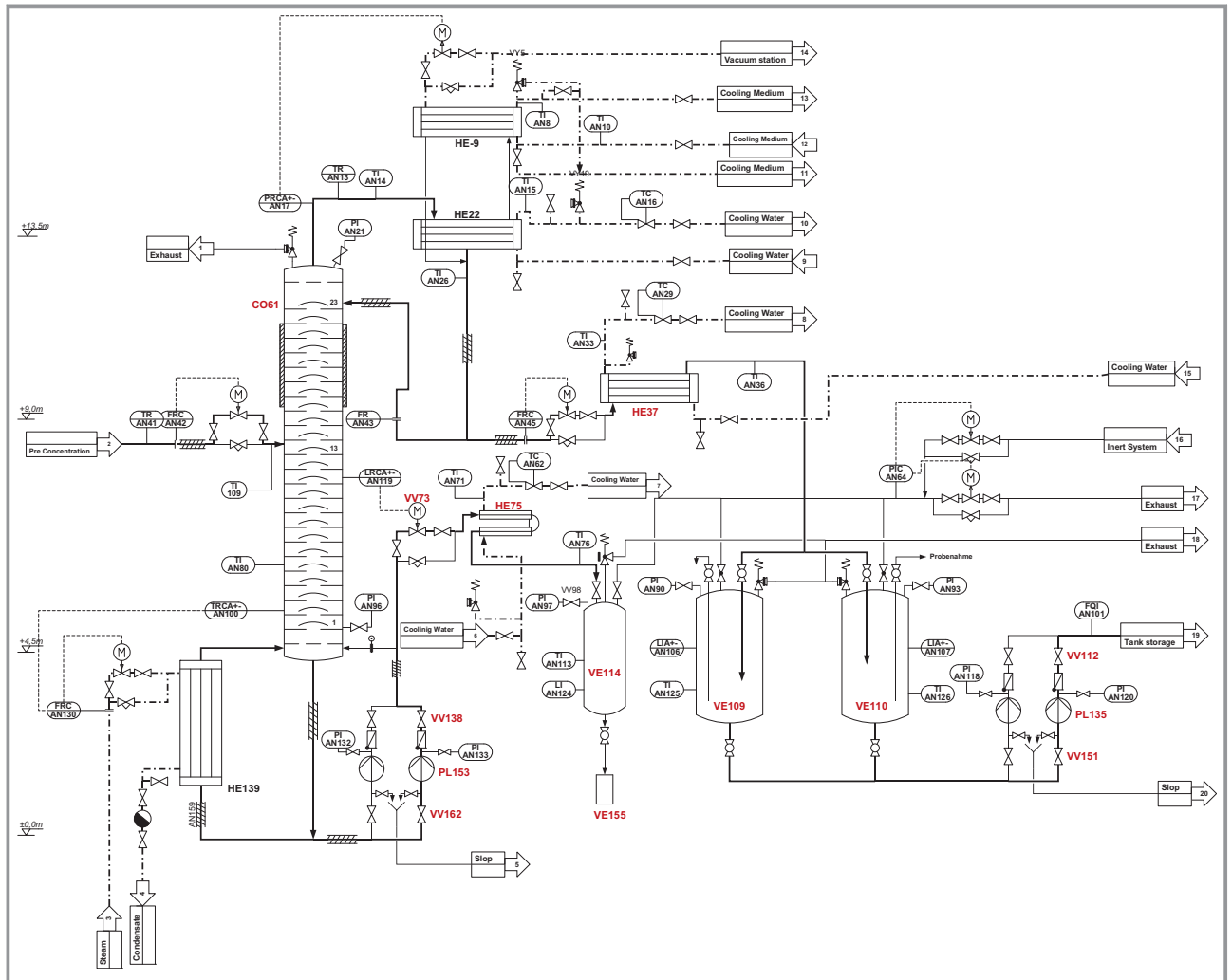


Figure 9. P&ID of a distillation plant regarding [30] used for validation of the preHAZOP (red color shows equipment detected in preHAZOP scenarios).

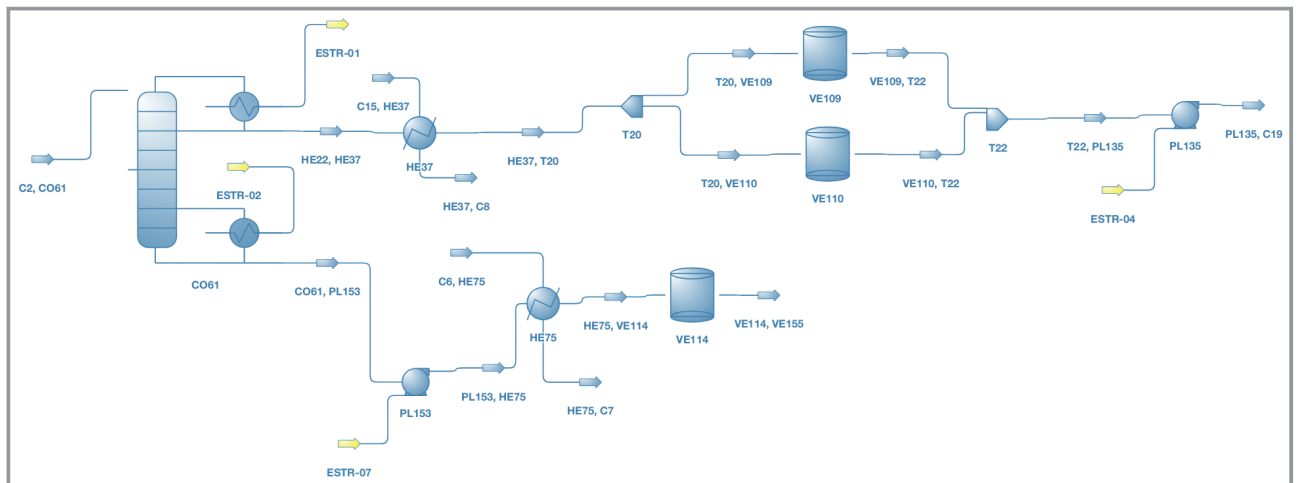


Figure 10. PFD simulation of the distillation plant used for validation.

Index	Description	Deviation	Cause	Consequence	Involved Equipment	Requirements	Suggested Safeguards	Existing Safeguards	Potencial Safeguards	Substances	F	S	Risk
1.1	Pump effects vacuum cause of inert failure, resulting in vessel demolition	Lower Pressure	Inert system failure	Damage	VE109, PL135	Check min. operation pressure of vessel	Pressure Low Shut down		VE109: Safety valve, Inert system, Level Low Alarm, ... // PL135: Redundant		F1	S2	High
1.2	Pump effects vacuum cause of inert failure, resulting in vessel demolition	Lower Pressure	Inert system failure	Damage	VE110, PL135	Check min. operation pressure of vessel	Pressure Low Shut down		VE110: Safety valve, Inert system, Level Low Alarm, ... // PL135: Redundant		F1	S2	High
2.1	Pumping against wrongly closed valve results in damage of pump	Higher Temperature	Valve wrongly closed	Damage	PL153, VV138		Bypass		PL153: Redundant		F1	S3	Tolerable
2.2	Pumping against wrongly closed valve results in damage of pump	Higher Temperature	Valve wrongly closed	Damage	PL135, VV112		Bypass		PL135: Redundant		F1	S3	Tolerable
3.1	Pumping against closed level control valve results in damage of pump	Higher Temperature	Valve closed by control	Damage	CO61, PL153, VV73		Bypass	Bypass (VV73)	CO61: Safety valve, Outlet Pressure Low Alarm, ... // PL153: Redundant// VV73: Level Control		F1	S3	Tolerable
4.1	Substance leaks out of pump cause of mechanical seal wear	Other Flow	Mechanical seal wear	Seal leakage	PL153		Maintenance		PL153: Redundant	Water	F1	S4	Tolerable
4.2	Substance leaks out of pump cause of mechanical seal wear	Other Flow	Mechanical seal wear	Seal leakage	PL135		Maintenance		PL135: Redundant	Ethanol, Water	F1	S1	High
5.1	Mechanical failure of the pump	No Rotation	Mechanical failure	Damage	PL153		Maintenance, Redundant	Redundant (PL153)	PL153: Redundant		F1	S3	Tolerable
5.2	Mechanical failure of the pump	No Rotation	Mechanical failure	Damage	PL135		Maintenance, Redundant	Redundant (PL135)	PL135: Redundant		F1	S3	Tolerable
8.1	Cooling circuit failure leads to a higher inlet temperature, resulting in a seal leakage of vessel	Higher Temperature	Cooling circuit fails	Seal Leakage	HE37, VE109	Check max. operation temperature of vessel	Inlet Temperature High Alarm		HE37: Safety valve, Inlet Flow Control, ... // VE109: Safety valve, Inert system, ...	Ethanol, Water	F1	S1	High
8.2	Cooling circuit failure leads to a higher inlet temperature, resulting in a seal leakage of vessel	Higher Temperature	Cooling circuit fails	Seal Leakage	HE37, VE110	Check max. operation temperature of vessel	Inlet Temperature High Alarm		HE37: Safety valve, Inlet Flow Control, ... // VE110: Safety valve, Inert system, ...	Ethanol, Water	F1	S1	High
8.3	Cooling circuit failure leads to a higher inlet temperature, resulting in a seal leakage of vessel	Higher Temperature	Cooling circuit fails	Seal Leakage	HE75, VE114	Check max. operation temperature of vessel	Inlet Temperature High Alarm		VE114: Safety valve, Inert system, Outlet Pressure Control	Water	F1	S4	Tolerable
10.1	Pump is turned on when there is no medium in vessel, so it runs dry	No Flow	Wrongly turned on	Damage	VE109, PL135			Level Low Alarm (VE109)	VE109: Safety valve, Inert system, Level Low Alarm, ... // PL135: Redundant		F1	S3	Tolerable
10.2	Pump is turned on when there is no medium in vessel, so it runs dry	No Flow	Wrongly turned on	Damage	VE110, PL135			Level Low Alarm (VE110)	VE110: Safety valve, Inert system, Level Low Alarm ... // PL135: Redundant		F1	S3	Tolerable
11.1	Pump is turned on when there is no medium in Column, so it runs dry	No Flow	Wrongly turned on	Damage	CO61, PL153			Level Low Alarm (CO61)	CO61: Safety valve, Outlet Pressure Low Alarm, Temperature High Alarm... // PL153: Redundant		F1	S3	Tolerable
14.1	Pump runs dry cause of a wrongly closed valve in front	No Flow	Wrongly closed valve	Damage	VV162, PL153		Temperature High Shut down		PL153: Redundant		F1	S3	Tolerable
14.2	Pump runs dry cause of a wrongly closed valve in front	No Flow	Wrongly closed valve	Damage	VV151, PL135		Temperature High Shut down		PL135: Redundant		F1	S3	Tolerable
15.1	An external fire effects a vessel rupture	Higher Pressure	External fire	Rupture	VE109	Check content and max. operation pressure		Safety valve (VE109)	VE109: Safety valve, Inert system, Level Low Alarm, Outlet Pressure Control, Level High Alarm	Ethanol, Water	F3	S0	High
15.2	An external fire effects a vessel rupture	Higher Pressure	External fire	Rupture	VE110	Check content and max. operation pressure		Safety valve (VE110)	VE110: Safety valve, Inert system, Level Low Alarm, Outlet Pressure Control, Level High Alarm	Ethanol, Water	F3	S0	High
15.3	An external fire effects a vessel rupture	Higher Pressure	External fire	Rupture	VE114	Check content and max. operation pressure		Safety valve (VE114)	VE114: Safety valve, Inert system, Outlet Pressure Control	Water	F3	S2	Tolerable
15.4	An external fire effects a vessel rupture	Higher Pressure	External fire	Rupture	VE155	Check content and max. operation pressure		Safety valve		Water	F3	S2	Tolerable

Figure 11. preHAZOP results with detected scenarios for the investigated distillation plant with a mixture of ethanol and water.

fact that these scenarios were previously excluded by the expert tool becomes clear upon closer examination of the omitted scenarios (index 6, 7, 9, 12, 13) in the database. The two similar scenarios 8 and 9 can be used as examples for

the justified selection of scenarios. Both require a heat exchanger upstream of a vessel, but one scenario calls for a heating operation while the other calls for a cooling operation. The expert tool recognizes the upstream heat exchang-

ers and identifies each as a cooling operation, which means that only the scenario with index 8 is listed as a potential hazard in the results table (see Fig. 11).

Risk Assessment – Validation

In the following, a closer look at the entries that can be traced back to the risk assessment is made. Each detected scenario has been evaluated in terms of risk and its magnitude has been estimated. The assessment provides information about the necessity and required reliability of the safety measures by describing the risk with the entry tolerable or high. The central function on which the assessment is based is the automated classification of the extent and is verified in the following by a close examination of the results (see Fig. 11). This is aided by the tables, which are used by the expert tool and shown previously or in the SI: evaluation matrices, Tabs. 2, 3), substance database (Tab. S4), equipment cost list (Tab. S5) and financial limits (Tab. 4).

If equipment is damaged, the extent is determined based on the financial loss. In the hazard scenarios detected, the damage is mostly limited to pumps, which turn out to be comparatively inexpensive based on the equipment cost list. Accordingly, the extent and thus the risk is low after observing the financial limits. The situation is different for the scenarios with the index 1.1 and 1.2, in which a vessel is damaged in each case.

In addition to equipment damage, scenarios involving substance leakage are also represented in the results. The decisive factors are the type of leakage, the possible leakage quantity and the substances. It must be proved that the preHAZOP tool considers the mentioned influencing factors accordingly. While the plant shows the hazardous mixture of ethanol and water in the distillate stream after the distillation, it carries negligible amounts of ethanol in the sump stream. Accordingly, the substances that leak out are listed in the results table (for the sake of clarity, the hazard classes are omitted). A clear correlation between the leaking substances and the resulting risk assessment can be seen. The hazardous substance ethanol is recognized as such based on the substance database and considered in the assessment, which justifies the high risk. A closer look at the extent reveals the influence of leakage quantity and form. The bursting of a container (index 15.1–15.4) entails a high hazard regardless of the substance that escapes. Accordingly, the magnitude is relatively high even for the escape of water with S2. At this point, the influence of the probability of occurrence is also noticeable, which means that the risk is not considered high. The influence of the quantity becomes clear in the case of the leakage of ethanol via seal leakage (index 4.2, 8.1, 8.2). A seal leak initially suggests a comparatively low magnitude, but here the potential leakage quantity is determinant. Quantities calculated based on mass flow are in the range of 450–4500 kg. In combination with the hazardous substance ethanol, this results in the critical hazard level S1 (see Tab. 2). The method for determining the leakage quantity is described in detail in Sect. 3.2.3.

There it is also described that in the event of a larger leak, as is the case with the bursting of the containers, the financial extent due to the associated damage is additionally considered by the preHAZOP. However, as in this case, the substance leakage remains the determining factor.

5 Conclusion and Outlook

Performing safety assessments and HAZOP studies is an arduous and difficult task for process engineering teams during process and plant design. The expert tool preHAZOP can perform an automated safety analysis based on existing knowledge prepared by AI methods. For this purpose, graph-based DEXPI flow diagrams are used and expanded with process information such as pressure, temperature, material mixture, etc. from DWSIM PFD simulation data. In comparison with a previously defined and expandable database, it was shown that safety-critical scenarios can be identified in the P&ID. With the help of substance data and definitions of limits regarding extent of damage and costs, it is also possible to carry out an initial risk assessment. The potential of the preHAZOP tool unfolds mainly in the early engineering phase. While an initial P&ID is being prepared, preHAZOP can already identify potential hazards and reduce design errors, thus, costs, too.

It is important for the authors to clarify that the present publication is a prototype for feasibility studies. To achieve reliable results, the scenario database needs to be extended by expert knowledge, therefore, care was taken during development to ensure easy accessibility by the user. Currently, the safety of the plant is often analyzed at equipment level. Therefore, it is not guaranteed that the weakest component will always be reliably detected. For this reason, it would be advantageous in the future to identify and treat integral pressure chambers in addition to individual components. In this way, it would be possible to better detect the weakest component of the pressure compartment in the event of a developing overload pressure and to recommend appropriate countermeasures at this point.

Supporting Information

Supporting Information for this article can be found under DOI: <https://doi.org/10.1002/cite.202200222>.

Acknowledgment

The research work has been developed within the KEEN project (support code: 01MK20014S) and has been funded by the BMWK (Federal Ministry of Economic Affairs and Climate Action). Open access funding enabled and organized by Projekt DEAL.

Abbreviations

AI	Artificial Intelligence
DEXPI	Data Exchange in Process Industry
GUI	Graphical User Interface
HAZOP	Hazard and Operability Analysis
PFD	Process Flow Diagram
P&ID	Piping and Instrumentation Diagram
NRTL	Non-random two-liquid mode

References

- [1] H. R. Greenberg, J. J. Cramer, *Risk Assessment and Risk Management for the Chemical Process Industry*, John Wiley & Sons, New York 1991.
- [2] N. Kockmann, P. Thenée, C. Fleischer-Trebes, G. Laudadio, T. Noël, *React. Chem. Eng.* **2017**, 2 (3), 258–280. DOI: <https://doi.org/10.1039/C7RE00021A>
- [3] A. Vaughen, Bruce; Hurban, David; First, Kenneth; Ness, *Process Saf. Prog.* **2020**, 39 (2), e12142. DOI: <https://doi.org/10.1002/prs.12142>
- [4] J. I. Single, J. Schmidt, J. Denecke, *J. Loss Prev. Process Ind.* **2019**, 62, 103952. DOI: <https://doi.org/10.1016/j.jlp.2019.103952>
- [5] L. Cui, J. Zhao, R. Zhang, *Process Saf. Environ. Prot.* **2010**, 88 (5), 327–334. DOI: <https://doi.org/10.1016/j.psep.2010.04.002>
- [6] T. Schmidberger, T. Scherf, A. Fay, *Automatisierungstech. Prax.* **2007**, 49 (6), 46–53.
- [7] M. Wiedau, G. Tolksdorf, J. Oeing, N. Kockmann, *Chem. Ing. Tech.* **2021**, 93 (12), 2105–2115. DOI: <https://doi.org/10.1002/cite.202100203>
- [8] N. Kockmann, *ChemBioEng Rev.* **2016**, 3 (1), 5–15. DOI: <https://doi.org/10.1002/cben.201500025>
- [9] L. Hohmann, K. Kössl, N. Kockmann, G. Schembecker, C. Bramsiepe, *Chem. Eng. Process.* **2017**, 111, 115–126. DOI: <https://doi.org/10.1016/j.cep.2016.09.017>
- [10] IVSS Sektion Chemie, *Risikobeurteilung in Der Anlagensicherheit*, 5th ed., Berufsgenossenschaft Rohstoffe und chemische Industrie, Heidelberg **2020**.
- [11] International Electrotechnical Commission, *IEC 61882:2016 – Hazard and Operability Studies (HAZOP Studies) – Application Guide*, VDE Verlag **2016**.
- [12] IVSS Sektion Chemie, *Gefahrenermittlung Und Gefahrenbewertung in Der Anlagensicherheit - Praxisbewerte Methoden*, 2nd ed. **2012**.
- [13] T. Kletz, *Learning from Accidents*, 3rd ed., Routledge, New York **2007**.
- [14] T. Kletz, P. Amyotte, *What Went Wrong? – Case Histories of Process Plant Disasters and How They Could Have Been Avoided*, 6th ed., Butterworth-Heinemann, Oxford **2017**.
- [15] T. Kletz, *Hazop and Hazan*, CRC Press, Boca Raton, FL **2018**.
- [16] M. Theißen, M. Wiedau, *DEXPI – P&ID Specification*, **2021**. <https://dexpi.org/specifications/>
- [17] ISO 15926-2, *Industrial Automation Systems and Integration – Integration of Life-Cycle Data for Process Plants Including Oil and Gas Production Facilities – Part 2: Data Model*, Beuth Verlag, Berlin **2013**.
- [18] ISO 10628-2, *Diagrams for the Chemical and Petrochemical Industry – Part 2: Graphical Symbols*, International Organization for Standardization, Geneva **2012**.
- [19] IEC 62424, *Representation of Process Control Engineering – Requests in P&I Diagrams and Data Exchange between P&ID Tools and PCE-CAE Tools*, International Electrotechnical Commission, Geneva **2016**.
- [20] ISO 10209, *Technical Product Documentation – Vocabulary – Terms Relating to Technical Drawings, Product Definition and Related Documentation*, International Organization for Standardization, Geneva **2012**.
- [21] *Proteus Schema for P&ID Exchange*, **2017**. <https://github.com/ProteusXML/proteusxml>
- [22] D. Wagner Oliveira de Medeiros, *DWSIM – The Open Source Process Simulator*, **2022**. <https://dwsim.org>
- [23] V. Turau, C. Weyer, *Algorithmische Graphentheorie*, De Gruyter, Berlin **2015**.
- [24] *GraphML specification*, **2017**. <http://graphml.graphdrawing.org/specification/dtd.html>
- [25] R. Tamassia, *Handbook of Graph Drawing and Visualization (Discrete Mathematics and Its Applications)*, CRC Press, Boca Raton, FL **2013**.
- [26] D. Wagner, *DWSIM – The Open Source Process Simulator*, **2022**. <https://dwsim.org>
- [27] J. Oeing, W. Welscher, N. Krink, L. Jansen, F. Henke, N. Kockmann, *Digital Chem. Eng.* **2022**, 4, 100038. DOI: <https://doi.org/10.1016/j.dche.2022.100038>
- [28] *12. Verordnung Zur Durchführung des Immisionsschutzgesetzes*, Bundesministerium der Justiz, Bundesanzeiger Verlag, Köln **2017**.
- [29] *Guideline 2012/18/EU*, European Parliament and European Council, Brussels **2012**.
- [30] A. Behr, D. W. Agar, J. Jörissen, A. J. Vorholt, *Einführung in die Technische Chemie*, 2nd ed., Spektrum, Heidelberg **2016**.
- [31] D. Wagner, *DWSIM Process Simulator*, **2022**.
- [32] H. Renon, J. M. Prausnitz, *AIChE J.* **1968**, 14 (1), 135–144. DOI: <https://doi.org/10.1002/aic.690140124>