

Dok-Nr. 114121

SUMMARY

There is one common aspect to the automated calculation of energy performance indicators, the estimation of the service life of a critical wearing part and the control and optimisation of the technological processes involved: all these digital applications are based on data. The adaptation of data-driven applications to the requirements of specific plants is however often a tedious business. The VDZ working group "Standardising data in cement plants" has been addressing this particular issue since September 2020 and has now produced a catalogue of requirements for the standardised storage of data in the future. This has been further supplemented by a theoretical study on the use of data in machine learning applications. A central aspect of this concept is the integration of machine-processable information into the database used for storage. Information on the process structure and on the relationships between individual data points facilitates access to the data and permits the use of a standard interface. More extensive assessment and evaluation of the data is also possible in combination with additional information. Initial potential solutions were developed and put to the test in a simple database. The catalogue of requirements and the approach employed are presented in the following article. ◀

ZUSAMMENFASSUNG

Egal ob es um die automatisierte Berechnung von Energiekennzahlen geht, die Lebensdauerabschätzung eines kritischen Verschleißteils oder die Regelung und Optimierung der verfahrenstechnischen Prozesse: Daten stellen die Grundlage für alle diese digitalen Anwendungen dar. Oftmals müssen dabei jedoch datengetriebene Anwendung mühsam an Werksstandorte angepasst werden. Seit September 2020 setzt sich der VDZ-Arbeitskreis „Standardisierung von Daten im Zementwerk“ gezielt mit diesem Problem auseinander und hat ein Anforderungsprofil für die künftige, standardisierte Speicherung von Daten erarbeitet. Dieses wurde im Weiteren durch eine theoretische Studie zur Nutzung von Daten in Anwendungen des maschinellen Lernens erweitert. Ein Kerngedanke dieses Konzepts ist die Integration von maschinenverwertbaren Prozessinformation in die zur Speicherung eingesetzte Datenbank. Informationen zum Prozessaufbau und zu den Bezügen einzelner Datenpunkte erleichtern den Zugriff auf Daten und ermöglichen den Einsatz einer einheitlichen Schnittstelle. In Kombination mit zusätzlichen Informationen wird zudem eine erweiterte Be- und Auswertung der Daten möglich. Erste Lösungskonzepte wurden entwickelt und in einer einfachen Test-Datenbank überprüft. Der Beitrag stellt das Anforderungsprofil sowie den entwickelten Lösungsansatz vor. ◀

Standardising data in cement plants

Standardisierung von Daten im Zementwerk

1 Introduction

Projects concerning the generation and handling of digital data and in particular the evaluation and utilisation of such data are currently being implemented in many areas of industry and society under the broad headings of "Digitalisation" and "Industry 4.0". The dramatic increase in the volume of data, together with the computing power and sophisticated evaluation methods now available, has created new options which could bring about fundamental changes in many areas of industry. The status of digitalisation in the cement industry was examined and assessed in a study performed by RWTH Aachen University back in 2019 [1]. This revealed that the cement production processes are already virtually fully digitised. There is however still potential for development with regard to utilisation of the data. A key aspect of digitalisation is the automation of data processing to reduce the number of repetitive and non-creative operations, for instance in the form of the software-assisted analysis of data and creation of indicators. Of central importance with regard to subsequent predictive capability are the fields of Artificial Intelligence (AI) and machine learning methods. Such algorithms provide possible solutions for many operations within the cement industry, including process control, digitally assisted maintenance and product quality monitoring for example.

Data management, that is the compilation and provision of data from various sources, is a particularly demanding process. It involves extracting data from source systems by way of interfaces, transforming it to the format of the target application and then loading it via a suitable interface [2]. At this point the quality of the data is already subject to the requirements of the subsequent applications, specifically in terms of "completeness, freedom from error, unequivocal interpretability, current status, precise accuracy of time stamps and standardised presentation wherever technically feasible" [2]. The specialist knowledge of the process owners is required for both quality assurance and structuring of the data.

At present, relevant data is compiled in different databases by the cement plants themselves, by equipment suppliers or by service providers. These often only apply to specific parts of a plant and do not exhibit a uniform standard. In addition, unrestricted access to the (raw) data may not always be possible depending on the provider. Another point to note is that the extraction of the data from the processes, its transformation and provision for a given application have to be individually implemented for every single solution. This also makes it difficult to achieve an integrated quality assurance concept for the data. The development of a uniform standard tailored to the technological processes is therefore of crucial significance for data management at cement plants. Furthermore it would be desirable for companies to be able to fully comprehend the data stored and control the provi-

1 Einleitung

Unter den weit gefassten Überbegriffen der „Digitalisierung“ und „Industrie 4.0“ werden aktuell an vielen Stellen in Industrie und Gesellschaft Projekte vorangetrieben, die die Generierung und den Umgang mit digitalen Daten sowie insbesondere deren Auswertung und Nutzung betrachten. Die rasante Zunahme der Datenmenge in Verbindung mit der Verfügbarkeit von Rechenleistung und ausgereiften Methoden zur Auswertung haben dabei Möglichkeiten geschaffen, die in vielen Bereichen der Industrie grundlegende Veränderungen herbeiführen können. Bereits im Jahre 2019 wurde im Rahmen einer Studie der RWTH Aachen der Stand der Digitalisierung in der Zementindustrie untersucht und bewertet [1]. Es zeigte sich, dass die Prozesse der Zementherstellung bereits nahezu vollständig digitalisiert sind. Bei der Nutzung der Daten besteht jedoch noch Entwicklungspotenzial. Ein Kerngedanke der Digitalisierung stellt dabei die Automatisierung der Datenverarbeitung dar, um repetitive und nicht-kreative Arbeitsschritte zu reduzieren, z.B. dass Daten softwaregestützt analysiert und zu Kennzahlen verdichtet werden. Zentrale Bedeutung kommt in Hinblick auf die daran anschließende Prognosefähigkeit insbesondere dem Thema der Künstlichen Intelligenz (KI) bzw. den Methoden des Maschinellen Lernens zu. Diese Algorithmen bieten Lösungsansätze für viele Fragestellungen der Zementindustrie, darunter z.B. zur Prozesssteuerung, der digital unterstützten Instandhaltung oder der Überwachung der Produktqualität.

Besonders aufwändig ist die Datenbewirtschaftung, also das Zusammenführen und Bereitstellen von Daten aus unterschiedlichen Quellen. Dabei werden die Daten über Schnittstellen aus den Quellsystemen extrahiert, in das Format der Zielanwendung überführt und dort über eine geeignete Schnittstelle eingeladen [2]. Dabei sind bereits die Anforderungen späterer Anwendungen an die Qualität der Daten, konkret an „die Vollständigkeit, die Fehlerfreiheit, die eindeutige Auslegbarkeit, die Aktualität, die exakte Richtigkeit von Zeitstempeln und die einheitliche Darstellung, soweit dies technisch realisierbar ist“ [2]. Die Qualitätssicherung wie auch die Strukturierung der Daten erfordern dabei das fachliche Wissen der Prozesseigner.

Relevante Daten werden aktuell von den Zementwerken selbst, von Anlagenbauern oder Dienstleistern in unterschiedlichen Datenbanken zusammengeführt. Diese decken in vielen Fällen nur Teilbereiche des Werkes ab und weisen keinen einheitlichen Standard auf. Hinzu kommt, dass je nach Anbieter der Zugriff auf die (Roh-)Daten nicht uneingeschränkt möglich ist. Weiterhin ist zu bedenken, dass die Extraktion der Daten aus den Prozessen, die Transformation und die Bereitstellung für eine Anwendung bei jeder Einzelschritt individuell realisiert werden muss. Vor diesem Hintergrund ist es zudem schwierig, die Qualität der Daten mit einem ganzheitlichen Konzept sicherzustellen. Daher kommt

sion of this – particularly with regard to external partners. The added value of a corresponding solution should however mean more than just standardisation. Which was the reason why the VDZ working group "Standardising data in cement plants" supported the development of a concept for the storage, commenting and networking of data at cement plants [3].

2 Requirements for the standardisation of data

The principal challenge associated with the standardisation of data at cement plants is the fundamental conflict between new requirements, from machine learning applications for example, and the established access paths of existing users. New applications involve new and modified requirements, whereas existing users have to rely on finding data at familiar locations and under familiar names. The retention of established identifiers is thus an important criterion for new storage concepts, as user acceptance is an essential prerequisite for the successful introduction of a standard. Careful consideration was therefore given to both existing identification systems and databases as well as future data utilisation scenarios with a view to developing a storage solution which is suitable for everyday use but at the same time future-proof.

2.1 Analysis of existing storage solutions

The main aim of standardised data storage is to create uniform access to the data. The identification of data points is thus a crucial factor. Great significance is however also to be attached to the incorporation of additional information to simplify access to and the interpretation of the data and to permit sustainable quality assurance. This requires in-depth consideration of not just the identification but also of the structure of the databases used for storage.

2.1.1 Identification systems

One of the best known process technology identification systems is the identification system for power stations (KKS). This is based on combinations of 15 to 17 letters and numbers which use breakdown levels ranging from the total plant to individual items of equipment to provide systematic identification [4]. The system offers a wide range of options but is scarcely intuitive and also extremely extensive. The identification systems used in the cement industry have a similar hierarchical structure, but are often based on numerical codes alone which number the identified elements in ascending order in the direction of the material flow. ► Fig. 1 (top) shows the structure of a typical identification system based on the example of a raw meal grinding system [3]. The plants are generally described on a "storage to storage" basis. Material is removed from a silo, for example, to be followed by transportation, processing (e.g. grinding) and further transportation of the material before arriving at a silo again. The storage of materials thus typically represents the boundary between the plant sections. The individual elements are usually numbered in ascending order in the direction of the material flow on each of the hierarchy levels.

This fundamental structure is common to numerous identification systems. Differences are however found primarily in cases where this identification concept is not unequivocal. One typical example of this is the preheater: The plant elements are numbered in ascending order in the direction of the material flow downwards along the preheater and via

der Entwicklung eines einheitlichen, auf die verfahrenstechnischen Prozesse zugeschnittenen Standards für die Datenbewirtschaftung im Zementwerk eine große Bedeutung zu. Weiterhin ist es wünschenswert, dass Unternehmen ihre gespeicherten Daten vollständig verstehen und ihre Bereitstellung vor allem hin zu externen Partnern – kontrollieren können. Der Mehrwert einer entsprechenden Lösung sollte aber nicht allein in der Standardisierung liegen. Daher wurde unterstützt durch den VDZ-Arbeitskreis „Standardisierung von Daten im Zementwerk“ ein Konzept für die Speicherung, Kommentierung und Vernetzung von Daten im Zementwerk entwickelt [3].

2 Anforderungen an die Standardisierung von Daten

Die wesentliche Herausforderung bei der Standardisierung von Daten im Zementwerk stellt der grundsätzliche Konflikt zwischen neuen Anforderungen wie z.B. aus Anwendungen des Maschinellen Lernens und den etablierten Zugriffswegen der Bestandsnutzer dar. Neue Anwendungen stellen neue und veränderte Anforderungen, während Bestandsnutzer darauf angewiesen sind, Daten an bekannten Stellen und unter bekannten Namen wiederzufinden. Entsprechend ist der Erhalt von etablierten Bezeichnungen eine wichtige Randbedingung für neue Speicherkonzepte, denn die Akzeptanz der Nutzer ist eine wesentliche Voraussetzung für die erfolgreiche Einführung eines Standards. Um entsprechend eine „alltagsverträgliche“ aber dennoch „zukunftsorientierte“ Speicherlösung zu entwickeln, wurden sowohl bestehenden Bezeichnungssysteme und Datenbanken wie auch Szenarien zur künftigen Nutzung von Daten sorgsam betrachtet.

2.1 Analyse von bestehenden Speicherlösungen

Das Ziel einer standardisierten Speicherung von Daten ist im Wesentlichen, einen einheitlichen Zugriff auf die Daten zu schaffen. Damit spielt die Bezeichnung von Datenpunkten eine entscheidende Rolle. Von großer Bedeutung ist aber auch die Einbindung von zusätzlichen Informationen, die den Zugriff und die Interpretation der Daten vereinfachen und eine nachhaltige Qualitätssicherung ermöglichen. Dazu muss neben der Bezeichnung auch die Struktur der Datenbanken zur Speicherung grundlegend betrachtet werden.

2.1.1 Bezeichnungssysteme

Eines der bekanntesten prozesstechnischen Bezeichnungssysteme ist das Kraftwerks-Kennzeichensystem (KKS). Dieses basiert auf 15- bis 17-stelligen Buchstaben und Ziffern-Kombinationen, die über Gliederungsstufen von der Gesamtanlage bis hin zum Betriebsmittel eine systematische Bezeichnung erlauben [4]. Das System bietet umfangreiche Möglichkeiten ist aber wenig intuitiv und sehr umfangreich. Die in der Zementindustrie eingesetzten Bezeichnungssysteme folgen grundsätzlich einem ähnlichen hierarchischen Aufbau, basieren aber oft nur auf numerischen Codes, die die bezeichneten Elemente in Richtung des Materialflusses hochzählen. ► Bild 1 zeigt oben am Beispiel einer Rohmahlanlage den Aufbau eines typischen Bezeichnungssystems [3]. Dabei werden die Anlagen in der Regel von „Lager zu Lager“ beschrieben. Material wird z.B. einem Silo entnommen, es folgen ein Transport, eine Verarbeitung (z.B. Mahlung) und ein weiterer Transport, bevor das Material wieder ein Lager erreicht. Die Lagerung von Materialien stellt somit typischerweise die Abgrenzung zwischen den Anlagenteilen dar. In den jeweiligen Hierarchieebenen werden die einzelnen Elemente üblicherweise in Richtung des Materialflusses hochgezählt.

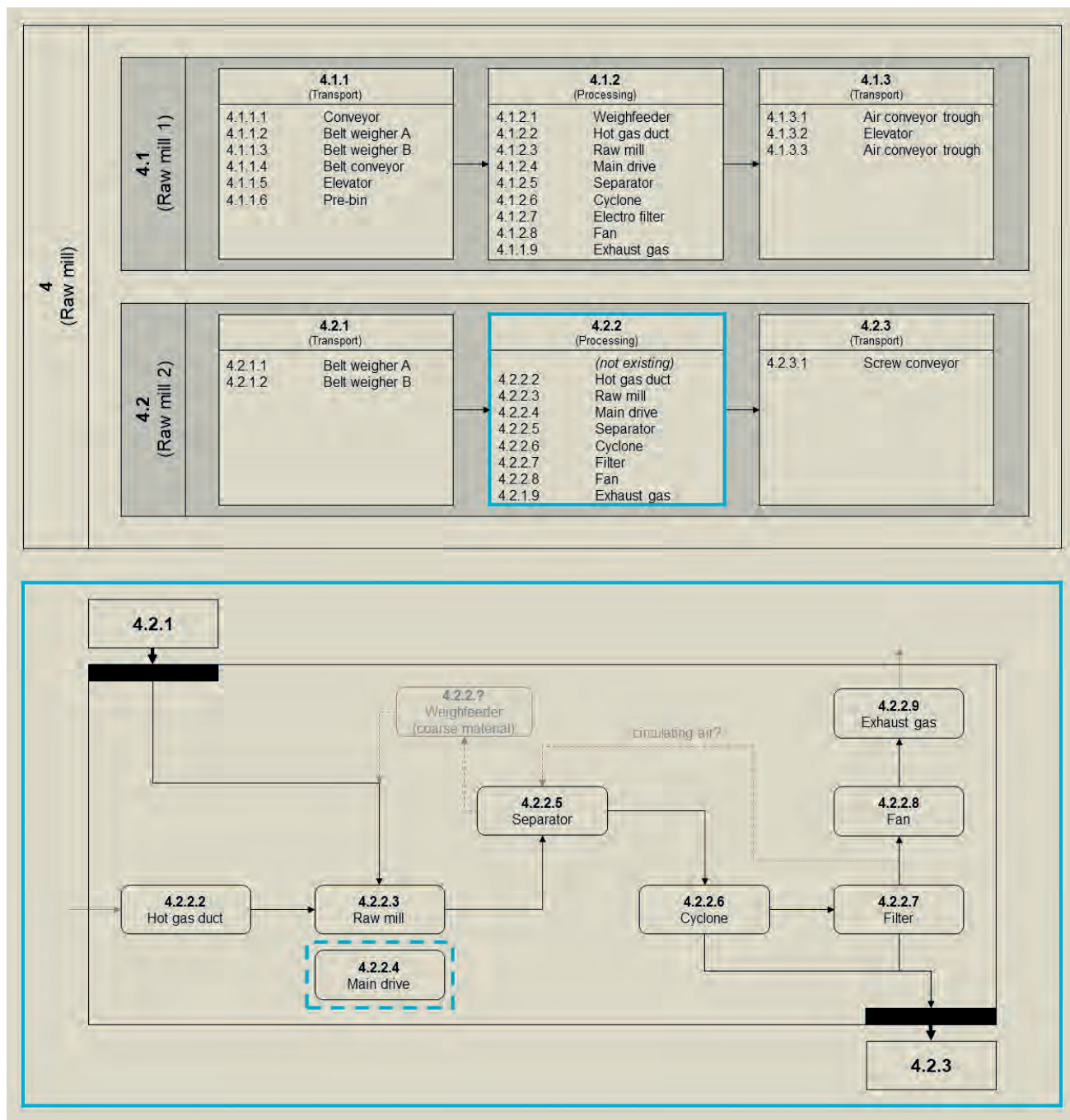


Figure 1: Example of numerical identifiers for plant components (top) including exemplary visualisation of cement mill (bottom) [3]

Bild 1: Beispiel für die numerische Bezeichnung von Anlagenteilen und -elementen (oben) mit exemplarischer Visualisierung der Zementmühle (unten) [3]

the kiln to the cooler. The path downstream of the preheater followed by the counter-current exhaust gas flow cannot however be definitely identified using this logic. The numbering of these elements then often continues in ascending order following on from the numbering of the cooler. Further examples are to be found in the case of grinding plants operating with a classifier in a closed circuit. Here again the simple counting method does not provide consistent identification of the transportation of the returned material as shown in Fig. 1 (bottom). The unequivocal numbering of elements not located in the material flow is also not possible. One example of this would be the positioning of the mill main drive in relation to the mill itself. As a result of these many necessary exceptions, the identification systems used in practice differ greatly from plant to plant despite having a very similar structure. It would be possible to create a uniform system by defining a broad set of rules. Such a solution

Diese grundlegende Struktur ist vielen Bezeichnungssystemen gleich. Unterschiede treten jedoch vor allem dort auf, wo dieser Bezeichnungsansatz nicht eindeutig ist. Ein typisches Beispiel ist der Vorwärmer: Die Anlagenelemente werden in Richtung des Materialflusses entlang des Vorwärmers nach unten und über den Ofen zum Kühler hochgezählt. Der Gasweg nach Vorwärmer, dem das im Gegenstrom fließende Abgas folgt, kann von dieser Logik jedoch nicht eindeutig erfasst werden. Diese Elemente werden dann oft im Anschluss an die Nummerierung des Kühlers weiter hochgezählt. Weitere Beispiele finden sich auch bei Mahlanlagen, die mit Sichter im geschlossenen Kreislauf arbeiten. Auch hier liefert die einfache Zählweise keine konsistente Bezeichnung des Rückguttransports wie in Bild 1 unten dargestellt. Weiterhin können Elemente, die sich nicht im Materialfluss befinden, ebenfalls nicht eindeutig nummeriert werden. Ein Beispiel hierfür wäre die Positionierung des Mühlenhaupt-

would not however offer any added value above and beyond standardisation.

2.1.2 Incorporation of storage solutions

Both the location and method of data storage also differ from plant to plant. In many cases process, laboratory and energy data is stored and called up separately in different systems. Solutions in which data from various sources is stored in a single database independently of the process are already state of the art. These have however not yet been implemented at all plants. At present no standard specifications exist with regard to the structure of such a database or for the data input and output interfaces required.

These databases continue to make use of familiar identification systems. For example, specific known data points are usually accessed. Furthermore, effective automated evaluation tools already exist which are able to perform rule-based calculation of indicators from the data. For instance it is possible to determine the specific energy demand for cement grinding according to type by combining laboratory, process and energy data.

2.2 Potential utilisation of data in the future

The data required for various applications was investigated by VDZ within the context of a theoretical study. As well as simple applications for the presentation and analysis of data, this also considered machine learning applications in the field of process optimisation and control as well as for maintenance. Analysis yielded three central requirement categories: Data quality, data interpretability and data networking.

Data quality: The first aspect of this is the accuracy of the individual values, confirmed by a separately transmitted "Quality Code" for example. Further criteria are the completeness and clarity of the data. A classic example in this regard is the question as to what is to be transmitted when a system is at rest: "0", no value at all or a minimum value > 0 .

Interpretability: The capacity for interpretation of the data (e.g. being able to assign step-changes in measured values to calibration or maintenance) is one of the requirements for data quality. The associated provision of additional, commenting information – referred to in the following as meta-data – goes one step further still. Metadata refers to any information necessary for the assessment and evaluation of data. Depending on the level, this encompasses data on the plant, the sensor or the actuator itself, data on changes in the system and, if applicable, references to further data sources. Relationships with other plant sections and elements may also be stored in the metadata. This forms the link with the third requirement category.

Networking: To be able to combine data from different sources, it generally has to be possible to link the data by way of a uniform time stamp. In doing so, particular attention has to be paid to the influence of possible temporal aggregation of the data. Data networking in the context of cause and effect or order in the process flow (predecessor and successor) would however also be conceivable. As with the metadata, this makes it possible to incorporate process knowledge into storage of the data.

Alongside Artificial Intelligence (AI) applications, a further priority should be efficient user access to the stored data.

antriebs relativ zur Mühle. Durch diese vielen notwendigen Ausnahmen unterscheiden sich die in der Praxis genutzten Bezeichnungssysteme deutlich von Werk zu Werk, auch wenn sie strukturell sehr ähnlich aufgebaut sind. Durch die Definition einer Vielzahl von Regeln könnte hier ein einheitliches System geschaffen werden. Diese Lösung bietet jedoch über die Standardisierung hinaus keinen weiteren Mehrwert.

2.1.2 Einbindung von Speicherlösungen

Wo und wie Daten gespeichert werden unterscheidet sich ebenfalls von Werk zu Werk. Es gibt viele Fälle in denen Prozess-, Labor oder Energiedaten in unterschiedlichen Systemen unabhängig voneinander gespeichert und abgerufen werden. Stand der Technik sind bereits Lösungen, die Daten aus verschiedenen Quellen prozessunabhängig in einer Datenbank speichern. Diese sind aktuell aber noch nicht in allen Werken implementiert. Für den Aufbau einer derartigen Datenbank sowie für die Schnittstellen zum Dateninput und -output existieren aktuell keine einheitlichen Vorgaben.

Innerhalb dieser Datenbanken kommen weiterhin bekannte Bezeichnungssysteme zum Einsatz. So erfolgt der Zugriff zumeist gezielt auf bekannte Datenpunkte. Weiterhin gibt es bereits effektive automatisierte Auswertungswerkzeuge, die regelbasiert Kennzahlen aus den Daten errechnen können. So kann z.B. der spezifische Energiebedarf für die Zementmahlung durch Kombination von Labor-, Prozess- und Energiedaten sortenbereinigt bestimmt werden.

2.2 Perspektivische Nutzung von Daten

Im Rahmen einer theoretischen Studie wurde der Datenbedarf verschiedener Anwendungen durch den VDZ untersucht. Betrachtet wurden einfache Anwendungen zur Darstellung und Analyse von Daten aber auch Anwendungen des maschinellen Lernens im Bereich der Prozessoptimierung und -steuerung sowie in der Instandhaltung. Aus der Analyse ergeben sich drei zentrale Anforderungsbereiche: Datenqualität, Interpretierbarkeit von Daten und die Vernetzung von Daten.

Datenqualität: Zunächst fällt hierunter die Richtigkeit der Einzelwerte z.B. bestätigt durch einen separat übermittelten „Quality Code“. Hinzu kommt aber auch die Vollständigkeit und die Eindeutigkeit der Daten. Klassisches Beispiel ist hier die Frage, was übermittelt wird, wenn eine Anlage steht: „0“, kein Wert oder ein Minimalwert > 0 .

Interpretierbarkeit: Die Möglichkeit Daten interpretieren zu können (z. B. Sprünge in Messwerten einer Kalibrierung oder Wartung zuordnen zu können) gehört zu den Anforderungen an die Datenqualität. Die damit verbundene Bereitstellung von zusätzlichen, kommentierenden Informationen – im Folgenden als Meta-Daten bezeichnet – geht aber noch einen Schritt weiter. Meta-Daten können alle für die Be- und Auswertung von Daten notwendigen Informationen sein. Dies umfasst je nach Ebene Daten über die Anlage, den Sensor oder Aktor selbst, Daten zu Veränderungen im System und ggfs. auch Verweise zu weiterführenden Datenquellen. Weiterhin können in den Meta-Daten auch Bezüge zu anderen Anlagenteilen und -elementen hinterlegt sein. Dies stellt die Verbindung zum dritten Anforderungsbereich dar.

Vernetzung: Um Daten aus unterschiedlichen Quellen verbinden zu können, müssen diese in der Regel über einen einheitlichen Zeitstempel miteinander verknüpft werden kön-

The introduction of new data viewing perspectives above and beyond current practice would be an added advantage. At present data tends to be organised on the basis of process technology aspects (e.g. structured in accordance with the plant and its sub-systems and numbered in line with the material flow). Other viewing perspectives (points of view) are however conceivable at cement plants. In the context of energy management, consideration is from the point of view of the power supply and looks at the electric meters and then the possible factors influencing energy demand. Maintenance is considered from the point of view of the plant and looks at assemblies and wearing parts, as well as possible indicators as to the service life of these. Such considerations call for new hierarchy levels and relationships between the data points.

2.3 Catalogue of requirements

Altogether the preceding observations thus give rise to a series of requirements for a standardised data storage solution for cement plants:

- ▶ The aim is a solution that is independent from the process, i.e. stand-alone database(s) in accordance with the recommendations of VDI/VDE 3714 [2].
- ▶ It should be possible to access the database by way of a clearly defined, uniform open interface.
- ▶ Established identification systems should be retained for existing users (possibly in parallel with a standardised means of identification).
- ▶ In combination with appropriate data analysis applications, the storage solution should permit simplified access paths without the need for detailed knowledge of the plant or of the identification system.
- ▶ Access to the data should be possible from various points of view (such as process technology, energy management, maintenance).
- ▶ For this purpose as well as for further applications, the aim is to achieve the integration of machine-readable process information (plant structure and relationships between elements).
- ▶ An important general requirement for the storage solution is assurance of the data quality. This includes the necessity for completeness, freedom from error, unequivocal interpretability and the current status of the data.
- ▶ The precise accuracy of time stamps also has to be ensured to achieve clear networking of the data. In this context rules are also required for the correct and meaningful temporal aggregation of the data.
- ▶ Metadata have to be stored in an appropriate manner for the purposes of interpretation, quality assurance and further applications. This refers in particular to the documentation of changes in the process or to the sensors (calibration). Plant data required for further calculations should additionally be stored.
- ▶ Data on relevant external context information from other data sources (e.g. engineering drawings or illustrations) should be accessible via corresponding links.
- ▶ All the requirements outlined here are to be assessed with regard to outlay and benefits in the course of technical implementation.

3 Structure of a storage solution

The standardisation of data at cement plants requires more than just a uniform identification system. The ultimate goal must be to achieve a standard interface allowing the straight-

nen. Dabei ist insbesondere der Einfluss einer eventuellen zeitlichen Aggregation von Daten zu beachten. Eine Vernetzung von Daten ist jedoch auch im Sinne von Ursache und Wirkung oder einer Ordnung im Prozessfluss (Vorgänger und Nachfolger) denkbar. Hier besteht, wie auch bei den Meta-Daten, die Möglichkeit, Prozesswissen in die Speicherung der Daten einfließen zu lassen.

Neben Anwendungen der Künstlichen Intelligenz (KI) sollte vor allem ein effizienter Nutzerzugriff auf gespeicherte Daten im Vordergrund stehen. Ein Mehrwert über aktuelle Betrachtungsweisen hinaus stellt dabei die Einführung neuer Sichtweisen auf die Daten dar. Aktuell werden Daten grundsätzlich nach verfahrenstechnischen Gesichtspunkten geordnet (z. B. über die Anlage und ihre Teilsysteme strukturiert und im Materialfluss nummeriert). Im Zementwerk sind aber auch andere Sichtweisen denkbar. Im Sinne des Energiemanagements wird von der Energieversorgung über die Zähler zu den Verbrauchern und weiter auf die möglichen Einflussgrößen auf den Verbrauch geschaut. Die Instandhaltung blickt von der Anlage auf Baugruppen und deren Verschleißteile sowie auf mögliche Indikatoren für deren Lebensdauer. Diese Betrachtungen erfordern neue Hierarchieebenen und Bezüge zwischen den Datenpunkten.

2.3 Anforderungsprofil

Zusammenfassend ergibt sich daher aus den vorangehenden Betrachtungen eine Reihe von Anforderungen an eine standardisierte Lösung zur Speicherung von Daten im Zementwerk:

- ▶ Es wird eine prozessunabhängige Lösung, d.h. eigenständige Datenbank(en) gemäß den Empfehlungen der VDI/VDE 3714 [2] angestrebt.
- ▶ Der Zugriff auf die Datenbank soll über eine klar definierte, einheitliche und offene Schnittstelle möglich sein.
- ▶ Etablierten Bezeichnungssysteme sollen für Bestandsnutzer weiterhin erhalten bleiben (ggfs. parallel zu einer standardisierten Bezeichnungsweise).
- ▶ Die Speicherlösung soll in Verbindung mit entsprechenden Anwendungen zur Datenanalyse vereinfachte Zugriffswege ohne Detailkenntnis der Anlage und der Bezeichnungssystematik möglich machen.
- ▶ Der Zugriff auf die Daten soll aus verschiedenen Sichtweisen möglich sein (u.a. Verfahrenstechnik, Energiemanagement, Instandhaltung).
- ▶ Hierzu und auch für weiterführende Anwendungen wird die Integration von maschinenlesbaren Prozessinformationen (Anlagenstruktur und Beziehungen zwischen Elementen) angestrebt.
- ▶ Die Sicherung der Datenqualität stellt eine wesentliche allgemeine Anforderung an die Speicherlösung dar. Dazu zählen Anforderungen an die Vollständigkeit, die Fehlerfreiheit, die eindeutige Auslegbarkeit und die Aktualität der Daten.
- ▶ Zur eindeutigen Vernetzung der Daten ist ferner die exakte Richtigkeit von Zeitstempeln sicherzustellen. Vor diesem Hintergrund bedarf es auch Regeln für die richtige und sinnvolle zeitliche Aggregation der Daten.
- ▶ Zur Interpretation, Qualitätssicherung und für weiterführende Anwendungen sind in sinnvoller Art und Weise Meta-Daten zu speichern. Hierzu zählt vor allem die Dokumentation von Veränderungen im Prozess oder an den Sensoren (Kalibrierung). Weiterhin sollten auch Anlagen-daten, die für weitere Berechnungen erforderlich sind abgelegt werden.

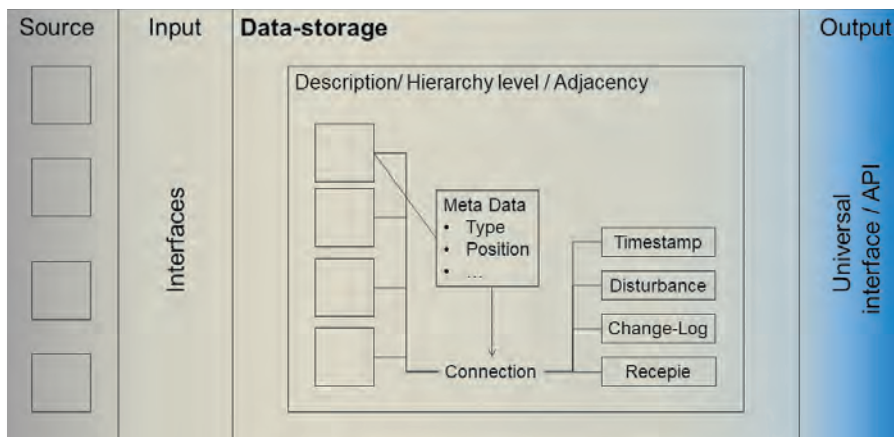


Figure 2: Concept for data storage in cement plants

Bild 2: Struktur einer Lösung zur Speicherung von Daten im Zementwerk

forward start-up of applications regardless of the location. As shown in Fig. 2 this would however necessitate not just an interface, but also solutions for data input and an intelligent concept for the underlying organisation of the data. In the course of the studies a test database was therefore set up to successively conduct trials on the individual solution modules [3]. Use was made of PostgreSQL [5] for this purpose. The basic principles of the solution, but not the implementation details, are explained in the following.

3.1 Hierarchy levels

In a manner similar to that of the identification systems analysed (s. Fig. 1), the plant structure is modelled in the database over several hierarchy levels. By interlinking various levels, an access path initially independent of any concrete identification of the data point is created to the data. Specific elements (subsequently also referred to as nodes in keeping with the nomenclature of graph theory) additionally make it possible to store information on each hierarchy level. Sensors or actuators, the measured data of which are ultimately stored in the form of time series, can be flexibly assigned to different levels. A joint query over several hierarchy levels is

» Daten zu relevanten externen Kontextinformationen aus anderen Datenquellen (z.B. technische Zeichnungen oder Bilder) sind durch entsprechende Links zu verknüpfen.

» Alle hier skizzierten Anforderungen sind im Zuge einer technischen Umsetzung hinsichtlich Aufwand und Nutzen zu bewerten.

3 Struktur einer Speicherlösung

Für die Standardisierung von Daten im Zementwerk wird mehr als nur ein einheitliches Bezeichnungssystem benötigt. Das Ergebnis muss

am Ende eine einheitliche Schnittstelle sein, die es erlaubt, Anwendungen unabhängig vom Standort ohne großen Aufwand in Betrieb zu nehmen. Hierzu bedarf es aber wie in Bild 2 dargestellt neben der Schnittstelle auch Lösungen für den Dateninput und auch ein intelligentes Konzept für die dahinterliegende Organisation der Daten. Im Zuge der Untersuchungen wurde dazu eine Test-Datenbank aufgebaut, um schrittweise die einzelnen Lösungsbausteine zu erproben [3]. Hierzu wurde PostgreSQL [5] verwendet. Nachfolgend werden die Grundideen der Lösung, jedoch nicht die Details der Implementierung erläutert.

3.1 Hierarchieebenen

Analog zu den analysierten Bezeichnungssystemen (s. Bild 1) erfolgt innerhalb der Datenbank eine Modellierung des Werksaufbaus über mehrere Hierarchieebenen. Dabei entsteht durch die Verknüpfung verschiedener Ebenen ein Zugriffsweg auf die Daten, der zunächst unabhängig von einer konkreten Bezeichnung des Datenpunkts ist. Weiterhin wird es durch konkrete Elemente (später im Sprachgebrauch der Graphentheorie auch Knoten genannt) möglich, in jeder Hierarchieebene Informationen zu hinterlegen. Sensoren oder Aktoren, deren Messdaten

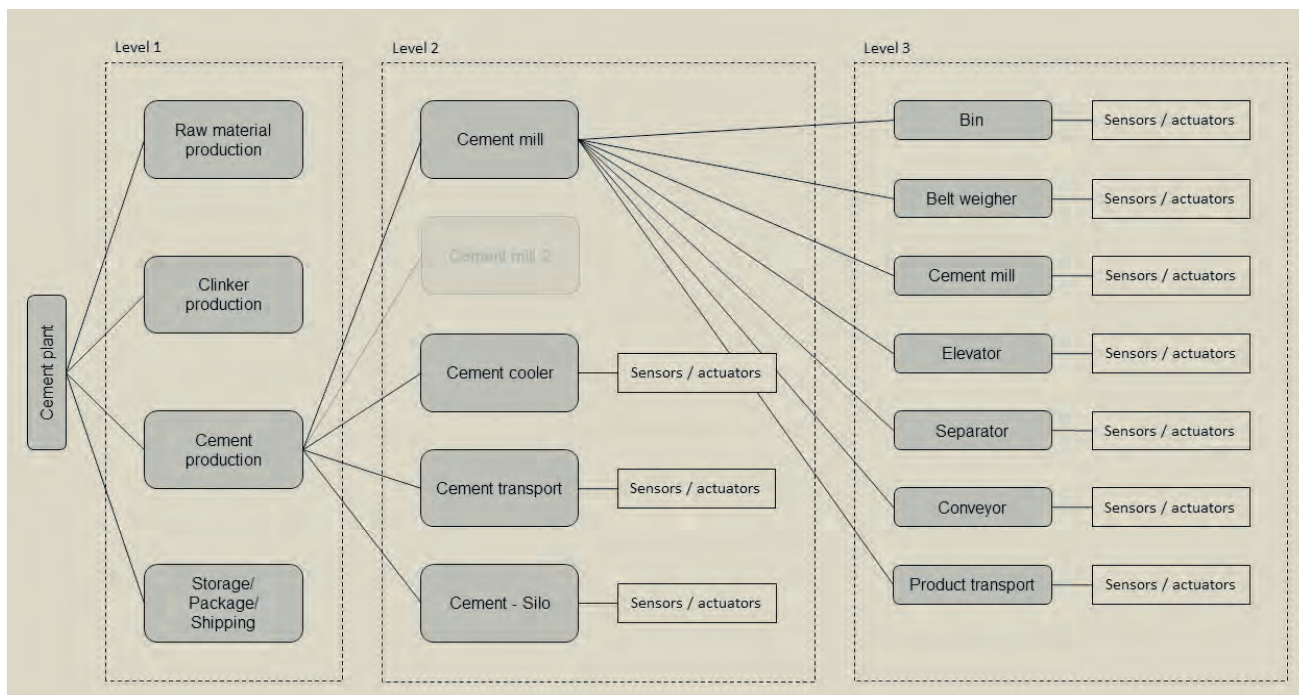


Figure 3: Structure of cement grinding plant across several hierarchy levels

Bild 3: Aufbau einer Zementmahanlage über mehrere verfahrenstechnische Hierarchieebenen

also possible for entire groups of elements (e.g. all temperatures in the area of the preheater).

» Fig. 3 illustrates the representation structure of a cement grinding plant in the test database. The various levels, from the organisation of the entire plant right down to the sections of the cement mill, represent the physical design of the installations and the structure of the plant according to process technology aspects.

3.2 Points of view

The catalogue of requirements proposes additional viewing perspectives above and beyond the traditional method of structuring the data on the basis of process technology. This can be implemented through the creation of additional parallel hierarchy levels in the same way as described in the preceding section. » Fig. 4 illustrates this based on the example of a preheater. The process technology structure right down to a selection of sensors can be seen on the left. Assignment to a higher level generally also describes a physical assignment. The relationship described is thus "belongs to" or conversely "consists of". Other dependencies such as "influences" and "is influenced" are however also conceivable. For example, pressure measurements could be assigned to the fan or to the power meter. The Energy and Maintenance viewing perspectives are therefore also shown in Fig. 4 (right). Levels other than the process technology aspects accordingly appear on the left. From the energy point of view, it would be conceivable to look from the power supply to the meters and loads and from there to the sensors which can be used to assess the energy demand. Maintenance looks first at the installations and the functional sections of these and then at wearing and spare parts. Here again sensor values are of interest, as is in some cases process data for assessment of the condition of plant sections. Irrespective of the viewing perspectives and higher-ranking structures, the actual measurement data only ever have to be present once. Only the access paths differ. The complexity and thus access time must however be assessed separately.

schließlich in Form von Zeitreihen gespeichert werden, können flexibel unterschiedlichen Ebenen zugeordnet werden. Die gebündelte Abfrage ganzer Gruppen von Elementen ist auch über mehrere Hierarchieebenen hinweg möglich (z.B. alle Temperaturen im Bereich des Vorwärmers).

» Bild 3 visualisiert die in der Test-Datenbank aufgebaute Repräsentation einer Zementmahanlage. Die verschiedenen Ebenen von der Einteilung des gesamten Werks bis hinunter zu den Anlagenteilen der Zementmühle repräsentieren den physischen Aufbau der Anlagen bzw. nehmen eine verfahrenstechnische Strukturierung der Anlage vor.

3.2 Sichtweisen

Im Anforderungsprofil wurden zusätzliche Sichtweisen über die klassisch-verfahrenstechnische Strukturierung der Daten hinaus vorgeschlagen. Dies kann durch den Aufbau von zusätzlichen, parallelen Hierarchieebenen realisiert werden und erfolgt analog zu dem im vorherigen Abschnitt beschriebenen Vorgehen. » Bild 4 stellt dies exemplarisch am Beispiel des Vorwärmers dar. Links ist die verfahrenstechnische Struktur bis hin zu einer Auswahl von Sensoren erkennbar. Dabei beschreibt die Zugehörigkeit zu einer höheren Ebene in der Regel auch eine physische Zugehörigkeit. Das beschriebene Verhältnis ist also „gehört zu“ oder anders herum „besteht aus“. Denkbare sind aber auch andere Abhängigkeiten wie „beeinflusst“ und „wird beeinflusst“. So könnten z.B. die Druckmessungen dem Gebläse oder dem Energiezähler zugeordnet werden. Auf der rechten Seite sind in Bild 4 daher auch die Sichtweisen Energie und Instandhaltung skizziert. Entsprechend ergeben sich andere Ebenen als die verfahrenstechnischen auf der linken Seite. Aus Sicht der Energieversorgung ist der Blick von der Spannungsversorgung über die Zähler hin zu den Verbrauchern und weiter zu Sensoren, die zur Bewertung der Energiebedarfe herangezogen werden können, vorstellbar. Die Instandhaltung blickt auf die Anlagen und ihre Funktionsbereiche und dann weiter auf Verschleiß- und Ersatzteile. Auch hier sind schließlich Sensorwerte interessant, zum Teil auch Prozessdaten zur Bewertung des Zustands von Anlagenteilen. Unabhängig von

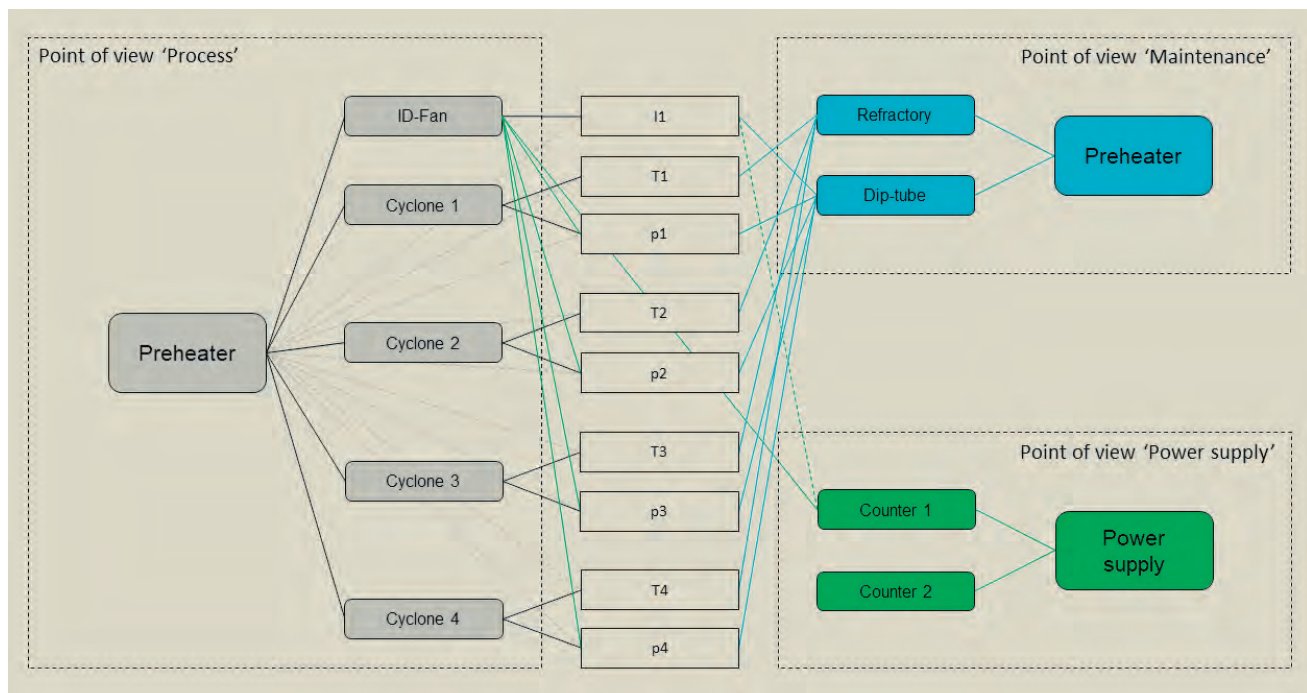


Figure 4: Visualization of different points of view on data by parallel hierarchies

Bild 4: Darstellung von unterschiedlichen Sichtweisen durch parallele Hierarchien

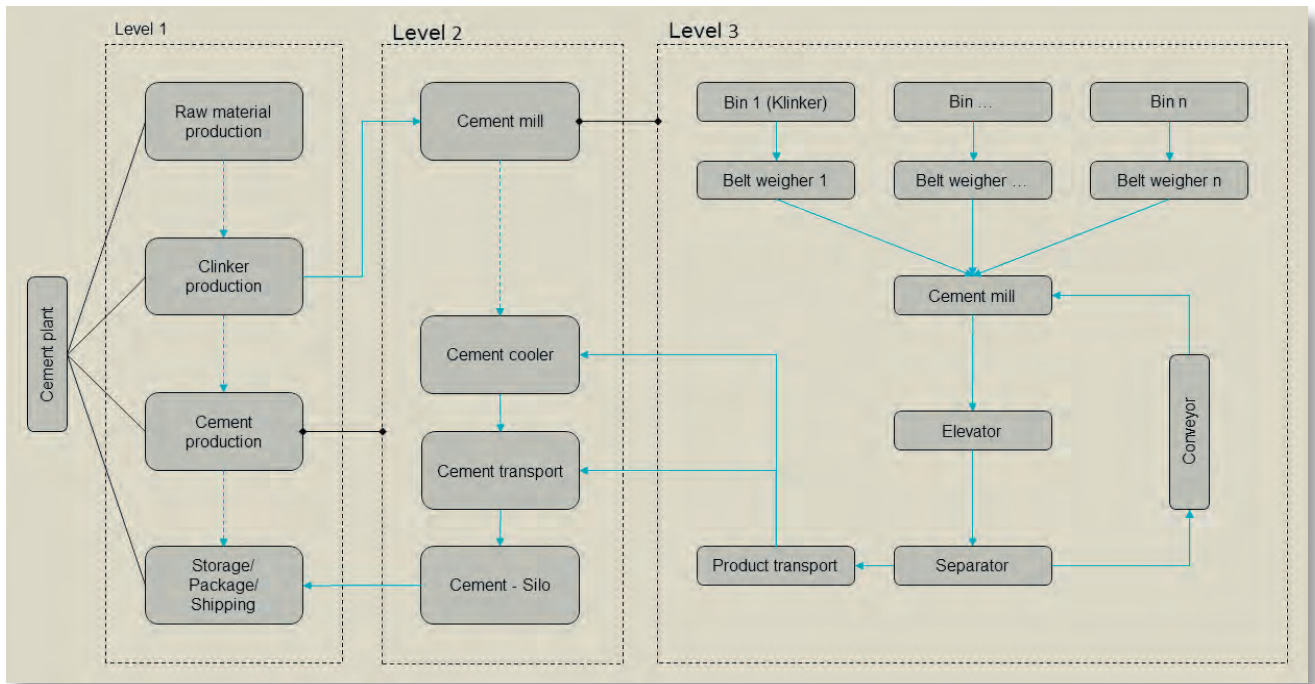


Figure 5: Representation of cement grinding plant across several hierarchy levels including information on adjacency

Bild 5: Darstellung einer Zementmahanlage über mehrere Hierarchieebenen und unter Beachtung der Adjazenzinformationen

3.3 Adjacency

In graph theory, the term Adjacency refers to the property of two "nodes" being linked in one graph. These adjacency relationships can be documented by way of so-called adjacency matrices. This abstract construction can however be applied to process technology installations. One example of this is the machine-readable documentation of piping systems for the calculation of characteristic curves for fans as part of modular process simulations [5]. Process technology installations are generally interlinked systems arranged along a material or gas flow. The adjacency relationship can be reduced to "Predecessor and Successor". This means that the adjacency relationships for each element of the plant ("node") can be documented using an adjacency list – a list of all predecessors and successors. It is possible to keep several lists in parallel, e.g. for the material and gas flow or for other viewing perspectives. Adjacency lists can also be used to document more complex plant structures over several hierarchy levels with branches and return flows of material. Changes such as the insertion of a new sensor can be readily implemented. ▶ Fig. 5 again shows a cement grinding plant [3]. In this case however, not only the hierarchy levels are visualised but also the adjacency information. The blue arrows indicate the material flow through the process.

Implementability was examined for this concept as well in the test database. The adjacency lists are classed as additional information belonging to the metadata.

3.4 Metadata

In addition to the adjacency information, numerous other items of information are required for evaluation of the measurement data or to ensure the quality of the data. ▶ Fig. 6 illustrates this for a sensor by way of example [3]. The actual measurement data is made up of measured values, the associated time stamps, information on the unit of the variable and possibly also a quality code. Further information is however also assigned to the sensor. This includes one or possibly several historical identifiers, details of the sensor

den Sichtweisen und übergeordneten Strukturen müssen die eigentlichen Messdaten immer nur einmal vorliegen. Es verändern sich nur die Zugriffswege. Die Komplexität und damit die Dauer von Zugriffen muss aber separat bewertet werden.

3.3 Adjazenz

Unter Adjazenz („Nachbarschaft“) versteht man in der Graphentheorie die Eigenschaft zweier „Knoten“, in einem Graphen verbunden zu sein. Diese Nachbarschaftsverhältnisse lassen sich durch so genannte Adjazenzmatrizen dokumentieren. Dieses abstrakte Konstrukt lässt sich doch auf verfahrenstechnische Anlagen übertragen. Ein Beispiel dafür ist die maschinenlesbare Dokumentation von Rohrleitungssystemen für die Berechnung von Anlagenkennlinien im Rahmen modular aufgebauter Prozesssimulationen [5]. Verfahrenstechnische Anlagen sind in der Regel verkettete Systeme, die sich entlang eines Material- oder Gasstroms aufreihen. Die Nachbarschaftsbeziehung lässt sich auf „Vorgänger und Nachfolger“ reduzieren. Daher können die Nachbarschaftsverhältnisse für jedes Element der Anlage („Knoten“) durch eine Adjazenzliste – eine Liste aller Vorgänger und aller Nachfolger – dokumentiert werden. Es ist möglich mehrere Listen parallel zu führen z.B. für den Material- und Gasstrom oder für andere Sichtweisen. Durch die Adjazenzlisten können auch komplexere Anlagenstrukturen über mehrere Hierarchieebenen mit Verzweigungen und Materialrückführungen dokumentiert werden. Änderungen, wie das Einfügen eines neuen Sensors, können ohne Probleme realisiert werden. ▶ Bild 5 zeigt nochmals eine Zementmahanlage [3]. Hier sind aber nicht nur die Hierarchieebenen sondern auch die Adjazenzinformationen visualisiert. Die blauen Pfeile zeigen den Materialfluss durch den Prozess an.

Auch für dieses Konzept wurde die Implementierbarkeit in der Test-Datenbank überprüft. Die Adjazenzlisten gehören als zusätzliche Information zu den Meta-Daten.

3.4 Meta-Daten

Neben den Adjazenz-Informationen existiert eine Vielzahl weiterer Informationen, die für die Auswertung der Messda-

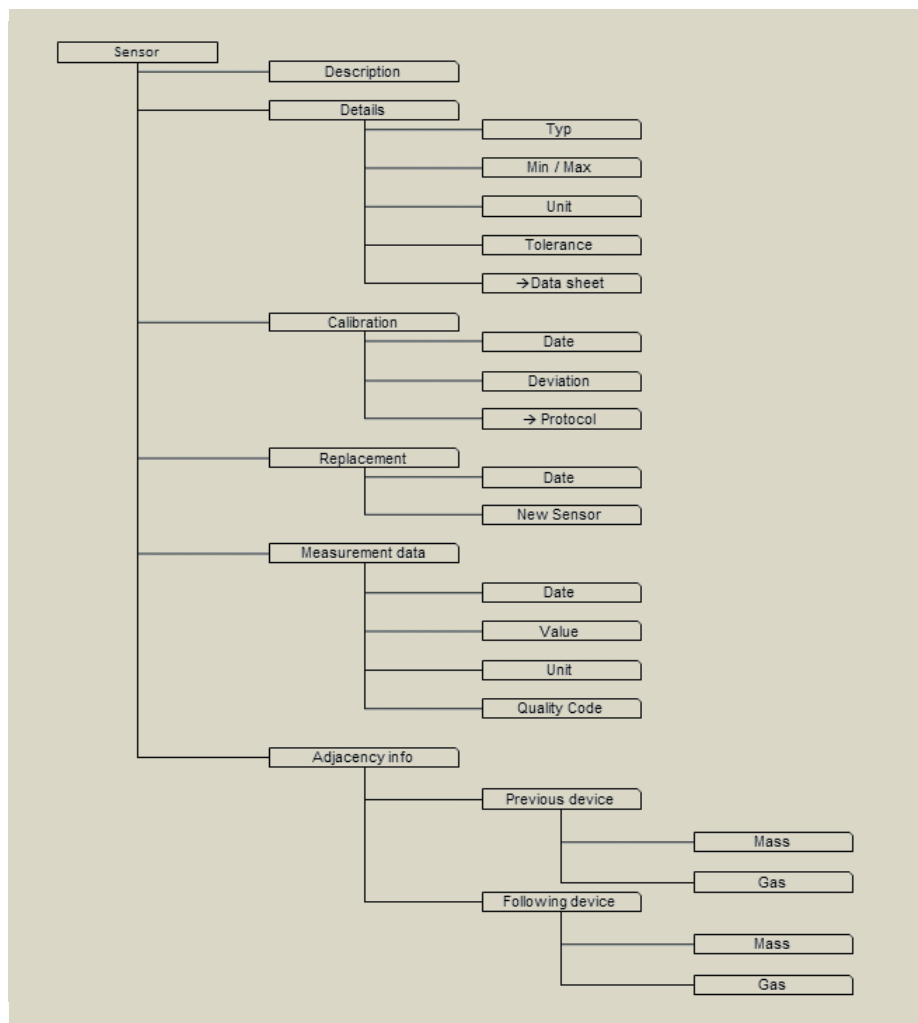


Figure 6: Example of data structure and possible metadata of a sensor

Bild 6: Beispiel für Datenstruktur und mögliche Meta-Daten eines Sensors

itself, information on last calibration, links to external data and a list of relevant changes, such as replacement of the sensor as well as maintenance work which could provide an explanation for any alterations to the measurement data.

The principle of also storing plant sections on higher-ranking hierarchy levels as specific elements in the database makes it possible to assign metadata to these as well, which can then be used for evaluation of the assigned measurement data. » Fig. 7 shows the possible metadata for a plant or a plant section. Of particular interest is the storage of dimensioning data in numerical form to make it available for automated indicator generation from a source if required.

4 Final remarks and outlook

This article presents a catalogue of requirements for the standardised storage of data at cement plants, taking into account both the needs of existing users and the projected future requirements of data-driven applications. It outlines a solution offering tangible additional benefits above and beyond mere standardisation. Furthermore, it

ten oder für die Sicherstellung der Datenqualität erforderlich sind. » Bild 6 illustriert dies am Beispiel eines Sensors [3]. Die eigentlichen Messdaten bestehen aus Messwerten, den zugehörigen Zeitstempeln, einer Information zur Einheit der Größe und ggfs. einem Quality Code. Dem Sensor sind aber weitere Informationen zugeordnet. Dazu zählen eine oder ggfs. mehrere historische Bezeichnungen, Details zum Sensor selbst, Informationen zur letzten Kalibrierung, Verlinkungen zu externen Daten und eine Liste relevanter Änderungen wie z.B. der Austausch des Sensor aber auch Wartungseingriffe, die eventuelle Veränderungen in den Messdaten erklärbar machen können.

Das Konzept, auch Anlagenteile in übergeordneten Hierarchieebenen als konkrete Elemente in der Datenbank zu hinterlegen, erlaubt es, diesen auch Meta-Daten zuzuordnen, die dann auch wiederum bei der Auswertung der zugeordneten Messdaten nutzbar sind. » Bild 7 zeigt mögliche Meta-Daten einer Anlage oder eines Anlagenteils. Besonders interessant ist die Speicherung von Dimensionierungsdaten in numerischer Form, damit diese

ggfs. für eine automatisierte Kennzahlenbildung aus einer Quelle zur Verfügung stehen.

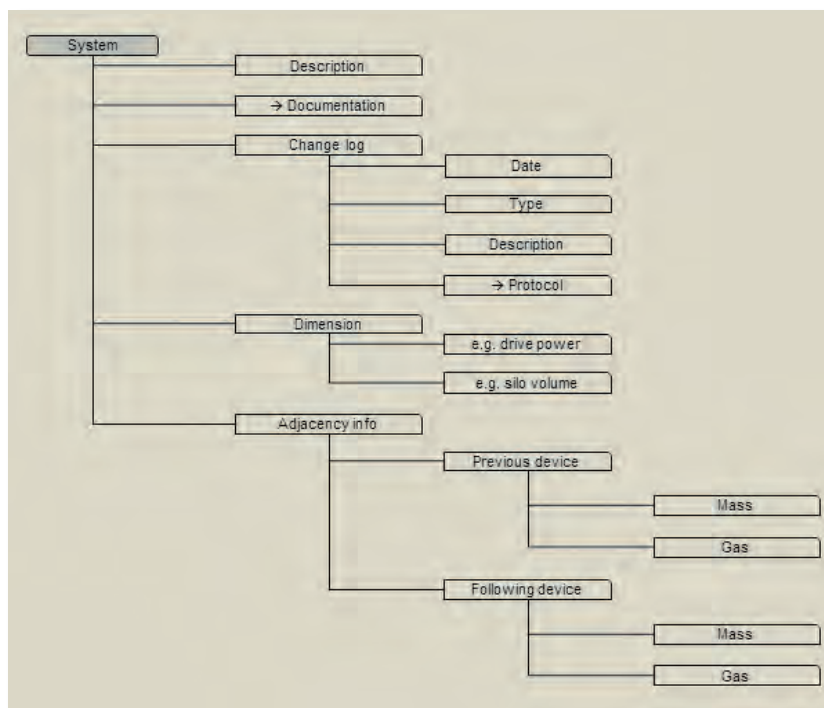


Figure 7: Example of metadata of a plant component

Bild 7: Beispiel für Meta-Daten eines Anlagenteils

describes the creation of a concrete concept for the storage of data which is designed to permit modelling of the process structure and the relationships between the individual process elements represented by data, employing the means of databases. This approach overcomes the limitations of traditional identification systems. In addition it is possible to create structures permitting different points of view (e.g. process technology or maintenance) and thus variable access paths to the data without reproduction of the data. In combination with proposals for supplementing the data with metadata, which is very important for evaluation and assessment purposes, significant items of process information are integrated into the storage solution in machine-readable form.

The trial implementation of a corresponding test database provided the assurance that the concepts under consideration are basically feasible in practice. In the course of further development it must however be established whether these are scalable to the data volumes and complexity of real installations. To this end, the VDZ working group "Standardising data in cement plants" is currently studying trial implementation on an industrial scale. Specifications for a standard open interface and recommendations for the storage of data are then to be derived from the expected findings. Finally, the results will be compiled and presented in the form of traditional VDZ process technology leaflets.

Many questions are still unresolved with regard to this planned further implementation. One such example is consideration of the data volumes to be stored. As already mentioned, appropriate (temporal) aggregation of the data is an important means of reducing the volume of data. The frequently voiced recommendation to store all the data occurring – just to be on the safe side – would lead to an enormous volume of data. Over the course of time it therefore becomes standard practice to replace old data by mean values from ever longer time intervals (e.g. second mean values to start with, then minute mean values). In other cases data is only stored in the event of significant changes. The complete raw data is however usually desirable for machine learning applications. The formulation of standard recommendations for the limitation of data volumes would require a better understanding of the data content (which items of data and relationships are of relevance for process control and optimisation?) as well as more experience with regard to AI applications. Useful findings in this respect are to be expected from ongoing research projects into the robustness of such applications in the cement production process, for example [7].

The working group "Standardising data in cement plants" will continue to perform pre-competitive studies solely on the storage of data up to and including a standard interface. The use of data for further purposes, e.g. in the form of specific machine learning applications, is not envisaged. ◀

4 Schlussbetrachtungen und Ausblick

Vorgestellt wurde ein Anforderungsprofil für die standardisierte Speicherung von Daten im Zementwerk, das sowohl die Anforderungen von Bestandsnutzern als auch den perspektivischen Datenbedarf von datengetriebenen Anwendungen berücksichtigt. Dabei wurde eine Lösung skizziert, die einen fassbaren Mehrwert über die reine Standardisierung hinaus liefert. Es wurde weiterhin ein konkretes Lösungskonzept zur Speicherung von Daten entworfen, dass es ermöglichen soll, den Prozessaufbau und die Bezüge zwischen den einzelnen durch Daten repräsentierten Prozesselementen mit den Mitteln von Datenbanken zu modellieren. Die Einschränkungen klassischer Bezeichnungssysteme werden dabei aufgehoben. Ferner können ohne Vervielfältigung der Daten Strukturen geschaffen werden, die unterschiedliche Sichtweisen (z.B. Verfahrenstechnik oder Instandhaltung) und damit variable Zugriffspfade auf die Daten zulassen. In Kombination mit Vorschlägen für die Ergänzung der Daten durch Meta-Daten, die für die Aus- und Bewertung von großer Bedeutung sind, werden wichtige Prozessinformationen in maschinenlesbarer Form in die Speicherlösung integriert.

Durch die beispielhafte Implementierung einer entsprechenden Test-Datenbank wurde sichergestellt, dass die angedachten Konzepte grundsätzlich realisierbar sind. Es ist aber im Rahmen der weiteren Entwicklung zu prüfen, ob diese auch auf die Datenmengen und die Komplexität realer Anlagen skalierbar sind. Hierzu prüft der VDZ-Arbeitskreis „Standardisierung von Daten im Zementwerk“ aktuell die exemplarische Implementierung im industriellen Maßstab. Auf Basis der hierdurch zu erwartenden Erkenntnisse sollen dann Spezifikationen für eine einheitliche und offene Schnittstelle und Empfehlungen für die Speicherung von Daten abgeleitet werden. Diese werden abschließend in der Tradition der verfahrenstechnischen Merkblätter des VDZ aufgearbeitet und bereitgestellt.

Diese anvisierte weiterführende Umsetzung ist mit vielen offenen Fragen verbunden. Hierzu zählt beispielsweise auch die Betrachtung der zu speichernden Datenmengen. Wie bereits erwähnt, stellt eine geeignete (zeitliche) Aggregation der Daten ein wichtiges Mittel zur Reduktion der Datenmenge dar. Die an vielen Stellen geäußerte Empfehlung, alle anfallenden Daten sicherheitshalber in vollem Umfang zu speichern führt zu sehr großen Datenmengen. Üblicherweise werden daher mit fortschreitender Zeit alte Daten durch Mittelwerte immer größer werdender Zeitintervalle ersetzt (z.B. zunächst Sekundenmittelwerte, dann Minutenmittelwerte). In anderen Fällen werden Daten nur bei signifikanten Veränderungen gespeichert. Für Anwendungen des maschinellen Lernens sind jedoch zumeist die vollständigen Rohdaten wünschenswert. Um einheitliche Empfehlungen für die Begrenzung der Datenmengen zu entwickeln, bedarf es eines besseren inhaltlichen Verständnisses der Daten (Welche Daten und Wirkzusammenhänge haben Relevanz für die Prozessführung und -optimierung?) wie auch mehr Erfahrung in Bezug auf KI-Anwendungen. Hilfreiche Erkenntnisse dazu sind u.a. aus laufenden Forschungsvorhaben zur Robustheit dieser Anwendungen im Zementherstellungsprozess zu erwarten [7].

Der Arbeitskreis „Standardisierung von Daten im Zementwerk“ wird weiterhin nur die Speicherung von Daten bis hin zu einer einheitlichen Schnittstelle vorwettbewerblich betrachten. Die weiterführende Nutzung von Daten z.B. in Form von konkreten Anwendungen des maschinellen Lernens ist nicht vorgesehen. ◀

- [1] Industrie 4.0 in der Zementindustrie Status Quo und Perspektiven (Industry 4.0 in the Cement Industry Status Quo and Outlook). Study on behalf of the Sociopolitical Working Group of the German Cement Industry conducted by the RWTH Aachen and published by the Construction Robotics GmbH, Aachen 2019.
- [2] Implementierung und Betrieb von Big-Data-Anwendungen in der produzierenden Industrie – Durchführung von Big-Data-Projekten (Implementation and operation of Big Data applications in the manufacturing industry – implementation of Big Data projects). VDI/VDE Directive 3714 Sheet 1 draft (2019).
- [3] Kalkbrenner, D.: Entwicklung einer standardisierten Lösung für die Datenspeicherung im Zementwerk (Development of a Standardised Solution for Data Storage at Cement Plants). Bachelor thesis, University of Applied Sciences Duesseldorf (as yet unpublished), Duesseldorf 2021.
- [4] KKS Kraftwerk-Kennzeichensystem. 8th revised edition, VGB PowerTech e.V., Essen 2018.
- [5] PostgreSQL (2021), Licence, <https://www.postgresql.org/about/licence/>, called up on 03.03.2021.
- [6] Emissionsreduktion in der Steine- und Erden-Industrie durch modellbasierte Prozessoptimierung (EMREDPRO) (Reduction of Emissions in the Pit and Quarry Industry through Model-Based Process Optimisation). Concluding report of the BMBF Research Project Reference no.: 01LY1611B, German Cement Works Association (VDZ), Duesseldorf 2020.
- [7] Study into the Robustness of AI Applications for Process Optimisation in Cement Manufacturing. Research Project AIF 21152 N, German Cement Works Association (VDZ), Duesseldorf 2021, <https://www.vdz-online.de/en/knowledge-base/research-projects/investigation-of-the-robustness-of-ai-applications-for-process-optimisation-in-cement-manufacturing>, called up on 26.03.2021.

Fast ■ Fair ■ Flexible ■



InterMod – Connect and Cement

InterMod, the individually combinable module system for smaller capacities up to 30 tph can be moved and removed quickly and ultra-flexibly and thus enables a quick location change. Connect and produce cement exactly where it is needed. In sufficient quantities, top quality and round the clock.

InterMod: the modular system makes it possible!

The modules are pre-assembled, individually combinable, quickly removable, ultra-flexible and designed for rapid change of location. They are equipped with proven single products and electronically pre-installed.

And all that: **Fast ■ Fair ■ Flexible ■**



Intercem Engineering GmbH ■ Intercem Installation GmbH

Carl-Zeiss-Straße 10 | 59302 Oelde | Germany

Tel.: +49 2522 92058-0 | E-Mail: info@intercem.de | www.intercem.de

