Alexander Mankel, Madlyn Kowalczyk

Data + Context = Invaluable Information

Plantsight for Process Industries

How data can add value to the process industries

Long before the topic of digitalization gained momentum, data played a decisive role in day-to-day plant operations of the process industries: product quality, plant efficiency, process reliability and much more can - then as now - only be reliably determined on the basis of data. In the context of global digitalization efforts, the topic of data, its management and analysis is becoming even more of a focus.

In this context, technological progress means on the one hand the breathtaking increase of generated data, but on the other hand it offers both techniques and tools to use this data profitably. This white paper aims to provide an overview of the value-adding potential of data in the process industries.

Table of contents

Data in process industries

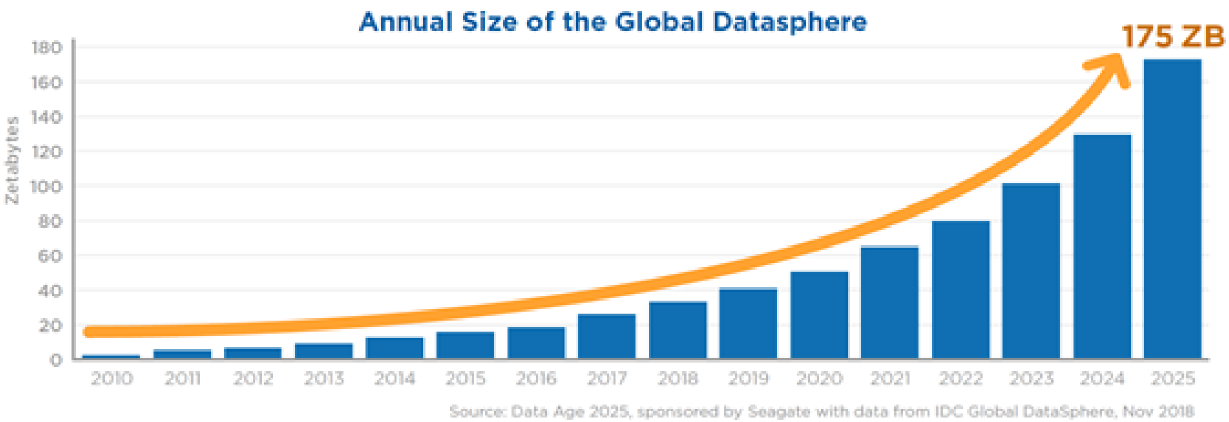
In 2018, international market research and consulting firm International Data Corporation (IDC) defined three key crystallization points for digitalization where digital content is created: the core (traditional and cloud data centers), the edge (hardened network infrastructure such as cell towers and network nodes), and the endpoints (PCs, smartphones, and cyber-physical systems (CPS) or participants in the Internet of things (IoT)). The sum of all this data, whether generated, captured or replicated, is referred to by IDC as the 'Global Datasphere', and it is experiencing tremendous growth. IDC predicts that the Global Datasphere will grow from 33 zettabytes (ZB) in 2018 to 175 ZB in 2025.[1] Data volumes in the process or manufacturing industries are also growing accordingly.

To give an example, the machines involved in the production of Simatic devices at the Siemens plant in Amberg generate a terabyte of data per hour with their sensors and actuators.[2]

Let's assume a plant has 3000 sensors and each one delivers one value per second. How many data sets are accumulated in a year? Hundreds of thousands? Millions? Admittedly, in this theoretical example it would be 94,608,000,000 - almost 95 billion data records! In real process plants, measured values are sometimes generated every millisecond, plus acyclic data such as alarms, messages, laboratory measurements, maintenance data or batch information. So, it's not just the enormous volumes that pose a challenge to create value from this so-called Big Data, other characteristics also have a decisive influence.

< tl; dr>[3]

The 'Global Datasphere', the sum of all data generated, captured or replicated worldwide, is growing rapidly. With the Internet of Things, the speed of growth will continue to increase.



The sum of all data worldwide (generated, captured, or replicated) is growing inexorably. [Source: see [1] page 6]

[1] David Reinsel, John Gantz, John Rydning: "The Digitization of the World - From Edge to Core. " An IDC Whitepaper, November 2018

[2] Jürgen Hill, "Siemens erweitert Digital-Enterprise-Angebot", Computerwoche, Feb. 25, 2019. https://www.computerwoche.de/a/siemens-erweitert-digital-enterprise-angebot,3546601,2 last accessed 2020/09/01

[3] tl: too long, dr: didn't read

Big Data / Dark Data / Smart Data - Definitions

So far, there is no clearly defined definition for the term **'Big Data'**. This is due to the fact that the topic is very complex and can be viewed from many different perspectives. While the quantity of data was originally the main criterion for Big Data [3] - without specifying when a quantity of data can be called Big Data - the term was later expanded to include the complexity of data, data processing and analysis, and economic potential[4]. The characteristics of Big Data will be examined in more detail in the next chapter.

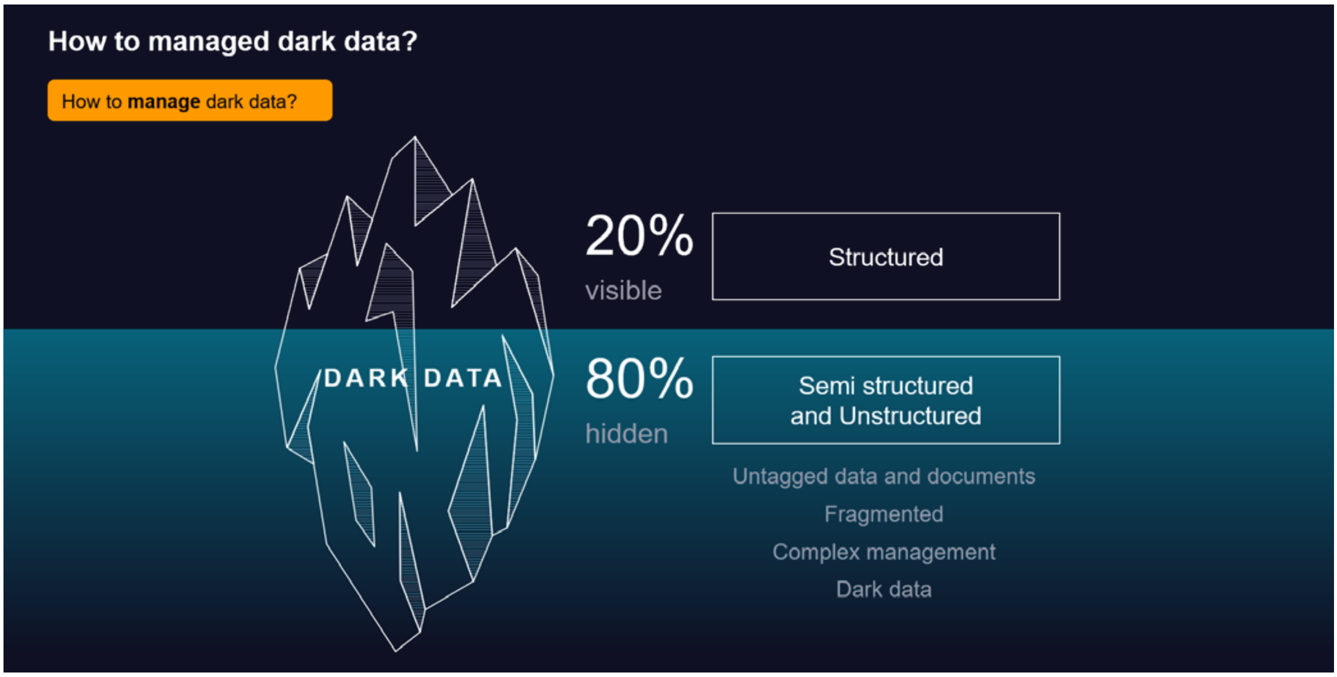
In this white paper, Big Data is understood to mean large volumes of structured, semi-structured and unstructured data that are so extensive and complex that traditional techniques and methods for data processing and use can no longer be applied in a way that adds value or is efficient. For the efficient processing of Big Data, new technologies, IT systems and IT architectures are therefore needed as an alternative to conventional relational databases (e.g., NoSQL databases). At the same time, it is assumed that the use of Big Data represents a relevant production and competitive factor and, therefore, opens up a new value creation potential.[5]

In the following, the term **'Dark Data'** is also used. This forms a subset of Big Data. Gartner defines it as the information assets that companies collect, process and store in the course of their regular business activities, but generally do not use for further purposes (e.g. for analysis, business relations and direct commercialization)[6]. Dark Data is inaccessible or hidden data that defies easy electronic processing because it is not in digital form.

Finally, **'Smart Data'** in this white paper refers to the result of transforming Big Data (including Dark Data) into structured and usable knowledge. The prerequisite is the integration of this data from heterogeneous systems and the extraction of insights at high speed from this very differently structured data using scalable methods and techniques.

< tl; dr>

The definition of 'Big Data' is fuzzy and evolving. Today, the term implies not only data volume, but also complexity and thus challenges in dealing with data.



Much of the data in companies is inaccessible but no less valuable

[3] Doug Laney: "3D Data Management: Controlling Data Volume, Velocity, and Variety". META Group Whitepaper, 2001

[4] BITKOM Study: „ Big Data im Praxiseinsatz - Szenarien, Beispiele, Effekte". Berlin, 2012

[5] For a current overview of the various definitions of Big Data, see Philipp Gölzer: "Big Data in Industrie 4.0. Eine strukturierte Aufarbeitung von Anforderungen, Anwendungsfällen und deren Umsetzung". Dissertation submitted to Friedrich-Alexander-Universität Erlangen-Nürnberg, 2016. https://www.researchgate.net/publication/312522237_Big_Data_in_Industrie_40_-_Eine_strukturierte_Aufarbeitung_von_Anforderungen_Anwendungsfallen_und_deren_Umsetzung last accessed 2020/09/04

[6] Cf. Gartner Glossary 'Dark Data': "Gartner defines dark data as the information assets organizations collect, process and store during regular business activities, but generally fail to use for other purposes (for example, analytics, business relationships and direct monetizing)." https://www.gartner.com/en/information-technology/glossary/dark-data last accessed 2020/09/04

Basic characteristics

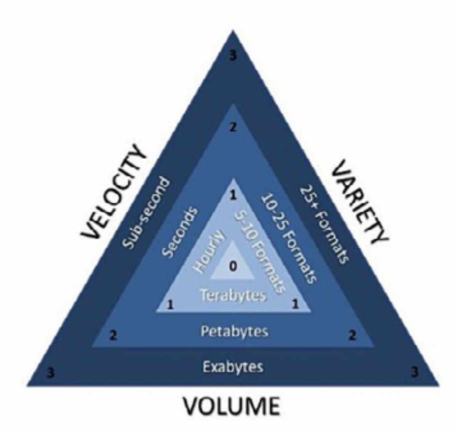
In a high-profile editorial in the Harvard Business Review in 2012, authors Andrew McAffee and Erik Brynjolfsson proclaimed a management revolution that would be sparked by Big Data: "Exploiting vast new flows of information can radically improve your company's performance. But first you'll have to change your decision-making culture." According to their criteria, Big Data defines itself as follows:[7]

Volume - When dealing with Big Data, the quantity or mass of data is particularly emphasized. There are different interpretations of the term 'volume', meaning either the size of the data volume or the amount of data generated in one day.

Velocity - The next core property of Big Data: Velocity refers to the speed at which data is generated, stored, processed, and visualized.

Variety - The third main characteristic identified describes the variety of data. These are often not in structured form, so that established models of data processing fail. But 'variety' also refers to the diversity of formats, structures, and storage in different systems.

These core properties were expanded in the following years. Thus, the data quality and therefore the reliability of the underlying data is referred to as **'veracity'** (correctness, truthfulness), regardless of the data type. Veracity is often interpreted in literature in terms of consistency, i.e. statistical reliability, and trustworthiness. Whereas trust is influenced by several factors, such as origin, collection criteria and processing methods, as well as the security of the data sources, including immutability and protection against unauthorized data access.[8] The worth of the data, labeled **'value'**, is now widely accepted as another addition. This property refers to the economic value, of the data to be processed. Particularly in the case of semi-structured and unstructured data, which differs from traditional data structures, information can lie hidden and, if identified and analyzed appropriately, represents a high value for companies. Also moving into the general definition is the change in the rate of data flow over time, known as **'variability'**. It was introduced to flank velocity and recognizes the fact that the velocity of data and its processing is not always constant; event-driven load peaks often occur. Finally, the term **'viability'** is often found, which refers to the usefulness of data attributes. In analyses, the attributes of data are not considered equal and, depending on the context, some attributes may be neglected. In terms of resource-efficient analysis or storage of data, it is important to identify only those attributes that have the greatest value and are, therefore, the most useful.



The 3V model complemented by Doug Laney's Gartner Data Magnitude Index (DMI).

0: without effect

1: low impact

2: medium impact

3: high impact

Graphic based on: Matthias Volk and Stefan Willi Hart: " Big-Data: Bestimmung der Big-Data-artigkeit von Projekten", 2016.

[7] Andrew McAffee, Erik Brynjolfsson: "Big Data. The Management Revolution. Harvard Business Review, October 2012

[8] Yuri Demchenko, Paola Grosso, Cees de Laat, Peter Membrey: "Addressing big data issues in Scientific Data Infrastructure." In: 2013 International Conference on Collaboration Technologies and Systems (CTS)

Challenges in dealing with Big Data in process industries

The aforementioned characteristics give rise to a large number of challenges when dealing with data. In industrial environment, huge quantities of data sets are created from different original sources and have correspondingly different structures, semantics and specifications, therefore, not easy to merge: "The data from sensors, from engineering and other databases, from process information management systems (PIMS), as well as from shift logs and operating procedures are very heterogeneous and must be transferred to a common semantic basis for linking in a central data warehouse".[9]

The fact that extensive operating data as well as historical data are available via the process control system is an advantage for process industries. By analyzing this available data, more economical plant operation modes can be determined and implemented without changing the production plant itself. However, suitable methods for data analysis for ongoing plant operation must be found, tested, and established. For example, data in the industrial environment is characterized not only by high velocity and variety, but also by noise and high redundancy. If the analysis of Big Data is to identify machine anomalies, then in addition to sufficient computing capacity, algorithms must also be developed that use specific search criteria and filters to ensure that knowledge is gained both effectively and efficiently.

In every process plant, there are also concepts for data storage which are mostly based on the respective documentation obligations. The use of this data for statistical evaluations is not possible without further ado, since the storage of data records in accordance with documentation makes completely different demands on the infrastructure than statistical evaluation and where access times, transfer rates, etc. are decisive. Thus, the question of centralized or decentralized data storage must also be considered in a differentiated manner for both use cases.

The increasing volume of data means that established storage systems and concepts must be reconsidered. Especially from the point of view that the up-to-date of the data is crucial when analyzing Big Data. Therefore, close attention must be paid to how data quality assurance can be achieved, and decisions must be made about which data can be archived and which can be deleted after use. The completeness of the data base presents a wide range of challenges. Here, aspects such as physical accessibility, electronic availability and transmission rates play just as important a role as information extraction from unstructured data.

The question of how to present the results of data analysis should not be neglected. Only if this is intuitive can it contribute to effective decision-making. The results must be prepared and presented differently depending on the user role and should be tailored to the respective information recipient in terms of depth of detail, scope, frequency, and form of visualization.

Another problem is rooted in the nature of process plants as they are subject to constant change. When optimizing processes, expanding or modifying plant (or parts of), and during scheduled maintenance shutdowns, components are replaced or supplemented and process control strategies are adapted. This erodes the learned models for data analysis and thus a constant learning process of Big Data models becomes a fixed component in the life cycle of plants and devices.

The challenges mentioned are only examples and must be supplemented individually.[10] This requires both expert knowledge and tools, which leads not least to potential security risks. Operational data for controlling important business processes contains sensitive information and therefore Big Data analyses may only be passed on to third parties for processing if appropriate preventive measures are taken to protect this sensitive data.

< tl; dr>

An accompanying strategy and professional software systems are essential to master the manifold challenges in the value-adding use of data in process industries.

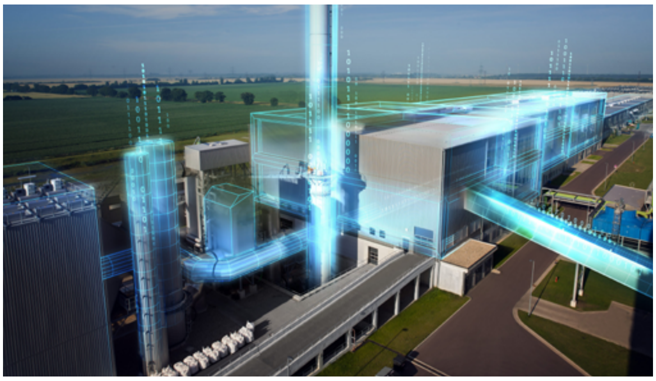
[9] Jens Folmer et al.: "Big and Smart Data. Challenges in the Process Industry" In: atp magazin, pp. 58-69, March 2017

[10] For an example of methodological approach, see JunPing Wang, WenSheng Zhang et al: "Industrial Big Data Analytics: Challenges, Methodologies, and Applications," 2018. https://www.researchgate.net/publication/326171566_Industrial_Big_Data_Analytics_Challenges_Methodologies_and_Applications last accessed 2020/09/04

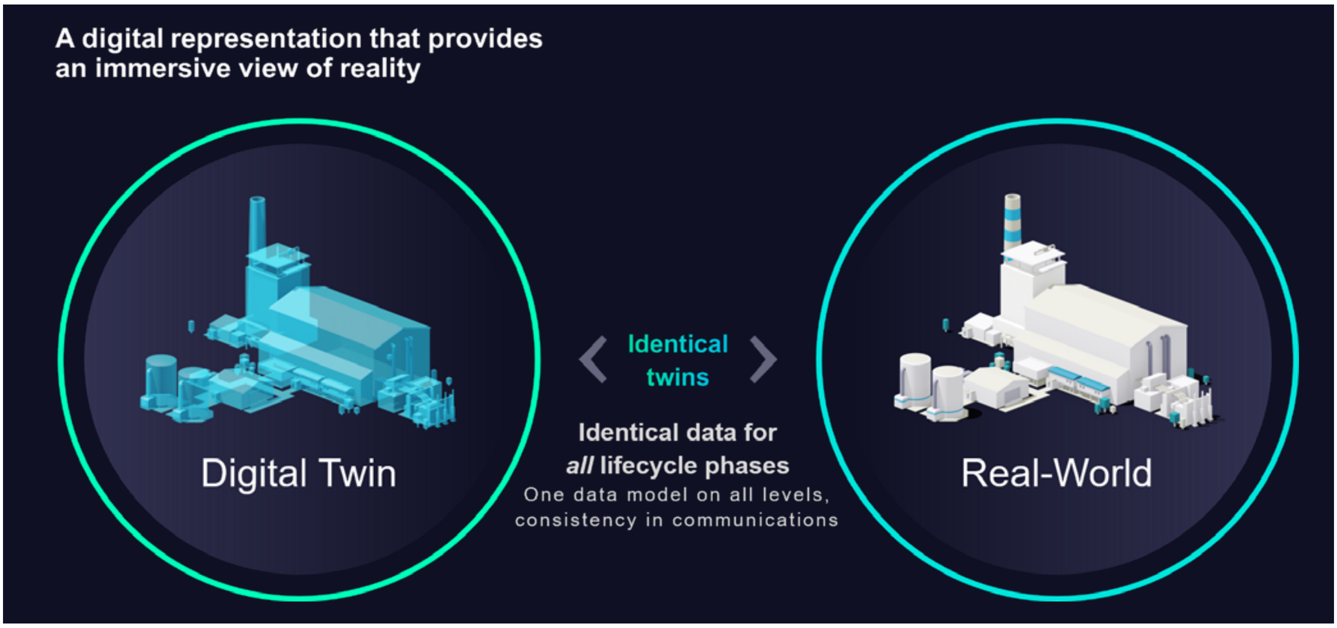
The digital twin

How can the integration of large, diverse, widely distributed and sometimes hidden volumes of data be achieved?

One approach much discussed today is the digital twin, the digital image of an object. This is not a static electronic copy of a real object, but a dynamically synchronized system that must maintain a minimum level of completeness and accuracy over time and should map object states in near real time. As already pointed out with Big Data, there is no clear, generally accepted definition for the digital twin. For this white paper, the following quote should serve as a first approximation:



"A key task of the digital twin is to organize and coordinate the networking of partial models into an integrated information space […] and to permanently store (raw) data and algorithms and make them readily available to people and technical systems with access rights. Metadata, i.e., information that provides the necessary context to make it easier to find relevant content, evaluate it, and act on it in open innovation processes, is of particular importance in this context." [11]



[11] Andreas Bamberg, Leon Urbas et al.: " Was den Digitalen Zwilling zum genialen Kompagnon macht ". In: Chemie Ingenieur Technik, 3/2020: „Eine Schlüsselaufgabe des Digitalen Zwillings ist die Vernetzung der Partialmodelle zu einem integrierten Informationsraum […] zu organisieren und zu koordinieren sowie (Roh-) Daten und Algorithmen dauerhaft vorzuhalten und für die zugriffsberechtigten Menschen und tech-nischen Systeme leicht verfügbar zu machen. Eine besondere Bedeutung haben dabei Meta-daten, also Information, die den notwendigen Kontext liefern, um relevante Inhalte leichter finden, bewerten und in offenen Innovationsprozessen handeln zu können."

Definition

The most important characteristics have already been brought into play. The ARC strategy paper[12] "Digital Twin Demystified" published in 2020 provides valuable information for further consideration. According to this paper, the digital twin consists of three core elements:

Context and characteristic data that define the properties of the real asset or process and largely arise during the planning phase. Typical data are process diagrams, simulations, 1D, 2D and 3D models, parts lists, maintenance instructions or product specifications.

Real-time and operational data that arise during the life cycle of the plant or process. This includes process values, messages or alarms that occur in real time or are available as historical data.

An information model used to integrate all the aforementioned data. It formalizes the properties, relationships and operations that can be performed on each data type that is part of the digital twin. Implemented as a database, it aggregates data from the disparate systems and acts as a 'single source of truth' - the only valid access point to consistent data - for each application that accesses the digital twin.

The strategy paper also distinguishes between two basic types of digital twins: static project-related digital twins on the one hand and dynamic performance-related digital twins on the other. The former is used for the development and deployment of new products, plants, or processes, and the latter is used to optimize plant performance and improve business functions related to a product, plant, or process. This includes all maintenance and servicing operations.

Requirements

If the digital twin is to serve as the basis for integrated data management and evaluation to optimize decision-making processes and improve plant performance, plant information management must fulfill several framework conditions. These will be briefly outlined and take into account special requirements of the process industries.

A precise plant identification system is a matter of course for process plants. Especially when existing plants, so-called brownfield plants, are to be transferred to the digital age, inconsistencies with regard to the tagging of the technical locations or in the structure of the plant identification pose major challenges because the digital twin absolutely requires a plant tagging based on relevant standards. Optimally, the labeling is derived from the firmly defined structure stored in the engineering or ERP (enterprise resource planning) system, however, another important requirement plays into this in that the plant information model must be valid throughout the entire life cycle. This also means that engineering models should be transitioned smoothly into the operating phase at project handover. The following addition is valuable here: "It is inaccurate to speak of the one life cycle of a plant. Rather, there are three different life cycle aspects that are considered independently […]. These are the life cycle of the process, the life cycle of the plant structure, and the life cycle of the actual assets of the plant." [13]

< tl; dr>

The digital twin can serve as an integrated information space and thus as a basis for consistent data management and evaluation. Information derived from this improves decision-making processes.

[12] Founded in 1986, ARC Advisory Group is a leading technology research and advisory firm for industry and infrastructure headquartered in the United States. ARC is characterized by an in-depth engagement with both information technologies (IT) and operational technologies (OT) and related business trends.

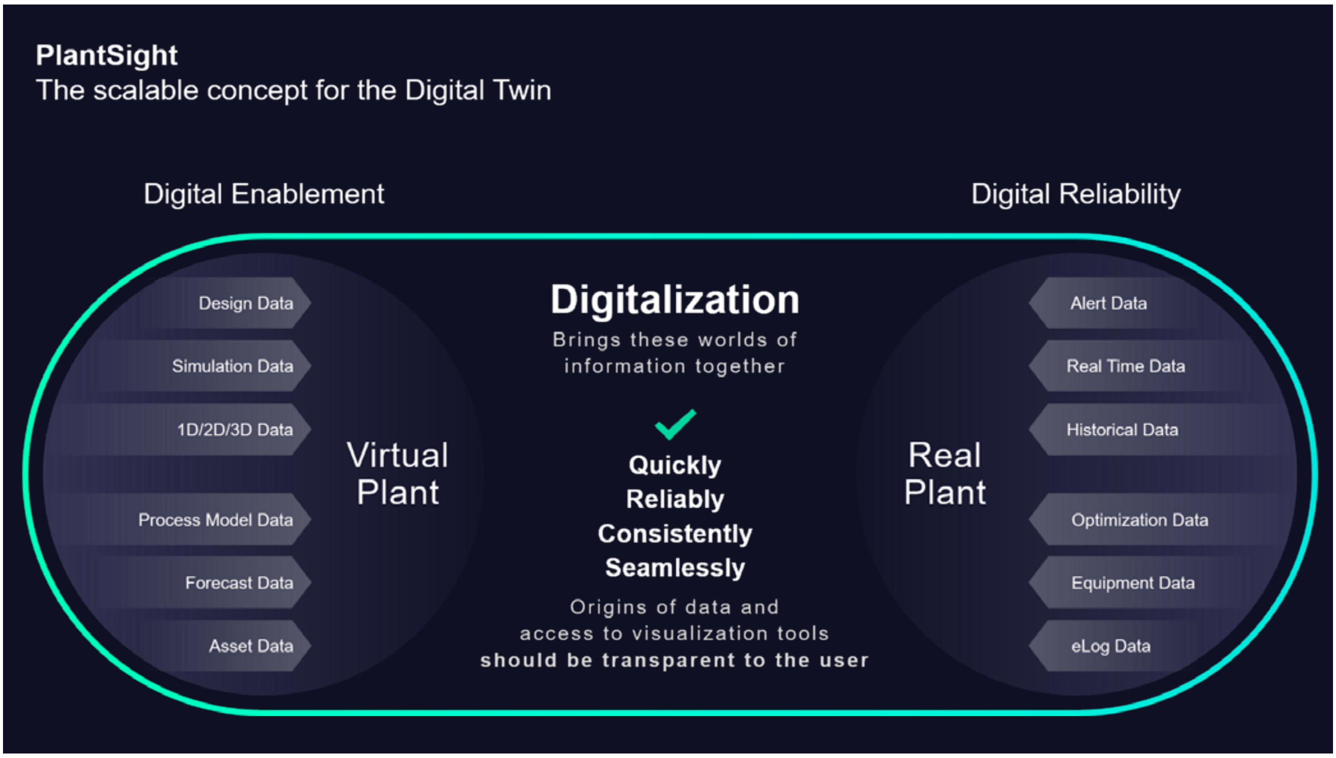
[13] Andreas Schüller et al.: „ Der Digitale Zwilling in der Prozessindustrie". In: atp magazine, pp. 70-81, 2019

The utility value of a digital twin depends directly on the availability of information from all value chains and all life cycles. Crucial are

- a constant information feedback from the knowledge gained during operation into the planning for the optimization of the plant performance as well as
- an automated continuous data reconciliation between model(s) and reality.

These data reconciliations require correspondingly standardized interfaces, defined methods and processes, and suitable electronic systems. From the user's point of view, such a multidimensional implementation of a digital image of a plant represents an enormous expenditure of time and money. This explains the desire for scalable approaches and which is economically indispensable. Questions, such as, can modernizations of plant parts / subsystems represent a first step towards digitalization? Also, is it possible to initially integrate only the most important assets? Or, can the introduction of maintenance and repair management be used to establish the digital twin?[14]

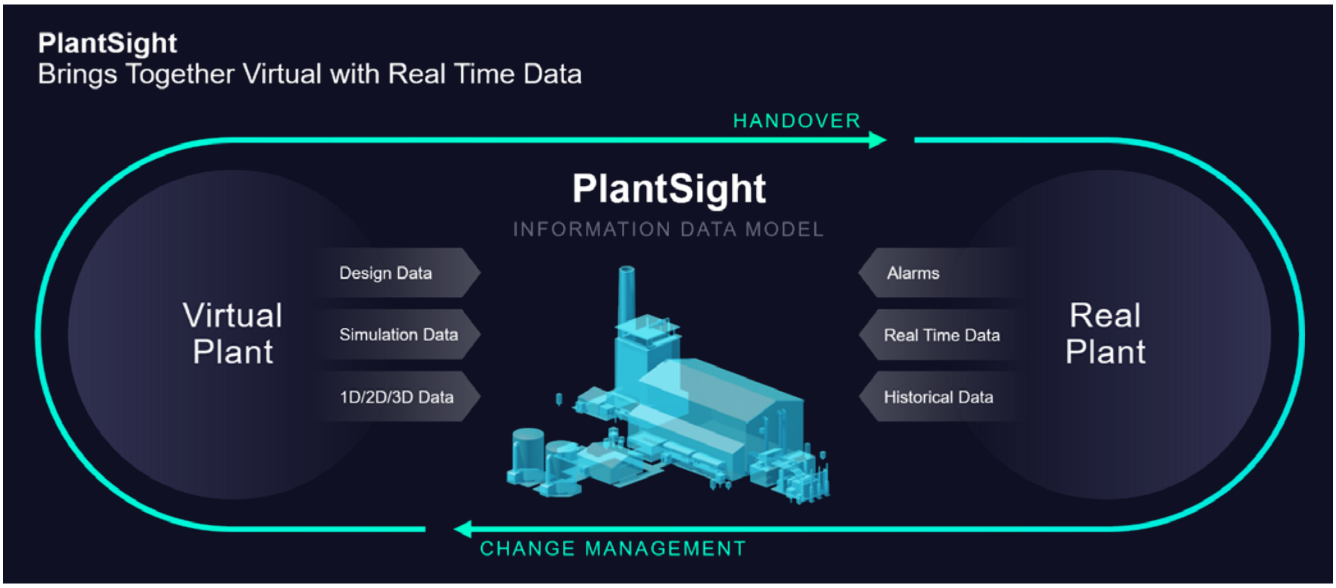
As mentioned at the beginning, the planning, operation, and maintenance of plants are based on processes with a strong division of labor, in which many specialized disciplines are involved. If the digital twin is to actually create an integrated information space, then another requirement is that information is to be made accessible, editable, and vividly visible in a role-specific manner. This also includes the unrestricted provision of data via mobile end devices for commissioning or maintenance personnel.



Digitalization connects the virtual and the real plant

[14] Many other important questions can be found in: Andreas Bamberg, Leon Urbas et al : „ Was den Digitalen Zwilling zum genialen Kompagnon macht". In: Chemie Ingenieur Technik, 3/2020, p. 197

The digital twin in action



Cross-format integration of all data sources thanks to Plantsight's uniform information model

With its comprehensive digital enterprise portfolio, Siemens AG offers industrial companies of all sizes the key to integrating and digitizing relevant business processes so that they can exploit the full potential of digitalization for themselves and reliably reach the next level of digital transformation.[15]

Basics

With Plantsight, Siemens is creating the technological basis for the complete digital twin, which is continuously updated and closely models its physical counterpart in both behavior and information context, to provide users with required information on a situation-specific basis. Plantsight combines static 1D, 2D and 3D data with dynamic real-time data to create a single, clear, and up-to-date representation of a process plant. Tag numbers provide the link between all data types and ensure cross-format integration of all data sources. These sources can include existing engineering or maintenance databases, 3D models, photogrammetric information, as well as data from enterprise resource planning (ERP) systems, recipe databases, project and portfolio management systems, or laboratory information management systems (LIMS). An open system architecture allows users to build the digital twin from data created with vendor A's tools with vendor B's 3D models and vendor C's simulation models. This is implemented through micro-services and connectors that can act as containers for almost any type of data and are thus able to create a connected data environment (CDE). This approach to distributed data management (where data points are intelligently linked) avoids unnecessary data replication while providing consistency and accessibility. Such functionality can significantly reduce the effort required to build the digital twin and increase the percentage of documented asset information.

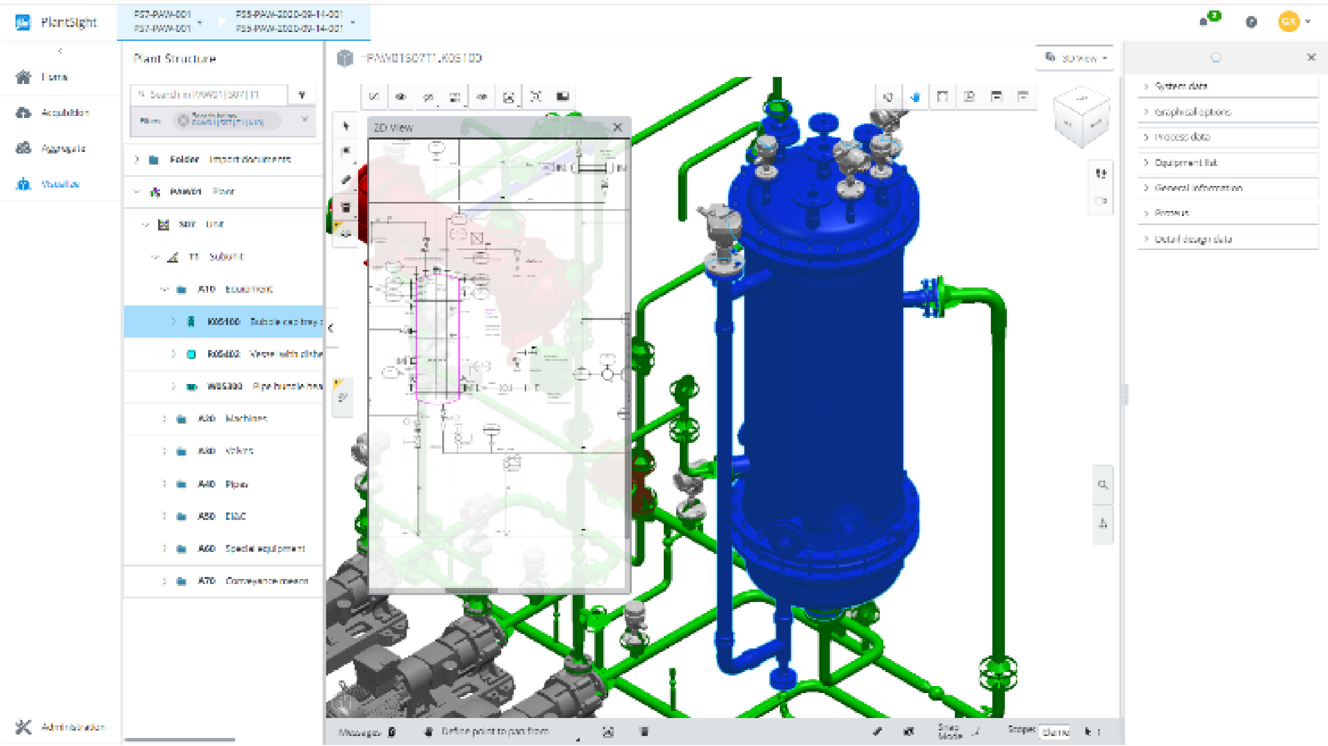
Services for validating and linking data with other information, combined with change tracking, improve the accuracy, completeness and trustworthiness of asset data (cf. 'Veracity' from chapter 2.2).

The easiest entry point for creating the digital twin, for both greenfield and brownfield projects, is the Piping and Instrumentation Diagram (P&ID). P&IDs should generally be up to date, even for older plants. In them, the real plant components are represented with appropriate tagging, and from here they can be supplemented with further data, attributes, and work processes. If a CAD model is available, it can also be coupled using connectors. By incorporating the results of laser scans or photogrammetry, existing assets can also be recorded and linked to the P&ID and any 3D models. Once Plantsight has captured and connected the data, these results are visualized and can then be verified and further contextualized through links, for example, by connecting to the ERP system, maintenance management or other sources.

[15] Clear entry: https://siemens.com/digital-enterprise

Application examples and scenarios

In all plants, the creation of a digital twin will always be associated with a specific intention. The type of challenge that is to be mastered with its help in the future is determined by the individual approach of the general methodology just described. In the subsequent sections, typical use cases are outlined.



Based on the centrally managed data, plant assets can be visualized in 2D or 3D view including associated data such as process values.

Holistic, up-to-date plant overview

Data integrity and the most complete integration possible of all sources are the keys to reducing unplanned downtime and increasing efficiency in various forms, e.g., lower costs, higher production levels and overall quality or less maintenance effort. If you want to reap the benefits of the digital twin as comprehensively as possible, you must build it on a stable (data) foundation. For this purpose, necessary / meaningful data aggregation and validation with integrated change management are essential. Comparative static information such as 1D, 2D, and 3D data from engineering must be consolidated and contextualized with dynamic data from operations, such as, historical and current process values, job information, maintenance and inspection data, etc.
The digital representation of the plant can then accompany the actual plant for decades and provide optimum support at all times. The goal is the constant traceability and simple visualization of all modifications made to plant components or processes. To achieve this, Plantsight creates integrated functional and spatial modeling that can ensure all changes made during operations are promptly and accurately captured and recorded. More importantly, Plantsight provides easy, role-specific access to the information that can be gleaned from the data and its context. This generates real added value to enable faster and more efficient decisions to be made by all the stakeholders in a plant - on the basis of a secure and comprehensible data / information situation.

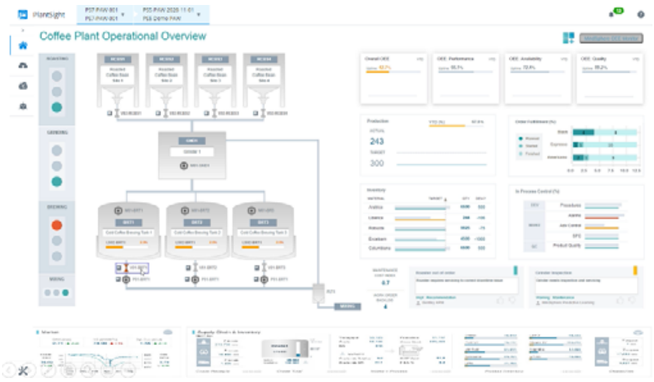
< tl; dr>

The structure and use of the digital twin varies depending on the objective. A step-by-step introduction and expansion is flexibly possible, and integrated databases can be used for optimization in a variety of ways.

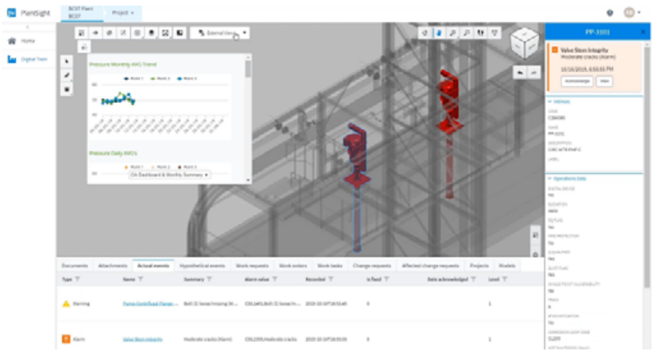
Discipline-specific asset information

The basis for fast and precise decision-making processes is in place when every player always knows the current state of his technical environment. All information that is crucial for the instantaneous activity in the corresponding situation must be clearly visualized and measures for proactive action must be just as easy to derive. A centralized digital asset portal enables this efficient decision-making and allows proactive operational management of the asset through access to real-time data, monitoring and alerts, and access to relevant systems, including supporting documentation. Typically, the following disciplines are covered, operations management, health and safety, process / production, reliability, maintenance, inspection and planning.

Plants facing extensive modernization work benefit from the digital twin as a platform for smooth collaboration between different disciplines. For example, process engineers can contribute their knowledge just as much as design teams. Everyone works with the proven and specialized tools, but the digital twin provides the basis for collaboration by using all the information generated. Role-specific access to plant information via 3D and VR representations can significantly reduce the time normally spent in the plant during preparation and planning for inspections and maintenance work in brownfield projects. In greenfield projects, on the other hand, virtual walk-throughs, immersive training, etc., are possible even before the plant is built and commissioned.



At a glance, Overall Equipment Effectiveness (OEE) can be captured. The dashboards can be customized.



The key figures of individual assets including current and historical values as well as the complete documentation can be accessed quickly and intuitively.

< tl; dr>

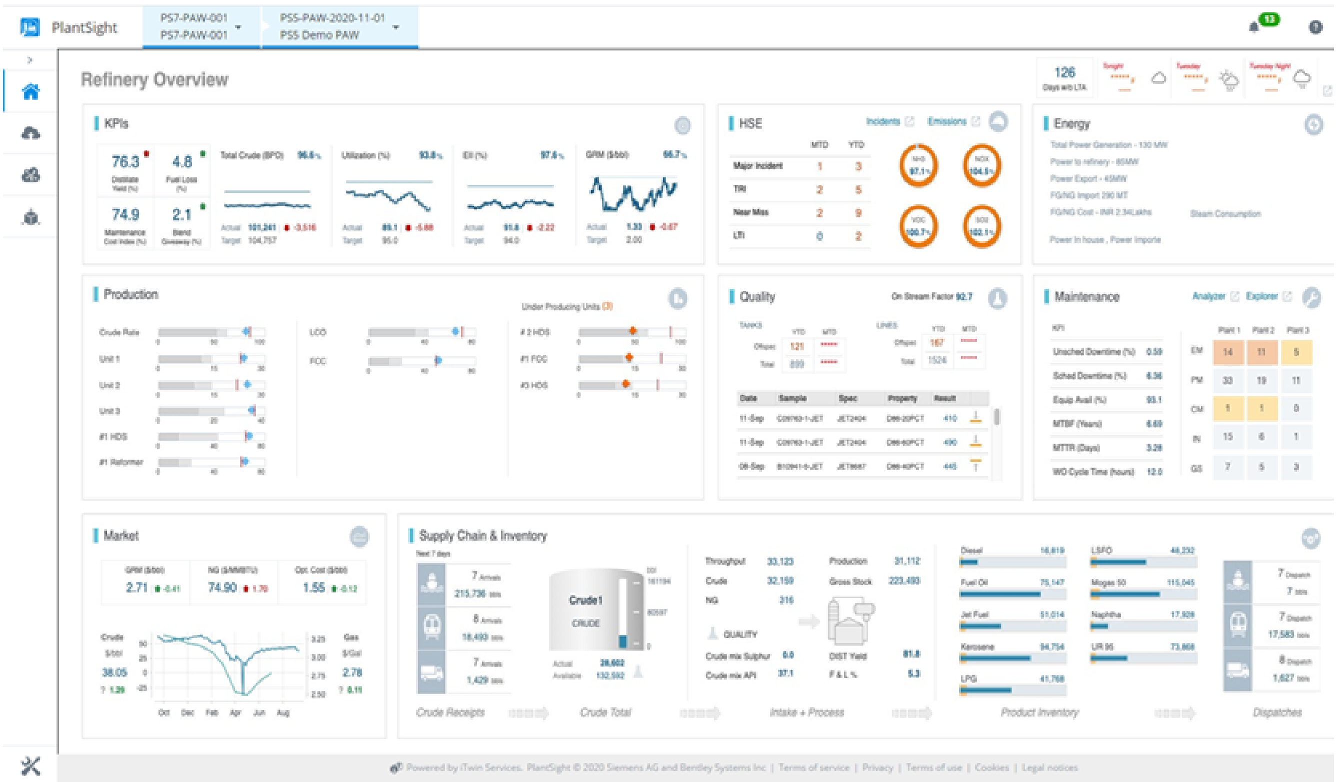
With the help of consistent data from the central asset portal, the asset status can be visualized holistically. This allows efficient decision making and proactive operational management.

Monitoring and management of critical assets

If the most important plant components, such as critical equipment or particularly energy-intensive units, are to be monitored first, the digital twin can be used to map detailed condition monitoring with extended analysis options for diagnostic data, including predictive maintenance strategies. Consolidated information from recorded time series allows to derive trends in terms of failure probabilities. In this way, risk-based maintenance measures can be established. Through continuous monitoring of turbines, compressors, pumps or other critical assets, valuable insights can result from comparing current performance data with the manufacturer's specifications and better decisions can be made. However, the digital twin also offers more far-reaching possibilities for action, such as, enabling different scenarios to be evaluated virtually. Also, which operating mode optimizes yield, which extends maintenance times to the maximum, etc.? Even more, like how do changes in operating conditions affect the equipment, the process or the performance?

Process and asset performance management

The digital twin can also be used for real-time production optimization. The constant feedback of information from the plant allows control loops to be optimized, energy and quality management to be carried out, and impact analyses (FMEA - Failure Mode and Effects Analysis) to be performed. The integration of simulations into the digital image holds enormous potential for online production optimization. What was previously initiated by manual intervention in the process can in future be carried out virtually and independently by the system on the basis of the simulation and with the aid of machine learning methods, or serve as action suggestions for the operator.
Asset performance management with holistic access to all relevant static and dynamic operating data (including historical data) uncovers mismatches between target and actual, and simultaneously supports root cause identification including investigation, diagnosis and remedial action planning for all connected assets in the production plant.



The basis for all optimization measures - at process, equipment, or operational level - is validated and contextualized data and the derivation of corresponding key figures.

Increase in operating performance

By storing production KPIs, a digital twin can also be used to introduce effective operational performance management: quasi-real-time operational decision support that links the objectives of the management level in its business context with the corresponding operational drivers creates the possibility for sustainable and continuous improvement of overall operational performance and allows a target-oriented response to exceptional cases.

In addition to this KPI-driven performance influence, both current and future planning (day, week, month, year, 5 years) can be compared with current operational performance via the holistic definition and management of operational objectives and their logical linkage with warnings of soft and hard limit violations (high-high, high, low, low-low) and driving modes can be adjusted accordingly.

Optimization of health, safety and environmental protection measures

The information provided by the digital twin also contributes to acute error prevention during ongoing processes. Since insights into irregularities can be derived from the data sets, the future design of work processes can be optimized. In addition, the quality of work can be improved by identifying critical factors and detecting errors or risks in good time through automatic analyses, thereby avoiding consequential damage. Immersive training simulations in a virtual plant environment help to train situational awareness and increase familiarity with the plant before personnel are even on the physical plant.

A digital twin can also be used to record all events that affect the environment. Thus, all relevant information related to such events, for example, venting of pressure systems, flaring of gas, disposal of waste materials, etc., is recorded in accordance with regulatory requirements. It will ensure that appropriate reports can be delivered to regulatory authorities, but can also serve as an audit trail should incidents / accidents occur.

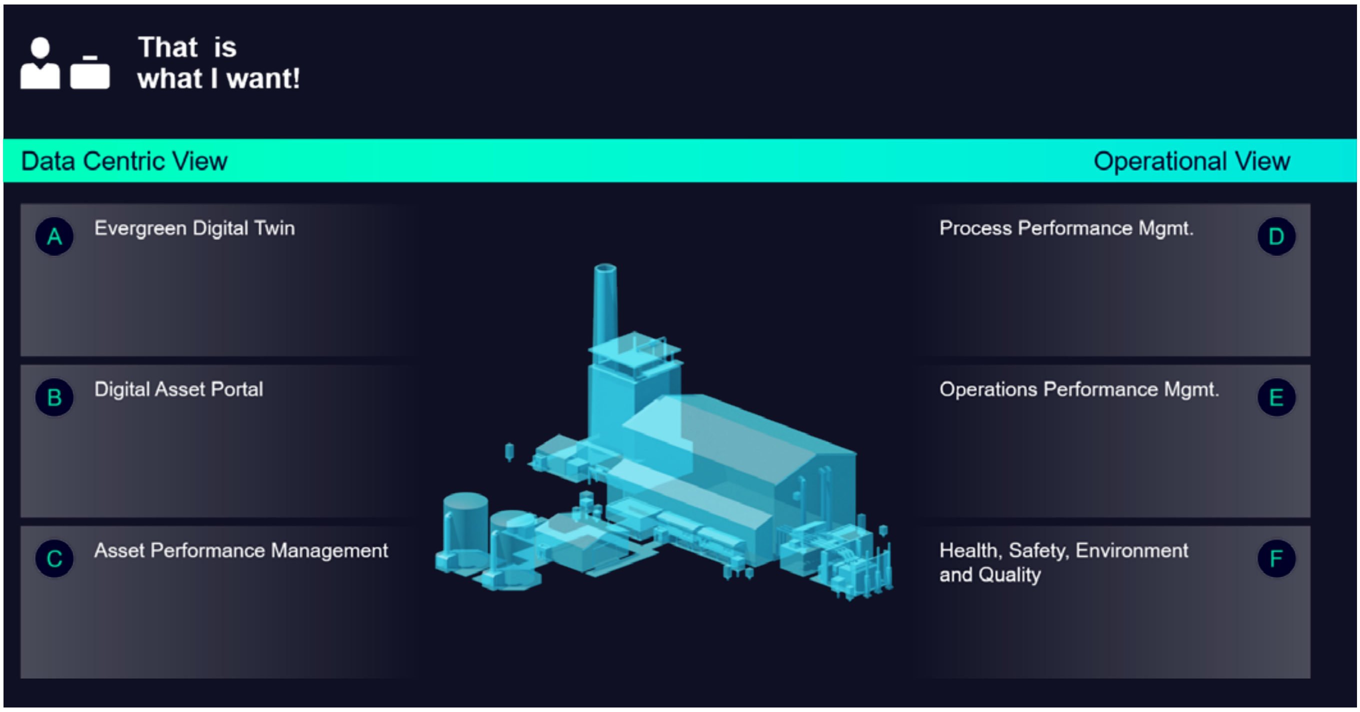
Whatever the starting point or the desired goal, the digital twin can be used profitably in many places. With Plantsight, it becomes possible to first tackle a specific challenge and then target new scenarios down the road - once integrated, data assets can be used in a variety of ways, and the more data that is made accessible, the more numerous the possibilities.

Virtual training simulations based on integrated data help to increase plant safety.

Conclusion

Most plants in the process industry have massive data resources from which insights could be drawn in a variety of ways that would help optimize processes and improve operational performance, safety, and product quality. A digital twin carefully set-up creates added value from these data treasures, which are accessible, unified, contextualized and made available via a web portal in a role-based manner with the help of Plantsight. 1D, 2D and 3D data brought together from many disparate sources provides integrated consistent data. Model-based data analysis uncovers relevant correlations and provides well-founded insights. In Plantsight, the real and virtual worlds can be reconciled through an all-encompassing digital twin platform. This also means consistent plant data and documentation throughout the lifecycle of the plant and this creates a better starting point for simulations, optimizations, expansions or new planning.

In addition to the aspect of predictive, demand-oriented maintenance, resource-saving process control is achieved particularly by using algorithms for (dynamic) online optimization. Feedback from the real to the virtual world allows the calculation of utilization and effectiveness data, and by means of integrated artificial intelligence and machine learning technologies in the simulation of the virtual plant, decisions for future plant modifications or operation can be verified in advance. Plantsight transforms raw data into a digital twin and provides all stakeholders with consistent reliable information that is user and situation specific, enabling informed decision making at all times.



Plantsight supports you in the realization of a wide variety of use cases and, with its integrated information space, ensures consistently well-founded decision-making.

No matter at what point the creation of the Plantsight digital twin is started, the inherent processes provide a self-reinforcing dynamic. Merging data into an integrated information space provides functional context to physical representations and vice versa. The more Dark Data that can be made visible, tagged, validated, and linked to other available information, then the more valuable and context-rich the information becomes. The value of data is also enhanced by ease of access, because the more intuitive it is to update and add to asset information, the higher the level of asset documentation and the more reliable and complete the information will be for other stakeholders. To make long-term success measurable, integrated technologies ensure that the quality and completeness of asset information is always monitored and that changes in asset performance are comparable in terms of asset characteristics or asset changes.

< tl; dr>

Digital twins open up enormous optimization potential for optimization throughout the entire plant life cycle. Plantsight as an overarching digital twin platform brings the real and virtual worlds into harmony.

Bibliography

Bamberg, Andreas; Urbas, Leon et al.: „Was den Digitalen Zwilling zum genialen Kompagnon macht". In: Chemie Ingenieur Technik, 3/2020.

BITKOM-Studie: „Big Data im Praxiseinsatz - Szenarien, Beispiele, Effekte". Berlin, 2012.

Demchenko, Yuri; Grosso, Paola et al.: "Addressing big data issues in Scientific Data Infrastructure": In: International Conference on Collaboration Technologies and Systems (CTS), 2013.

Folmer, Jens et al.: „Big und Smart Data. Herausforderungen in der Prozessindustrie". In: atp magazin, S. 58-69, März 2017.

Gölzer, Philipp: „Big Data in Industrie 4.0. Eine strukturierte Aufarbeitung von Anforderungen, Anwendungsfällen und deren Umsetzung". Dissertation vorgelegt an der Friedrich-Alexander-Universität Erlangen-Nürnberg, 2016.

Hill, Jürgen: „Siemens erweitert Digital-Enterprise-Angebot". In: Computerwoche vom 25.02.2019: https://www.computerwoche.de/a/siemens-erweitert-digital-enterprise-angebot,3546601,2 last accessed 2020/09/01.

Laney, Doug: „3D Data Management: Controlling Data Volume, Velocity, and Variety". META Group Whitepaper, 2001.

McAffee, Andrew; Brynjolfsson, Erik: "Big Data. The Management Revolution". Harvard Business Review, Oktober 2012.

Reinsel, David; Gantz, John; Rydning, John: "The Digitization of the World – From Edge to Core". An IDC Whitepaper, November 2018.

Schüller, Andreas et al.: „Der Digitale Zwilling in der Prozessindustrie". In: atp magazin, S. 70-81, 2019.

Volk, Matthias und Hart, Stefan Willi: „Big-Data: Bestimmung der Big-Data-artigkeit von Projekten" In: Schenk, Michael (Hg.): „Nutzung digitaler Methoden und Modelle in Engineering and Construction im Anlagenbau. 25. Industriearbeitskreis »Kooperation im Anlagenbau« 21. Juni 2016", S. 29-43.

Wang, JunPing; Zhang, WenSheng et al.: „Industrial Big Data Analytics: Challenges, Methodologies, and Applications". In: In IEEE Transactions on Automation Science and Engineering, 2018.

Siemens AG

Digital Industries
Process Automation
Oestliche Rheinbruckenstrasse 50
76187 Karlsruhe, Germany

Subject to change without prior notice
Article No. DIPA-B10189-00-7600
Printed in Germany