Big Data – Industry 4.0 – Quality 4.0?

Dr.-Ing. Edgar Dietrich, Q-DAS® GmbH & Co. KG

Nowadays, topics like big data, big smart data, smart data, predictive analytics, Industry 4.0 or smart factory are discussed more and more frequently. How do these topics and the corresponding developments affect the field of quality management? Is there something like quality 4.0 as we might therefore assume? Q-DAS® has an answer to this question: "quality4industry"!

These terms represent subject areas that are inextricably linked. For a better explanation of these relations, this article first focuses on the meaning of these terms and their relevance to quality in industrial production. It is also clear that these terms just mark the beginning of a whole new development. We cannot foresee the impacts of this development completely for the same reasons that few could have thought that the commercialization of the Internet 20 years ago would affect us in such profound ways today. It is the same with the idea of Industry 4.0 - only time will tell the impacts resulting from these new opportunities. We can assess the benefits, advantages and disadvantages only in the future. However, we already have to deal with all that is known and venture a glimpse into the future, develop possible scenarios and estimate their probability of occurrence. This is the only way to set the right course early.

This article is an attempt to explain the topics, terms and issues linked to this new development from a Q-DAS® perspective, to provide answers to the questions and to give helpful suggestions. Feel free to contact us for a personal discussion of these topics.

Big Data

The German Bitkom [1] association defines big data in an article as follows: "Big data refers to the economically reasonable acquisition and application of decision-relevant knowledge gained from qualitatively versatile information structured in different ways. These pieces of information undergo rapid changes and are provided in an unknown amount." According to this definition, big data means that companies try to measure, analyze, calculate, assess and evaluate data from many different unstructured sources that frequently change, so far.

Why does Big Data exist at all?

The fundamental condition is the possibility to store plenty of information cost-efficiently in the smallest space. Further reasons are the many automated sensors and recording systems as well as the high transfer rate and processing speed providing and transferring real-time information generally without any manual input. These aspects lead to the creation of very large data volumes whose extent is currently unknown. As an example, the generation of five exabytes (one exabyte is one quintillion bytes) took

- more than 1000 years by 2003 [2]
- 2 days in 2011
- 10 minutes in 2013

This trend will continue and there is no end in sight. Figure 1 shows an example of the data traffic over time in conjunction with the respective technology. In 2020, this data traffic will be more than 50 times the traffic we have today.

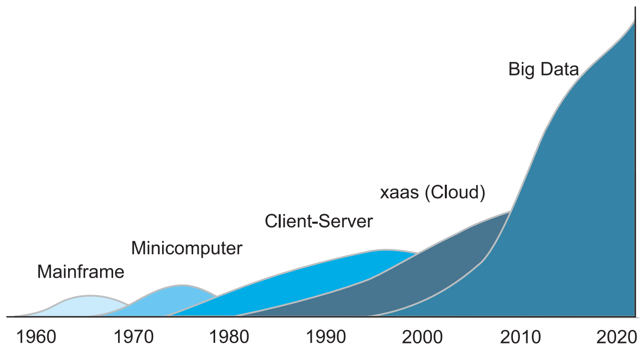

Figure 1: Growth of Data [2]

It is beyond all question that the relations between single aspects of the economic life have become more and more complex due to globalization, transport and communication options - and the level of complexity still rises. This is the reason why people and companies need high-quality decisions in order to control processes and operations. Figure 2 shows a simple diagram illustrating how to find high-quality solutions to the respective tasks based on the generated and stored information.

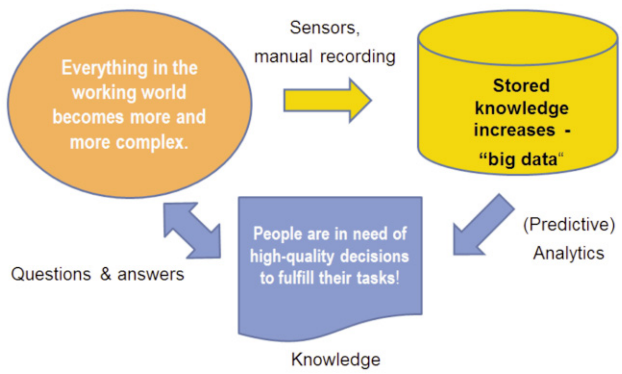

Figure 2: Valid Knowledge Acquisition

These answers are gained from the results of statistical analyses. The recorded data are analyzed and the result of these analyses is provided in the form of statistics and in graphics. Q-DAS® refers to this process as knowledge acquisition.

There is still another question that might arise: "What is so new about big data?" Even in the past we collected, stored and evaluated plenty of data after all. Figure 3 shows the main data sources and the corresponding data management systems in industrial production and from a Q-DAS® perspective. Additionally, the diagram includes the different kinds of data processing and the different displays of results. In case of big data, we have to add the "Internet of things and services" (new data source) to the existing data sources.

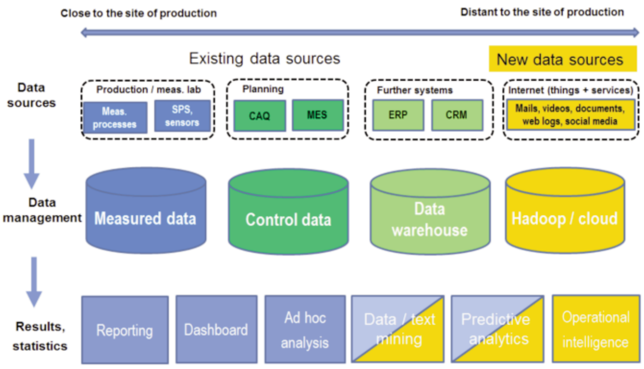

Figure 3: Big Data = something traditional + something new

From this point of view, big data refers to the recording and storage of huge amounts of data provided by various sources and the "real-time" (near real-time) analysis of these data in order to cope with a current task or answer one of the latest questions.

How to Use Big Data Smartly

What is the problem with big data today? Collecting and storing data is just one side of the coin. The other side of the coin is much more exciting. It is the beneficial analysis and evaluation of the stored information. (Predictive) Analytics is the catchword. Predictive analytics apply statistical procedures, prediction models, optimization algorithms, data mining, text mining and image mining in order to achieve the desired results. How well these procedures work, however, depends on how well the data basis agrees with reality, i.e. on how representative it is. It is assumed today that the major part of all data is unstructured; experts even estimate that more than 80 percent of all data is unstructured [3]. A vast amount of collected and stored unstructured data from different data sources usually become a data graveyard. This is not a useful basis of data for an analysis leading to the desired results; on the contrary, the gained results are incomplete and fuzzy. There will not be any confidence in the acquired knowledge and people will always have doubts about the benefit. For this reason, you have to create a corporate infrastructure in order to record a vast amount of data validly from various data sources and to manage and evaluate them in a structured way. Instead of big data, you have to create "(big) smart data" right from the start. This is the only way to evaluate data in real time and to provide validated results within seconds. By creating big smart data you acquired knowledge in order to extract the desired added value from this new development.

You meet this requirement by performing a "data consolidation" (see Figure 4) after recording the data but before storing them.

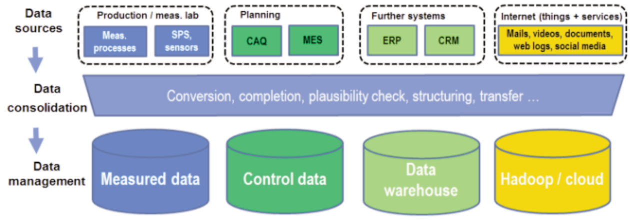

Figure 4: Data Consolidation

Smart Data through Data Consolidation

Nowadays, there are various types of data sources; however, these sources normally do not communicate with other systems, least of all with one another. In order to structure the data from the different sources, you have to specify which data you want to store to which storage location, how you want to store it and which amount of data you want to store. Currently, there is a lack of standards defining the communication of the different sources with their environment. Standards regulating the transfer and storage of data would stop the proliferation of data provided by different data systems and would lead to data consistency. These standards would be the basis for reliable evaluations.

The Q-DAS® ASCII Transfer Format covers some of these Requirements

The "Q-DAS® ASCII transfer format" made by Q-DAS® has become a worldwide standard regarding the transfer of measured data in a specific file format. The AQDEF document (Advanced Quality Data Exchange Format) describes some aspects of this format. And meanwhile, many companies are required to fulfill the demands of this document when providing newly purchased measuring instruments. These demands ensure that the respective data fields are available for the required information in order to store it to the database in a well-structured way.

However, even if these fields are available, there is no guarantee that they are completed correctly. This is the reason why the plausibility check is an important aspect of data consolidation since it evaluates the completeness and meaningfulness of the transferred information. In order to run a plausibility check, you have to introduce some mechanisms fulfilling this task. In general, you can only detect erroneous data or missing information and, if required, add some further information directly upon recording data from the source.

We will develop the Q-DAS® ASCII transfer format step by step with regard to new data sources (see Figure 4) and extended applications like traceability, grouping of parts and storage of process parameters with or without corresponding measured values. Based on the requirements of new data sources, we will add new fields for storing data to the database. In case a data source does not support the Q-DAS® transfer format, the respective data has to be converted completely into the respective structure by using a converter.

(Predictive) Analytics

In order to fulfill certain tasks and answer important questions, all your data must be analyzed automatically and, if possible, in real time with the help of statistical procedures. These procedures help you to detect specific patterns and relations. You have to process the results according to the task and then provide them in a suitable form to the respective recipient (see Figure 5).

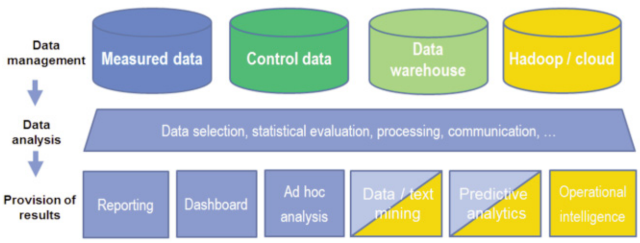

Figure 5: Data analysis

The following two examples are quite striking and illustrate the new opportunities.

- Amazon was awarded patent no. US 8.615.473.B2 (method and system for anticipatory package shipping) in the USA with the aim to ship a product (i.e. a parcel) even before the customer places an order. This system predicts the probability of a possible order based on the customer behavior patterns and ships the product in case of a good chance of success. If customers order the respective product, they will receive it within 2 hours.
- Some navigation systems send – if enabled – information about position and speed to the central data pool. In case the police gain access to this type of information – and this is exactly what happened in an EU country – they can find out at which neuralgic location the number of speeding violations is significant in order to make radar speed checks more efficient.

Validated Evaluation Strategies

Q-DAS® developed evaluation strategies referring to the processing of measured values in industrial production. These strategies fulfill tasks and answer questions from the field of quality assurance. This is the reason why companies created guidelines or process instructions included in the Q-DAS® software so that the software performs an automated statistical evaluation of the stored smart data.

Based on the consistent information (smart data) stored in a well-structured way you may select the desired or required queries for data analysis with the help of flexible filter criteria. Afterwards, the provided data is evaluated according to the specified evaluation strategy. These strategies lead to valid results provided in the required form (reports, dashboard …), e.g. in tables and graphics, depending on user requirements.

Industry 4.0

Industry 4.0 is also referred to as an intelligent factory, smart factory or digital factory. In terms of production, it is characterized by adaptability, resource efficiency and ergonomics [4]. Additionally, there is the extensive integration of customers and suppliers especially in order to include information from the supply chain already in the manufacturing process of a product and to ensure traceability.

The technological basis of Industry 4.0 are cyber-physical systems, also referred to as CPS or embedded systems, linking the Internet to the perception of the environment by means of sensors. In other words, the real world and the virtual world are growing closer and closer together. Particular aims of these systems are the

- self-configuration
- self-optimization
- self-diagnosis and
- cognition

of the entire factory in order to produce unique, strongly individualized products under the conditions of highly flexible large series production. Industry 4.0 makes it easier to consider single customer requirements and to produce individual items profitably. The batch size one competes with mass production.

As an example, a supporter of Industry 4.0 claims:

"Nowadays, production has already fallen into the capacity trap. Clocked and synchronized systems (e.g. assembly lines) are to blame for this development, i.e. the achievement of the last industrial revolution. In case of push markets, these systems work as long as the markets are overstocked with products and customers buy them. However, the situation has changed considerably!"

In case you want to implement Industry 4.0 as described above, a product as we know it today has to become an "object with a memory" (digital object memory) that does not only know its own status but also selects the next evaluation and processing steps. So Industry 4.0 is part of a networked, intelligent world tracking the entire life span of a product from its design to its creation and assembly to its benefit in everyday life. Industry 4.0 helps to control processes explicitly and to optimize these processes, which is not yet feasible today. A more extensive supplier integration unleashing a new wave of opportunities for customer and supplier benefits is also required. All these facts make completely new production methods and value-added networks possible.

The requirements for this development are the clear identification of individual parts and components, the increasing performance of sensors to observe the real world, the storage options and a network of different systems. The vision of Industry 4.0 are highly connected structures available for the storage, evaluation and respective application of any kind of data. This is considered to be the production system of the future since it will be very important to control mass data in the future. In other words, big data relating to industrial production - this is our vision today!

Industry 4.0 is Part of Big Data

Industry 4.0 can surely be considered as part of big data since not all stored data is relevant for the industrial production. However, new data sources like the Internet of things and services are also part of Industry 4.0. Even when creating products in a world of Industry 4.0, quality plays an important role since a highly automated production does not ensure at all that the required characteristic specifications are met. Quality assurance measures as known from the "old" world still apply, they are just adapted, if necessary.

Today, companies distinguish between an internal network (close to the site of production) in production (see Figure 4) without any possible access to the Internet and an external network (distant to the site of production) possibly communicating on the Internet. So the following question provides more topics to discuss. "Where do you store and process which data?" Opinions differ considerably as the following two statements issued by experts show

"The amount of data will rise sharply in such cyber-physical systems which will make the management and processing of data in a company from the manufacturing industry basically impossible. Companies will be bound to move their information and computing power to the cloud, even though the rights to the data and information currently remain unclear!"

By contrast:

"In my opinion, the recording and direct processing of measured values does not take place in the cloud, even in the future!"

There is no difference between these two statements when distinguishing between a "public cloud" and a "private cloud" (Figure 6). The private cloud grants access to the virtual IT infrastructure of your own organization. As an example, the company provides a terminal server environment in the cloud. The company does not only store the data but also runs the application there. As opposed to the private cloud, the public computer cloud provides Internet access to the virtual IT infrastructure for the general public. Customers may book IT infrastructure from public cloud service providers. The price depends on how much the infrastructure is actually applied or how much storage capacity you need. At the same time you do not have to invest any capital in an infrastructure of computers and datacenters.

From a present-day perspective, it is hard to believe that companies from the manufacturing industry will move their production data to the public cloud. This is what our experience with customers proves. Confidentiality is more essential than ever. For the exchange of data between supplier and customer, companies will provide at least some part of the information in the private cloud; however, only a specified group of participants will have access to the cloud.

Q-DAS® offers solutions even for this situation. As an example, suppliers record measured data by using the web application of the customer and are able to provide it immediately to the respective customer. At the same time customers make the Q-DAS® Statistical Server available to their suppliers and the suppliers use it to evaluate the measured data. The supplier records the measured data with the help of a Q-DAS® product and sends it to the Q-DAS® Statistical Server in the AQDEF format. The Statistical Server evaluates the data and sends it back in the specified form. This is also referred to as a SaaS (Software as a Service) application. In terms of Industry 4.0, the Q-DAS® Statistical Server may be considered as an embedded system today and thus it is regarded as part of the overall system.

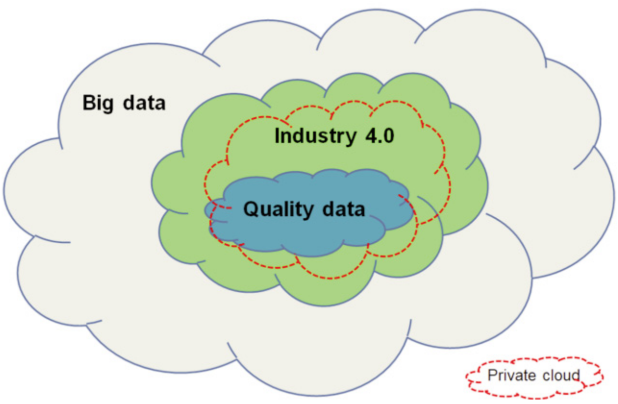

Figure 6: Internal network - private cloud

Critical Comments on Industry 4.0

The idea of computer integrated manufacturing (CIM) was created more than twenty years ago. In hindsight, this idea is quite similar to today's concept of Industry 4.0; however, there is hardly anything left of the CIM idea. We have to admit though that the sensor technology, the transfer and storage of data and the high-performance evaluation were not possible back then with the previously available systems. Under these circumstances, the odds are more in favor of Industry 4.0 today.

While the supporters of Industry 4.0 are euphoric and claim that Industry 4.0 is already a reality, critics argue that Industry 4.0 is nothing but an idea. However, it will be a bumpy road to implement the vision of Industry 4.0 demanding lots of patience (instead of euphoria). Different kinds of systems will not communicate without the standards and clear regulations mentioned above. Even though we are able to meet the technical requirements today, Industry 4.0 will fail without these standards and regulations. A nationwide implementation of the Industry 4.0 idea will not be realized before the required standards are developed and accepted. The problem is that the development and acceptance of these standards are the aspects that turn out to be difficult in practice.

As is the case with anything new, we will only see the true benefits in the next few years after gaining some experience. As is often the case in life, the happy medium might be the right way to go: "A specific amount of products, especially mass-produced articles, will be produced conventionally, even in the future, whereas products produced in small numbers or even with batch size one will be manufactured in a CPS environment!"

Q-DAS® and Industry 4.0

By implementing the idea of Industry 4.0, different tasks available in industrial production will grow closer and closer together. Isolated applications will become networked applications. The processing of information will gain in importance. Industry 4.0 cannot succeed without software-intensive applications embedded in an overall system. In terms of Industry 4.0, the Q-DAS® software has been on the right track for a long time.

The Q-DAS® products meet all requirements – at least with respect to the field of application close to the site of production (see Figure 7) – beginning with the generation of information (data sources) to data consolidation and data management to the analysis of data and the provision of results.

Our software communicates with the data sources in real time via RS-232, USB, Ethernet or Wi-Fi, depending on the type of interface or based on the Q-DAS® ASCII transfer format. The options available to communicate

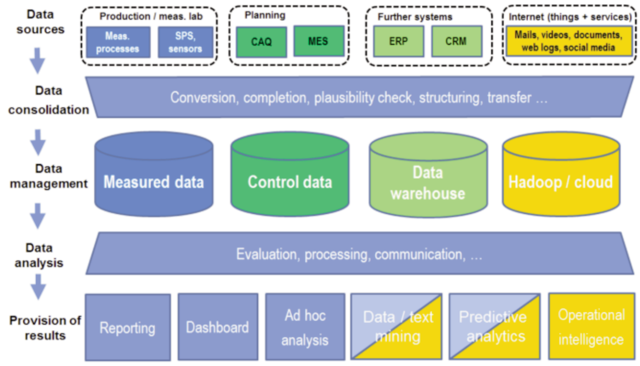

Figure 7: From the data source to the display of results

with different measurement processes (coordinate measuring machines in particular), SPS controls, SPC, CAQ and MES systems and the ERP system of SAP based on the Q-DAS® ASCII transfer format are a worldwide accepted standard now.

In addition, Q-DAS® provides software tools for important fields of "data consolidation" and "data transfer" to a "(Q-DAS®) database" in order to create smart data from the unstructured mass data. These tools help to structure, validate and complete the data.

The Q-DAS® statistics library (Q-DAS® Statistical Server) includes the required analysis tools (data analysis) available as an embedded system. As an example, the Q-DAS® Statistical Server has been used as an embedded system via the STI interface of SAP for a long time in order to display graphics and results directly in SAP. Several manufacturers of SPC systems also use this option.

Reporting tools are available for the display of results. They display results in different tables or graphics depending on the respective task and target groups, communicate them to the recipients and store and process compressed results.

Q-DAS® CAMERA Concept as Part of Industry 4.0

On the one hand, there is the availability of the single tools but on the other hand there is the interaction between these tools required in order to create an automatism. For this reason, we developed the Q-DAS CAMERA® Concept (see Figure 8) helping companies to create a well-structured, comprehensive system for the data flow reaching from the data source to the management of data to the processing of results.

By adding the aspects of the Q-DAS CAMERA® Concept to the contents of Figure 3, you gain the diagram shown in Figure 9. The single packages of data are transferred from the different data sources to the data consolidation and stored to the database structure there. Depending on the respective query (task), the analysis tools select the data from the database and evaluate them statistically according to the respective evaluation strategy (e.g.

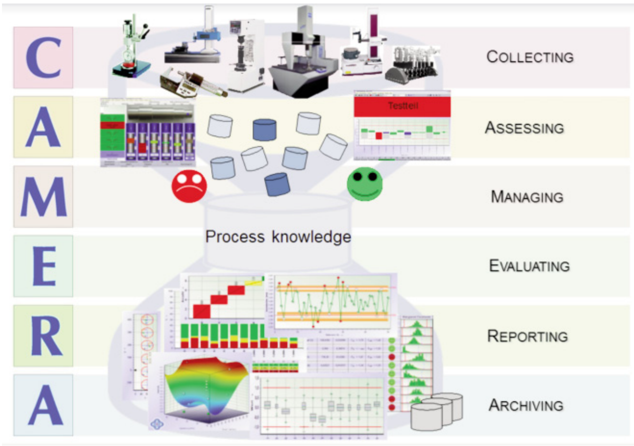

Figure 8: Q-DAS CAMERA® Concept

company guidelines). The results are processed respectively and provided to the responsible recipient. In terms of quality assurance in industrial production, the Q-DAS CAMERA® Concept even implements the most basic approach of the big data and Industry 4.0 ideas.

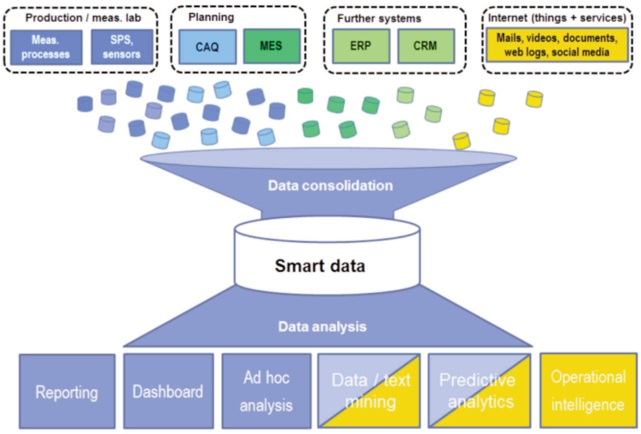

Figure 9: Q-DAS CAMERA® Concept as part of Big Data and Industry 4.0

Of course, we have not been able to integrate all available data sources into the Q-DAS® software yet, particularly the "new data sources" are still missing. However, we continuously develop new interfaces according to customer requirements and on customer request. Step by step, we develop and include additionally required analysis tools in our software, too.

Digital Object Memory from a Quality Perspective

In terms of Industry 4.0, products become more and more intelligent from the produced component to the finished product, i.e. they will have a digital memory in the future in order to control their way through the (production) environment. Quality will play a central role. A high level of automation alone does not ensure at all that the processed product complies with the characteristic specifications demanded by and arranged with the customer. In order to be able to meet these requirements, the object memory has to include sufficient information assuring the quality of the product.

Typical question a product has to answer in terms of Industry 4.0:

- Which measurement process is capable of evaluating the respective characteristics?
- What are the capabilities of the respective machines and production facilities?
- What is the condition of the available processing tools?
- What is the position of the characteristic within the specification?
- Does the outer diameter match a bore hole?
- What is the control limit of the characteristic while it is analyzed?
- How is the quality of a part in the overall evaluation of all characteristics?
- What are the process parameters needed to adjust a process step?
- What is the expected life span of the product?
- ...

In order to answer all these and many more questions in the respective situation from the product's perspective, sufficient information have to be stored in the object memory. How are these pieces of information generated? Figure 10 shows typical tasks indicating how stored data (smart data) are evaluated statistically based on specified evaluation strategies and how to store these results in one or several databases of results. You may perform this kind of analyses continuously by using Q-DAS® products in order that the stored results are always up to date. The digital object memory feeds on this information.

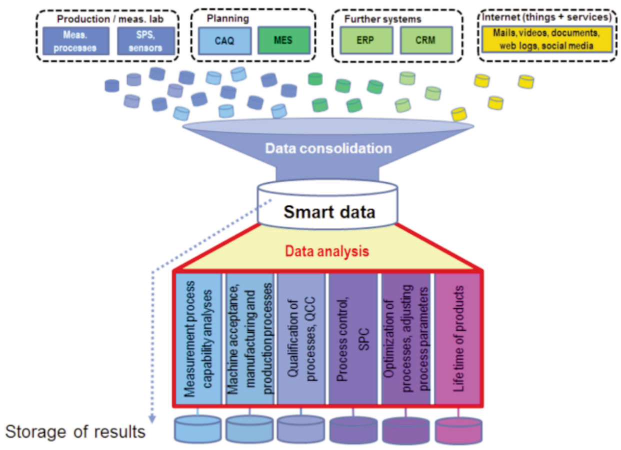

Figure 10: Basis of Information for the digital product memory

The product may retrieve e.g. the following information from the object memory:

- A product characteristic knows when it has to be inspected and knows the suitable measurement process, too.
- The results of the capability analyses of the respective machines, manufacturing and production facilities are available.
- Processes are qualified. You may find out which product characteristic has been produced by which facility.
- The process characteristic is evaluated in order to determine whether it meets the requirements with due regard to the uncertainty of measurement.
- The optimized parameters are known for the following process steps and will be adjusted automatically.
- The expected life span of the product is known from reliability analyses.
- …

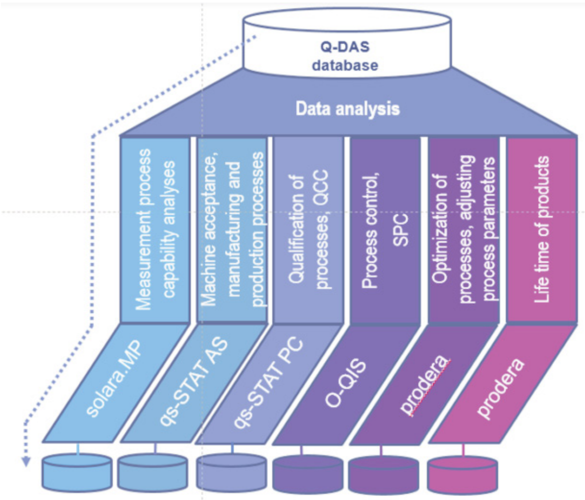

Figure 11: Q-DAS® products for creating a digital object memory

This leads to the knowledge acquisition mentioned in Figure 2.

The article by Michael Radeck (see page 9) shows possible Industry 4.0 scenarios.

The article of Stephan Sprink (see page 11) deals with the following issue: "Agile systems already pursue the idea of Industry 4.0".

Figure 11 shows the Q-DAS® products helping to create the basis of information for a digital object memory in terms of quality. Moreover, version 11 of the Q-DAS® products provides a project manager in order to control single tasks.

A product attached to a memory must be able to communicate with its environment (see Figure 12), of course. Unfortunately, we still have a long way to go until we will reach this goal. The standards mentioned before are still missing. Only they can specify how these systems communicate with one another and define which types of information are available and where.

Discuss this seminal and exciting topic with us and get involved in quality4industry!

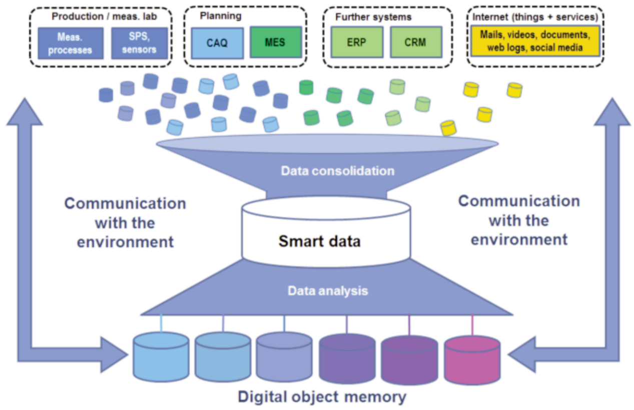

Figure 12: Product communicating with its environment

We will be pleased to contribute to the implementation of Industry 4.0 in your company with the help of the Q-DAS CAMERA® Concept. I look forward to receiving your questions, notes and critical comments by e-mail (edgar.dietrich@q-das.de).

References

[1] Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte, Bitkom, Berlin 2012.

[2] Digitale Unternehmen – Herausforderungen und Potenziale, Software AG Darmstadt 2013.

[3] IDC Digital Universe Study, Juni 2010.

[4] Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0, acatech, Frankfurt 2013.

In order to avoid irritations, we decided against mentioning the authors of the quotes. However, the names are known to the author of this article.