

Bernd Valeske*, Ralf Tschuncky, Frank Leinenbach, Ahmad Osman, Ziang Wei, Florian Römer, Dirk Koster, Kevin Becker, and Thomas Schwender

Cognitive sensor systems for NDE 4.0: Technology, AI embedding, validation and qualification

Kognitive Sensorsysteme für ZFP4.0: Technologie-Entwicklungen mit eingebetteter KI und Konzepte zur Validierung

<https://doi.org/10.1515/teme-2021-0131>

Received December 14, 2021; accepted March 6, 2022

Abstract: Cognitive sensor systems (CSS) determine the future of inspection and monitoring systems for the non-destructive evaluation (NDE) of material states and their properties and key enabler of NDE 4.0 activities. CSS generate a complete NDE 4.0 data and information ecosystem, i. e. they are part of the materials data space and they are integrated in the concepts of Industry 4.0 (I4.0). Thus, they are elements of the Industrial Internet of Things (IIoT) and of the required interfaces. Applied Artificial Intelligence (AI) is a key element for the development of cognitive NDE 4.0 sensor systems. On the one side, AI can be embedded in the sensor's microelectronics (e. g. neuromorphic hardware architectures) and on the other side, applied AI is essential for software modules in order to produce end-user-information by fusing multi-mode sensor data and measurements. Besides of applied AI, trusted AI also plays an important role in CSS, as it is able to provide reliable and trustworthy data evaluation decisions for the end user. For this recently rapidly growing demand of performant and reliable CSS, specific requirements have to be fulfilled for validation and qualification of their correct function. The concept for quality assurance of NDE 4.0 sensor and inspection systems has to cover all of the functional subsystems, i. e. data acquisition, data processing, data evaluation and data transfer etc. Approaches to these objectives are presented in this paper after giving an overview on the

***Corresponding author: Bernd Valeske**, Fraunhofer IZFP, Institute for Nondestructive Testing, 66123 Saarbruecken, Germany; and University of Applied Sciences – htw saar, 66117 Saarbruecken, Germany, e-mail: bernd.valeske@izfp.fraunhofer.de, ORCID: <https://orcid.org/0000-0002-2758-9754>

Ralf Tschuncky, Frank Leinenbach, Florian Römer, Dirk Koster, Kevin Becker, Thomas Schwender, Fraunhofer IZFP, Institute for Nondestructive Testing, 66123 Saarbruecken, Germany

Ahmad Osman, Fraunhofer IZFP, Institute for Nondestructive Testing, 66123 Saarbruecken, Germany; and University of Applied Sciences – htw saar, 66117 Saarbruecken, Germany

Ziang Wei, University of Applied Sciences – htw saar, 66117 Saarbruecken, Germany

most important elements of CSS for NDE 4.0 applications. Reliable and safe microelectronics is a further issue in the qualification process for CSS.

Keywords: Cognitive sensor systems, NDE 4.0, materials data space, IIoT, OPC UA, DICONDE, advanced microelectronics, compressed sensing, trusted AI, validation and qualification process.

Zusammenfassung: Kognitive Sensorsysteme (KSS) sind maßgebliche Wegbereiter für Entwicklungen zur zerstörungsfreien Erfassung und Bewertung von Materialzuständen. Erst durch KSS und durch Einbettung der von ihnen generierten Informationen in digitale Datenplattformen im Konzept von Industrie 4.0 lässt sich der Materialzyklus vollständig beschreiben. Die so erzeugte digitale Produktakte erlaubt die Optimierung in jeder Phase und entlang des gesamten Material- und Produkt-Lebenskreislaufs (vom Rohstoff bis zum Recycling/Wiederverwertung). Erst durch KSS und ihre Integration in das industrielle Internet der Dinge (IIoT) wird das Modell der (digitalisierten) Material-Kreislaufwirtschaft ermöglicht. Die klassische zerstörungsfreie Mess- und Prüftechnik (ZFP) erfährt dabei einen dramatischen technologischen Wandel. Eingebettete künstliche Intelligenz (KI) sowohl auf der Mikro-Elektronikebene (neuromorphe Hardware) als auch auf Software-Ebene ist neben der beherrschten Prüfphysik bestimmend für zukunftsweisende KSS. Für den industrietauglichen und rechtssicheren Einsatz solcher Sensorsysteme sind Validierungsvorschriften sowie Normen und Standards zur Überprüfung bzw. (Bescheinigung) der korrekten Funktion (d. h. der integrierten Datenbewertung und Material-Zustandsinformation) unabdingbar. Dazu müssen alle Teilsysteme beginnend von der Datenakquisition (Messung am Sensor) über die Signal- und Datenverarbeitung bis hin zur anwendungsspezifischen Informationsaufbereitung (Bewertung) nach objektiven Vorgehensweisen überprüfbar sein und zertifiziert werden. Der vorliegende Beitrag stellt die wichtigsten Technologie-Module von KSS und aktuelle Konzepte zu ihrer Validierung und Zertifizierung vor.



Figure 1: Material data space: Data ecosystem along the complete lifecycle of materials and products (enabled by CSS and NDE 4.0).

Schlagwörter: Kognitive Sensorsysteme, ZFP4.0, neuro-morphe Mikroelektronik, IIoT, OPC UA, DICONDE, erklärbare KI, Compressed Sensing, Validierung, Qualifizierung.

1 Introduction to cognitive sensor systems and NDE 4.0

Cognitive sensor systems (CSS) determine the future of inspection and monitoring systems for the nondestructive evaluation (NDE) of materials states and properties and key enabler of NDE 4.0 activities. NDE 4.0 enables acquiring data for describing and predicting the materials state along its complete lifecycle [92]. CSS are elements of the Industrial Internet of Things (IIoT) [89, 91]. The holistic data along the lifecycle of materials is stored in the corresponding data cloud [93], cp. Fig. 1. The most common platform is the material data space, which is part of the international data space [40]. Interfaces (i. e. protocol specifications and data formats) are essential for reliably transferring and supplying data from the sensor devices to the eco-system platforms (cp. Section 2).

Applied artificial intelligence (AI) is a further key element for progress in cognitive NDE 4.0 sensor systems (cp. Section 4). On the one side, AI is embedded in the sensor's microelectronics (e. g. neuromorphic hardware architectures) and on the other side, applied AI can be used for software modules in order to produce end-user-information by fusing multi-mode sensor data and measurements (cp. Section 3).

In principle, CSS are working in a similar manner like it is known for the human sensing and cognition process. Here, the brain is in real-time evaluating and combining the acquired data supplied by our human senses (sight, smell, touch, taste, hearing), potentially amplifying and

prioritizing the data source with the most relevant input or adding information from further sources, if needed. At the same time, our brain is the control center for adaptation and re-orientation (positioning) of our senses. Our memory helps to learn from events and comparable situations in the past.

This concept is imitated on a technological level by CSS. They use elaborated multi-mode physical sensor systems like they are used in nondestructive inspection technologies (e. g. acoustics/ultrasound, electromagnetics, thermography, X-ray etc.) for data acquisition. Applied AI is used for advanced data processing and for learning by applying the power of deep neural networks and other deep learning approaches. This is the brain of CSS, which is implemented on the hardware level (i. e. by microelectronics like neuromorphic hardware architectures) as well as on the software level. The material data space represents the complete data ecosystem with information about all stages in the lifecycle of a material (i. e. the complete history). It serves as the container for the product and material data and therefore represents the memory in this comparison. Data and information transfer is accomplished via defined interfaces and communication protocols that are becoming more and more standardized within the network of the internet of things. This is similar to our nerve system with its nervous pathways. The idea is illustrated in Fig. 2.

For this recently rapidly growing demand for performant and reliable CSS, specific requirements have to be fulfilled for validation and in order to guarantee their correct function. This is especially relevant as CSS are used in modern nondestructive inspection and evaluations systems, which give access or at least derive information about the materials state for safe and reliable operation of technical products. They are based on multi-mode physical measurements, i. e. a combination of inspection meth-

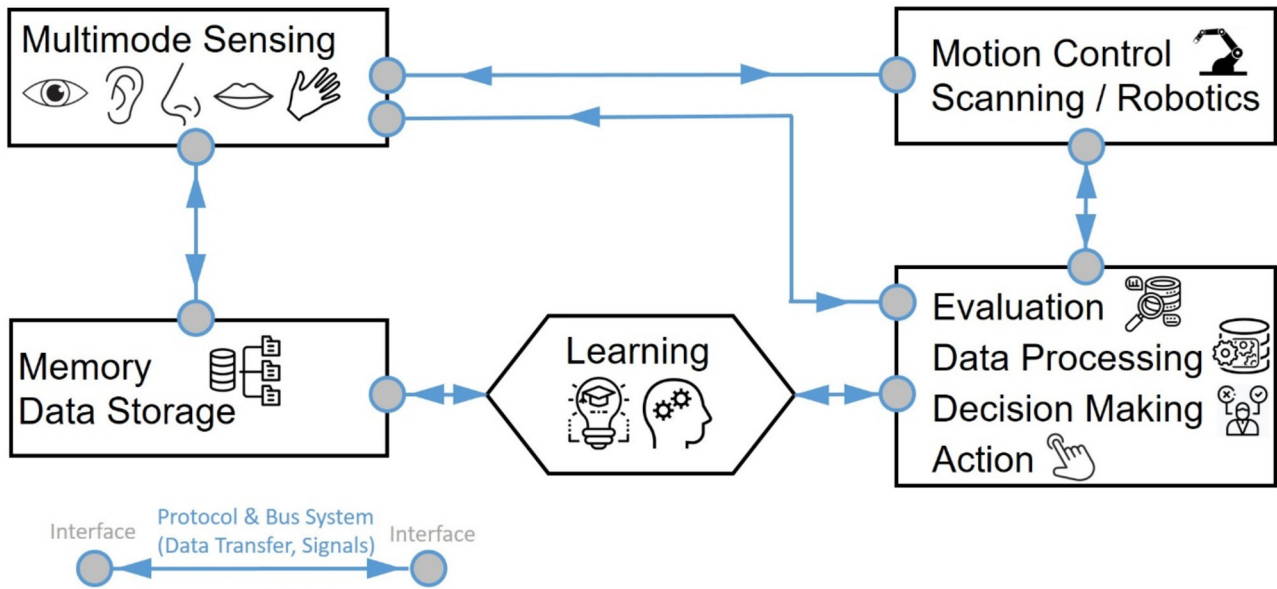


Figure 2: Concept of CSS in analogy to the human cognition with corresponding connectivity and interaction of basic modules and elements.

ods like acoustics, ultrasound, microwaves, electromagnetic and eddy current inspection systems, thermography etc. Furthermore, CSS help to predict the residual lifetime and they continuously monitor the integrity and the state of materials and products for safe operation by avoiding hazards and accidents for critical components, e. g. in the transport and traffic sector (railway, aviation, automotive), in the chemical apparatus industry, for pressure vessels, for power-plants or for the civil infrastructure. NDE 4.0 sensor systems are also quite common for the use in predictive maintenance and for structural health monitoring. For example, they are used for corrosion monitoring, reinforced concrete, bridges, planes, wings, renewable energy plants etc.

As CSS are using advanced algorithms for data and information generation, reliability and qualification of their specified function is one of the most relevant prerequisites for trusted applications and for the legal and psychological acceptance in industry [8]. As we discuss in this paper, this process can be based on validation procedures for classical NDE (Receiver Operation Characteristics – ROC, Probability Of Detection – POD, etc.) [67] which are combined with advanced procedures for specific validation methods within the concept of trusted AI. State-of-the-art and current activities in validation and qualification with a focus on NDE 4.0 are summarized in Section 5.

The concept for quality assurance of NDE 4.0 sensor and inspection systems has to cover all of the functional sub-systems, i. e. data acquisition, data processing and evaluation, data transfer etc. Approaches to these objec-

tives are presented in this paper after giving an overview on the crucial elements (modules) of CCS for NDE 4.0 applications in Sections 2–4. Some exemplary results for validated R&D technologies are shown in each section in order to illustrate the qualification strategy of CSS for NDE 4.0. Of course, for the integration into the IIoT, special aspects of safe and secure data handling have to be considered in addition. Reliable and safe microelectronics is a further issue in quality assurance for CSS (cp. Section 5.4).

This paper is organized as follows: Sections 2, 3 and 4 describe essential building blocks of CSS for NDE 4.0. In particular, we describe the NDE 4.0 data ecosystem with respect to interfaces and data formats (OPC UA, DICONDE+) in Section 2. Moreover, Section 3 introduces smart microelectronics and advanced signal acquisition concepts with a focus on neuromorphic hardware as well as compressed sensing concepts. Section 4 discusses AI for NDE 4.0 with a particular spotlight on explainability. Each of the Sections 2 to 4 present their own prototypical implementation examples. Section 5 then introduces standards and procedures for validation of entire CSS composed of the presented building blocks with a discussion on standardization activities, performance measures, as well as trusted microelectronics. Finally, Section 6 provides a summary and outlook of the paper. Note that, to improve the accessibility of the manuscript, the state of the art is discussed in the separate subsections related to the actual technical contributions, in particular Sections 2 to 4.

2 NDE 4.0 data ecosystem

2.1 The need for digitalization

Data ecosystems are an important pillar when we talk about the realization of NDE 4.0. While in other industries the digital transformation has already taken place, this step is still pending in the case of NDE. The reason for this is, among other things, the diversity of the NDE. Many procedures are automated, partially automated, or performed manually. In many cases, the inspector performs his evaluation visually and documents results in handwriting. The documentation and archiving of data often takes place without interaction with document management systems (e. g. Enterprise Resource Planning, ERP). The data is usually stored either on the inspection computer, the desktop PC of the person performing the inspection, or on a plant-internal server. Usually, the inspector is responsible for passing on both the inspection data and the inspection results.

Furthermore, the connection and data exchange to existing control and management systems or IIoT environments represents a special challenge for the NDE compared to other technologies of actuator and sensor technology. While, for example, industrial robots and vision systems transfer clearly defined logical or scalar information, there are various further requirements in the NDE like the documentation of results. Therefore, additional to the communication interfaces for the NDE devices, correlating processes must also be digitized and mapped. In addition, data must be shared with other systems in a controlled manner for processing and evaluation of results. This current industrial and global development offers the opportunity for nondestructive evaluation to implement the urgently needed further development with regard to digitalization, as required for use in IIoT networks. Potential NDE 4.0 software technologies for this are Open Platform Communications Unified Architecture (OPC UA) and Digital Imaging and Communications for Non-Destructive Evaluation (DICONDE) [90].

2.2 OPC UA interfaces

2.2.1 Key values of OPC UA

OPC UA is an open standard for cross-manufacturer and cross-platform communication across different levels and degrees of automation [52]. This standard is already used by various industrial companies and by individual associations such as the Mechanical Engineering Industry Association (VDMA) to implement the Industry 4.0 (I4.0)

aspects. Also, the communication between different machines of a production network is realized via OPC UA [87]. Within the framework of the committee work of the VDMA, manufacturer-independent information models for the areas of robotics and industrial image processing (machine vision) have already been developed and published [84, 85]. The specifications of such information models are called companion specification. With the OPC UA Companion Specification for Machinery [86] a specification also exists that can be used as a universal fundament for further companion specifications.

To understand the importance of OPC UA for I4.0 in general and NDE 4.0 in particular, it is necessary to take a brief look at the current state of the art for communication interfaces. Previous communication structures of the third industrial revolution are based on the model of the automation pyramid [45]. Depending on the communication level, different communication interfaces have been used to meet the requirements in terms of data volume and transmission speed. This resulted in the structure of the data and the choice of technology being manufacturer-dependent [26].

OPC UA builds on these existing technologies and extends them in a useful way. The technology can be used at all levels of the automation pyramid [9]. To realize deterministic communication requirements on the actuator/sensor level, the hardware standard Time-Sensitive Network is additionally required [10]. Furthermore, the structure of an OPC UA server is explicit, since its description is integrated on the server and does not have to be interpreted manually by the end user. This is also the case without the use of companion specifications, whereby their uniform structure allows additional facilitation with regard to the integration and exchange of components into existing network structures.

2.2.2 OPC UA for NDE 4.0

For the use of NDE methods in IIoT networks, OPC UA represents an essential building block for the above reasons. Methods can communicate directly with MES (Manufacturing Execution System) and ERP systems, receive inspection instructions and provide their results in real time. However, in order to be able to use this technology effectively, a NDE-specific companion specification is required that fulfills the general requirements for the inspection methods and takes into account the respective degree of automation of each method. This includes, for example, the inspection instructions and the results of the inspection report. Such a companion specification is currently

under development within the OPC UA working group of the ZFP4.0 expert group of the German Society for Non-destructive Testing (DGZfP). Independent of the development of this information model, OPC UA is already used for the control of inspection tasks. [38] describes a robotic inspection of adhesively bonded joints using active thermography and OPC UA as a communication interface. The implementation of a digital twin for quality assurance of welding processes was also realized using OPC UA [51]. OPC UA thus enables the digital provision of information, which was previously only available in printed form or as a PDF file, as well as the use of the results as relevant data for other IIoT approaches, such as AI aspects or the digital twin.

2.3 Data format DICONDE+

2.3.1 The standardized data format DICONDE

In order to archive and use relevant information in a generally usable way for years, uniform data formats are an essential requirement. One appropriate data format that meets the requirements of NDE 4.0 is DICONDE. The technology DICONDE is derived from Association for Computing Machinery (ACM) and National Electrical Manufacturers Association (NEMA). DICONDE describes both a generic data format and a communication interface. It has been known in medicine since 1985 and was introduced in 1993 as Digital Imaging and Communications in Medicine (DICOM). The medical variant has already been implemented for storage of images procedures by e. g. ultrasound or X-ray. There are a number of extensions for DICONDE, which are based on medicine and are adapted in various fields, extend them or add new possibilities. Instead of the patient, the focus is now on the specific test object with its properties. Standards specify which values must be filled in by the user and which are optional or conditionally optional for example, for eddy current, ultrasound and X-ray inspection [1, 3, 4]. In parallel with the linking of test results with component parts, relevant metadata are also stored in the DICONDE format. Beside UIDs (Unique Identifier) and date of creation, this also includes, amongst others, component information such as material properties or, with regard to inspection procedures, serial numbers of the systems and device settings. This makes an inspection process completely traceable, even for further users, and enables further evaluations and analyses.

The implementation of a widely used protocol results in various advantages. Inspection systems can be embedded in larger industrial plants and offer new possibili-

ties for cross-modal data evaluation and audit-proof storage via standardized paths. This reduces both integration costs and the susceptibility to errors, since all participants use a structured unique generic data format. Thus, DICONDE enables that future evaluations or analyses do not necessarily have to be performed with the software with which the data were acquired. Since evaluated data are stored in DICONDE, the recording software can directly calculate all necessary images (e. g. cross-sectional images or C-scans) and store them directly into the DICONDE archive.

2.3.2 DICONDE+

Under the keyword DICONDE+, research is currently taking place that addresses the integration of CSS in DICONDE, but also pursues the intelligent use of data sets for NDE 4.0 aspects, for example by integrating also the raw data sets into DICONDE+ in general. On the one hand, the clear structure of DICONDE means that data relating to components and inspection results can be clearly interpreted, while at the same time this structure must be strictly applied in order to generate DICONDE-compliant data sets. This makes it difficult, for example, to integrate inspection data obtained with unconventional multi-sensor systems or newly developed hybrid inspection methods [12, 82]. By creating suitable standards for such methods, new inspection and evaluation methods and thus data from innovative novel sensor systems or fused data sets could also be integrated into the data format in the future. Furthermore, this standard could generally be extended to include destructive testing methods.

With regard to the second aspect, DICONDE is also suited for intelligent data processing using approaches that go beyond those of curated data hosting. This allows DICONDE data sets to be used not only for archiving but also as an example in AI applications by means of a semantic search. To enable such a semantic search, the metadata of the DICONDE data must be mapped in a suited database using an adequate ontology. Such an implementation could be realized, for example, by a SPARQL server (SPARQL Protocol and RDF Query Language) [73]. In this way, complex search queries are possible, which allow the user to search specifically for data sets for his current application. An example of this would be the search for additional data sets as training data for the implementation, qualification and validation of an AI algorithm. This algorithm is more robust with additional data from a DICONDE archive and allows more extensive qualification and validation of the approaches developed. At the same time, new



Figure 3: NDE 4.0 demonstrator consisting of a dual robot unit (left: KUKA LBR iiwa, right: ABB YuMi). The B-pillar base is located on the worktable of the ABB YuMi. The three displays show raw data of a sensor (right), the common display of all current sensor results (middle) and the documentation in DICONDE (left).

data sets, which were only possible using the semantic search, can be added to the DICONDE. In this way, generic data formats are not only an effective means for archiving data, but also pave the way for the implementation, qualification and validation of AI applications and they allow new data-driven business models.

2.4 Applications and prototypical implementations

To show the interaction of these two technologies, a demonstrator was set up at Fraunhofer IZFP [41]. For this purpose, inspection methods are coupled with an OPC UA interface as well as with a DICONDE archive in order to exchange information with existing actuators, but also with each other and to archive the results in DICONDE format. The inspection is carried out at the base of a B-pillar of an automotive car body structure. In a first step, the metallic material is identified with an eddy current system. This information is passed on to the other sensor systems via OPC UA, which adapt their test parameters to the material. This is followed by a wall thickness measurement using ultrasound and a characterization of the material properties using the micromagnetic inspection method 3MA (Micromagnetic Multi-parameter Microstructure and Stress Anal-

ysis) [98]. All results are made available to the user on a single human machine interface, but are also added directly to the DICONDE file.

By using the technologies mentioned, additional sensor systems and actuators can be integrated into the demonstrator via OPC UA, but the report created in DICONDE can also be supplemented with additional test results and corresponding meta information. Figure 3 shows the setup of the demonstrator with the three sensor systems, the industrial robots used and the component to be inspected.

3 Smart microelectronics and advanced signal acquisition for NDE 4.0

3.1 Neuromorphic hardware

3.1.1 Computing in sensor systems

Cognition in terms of sensor systems can be understood as the ability of the system itself for learning and/or inferencing. In simplified terms, each sensor system consists of the individual components shown in Fig. 4. It con-

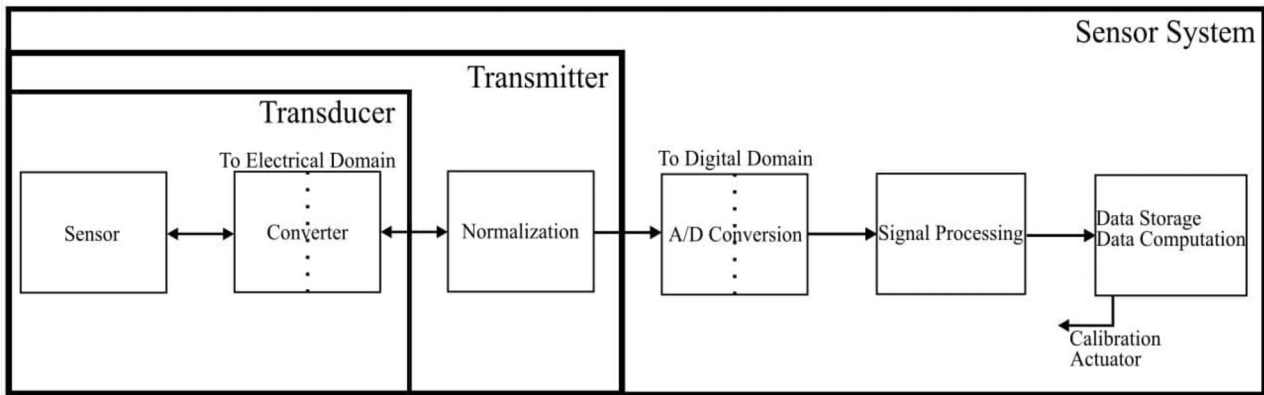


Figure 4: Simplified representation of the schematic structure of sensor systems consisting of sensor, analog-to-digital conversion and data processing unit.

sists of the sensor itself and a converter to the electrical domain. Both elements are referred to as the transducer. Together with a measurement amplification and/or normalization, the so-called transmitter is created. After the signal normalization, the conversion from the analog to the digital domain can take place [78]. Before analog-to-digital conversion, the sensor system does not yet have any cognitive capabilities. The discretization of the analog signal and digital signal processing improve the robustness of signal processing and transmission. In addition, microprocessors can be used to solve more complex tasks with the existing sensor signals. The data volume when processing sensor signals is mainly dependent on resolution, sampling rate as well as the number of signals to be digitized. With the increasing use of data-driven signal processing approaches such as machine-learning algorithms to build complex models for decision-making and classification, logically the amount of data is also increasing. With the growing demand for sensor systems that can understand complex processes and make conclusions based on incoming sensor data, as well as the increase in miniaturization and energy efficiency, it is necessary to develop new processor structures. A well-known limitation in processing large amounts of data from microprocessors is the so-called memory-wall. This problem arises from embedded systems up to high performance systems (e. g. servers, supercomputers) [80]. One approach is to follow a biologically inspired data processing, as it takes place in the brain of living beings, known as neuromorphic computing. Among this, especially the approach to process data within the memory is very promising. The creation of complex models or processing of data in memory is also referred as In-Memory Computing (IMC) [36]. It has been shown that this technique can improve the energy efficiency up to multiple hundreds of Tera Operations Per Sec-

ond per Watt (TOPS/W) [105] and up to 10000 TOPS/W [53]. New technologies such as resistive RAMs (Random-Access Memory) can be used for data processing, but also familiar memory technologies such as SRAM (Static RAM), DRAM (Dynamic RAM) and flash memory. Each of these technologies has its technical and economic advantages and disadvantages [36]. Not only the systematic structure and the way of data to be transported within the processor changes, but also the way of signal processing. Using IMC units as an example, analog data processing is implemented. This can lead to advantages in terms of energy efficiency and data throughput. On the other hand, the analog nature can have the disadvantage that the accuracy can be reduced considering classification algorithms.

3.1.2 Energy efficient sensor nodes

The selection of a suitable hardware platform for a specific application is an essential part in the development of sensor systems. Energy efficiency, miniaturization, cost, as well as usability are only a few parameters that can play a role. While an Application-Specific-Integrated-Circuit (ASIC) implementation is probably the most efficient solution, microcontroller as well as Field-Programmable-Gate-Array (FPGA) solutions show a much higher flexibility. Therefore, when designing a sensor system, it is important to know all parameters and requirements for the target application. In the case of applications with machine-learning algorithms on sensor nodes, the efficient use of available resources plays an increasingly important role, as also briefly explained in Section 3.1.1. The design of dedicated circuits for the execution of previously trained algorithms are an important part of achieving an optimal solution.

3.2 Advanced signal processing: Compressed sensing and sparse signal recovery

3.2.1 The role of advanced signal processing for NDE 4.0

There is unanimous agreement that realizing the vision of NDE 4.0 requires the deployment of large quantities of sensor nodes to deliver the required digital data [88]. In terms of the scalability, the development of low-cost IIoT sensor devices is particularly attractive, as it allows their disposition in large quantities [89]. If we manage to solve the challenges attached to making such systems interoperable (cp. Section 2), we can aggregate large streams of data about our processes and products. On the one hand, this allows sophisticated analyses and a robust training of data driven methods (cp. Section 4). On the other hand, collecting, maintaining, and accessing these data sets may turn into a severe challenge on its own [104]. It requires not only continuous significant investment in IT resources but also personnel (i. e. data scientists), able to retrieve the relevant information from the collected data. In many cases, this approach may be wasteful as an unfiltered collection of raw sensor data leads to a high degree of redundancy and it would be much more efficient to focus on the relevant data in the first place.

This is the promise of the Compressed Sensing (CS) paradigm [25]. In its core, it states that we can sample signals significantly below the Shannon-Nyquist rate without loss of relevant information, provided that prior knowledge on the signals is available [17]. This is very often the case in industrial applications. Hence, if we manage to encode this information into the sensors themselves, their recorded signals contain significantly less redundancy and deliver much more relevant data. CS has been successfully applied in a number of research fields, including fields like magnetic resonance imaging [50], radar [15], imaging [19], and many others. Recently, we have started to explore the intimate connections between CS and the field of NDE 4.0 [68]. There are at least two aspects of CS that can be applied beneficially: novel sub-Nyquist sampling architectures for reducing data redundancy on the one hand and an advanced algorithmic framework for data processing on the other hand. Since we do not aim at a comprehensive review of all NDE applications of CS in this section, let us instead focus on a few relevant examples.

As a first example, let us consider industrial X-ray Computed Tomography (CT). X-ray CT is used in a large and increasing number of industrial applications [13]. It delivers representations of the interior structure of a wide range of materials, allows identification and measurement

of geometrical features (including defects), and delivers material information through the X-ray absorption [96]. A bottleneck of 3D X-ray CT systems is their inspection speed, since the spatial Nyquist theorem requires a certain number of projections, each of which needs a certain minimum exposure time [62]. CS offers a potential solution since it admits accurate reconstructions from sub-sampled data, provided that prior knowledge about the objects is available. In industrial settings, such knowledge comes in multiple forms: manufactured pieces typically consist of a limited number of known materials, rendering the reconstructed images piece-wise constant. Also, in some cases (such as inline inspection), the geometry of the piece is known up to deviations due to manufacturing and positioning tolerances as well as the defects that we aim to find. Based on this prior knowledge, we can carry out 3D scans with significantly fewer projections without compromising image quality, provided that the algorithmic framework for the reconstruction is suitably adjusted. Adapting methods first proposed in a medical context [18, 76], we were able to show that in inline X-ray CT applications, a ten-fold reduction of the projections is possible without compromising the capabilities of finding even small defects [77]. As a second example, let us consider ultrasound. Ultrasonic signals possess high degrees of redundancy in many industrial settings for two reasons: firstly, their time signatures often consist of isolated echoes with (partially) known pulse shapes. Secondly, the wavefield typically does not change abruptly over space, rendering scans taking at adjacent points highly correlated. Both factors have been used to devise systems that incorporate temporal CS (e. g., through the modulated wideband converter architecture [14] or through (scrambled) Fourier sensing [49]), combined with spatial subsampling [63]. In [46] it is shown how spatial resolution can be obtained from a single transducer based on CS principles by placing pseudo-random masks into the acoustic path. For lamb waves, jittered subsampling is considered in [28]. Subsampling is usually combined with advanced reconstruction algorithms (which can also be applied on Nyquist sampled data). A prominent example is the use of sparse recovery algorithms such as FISTA (Fast Iterative Shrinkage-Thresholding Algorithm) [11] or STELA (Soft-Thresholding with Exact Line search Algorithm) [103], which yield superior image quality, even when applied to Nyquist-sampled data [48], but also facilitate reconstruction from CS measurements with very high compression rates [49]. As a final example, let us consider Terahertz imaging, i. e., the usage of electromagnetic waves in frequency ranges above 100 GHz. The interaction of such waves with matter permits scans with high sensitivity and high spatial resolution

in a number of specific applications, such as the inspection of polymers, thermoplastics, as well as certain composite materials, alongside the measurement of coating thickness, the evaluation of adhesions and welds, and others [27]. While hazard-free and flexible to apply, the main drawback of Terahertz scans is the fact that they are typically time-consuming due to the required raster scans [43]. This is an interesting application of CS where the combination of Fourier optics [32] and the principle of the single pixel camera [19] have enabled the conception of systems capable of spatially resolved Terahertz imaging without the necessity to carry out mechanical movements [16]. These examples show that CS-based acquisition devices can lead to quite unconventional designs.

Overall, CS allows significantly increased flexibility on the design of our sampling architectures, to take into account prior knowledge and thereby focus more on sampling relevant data directly. Such innovations can lead to reduced power consumption and data rate, which is particularly relevant for wireless sensors, but may also be a key enabler for high-rate applications (such as scaling up the number of channels in a Multi-Input-Multi-Output setup). At the same time, the algorithmic framework of the CS field can be beneficially applied for enhanced image reconstruction, giving rise to families of algorithms with adjustable complexity vs. performance trade-off.

3.2.2 Quality assessment of advanced reconstruction schemes

As CS introduces new concepts into the acquisition and processing of sensor signals, including concepts of randomness, it is not always easy to characterize the achievable quality deterministically. Yet, for NDE applications, a reliable quality assessment is of utmost importance. Hence, this section discusses potential performance metrics along with their advantages and drawbacks. There are two key ingredients to a successful reconstruction method: its achievable quality and the computational complexity to obtain this result. More often than not, these two are conflicting goals, and as argued above, the algorithmic framework of the CS field provides methods to control the trade-off flexibly. Hence, we need to evaluate both: the quality as well as the complexity. In this section, we want to discuss metrics for these two aspects, using CS in ultrasound as an example.

Regarding quality metrics for ultrasonic images, the first and most obvious choice is a subjective comparison of the image quality. As state of the art NDE is still often manual, practitioners are quite used to assessing unprocessed

and processed data (such as A-Scans or B-Scans) by visual inspection. It has been shown that parametric reconstruction algorithms can lead to better visual image quality by virtue of a more precise image of the defect geometry [48], even when combined with severe sub-Nyquist sampling through CS [49]. However, the evaluations of the visual image quality is subjective and may be misleading. As a more objective alternative, it is quite common to apply the Contrast to Noise Ratio (CNR) for imaging features in regions of interest, such as lesions in medical settings. However, for modern reconstruction algorithms, the CNR may not be appropriate since their amplitudes do not necessarily represent backscattering intensities anymore and the CNR is susceptible to amplitude variations such as a dynamic range compression that do not improve detectability, as argued in [65]. The paper also suggests a more stable metric given by the generalized CNR (gCNR), which is directly related to the probabilities of false/missed detections and always normalized to the interval $[0, 1]$. The gCNR has proven to be a useful tool when analyzing imperfections in reconstruction methods in NDE, e. g. due to positioning uncertainties in manual ultrasonic testing [44]. Likewise, to study the degradation of image quality due to changes in the measurement setup, such as sub-Nyquist sampling, image quality metrics such as the structural similarity index measure (SSIM) [70] have been applied.

Finally, a common approach in the field of NDE is to study the abilities of a given imaging approach to detect and accurately size defects [95]. The corresponding metric is the POD, which measures the chance of a given detector to find defects as a function of their size. While it is common to apply POD analysis to data acquired from reference specimen, it can also be used in a simulative manner by the following procedure: (1) using a forward model that can generate realistic ultrasound signals from defects with varying sizes; (2) emulating measurement effects (such as compressed sensing and additive noise) in software; (3) feeding the synthetically generated measurement data through the reconstruction framework; (4) extracting relevant features from the obtained images (such as a defect amplitude); (5) performing a regression between defect size and the chosen feature from a set of noisy realizations; (6) using the empirical statistics to compute a POD curve, e. g. described in [7]. While the absolute numbers obtained with this approach may not always be entirely realistic (since simulations can be too ideal), it provides a way to compare the influence of system parameters (such as compression rates and strategies) and the used algorithmic framework on the POD curve. An example of this procedure is shown in Fig. 5 where we compare the POD

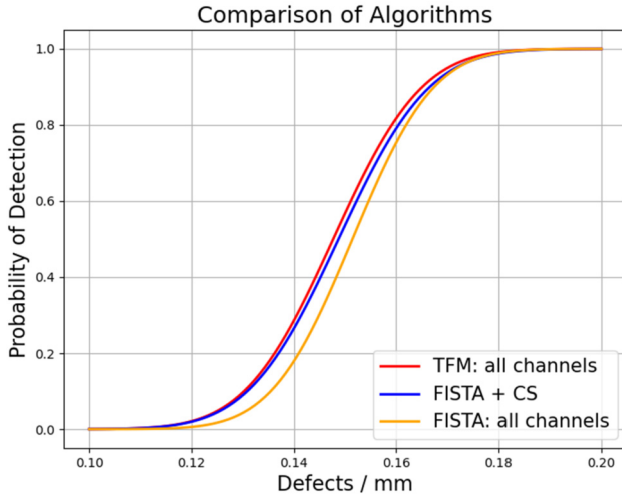


Figure 5: Sample POD result for a synthetic 16×16 FMC data set with a single defect of varying size, comparing classical reconstruction via the TFM (red curve) with FISTA using $5 \times$ temporal compression (orange curve) and $50 \times$ spatio-temporal compression ($5 \times$ temporal and $10 \times$ spatial, blue curve).

in a synthetic 16×16 Full Matrix Capture data set (FMC) with a single defect of varying size. We compare the Total Focusing Method (TFM, red curve) with parametric reconstructions through FISTA. For FISTA we employ temporal compression [49] (factor 5, orange curve) and spatio-temporal compression [63] ($5 \times$ temporal, $10 \times$ spatial, blue curve). Figure 5 shows that in the simulation FISTA always outperforms the TFM, even when significant data reduction (factor 50) is applied. These results should be verified based on a real specimen with varying defect sizes in future investigations.

Alongside the performance, we also need to quantify the complexity of given algorithms. While in the past, it was common to simply count the number required of floating point operations, this is no longer sufficient since computations are not necessarily linear any longer. With a variety of new computing paradigms, allowing for an increasing amount of parallelization and optimized processing on specific platforms such as GPUs, the complexity analysis is not an easy task anymore. Often, we resort to studying asymptotic complexities that give an indication of the expected scaling of the run time with a change in the input dimension. As an example, while algorithms such as Synthetic Aperture Focusing Technique (SAFT) possess a linear scaling of the complexity with respect to its input dimensions, iterative algorithms such as FISTA can scale quadratically, which may render them infeasible for larger input sizes. At the same time, as the available computing power keeps increasing due to Moore's law and the ad-

vent of quantum computing promises the availability of unprecedented computing resources [94], interest in more complex processing algorithms is now bigger than it ever was before.

3.3 Applications and first prototypical implementations

3.3.1 IZFP generic AI-accelerator

Currently, Fraunhofer IZFP is developing solutions for CSS that are more energy-efficient and more ecological friendly and at the same time scalable in terms of data volume, and data processing. The first approach here is an all-digital FPGA solution that is currently in development. The system is developed to process eddy current data and consists of a RISC-V (Reduced Instruction Set Computers – five) co-processor as well as dedicated AI accelerators. The key idea of AI accelerators is to take advantage of the nature of neural networks. These consist of several layers, where it is possible to execute every operation of the neurons of a single layer simultaneously. Since each FPGA has different resources, accelerators have been developed that are not only adaptable in bit width and thus adaptable in the accuracy of the algorithm to be executed, but also configurable in terms of parallelism of data processing and number of accelerator stages. With this approach, it is possible to easily adapt the hardware with software hardware co-design is facilitated. In addition, RISC-V processors allow the underlying instruction set architecture to be extended, which in future works should lead to even greater flexibility in processing data.

3.3.2 Assistance in NDE by augmented reality

In the context of NDE 4.0, an active area of research is the development of assistance systems in order to support the operating personal during inspection tasks. The main goals are the automated documentation, the visualization of intermediate inspection results, the automatic verification of the correct scanning and data acquisition process [22, 57]. In addition, they facilitate novel digital workflows for manual inspection tasks. Generally, all systems rely on real-time combination of measurement data with the tracked position of the probe [89]. An additional aspect is the improvement of the training of inspection personnel, e. g. through mixed reality concepts as discussed in [58]. The recorded data can also be integrated into the framework of digital building information modeling (BIM), as



Figure 6: Assistance system with NDE 4.0 features and Augmented Reality visualization (AR), image has been edited for illustration purposes [89].

described in [75]. A survey of the recent state of the art for these technologies can be found in [99].

To give a specific example, we briefly describe the features of the smart inspection systems from [89]. The presented laboratory prototype system acts as a complete platform to gain R&D experience for future NDE 4.0 sensor systems. Further aspects consider the impact of usability on human-machine interaction and the improvement on acquiring proficiency and the practical use of such technologies. A picture of this system is shown in Fig. 6.

The smart assistance system shown in Fig. 6 is suitable for various sensor modalities (e. g., ultrasound or eddy current testing). It implements a camera-based tracking of the probe and supports the inspector during the scanning procedure (scanned area, non-scanned area, signal validity, etc.). The inspection data and results are visualized via AR in real-time.

It combines several features relevant to future NDE 4.0 systems, including:

- advanced signal processing and progressive image reconstruction techniques (inspired by CS concepts, cp. Section 3.2) as well as AI-based advanced pattern recognition techniques (cp. Section 4) for generating high quality feedback in real-time,
- unified interfaces and data formats to facilitate the IIoT integration (cp. Section 2),
- autonomous signal control and self-monitoring to ensure the safety and the validity of the recorded data,

- augmented reality-based real-time visualization,
- remote control and cloud-based operation.

A current research item is the robust processing of manually acquired data, which has to cope with irregular sampling densities and imperfectly known sampling positions. Based on the combination of statistical interpolation techniques such as spatio-temporal Kriging [55] and data driven techniques [64], we can achieve progressive image reconstructions with a quality comparable to time-consuming automatic scans on a fine grid.

4 AI empowered NDE 4.0

4.1 AI for NDE 4.0

AI in its forms covering the image processing & Machine Learning (ML) and the modern data based Deep Learning (DL) methods have further potential to improve NDE methods [59, 60]. The DL methods that have already proven successful in many data processing tasks and their applications can no doubt provide the entry card of NDE into the industrial schematic I4.0. AI algorithms especially in their traditional ML-based form have a long history in NDE. They help to solve automatic recognition of patterns in NDE data such as automated defect recognition. The application of DL-approaches in NDE emerged in the recent

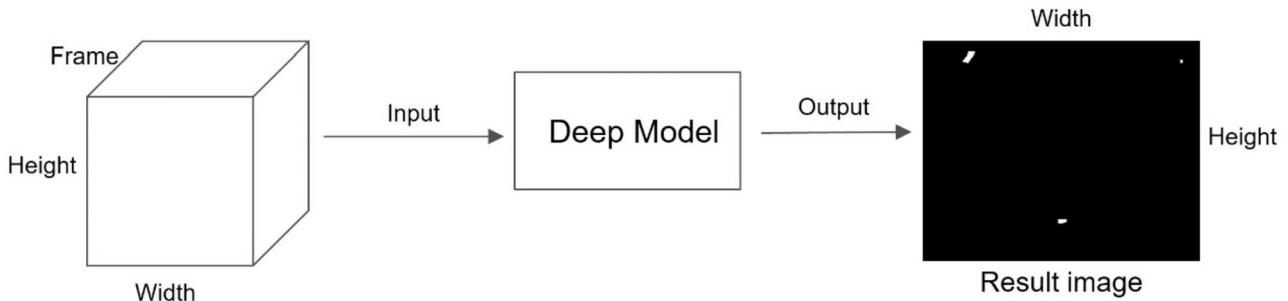


Figure 7: U-net defect prediction from [100].

years and is attracting more and more attention. In the first line, AI is an assistance technology to the human experts in their cognitive conduct when interpreting NDE signals and images. This definition is crucial for easing the acceptance of the AI-methods in the NDE scene, as people are afraid of the enormous modeling capabilities of DL methods in data interpretation and mapping into discrete or continuous domains. Thus, for a trusted, accepted and stable deployment of the AI-based NDE systems in the NDE community and by the NDE users, the outcomes of the AI models must be explainable and understandable for the humans.

4.2 Explainable AI for NDE 4.0

Explainable AI (XAI) also called trusted AI develops a set of mechanisms and approaches that can produce high-quality explainable, intuitive, human-understandable explanations of the outputs of the AI models. Even though the deep models have achieved high performance in image processing tasks [39, 47], the end users of those models are always concerned about their lack of transparency and explainability. Therefore, several approaches have been proposed in recent years to address this issue. There are mainly two popular categories of XAI approaches: feature attribution approaches [6, 20, 72] and activation maximization approaches [81, 102]. For the feature attribution approaches, authors in [6] have introduced a layer-wise relevance propagation method, where the output activation is propagated throughout the whole neural network down to the input data, yielding a heat map in the input data space. For the output maximization approaches, authors in [81] have proposed an image-specific class saliency visualization method, where for each class a computed image is generated to maximize its corresponding output class activation. This method provides an insight of what kind of features the trained model has really learned.

For the NDE area, besides of the XAI of the AI methods, there is one more challenge for the explainability of the NDE data itself. Unlike optical images, the NDE data are often hard to be interpreted directly. Such as the ultrasound data and thermographic data, different kinds of transformations must be applied to facilitate a better contrast for data visualization. Afterwards, the data annotation is based on this enhanced visualization. Moreover, for a supervised machine learning method, the quality of data annotation is crucial. As a result, the explainability of the NDE data should also be paid more attention. For the enhancement of thermographic data, there are some conventional methods, such as principal component analysis (PCA) [66], pulsed phase thermography (PPT) [56]. There are also some methods using neural networks, such as the method proposed by [101], where an autoencoder is used for thermal profiles.

4.3 Application and prototypical implementation

The first application we would like to notify about is to explain the NDE data. In [97], we have used stacked denoising autoencoder to preprocess the thermal profiles before PCA is applied, which is able to enhance the contrast of certain defects in specimens of carbon fiber reinforced plastics without weakening the contrast of the other relevant defects. This helps the expert to better analyze the NDE data.

The second application of the XAI in NDE 4.0 aims to explain a deep model for thermographic data segmentation. The model is proposed in our previous work, where three ML methods are experimented for defect detection on the thermographic data for solid oxide fuel cells [100] (cp. Fig. 7). Among the three methods, U-net is able to outperform the other two conventional ML methods (random forest method and support vector machine method). The

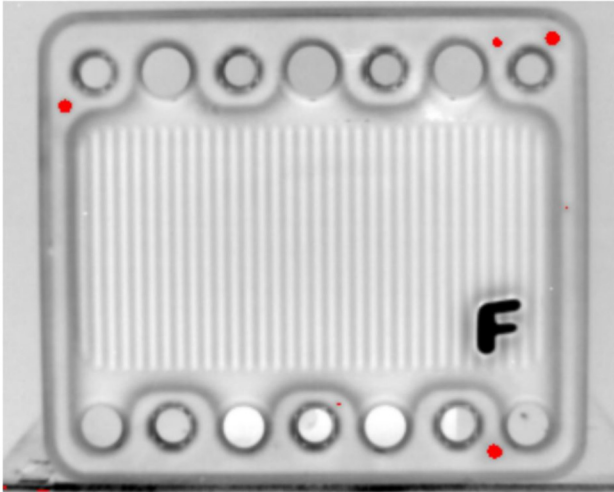


Figure 8: Visualization of one result image on the thermographic data. The red dots indicate the predicted defect regions, and the background image is an image slice from results of PPT applied on the raw thermographic data.

size of the input data is $297 \times 256 \times 320$. The model then produces a 256×320 output (cp. Fig. 8).

To explain the prediction of the model's decision, one pixel from the predicted region is selected and the feature attribution method is applied [81]. Afterwards, a $297 \times 256 \times 320$ heat map is generated. The heat map represents the contribution of each input data part to that one-pixel prediction. In Fig. 9, the location of the predicted pixel and one slice of the calculated heat map is shown.

Comparing to the rest of the infrared sequence, the first 50 frames have contributed more, see the bottom left image in Fig. 9. Besides, in the top right image in Fig. 9, it is obvious that the region around the predicted pixel has contributed more to the prediction pixel than the other regions, which is quite similar to human perception. With such explainability, the trust and acceptance of the AI in NDE can be boosted.

5 Standards and procedures for validation and for safe operation of NDE 4.0

5.1 Roadmap for standardization of AI and the German expert group NDE 4.0

In recent years, the German expert group for AI was founded, which is a joined initiative of the main organizations in this field with a close link to German authorities

(ministries for research and economics BMBF and BMWI, DIN e. V., the German commission for electrical and information science and engineering DKE and in addition the national R&D top level experts for AI) The objectives of this expert group are focused on supporting activities for innovation and for the practical transfer of applied AI technologies into the industry by procedures for standardization and by establishing corresponding certification processes in the future on a national and international level [23]. That is why the German DIN e. V. and similar European initiatives are playing a major role in this context. Two aspects within the whitepaper for the standardization roadmap are most relevant. The first aspect deals with the demand for smart standards and written procedures which state how to systematically explore and exploit existing, classical standards by application of AI tools. They are applied to databases and to existing documents for the variety of already accepted standards defining the state of the art in nearly all segments of industry, i. e. application domains for upcoming AI empowered technologies, cp. Section 5.2 in [24]. The second most important aspect for qualifications standards concerning NDE 4.0 technologies refers to general guidelines and existing standards that should be taken into account as a starting point for the development of specific validation and certification procedures, cp. Section 6 in [24].

At the same time, the DGZfP established the expert group for NDE 4.0, which is focused on the specific roadmaps for future NDE systems with a strong impact of applied AI and of requirements for integration into the IIoT, [21]. An overview on the goals and activities of this group is presented in [83]. Altogether, both initiatives define the overall framework and supply guidelines for qualification (i. e. validation and certification according to standards) which are exactly matching the specific requirements for transferring NDE 4.0 into industrial application. Furthermore, the corresponding regulations and standards are mandatory when dealing with and when validating first prototypical applications, which should be accepted by the industry (legal conformity) [8]. The following sections present some more detailed aspects on validation and qualification for NDE 4.0 technologies.

5.2 Guidelines and standards for NDE/NDE 4.0

NDE systems must be qualified to prove that they are suitable for the inspection task before they can be used in an operational environment. For routine inspections described in national or international standards, this qual-

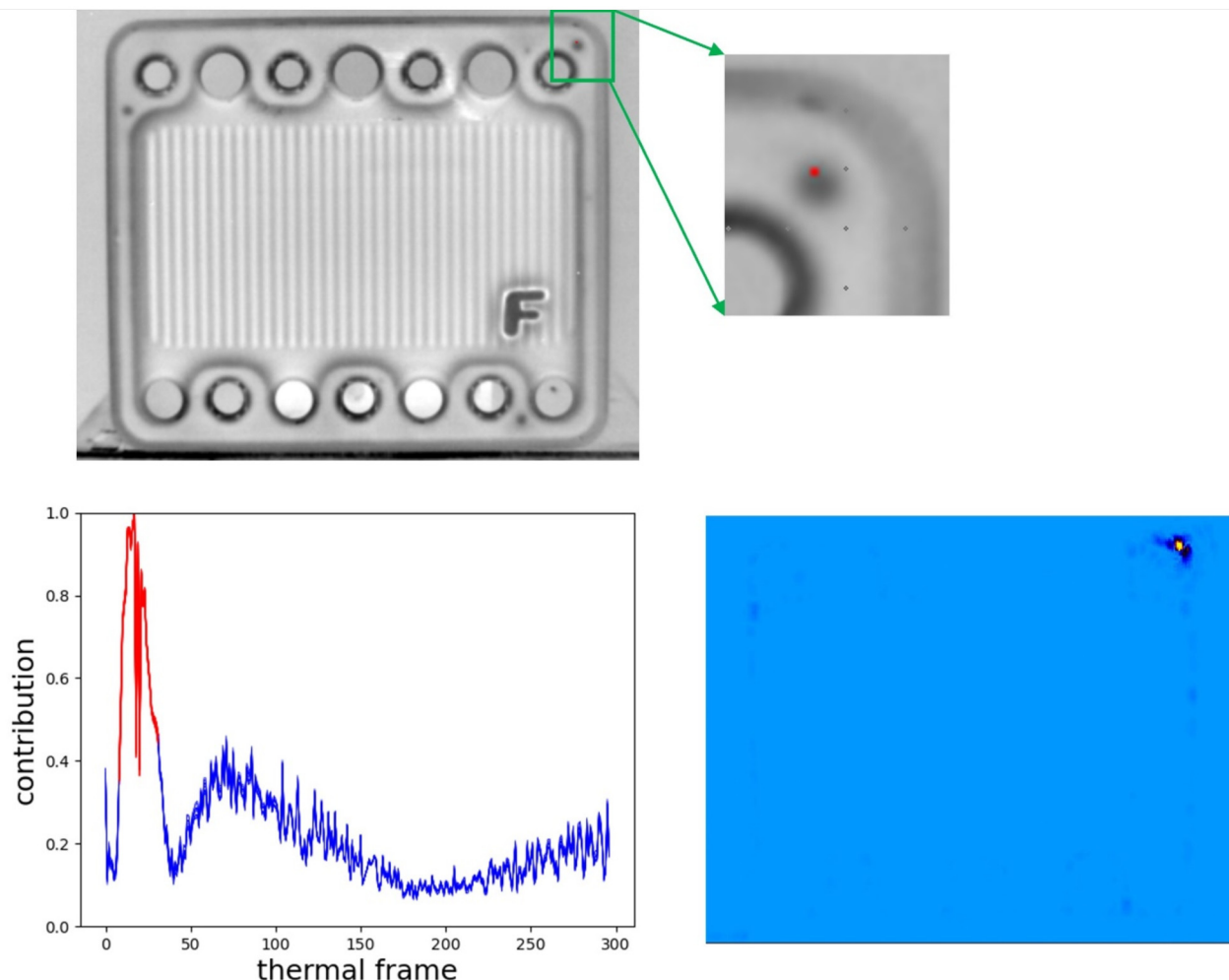


Figure 9: The top image indicates the prediction pixel's location, the bottom left image shows the contribution along the infrared sequence, the bottom right image is a slice (256×320) of the 3D heat map.

ification can be ignored. Qualification must be considered when a novel NDE technique is introduced or when the safety or economic consequences of a potentially poor NDE performance and/or the difficulties in applying the NDE technique are very high. In these cases, it is necessary to provide additional validation that the NDE technique can fulfil the requirements. The adoption of novel NDE technology particularly affects NDE 4.0 systems, especially when ML analysis and classification algorithms are used, as there is currently no rulebook or standard, which describe the entire qualification process in the field of NDE for such applications. At this point, potential users of NDE 4.0 systems are challenged with how ML applications can be integrated into the qualification process and how their reliability can be verified.

The European Network for Inspection & Qualification (ENIQ) provides a methodology for the qualification of

nondestructive testing procedures in the field of nuclear power plants [30]. The method described can also be transferred to applications beyond nuclear technology. Qualification as described by ENIQ includes a combination of technical justification (TJ), as well as inspections on specimens with real and artificial defects. The TJ includes the compilation of all evidence for the suitability of the inspection method (results of proficiency tests, feedback from field experience, applicable and validated theoretical models, and physical considerations). An overview of the major steps of a test system qualification is shown in Table 1.

5.2.1 Qualification of ML-based NDE systems

The performance of an NDE system is proven within the framework of the ENIQ qualification methodology by

Table 1: Major steps to be followed prior and during an inspection qualification [30].

Step no.	Task
1	Make available all required input information concerning component, defects, inspection and qualification objectives
2	Optimize inspection procedure using typical reference/training test pieces
3	Prepare inspection procedure and TJ
4	Assess submitted inspection procedure and TJ
5	Propose qualification plan including open and blind test piece trials, as required
6	Accept/refuse qualification plan
7	Conduct open trials for inspection procedure/equipment, if required
8	Issue/refuse qualification certificate for procedure/equipment
9	Conduct complementary qualification of personnel using qualified inspection procedure/equipment
10	Issue/refuse qualification certificate for personnel
11	Compile and finalise qualification dossier
12	Acceptance of qualified inspection by plant operator (and/or regulator)

means of a technical justification and practical trials on suitable test specimens. For systems that use ML-based methods for data analysis, the proof of performance can be performed in the same manner. In the following aspects, the ML components of an NDE 4.0 system must be included in the qualification process or require new activities that do not have to be considered for conventional NDE systems [31].

- When designing specimens with known and unknown defects for procedure validation within the TJ and practical trials, the design of training data sets and test data sets must be considered for ML components.
- Selection and training of ML algorithms
- Generation of qualification data sets for ML
- Consideration of ML during procedure validation
- Qualification of the inspection procedure and the ML algorithms (cp. Section 5.4)

5.2.2 Requirements for technical justification when using ML concepts

The contents and structure of a technical justification are documented in the guide Strategy and Recommended Contents for Technical Justifications [29]. As part of the approach described in this document, ML components must be considered in various sections of the TJ [31]. A summary of the ML system, including the justification for its

use of the selected algorithms and the expected results, should be provided. Furthermore, it should be presented how the system will be put into operation and how version traceability can be ensured. Finally, a summary should be provided detailing the capabilities of the ML system and indicating any recommendations for improvement if the full specification could not be achieved. The explanation of influencing factors as well as the experimental proof of the performance in terms of reliability of the ML components as well as the overall system will be discussed in more detail here. When considering influencing factors, it is assumed that the performance of ML algorithms is based more on experimental results than on clearly comprehensible physical relationships. At this point, transparency about the algorithms used plays an essential role (cp. Section 4.2). The decision to select the algorithms used must be presented as part of the qualification process. The selected ML algorithms should be documented in sufficient detail to assess the potential for error and the required justification. It is expected that the quality of the data sets used for training as well as for validation of the ML algorithms will have a significant impact on the performance of the overall system. Therefore, the data sets used should cover the range of defects necessary to achieve the test objective, as well as contain representations that lead to possible false calls. In addition, it must be demonstrated that the data sets do not contain bias that negatively influences the test results. Another influencing factor is the distribution of the data sets used for the validation and qualification of the ML algorithms. Here, care must be taken to ensure a clear separation of training data sets that are used directly for training the ML model. Test data sets used for monitoring and evaluation of the training and qualification data sets used for validation of the final performance of the model must be clearly separated. Another important point is the validation of the performance of the ML components and the overall system by determining the reliability. Known statistical performance indicators can be used here. It is important to select the statistical methods so that they are suitable for the NDE system to be qualified [31, 42]. Here a subdivision can be made. On the one hand, into procedures for the evaluation of faults such as the probability of detection (POD) as a function of fault characteristics or the receiver operating characteristic (ROC) described in more detail in Section 5.3 [5, 33] and into procedures which determine the suitability of a procedure for the determination of fault characteristics with the aid of a modular model where, in addition to the intrinsic capabilities, application parameters, human factors and the organizational context are taken into account [54, 67]. Another

Table 2: Ground truth and results of three exemplary classifiers.

Specimen	1	2	3	4	5	6	7	8	9	10
Ground truth	OK	OK	OK	OK	OK	OK	OK	OK	Not OK	Not OK
Classifier 1	OK	OK	OK	OK	OK	Not OK	Not OK	Not OK	Not OK	Not OK
Classifier 2	OK	OK	OK	OK	OK	OK	OK	Not OK	Not OK	Not OK

issue when considering the reliability of an NDE 4.0 system is the model variability. Estimation of this variability can be done either by using probabilistic ML models or by validation of different ML models for a given training data set (e. g. by a cross-validation procedure) [31]. During the qualification process of the ML components, care should be taken not only to consider the target variable such as defect size/defect class, but also to consider the false calls that occur. The false call rates can be used to draw conclusions about performance at an early stage during the development and qualification process. Possible overfitting, i. e. the ML model delivers good results for training and validation data used during development, but fails when generalized to unknown defects, must be excluded by testing on unknown data sets.

5.2.3 Qualification in practical trials

For traditional NDE systems, practical trials are performed as part of the qualification process. These requirements are described in detail in [30]. Essentially, the tests are divided into open tests and blind tests. The open trials are usually used to validate the inspection technique and the process procedures. In the blind tests, the performance of the operators in interaction with the inspection technique is considered.

For NDE 4.0 systems that perform data evaluation based on ML components, the operator's tasks change. The operator may only be involved in the data evaluation in a supervisory role. However, the shift of data evaluation to the test system now also requires blind tests on unknown real components or specimens for the ML components. This means that separating the evidence of process performance and operator performance becomes less important. A recommendation for the procedure is described in [31].

5.2.4 Essentials for guidelines and standards of NDE 4.0

ML is important for NDE 4.0 systems. They offer the potential to improve the performance as well as the reliability

of currently used inspection techniques. NDE 4.0 systems allow to reduce human factors that lead to high variation in inspection results. In contrast, however, new challenges arise that must be taken into account in qualification processes for such systems. These include the very reduced focus and experience compared to humans with regard to the influencing factors to be considered during an inspection. Due to this, as well as the low level of experience and the resulting lack of trust, ML systems are expected to be highly repeatable and consistent in their application. This means that qualification processes such as those described in [30] must be extended for NDE 4.0 systems. This includes e. g. the development and maintenance of the corresponding qualification specimen and data sets in order to be able to prove reliability in a resilient way.

The ENIQ document [31] provides a clear recommendation for the extension of existing qualification processes in the field of nuclear facilities. The recommendations set out in this document can also be transferred to applications beyond nuclear technology and can be used as a basis for a generalized procedure.

5.3 Performance measures for AI classifiers in NDE 4.0

Selecting the evaluation metric is the key ingredient of measuring the effectiveness of a classifier. Corresponding procedures are described in [59]. Let us consider a binary classification problem, where the system predicts whether a specimen is diagnosed a defective or not. The performance of the three exemplary models will be evaluated in the given test set, where only 2 out of 10 specimens are in reality defective (ground truth). The predictions for each exemplary fictitious model are shown in Table 2.

A specimen which is defective (positive) and is classified as defective is referred to as True Positive (TP). A specimen which is not defective (negative) and is classified as not defective is referred to as True Negative (TN). A specimen which is defective (positive) and classified as not defective is referred to as False Negative (FN). A specimen which is not defective (negative) and classified as defective is referred to as False Positive (FP).

Table 3: Scheme of confusion matrix.

Confusion matrix		Ground Truth	
		Positive	Negative
Predicted	Positive	TP	FP
	Negative	FN	TN

The accuracy of a classifier, i. e. model, is defined as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

For the above-mentioned example, the accuracy of classifier 1 is 0.7, the accuracy of classifier 2 is 0.9 and of classifier 3 is 0.8. Thus after evaluating the accuracy of each model, the best performance is achieved by model 2.

Usually, in the case of imbalanced data sets, accuracy is no longer the proper metric because the classifier tends to be more biased towards the majority class. Another problem with accuracy is that it does not give much detail about the performance of the classifier (i. e. it does not directly tell how many specimens are classified as defective while being intact and vice-versa).

A confusion matrix gives more insights not only about the performance of the classifier but also about the type of errors that are being made. As the name suggests the confusion matrix outputs the performance of the classifier in the form of a table. The columns correspond to the ground truth and the rows to the model predictions [59].

The confusion matrix is used to derive other evaluation metrics such as sensitivity (recall), specificity, precision, F1-score as well as the ROC curve.

Sensitivity (recall) and specificity are two metrics used to assess the class-specific performance of a classifier.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

Sensitivity is the proportion of the positive instances that are correctly classified, while specificity is the proportion of the negative instances that are correctly classified. In our example, the sensitivity shows us how good our model is in correctly predicting all specimens which are actually defective, while specificity shows how good the model is in correctly predicting all non-defect specimens.

Precision is another metric that describes which proportion of predicted positives is truly positive. Out of all specimens that the model diagnosed with defects, how many of them actually are defective.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

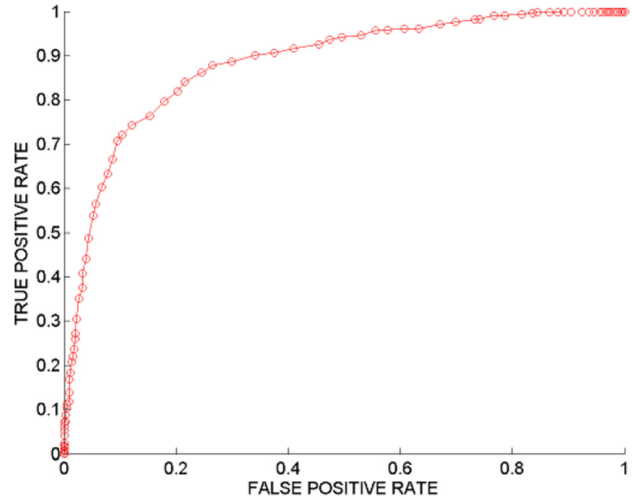


Figure 10: Example of a ROC curve. Each point refers to a specific decision threshold above which all samples are considered as defective [59, 61].

A metric that combines both precision and recall into a single formula is F1-score. It can be interpreted as a harmonic mean between recall and precision values, where an F1-score reaches its best value at 1 and the worst value at 0. A good F1-score means that we have a low number of False Positives and False Negatives.

$$F_1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

ROC curve is a graphic representation of false-positive rate (1- specificity) versus true-positive rate (sensitivity) for all possible decision thresholds (cp. Fig. 10). The overall performance of the classifier summarized over all possible thresholds is given by the Area Under the ROC Curve (AUC). Since ROC curves consider all possible thresholds, they are quite helpful in comparing different classifiers. The choice of the threshold is task-related. For instance, in a cancer prognosis system, the threshold is preferably chosen to be smaller than 0.5. An ideal classifier will have an AUC of 1, which means the ROC curve will touch the top left corner. In addition to being a generally useful performance graphing method, they have properties that make them particularly useful for domains with skewed class distribution and irregular classification error costs. These characteristics have become increasingly important as research continues into the areas of cost sensitive learning and learning in the presence of unbalanced classes.

Moreover, regarding the accuracy of an object detection algorithm, the primary metrics used to evaluate the performance of the system are Intersection over Union (IoU) and mean Average Precision (mAP).

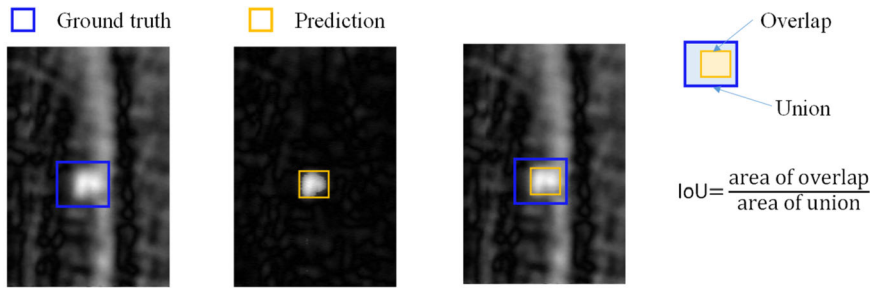


Figure 11: Example of Intersection over Union for the defect detection in a metal sheet using eddy current technique.

IoU can be used in any object detection algorithm that predicts bounding boxes as output and has ground-truth bounding boxes. IoU is a ratio of the overlapped area between the predicted bounding box and the ground-truth bounding box to the common area enclosed by both the predicted bounding box and the ground-truth bounding box (cp. Fig. 11). The higher the IoU, the better the fit.

To better visualize the performance of an object detection or a classification algorithm, a precision-recall curve, can be used for each threshold value, similarly to ROC curve. If we want to represent the performance of the detector using only one metric, the area under this curve is similarly calculated. This metric is referred to as Average Precision (AP).

The mAP is actually the average of all APs. We compute the AP for each class and image and average them, giving us a single metric.

5.4 Trusted microelectronics

As indicated in the introduction in Section 1, there is another very important aspect that must be considered for future applications on the industry level for NDE 4.0. This is about the trustworthiness of microelectronics hardware and how to guarantee that no technical hidden fraud may happen to the hardware architectures. We see two very promising approaches for trusted and secure microelectronics.

5.4.1 Global context and the necessity for trusted microelectronics

The worldwide drive for digitalization in civil and industrial environments, as well as the desire for a networked world, are causing a global increase in the demand for electronic components. Solutions for current trends such as smart mobility/city as well as IIoT and 5G are technology drivers here. However, the increased use of micro-

electronic components and their data traffic to cloud systems is significantly increasing global energy consumption. To counteract this development, an increasing shift of decision-making from the cloud level towards the end device is being strived for. In this way, decisions can be made directly at the point of action and energy-intensive data traffic can be saved [37]. The consequence is the use of high-performance microelectronic circuits in the most diverse areas of daily life and opens up a wide range of possibilities for malicious attacks. In principle, malicious attacks can pursue different goals such as spying on personal data or company secrets, manipulating or deleting files or sabotaging production processes and security-relevant infrastructure. In addition to the software technologies that have been established for many years, such as computer viruses, Trojan horses, or computer worms, the use of so-called hardware Trojans is becoming increasingly relevant.

A hardware Trojan is the manipulation of hardware or firmware of integrated circuits and electronic modules. The manipulation can take place at different points of the value chain and remain unnoticed for a long time [35]. In addition to being considered in the development phase of microelectronic circuits, hardware Trojans can also be implemented unnoticed within chip or module production. Another point of attack is the transport of electronic components. FPGA devices offer special attack possibilities here by manipulating the programming process [74]. The potential risk of attack is further exacerbated by the globally interwoven supply and value chains. While products from the consumer sector, with the exception of communication devices such as mobile phones, routers, etc. [79], are less in focus, components in security-relevant areas such as power plants, internet nodes or communication systems in intelligence agencies must be reliable and trustworthy.

Sensor systems in NDE 4.0 often provide safety-relevant data, are used to optimize processes or determine

the quality of products. Ensuring the trustworthiness of NDE sensors is therefore a particular challenge.

5.4.2 Technological solutions for security of microelectronics in NDE 4.0

Security on low-level hardware is crucial as a solution for NDE 4.0 security. This implies among others, a transparency and openness of the underlying processor structure. RISC-V as an open-source instruction set architecture is believed to offer a level of security that is not yet available with other solutions. There are groups working at security extensions considering cryptographic extensions as well as trusted execution environment. Because of the open-source nature, everyone can get involved to fill security vulnerabilities.

In addition to the described technologies for a precautionary defense against tampering, microelectronic circuits and assemblies can also be examined for the presence of hardware Trojans. For the classification of hardware Trojans, the characterization of the physical implementation, the type of activation as well as the action during the active phase can be used [71]. Different destructive and nondestructive methods can be used for detection. The detection of manipulated circuit parts on chip level can be done with the help of microscopic methods (e. g. scanning optical microscopy, scanning electron microscopy, and others) after opening the component package [71]. With the scanning acoustic microscopy method, chip structures can be imaged without package manipulation [34]. Functional tests, e. g. applying defined test patterns to the input pins, can indicate manufacturing faults or hardware Trojans by comparing the output signals. Side channel analysis is able to analyze different signals caused by electrical activity in electronic circuits. Side-channel signals include electromagnetic radiation, timing information, thermal radiation, and power consumption [71]. The use of imaging thermographic cameras is considered to be promising [69]. To increase the detection probability of different classes of hardware Trojans, the combination of several detection methods can be used.

6 Summary and outlook

In this publication we discuss the aspects of technology, AI embedding, validation and qualification of CSS for NDE 4.0. After an introduction to the notion of CSS and their relation to NDE 4.0, we discuss three major building blocks

of CSS: the data ecosystem, advanced microelectronics, and AI are examined in separate sections, respectively.

Regarding the data ecosystem, Section 2 emphasizes the need to digitize NDE data for integration into IIoT networks. OPC UA in combination with a uniform companion specification is introduced as a possible communication interface. DICONDE is presented as an example for a universal and manufacturer-independent data format. For its adoption in the NDE 4.0 world, further developments should be implemented. These are being pursued in the form of DICONDE+, which will enable amongst others a closed data loop.

In terms of the advanced microelectronics, a basic concept for the development of highly energy-efficient sensor systems with all its necessary components is illustrated in Section 3. New types of computing units, capable of directly classifying the sensor data, are able to reduce the requirements on downstream microcontrollers on the one hand and the maximum energy consumption on the other hand. The section also discusses the role of advanced signal processing for acquiring data more efficiently and provides quality assessment methodologies for advanced reconstruction schemes. We conclude with concrete examples for NDE 4.0 sensor systems, including a prototype FPGA-based sensor system implementation as well as examples for assistance systems for manual NDE.

In Section 4, AI methods (in particular, trusted AI methods) and their applications in NDE 4.0 for automated data interpretation are discussed. Here, two examples are presented, one for NDE 4.0 data explanation and another for DL model decision's explanation.

After presenting the essential building blocks, the final Section 5 examines aspects of the entire CSS. Firstly, the standardization, qualification and technical justification of AI-based sensor systems for NDE 4.0 are shown. Secondly, different performance measures for AI-based systems are presented, such as accuracy, sensitivity, precision and F1-score. Finally, trusted microelectronics, as a very important aspect for future applications on the industry level for NDE 4.0, is discussed.

Altogether, in this paper we present an overview on different fields and on the current state of strategies and procedures for qualification and validation of advanced NDE systems, which are closely linked to the upcoming world of NDE 4.0. Completely new philosophies have to be considered in order to assure a safe and trusted technology. Multidisciplinary aspects arise from the fusion and combination of approaches for AI especially trusted AI empowered CSS. Special demands for digital embedding into NDE 4.0 data ecosystems and qualification requirements

for future microelectronic hardware architectures must be taken into account.

CSS and their application in the field of NDE 4.0 have an enormous potential to introduce disruptive technologies and they allow to raise the degree of information about materials. In the near future, completely new scenarios for circular economy of materials and for digital life-cycle optimization may be realized. Legalization and approval of these brand-new technologies according to certification guidelines are prerequisites for this next generation sensor systems and for being accepted in industry. In addition, the autonomous operation requires acceptance by the responsible inspection personnel from a psychological point of view.

References

1. ASTM E2339-15: Standard Practice for Digital Imaging and Communication in Nondestructive Evaluation (DICONDE), 2015.
2. ASTM E1820-15a: Standard Test Method for Measurement of Fracture Toughness, ASTM International, West Conshohocken, PA, 2015, <http://dx.doi.org/10.1520/E1820-15A>.
3. ASTM E2663-14: Standard Practice for Digital Imaging and Communication in Nondestructive Evaluation (DICONDE) for Ultrasonic Test Methods, 2018.
4. ASTM E2767-13: Standard Practice for Digital Imaging and Communication in Nondestructive Evaluation (DICONDE) for X-ray Computed Tomography (CT) Test Methods, 2018.
5. ASTM E2862-18: Standard Practice for Probability of Detection Analysis for Hit/Miss Data, American Society for Testing of Materials (ASTM), 2018.
6. S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, W. Samek, On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation, in *PloS one* 10, e0130140, 2015.
7. M. Baudin, A. Dutfoy, B. Iooss, A. L. Popelin, OpenTURNS: An Industrial Software for Uncertainty Quantification in Simulation, in Ghanem R., Higdon D., Owhadi H. (eds.), *Handbook of Uncertainty Quantification*, Springer, Cham, 2015, https://doi.org/10.1007/978-3-319-11259-6_64-1.
8. M. Bertovich, S. Feistkorn, D. Kanzler, B. Valeske, J. Vrana, ZfP 4.0 aus der Sicht der ZfP-Community: Umfrageergebnisse, Herausforderungen und Perspektiven, in *ZfP-Zeitung* 174, pp. 43–49, 2021. ISSN: 0948-5112.
9. A. Bildstein, J. Seidelmann, Migration Zur Industrie-4.0-Fertigung, in *Handbuch Industrie 4.0 vol. 1*, pp. 227–242. Springer, 2016.
10. D. Bruckner, M. P. Stanica, R. Blair, S. Schriegel, S. Kehrer, M. G. Seewald, T. Sauter, An Introduction to OPC UA TSN for Industrial Communication Systems, in *Proceedings of the IEEE* 107(6), pp. 1121–1131, 2019.
11. A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, in *SIAM Journal on Imaging Sciences* 2(1), pp. 183–202, 2009.
12. M. Borsutzki, R. G. Thiessen, I. Altpeter, G. Dobmann, G. Hübschen, R. Tschuncky, R. Szielasko, Nondestructive characterisation of damage evolution in advanced high strength steels, in *ECF 2010, 18th European Conference on Fracture*. CD-ROM: *Fracture of Materials and Fractures from Micro to Macro Scale*, 30.08.–03.09.2010, Dresden, 9 pp., 2010.
13. L. De Chiffre, S. Carmignato, J. P. Kruth, R. Schmitt, A. Weckenmann, Industrial applications of computed tomography, in *CIRP annals* 63(2), pp. 655–677, 2014.
14. T. Chernyakova, Y. C. Eldar, Fourier Domain Beamforming: The Path to Compressed Ultrasound Imaging, in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 61(8), pp. 1252–1267, 2014.
15. D. Cohen, Y. C. Eldar, Sub-nyquist radar systems: Temporal, spectral, and spatial compression, in *IEEE Signal Processing Magazine* 35(6), pp. 35–58, 2018.
16. W. L. Chan, L. Matthew, L. Moravec, R. G. Baraniuk, D. M. Mittleman, Terahertz imaging with compressed sensing and phase retrieval, in *Optics Letters* 33(9), pp. 974–976, 2008.
17. E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, in *IEEE Transactions on Information Theory* 52(2), pp. 489–509, 2006.
18. G.-H. Chen, J. Tang, S. Leng, Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets, in *Medical physics* 35(2), pp. 660–663, 2008.
19. M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, R. G. Baraniuk, Single-pixel imaging via compressive sampling, in *IEEE Signal Processing Magazine* 25(2), pp. 83–91, 2008.
20. P. Dabkowski, Y. Gal, Real time image saliency for black box classifiers, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 6970–6979, 2017.
21. DGZfP Fachausschuss ZfP4.0, Profile and Fields of Activities, <https://www.dgzfp.de/Fachausschuss%3BC3%BCsse/ZfP-40> (as consulted online on 09 December 2021).
22. R. Deppe, O. Nemitz, J. Herder, Augmented reality for supporting manual nondestructive ultrasonic testing of metal pipes and plates. *Virtuelle und Erweiterte Realität – 15, Workshop der GI-Fachgruppe VR/AR*, 2018.
23. German expert group for artificial intelligence, internet platform: <https://din.one/site/ki> (as consulted online on 09 December 2021).
24. German standardization roadmap artificial intelligence/Deutsche Normungsroadmap Künstliche Intelligenz, <https://www.dke.de/resource/blob/2017010/99bc6d952073ca88f52c0ae4a8c351a8/nr-ki-english---download-data.pdf> (as consulted online on 09 December 2021).
25. D. L. Donoho, Compressed sensing, in *IEEE Transactions on Information Theory* 52(4), pp. 1289–1306, 2006.
26. L. Duerkop, Stand der Technik und der Forschung, in *Automatische Konfiguration Von Echtzeit-Ethernet*, pp. 7–74, Springer, 2017.
27. F. Ellrich, M. Bauer, N. Schreiner, A. Keil, T. Pfeiffer, J. Klier, S. Weber, J. Jonuscheit, F. Friederich, D. Molter, Terahertz Quality

- Inspection for Automotive and Aviation Industries, in *Journal of Infrared, Millimeter, and Terahertz Waves*, 41, 470–489, 2020, <https://doi.org/10.1007/s10762-019-00639-4>.
28. Y. K. Esfandabadi, L. De Marchi, N. Testoni, A. Marzani, G. Masetti, Full wavefield analysis and damage imaging through compressive sensing in lamb wave inspections, in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 65(2), pp. 269–280, 2017.
 29. ENIQ Recommended Practice 2: Strategy and Recommended Contents for Technical Justifications – Issue 3, ENIQ report no. 54, The NUGENIA Association, 2018.
 30. The European Methodology for Qualification of Non-Destructive Testing – Issue 4, ENIQ report no. 61, The NUGENIA Association, 2019.
 31. Qualification of Non-Destructive Testing Systems that Make Use of Machine Learning – Issue 1, ENIQ report No. 65, The NUGENIA Association, 2021.
 32. O. K. Ersoy, *Diffraction, Fourier Optics and Imaging*, Vol. 30, John Wiley & Sons, 2006.
 33. T. Fawcett, An introduction to ROC analysis, in *Pattern Recognition Letters* 27, pp. 861–874, 2015.
 34. J. Flannery, G. M. Crean, S. C. O. Mathuna, Imaging of Integrated Circuit Packaging Technologies Using Scanning Acoustic Microscopy, in: Ermert H., Harjes H. P. (eds.), *Acoustical Imaging. Acoustical Imaging*, vol. 19, Springer, Boston, MA, 1992, https://doi.org/10.1007/978-1-4615-3370-2_113.
 35. J. Francq, F. Frick, Introduction to Hardware Trojan Detection Methods, in *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 9–13 March 2015, Grenoble (France), <https://doi.org/10.7873/DATE.2015.1101>.
 36. D. Fujiki, X. Wang, A. Subramaniyan, R. Das, *Synthesis Lectures on Computer Architecture*, Vol. 16, No. 2, pp. 1–140, 2021, <https://dx.doi.org/10.2200/S01109ED1V01Y202106CAC057>.
 37. J. Hartmann, Global Trends in Microelectronics and how Europe can address them, in *ESSCIRC/ESSDERC 2020 Conference*, Sep 2020, Grenoble (France), <https://www.ipcei-me.eu/wp-content/uploads/2020/11/13ahartmannslides1598725993144.pdf> (as consulted online on 10 December 2021).
 38. M. Hill, B. Faupel, A Robotized Non-destructive Quality Device for the Inspection of Glue Joints by Active Thermography, in *Journal of Nondestructive Evaluation* 39(3), 2020.
 39. K. He, G. Gkioxari, P. Dollar, R. Girshick, Mask R-CNN, in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017.
 40. International Data Space Association – White paper and rule book, Version 1.0, December 2020, Open Internet Publication: https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-White-Paper-IDSA-Rule-Book.pdf (as consulted online on 13 December 2021).
 41. Fraunhofer IZFP, NDE 4.0 – digitale Transformation der zerstörungsfreien Prüfung/Digital Transformation of NDE, <https://www.youtube.com/watch?v=74Kn4LDwkE0> (as consulted online on 13 December 2021).
 42. A. Jüngert, S. Dugan, G. Wackenhut, R. Lambert, M. Spies, H. Rieder. Bewertung der Zuverlässigkeit geschweißter Komponenten unter Einbeziehung von Ultraschallprüfungen an realistischen Testfehlern, DGZfP (Hg.), Jahrestagung, 2017.
 43. C. Jansen, S. Wietzke, O. Peters, M. Scheller, N. Vieweg, M. Salhi, N. Krumbholz, C. Jordens, T. Hochrein, M. Koch, Terahertz imaging: applications and perspectives, in *Applied Optics* 49(19), pp. E48–E57, 2010.
 44. F. Krieg, S. Kodera, J. Kirchhof, F. Römer, A. Ihlow, S. Lugin, A. Osman, G. Del Galdo, 3D reconstruction of handheld data by SAFT and its impediment by measurement inaccuracies, in *Proceedings of the 2019 IEEE International Ultrasonics Symposium*, Oct 2019, Glasgow, UK, <http://dx.doi.org/10.1109/ULTSYM.2019.8926018>.
 45. M. Kleinemeier, Von Der Automatisierungspyramide Zu Unternehmenssteuerungs-Netzwerken. in *Handbuch Industrie 4.0*, Bd. 1, pp. 219–226, Springer, 2016.
 46. P. Kruizinga, P. van der Meulen, A. Fedjajevs, F. Mastik, G. Springeling, N. de Jong, J. G. Bosch, G. Leus, Compressive 3D ultrasound imaging using a single sensor, in *Science advances* 3(12), e1701423, 2017.
 47. A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet classification with deep convolutional neural networks, in *Commun. ACM* 60(6), pp. 84–90, 2017, <https://dx.doi.org/10.1145/3065386>.
 48. J. Kirchhof, S. Semper, F. Römer, GPU-Accelerated Matrix-Free 3D Ultrasound Reconstruction for Nondestructive Testing, in *2018 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, 2018, <https://dx.doi.org/10.1109/ULTSYM.2018.8579936>.
 49. J. Kirchhof, S. Semper, C. Wagner, E. Pérez, F. Römer, G. Del Galdo, Frequency Sub-Sampling of Ultrasound Non-Destructive Measurements: Acquisition, Reconstruction and Performance, in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 68(10), pp. 3174–3191, 2021, <https://dx.doi.org/10.1109/TUFFC.2021.3085007>.
 50. M. Lustig, D. L. Donoho, J. M. Santos, J. M. Pauly, Compressed sensing MRI, in *IEEE Signal Processing Magazine* 25(2), pp. 72–82, 2008.
 51. L. Li, D. Liu, J. Liu, H. Zhou, J. Zhou, Quality Prediction and Control of Assembly and Welding Process for Ship Group Product Based on Digital Twin, in *Scanning* 2020, 3758730, 2020.
 52. S. Lehnhoff, S. Rohjans, M. Uslar, W. Mahnke, OPC Unified Architecture: A Service-oriented Architecture for Smart Grids, in *First International Workshop on Software Engineering Challenges for the Smart Grid (SE-SmartGrids)*, pp. 1–7, 2012, <https://dx.doi.org/10.1109/SE4SG.2012.6225723>.
 53. N. Laleni et al., In-Memory Computing exceeding 10000 TOPS/W using Ferroelectric Field Effect Transistors for EdgeAI Applications, in *Mikrosystemtechnik Kongress, Stuttgart-Ludwigsburg*, 2021.
 54. C. Müller, M. Bertovic, M. Gaal, H. Heidt, M. Pavlovic, M. Rosenthal, K. Takahashi, J. Pitkänen, Ulf Ronneteg, Progress in Evaluating the Reliability of NDE Systems – Paradigm Shift, in *4 European-American Workshop on Reliability of NDE-WE.1.A.2*, Berlin 2009.
 55. J. M. Montero, G. Fernández-Avilés, J. Mateu, Spatial and spatio-temporal geostatistical modeling and kriging, Vol. 998, John Wiley & Sons, 2015.
 56. X. Maldague, S. Marinetti, Pulse phase infrared thermography, in *Journal of applied physics* 79(5) pp. 2694–2698, 1996.
 57. J. Meyer, J. Rehbein, J. de Freese, J. Holtmannspötter, Visualisation of ultrasonic testing data using augmented reality, in *7th International Symposium on NDT in Aerospace*,

- Bremen, Germany, 2015.
58. T. V. Nguyen, S. Kamma, V. Adari, T. Lesthaeghe, T. Boehnlein, V. Kramb, Mixed reality system for nondestructive evaluation training, in *Virtual Reality* 25(3), pp. 709–718, Springer, 2021.
 59. A. Osman, Y. Duan, V. Kaftandjian, Applied Artificial Intelligence, in NDE. In: Meyendorf N., Ida N., Singh R., Vrana J. (eds.), *Handbook of Nondestructive Evaluation 4.0*, Springer, Cham, 2021, https://doi.org/10.1007/978-3-030-48200-8_49-1.
 60. R. S. Fernandez, K. Hayes, F. Gayosso, Artificial Intelligence and NDE Competencies, in *Handbook of Nondestructive Evaluation 4.0*, pp. 1–53, 2021.
 61. A. Osman, Automated evaluation of three dimensional ultrasonic datasets, Diss. INSA de Lyon, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2013.
 62. A. du Plessis, W. P. Bosh, A review of X-ray computed tomography of concrete and asphalt construction materials, in *Construction and Building Materials* 199, pp. 637–651, 2019.
 63. E. Pérez, J. Kirchhof, F. Krieg, F. Römer, Subsampling Approaches for Compressed Sensing with Ultrasound Arrays in Non-Destructive Testing, in *MDPI Sensors* 20(23), 2020, <https://dx.doi.org/10.3390/s20236734>.
 64. R. Pandey, J. Kirchhof, F. Krieg, E. Pérez, F. Römer, Preprocessing of Freehand Ultrasound Synthetic Aperture Measurements using DNN, in *Proceedings of the 29th European Signal Processing Conference (EUSIPCO-2021)*, Dublin, Ireland, 2021.
 65. A. Rodriguez-Molares, O. M. Hoel Rindal, J. D'hooge, S. Måsøy, A. Austeng, H. Torp, The Generalized Contrast-to-Noise Ratio, in *2018 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, 2018, <https://dx.doi.org/10.1109/ULTSYM.2018.8580101>.
 66. N. Rajic, Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures, in *Composite Structures* 58(4), pp. 521–528, 2002.
 67. V. K. Rentala, D. Kanzler, P. Fuchs, POD Evaluation: The Key Performance Indicator for NDE 4.0, in *Journal of Nondestructive Evaluation* 41(20), 2022, <https://doi.org/10.1007/s10921-022-00843-8>.
 68. F. Römer, J. Kirchhof, F. Krieg, E. Pérez, Compressed Sensing: From Big Data to Relevant Data, in Meyendorf N., Ida N., Singh R., Vrana J. (eds.), *Handbook of Nondestructive Evaluation 4.0*, Springer, Cham, 2021.
 69. C. Rooney, A. Seeam, X. Bellekens, Creation and Detection of Hardware Trojans Using Non-Invasive Off-The-Shelf Technologies, in *Electronics* 7, 2018, <https://doi.org/10.3390/electronics7070124>.
 70. A. Ramkumar, A. K. Thittai, Strategic undersampling and recovery using compressed sensing for enhancing ultrasound image quality, in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 67(3), pp. 547–556, 2019.
 71. B. Sanno, Detecting Hardware Trojans, Ruhr-University Bochum, July 2009, https://www.emsec.ruhr-uni-bochum.de/media/crypto/attachments/files/2011/03/benjamin_sanno.semembsec_termpaper_20090723_final.pdf (as consulted online on 10 December 2021).
 72. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 618–626, 2017.
 73. A. Styperek, M. Ciesielczyk, A. Szwab, P. Misiorek, Evaluation of SPARQL-compliant Semantic Search User Interfaces, in *Vietnam Journal of Computer Science* 2(3), pp. 191–199, 2015.
 74. P. Swierczynski, M. Fyrbiak, P. Koppe, A. Moradi, C. Paar, Interdiction in Practice – Hardware Trojan Against a High-Security USB Flash Drive, in *J. Cryptogr. Eng.* 7, pp. 199–211, 2017, <https://doi.org/10.1007/s13389-016-0132-7>.
 75. M. Schickert, C. Koch, F. Bonitz, Prospects for Integrating Augmented Reality Visualization of Nondestructive Testing Results into Model-Based Infrastructure Inspection, in *NDE/NDT for Highways & Bridges*, New Brunswick, NJ, USA, 2018.
 76. E. Y. Sidky, C.-M. Kao, X. Pan, Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT, in *Journal of X-ray Science and Technology* 14(2), pp. 119–139, 2006.
 77. T. Schön, F. Römer, S. Oeckl, M. Großmann, R. Gruber, A. Jung, G. Del Galdo, Cycle time reduction in process integrated computed tomography using compressed sensing, in *Proceedings of the 13th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Fully 3D)*, Newport, RI, 2015.
 78. E. Schrüfer, L. Reindl, B. Zagar, Elektrische Messtechnik, Messungen elektrischer und nichtelektrischer Größen, Hanser, Carl GmbH + Co, 2014.
 79. SPIEGEL Staff: Inside TAO: Documents reveal top NSA hacking unit (2013). <http://www.spiegel.de/international/world/the-nsa-uses-powerful-toolbox-in-effort-to-spy-on-global-networks-a-940969.html> (as consulted online on 09 December 2021).
 80. X.-H. Sun, Remove the memory wall: from performance modeling to architecture optimization, in *Proceedings 20th IEEE International Parallel & Distributed Processing Symposium*, 2 pp., 2006, <https://dx.doi.org/10.1109/IPDPS.2006.1639621>.
 81. K. Simonyan, A. Vedaldi, A. Zisserman, Deep inside convolutional networks: Visualising image classification models and saliency maps, in *Workshop at International Conference on Learning Representations*, 2014.
 82. R. Tschuncky, K. Szielasko, I. Altpeter, Hybrid Methods for Materials Characterization, in *Materials Characterization Using Nondestructive Evaluation (NDE) Methods*, pp. 263–291, Woodhead Publishing, Cambridge, 2016.
 83. B. Valeske, International Virtual Conference on NDE 4.0, 14/15 & 20/21 April 2021, abstract book, page 29, https://2021.nde40.com/Portals/ndepre2021/Dokumente/NDE40_Abstracts.pdf, see also: <https://www.youtube.com/watch?v=MQbQVHm-i3E> (as consulted online on 13 December 2021).
 84. VDMA 40010-1: OPC UA Companion Specification for Robotics (OPC Robotics) – Part 1: Vertical integration, 2019.
 85. VDMA 40100-1: OPC UA for Machine Vision (OPC Machine Vision) – Part 1: Control, configuration management, recipe management, result management, 2019.
 86. VDMA 40001-1: OPC UA for Machinery (OPC Machinery) – Part 1: Basic Building Blocks, 2021.
 87. A. Vick, J. Krueger, Using OPC UA for Distributed Industrial Robot Control, in *ISR 2018; 50th International Symposium on Robotics*, pp. 1–6, 2018.

88. J. Vrana, N. Meyendorf, N. Ida, R. Singh, Introduction to NDE 4.0, in *Handbook of Nondestructive Evaluation 4.0*, pp. 1–28, Springer, 2021.
89. B. Valeske, A. Osman, F. Römer, R. Tschuncky, Next Generation NDE Sensor Systems as IIoT Elements of Industry 4.0, in *Research in Nondestructive Evaluation 31*(5-6), pp. 340–369, 2020, <https://dx.doi.org/10.1080/09349847.2020.1841862>.
90. J. Vrana, NDE Perception and Emerging Reality: NDE 4.0 Value Extraction, in *Materials Evaluation 78*(7), pp. 835–851, 2020.
91. J. Vrana, The Core of the Fourth Revolutions: Industrial Internet of Things, Digital Twin, and Cyber-Physical Loops, in *Journal of Nondestructive Evaluation 40*(46), 2021. <https://doi.org/10.1007/s10921-021-00777-7>.
92. J. Vrana, R. Singh, The World of NDE 4.0, amazon fulfillment book print, 2021, ISBN: 979-8-4625-1421-0.
93. J. Vrana, R. Singh, Cyber-Physical Loops as Drivers of Value Creation in NDE 4.0, in *Journal of Nondestructive Evaluation 40*(61), 2021, <https://doi.org/10.1007/s10921-021-00793-7>.
94. J. Vrana, R. Singh, NDE 4.0—A Design Thinking Perspective, in *J. Nondestruct. Eval. 40*, p. 8, 2021, <https://doi.org/10.1007/s10921-020-00735-9>.
95. J. Vrana, K. Schörner, H. Mooshofer, K. Kolk, A. Zimmer, K. Fendt, Ultrasonic Computed Tomography – Pushing the Boundaries of the Ultrasonic Inspection of Forgings, in *Steel Research Int. 89*, 1700448, 2018, <https://doi.org/10.1002/srin.201700448>.
96. M. Wang, *Industrial tomography: systems and applications*, Elsevier, 2015.
97. Z. Wei, H. Fernandes, J. R. Tarpani, A. Osman, X. Maldague, Stacked denoising autoencoder for infrared thermography image enhancement, in *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, pp. 1–7, 2021.
98. B. Wolter, Y. Gabi, C. Conrad, Nondestructive Testing with 3MA—An Overview of Principles and Applications, in *Applied Sciences 9*(6), p. 1068, 2019, <https://doi.org/10.3390/app9061068>.
99. A. Wilken, F. Hellemann, L. Turgut, G. Helfrich, Concept for digitalisation of an inspection process using hybrid tracking of part and probe for future maintenance and digital twins, in *Deutscher Luft- und Raumfahrtkongress 2021*, Bremen, Germany, 2021.
100. Z. Wei, A. Osman, D. Gross, U. Netzelmann, Artificial Intelligence for Defect Detection in Infrared Images of Solid Oxide Fuel Cells, in *Infrared Physics & Technology 103*815, 2021, <https://dx.doi.org/10.1016/j.infrared.2021.103815>.
101. C. Xu, J. Xie, C. Wu, L. Gao, G. Chen, G. Song, Enhancing the Visibility of Delamination during Pulsed Thermography of Carbon Fiber-Reinforced Plates Using a Stacked Autoencoder, in *Sensors (Basel, Switzerland) 18*(9), 2018, <https://dx.doi.org/10.3390/s18092809>.
102. J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, H. Lipson, Understanding neural networks through deep visualization, arXiv preprint arXiv:1506.06579, 2015.
103. Y. Yang, M. Pesavento, A unified successive pseudoconvex approximation framework, in *IEEE Transactions on Signal Processing 65*(13), pp. 3313–3328, 2017.
104. J. Yan, Y. Meng, L. Lu, L. Li, Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance, in *IEEE Access 5*, pp. 23484–23491, 2017, <https://dx.doi.org/10.1109/ACCESS.2017.2765544>.
105. J. Zhang, Z. Wang, N. Verma, A machine-learning classifier implemented in a standard 6T SRAM array, in *IEEE Symposium on VLSI Circuits*, pp. 1–2, 2016.

Bionotes



Prof. Dr. Bernd Valeske

Fraunhofer IZFP, Institute for Nondestructive Testing, 66123 Saarbruecken, Germany
University of Applied Sciences – htw saar, 66117 Saarbruecken, Germany
bernd.valeske@izfp.fraunhofer.de

Prof. Dr.-Ing. Bernd Valeske is managing director of Fraunhofer IZFP and professor for quality control and maintenance at the Saarland University of Applied Sciences. His scientific focus is in research for intelligent sensor systems and in developments for tailored sensing and monitoring of materials, products and production processes with almost 20 years of experience in NDE. He is chairman of the NDE4.0 expert group of the German Society for Nondestructive Testing (DGZfP).



Dr. Ralf Tschuncky

Fraunhofer IZFP, Institute for Nondestructive Testing, 66123 Saarbruecken, Germany
ralf.tschuncky@izfp.fraunhofer.de

Dr. Ralf Tschuncky holds a PhD in Material Science from University of Saarland. He is Chief Scientist at the Fraunhofer IZFP and he has the position of the institute's research manager. He started in 1999 to work at the Fraunhofer IZFP. Until today, he works in the development, evaluation and application of nondestructive testing and evaluation technologies including corresponding data processing and analysis methods at Fraunhofer IZFP.



M. Sc. Frank Leinenbach
Fraunhofer IZFP, Institute for
Nondestructive Testing, 66123
Saarbruecken, Germany
frank.leinenbach@izfp.fraunhofer.de

Frank Leinenbach studied electrical engineering at the University of Applied Sciences in Saarbruecken, Germany, where he received the Master of Science degree in electrical engineering in September 2013. In July 2015 he joined the Fraunhofer Institute for Nondestructive Testing IZFP in the department for production integrated NDT with a research focus on automation, control and communication for nondestructive testing. He leads the working group OPC UA in the NDE4.0 expert group of the German Society for Nondestructive Testing (DGZfP) since October 2021.



Dr. Florian Römer
Fraunhofer IZFP, Institute for
Nondestructive Testing, 66123
Saarbruecken, Germany
florian.roemer@izfp.fraunhofer.de

Florian Römer studied computer engineering at the Ilmenau University of Technology, Germany, and McMaster University, Hamilton, ON, Canada. He received the Diplom-Ingenieur degree in communications engineering and the doctoral (Dr.-Ing.) degree in electrical engineering from the Ilmenau University of Technology in October 2006 and October 2012, respectively. In January 2018 he joined the Fraunhofer Institute for Nondestructive Testing IZFP where he is leading the SigMaSense group with a research focus on innovative sensing and signal processing for material diagnostics and nondestructive testing.



Prof. Dr. Ahmad Osman
Fraunhofer IZFP, Institute for
Nondestructive Testing, 66123
Saarbruecken, Germany
University of Applied Sciences – htw saar,
66117 Saarbruecken, Germany
ahmad.osman@izfp.fraunhofer.de

Prof. Dr.-Ing. Ahmad Osman is group manager and vice department leader at Fraunhofer IZFP. Osman is professor for Artificial Intelligence, Sensing Sensors and Systems at the Saarland University of Applied Sciences and adjunct professor at Laval university, Canada. His scientific focus is in research and application of modern artificial intelligence methods to develop intelligent NDE systems for structural health assessment and quality control of processes, products and structures. Osman is a senior IEEE member and chairman of the AI expert group of the German Society for Nondestructive Testing.



M. Sc. Dirk Koster
Fraunhofer IZFP, Institute for
Nondestructive Testing, 66123
Saarbruecken, Germany
dirk.koster@izfp.fraunhofer.de

Dirk Koster studied electrical engineering at the University of Applied Sciences in Saarbruecken, Germany. He received the Diplom-Ingenieur (FH) degree in communications engineering in October 2007 and the Master of Science degree in micro- and telecommunication electronics in September 2013. In November 2007, he joined the Fraunhofer Institute for Nondestructive Testing IZFP in the department electronics for NDT systems with research focus on electronic development for nondestructive testing. He is Topic Coordinator for the eddy current technique and since January 2021, he is Technology Manager for Sensor-Intelligence Devices.



M. Sc. Ziang Wei
University of Applied Sciences – htw saar,
66117 Saarbruecken, Germany
ziang.wei@htw-saarland.de

Ziang Wei is an AI researcher at University of Applied Sciences in Saarbrücken (htw saar). He received master's degree from RWTH Aachen University, Aachen, in 2017. He was a data scientist at Fraunhofer IZFP in Saarbrücken Germany. His research interests include machine learning, data processing, Explainable AI in NDE4.0 and thermography.



M. Sc. Kevin Becker
Fraunhofer IZFP, Institute for
Nondestructive Testing, 66123
Saarbruecken, Germany
kevin.becker@izfp.fraunhofer.de

Kevin Becker studied electrical and information engineering at Technical University of Darmstadt, Germany, where he received his Master of Science in electrical and informational engineering with focus on electronics in April 2020. In 2020 he joined the Fraunhofer Institute for Nondestructive Testing IZFP in the department of Sensor Intelligence Devices, where he is currently working on his PhD.

**Dipl.-Ing. Thomas Schwender**

Fraunhofer IZFP, Institute for
Nondestructive Testing, 66123
Saarbruecken, Germany

thomas.schwender@izfp.fraunhofer.de

Thomas Schwender is head of the department for the development of NDE systems for the quality control of components and parts and technical manager of the DIN EN ISO 17025 accredited test laboratory at Fraunhofer IZFP. On the scientific side, the focus is on the further development of ultrasonic testing and evaluation techniques, which also includes the application of advanced reconstruction methods and machine learning concepts. In his more than 15 years of work in the field of NDE, he has application experience in many industrial sectors from power generation to railroad applications. Furthermore, he is engaged in the field of innovation management with the task of integrating general technological trends in the future development of NDE Systems.