Internet of Things (IoT) in the printing industry

Stefan Meissner[1]

[1]TU Dublin

Keywords: Internet of Things (IoT), Industrial Internet of Things (IIoT), Industry 4.0, Print 4.0, JDF, XJDF

Abstract
Internet of Things (IoT), Industry 4.0, Big Data, Industrial Internet, etc. – these are all terms that are increasingly hitting the printing industry. Printers might ask how these technologies affect their daily production business and how they collaborate with their current technological environment? This research elaborates the ideas and concepts of the Internet of Things (IoT), Industry 4.0, Big Data, Industrial Internet, etc. It explores how these technologies can seamlessly integrate into print production lines. This publication aims to work out a general basic scientific understanding of the newly emerging technologies from the printing industry and shows how these technologies collaborate with existing print production equipment. Finally, the paper ends with an example of implementing IoT/Industry 4.0 in a printing house.

Introduction

Internet of Things (IoT), Industry 4.0, Big Data, Industrial Internet, etc. – these are all terms that are increasingly hitting the printing industry. Printers might ask how these technologies affect their daily production business and how they collaborate with their current technological environment? As many devices and applications already provide a standardized interface such as Job Definition Format (JDF) / Job Messaging Format (JMF), another question might be if there is a way to use these existing interfaces in the new context? In the industrial context, new interface technologies cannot be implemented as fast as in the pure IT context as the investment cycle is longer. Many industrial production systems still produce high-quality products to competitive costs - but from an IT perspective, they are considered legacy systems. Based on that background, the critical question is not IF there is a way to use the existing interfaces but HOW to get it working. One key to a successful introduction of IoT, Industry 4.0, etc., in the printing industry, is to bring the status quo production devices and -applications into the Cloud without changing their existing interfaces.

This research elaborates the idea and concepts behind the Internet of Things (IoT), Industry 4.0, Industrial Internet, etc., and explores how these technologies can seamlessly integrate into print production lines. The research starts with a thorough examination of IoT to manifest a general understanding and identify its key components. Next, the study will do a deep dive into the existing printing technologies relevant in the IoT context. The research aims to identify the critical elements needed from both fields, IoT and printing. Further, it provides a first approach of how to build an IoT-enabled print production line. The research intends to be an entry point for printing companies that want to use IoT in their production line. Thus, this publication follows this research question: **How to start to integrate Internet of Things (IoT) / Industry 4.0 technologies and methods seamlessly in a print production line considering the existing printing industry's standard technologies?**

The scope of this publication includes basic technology research and a theoretical example of a simple application of how to integrate IoT/ Industry 4.0 technologies in print production. The knowledge generation is based on scientific papers, industry standards, and technical documentation of leading organizations/ experts of the appropriate field. However, this publication does not include the physical implementation of a real IoT/Industry 4.0 application. The paper focuses on the theoretical technical components and their required capabilities rather than on specific products. For instance, this research would elaborate on the needed characteristics of a database rather than on a particular product. This document claims to be essential research in its field to support the reader in understanding and providing food for thought about implementing IoT/ Industry 4.0 in print production. However, an actual technical implementation of IoT/Industry 4.0 requires the reader to study additional, more specific technical documentation. As previously mentioned, this text intends to provide a solid entry point into this topic and not to be a complete technical guideline for implementation.

IT Technologies

According to Santucci (2009, p. 2), the term 'Internet of Things' (IoT) has been introduced by Kevin Ashton in 1998. Ashton (2016) self elaborated on his blog that the term came up while working on a presentation with many visionaries in spring 1999, and he cannot exclude that someone else in the team came up with it first. Since then, the term 'Internet of Things' (IoT) has gained global adoption. Sisinni et al. (2018, p. 4724) describe IoT as turning ordinary objects into connected smart devices capable of sensing the surrounding environment, transmitting and processing acquired data, and then feedback to the environment. Bandyopadhyay and Sen (2011, p. 50) parallel Sisinni's view but from a more differentiated perspective on IoT. They distinguish IoT between the 'Internet' and the 'Things': "The first one pushes towards a network-oriented vision of IoT, while the second tends to move the focus on generic objects to be integrated into a common framework." The INFSO (2008, p. 4) describes IoT as "a worldwide network of interconnected objects uniquely addressable, based on standard communication protocols." In today's world, IoT is playing an increasingly significant role. Examples of IoT devices include smart sensors for measuring temperature, humidity, pressure, etc. (BehrTech, 2020), smart speakers such as 'Google Home' or 'Amazon Echo', and also wearables such as smartwatches.

In the industrial environment, the term 'Industrial Internet of Things' (IIoT) is increasingly emerging. Sisinni et al. (2018, p. 4724) see the IIoT as a subset of IoT focusing on machine-to-machine communication to understand the manufacturing process better and ensure efficient and sustainable production. In opposition to IIoT, Sisinni et al. (2018, p. 4725) call the usual IoT 'Consumer IoT' (CIoT) and describes it as machine-to-user communication targeted to improve human awareness of the surrounding environment to save time and money. Figure 1 provides a qualitative comparison between CIoT and IIoT. The classification of IoT in CIoT and IIoT is also supported by Wurm et al. (2016, p. 519); however, they call CIoT 'Commercial IoT' rather than 'Consumer IoT.' The term IIoT is often used in the context of 'Industry 4.0' - the next level of the industrial revolution. Wan et al. (2016, p. 7373) see IIoT as the fundamental premise for 'Industry 4.0' to obtain continuous information from the shop floor's sensors and objects. Sisinni et al. (2018, p. 4724) state from a more general perspective that IoT enables industries (and society) to interact efficiently between the physical world and its digital counterpart.

	Consumer IoT	Industrial IoT
Impact	Revolution	Evolution
Service Model	Human-centered	Machine-oriented
Current Status	New devices and standards	Existing devices and standards
Connectivity	Ad-Hoc (infrastructure is not tolerated; nodes can be mobile)	Structured (nodes are fixed; centralized network management)
Criticality	Not stringent (excluding medical applications)	Mission critical (timing, reliability, security, privacy)
Data Volume	Medium to High	High to Very High

Figure 1: Comparison between Consumer IoT and Industrial IoT by Sisinni et al. (2018, p. 4725)

In Figure 1, Sisinni et al. (2018, p. 4725) provide a comparison between Consumer IoT (CIoT) and Industrial IoT (IIoT). They consider CIoT as a revolution as it emphasizes novel devices and new communication standards. In contrast, IIoT is about integrating existing production floors based on interfaces that might have been designed a long time before IoT has emerged. Sisinni et al., therefore, see the latter approach as an evolution. Breivold and Sandström (2015, p. 536) parallel the evolutionary perspective in an industrial environment: They describe industrial systems as long-living legacy systems with a standard lifetime of ten to thirty years and continuously subjected to a substantial amount of evolutionary changes. Druckmarkt (2002, p. 109) writes about the printing industry that offset presses have a life cycle between seven and twelve years typically. Examples of communication standards in the Graphic Arts Industry include the Portable Document Format (PDF) and the Job Definition Format (JDF). A PDF is a standard data format to specify the graphical artwork itself, while a JDF determines the product- and job metadata such as substrate, amount, and delivery time (Meissner, 2019, p. 12). Both file formats have their origins in the nineties and are still heavily used in many print production lines. Bringing IIoT into the printing industry probably means finding an evolutionary approach to using PDF/JDF in the new context.

Internet of Things in general (CIoT and IIoT) is a crucial component of Industry 4.0. The center of Industry 4.0 is a large data volume that is continuously being organized, analyzed, and mined to obtain valuable knowledge for judging, decision making, or reasoning. IoT is used to create and constantly extend the large data volume with new structured, semi-structured, and unstructured information reflecting the physical production line (Qi & Tao, 2018). IoT allows a real-time interconnection between the physical objects on the production floor and their digital representations in the data storage. Sisinni et al. (2018, p. 4726) refer to the objects' digital representation as to their second identity or 'Digital Twin.' 'Digital Twin' is a commonly used term in the context of Industry 4.0: Rod'ko et al. (2020, p. 423) describe 'Digital Twin' as a "virtual prototype of a real physical instrument, group of devices, or process, which models the technical characteristics, internal processes, and behavior of the real object in its surroundings." Anecdotally many scientific sources support the definition of Rod'ko et al. However, sources outside the scientific community sometimes use the term 'Digital Twin' in a different way. For instance, General Electric (2016) uses the term 'Digital Twin' for an entire highly integrated system rather than just for the digital model - what it actually should be. It seems that the term 'Digital Twin' prone to misunderstandings. Maybe it is reasonable to use more precise terms such as 'digital replica' or 'digital model' instead of 'Digital Twin.' The proceeding document uses the term 'digital replica'.

Cyber-Physical Systems (CPS) is often used in the context of the Industrial Internet of Things and Industry 4.0. The National Science and Technology Council defines CPS as "networked computing systems – interconnected software, microprocessors, sensors, and actuators – deeply integrated within engineered physical systems to monitor and control capabilities and behaviors of the physical system as a whole" (NITRD, 2012, p. 15). Cyber-Physical Systems need to be distinguished from embedded systems: Whereas embedded systems are designed as stand-alone devices, CPS'es focus on the network of those (Jazdi, 2014, p. 1). Lee (2008, p. 363) describes Cyber-Physical Systems as "integrations of computation and physical processes." Lee further elaborates on CPS'es as "embedded computers and networks monitor and control the physical processes, usually with feedback loops where physical processes affect computations and vice versa." In the context of the printing industry, a single prepress, press-, or post-press device applies to an embedded device. In contrast, the entire production process applies to a Cyber-Physical System. In Figure 2, Sisinni et al. (2018) illustrated how CPS relates to IoT, IIoT, and Industry 4.0: They see Industry 4.0 due to the integration of the Internet of Things, Industrial Internet of Things, and Cyber-Physical Systems.

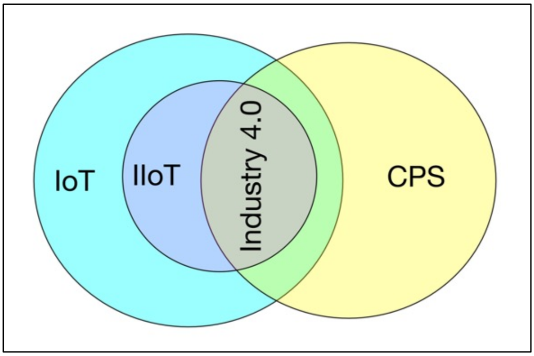


Figure 2: IoT, CPS, IIoT, and Industry 4.0 in Venn diagram by Sisinni et al. (2018, p. 4724)

As previously mentioned, one base concept of Industry 4.0 is to create a digital replica of a production line by continuously recording production and environmental parameters using '(Industrial) Internet of Things.' Mourtzis et al. (2016, p. 290) examined this IoT adaption in manufacturing, which empowers companies to foster data-driven strategies. They wrote about a transition from 'Industrial Data' into 'Industrial Big Data' due to an increased total data volume. Mourtzis et al. (2016, p. 292) further distinguish strictly between 'Big Data' and 'Industrial Big Data' whereas the latter is about machine-generated information within an industrial environment to realize Industry 4.0. Lee et al. (2014, p. 4) support this distinction between 'Big Data' and 'Industrial Big Data.' They further observed that most of the research about big data focuses on social or commercial mining based 'human-generated' or human-related information rather than on machine-generated or industrial information in the industrial context. Snijders et al. (2012, p. 1) consider 'Big Data' in general as a loosely defined term to describe large and complex data sets that are "awkward" to be processed by standard statistical software. Mauro et al. (2016) further examined the definition of 'Big Data' and finally published a proposal which underpins Snijders et al. consideration. They proposed to define 'Big Data' as "the Information asset characterized by such a High Volume, Velocity, and Variety to require specific Technology and Analytical Methods for its transformation into Value." Many scientific papers support this definition.

'Data Mining' is a term frequently used in combination with 'Big Data.' Wang and Wang (2015, p. 143) describe 'Data Mining' as a technique to discover patterns and models in the context of large-scale data. Leskovec et al. (2014, p. 17) parallel the meaning of model extraction from data, but they do not limit 'Data Mining' to 'Big Data. Leskovec et al. (2014, p. 21) rather see 'Big Data' as one (less challenging) application of 'Data Mining.' Ratner (2017, p. 9) states that today's statisticians define 'Data Mining' as any process that finds unexpected structures in data. Ratner further sorts out that 'Data Mining' is the process of exploring the data, not exploiting it. 'Data Mining' can be distinguished between descriptive and predictive. Descriptive mining characterizes properties in a data set, while predictive mining performs induction on a data set to make predictions (Han, Kamber, & Pei, 2012, p. 15). Data Scientists can use various techniques concerning 'Data Mining' including Classification Analytics, Association Rule Learning, Anomaly or Outlier Detection, Clustering Analysis, Regression Analysis, or Machine Learning and Artificial Intelligence (talend, 2021). However, the quality of the results depends strongly on the quality of the underlying data set. According to Alasadi and Bhaya (2017, p. 4102), raw data typically susceptible to missing values, noise, incompleteness, inconsistency, and therefore have to be preprocessed. Singh (2020) supports Alasadi's and Bhaya's observations. Singh further writes that data preprocessing is an essential task as it transforms raw data into a more understandable, practical, and efficient format.

Printing Technologies and Trends

The term 'Print 4.0' is an analogy of 'Industry 4.0' and was initially defined by the Bundesverband Druck und Medien (BVDM), a german printing association, in 2015. According to Belz et al. (2018, p. 6) 'Print 4.0' is used to communicate the adaption of Industry 4.0 concepts and techniques to the printing industry. Concepts and techniques in the context of 'Print 4.0' mean Cloud-Computing, Smart Factory, Augmented/Mixed Reality, Artificial Intelligence, Internet of Things, Big Data, Mass Customization (Belz, Zacharias, & Beinhauer, 2018, p. 5). Ternay (2020) also sees 'Print 4.0' as the adaption of 'Industry 4.0' in the printing industry. Ternay listed almost the same concepts and techniques as Belz et al. However, Ternay only sees 3D printing and digital printing in tomorrow's smart factories (not conventional/offset printing). Ternay does not constitute why he is excluding offset printing in smart factories. Kleeberg (2019, p. 42) emphasizes cross-linking between production components, suppliers, and customers as the centralistic topic in 'Print 4.0'. From this perspective, Kleeberg focuses on the importance of communication standards in the printing industry, such as the Exchange Job Definition Format (XJDF) and Adobe's Portable Document Format (PDF). Belz et al. (2018, p. 25) also see communication standards, such as XJDF and PDF, as a key in 'Print 4.0' to set up integrated networks interconnecting devices and customers.

The Portable Document Format (PDF) is a file format representing 'final form' formatted documents. PDF page content is not limited to text, images, graphics but can also contain multimedia assets and auxiliary structures, including metadata, bookmarks, and annotations (IETF, 2017, p. 2). The idea of PDF "is to enable users to exchange and view electronic documents easily and reliably, independently of the environment in which they were created" (Adobe, 2007, p. 25). PDF initial version has been published by Adobe Systems Incorporated in 1993. Since then, PDF has become a significant file format for capturing and exchanging formatted documents. PDF 1.7 has become an ISO standard (ISO 32000-1:2008) in 2008 (IETF, 2017, p. 3). The document format has also become a leading role in the printing industry. Kodak, for instance, has built an entire prepress workflow system around PDF. According to Kodak's Workflow Documentation (2021), all artwork files processed by their workflow are being converted internally to single pages PDFs when entering the system. Heidelberger Druckmaschinen AG (2017, p. 10) also uses PDF as their internal file format in prepress processing. Besides, PDF is the preferred file format for document exchange of many online printers, including Vistaprint, flyeralarm, and Saxoprint. PDF is the defacto standard file format in the printing industry for customer/supplier artwork transmissions and document exchange between production components.

The Exchange Job Definition Format (XJDF) and the Exchange Job Messaging Format (XJMF) are industry standards of the printing industry published by the CIP4 Organization. An XJDF describes the job and process details transferred between controllers and devices (CIP4, 2020 a, p. 21), while XJMF messages are the communication protocol (CIP4, 2020 a, p. 263). XJDF and XJMF are major version updates of the Job Definition Format (JDF) and the Job Messaging Format (JMF). Many vendors in the printing industry have implemented CIP4 Standards in their products. Heidelberger Druckmaschinen AG (2012, p. 2), for instance, states in a product brochure that Prinect Integration Manager fully utilizes the benefits of JDF. EFI (2020) announces on a product website that the "JDF technology provides an unmatched digital workflow experience." Also, HP (2020) and Koenig & Bauer (2020) support JDF/JMF in their products. Further research revealed that many commonly known companies in the printing industry support the CIP4 standard data formats. The CIP4 Organization (2020 c) writes that the organization consists of over 1,600 individuals from approximately 300 member companies that are organized in 17 working groups. CIP4 Standards seem to be an excellent approach for a unified and vendor-independent integration of devices and applications on a shop floor.

XJMF Messages are inherently different Extensible Markup Language (XML) structures as XJDF Documents and represent the spine of a well-defined communication between devices and applications (CIP4, 2020 a, p. 263). The CIP4 Organization (2020 a, p. 265f) classified XJMF Messages into four families: Queries, Commands, Signals, and Responses. Queries retrieve information from the receiver without status change, while Commands messages change the status. Signals are XJMF messages sent asynchronously to a subscriber, and Responses are synchronous replies to Queries and Commands. HP (2015, p. 11f) describes in their JDF Developer's Guide how to manage print jobs on their presses using JMF Commands and how to monitor status changes using JMF Queries and JMF Signals. Xerox (2009, pp. 3-1f) also explained in their JMF User Guide how to interact with Xerox printing systems using JMF Commands, -Queries, and Signals. It appears that the concept of JMF Messages is not only well specified by the CIP4 Organization but has also been adopted by vendors in the Graphic Arts Industry. However, investigations revealed that most vendors do (still) only support JMF, although XJMF has been available since 2018. The author has not found any vendor's developer- or user guide explaining how to integrate a printing system using XJMF.

Although JDF/JMF are official industry standards in the Graphic Arts Industry, many JDF/JMF enabled devices are not compatible out-of-the-box. Therefore, the CIP4 Organization (2020 d) provides a JDF Integration Matrix, which outlines the compatible devices and applications. One reason for that are dialects emerged because of ambiguities in the JDF Specification. So, the JDF Specification specifies the resources 'LayoutPreparationParams' (CIP4, 2020 b, p. 488), 'StrippingParams' (CIP4, 2020 b, p. 586), and 'Layout' (CIP4, 2020 b, p. 461), which all can be used to define the positioning of the customer's artwork PDFs on a print master. In the XJDF Specification, the three resources 'LayoutPreparationParams', 'StrippingParams', and 'Layout' have been migrated to 'Layout' only (CIP4, 2020 a, p. 184). Another reason for the dialects is differences in the interpretation of the JDF Specification. For instance, when subscribing to JMF Status Signals, an attribute 'RepeatTime' can be defined, which causes a periodical time-based status signal. A test implementation by the author revealed that Heidelberg's interpretation of 'RepeatTime' produces periodical status signals in addition to event-triggered signals such as for a device's status change. HP's interpretation of the attribute 'RepeatTime' is that signals are only sent time-based. Once 'RepeatTime' is specified, event-triggered signals are no longer provided and are getting lost. According to the current JDF Specification (CIP4, 2020 b, p. 141), Heidelberg's interpretation is correct. However, a previous version of the JDF Specification (CIP4, 2009, p. 189) was not that precise about the proper way of implementing 'RepeatTime'.

Albeit there are dialects in JDF (and might occur in XJDF), it is recommended to use it rather than most proprietary standards. Having a fully functional out-of-the-box integration is only the tip of an iceberg and is not that important. A far more crucial aspect of JDF/XJDF is that it can be considered a common Domain Specific Language (DSL) for the printing industry. Kosar et al. (2008, p. 390) describe DSL as "a language designed to provide a notation tailored toward an application domain, and [which] is based only on the relevant concepts and features of that domain." Kosar et al. (2008, p. 390) wrote further about the disadvantages of a DSL that it is costly to develop as it requires both domain and language development expertise. Kosar et al. see this as one reason DSLs are rarely used in solving software engineering problems. Managoli (2020) sees DSLs as a powerful method to capture domain attributes and to have a common language for domain experts and developers. As a contra Managoli (2020) wrote, it is challenging to develop a DSL as individuals are required who have domain knowledge and language-development knowledge. Using the JDF/XJDF Specifications, companies can benefit from a DSL designed and maintained by many domain experts across the printing industry based on over two decades of expert discussions for free. The specifications identify, define, and structure any kinds of products (CIP4, 2020 a, pp. 35-56), processes (CIP4, 2020 a, pp. 57-106), and resources (CIP4, 2020 a, pp. 107-261) in the printing industries domain.

5.6.15 Folding

Buckle folders or knife folders are used for **Folding** sheets. One or more sheets can be folded at the same time. Web presses often provide in-line **Folding** equipment. Longitudinal **Folding** is often performed using a former, a plow folder or a belt. Jaw folding, chopper folding or drum folding equipment is used for folding the sheets that have been divided.

The **XJDF** **Folding** process covers both operations done in stand-alone **Folding** machinery—typically found when processing printed materials from sheet-fed presses—and in-line equipment of web presses. Creasing and/or slot perforating are sometimes necessary parts of the **Folding** operation that guarantee exact process execution. They depend on the folder used, the ***Media*** and the folding layout. These operations are specified in ***FoldingParams/Crease*** and ***FoldingParams/Perforate*** respectively.

Table 5.93: Folding – Input Resources

NAME	DESCRIPTION
Component	***Component*** resources, including a printed sheet or a pile of sheets, are used in the **Folding** process.
FoldingParams	Specific parameters to set up the machinery. Any process coordinate transformations that apply to **Folding** SHALL be specified in the respective parent ***Resource***/@*Orientation* or ***Resource***/@*Transformation*.
Generic Input Resources*	See Table 5.1 Generic Input ResourceSets for additional input resources that are valid for all process types.

Table 5.94: Folding – Output Resources

NAME	DESCRIPTION
Component	The process produces a ***Component***, which in most cases is a folded sheet.

Figure 3: Screen of the Folding Process definition in the XJDF Specification 2.1 (CIP4, 2020 a, p. 91)

Figure 3 illustrates how, for instance, a process (here: Folding Process) is specified in the XJDF Specification (CIP4, 2020 a, p. 91). Folding means applying folds on a printed sheet (see Figure 4). In the context of JDF/XJDF a process is defined as an individual workflow step “assumed to be executed by a single purpose device. Some [...] devices are able to combine the functionality of multiple single-purpose devices and execute more than one process type” (CIP4, 2020 a, p. 58). CIP4 has specified the scope of a process and all processes that might occur in a print production process. A standardized name defines each process and an accurate description (see Figure 3). As processes are designed to be organized in a workflow, each process consumes input resources to produce output resources. The appropriate resources are also part of the process specification. The Folding process, for instance, consumes obligatory Component- and FoldingParams resources and may receive optional generic input resources. The output of the Folding process is a Component resource again (see Figure 3). The specifications define not only the processes but also the resources. A “Component is used to describe the various versions of semi-finished goods in the press and postpress area, such as a pile of folded sheets that have been collected and are then to be joined and trimmed” (CIP4, 2020 a, p. 145). Domain experts and developers can use these process and resource definitions immediately for a domain-specific language.

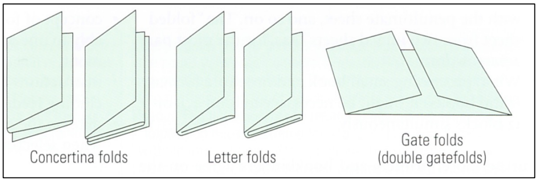
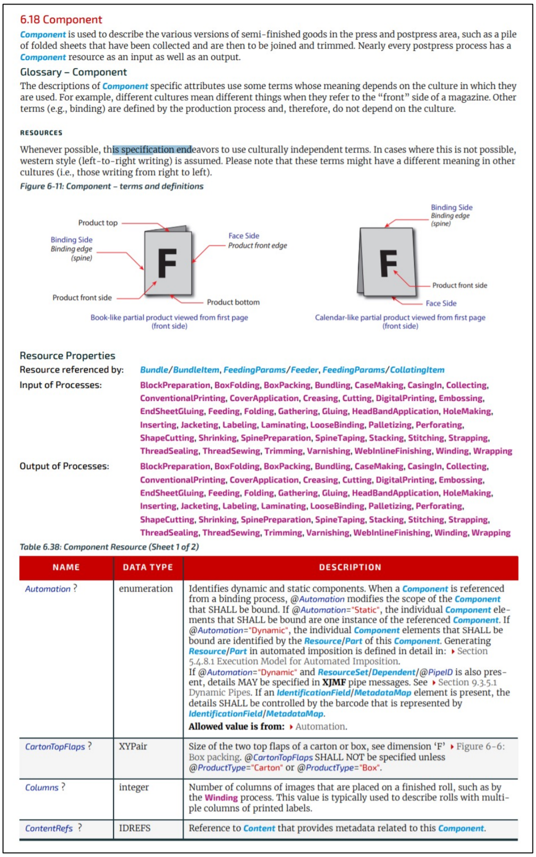


Figure 4: Examples of folded sheets (Kipphan, 2001, p. 809)



6.18 Component

Component is used to describe the various versions of semi-finished goods in the press and postpress area, such as a pile of folded sheets that have been collected and are then to be joined and trimmed. Nearly every postpress process has a Component resource as an input as well as an output.

Glossary – Component

The descriptions of Component specific attributes use some terms whose meaning depends on the culture in which they are used. For example, different cultures mean different things when they refer to the "front" side of a magazine. Other terms (e.g., binding) are defined by the production process and, therefore, do not depend on the culture.

RESOURCES

Whenever possible, this specification endeavors to use culturally independent terms. In cases where this is not possible, western style (left-to-right writing) is assumed. Please note that these terms might have a different meaning in other cultures (i.e., those writing from right to left).

Figure 6-11: Component – terms and definitions

Product top
Binding Side
Binding edge (spine)
Face Side
Product front edge
Product front side
Product bottom
Book-like partial product viewed from first page (front side)

Binding Side
Binding edge (spine)
Product front side
Face Side
Calendar-like partial product viewed from first page (front side)

Resource Properties

Resource referenced by: Bundle/BundleItem, FeedingParams/Feeder, FeedingParams/CollatingItem

Input of Processes: BlockPreparation, BoxFolding, BoxPacking, Bundling, CaseMaking, CasingIn, Collecting, ConventionalPrinting, CoverApplication, Creasing, Cutting, DigitalPrinting, Embossing, EndSheetGluing, Feeding, Folding, Gathering, Gluing, HeadBandApplication, HoleMaking, Inserting, Jacketing, Labeling, Laminating, LooseBinding, Palletizing, Perforating, ShapeCutting, Shrinking, SpinePreparation, SpineTaping, Stacking, Stitching, Strapping, ThreadSealing, ThreadSewing, Trimming, Varnishing, WebInlineFinishing, Winding, Wrapping

Output of Processes: BlockPreparation, BoxFolding, BoxPacking, Bundling, CaseMaking, CasingIn, Collecting, ConventionalPrinting, CoverApplication, Creasing, Cutting, DigitalPrinting, Embossing, EndSheetGluing, Feeding, Folding, Gathering, Gluing, HeadBandApplication, HoleMaking, Inserting, Jacketing, Labeling, Laminating, LooseBinding, Palletizing, Perforating, ShapeCutting, Shrinking, SpinePreparation, SpineTaping, Stacking, Stitching, Strapping, ThreadSealing, ThreadSewing, Trimming, Varnishing, WebInlineFinishing, Winding, Wrapping

Table 6.38: Component Resource (Sheet 1 of 2)

NAME	DATA TYPE	DESCRIPTION
Automation ?	enumeration	Identifies dynamic and static components. When a Component is referenced from a binding process, @Automation modifies the scope of the Component that SHALL be bound. If @Automation="Static", the individual Component elements that SHALL be bound are one instance of the referenced Component. If @Automation="Dynamic", the individual Component elements that SHALL be bound are identified by the Resource/Part of this Component. Generating Resource/Part in automated imposition is defined in detail in: Section 5.4.8.1 Execution Model for Automated Imposition. If @Automation="Dynamic" and ResourceSet/Dependent/@PipeID is also present, details MAY be specified in **XJMF** pipe messages. See Section 9.3.5.1 Dynamic Pipes. If an IdentificationField/MetadataMap element is present, the details SHALL be controlled by the barcode that is represented by IdentificationField/MetadataMap. **Allowed value is from:** Automation.
CartonTopFlaps ?	XYPair	Size of the two top flaps of a carton or box, see dimension 'F' Figure 6-6: Box packing. @CartonTopFlaps SHALL NOT be specified unless @ProductType="Carton" or @ProductType="Box".
Columns ?	integer	Number of columns of images that are placed on a finished roll, such as by the Winding process. This value is typically used to describe rolls with multiple columns of printed labels.
ContentRefs ?	IDREFS	Reference to Content that provides metadata related to this Component.

Figure 5: Screen of the Component Definition (Part 1) (CIP4, 2020 a, pp. 145-146)

The JDF/XJDF Specifications are even more profound than just providing a domain-specific language. Besides the definitions of products, processes, and resources, the specifications also define their properties and relations. Figure 4 visualizes the first half of the Component resource specification. The Component specification starts with an explanation of what 'Component' means in the context of JDF/XJDF and the definition of terms used to specify a Component. The second part of the Component specification lists all properties needed to define a JDF/XJDF Component organized as attributes and subelements. Each property is represented by a table row specifying the standardized name, the data type, and a detailed description of how to use this property. Each property's cardinality is expressed using a simple Extended Backus-Naur Form notation ('', '?', '*', '+') right after its name (CIP4, 2020 a, pp. 7-8). The properties tables, in general, defines how all entities are related to each other and so describe a standard data model for the printing industry. West (2003, p. 2) defines the term 'data model' as the definition of structure and meaning of the data. West (2003, p. 7) further states that a good data model supports businesses in reducing risk and costs in software projects and achieving new business opportunities, increased effectiveness, and higher responsiveness to changes (see Figure 5). The Princeton University (2021) concurs with West's definition: The university determines that data models organize data elements and standardize how they relate to one another. In general, data models describe the data itself but not how to encode it.

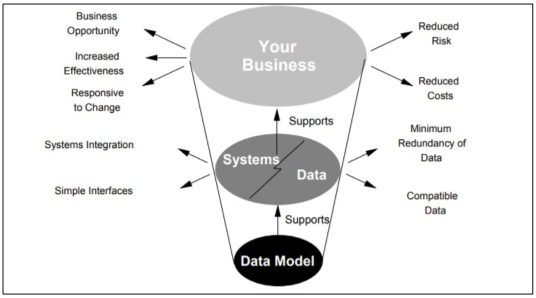


Figure 6: How data model delivers benefit (West, 2003, p. 7)

The JDF/XJDF Specification uses the Extensible Markup Language (XML) to encode its data model (CIP4, 2020 a, p. 4). XML was first standardized in 1998 (W3C, 1998) and has been designed in documents (called 'XML Documents') made up of entities as storage units and markup to describe the document's storage layout and logical structure (W3C, 2008). W3schools.com (2021) sees XML as "a software- and hardware-independent tool for storing and transporting data." XML was developed for usage in the World Wide Web (W3C, 2008) and was the dominating technology for data exchange between web applications/devices for a long time. Wilde (2020) observed XML's dominance in the past – however, today, he sees JSON (JavaScript Object Notation) instead of becoming the mainstay of Application Programming Interface (API) technologies. Google Trends (2021) acknowledge Wilde's observation regarding the emerge of JSON as a major API technology (see Figure 6). The International Organization for Standardization (2017) defines JSON as a lightweight, text-based, language-independent syntax for describing data interchange formats. Joshi (2017) poses JSON to be less verbose, faster, and more readable than XML. CIP4 (2021) is currently working on a JSON representation of JDF/XJDF to make its industry standards more compliant with common cloud standards.

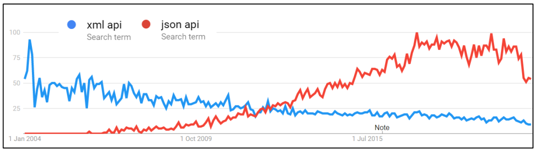


Figure 7: Comparison of search terms 'xml api' and 'json api' in Google Trends (2021)

Conclusion

The printing industry is moving towards Industry 4.0. The German printing association BVDM, for instance, has established the term 'Print 4.0' as an analogy to 'Industry 4.0' in 2015 to foster this progression of converting traditional printing houses to 'Smart Print Factories.' Since then, printing houses, system- and device vendors have started to implement the Industry 4.0 concepts and techniques into their systems. One example of a printing house working actively on this progression is the 'media print solutions' (2021) who built a smart factory from a greenfield perspective next to the traditional shopfloor. The transition to 'Smart Print Factories' is accelerated by the system- and device vendors in the printing industry. Canon (2018) describes this progression as "what was once a craft-based business is now moving to a widely connected and process-driven industry that is paving the way for Industry 4.0". Canon (2018) further sees its mission in ensure its customers perform in the most efficient manner possible. As integration is a fundamental part of Industry 4.0, interfaces and open standards such as PDF, JDF/XJDF are playing a significant role in that field – which conforms to Belz et al. (2018, p. 25) observation. Hoffmann-Walbeck (2018, p. 9) wrote about JDF that it anticipated the vision of Industry 4.0 in parts. All in all, there can be recognized an active development towards Industry 4.0 in the printing industry.

When implementing Industry 4.0 in a (print-) production line, the Internet of Things (IoT) is undoubtedly one of the first steps required. IoT's task is to acquire (real-time) data from the shop floor, to build a 'digital replica,' representing the foundation of many Industry 4.0 applications. IoT in the context of a production line is not limited to the Industrial Internet of Things (IIoT) and includes the Consumer Internet of Things (CIoT). In the first step, of course, the integration of the production equipment using IIoT has priority. The key of IIoT is to identify, use, and extend existing (standard) interfaces that might be designed a long time before the idea of Industry 4.0 and IoT has emerged. Sisinni et al. (2018, p. 4725) called this development approach 'evolution,' as legacy technologies needed to be integrated seamlessly in today's cloud infrastructure. Once the production equipment metrics are captured using IIoT, the 'digital replica' should be extended by environmental parameters such as weather, humidity, temperature, energy consumption, etc. Companies can efficiently track these parameters by installing arbitrary CIoT devices such as smart sensors (BehrTech, 2020) into the shop floor. Connecting public smart sensors can also measure environmental parameters such as the outside temperature and air pressure. One provider of weather data is, for instance, OpenWeather (2021). IoT's key is to build a highly detailed 'digital replica' of the production line and environment. The better the 'digital replica,' the better the Industry 4.0 applications built hereon. To make enterprises achieving the best results from their 'digital replica,' they need to find a way to use IIoT- and CIoT technologies seamlessly hand in hand.

One major IIoT Technology in the printing industry is JDF/JMF or XJDF/XJMF. JDF/JMF has its roots in the 1990'ties - a long time before the idea of Industry 4.0 appeared. In the 1990'ties, even the internet itself was in a very early stage. When the structure of JDF/JMF has been designed, the focus was on connectivity and integration within a single production site. Neither communication between production sites nor the connectivity with a third-party unit like 'The Cloud' was in the scope. The first glimpse of widening the scope by connecting companies within the Graphic Arts Industry came up with the PrintTalk Technical Briefing Paper (PrintTalk, 2000) in 2000. JDF/JMF might appear awkward or obsolete in today's world. Nevertheless, JDF/JMF is an open interface technology implemented by many active applications and devices. Beyond, the CIP4 Specifications can be seen as an integrated global Domain Specific Language (DSL) and data model for the printing industry – the fundamental pieces in globalization times. Things making JDF/JMF appearing obsolete are probably its internal structure and its encoding as Extensible Markup Language (XML) rather than JavaScript Object Notation (JSON). JDF/JMF's internal structure has been optimized significantly during the major release update to XJDF/XJMF (CIP4, 2020 a, pp. 1-6) – the XML encoding is undergoing a revision by a newly established Cloud Workgroup (CIP4, 2021), currently working on a JSON representation of JDF/XJDF. The CIP4 standards are underlying an active evolutionary development process described by Sisinni et al. (2018, p. 4725) to increase today's cloud technologies' compliance.

Once IoT devices are in place, connected, and produce data to feed the digital replica, the actual work begins: Data Mining. Data Mining is the approach to discover patterns and models in the digital replica context (Wang & Wang, 2015, p. 143). Next to the equipment performance data acquired using JMF/XJMF, tracking the interior climate (temperature/humidity) in a print production site could be an insightful measurement series of the digital replica. Studies as "Climate, Paper, and Print" (Falter, 1998) elaborated on the influence of climate on shop floors productivity and quality. Using the digital replica, someone can check the current climate on the shop floor and how it behaves over time. Having additional sensors indicating open doors and tracking the outside weather conditions might reveal the effect of open doors on temperature/humidity over the year. A next enlightening experience could be the correlation between the shop floor climate and the presses' performance. Is there a level where the number of paper-run issues increases, or is there even a minimum? Such a minimum value would result in the ideal (individual) climate condition for a specific production floor. Once a target climate condition is defined, the digital replica can be used to measure and sustain it constantly. This was just one first example of how Industry 4.0 might influence production in future. Other Industry 4.0 topics include predictive maintenance or automated production planning. Industry 4.0 is still in its beginnings in the printing industry.

References

1. Adobe. (2007). PDF Reference, sixth edition: Adobe Portable Document Format 1.7. Adobe Systems Incorporated.

2. Alasadi, S., & Bhaya, W. (2017). Review of Data Preprocessing Techniques in Data Mining. Journal of Engineering and Applied Sciences 12 (16), 4102-4107.

3. Ashton, K. (2016, March 18). Beginning the Internet of Things. Retrieved from https://medium.com/@kevin_ashton/beginning-the-internet-of-things-6d5ab6178801

4. Bandyopadhyay, D., & Sen, J. (2011). Internet of Things: Applications and Challenges. Wireless Pers Commun, 49-69.

5. BehrTech. (2020). Top 10 IoT Sensor Types. Retrieved October 27, 2020, from https://behrtech.com/blog/top-10-iot-sensor-types/

6. Belz, H., Zacharias, W., & Beinhauer, W. (2018). Druckindustrie 4.0 - Impulse zur Strategieentwicklung. Berlin: Bundesverband Druck und Medien.

7. Breivold, H. P., & Sandström, K. (2015). Internet of Things for Industrial Automation – Challenges and Technical. IEEE International Conference on Data Science and Data Intensive Systems, 532-539.

8. Bryman, & Bell. (2015). Business Research Methods. Oxford: Oxford University Press.

9. Canon. (2018, December). INDUSTRY 4.0: THE SMART PRINT FACTORY. Retrieved from https://pps.csa.canon.com/newsletter-article/book-printers-publishers/2018-12-01-000000/industry-40-smart-print-factory2018-12

10. CIP4. (2009). JDF Specification Release 1.4a. Zurich: CIP4 Organization.

11. CIP4. (2020 a). XJDF Specification 2.1. Zurich: CIP4 Organization.

12. CIP4. (2020 b). JDF Specificaion 1.7. Zurich: CIP4 Organization.

13. CIP4. (2020 c). Who is CIP4. Retrieved October 01, 2020, from https://cip4.org/who-is-cip4.html

14. CIP4. (2020 d). The JDF Integration Matrix. Retrieved October 26, 2020, from https://cip4.org/jdfwork/jdf-integration-matrix/jdf-matrix-application.html

15. CIP4. (2021). JSON mapping for XJDF datatypes. Retrieved January 20, 2021, from https://confluence.cip4.org/display/WJT/JSON+mapping+for+XJDF+-datatypes

16. Druckmarkt. (2002). Wie flexibel können Druckmaschinen sein? Druckmarkt, pp. 109-111.

17. EFI. (2020). Fiery JDF. Retrieved October 1, 2020, from https://www.efi.com/products/fiery-servers-and-software/fiery-integration/fiery-jdf/overview/

18. Falter, K.-A. (1998, September). Klima, Papier, und Druck. Fogra Praxis Report 50.

19. Fisher, C. (2010). Researching and Writing a Dissertation. Harlow: Pearson Education Limited.

20. General Electric. (2016, November 18). Minds + Machines: Meet A Digital Twin. Retrieved from https://www.youtube.com/watch?v=2dCz3oL2rTw

21. Google Scholar. (2020, December 1). Google Scholar. Retrieved from https://scholar.google.com/

22. Google Trends. (2021, January 20). Google Trends. Retrieved from https://trends.google.com/

23. Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques. Waltham: Elsevier.

24. Heidelberger Druckmaschinen AG. (2012). Prinect Integration Manager. Heidelberg: Heidelberger Druckmaschinen AG. Retrieved October 1, 2020, from https://www.heidelberg.com/global/media/en/global_media/products___prinect_modules/pdf/prinect_integration_manager.pdf

25. Heidelberger Druckmaschinen AG. (2017). User's Guide. Prinect DFE 2017. Heidelberg. Retrieved from https://onlinehelp.prinect-lounge.com/App/App_pdfs/pdfs/DFE_Workflow_2017_en.pdf

26. Hoffmann-Walbeck, T. (2018). Smart Factory: JDF and XJDF. Journal of Graphic Engineering and Design, Volume 9 (1), 5-9.

27. HP. (2015). HP Integrated Digital Printing Worker JDF Developer's Guide Version 6.0.1. HP Development Company.

28. HP. (2020). Retrieved October 1, 2020, from Commercial Printing SDKs: https://developers.hp.com/smart-stream-sdks/commercial-printing-sdks

29. IETF. (2017, March). RFC 8118 - The application/pdf Media Type. (Internet Engineering Task Force) Retrieved from https://tools.ietf.org/html/rfc8118

30. INFSO. (2008). Internet of Things in 2020.

31. International Organization for Standardization. (2017). Information technology — The JSON data interchange syntax (ISO/IEC 21778:2017). Zurich: International Oranization for Standardization.

32. Jazdi, N. (2014). Cyber Physical Systems in the Context of Industry 4.0. IEEE.

33. Joshi, V. (2017, January 18). WHY YOU SHOULD BE USING JSON VS XML. Retrieved from https://blog.cloud-elements.com/author/vineet-joshi

34. Kipphan, H. (2001). Handbook of Print Media. Berlin: Springer Verlag.

35. Kleeberg, D. (2019, October). PRINT 4.0, KÜNSTLICHE INTELLIGENZ UND DEREN UMSETZUNG. Druckmarkt, pp. 42-49. Retrieved from https://www.druckmarkt.com/archiv/pdf/123/42_Print_4.0.pdf

36. Kodak. (2021). Kodak Workflow Documentation. Retrieved January 4, 2021, from https://workflowhelp.kodak.com/display/PRIN75/Workflow

37. Koenig & Bauer. (2020). Workflow Solutions. Retrieved October 1, 2020, from https://www.koenig-bauer.com/en/service/sheetfed/productivity-service-complete/workflow-solutions/

38. Kosar, T., Lopez, P., Barrientos, P., & Mernik, M. (2008). A preliminary study on various implementation approaches of domain-specific language. Information and Software Technology 50 , 390–405.

39. Lee, E. (2008). Cyber Physical Systems: Design Challenges. 11th IEEE Symposium on Object Oriented Real-Time Distributed Computing (ISORC), 363-369.

40. Lee, J., Kao, H., & Yang, S. (2014). Service innovation and smart analytics for Industry 4.0 and big data environment. Product Services Systems and Value Creation. Proceedings of the 6th CIRP Conference on Industrial Product-Service Systems, 3-8.

41. Leskovec, Rajaraman, & Ullmann. (2014). Mining of Massive Datasets. Standford University. Retrieved from http://infolab.stanford.edu/~ullman/mmds/book.pdf

42. Managoli, G. (2020, February 24). What developers need to know about domain-specific languages. opensource.com. Retrieved from https://opensource.com/article/20/2/domain-specific-languages

43. Mauro, Greco, & Grimaldi. (2016). A formal definition of Big Data based on its essential features. Emerald Inisght Library Review Volume 65 Issue 3.

44. mediaprint solutions GmbH. (2021). media print smart factory. Retrieved from https://mediaprint.de/mediaprint-smart-factory/

45. Meissner, S. (2019). XJDF - Standardisierter Datenaustausch ZWISCHEN Unternehmen. Retrieved from https://www.slideshare.net/stefanmeissner/xjdf-datenaustausch-zwischen-unternehmen

46. Mourtzis, D., Vlachou, E., & Milas, N. (2016). Industrial Big Data as a result of IoT adoption in Manufacturing. 5th CIRP Global Web Conference Research and Innovation for Future Production, 290-295.

47. NITRD. (2012). The Networking and Information Technology Research and Development (NITRD) Program 2012 Stratic Plan. National Science and Technology Council.

48. OpenWeather. (2021). Weather API. Retrieved from https://openweathermap.org/api

49. Princeton University. (2021). What is a Data Model? Retrieved January 20, 2021, from https://cedar.princeton.edu/understanding-data/what-data-model

50. PrintTalk. (2000, September 24). PrintTalk Technical Briefing Paper. Reston VA: PrintTalk Consortium. Retrieved from https://web.archive.org/web/20030309073938/http://www.printtalk.org/briefpaper04.pdf

51. Qi, Q., & Tao, F. (2018). Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. IEEE Access, 3585-3593.

52. Ratner, B. (2017). Statistical and Machine-Learning Data Mining. Boca Raton FL: CRC Press.

53. Rod'ko, Badanova, Smirnov, Berestov, Zhukovskii, Evteev, & Shvedunov. (2020). The Concept of a Digital Twin of a Radiotherapy System. Biomedical Engineering, Vol. 53, No. 6, 421-424.

54. Santucci, G. (2009). The Internet of Things The Internet of Things – when your when your. Luxembourg: International Conference on Future International Conference on Future. Retrieved from https://www.ipv6council.lu/docs/G_Santucci.pdf

55. Singh, H. (2020, May 13). Understanding Data Preprocessing. Retrieved January 4, 2021, from https://towardsdatascience.com/data-preprocessing-e2b0bed-4c7fb

56. Sisinni, E., Saifullah, A., Han, S., Jennehag, U., & Gidlund, M. (2018). Industrial Internet of Things: Challenges, Opportunities, and Directions. IEEE Transactions on Industrial Informatics, vol. 14, no. 11, 4724-4734.

57. Snijders, Matzat, & Reips. (2012). "Big Data": Big Gaps of Knowledge in the Field of Internet Science. International Journal of Internet Science, 1-5.

58. talend. (2021). 16 Data Mining Techniques: The Complete List. Retrieved January 4, 2021, from https://www.talend.com/resources/data-mining-techniques/

59. Ternay, Y. d. (2020). THE BUILDING BLOCKS OF PRINT 4.0. inkish.news. Retrieved from https://inkish.news/article/NDU1/the-building-blocks-of-print-4.0

60. W3C. (1998). Extensible Markup Language (XML) 1.0. W3C. Retrieved from https://www.w3.org/TR/1998/REC-xml-19980210

61. W3C. (2008). Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C. Retrieved from https://www.w3.org/TR/2008/REC-xml-20081126

62. w3schools.com. (2021). Introduction to XML. Retrieved January 20, 2021, from https://www.w3schools.com/xml/xml_whatis.asp

63. Wan, J., Tang, S., Shu, Z., Wang, S., Imran, M., & Vasilakos, A. (2016). Software-Defined Industrial Internet of Things in the Context of Industry 4.0. IEEE Sensors Journal, Vol. 16, No. 20, , 7373-7380.

64. Wang, L., & Wang, G. (2015). Data Mining Applications in Big Data. Computer Engineering and Applications Vol. 4, No. 3, 143-152.

65. West, M. (2003). Developing High Quality Data Models. London: Shell International Limited. Retrieved from https://web.archive.org/web/20060718132354/http://www.matthew-west.org.uk/documents/princ03.pdf

66. Wilde, E. (2020, November 11). API Formats: Why JSON won over XML. Retrieved from https://apifriends.com/api-management/why-json-won-over-xml/

67. Wurm, J., Hoang, K., Arias, O., Sadeghi, A., & Y, J. (2016). Security Analysis on Consumer and Industrial IoT Devices. IEEE, 519-524.

68. Xerox. (2009). FreeFlow® JMF Service User Guide. Webster, NY 14580: Xerox.

Stefan Meissner

Stefan Meissner, Dipl.-Ing. (FH)
PhD Candidate TU Dublin

Dublin, Ireland

d06114283@mytudublin.ie