A flexible Security Analytics Service for the Industrial IoT

Philip Empl
Günther Pernul
philip.empl@ur.de
guenther.pernul@ur.de
Chair of Information Systems, University of Regensburg
Regensburg, Germany

ABSTRACT

In cloud computing, the cloud serves as a central data hub for the Industrial Internet of Things' (IIoT) data and is deployed in diverse application fields, e.g., Smart Grid or Smart Manufacturing. Therefore, the aggregated and contextualized data is bundled in a central data hub, bringing tremendous cybersecurity advantages. Given the threat landscape in IIoT systems, especially SMEs (Small and medium-sized enterprises) need to be prepared regarding their cybersecurity, react quickly, and strengthen their overall cybersecurity. For instance, with the application of machine learning algorithms, security-related data can be analyzed predictively in order to be able to ward off a potential attack at an early stage. Since modern reference architectures for IIoT systems, such as RAMI 4.0 or IIRA, consider cybersecurity approaches on a high level and SMEs lack financial funds and knowledge, this paper conceptualizes a security analytics service used as a security add-on to these reference architectures. Thus, this paper conceptualizes a flexible security analytics service that implements security capabilities with flexible analytical techniques that fit specific SMEs' needs. The security analytics service is also evaluated with a real-world use case.

CCS CONCEPTS

• **Security and privacy** → **Security services**; *Domain-specific security and privacy architectures*; • **Applied computing** → *IT architectures*; *Reference models*; Service-oriented architectures; • **Information systems**;

KEYWORDS

Industrial IoT, Security as a Service, Security Analytics

ACM Reference Format:
Philip Empl and Günther Pernul. 2021. A flexible Security Analytics Service for the Industrial IoT. In *Proceedings of the 2021 ACM Workshop on Secure and Trustworthy Cyber-physical Systems (SAT-CPS '21), April 28, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3445969.3450427

SAT-CPS '21, April 28, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8319-6/21/04…$15.00
https://doi.org/10.1145/3445969.3450427

1 INTRODUCTION

The emerging technologies around the Industrial Internet of Things (IIoT) provide considerable advantages in many application fields. For instance, distributed sensors can be utilized in IIoT systems to optimize processes and escort the industry to Smart Manufacturing's vision by applying machine learning algorithms. Through the coincide of domains, new insights can be generated, such as optimizing a machine's production parameters. The alignment of information technology (IT) and operations technology (OT) requires explicit knowledge in both domains and extensive cybersecurity knowledge. While IT systems are processing tons of OT systems data (e.g., from Industrial Control Systems) in decentralized architectures, new security threats arise, e.g., Sybil attacks [27], Wormhole attacks [10] or Distributed Denial of Service attacks [21]. Small and medium-sized enterprises (SME) are not up to the current state of the art regarding cybersecurity, making them vulnerable to new attack vectors. They often exhibit outstanding knowledge around their supply chain and production processes but cannot compete with the steadily evolving attack vectors. This gap can be traced back to a lack of financial, technical, or even personnel resources for innovation [19] and the necessary knowledge about information security and vulnerabilities [8].

Knowledge is powerful and essential to be resilient against novel attack vectors, such as the Distributed Denial of Service attack. To what extent data and wisdom are related is illustrated in the Data-Information-Knowledge-Wisdom (DIKW) hierarchy by Ackoff [1]. In it, data generates information, information creates knowledge, and knowledge makes wisdom. It might not always be necessary to climb the hierarchy latter to the top, but a response to an incident should be triggered quickly concerning cybersecurity. Thereby, quick reactions build upon wisdom. For example, merely collecting security-related data would not benefit the industry, such as network traffic capture. Only those who process this network data, recognize this information's context and make predictions can become game-changers. In the best case, the systems are sealed off at an early stage to prevent severe damage or ward off the attack altogether. Therefore, security analytics must be established in the industrial sector to tackle the current threat landscape.

Security analytics must necessarily be embedded in an industrial architecture that allows OT to be linked to IT and vice versa. SMEs represent the German industry's backbone and do not view cybersecurity as an integral part of IIoT systems but rather as an add-on. As most of their production systems exist over decades, security by design considerations have changed, especially in this interconnection of IT and OT. Therefore, SMEs offer a vulnerable target for attack. Thus, system functionality is prioritized over cybersecurity

in SMEs, which is partly because many SMEs lack the financial resources and the necessary knowledge.
Nevertheless, SMEs can thereby implement different architectures, e.g., Reference Architectural Model Industry 4.0 (RAMI 4.0) [24] or Industrial Internet Reference Architecture (IIRA) [14]. Due to the lack of standardization, SMEs run the risk of betting on the wrong horse when selecting them or implement their versions. Nonetheless, all of these architectures rely on the application of the cloud. The cloud established itself as a data hub that enables manufacturing enterprises of all sizes to quickly connect their OT systems to cloud-dependent IT systems. The IIoT architectures use middleware layers in general, which can be built on service-oriented architectures [2], enabling the communication between individual heterogeneous systems. A system offers numerous services made available by the cloud service provider (CSP) to the cloud service consumer (CSC) and billed to the CSC via subscriptions. Many SMEs align their systems in the cloud because embedded services yield tremendous advantages of compatibility and interoperability. Thus, integrating new services into an existing IT architecture is made comfortable.
This paper conceptualizes a security analytics service that is compatible with well-known reference architectures. As those reference architectures take the right step into standardization, they represent a further obstacle for SMEs as they do not offer concrete recommendations in terms of security analytics. Our security analytics service follows those security principles but offers concrete security capabilities (see DIKW) that can be instantiated as flexible analytical techniques. That means that every security capability might encompass several analytical techniques (e.g., detective security capability is instantiated by Complex Event Processing) that can be bundled or standalone. This flexibility of analytical techniques is significant because IoT architectures differ in their respective industry regarding their application, middleware, and protocols [26], which results in different requirements. In this way, the security analytics service is flexible and generic, which is accomplished by providing analytical capabilities. Thus, flexibility is leading to a straightforward application in the respective industry.

1.1 Contribution

As this paper is addressing cross-domain knowledge to accomplish knowledge gaps of SMEs in cybersecurity, the contribution of our paper is threefold:

- We are introducing a novel and flexible security analytics service, which fits the needs of industries and their use cases. The security analytics service exhibits security capabilities, which can yield a set of analytical techniques.
- We disclose how to generate wisdom from data to bridge the knowledge gap. Thereby, the security analytics service assists in complying with all capabilities in the DIKW hierarchy.
- We are providing multi-disciplinary research that transfers knowledge in different domains and knits cybersecurity to IIoT.

1.2 Structure

This paper is structured as follows. Section 2 provides the necessary background knowledge in reference architectures, Security as a Service (SECaaS), and security analytics. Section 3 determines the preliminaries towards a conceptual service model. Section 4 provides the conceptual model for the security analytics service and defined relevant service characteristics. Section 5 shows the architectural module, which knit the security analytics service to the IoT architecture (IoT system) and vice versa. Section 6 evaluates the architectural module with a real-world use case, and Section 7 provides the conclusion, limitations, and further research.

2 BACKGROUND AND RELATED WORK

This section presents well-known reference architectures for the IIoT and describes SECaaS and security analytics. Besides, the related work section provides knowledge to which this paper links. These fundamentals are essential as a flexible security analytics service for IIoT systems is being developed that operates according to the principles of SECaaS.

2.1 IIoT Reference Architectures

Reference architectures are utilized in the IIoT to provide an architectural construct that builds on existing standards and norms. Therefore, they are the first step towards standardization, as these reference architectures explicitly refer to gaps or problems in current standards. In the industrial context, these include RAMI 4.0 [24] and IIRA [14].
Although those architectures appear different at first glance, both contain similar views of an IIoT architecture. RAMI 4.0 registers six different layers (business, functional, information, communication, integration, and asset) designed to fulfil various architectural tasks. Each layer in this architecture is viewed from a life-cycle and value stream perspective (cf. IEC 62890) and from the hierarchical level s(cf. IEC 62264/IEC 61512). Unlike RAMI 4.0, IIRA defines four viewpoints (implementation, functional, usage, and business) for different stakeholders, which can be seen in terms of the industrial sector and the product's life cycles. Both reference architectures' advantage is their interoperability (see Fig. 1) since the functional domains (cf. functional view) of IIRA can be mapped with the reference architecture RAMI 4.0 [15].
In the following, the layers and their functionalities are explained. According to RAMI 4.0, physical assets (1) involve all physical actors, such as documents, software, and human actors. Besides, the individual components of a cyber-physical system can also be counted as physical assets. The integration layer (2) enriches the aforementioned physical assets with data transmitted in the communication layer (3), e.g., with the ISO/OSI model. In the information layer (4), the transmitted data is getting contextualized and semantically categorized. The function layer (5) contains relevant processes that support the business process layer (6).
As almost every industrial reference architectures is based on middlewares or service-oriented architectures (SOA) [2], the relevant layer for this paper's consideration is the information layer. This layer also defines relevant services and ensures the technical functionality of an IIoT architecture. Additionally, the information layer also

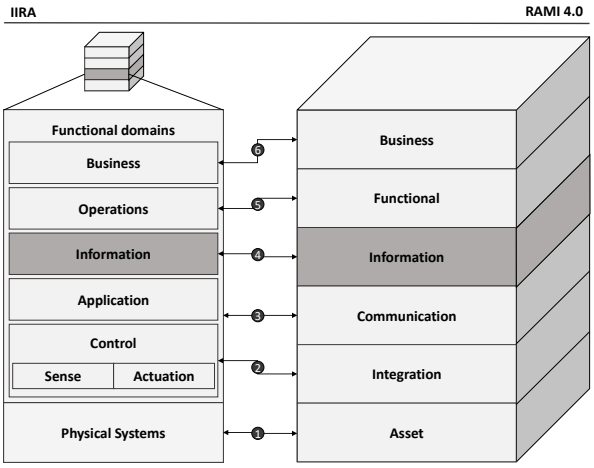

Figure 1: Mapping of functional aspects in both reference architectures.

provides reactions to specific events, such as an actionable to an empty machine magazine.

2.2 Security as a Service (SECaaS)

In cloud computing, there exist four main deployment models (public, private, hybrid, community) and three service delivery models (Infrastructure as a service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS)) [18].

In general, security can be delivered in three alternate models, so-called security service delivery models. These security service delivery models involve 1) on-premises security, 2) managed security, and 3) SECaaS [25]. While on-premise security is deployed on the client-side, managed security involves a CSP via service level agreements and is set up for the enterprise. Both models cannot be aligned with typical cloud computing advantages [25], such as elasticity, broad network access, or pay-as-you-go subscriptions [18].

Besides the service delivery models, there exist plenty of different services. In the SECaaS, the overarching goal is to provide consumers with security services. Security services are wide-ranged and can be categorized as identity and access management, data loss prevention, web security, email security, security assessment, intrusion management, security information and event management (SIEM), encryption, network security and business continuity, disaster recovery, and continuous monitoring [5].

SECaaS benefits from the broad application of cloud computing and thus receives extraordinary encouragement from the scientific community [25]. A meta-model is also proposed by Furfaro et al. [9] to support the modelling of SECaaS in cloud environments. In this meta-model, a distinction is made about whether a service acts in support of another service's security or whether it offers security features. This methodological approach is detailed and allows different perspectives to SECaaS. We will come back to the meta-model when discussing our flexible security analytics service in Section 3 of this paper.

2.3 Security Analytics

As the amount of structured and unstructured security-related data is grown, traditional SIEM systems reach their limits [4]. In the era of big data, new technologies arise which can support efficient real-time data processing and analysis, leading to the paradigm of big data analytics [31]. Big data processing technologies are nothing new in the IoT because of the vast amount of data generated by heterogeneous IoT devices. The application of big data processing technologies, e.g., Apache Spark [33], is therefore not far distant to support security. In this context, (big data) security analytics is used to analyze contextual data in cybersecurity [16]. In general, analytical techniques exist that fulfil single steps of the DIKW hierarchy by Ackoff [1] in the context of IoT [28]. Therefore, these analytical techniques can be classified as descriptive, diagnostic, detective, predictive, and prescriptive. In this paper, we refer to those categories as security capabilities. While descriptive and diagnostic capabilities are strictly retrospective, a detective capability reflects interactions in real-time, and predictive and prescriptive capabilities are prospective.

2.4 Related Work

The provisioning of security services in on-premise solutions (SECaaS) was already carefully elaborated in past research, e.g. Zarca et al. [20] focuses on appropriate methods for enhancing security in IIoT by centrally orchestrating security mechanisms and controllers. Therein, modules interact with each other to satisfy a quick reaction and ease of control in the IoT landscape. However, these approaches omit an integral component of cybersecurity: the generation and exchange of knowledge. Ackoff [1] defined the first approach to gather wisdom from data by providing the DIKW-hierarchy. Decades later, the IoT evolves, and authors rethink the existing DIKW by defining new layers appropriate for the IoT [11]. With the rise of big data, new opportunities regarding data analytics appeared, and analytical methods are categorized along with their outcome, e.g. predictive analytics [28]. Based on those methods, the cybersecurity state of a system or architecture can be evaluated, e.g., cybersecurity dynamics provides a metric-based approach used to compare several security states of a system [32]. Furthermore, cybersecurity dynamics' objectives provide valuable insights into the interaction between descriptive, predictive and prescriptive capabilities. Necessarily, assessing the current cybersecurity state requires the judgment of a security expert. Likewise, also for the implementation of security solutions. Thus, the exchange of knowledge between experts and novices os crucial. The creation and exchange of knowledge result from collaborations between security experts and novices and must be necessarily supported by novel technological approaches. [3]. Past research focused on the assessment of security analytics methods, the exchange and creation of knowledge in security analytics and concrete instances of security analytics components, e.g. incident response. This research is knitting past research together by meaningfully aggregating and assembling their findings.

3 PRELIMINARIES

This paper approaches a flexible security analytics service concerning the meta-model described by Furfaro et al. [9]. This meta-model involves three phases, which are namely security services identification (SSI), design solutions definition (DSD), and design solutions analysis (DSA). The SSI-phase needs a definition of the security delivery model and the security requirements, which leads to security service conceptual models (SSCMs). With the exploitation of those SSCM, security service design solutions (SSDSs) are approached in the DSD-phase, evaluated, and selected in the DSA-phase, leading to an evaluated selected design solution (SDS).
As this security analytics service offers security capabilities with flexible analytical techniques, the delivery scenario is considered a standalone service, which does not provide security services to existing services. The next step towards SSCMs is to define appropriate security requirements. Those requirements are described in the following few paragraphs by determining the involved entities and defining the threat model. Afterwards, the necessary security requirements are derived.

3.1 Involved Entities

The security analytics service should be applied within IIoT systems. Therefore, we define a standard technological setup. According to IIRA or RAMI 4.0, we instantiate an architecture based on a middleware, which yields machines and their edge nodes in an Industrial Control System (ICS) and a cloud to process and analyze the gathered data. Therefore, we derive the following entities that are involved within the architecture, i.e., life cycle parties of machines [7] and Cloud Computing entities: CSP, CSC, network service provider (NSP), internal IT department, manufacturer, distributor, owner, and maintainer.

3.2 Threat Model

Threat modelling methods like Octave, Coras, Mehari, or attack trees lack a holistic methodology as detailed threat lists tend to be incomplete or subjective. In contrast, STRIDE (Spoofing, Tampering, repudiation, information disclosure, denial of service, and elevation of privileges) modelling by Microsoft leads to a classification of threats given constant attack patterns. In contrast, LINDDUN is another threat modelling language that attempts privacy modelling. Since we address a service that is not designed to incorporate privacy-sensitive data, we do not consider privacy leakage and exclude, therefore, LINDDUN. STRIDE ensures a mapping of elements (external entity, processing node, data store, and data flow) to a particular category [22]. Depending on those categories, threats can be identified. For the STRIDE model, a context diagram and a data flow model are required, shown in Fig. 2.
Level-0 defines the context diagram for our threat model, in which all previously described entities are included. These entities interact with a so-called IIoT system, which represents the combination of IT and OT. If we complement this IIoT system with a data flow model (level-1), three central components of the generic IIoT system are defined: machine, edge, and cloud. This data flow diagram could now be abstracted into further levels, e.g., the cloud may contain various information systems in level-2 that communicate with each other.

We assume that the internal IT department of an SME is using resources of a CSP and is therefore partly considered as CSC. Consequently, we assume that the internal IT department is developing somewhat cloud applications for the IIoT system. We consider the cloud instance trustworthy as the cloud provides robust authentication mechanisms for the edge nodes and the internal IT department. Furthermore, we assume the data, which is transferred from the edge node to the cloud, to be encrypted.
Given the assumption, we derive the following threat landscape:

- Spoofing attacks are realistic at the edge node and the machine, as unauthorized entities might have physical access to interact with them.
- Tampering attacks are changing the integrity of data or machine commands. Every entity in the IIoT system can manipulate data flows independent of physical or virtual access.
- Non-repudiation is reached by a false logging-mechanism or missing signature of data streams.
- The elevation of privileges is possible at the machine, the edge node, or the cloud.
- Furthermore, all data streams need to be encrypted and authorized to not reveal sensitive data to possible attackers (entities) and therefore, disclosure sensitive information.
- Denial of Service (DoS) attacks are assumed to be located at the edge node as the cloud provides mechanisms to balance immense loads, e.g., load balancing or horizontal or vertical scalability.

3.3 Security Requirements

In ISO/IEC-27000, there exist several security requirements for an architecture: confidentiality, integrity, and availability, which other requirements like non-repudiation or authenticity can complement. In the threat model mentioned above, various threats are apparent to an IIoT system. This variety of threats can be traced back to the amount of participating entities.
As the operational technology aims instead at a high availability than confidentiality, we consider the availability as the security requirement with the highest priority. The availability must be reached for the OT (e.g., ICS) and the IT, as both components are interconnected. Second, data need to be encrypted regarding confidentiality as the machine's data is classified as confidential. Last, considering the data quality of machines' data, data need to exhibit integrity. When data feature integrity, this also leads to a better quality of security analytics. Security analytics in IIoT systems should aim to preserve a high availability, integrity in its data streams and encrypt the data efficiently. In order to achieve these security requirements, we include the security capabilities mentioned above. In this paper, these capabilities are used for security analytics and referenced as security capabilities. We consider these security capabilities to cover three temporal dimensions (hindsight, insight, and foresight) and allow a holistic view of data. While descriptive and diagnostic capabilities are aimed at hindsight, detective capabilities are for insight and predictive plus prescriptive capabilities for foresight.
Thus, the security capabilities are follows: (i) descriptive, (ii) diagnostic, (iii) detective, (iv) predictive, and (v) prescriptive. We refer

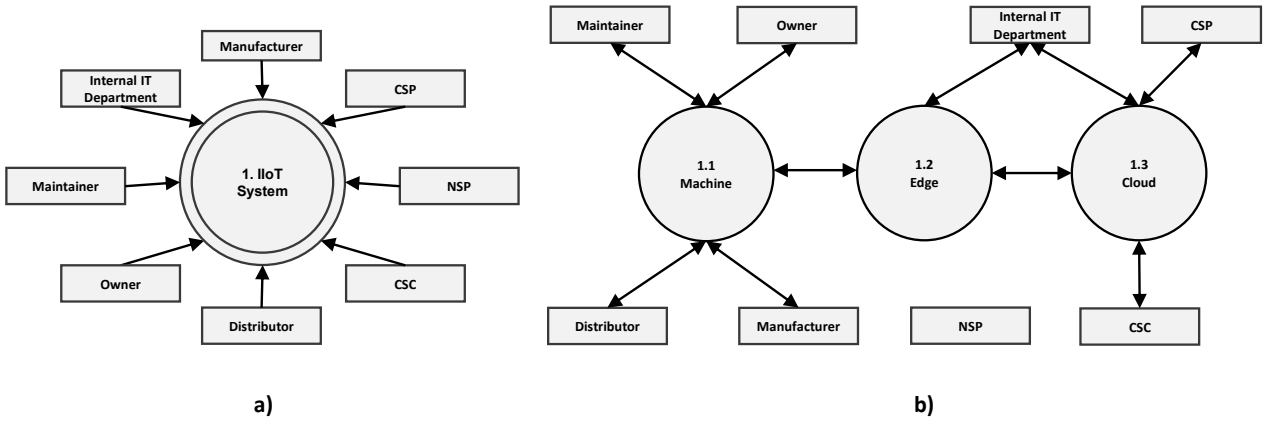

Figure 2: context diagram on level-0 (a) and data flow model on level-1 (b).

to them as service requirements in the service context and security capabilities in security analytics.

3.4 Framing the Problem

The threat model has shown the threat landscape in IIoT systems on level-0 and level 1. The threat landscape's complexity rises by abstracting more detailed threat model levels (e.g., level-2 or level-3). The higher the threat model's level, the more detailed the threat landscape is described, i.e., the more knowledge is required to describe the relationships and dependencies between systems. Based on the arising complexity, extensive cybersecurity knowledge is required to set up even more detailed threat models. Concerning SMEs, they need to accomplish this threat landscape with appropriate knowledge. As knowledge is generated from data and information, SMEs need appropriate security analytics methods to generate wisdom from data.

Common reference architectures like IIRA or RAMI 4.0 provide dedicated security frameworks, e.g., the Industrial Internet Security framework [17]. As can be inferred from frameworks, those security frameworks provide generically functional breakdowns for security components in IIoT systems. For instance, the security component "security monitoring and analysis" is described as three constituents: monitor, analyze, and react. Those components are described with their shared attributes but are not specified in depth nor within the context of a cloud.

This research paper builds upon this gap and describes a security analytics service that follows those security framework proposals and offers relevant security capabilities to provide SMEs with the necessary knowledge. Moreover, these security capabilities may yield different analytical techniques that can be applied flexibly, which means SMEs can swap or bundle them depending on their desired outcome (e.g., proactive reaction or forensics). In summary, the security analytics service offers five security capabilities, and each of them can flexibly yield different analytical techniques, e.g., SMEs can implement Complex Event Processing as one instance of a detective security capability. Thus, this service is generically applicable to different industries. Regarding a particular industry, analytical techniques lead to wisdom, i.e., a better comprehension of potential threats in the IIoT system and, therefore, to a more efficient incident response.

4 SERVICE MODELING

After specifying the service requirements and the security analytics service delivery model, this section approaches the SSCM, possible design solutions, and the selection of an appropriate conceptual model.

4.1 Conceptual Model

First, the meta-model's main attributes are specified, including the description, security concerns, security level, the category, and relevant service attributes. These entries are shown in Tab. 1.To satisfy the security requirements, we have defined security capabilities as service requirements. Thus, all dimensions of security analytics are addressed and included in the security analytics service.

Although the security analytics service addresses all security capabilities, the efficacy depends heavily on the implemented analytical techniques as humans need to implement analytical techniques first. Additionally, the security analytics service faces a broad threat landscape and thus, we conclude the efficacy of detecting possible threats as a possible security concern. Furthermore, the abuse of the rights of a user is another security concern.

SIEM and intrusion management are considered the main categories for the security analytics service because security analytics is a combination of both categories [30]. The security analytics service is applied as an IaaS to simplify analytical techniques to a user or group. Thus, the deployment is straightforward and can be conducted without much effort since most SMEs have outsourced their computing resources to the cloud. Furthermore, the service should be independent of any service providers and is targeting a minimum of service costs.

4.2 Design Solution

The considered SSCM provides five distinct service requirements, which need to be involved within the design solution, i.e., SSDS. A

Table 1: Conceptual model of the security analytics service.

Meta model concept		Description
SECaaS	Service requirements	*SR1: descriptive security capability* *SR2: diagnostic security capability* *SR3: detective security capability* *SR4: predictive security capability* *SR5: prescriptive security capability*
	Security service	*This paper is facing a flexible security analytics service for IIoT systems by satisfying the service requirements SR1-SR5.*
	Security concern	*Efficacy, misuse of rights*
	Security level	*This service's primary focus is the generic application to instantiate analytical techniques dependent on the desired outcome.*
Category		*The security analytics service is categorized as Security Information and Event Management & Intrusion Management.*
Service	Delivery modality	*Infrastructure as a Service*
	Deployment modality	*Public, private*
	Service provider	*Any*
	Cost of service	*Target: minimal/ zero additional infrastructure*
	Period of validity	*Not specified*

possible design solution needs to integrate these service requirements and thus, needs an opportunity to allow multiple analytical techniques as instances of one security capability. As these service requirements are holistically addressed, the security requirements are satisfied. Based on these findings, we instantiate one design solution. The design solution is represented by policies and security mechanisms (SM) that fulfil the SSCM (see Tab. 2).

The SSCM under design faces strong security concerns regarding its efficacy as SMEs lack knowledge in analytical techniques, and the given threat landscape is increasing. Security novices cannot implement an analytical technique without any knowledge about the desired outcome. Thus the efficacy of the analytical technique is questionable.

This problem can be minimized with the usage of a peer-reviewed repository. Repositories have established themselves in many areas as they offer peer-reviewed functional patterns, which can be adapted by the community and tailored to one's needs. Thus, we consider, similar to [13], a repository that yields analytical techniques. Of course, repositories can also be managed internally (private), but we explicitly refer to public repositories. Public repositories' usage fills the knowledge gap between SMEs and larger companies and leads to high-quality analytical techniques. Moreover, this fosters the exchange of knowledge and enables a thorough elaboration of a system's current cybersecurity state.

The security analytics service is secured by a username and password authentication schema. A user or the user group needs to sign in to configure and adapt the analytical techniques stored within a public repository.

As we have only derived one design solution, this design solution is also our selected solution for further considerations. In the next section, we are sketching this solution with appropriate architectural modules.

Table 2: Design solution of the security analytics service.

Meta model concept			Description
Policy	Identifier		*Id-01*
	Description		*The security analytics service binds a public repository, i.e. security analytics repository, which yields descriptive information about specific analytical techniques to satisfy a security requirement.*
	Statement	Action	*The user can select or define analytical techniques, which are dependent on the desired security capability. These analytical techniques are stored in a public repository and can be accessed by other SMEs.*
		Effect	*The selected analytical technique can quickly be adopted and adjusted to fit the security requirements of the IIoT system.*
	User		*Group/ single*
SM	Authentication time		*Disable*
	Authentication location		*Disable*
	Credential based		*Username and strong password*

5 ARCHITECTURAL MODULES

A design solution is selected in the previous section that addresses the service requirements for security analytics service. Based on this design, the anchoring in IIoT systems and the relevant modules are presented in this section. In general, security analytics should be included in OT and IT concerning reference architectures. Zarca et al. [20] provide, therefore, a security orchestration plane that enables a holistic orchestration of security task over the IIoT system. This conception is also underlining that security analytics is often perceived as a monitor in such architectures [29]. Thus, we align our security analytics service holistically over the IIoT system and expand the monitoring functionality to respective security modules and analytical techniques. The result is presented by Fig. 3 and the modules are described in the following beginning at the bottom.

5.1 IIoT System

The IIoT system in Fig. 3 reflects all IT and OT domains and their interplays. Based on the IIRA, we derived the five functional domains and their interplays. The control domain is located at the edge, the information and operation domain in the cloud, and the in-house server's business and application domain. We explain the single domains shortly as they have been described in the background

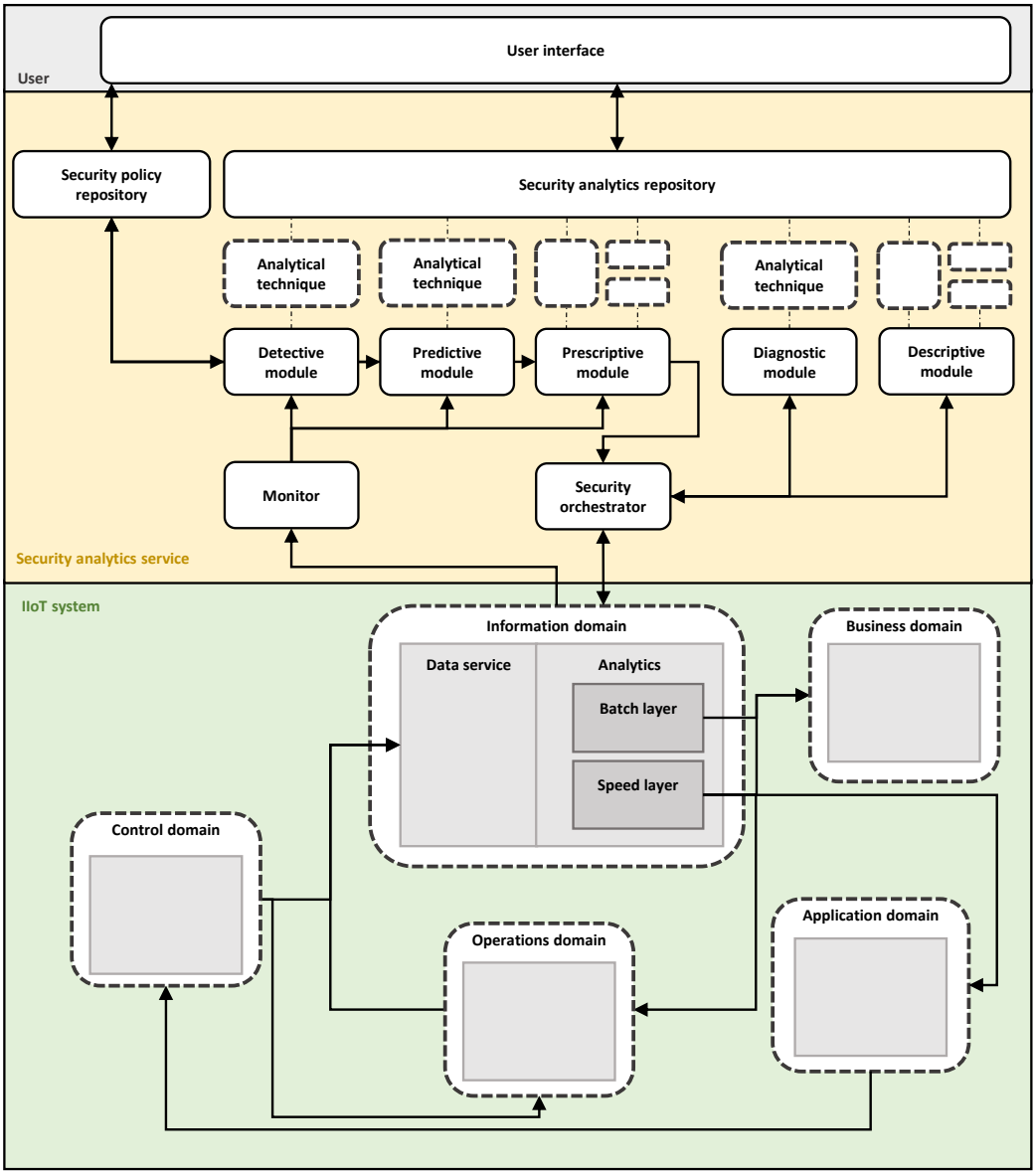

Figure 3: Architectural modules for the security analytics service in IIoT systems.

section.

The control domain is aggregating sensor and machine data and sends the data to the information domain. The information domain is transforming and cleaning the data and accelerates the information based on the lambda architecture by Kiran et al. [12] in a speed layer. The business domain visualizes the data and connects it to various databases, e.g., Manufacturing Executive System (MES) or Enterprise Resource Planning (ERP). Depending on the business logic, the application domain is providing feedback to the control domain. The operation domain is responsible for the monitoring, optimization, and forecasting of machines' states. The operation domain also provides feedback to the control domain to adjust the machines' parameter. Data and information are fed into the information domain to log the communication between domains and analyze their behaviour.

5.2 Security Analytics Service

As also can be inferred from Fig. 3, the security analytics service is connected to the information domain and thus, connected to the databases and the streaming data. This security analytics service needs strong authentication regarding the streaming data and the databases. We omitted the interconnections between the security

analytics service and the control domain, as this is not the main focus of this paper. The previously defined service requirements SR1-SR5 are applied as security modules (with security capabilities) in the security analytics service. Those security modules are arranged according to the DIKW hierarchy, i.e., they are nearly ordered from the data (right) to the wisdom (left). In the following, the single parts of the security analytics service are described.

Security orchestrator. The security orchestrator is the bridge to the IIoT system and orchestrates and automates the response to a possible threat. The data is partly forwarded to the monitor (not depicted), but it should also be achievable to control networks from the control domain with the aid of higher-level software-defined network controllers. The security orchestrator logically receives feedback from the predictive, detective, and prescriptive security modules to derive actions. We provided only one connection (prescriptive module to security orchestrator) to enhance clarity.

Monitor. The monitor is also proposed in the recommendations for improving cybersecurity in the Industrial Internet Security Framework, among others. The monitor aggregates data from all sources and captures incoming data in real-time. The amount of data makes it possible to monitor systems and pass them to the individual security modules. The data sources are not limited to the information domain; e.g., databases from the business domain can also be connected to gather information about ERP or MES.

Descriptive module. The descriptive analytics module provides possibilities for finding out what happened in a system, thereby providing hindsight. Mostly data is summarized or presented in raw data format. Common techniques are visual analytics or business analytics, where the data or information is presented graphically.

Diagnostic module. The diagnostic module also creates hindsight but goes beyond descriptive analysis and investigates the cause of a particular discrepancy or anomaly. For instance, within digital forensics, security experts can investigate data sources and causalities.

Detective module. The detective module provides a statement about what is currently going on in the systems, and thus, it provides insights. Therefore, the analyses are only related to the current systems state and are performed based on the monitor's data. Detection mechanisms are rule-based, time-based, or anomaly-based to detect incidents in real-time and support decision-making. However, correlations or clustering of information can also be utilized to detect inconsistencies.

Predictive module. In the predictive module, future system states are predicted based on the real-time data received. If irregularities are discovered in the data, potential incidents might be averted at an early stage. This module includes analytical techniques such as regressions or fuzzy logic.

Prescriptive module. The prescriptive module, in comparison to the predictive module, also predicts states of a system. Nevertheless, the prescriptive module goes one step further and, unlike pure forecasting, also derives recommendations for actions. Thus, possible actions can be determined for the incident response process, such as the systems' encapsulation. Furthermore, recommender systems or Monte Carlo simulations are considered analytical techniques.

Analytical technique. An analytical technique is understood as an instance of a particular security module (security capability). As mentioned in the individual security modules, concrete algorithms or techniques are applied that are loose or coupled. Moreover, analytical pipelines can be created by bundling multiple analytical techniques together. The analytical techniques are considered security modules and are originally located in the security analytics repository.

Security policy repository. Various policies can be defined in the security policy repository, classified according to different security levels (e.g., low, medium, or high). The security policy repository offers two interfaces: one to the user and another to the detective, predictive and prescriptive modules. By applying these security policies, decisions are derived for the incident response.

Security analytics repository. One main part of the security analytics service is the security analytics repository. Instances of security modules (analytical techniques) are stored in this repository and shared within and across a companies' borders. This information sharing enables a knowledge transfer and peer review of the instantiated analytical techniques.

5.3 Security Expert

As described in the last section, a user needs to communicate with the security analytics service. As this interaction dives deep into the subject, the user needs to exhibit outstanding domain knowledge. Therefore, we are considering the user as a security expert. We manifest a user interface that enables interaction with the security analytics repository and the security policy repository (see Fig. 3). A security expert first needs to sign in via a password-based authentication schema. In the next steps, the security expert can define or edit security policies for the IIoT system or create new analytical techniques for a particular security module.

The definition of security policies presents itself as a manageable task for a security expert. However, the implementation and application of analytical techniques in terms of cybersecurity is a mammoth task. Here, the security expert must have in-depth knowledge from various domains, namely cybersecurity, software development, and statistics.

6 EVALUATION

As this paper does not implement the concept, technical experiments can be excluded. However, this paper is dedicated to SMEs, and we will evaluate the concept with a real-world use case. We refer to SISSeC[1], an ongoing German research project.

SISSeC aims to securely link a printed circuit board (PCB) manufacturer's machine and sensor data via an edge gateway to a so-called sensor cloud. Within the edge gateway, intrusion detection mechanisms are applied, and access controls mechanisms authenticate machines and sensors. The sensor cloud collects and aggregates this data to determine the potential security state with security analytics (e.g., via a digital twin). The PCB manufacturer's overall objective is to collect relevant data and identify likely future states. In addition to the cybersecurity focus, such events are also production-related, e.g., the potential wear of a drill bit on the drilling and milling machine and the time remaining until its change.

Since the project itself is attempting a particular architecture, no

[1] Secure Industrial Semantic SEnsor Cloud

reference architectures are utilized. The main architectural building blocks of SISSeC (machine, edge, and cloud) can nevertheless be categorically classified in the architectural modules identified above. The control domain thus represents the machines, sensors, and the edge gateway. The sensor cloud in SISSeC is perceived as the operations domain and the information domain. The traditional IT systems (e.g., ERP) of the PCB manufacturer are understood as the business and application domains. In summary, the IIoT system is fully covered by the architectural modules.

The information domain is realized using various big data technologies (e.g., Apache Kafka) and equipped with two data pipelines (batch and stream). This bidirectional data processing enables the connection to the security orchestrator and the monitor, which create the link between these big data technologies and the security analytics service. All the relevant prerequisites for the security analytics service are given.

As one of the SISSeC goals is to conduct security analytics with a digital twin, we evaluate if our architectural module permits this instance of security analytics. Dietz & Pernul [6] state that digital twins can run in three different operation modes that benefit cybersecurity: simulations, replications, and analytics. The latter is relevant for generating foresight by applying statistical analyses or machine learning. Concerning the security analytics service's security modules, a digital twin represents the link from real-time data to feasible simulation or forecast scenarios. In this context, the digital twin is, therefore, regarded as a foresight security module. Since the digital twin can also derive recommendations for actions and forecasts, the digital twin is considered the prescriptive module. Our security analytics service stores digital twin models (e.g., AutomationML [23]) for the PCB manufacturer in the security analytics repository. Concerning this digital twin model, a security expert instantiates it as an analytical technique for the prescriptive module, enabling future system states' simulations. The PCB manufacturer must not necessarily create the digital twin itself but can also purchase it under certain circumstances from a machine's manufacturer.

This use case addresses only the prescriptive module but can be extended vertically regarding the remaining security modules. This use case demonstrates the general applicability of the security analytics service and the potential benefits for SMEs and research.

7 CONCLUSION

As the threat landscape continues to grow and the gap between larger companies and SMEs in terms of cybersecurity knowledge widens, this paper takes a step towards improving knowledge transfer. The knowledge transfer is provisioned by a conceptualization of a general-purpose flexible security analytics service for the IIoT suited for SME environments. Although this knowledge transfer has been tailored to SMEs, larger companies can also benefit from a shared knowledge base. In general, the shared knowledge basis is necessary because companies are also encountering non-standardized reference architectures in the IIoT and encounter advantages regarding their financial and personnel resources.

We presented a security analytics service that enables security capabilities through security modules. These security modules can be instantiated by diverse analytical techniques to address a particular industry's needs flexibly. Furthermore, we disclosed how to generate data to wisdom and included this knowledge into our security analytics service. Thus, we bridge the knowledge gap within SMEs and contribute to the multi-disciplinary research area concerning the IIoT. Moreover, we presented the security analytics repository, which transfers knowledge between SMEs and established peer-reviewed implementations of analytical techniques.

Besides the contribution, this paper exhibits two limitations. First, we do not specify any access control roles to our security analytics service, as we have only covered a single entity (security expert). Second, we excluded parts of the incident response. As the response opportunities to a detected or predicted threat increase with the complexity of IIoT systems, e.g., with software-defined networks for the ICS, a holistic sketch of interconnections within the architectural module is almost impossible.

In the future, we will work on implementing a prototype of this security analytics service. The design of a specific implementation can be structured in several ways. We recommend a design based on virtual machines, dockerized applications, and specific connectors for a straightforward adoption and application of the capability modules. As we admit in the introduction, data is vital to generate wisdom. By bundling the data sources and types of data (stream vs batch) required to fulfil a particular capability module (e.g., the descriptive module takes batch data), analytical techniques can be instantiated. Thus, each capability module should be represented by one virtual machine. The virtual machines can be interconnected with dedicated bridges to accomplish the dependencies between the capability modules. Each of the virtual machines is running on Docker to establish containerized and modular applications. There are many open-source software and tools for security analytics applications that need to be categorized depending on their functionality, e.g., Suricata is an intrusion detection tool and is, therefore, available in the detective module. In this context, the security analytics repository is considered a collection of dockerized applications. Besides provisioning dockerized applications, the security analytics repositories need to yield connectors for each application to bind multiple applications' sinks and sources. Additionally, machine learning models can be shared across the security analytics repository and integrated into the desired capability module. Those machine learning models might be integrated into the predictive module (stream data) and the descriptive module (batch data). The different analytical techniques can be tagged inside the security analytics repository (e.g. for detective module) and ranked by other users. These mechanisms are resulting in transparency as to which methods are successfully contributing to the cybersecurity.

Based on this prototype, an evaluation is planned, which measures the effectiveness and efficiency of the service. Quasi-experiments evaluate latter for each of the modules, e.g. the descriptive module is used by one group, which is working out an incident with security analytics service, while the other group does not have the security analytics service available.

In summary, we expect that this paper will contribute to the current state of research, as security capabilities have not been included in security analytics so far. Furthermore, this paper will also transfer knowledge to industry and, in the case of a successful prototype, will also contribute to the exchange and creation of knowledge.

ACKNOWLEDGMENTS

We want to acknowledge all project partners involved in the Secure Industrial Semantic Sensor Cloud (SISSeC) project funded by the German Federal Ministry of Economics and Energy (BMWi) as part of its central innovation program for small and medium-sized businesses.

REFERENCES

[1] R. L. Ackoff. 1989. From Data to Wisdom. *Journal of Applied Systems Analysis* 16 (1989), 3–9.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash. 2015. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communications Surveys Tutorials* 17, 4 (2015), 2347–2376.

[3] F. Böhm, M. Vielberth, and G. Pernul. 2021. Bridging Knowledge Gaps in Security Analytics. *Proceedings of the 7th. International Conference on Information Systems Security and Privacy (ICISSP 2021)*, 98–108.

[4] A. A. Cárdenas, P. K. Manadhata, and S. P. Rajan. 2013. Big Data Analytics for Security. *IEEE Security and Privacy* 11, 6 (2013), 74–76.

[5] Cloud Security Alliance. 2011. *Defined Categories of Security as a Service.* Technical Report.

[6] M. Dietz and G. Pernul. 2020. Unleashing the Digital Twin's Potential for ICS Security. *IEEE Security Privacy* 18, 4 (2020), 20–27.

[7] M. Dietz, B. Putz, and G. Pernul. 2019. A Distributed Ledger Approach to Digital Twin Secure Data Sharing. In *Proceedings of the 33th. Annual IFIP WG 11.3 Conference on Data and Applications Security and Privacy XXXIII (DBSec 2019).* Springer International Publishing, Cham, 281–300.

[8] S. Dojkovski, S. Lichtenstein, and M. Warren. 2007. Fostering Information Security Culture in Small and Medium Size Enterprises: An Interpretive Study in Australia. In *Proceedings of the 15th. European Conference on Information Systems (ECIS 2007).*

[9] A. Furfaro, A. Garro, and A. Tundis. 2014. Towards Security as a Service (SecaaS): On the modeling of Security Services for Cloud Computing. In *Proceedings of the 48th. International Carnahan Conference on Security Technology (ICCST 2014).* 1–6.

[10] M. Goyal and M. Dutta. 2018. Intrusion Detection of Wormhole Attack in IoT: A Review. In *Proceedings of the 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET 2018).* 1–5.

[11] M. E. Jennex. 2009. Re-Visiting the Knowledge Pyramid. *Proceedings of the 42ndn Hawaii International Conference on System Sciences (HICSS 2009)*, 1–7.

[12] M. Kiran, P. Murphy, I. Monga, J. Dugan, and S. S. Baveja. 2015. Lambda Architecture for cost-effective Batch and Speed Big Data Processing. In *Proceedings of the 2015 IEEE International Conference on Big Data (Big Data 2015).* 2785–2792.

[13] I. Kotenko, O. Polubelova, and I. Saenko. 2012. Data Repository for Security Information and Event Management in Service Infrastructures. In *Proceedings of the 9th. International Conference on Security and Cryptography (SECRYPT 2021).* 308–313.

[14] S. W. Lin, B. Miller, J. Durand, R. Joshi, P. Didier, A. Chigani, R. Torenbeek, D. Duggal, R. Martin, and G. Bleakley. 2015. *Industrial Internet Reference Architecture.* Technical Report.

[15] S. W. Lin, B. Murphy, E. Clauer, U. Loewen, R. Neubert, G. Bachmann, M. Pai, and M. Hankel. 2017. *Architecture Alignment and Interoperability: An industrial Internet Consortium and Plattform Industrie 4.0 joint whitepaper.* Technical Report.

[16] T. Mahmood and U. Afzal. 2013. Security Analytics: Big Data Analytics for Cybersecurity: A Review of Trends, Techniques and Tools. In *Proceedings of the 2nd. National Conference on Information Assurance (NCIA 2013).* 129–134.

[17] R. Martin, S. Schrecker, H. Soroush, J. Molina, J. P. LeBlanc, F. Hirsch, M. Buchheit, A. Ginter, H. Banavara, S. Eswarahally, K. Raman, A. King, Q. Zhang, P. MacKay, and B. Witten. 2016. *Industrial Internet Security Framework.* Technical Report.

[18] P. M. Mell and T. Grance. 2011. *SP 800-145. The NIST Definition of Cloud Computing.* Technical Report.

[19] S. Mittal, M. Ahmad Khan, D. Romero, and T. Wuest. 2018. A critical Review of Smart Manufacturing & Industry 4.0 Maturity Models: Implications for small and medium-sized Enterprises (SMEs). *Journal of Manufacturing Systems* 49 (2018), 194–214.

[20] A. Molina Zarca, J. Bernal Bernabe, I. Farris, Y. Khettab, T. Taleb, and A. Skarmeta. 2018. Enhancing IoT Security through Network Softwarization and virtual Security Appliances. *International Journal of Network Management* 28, 5 (2018), e2038.

[21] M. M. Salim, S. Rathore, and J. Park. 2019. Distributed Denial of Service Attacks and its Defenses in IoT: a Survey. *The Journal of Supercomputing* 76, 7 (2019), 5320–5363.

[22] R. Scandariato, K. Wuyts, and W. Joosen. 2013. A descriptive Study of Microsoft's Threat Modeling Technique. *Requirements Engineering* 20 (2013), 163–180.

[23] G. N. Schroeder, C. Steinmetz, C. E. Pereira, and D. B. Espindola. 2016. Digital Twin Data Modeling with AutomationML and a Communication Methodology for Data Exchange. *IFAC-PapersOnLine* 49, 30 (2016), 12–17.

[24] K. Schweichhart. 2016. *Reference Architectural Model Industrie 4.0 (RAMI 4.0).* Technical Report.

[25] C. Senk. 2013. Adoption of Security as a Service. *Journal of Internet Services and Applications* 4 (2013), 1–16.

[26] P. Sethi and S. Sarangi. 2017. Internet of Things: Architectures, Protocols, and Applications. *Journal of Electrical and Computer Engineering* 1 (2017), 1–25.

[27] H. Shafiei, A. Khonsari, H. Derakhshi, and P. Mousavi. 2014. Detection and mitigation of sinkhole attacks in wireless sensor networks. *J. Comput. System Sci.* 80, 3 (2014), 644–653.

[28] E. Siow, T. Tiropanis, and W. Hall. 2018. Analytics for the Internet of Things: A Survey. *Comput. Surveys* 51, 4 (2018).

[29] K. A. Torkura, M. I. H. Sukmana, F. Cheng, and C. Meinel. 2017. Leveraging Cloud Native Design Patterns for Security-as-a-Service Applications. In *Proceedings of the 2nd. International Conference on Smart Cloud (SmartCloud 2017).* 90–97.

[30] W. Wang and S. Yongchareon. 2017. A Survey on Security as a Service. In *Proceedings of the 18th. Web Information Systems Engineering (WISE 2017).* Springer International Publishing, Cham, 303–310.

[31] T. Y. Win, H. Tianfield, and Q. Mair. 2018. Big Data Based Security Analytics for Protecting Virtualized Infrastructures in Cloud Computing. *IEEE Transactions on Big Data* 4, 1 (2018), 11–25.

[32] S. Xu. 2019. Cybersecurity Dynamics: A Foundation for the Science of Cybersecurity. In *Proactive and Dynamic Network Defense,* C. Wang and Z. Lu (Eds.). Springer International Publishing, Cham, 1–31.

[33] M. Zaharia, R. S. Xin, P. Wendell, T. Das, M. Armbrust, A. Dave, X. Meng, S. Rosen, J.and Venkataraman, M. J. Franklin, A. Ghodsi, J. Gonzalez, S. Shenker, and I. Stoica. 2016. Apache Spark: A Unified Engine for Big Data Processing. *Commun. ACM* 59, 11 (2016), 56–65.