Asset logging in the energy sector: a scalable blockchain-based data platform

Alexander Djamali1*, Patrick Dossow1, Michael Hinterstocker1, Benjamin Schellinger2,3,4, Johannes Sedlmeir2,3,4, Fabiane Völter2,3,4 and Lukas Willburger2,3,5

From The 10th DACH+ Conference on Energy Informatics
Virtual. 13–17 September 2021

Abstract
Due to a steeply growing number of energy assets, the increasingly decentralized and segmented energy sector fuels the potential for new digital use cases. In this paper, we focus our attention on the application field of asset logging, which addresses the collection, documentation, and usage of relevant asset data for direct or later verification. We identified a number of promising use cases that so far have not been implemented; supposedly due to the lack of a suitable technical infrastructure. Besides the high degree of complexity associated with various stakeholders and the diversity of assets involved, the main challenge we found in asset logging use cases is to guarantee the tamper-resistance and integrity of the stored data while meeting scalability, addressing cost requirements, and protecting sensitive data. Against this backdrop, we present a blockchain-based platform and argue that it can meet all identified requirements. Our proposed technical solution hierarchically aggregates data in Merkle trees and leverages Merkle proofs for the efficient and privacy-preserving verification of data integrity, thereby ensuring scalability even for highly frequent data logging. By connecting all stakeholders and assets involved on the platform through bilateral and authenticated communication channels and adding a blockchain as a shared foundation of trust, we implement a wide range of asset logging use cases and provide the basis for leveraging platform effects in future use cases that build on verifiable data. Along with the technical aspects of our solution, we discuss the challenges of its practical implementation in the energy sector and the next steps for testing in a regulatory sandbox approach.

Keywords: Asset management, Distributed ledger, Energy asset, Merkle proof, Privacy, Self-sovereign identity

Introduction
In light of the ongoing energy transition, the energy sector is subject to significant changes, which are expected to further accelerate in the future. The ever-larger number of decentralized energy assets, most notably wind turbines and photovoltaic systems but also stationary batteries and battery electric vehicles, increases the complexity of the energy system at a challenging pace. As of today, some owners and operators of energy
assets struggle with these changes while others are keen to seize emerging opportunities.

A major field of development and change in the energy sector is the acquisition and usage of digital data for the documentation and verification of the state and operation of assets (Zeiselmair et al. 2019). Digitalization is a necessity to monitor and evaluate a large number of often decentralized assets and to ensure the proper functioning of the complex energy system (BDEW Bundesverband der Energie- und Wasserwirtschaft e.V. 2020). This is particularly important for assets of systemic relevance and assets that are newly integrated into already digitalized processes. In addition, reliable digital asset data can enable use cases that were previously infeasible or impractical. In this regard, complex processes with a high demand regarding up-to-date data availability and temporal resolution can be newly implemented through the appropriate provisioning and handling of high-quality data (Albrecht et al. 2018).

To achieve actual progress, we focus our research on a field of applications beyond the mere digitization of asset data. Through systematic dialogues and focus groups with experts of partner companies from the energy sector, we identified that most relevant use cases are characterized by a small yet challenging set of common requirements. These use cases, which we refer to as “asset logging” use cases, generally demand tamper-resistance and verifiability along the entire chain of data usage and traceability of the chronological sequence of collected data, while maintaining data privacy and sovereignty required by the data's respective owner. In turn, these specifications led to our research on device-specific machine identities and a decentralized, blockchain-based platform (Carminati et al. 2018). In this paper, we aim to create a strong understanding of the requirements, challenges, and potentials of the group of asset logging use cases. We present our novel digital platform architecture that we designed and implemented to realize the identified use cases and showcase how it meets the identified requirements. A discussion of prevailing issues and challenges for the intended practical implementation of selected use cases in a sandbox approach completes our contribution.

**Use case assessment in the energy sector**

In the energy sector, the transition to digitalization and new digital business approaches tends to be a relatively slow process (Bundesministeriums für Wirtschaft und Energie (BMWi) 2020). As for asset logging, we found that even though the documentation and verification of assets is a necessity for many applications (Balzer and Schorn 2014), there is still a lack of clear standards and a strong dependency on different degrees of digitalization (Zeiselmair et al. 2019), which opens up room for improvements. At the same time, an increasing overlap and interaction of different roles creates the need as well as the opportunity for new use cases in the field. One example of such new stakeholders in the energy system is the “prosumer”, who has emerged as a hybrid of the two conventional roles of electricity producer and electricity consumer as a result of the increasing electricity production by traditional end consumers through their own photovoltaic systems (Kotler 2010; Toffler 1980).

**Basics and preliminary investigation**

**Definition of asset logging**

To ensure a shared understanding, a clear definition of asset logging as a field of application is essential. According to our definition, asset logging comprises scenarios in which
data from registered assets is logged and stored for the later or ongoing verification of certain propositions or processes. Thus, asset logging use cases generally consist of three primary steps: data collection, tamper-resistant data storage, and the verification whether certain conditions, which were agreed upon ex-ante, are met on the basis of collected data.

Warranty management represents an exemplary use case. In order to assess in hindsight whether warranty conditions are met, the operation of relevant assets is continuously monitored. Asset data documenting its operation is periodically collected and stored in a tamper-resistant way. If a warrantee raises a warranty claim, the tamper-resistant data serves for assessing whether the asset has been operated according to the conditions defined in the warranty agreement. If this is the case, the warranty claim is valid. As the data is only shared in the case of a warranty claim, business secrets of the warrantee are typically preserved. Due to the tamper-resistant data storage, the warrantor can be certain that only genuinely valid warranty cases are confirmed, while the warrantee is assured that all warranty cases can be provably claimed on the basis of verifiable data.

Initial research

Previous research has already identified asset logging as one of the most promising fields of application for a digital, potentially blockchain-based platform (Bogensperger et al. 2018). Inspired by this research, we conducted a close, systematic exchange via bilateral workshops with focus groups of interested, relevant partner companies from the energy sector. In these interactions, we realized that various stakeholders with different roles in the energy sector own or operate a variety of often decentralized energy assets. On the one hand, we identified that existing challenges associated with a lack of digitalization in this field of application can be remedied by a fully automated platform solution. On the other hand, we determined that we can ensure tamper-resistance and data integrity in these cases through the usage of blockchain technology. Furthermore, use cases could be implemented that could previously not or hardly be realized, as the handling of large numbers of both different assets and different stakeholders could be enabled through a platform with automated documentation and verification.

Potential assessment

Building on our insights of the preceding project, we now took a more detailed and methodical look at the application field of asset logging, the relevant associated use cases and the requirements to realize such use cases. For a start, we identified the potential of asset logging use cases by estimating the number of eligible assets in Germany. For the numbers of today, we gathered status quo data, whereas for our estimation of 2030, we used political targets and projections. In line with Hinterstocker et al. (2020), we found renewable generation plants as well as electricity consumption and storage units to be the most relevant groups of assets in this regard.

Figure 1 displays the resulting list and numbers of potential assets in the energy sector. As of today, already more than three million assets are potentially of relevance for use cases in the field of asset logging, where photovoltaic systems and heat pumps account for the largest numbers of assets. Within the next decade, we estimate that the number of relevant assets is likely to multiply, resulting in over 17 million potentially applicable assets for 2030. With more than half of the potential assets, battery electric vehicles are
anticipated to be by far the most relevant group of assets for 2030, while photovoltaic systems continue to play a major role with a share of more than 20%. As the number of assets potentially suitable for logging solutions will most likely grow significantly in the future, we expect asset logging to increasingly gain relevance. For further investigations of the real potential and motivation of the energy sector to implement such use cases and to test our technical solutions developed in this context, we must identify, discuss and rank potentially relevant use cases.

**Use case analysis**

In our current research, we set out to methodically develop and select relevant use cases in the field of asset logging. For this purpose, we implemented a three-step process to identify, prioritize, and select use cases that could potentially be tested later in a sandbox approach. Our use case process allowed us to combine practical input from our partner companies from the energy sector and scientific expertise of the interdisciplinary project team in the fields of energy economics, computer science, and legal science.

**Identification**

To start with the identification and development of potentially relevant use cases, we conducted a total of eleven bilateral workshops with more than 70 expert participants from various business areas of the energy sector. Apart from public utility companies, partners from large energy providers, full-service providers, as well as from component manufacturers and both distribution network operators and transmission system operators were represented. In the first phase of the workshops, a total of 90 separate high-level discussions were held on 32 different use cases, 16 of which we assign to the field of asset logging. In a second phase, most of the workshop time was devoted to in-depth interviews concerning the specific benefits and pitfalls of the use cases considered to be most relevant for each group of experts.

Resulting from these focused examinations, we identified twelve asset logging use cases to be of high relevance for both commercial application and interdisciplinary research.
These use cases are characterized by a high business potential, where cost savings or additionally generated revenues through automation are frequently complemented by an added value in tamper resistance and privacy aspects. Additionally, it is precisely these high expectations on tamper-resistant processing and transfer of data for verification in combination with the prevailing demands for data privacy and data sovereignty that make the cases highly interesting from a scientific point of view. Finally, we clustered the identified twelve asset logging use cases based on their data requirements and created a comprehensive definition for each relevant use case, including user story, added value, required data, descriptions of the roles involved, and various representations of the process flow. An overview of the asset logging use cases as well as exemplary parts of the use case definitions are published in Hinterstocker et al. (2020).

**Prioritization**

The second methodical step consists of combining several use case rankings that were developed after the workshops. We created a first ranking of the use cases based on the insights of the workshop series, taking the gained assessment over stakeholder benefits, technical requirements, customer potential, and legal obstacles as perceived by the partner companies into account. Here, frequency as well as the depth of detail and recognized relevance were converted into a numerical rating.

A second ranking was compiled by incorporating the views of all participating research institutes. To do so, all research partners provided an expert assessment of the use cases from a legal, technical, and energy economics perspective. Again, we converted these assessments into a single numerical rating. Combining these two rankings, a conclusive ranking was created which thus reflects both the practice-oriented view of the partner companies and the research perspective of scientific experts. On the basis of this ranking, we concluded that eight of the previously identified asset logging use cases have a high relevance across all perspectives.

**Selection**

In a third and final step, we selected those use cases best suited for actual implementation under real conditions. To be selected, a use case must meet two conditions: First, it must have both high business potential and a high added value from a research perspective. This condition is represented by a high ranking in the conclusive ranking resulting from the prioritization step. Second, efforts arising from an initial implementation must be manageable, which implies easy access to asset data and the voluntary participation of stakeholders. From the pool of identified asset logging use cases, four use cases meet these conditions (Hinterstocker et al. 2020):

1) Service and maintenance models, where contractual agreements are verified through the tamper-resistant documentation of maintenance data
2) Warranty management, where a warranty claim is verified through the tamper-resistant documentation of asset data
3) Operation contracting, where the operation of an asset is outsourced and thus documented in a tamper-resistant manner to prevent conflicts between operator and owner
4) Regulatory requirements, where tamper-resistant asset data is transmitted for regulatory requirements.
These four use cases share similar data and infrastructure requirements, so implementing one use case makes implementing the other use cases easier due to emerging synergies.

Requirements for asset logging in the energy sector
In the course of the use case process, we found that a shared set of requirements must be met for all relevant asset logging use cases. The requirements we identified can be classified into two groups. In group one, requirements for data measurement and collection are defined. The second group specifies requirements for the actual technical solution i.e. the platform architecture. Here, we outline the requirements for the storage, processing, and further use of collected data on the platform.

Data measurement and collection requirements
To ensure data integrity in asset logging, the entire chain of data processing must be tamper-resistant, which includes the step of data collection. Hence, tamper-resistant data collection must be technically possible, which means that manipulation of data during or immediately after measurement must be prevented on all accounts (Albrecht et al. 2018). At the same time, the data must be clearly attributable to its origin, i.e. the corresponding asset, where the chronological sequence of the measurement must be transparent (Albrecht et al. 2018). We refer to this set of requirements as traceability. In addition, the asset owner or operator must permit the tamper-resistant collection of asset data. In this regard, most asset owners demand data privacy and data sovereignty, especially since the data might contain sensitive business secrets.

In the case of warranty management, for instance, the following types of data must be collected: maintenance and operation schedule, maintenance and availability reports, and operational data. These different types of data are measured and collected by different means at different intervals, e.g., maintenance reports are compiled at the discrete time of reporting, whereas operational data is continuously collected by sensors in a fixed temporal resolution. Regardless of these differences, all data must be collected in a tamper-resistant, traceable manner and their further use must be permitted by the asset owner.

Architecture requirements
As the collected data must be stored for later use or processing, requirements arise for the technical implementation of a suitable data platform. For asset logging, these architecture requirements are:

1) Ex-post verification of data integrity
2) Protection of business or trade secrets
3) Secure identification of participating stakeholders
4) Scalability of the platform.

As all relevant asset logging use cases involve some kind of data verification process, all data stored on and accessed from the platform must be both traceable and tamper-resistant to ensure data integrity. Similar to the requirements during data collection, a sufficient degree of data privacy is demanded to protect important business or trade secrets without impeding the verification process. Furthermore, all relevant stakeholders, which include public authorities in some cases, must be able to clearly identify or authen-
ticate themselves when accessing the platform. Finally, the platform must be designed in a way that guarantees scalability, i.e., that it can be easily expanded to additional use cases and participants while ensuring cost-effectiveness.

**Technical background**

To provide a basic understanding of our proposed technical solution for asset logging that meets all requirements, we must first outline the relevant technological concepts of metering infrastructure, relevant aspects of blockchain technology, and Self-sovereign Identity (SSI) in the following.

**Tamper-resistant metering infrastructure**

To guarantee data integrity along the entire data processing line, tamper resistance is a basic prerequisite right from the data collection stage. For this reason, the metering infrastructure directly at the asset must meet high data security and privacy standards for asset logging use cases. To discuss these requirements for metering, the smart metering infrastructure in Germany is considered as a relevant example since its technological development is well advanced and comprehensive standards for secure data processing and transmission exist (Bundesamt für Sicherheit in der Informationstechnik (BSI) 2018). Other data-collecting technology must meet the same criteria as defined in Use case assessment in the energy sector, for which no industrial standard as with smart meters has been defined yet.

The basic principle of smart metering infrastructure is to provide government-certified infrastructure for tamper resistance, privacy, and traceability of the data collection process (Bundesamt für Sicherheit in der Informationstechnik (BSI) 2018). In this context, the smart meter provides tamper-resistant and traceable measurements of at least electrical energy, with meter readings at 15-minute-intervals, electricity feed-in (for electricity-generating assets), and grid status (current, voltage, and phase angle) (Bundesregierung 2016). These measurements can be forwarded through an encrypted Local Metrological Network (LMN) to the Smart Meter Gateway (SMGW), which can transmit meter data towards authorized stakeholders via an encrypted Wide Area Network (WAN) (Bundesamt für Sicherheit und Informationstechnik 2013). The smart meter rollout in the European Union follows respective guidelines (Europäische Union 2009). Since the implementation of the guideline is the responsibility of the member states, the progress on meters, infrastructure, and rollout differs within the member states of the European Union (Alaton and Tounquet 2020).

**Blockchain technology**

A blockchain represents a particular type of distributed data structure that contains information grouped into blocks. Each node participating in the peer-to-peer network redundantly stores the data. Since each block refers to the previous block using a hash pointer, the data blocks are chronologically ordered and concatenated into a chain. Any modifications to data in the chain are detected, making the blockchain a tamper-resistant database (Beck et al. 2016). Upon creating a transaction, a user digitally signs their transaction using a private key. Prior to processing transactions, the nodes use the sender’s public key to verify the transaction. Thus, a blockchain relies on public-private key cryptography to ensure the legitimacy of transactions. When transactions are propagated on
the network, nodes must agree on the state of the system, i.e., the integrity of the transactions and the correct order of the blocks. The mechanisms for reaching consensus are manifold and imply different benefits and drawbacks (Wüst and Gervais 2018; Zhang et al. 2019; Kannengießer et al. 2020; Sedlmeier et al. 2020). Using consensus protocols eliminates the need for a central trusted party, which is seen as a core value of the technology (Fridgen et al. 2018). After the network reaches consensus, each node adds this new block to its own copy of the blockchain (Beck et al. 2016). Although the data recorded on the blockchain is cryptographically secured and thus tamper-resistant, this does not eliminate the possibility of altered input data in the first place (Sheldon 2020). Also, there exist different types of consensus mechanisms like Proof-of-Work (PoW), Proof-of-Stake (PoS), and Proof-of-Authority (PoA). PoW and PoS form the majority of consensus mechanisms in public permissionless blockchains, while PoA is used in permissioned blockchains (Beck et al. 2016; Xiao et al. 2020). In this context, the type of consensus mechanism and the number of nodes that redundantly store copies of the blockchain determine the energy consumption of a blockchain (Sedlmeier et al. 2020). While PoW-based blockchains are energy-intensive by design through the mining process and their energy consumption can be as large as the one of an entire industrialized country, the energy consumption of PoS-based permissionless blockchains and PoA-based permissioned blockchains is predominately caused by redundant storage and verification. Thus, despite being less efficient than a centralized ledger, these types of blockchains exhibit an energy consumption that is several orders of magnitude lower than that of PoW-based blockchains like Bitcoin (Sedlmeier et al. 2020). We can safely assume that by choosing a non-PoW-based blockchain design, the energy footprint of a blockchain-based platform for asset logging is not concerning. In fact, this choice of technology may even cause net energy or carbon savings if considerable amounts of paper-based documentation can be prevented.

Researchers and practitioners generally distinguish between the design parameters access restriction and reading/writing permissions (Wüst and Gervais 2018). The former addresses whether transactions are publicly visible or only visible to pre-defined parties. The latter design parameter defines whether participation in consensus and the validation of transactions is permissioned or permissionless. In general, blockchains are regarded to offer high security guarantees. However, both permissioned and permissionless blockchains still exhibit some vulnerabilities that need to be considered (Guggenberger et al. 2021).

The hash of a block is derived from the data stored in the block, e.g., transactions, a timestamp, or the difficulty. Transactions and their corresponding metadata are typically saved in a data structure using a Merkle tree. Merkle trees allow to efficiently represent an arbitrary number of transactions or data sets with any size in one hash. Using a binary Merkle tree, data points are repeatedly hashed in pairs to form a final hash. This last represents a cryptographic commitment to all the underlying transactions or “leaves”, also referred to as Merkle root. In addition, Merkle trees allow to verify the integrity of a data point using only a subset of the hashed data points (Merkle 1987), using a so-called Merkle Proofs (MPs) that contains the adjacent hashes of the path from the Merkle root to the data point under consideration. As a result, this enables verification with low computational overhead compared to hashing all data points in a stream. In this case, a verifier would need the hashes of all data points to verify that a particular data point was included in the processing. The depth of a binary Merkle tree (and hence the computational and
storage complexity of a MP) scale with \( \log_2(N) \), or in a \( p \)-ary tree with \( \log_p(N) \), where \( N \) represents the number of data entry points. In the case where every data point is individually hashed, the length of the hash list scales with \( N \). As pointed out above, verifiers require only a subset of hashes to verify the integrity of a data point when using MPs. This mechanism is illustrated in Fig. 2. In order to create a referencing structure along blockchains, each block points towards the Merkle root of the transactions of the previous block.

In addition to storing plain data such as transactions, it is also possible to integrate business logic into the blockchain via the deployment of Smart Contracts (SCs). SCs are computer programs that execute predefined code when certain conditions are met (Szabo 1997; Buterin and et al. 2014). Generally, external information, i.e. a transaction to a SC, triggers the inherent functions. Thus, with the introduction of the Ethereum blockchain, decentralized applications and digital tokens could be implemented using SCs for the first time which enables the creation of new ecosystems (Buterin and et al. 2014).

In sum, researchers and practitioners acknowledge several important characteristics of the technology (Beck et al. 2016; Zhang et al. 2019; Butijn et al. 2020; Xiao et al. 2020; Gudgeon et al. 2020; Amend et al. 2021). First, by using cryptographic hash references, blockchains establish tamper-resistant data records (Beck et al. 2016). Second, consensus mechanisms provide a single source of truth (Xiao et al. 2020). Third, its distributed manner and redundant data storage ensure relatively high resistance to malicious attacks and crashes (Zhang et al. 2019). Fourth, MPs facilitate the verification of data integrity (Merkle 1987). Fifth, SCs allow to implement arbitrary business logic on blockchains (Buterin and et al. 2014). As a result, blockchains are considered as highly trusted, which is beneficial for critical infrastructures such as the energy sector (Andoni et al. 2019; Bao et al. 2020), and enable the implementation of other technologies.

**Self-sovereign identity (SSI)**

Our discussions in the bilateral workshops with stakeholders revealed that siloed and outdated data are frequent problems in the energy sector. To achieve authenticated and End-to-End (E2E) encrypted bilateral communication channels between parties, verifying identities and their attributes is indispensable – regardless of analog or digital information exchange (Bernal Bernabe et al. 2019). Centralized or federated identity management solution can address some of these problems but may not leave the users in control of their own information and raise security, ethical, or economic concerns because of the data that they can aggregate. Der et al. (2017); Liu et al. (2020); Mühle et al. (2018). Given that situation, an emerging identity management paradigm called SSI can provide an architecture to leverage portable identities that are maintained in a decentralized manner (Wang and De Filippi 2020; Strüker et al. 2021). SSI can essentially be regarded the application...
of asymmetric cryptography and digital certificates for the identity management not only of servers on the web (as we have known it for 30 years now) but also for end users and smart devices.

SSI can best be described with an analogy from the real world: everyone possesses a wallet that contains multiple plastic cards like a driver’s license or a personal ID card. In the context of SSI, this storage relates to the digital wallet, which can be represented by an app on the owner’s smartphone (Hong and Kim 2020). The physical identification cards themselves only contain the relevant information for a certain context. Driver’s licenses may include the name of drivers and the range of vehicles they are allowed to drive, but not their birthplace, as it is not important in traffic control. The issuing authority, like the federal state, ensures the credibility, tamper-resistance, and uniqueness of the document and makes its underlying schema publicly available. Therefore, third parties can verify its integrity without contacting the issuer. SSI provides a similar approach to physical ID cards by using Verifiable Credentials (VCs) (Avellaneda et al. 2019; Mühle et al. 2018). VCs contain identity data about their owner, which are digitally signed by trustworthy authorities using cryptographic techniques (Sporny et al. 2019). Usually, they are stored in a dedicated digital wallet to which only the owner has access. In addition, the rise of agents provides a complementary solution by managing certain VCs without the need for the owner to be permanently available (Ferdous et al. 2019; Nauta and Joosten 2019). Either way, credentials are generally not transferred directly to other parties: The owner generates Verifiable Presentations (VPs) of one or more VCs, respectively as a subset of their properties, to present tamper-resistant evidence to a verifying party (Preukschat and Reed 2019; Sporny et al. 2019). VCs are not limited to information itself but include possibilities to also provide statements, e.g., whether a person is a resident of a certain city. The underlying cryptographic techniques include, amongst others, Zero-Knowledge Proofs (ZKPs). ZKPs solely guarantee the validity of a statement and do not disclose any additional, unnecessary information, thereby preserving privacy to the maximum extent (Goldwasser et al. 1989).

With the absence of physical interaction and the need for secure data transmission, parties in an SSI ecosystem can assign themselves unique identifiers, so-called Decentralized Identifiers (DIDs), to establish bilateral, E2E encrypted messaging channels. DIDs must hence be created decentrally and should be renewed for every interaction, especially when natural persons are participating, to ensure that correlations are impossible. A standard format developed by the W3C defines three mandatory components of a DID (Reed et al. 2020): The first part contains the underlying URI-schema followed by the DID method which specifies the chosen DLT and how operations shall be executed. The third block completes the DID by providing a method-specific identifier. A given DID resolves to a linked DID document consisting of related information like cryptographic details.

However, in order to achieve a fully integrated system, further infrastructural components are necessary. To verify the integrity of VPs, information about their underlying schema and their issuers is essential. In addition, credentials can be revoked at any given time by the original issuer. Therefore a registry must be established to verify that a VC is valid. Using such a public yet privacy-ensuring registry in combination with the verification of the issuer’s digital signature, holders can prove that a VC is not revoked without contacting the issuer. Against this background, blockchains are often considered well suited for an unbiased registry that does not require certificate authorities to
maintain the public key infrastructure and that provides additional services for ecosystem governance. Due to its decentralized, highly available, and tamper-resistant nature, a blockchain can thus perfectly facilitate the convergence to other technologies such as SSI, releasing synergies created by the combination of both (Ferdous et al. 2019; van Bokkem et al. 2019).

Platform architecture for asset logging
We propose a decentralized blockchain-based approach for the architecture of our platform. This can help to impede monopolization and increase stakeholder acceptance through direct participation. Since there are no alternative centralized approaches to the best of our knowledge, a comparison in this regard has yet to be made (Bogensperger et al. 2018).

Architecture development
In order to easily verify data integrity, a first approach could be to write plaintext data to the blockchain. However, storing plaintext data on the blockchain raises privacy concerns and may violate data protection requirements. To reconcile data verifiability and privacy, hashes of data are therefore stored on the blockchain. In addition, scalability is critical to anchor data efficiently on the blockchain. Given these requirements, MPs offer a suitable solution to achieve this goal.

Limitations of existing blockchain-based approaches
On the basis of our derived requirements, we propose an approach building on four essential components: Tamper-resistant data logging through certified components (e.g., sensors), digital signatures as well as the authenticated and E2E encrypted bilateral data exchange via SSI, blockchain technology, and Merkle trees. Overall, our architecture involves interactions of two types of players. Generally, asset owners take the role of provers while any other stakeholder might represent a verifier. In a typical use case, asset owners or operators aim to prove that their asset data is reliable – i.e., authentic at the time of generation, and unchanged since. Verifiers check ex-post that the data is trustworthy with high reliability. As data is shared in cases of disputes, provers might have conflicts of interest, which may spur manipulations. Under these circumstances, verifiers are interested in verifiably tamper-resistant data storage.

A simple and frequently suggested solution to achieve this goal would run blockchain clients on assets to directly push the data that they generated onto a distributed ledger. Due to a blockchain’s familiar characteristics of tamper-resistance and practical immutability, this would ensure a high level of trust and availability of data to third parties when they request it later; in fact, verifiers could directly query the relevant data from their own blockchain node or request it from a node that they trust. However, when following this approach, we would encounter two significant problems. On the one hand, data privacy is not guaranteed as any blockchain node would be able to access the data under consideration. This challenge could be solved by hashing the written data and providing plain-text data to third parties off-chain through a bilateral communication channel. Yet, on the other hand, we would quickly encounter scalability issues. Storing data on public blockchains such as Ethereum is typically highly expensive and only feasible to a very limited amount in the order of a few kilobytes per second. On the Ethereum
blockchain, a single block typically contains around 50 kB of data, and the average block time is around 13 s. An operation for only storing a few bytes of data in a SC consumes 20,000 gas, which is required for the network’s computational work. Thus, the theoretical maximum for a simple storing operation would be 50 transactions per second (or 650 transactions per block). At present, the empirical mean is 15 transactions per second on Ethereum, driven by more complex computations and consequently requiring larger amounts of gas (Etherscan 2021). In addition, the performance of permissioned blockchains that generally use better hardware is also limited to a few hundred up to a few thousand transactions per second (Sedlmeir et al. 2021). Even with solid hardware and when optimizing for upload capacity, a medium-sized Hyperledger Fabric network, which is a popular permissioned blockchain and one of the enterprise blockchains with the highest performance in a larger comparison (Sedlmeir et al. 2021), cannot upload more than 15 MB/s of data (Guggenberger et al. 2021). Taking into account that millions of assets in the energy sector are potentially relevant and a high temporal granularity would require each of them to send a sensor date every few minutes, this exceeds the capacity of permissionless blockchains by orders of magnitude and challenges the capacity of dedicated permissioned networks. In addition, the redundant storage of these amounts of data is expensive and wastes storage resources. Thus, solely registering asset data in plain text or in hashed form on distributed ledgers does not fulfill the requirements set out in Use case assessment in the energy sector.

Proposed architecture

To solve this problem, our architecture builds on hierarchical aggregation of data in the form of Merkle trees (Chod et al. 2020). The entire architecture is depicted in Fig. 3. Accordingly, we propose that data logging modules of assets make use of their private keys to sign their generated data. We propose so as digital signatures allow to ensure the authenticity of logged data. In specific, a third party can verify whether the provided data actually stems from the asset referred to. In case data remains unsigned, verifiers have no means to ensure whether the data was generated by an asset under contract or any other asset. Thus, no matter which processing mechanisms are applied, digital signatures should always be used to ensure data authenticity.
The entire process of the asset logging’s business logic is illustrated in the sequence diagram in Fig. 4. After signing the logged data for a certain epoch (e.g., every 15 minutes), assets send the data to their owner (or an aggregation service provider) using a bilateral, E2E-encrypted communication channel. Subsequently to receiving the signed data, owners aggregate the received data into a binary Merkle tree. Constructing a Merkle tree involves the subsequent hashing of signed transaction data. As a result, asset owners generate a Merkle root that consists of the asset’s logged data as displayed in Fig. 2. Asset owners run a client that emits an event on the blockchain containing the Merkle root. The resulting event logs are then written to the blockchain. The use of a smart contract and the included contract memory would also have been possible but would have led to higher gas costs.
Thus, the Merkle root stored on the blockchain represents a digital fingerprint of the logged asset data for a distinct epoch. In case of high data granularity, several epochs may also be aggregated.

In addition, asset owners locally save the asset’s logged data in plain text, which allows them to compute the full Merkle tree, including its root, at any time. They also note the block number in which the transaction that contained the Merkle root was added to the ledger. The block number is included in the transaction receipt when the block is stored on the blockchain. Against this backdrop, owners can provide third parties with relevant signed plain text data upon request over a bilateral E2E encrypted communication channel. After receiving the relevant plain-text data, third parties can request a MP to verify that the logged data has not been changed in the meantime. Based on the Merkle tree, asset owners generate a MP on demand (e.g., by a third party) containing the path to the corresponding root of the Merkle tree, which requires only a minimum number of hashes and configuration information. Then, third parties can first verify digital signatures using the public keys from corresponding DID documents. Second, using the MP it can be verified whether the data provided by the asset owner represents a valid input to the corresponding root. Third, verifiers check whether the MP is based on the same Merkle root as saved on the blockchain (see Fig. 4). Based on this information, the third party can verify the integrity of the data.

Requirements-based evaluation

The proposed approach bears three central advantages. First, by compressing large batches of data at once, MPs allow for a scalable solution. With an arbitrary amount of data points, the size of Merkle roots remains fixed (e.g. 256 bits when relying on SHA256). Thus, no matter how many data points are used as an input, the respective on-chain transaction data remains at a predetermined size. Second, as verifiers require only a subset of encrypted records, the verification of the Merkle tree (both the computational complexity for the verifier and the amount of communication that is necessary between the prover and the verifier) scales with \(\log_2(N)\). Third, in contrast to unordered batching of data points, MPs enable privacy by default. This is as verifying the integrity of MPs requires only surrounding hashes of the respective data input (see also Fig. 2). Thus, no additional data points must be revealed. In combination with the surrounding hashes, the users’ data points themselves are sufficient for verifying the Merkle Root’s correctness. In contrast, while unstructured batching of data to a single hash also is considered a scalable solution, it requires revealing all other inputs to verify the integrity of the resulting hash. Third, due to high entropy, Merkle roots also do not allow to trace back to the input data (as may happen for single hashes when there are only few reasonable options for the underlying data that can be tested by a brute-force approach). Hence, resistance to preimage attacks can be considered high (Merkle 1987).

As a result, by storing only Merkle roots on a blockchain that are associated with a large number of sensor values (from multiple sensors or multiple epochs), privacy and security risks as well as scalability issues are minimized. Thus, our approach can be considered scalable and privacy-preserving as it does not reveal any trade and business secrets or personal data. Notably, the blockchain does not guarantee that the original data, hence the logged data by an asset, has been altered by the owner or the operator at the point
of generation. However, the energy sector provides certified infrastructures, such that we may rely on trusted data logging devices (see Tamper-resistant metering infrastructure). Combining a blockchain-based approach, digital signatures at the point of trusted data creation, and resource-friendly verification mechanisms allows to check whether an alteration has taken place. Furthermore, our approach allows for a public, permissionless blockchain as it serves for tamper-resistant storage of Merkle roots only. No plain-text or single hashed data is written onto the ledger. The transactions themselves do not reveal any information without bilateral exchange of plain-text data from asset owners. We propose an open platform, which any stakeholder can access at any time and is private by design. As a result, our platform is not limited to an underlying use case but can be extended to further use cases and even domains in the future.

**Challenges for practical implementation**

The focus of designing our blockchain-based platform for asset logging use cases is not only on its academic contribution but also to be practically implemented and tested. While planning and implementing our field trials, we already encountered various challenges, gained first experiences in trial preparation, and collected feedback on the side of stakeholders such as commercial energy asset operators, energy service providers, or network operators.

**Prevailing challenges and limitations**

Some of the identified challenges can be bypassed or eliminated before long, others pose clear limitations in the field of asset logging.

**Digitalization and infrastructure**

To date, data collection and in particular data transfer and processing in Germany’s energy sector often remain a manual task (Bundesministeriums für Wirtschaft und Energie (BMWi) 2020). Therefore, the lack of digital infrastructure remains a severe impediment to the implementation of digital use cases such as asset logging. In some instances, asset data is even recorded analogously by employees and digitalized in a subsequent step only. Furthermore, numerous assets are not yet equipped with any measurement devices or sensors, which poses a major obstacle to the implementation of any digital use cases. In this respect, our solution for asset logging use cases can serve as an additional motivation for the involved stakeholders to digitalize, although this digitalization must take place before the platform can be implemented on a large scale.

In particular, the German SMGW rollout, which would provide a convenient way of tamper-resistant, privacy-ensuring data collection, proceeds slowly. Due to currently unresolved legal concerns, it is unclear when SMGWs will be wide-spread operational in Germany (Oberverwaltungsgericht Münster 2021). In contrast, similar infrastructure is already available in other countries. Yet, the slow rollout in Germany can be bypassed by using suitable measurement equipment that can communicate in an E2E encrypted way with an SSL-certified server. Such suitable measurement equipment could in some cases also be applied to collect other types of necessary data, which cannot be collected with a SMGW. Other metering solutions must, however, not only guarantee the tamper resistance of the meter but also that the meter is connected to the correct data generating process.
Regulation
Another potentially limiting factor for swift large-scale implementation represents the mandatory involvement of public authorities or regulators in some of the most promising asset logging use cases, such as the German grid regulator in the case of the verification of balancing services (Bogensperger et al. 2018). The digital transformation of regulatory processes and the necessary detailed review of compliance with regulatory guidelines require a great deal of time and effort, and involved authorities must first appreciate the added value of the platform before the use cases can be implemented in practice.

To implement the proposed digital use cases, these authorities must recognize the blockchain-based proof of data integrity in their guidelines and must be able to perform the necessary verification based on the provided data and MPs.

Governance
Furthermore, questions about blockchain governance prevail. This applies to both blockchains used for SSI as well as the blockchain infrastructure used for the storage of Merkle roots. Following Beck et al. (2018), stakeholders should agree on the dimensions of decision rights, accountability, and incentive structures. For example, the operation of nodes might not be efficient or feasible for small-scale stakeholders. However, asset owners will still need reading and writing access. In contrast, economic benefits caused by high reliability of underlying data might serve as a sufficient incentive for larger players to operate nodes themselves. Further design choices can be configured in dependence of stakeholders’ interests. In general, there are various design choices of how to process Merkle trees and store them on the blockchain. Moreover, alternative design choices with regards to the creation of Merkle trees need to be considered. For example, larger sizes of data can be committed by recursively creating Merkle trees. Nevertheless, the increase in commitment sizes comes with downsides regarding the efficiency and privacy of verification. Thus, before transforming our prototype to a productive solution, further governance- and design-related questions should be considered. Advantages and drawbacks should be carefully balanced off. In general, computations should be performed solely off-chain while blockchains are used for tamper-resistant documentation purposes only.

User acceptance
Lastly, for successful adoption of emergent technologies, user acceptance studies are crucial (Ostern 2018). This is especially so as the underlying context involves various different stakeholder groups. Thus, in order to successfully implement the proposed solution, stakeholders will need to be informed about the processing and mechanisms of the above-mentioned architecture. For example, workshop series and on-site training allow to showcase the inner workings of systems and the adherence to requirements, which was previously shown to be particularly important for acceptance (Ostern 2018).

Conclusion and outlook
Our research in the field of asset logging in the energy sector highlighted that many use cases offer both business relevance and scientific significance. With regards to the identified use cases and associated requirements, we draw the following conclusions: First, the future market potential for asset logging is high due to an already signifi-
cant, but also expected increase in the number of relevant energy assets. Second, the prospects of cost reduction and improved data protection represent the main drivers for stakeholders involved in the field of asset logging. Third, the applied methodical process of exchange with partner companies allowed us to identify eight asset logging use cases as highly relevant in the energy sector both from a business and a research perspective. Regarding the technical solution for realizing most asset logging use cases, we found that they must enable an ex-post verification of data integrity, while simultaneously protecting relevant business secrets and allowing for scalability. Furthermore, since most use cases require different types of data, a key challenge represents the collection of signed data that is tamper-resistant from the moment it is generated. Last, the lack of digitalization in the energy sector poses a considerable obstacle to swift implementation.

We presented a platform architecture as a suitable technical solution for implementing asset logging use cases. Any asset that can provide signed data can in principle be connected to the platform, which therefore enables the implementation of a range of use cases. The interplay of blockchain, Merkle proofs, and E2E-encrypted communication channels guarantees traceability, data integrity, and privacy for all participating stakeholders. Finally, the architecture is designed for cost-effective implementation and scaling, i.e. it is extendable to other use cases and stakeholders.

We selected two use cases – namely warranty management and regulatory requirements – to be implemented in a sandbox approach. These field trials are intended to demonstrate the functionality and acceptability of the solution to pave the way for implementing asset logging use cases at a larger scale in the future. This is especially so if we succeed in demonstrating the added value to relevant stakeholders including public authorities and regulators, which may increase participation in and widespread acceptance of our solution. To achieve this objective, complementing the SMGW infrastructure with tamper-resistant data collection processes for other types of data represents the key necessity to be addressed in the near future. Furthermore, our research highlights the importance of a widespread roll-out of reliable and certified data measurement infrastructures such as SMGWs. Regarding future research opportunities, we propose a detailed analysis of market potential, feedback effects on the energy system, and synergies in terms of use cases and collected data. Furthermore, future research should address how the proposed data infrastructure can be used for the processing of additional business logic. For example, our proposed architecture can build the foundation for processing subsequent warranty conditions. Also, as the proposed infrastructure enables verifiable data integrity, previously proposed use cases for further data processing and disbursement such as suggested in Mohanta et al. (2018) become feasible.

Acknowledgements
We thank Nicolas Ruhland for his help with implementing the network of aries-cloudagents, as well as Anna Walter and Marvin Ehaus for their help with bringing the implementation to a productive environment. We would also like to thank Nils Lütbach and Jens Strüker for their scientific guidance and our research partner Stiftung Umweltenergie recht as well as our partner companies Bayernwerk, BayWa re., E-Werk Schweiger, innogy/E.ON, Lechwerke, RheinEnergie, SMA, Stadtwerke Bad Tölz, Stadtwerke Dachau, Thüga, TransnetBW and VWEW for supporting the project.

About this supplement
This article has been published as part of Energy Informatics Volume 4 Supplement 3, 2021: Proceedings of the 10th DACH+ Conference on Energy Informatics. The full contents of the supplement are available online at https://energyinformatics.springeropen.com/articles/supplements/volume-4-supplement-3.
Authors’ contributions
AD, PD and MH contributed to the energy-specific chapters of the paper including introduction, use case assessment, and challenges. BS, JS, FV and LW provided the technology-specific chapters including the technical background, architecture, and challenges. All authors read and approved the final manuscript.

Funding
The project InDEED is funded by the Federal Ministry for Economic Affairs and Energy (BMWi) based on a resolution of the German Bundestag (funding code: 03E16026A). Publication funding was provided by the German Federal Ministry for Economic Affairs and Energy.

Availability of data and materials
There is no additional data and materials.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
1FfE, Am Blütenanger 71, 80995 München, GER. 2Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Wittelsbachererring 10, 95444 Bayreuth, GER. 3Research Center Finance & Information Management, Universitätsstraße 12, 86159 Augsburg, GER. 4University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, GER. 5University of Augsburg, Universitätsstraße 2, 86159 Augsburg, GER.

Published: 13 September 2021

References
Alaton C, Tounquet F (2020) Benchmarking Smart Metering Deployment in the EU-28. https://op.europa.eu/o/opportal-service/download-handler?identifier=b397ef73-698f-11ea-b735-01a77ed71a1&format=pdf&language=en&productionSystem=cellar&part=. Accessed 07 May 2021
Albrecht S, Reichert S, Schmid J, Strüker J, Neumann D, Fridgen G (2018) Dynamics of Blockchain Implementation - A Case Study From The Energy Sector. In: Proceedings of the 51st Hawaii International Conference on System Sciences. pp. 3527–3536. https://doi.org/10.24251/hicss.2018.446
Ammend J, Kaiser J, Uhlig L, Urbach N, Volter F (2021) What Do We Really Need? A Systematic Literature Review of the Requirements for Blockchain-based E-government Services. In: Wirtschaftsinformatik 2021 Proceedings. AIS, Duisburg
Andoni M, Robu V, Flynn D, Abram S, Geach D, Jenkins D, McCallum P, Peacock A (2019) Blockchain Technology in the Energy Sector: A Systematic Review of Challenges and Opportunities. Renew Sust Energ Rev 100:143–174
Avellaneda O, Bachmann A, Barbir A, Brenan J, Dingle P, Duffy KH, Maler E, Reed D, Sporny M (2019) Decentralized Identity: Where Did it Come From and Where is it Going? IEEE Communs Stand Mag 3(4):10–13
Balzer G, Schorn C (2014) Asset Management Für Infrastrukturverwaltungs – Energie und Wassert, 2nd ed (VDI-Buch). Springer Vieweg, Berlin/Heidelberg
Bao J, He D, Luo M, Choo KR (2020) A Survey of Blockchain Applications in the Energy Sector. IEEE Syst J1–12. https://doi.org/10.1109/jsyst.2020.2998791
BDEW Bundesverband der Energie- und Wasserwirtschaft e.V. (2020) Digital@EVU 2020: Unternehmen planen höhere Investitionen in Digitalisierung. https://www.bdew.de/service/publikationen/digital-evu-2020-unternehmen-planen-hohere-investitionen-in-digitalisierung/, Accessed 07 May 2021
Beck R, Czepuch JS, Lolike N, Malone S (2016) Blockchain — The Gateway to Trustfree Cryptographic Transactions. In: 24th European Conference on Information Systems. AIS, Istanbul
Beck R, Müller-Bloch C, King JL (2018) Governance in the Blockchain Economy: A Framework And Research Agenda. J Assoc Inf Syst 19(10):1020–1034. https://doi.org/10.17705/1jais.00518
Bernal Bernabe J, Canovas JL, Hernandez-Ramos JL, Torres Moreno R, Skarmeta A (2019) Privacy-Preserving Solutions for Blockchain: Review and Challenges. IEEE Access 7:164908–164940
Bogensperger A, Zetelmaier A, Hinterstocker M, Duffter C (2018) Blockchain – Chances for the Transformation of our Energy System. Report Section: Use Cases
Bundesamt für Sicherheit in der Informationstechnik (BSI) (2018) Das Smart-Meter-Gateway. Bonn. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Broschueren/Smart-Meter-Gateway.pdf?__blob=publicationFile&v=1. Accessed 07 May 2021
Bundesamt für Sicherheit und Informationstechnik (2013) Technische Richtlinie BSI TR-03109-1 – Anforderungen an die Interoperabilität der Kommunikationsanlage eines intelligentsen Messsystems. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeRichtlinien/TR03109/TR03109-1.pdf?__blob=publicationFile&v=1. Accessed 07 May 2021
Bundesministerium für Umwelt Naturschutz und Klima Protection (BMU) (2020) Förderrichtlinien für Energiewirtschaft. https://www.bmu.de/jetzt/nachhaltig/energene/wirtschaft/energie-energywirtschaft/foerderung/, Accessed 07 May 2021
Bundesministerium für Wirtschaft und Energie (BMWi) (2020) Digitalisierung der Wirtschaft in Deutschland – Langfassung eines Ergebnisses im Projekt “Entwicklung und Messung der Digitalisierung der Wirtschaft am Standort Deutschland”. In: Digitalisierungssicherheit. Bundesministerium für Wirtschaft und Energie, Berlin
Buterin V, et al. (2014) A Next-Generation Smart Contract and Decentralized Application Platform. https://github.com/ethereum/wiki/wiki/White-Paper. Accessed 07 May 2021
van Bokkem D, Hageman R, Koning G, Nguyen L, Zarin N (2019) Self-sovereign Identity Solutions: The Necessity of Blockchain Technology. http://arxiv.org/abs/1904.12816. Accessed 07 May 2021
Wang F, De Filippi P (2020) Self-Sovereign Identity in a Globalized World: Credentials-Based Identity Systems as a Driver for Economic Inclusion. Front Blockchain 2:28
Wüst K, Gervais A (2018) Do You Need a Blockchain? In: Crypto Valley Conference on Blockchain Technology. IEEE. pp 45–54. https://doi.org/10.1109/cvcbt.2018.00011
Xiao Y, Zhang N, Lou W, Hou YT (2020) A Survey of Distributed Consensus Protocols for Blockchain Networks. IEEE Commun Surv Tutorials 22(2):1432–1465
Zeiselmair A, Hinterstocker M, Bogensperger A, von Roon S (2019) Asset Logging – Transparent Documentation of Asset Data Using a Decentralized Platform. In: 8th DACH Conference on Energy Informatics. Springer, Salzburg
Zhang R, Xue R, Liu L (2019) Security and Privacy on Blockchain. ACM Comput Surv 52(3):1–34

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.