Blockchains for Achieving Data Awareness and Enabling Data Sharing

Robert Siegfried, Torsten Müller
Aditerna
Riemerling/Munich, Germany
robert.siegfried@aditerna.de, torsten.mueller@aditerna.de

Mario Pehla
German Air Force
Berlin, Germany
MarioPehla@bundeswehr.org

Hilmar Holland
German Air Force
Cologne, Germany
HilmarHolland@bundeswehr.org

ABSTRACT

Defense organizations worldwide are faced with growing volumes of data, provided by a variety of sources and transferred via multiple channels to the operational users. While amazing data generation techniques are introduced (to include autonomous sensors and unmanned platforms), data formats are all but standardized, data transfer often involves significant manual efforts (like sending hard drives, faxes or Email), and information about available data (to include basic metadata, but also background information and trust labels) is lacking too often. Improving Defense-wide data awareness, i.e., across organizational boundaries, and simplifying and accelerating data sharing are critical factors to succeed in rapidly changing future operating environments.

Distributed ledger technology (aka blockchain) has gained significant attention with the rise of crypto currencies like Bitcoin. This paper explores how blockchains can be leveraged to enable data sharing while at the same time providing a high degree of information security to achieve data awareness in complex organizations.

Specifically, this paper presents two use cases: Firstly, using blockchains to create situational awareness about simulation resources owned and managed by a variety of organizations. This use case picks up recent M&S as a Service (MSaaS) efforts that aim at providing users a federated discovery toolset for simulation resources across organizations and nations. Secondly, using blockchains to improve mission data provisioning for military air platforms. Both use cases are illustrated with practical experiences from the MSaaS efforts of the NATO Modelling and Simulation Group (NMSG) and recent German Air Force efforts to improve mission data provisioning.

ABOUT THE AUTHORS

Dr. Robert Siegfried is a Senior M&S Consultant and Managing Director of Aditerna. Robert is serving as Chair of the NATO Modelling and Simulation Group (NMSG) and leads the NMSG's efforts towards Modelling and Simulation as a Service (MSaaS). Robert is a member of the Executive Committee of the Simulation Interoperability Standards Organization (SISO) and is actively engaged in multiple SISO working groups.

Torsten Müller is Managing Director of Aditerna and certified project manager with 15+ years of experience in various international projects and is a recognized expert for IT project management in traditional and agile project environments. He has successfully led various projects in different domains, such as Telecommunications, Transport, and Defense.

LTC Mario Pehla is a staff officer in the Plans and Policy Branch of the HQ German Air Force. He served in various positions in higher national commands and also in NATO's HQ SACT. He leads numerous activities with IT/cyber development background as an IT expert and currently advises on the development of the Combat Cloud in the tri-national Next Generation Weapon System project.

CAPT Hilmar Holland is an IT officer at the German Air Force Forces Command. In his 30+ years of military service, Hilmar was involved in various projects with interoperability aspects for land-based and flying weapon systems. For example, Hilmar was member of Apollo Working Group (NLD/DEU) and the TOIWG (Three Nation Interoperability Working Group) of USA/NLD/DEU.

Blockchains for Achieving Data Awareness and Enabling Data Sharing

Robert Siegfried, Torsten Müller
Aditerna
Riemerling/Munich, Germany
robert.siegfried@aditerna.de, torsten.mueller@aditerna.de

Mario Pehla
German Air Force
Berlin, Germany
MarioPehla@bundeswehr.org

Hilmar Holland
German Air Force
Cologne, Germany
HilmarHolland@bundeswehr.org

1. IMPORTANCE OF DATA AWARENESS AND DATA SHARING

"Data is the new oil", who has not heard this phrase before? However, what exactly does it mean? Obviously not every organization with lots of data is per se rich. Stressing the analogy between oil and data further, we run into terms like "oil production" and "data mining", meaning just having oil and data is not sufficient. Both must be processed further to create value.

Gartner estimated in 2017 that by 2021 75% of all organizations have appointed a Chief Data Officer (Richardson, Moran, Logan, Edjlali, & Faria, 2017). This shows that digitalization strategies and data-centric structures are enforced throughout most industries. Growing volumes of data offer new opportunities and possibilities for organizations, provided that the data is not kept in silos but used and analyzed across the organization or even shared with partners.

Defense organizations worldwide are faced as well with growing volumes of data, provided by a variety of sources and transferred via multiple channels to the operational users. While amazing data generation techniques are developed and introduced (to include, for example, autonomous systems, intelligent sensors and unmanned platforms), data formats are all but standardized, data transfer often involves significant manual efforts (like sending hard drives, faxes or Email), and information about available data (to include basic metadata, but also background information and trust labels) is lacking too often. Improving Defense-wide data awareness, i.e., across organizational boundaries, and simplifying and accelerating data sharing are critical factors to succeed in rapidly changing future operating environments.

This paper explores the challenges connected to "Data Awareness" and "Data Sharing", and how Distributed Ledger Technology (DLT), more specifically blockchains, can be used to address these challenges. For this paper, we use the terms in the following meaning:

- "Data Awareness" refers to visibility and insight about data, preferably in real-time and regardless of file format or storage location.
- "Data sharing" refers to "the collection of practices, technologies, cultural elements and legal frameworks that are relevant to transactions in any kind of information digitally, between different kinds of organisations." (European Commission, 2021)

This paper presents two specific use cases and investigates how blockchains can be utilized. The first use case concerns the sharing of simulation resources like 3-d models, datasets, etc. across organizational, and potentially national, boundaries. The second use case concerns the provisioning of mission data in military aviation. In both use cases, data sharing and the associated data awareness are key to efficient and effective use of resources, and ultimately for efficient mission execution.

2. DISTRIBUTED LEDGER TECHNOLOGY

2.1 Overview And General Principles

A distributed ledger is, simply speaking, a special case of a distributed database that may be spread across multiple organizations, sites, or countries, and is typically public. (UK Government Office for Science, 2016) As such, Distributed Ledger Technology (DLT) is tightly coupled to distributed systems theory and opens a whole field of research. Figure 1 illustrates this and emphasizes that blockchains are a specific instantiation of DLT.

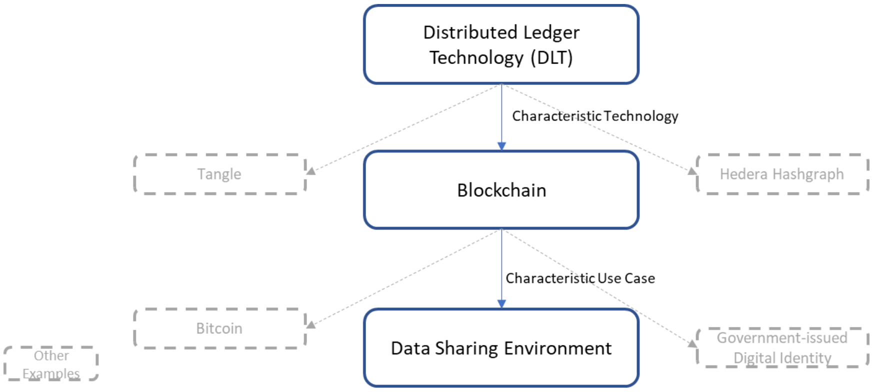

Figure 1: Blockchains are a subset of Distributed Ledger Technology (DLT).

Blockchains themselves come in different flavors with different characteristics, depending on the use case. There are permissioned blockchains, permission-less blockchains, blockchains based on proof-of-work, based on proof-of-stake", based on "proof of authorization" and other options. The actual setup and configuration of a blockchain is always depending on the specific use case.

2.2 Blockchain, Blocks, Transactions, And Nodes

Blockchains grow over time as new blocks with new information are added. The blocks carry the actual data along with hash values to ensure data integrity. The link between two blocks is created by including the hash value of the previous block when generating the hash value for the current block. This way, a chain of interconnected blocks is created, the blockchain (see Figure 2).

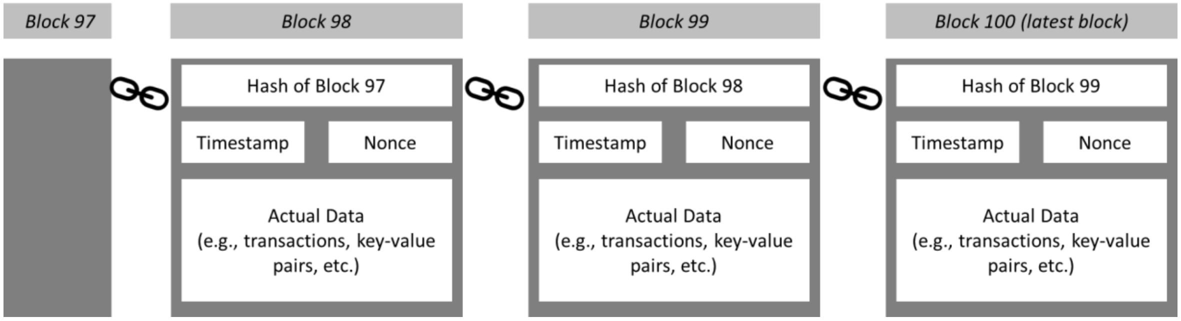
Block 97
Block 98
Block 99
Block 100 (latest block)
Hash of Block 97
Timestamp
Nonce
Actual Data (e.g., transactions, key-value pairs, etc.)
Hash of Block 98
Timestamp
Nonce
Actual Data (e.g., transactions, key-value pairs, etc.)
Hash of Block 99
Timestamp
Nonce
Actual Data (e.g., transactions, key-value pairs, etc.)

Figure 2: Illustration of basic principles of a blockchain.

New data (e.g., financial transactions) are added to a new block in the blockchain. Many blockchains use a consensus mechanism where the participating nodes in the blockchain network validate new blocks e.g., by checking if the originator of the new block is allowed to add blocks, and if the block adheres to the blockchains standard protocol. There are different technical ways for this validation. Some blockchains let nodes simulate the addition of a new block to a blockchain, and only if the simulated output of all nodes is the same and adheres to the standards, the originator node is authorized to add the new block to the blockchain.

The blockchain is not stored on a centralized server, but a copy of it is spread out over many servers (usually called "nodes" in this context) which participate in the blockchain network. Each copy of the blockchain can be locally validated (by each node itself) and is kept in sync with each new block added (by whatever node in the entire blockchain network). This makes the information available redundantly since a (full) copy of the blockchain is available on all nodes. Additionally, each node validates newly added blocks using a consensus mechanism. This way, an attempt to tamper the content on the blockchain is easily detected. Changing a block on only one node of the blockchain network is not sufficient to change the information throughout the whole blockchain network.

Additionally, modified (tampered) blocks will not synchronize throughout the blockchain network because the tampered block is easily identified as it violates the interconnected chain of hashed blocks and therefore is not accepted by other nodes.

2.3 Permission To Participate

Blockchains offer different ways to allow participants to take part. Obvious blockchains coming to mind are decentralized permission-less blockchains with thousands of nodes, as they are used for crypto currencies like Bitcoin. Anyone can participate in such blockchain networks, and anyone can verify the transactions on the blockchain (also called "passive participation") since the ledger is publicly accessible. For permission-less blockchains, the creation of new blocks ("active participation") is only possible with a certain "status", which is proven by e.g., work (Proof of Work (PoW) as in "mining" activities as for e.g., Bitcoin, or stake (Proof of Stake (PoS), as in holding a certain number of coins of a specific crypto currency, as for e.g. BNB or Dash.

Permissioned blockchains on the other hand have different ways to exclude non-authorized users from even a passive participation. One of the most common ways to accomplish this is to use a blockchain based on the "Proof of Authority" (PoA) approach. This type of blockchain is usually less centralized, using less nodes, as permission-less blockchains. They provide the option to employ authorization rules, for example via special nodes which act as Membership Service Provider (MSP). This permission-based approach enables blockchains to become part of industry solutions, for example supporting production or logistics-oriented use cases, where the information of the blockchain should only be visible to a limited group of participating members. It is important to understand that also participants of a permission-based blockchain can have different levels of authorization, e.g., distinguishing between seeing only specific blocks of the blockchain, seeing only specific information within the blocks, seeing all blocks with all information, being allowed to add new blocks etc.

2.4 Hashes And Blocks

The blocks in a blockchain are secured using hash algorithms. In general, different hash algorithms are available, like MD (Message Digest) 4, MD5, SHA (Secure Hash Algorithm)-1 or SHA-256. Hash algorithms are mathematical functions that turn an input into an output while guaranteeing certain characteristics of the output (the actual "hash value"). (Roshdy, Fouad, & Aboul-Dahab, 2013) Main characteristics of hash algorithms are:

- The same input leads always to the same output, i.e. hash algorithms are deterministic.
- Small changes in the input should lead to a significantly different output (hash value). The actual hash values should be uniformly distributed across the value space as to minimize collisions (where different inputs are mapped to the same hash value).
- The hash algorithm is (practically) irreversible, meaning a conclusion from the hash value (output) to the input is impossible, even knowing which algorithm has been used. Hash algorithms are essentially one-way functions.

In the context of blockchains, hash values are used to create persistent (essentially unmodifiable) links between the blocks on the blockchain.

2.5 Benefits Of Blockchain Technology

Simply speaking, a blockchain is a decentralized database where each block is linked to the previous block through hash values of the actual data in the individual blocks. The blockchain is redundantly stored on the nodes within the blockchain network, thereby supporting the availability of the information to all authorized participants. The integrity of information written into the blocks of the blockchain is secured by hash values. The blocks are interconnected with each other by including the hash value of the previous block into the computation of the hash of the current block, which ensures the integrity of the information on the blockchain. (Shah, 2019)

This means, that a block that has been added to the blockchain cannot be changed or removed by just one party. The majority or even the whole blockchain network (depending on the specific implementation) would have to approve such a step. This leads to a tamper-proof data structure containing non-disputable data.

Obviously, a blockchain is not a comprehensive information security solution. Blockchains should be understood as providing useful functionalities as a part of the information security toolbox, leveraging blockchain characteristics as:

- Decentralization
- Chain-linking
- Consensus mechanism
- Tamper-proofness
- Transparency within the blockchain network, yet confidentiality to the outside
- Nonrepudiation (Non-disputable)

However, blockchain provides additional benefits which are not purely information security-related.

Integrity
Availability
Information Security
Confidentiality

Figure 3: Fundamental aspects of information security.

Blockchain technology can leverage its strength and benefits very well in a distributed setup. Whenever (partner) organizations come together for a joint endeavor, there is often a need to share data. Sharing data, even with partners is still a challenge for many organizations. Blockchain provides a layer that is transparent for all participants and addresses the three major aspects of information security (see Figure 3), confidentiality, availability, and integrity of data (Ali & Afzal, 2018). Publishing only the metadata of available data on the blockchain enables data awareness and controlled data sharing. This approach provides an overview of available data to all participants. The access to the data itself can still be limited, e.g., to a "need-to-know" principle.

2.6 Drawbacks And Challenges Of Blockchain Technology

The complexity of blockchain technology and the need to operate multiple nodes with redundant data are some of the drawbacks of using blockchains. So, blockchain technology comes with a price tag. Realistically, it is more complex to develop and run a blockchain network in comparison to a common (potentially distributed) database. Hence, before introducing blockchain technology, the decision makers should validate the benefit of blockchain technology for the specific use case at hand. There must be solid evidence that the use of blockchain brings in benefits that outweigh its drawbacks and challenges.

There are several questionnaires and checklists available to make a case for or against blockchain usage. At least one of these or a self-drafted checklist should be considered and worked through before the decision for a blockchain implementation is done. Without the claim that this is the best source to use, having a look at the White Paper "Blockchain beyond the Hype – A Practical Framework for Business Leaders" published by the World Economic Forum in April 2018 (Mulligan, Zhu Scott, Warren, & Rangaswami, 2018) might be a good starting point for validating the decision for or against the usage of blockchain technology. We recommend considering other sources as well.

3. USE CASE 1: BLOCKCHAIN FOR FEDERATED M&S RESOURCE CATALOGUES

3.1 Statement Of Need

Military forces recognize that efficient and flexible utilization of simulation is a critical factor in sustaining an advantage over our adversaries. Recent developments in cloud computing technology and modular open system architectures offer opportunities to address these critical needs. Specifically, as Modelling and Simulation as a Service (MSaaS) capabilities mature, emerging software delivery combines modular orientation with an as-a-service delivery model of cloud computing to enable more composable simulation environments that can be deployed, adapted, and executed on-demand. The MSaaS paradigm supports stand-alone use as well as integration of simulated and real systems into a unified cloud-based simulation environment when and as needed to support the warfighter.

The MSaaS efforts of the NATO Modelling and Simulation Group (NMSG) (NATO STO, October 2018), Air Force Agency for Modeling & Simulation (AFAMS) Air Missile Space Training Environment (AMSTE) and the Office of the Secretary of Defense (OSD) Advanced Distributed Learning (ADL) Total Learning Architecture (TLA) have highlighted the improved cataloguing, sharing and re-use of models, data, and simulation resources as a key underpinning element. The requirements are to catalogue, store, search for and retrieve a wide range of M&S resources developed (and owned) by various organizations (potentially across several nations). These resources range from individual models to elements of simulation data such as terrain databases, component models, to run-time simulation services. The process of searching for, finding, filtering, and retrieving resources is known as 'discovery' and is highly dependent upon the underpinning information model used to describe simulation resources. The "Allied Framework for MSaaS" defines the MSaaS Portal as the user-facing application providing the discovery functionalities and serving as the front-door to advanced functionalities like on-demand deployment and execution. (NATO STO, 2019)

The vision of MSaaS within NATO context is to establish an MSaaS ecosystem, where MSaaS Capabilities from different nations are federated, see Figure 4. An MSaaS ecosystem is a network of organizations with MSaaS Capabilities that drives the creation, delivery and use of M&S services in NATO context. An MSaaS Capability consists of operational capabilities as well as technical capabilities needed to achieve this.

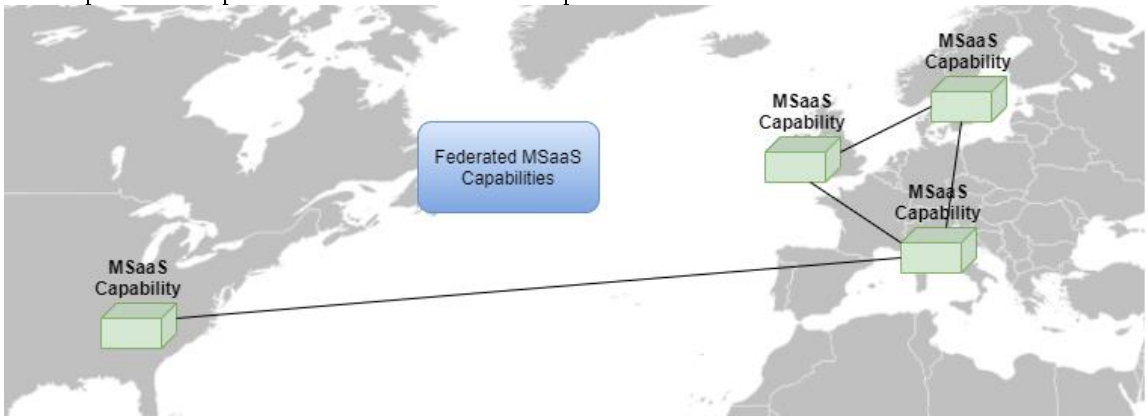

Figure 4: Federated MSaaS Capabilities. (NATO STO, 2021)

3.2 Federated Resource Catalogues

A key element of the "Allied Framework for MSaaS" is the idea of an "MSaaS Portal" that serves as an entry point for users to discover and assess available M&S resources. Figure 5 illustrates the idea of federated registries and repositories. In this context, "registry" refers to information systems that store, manage, and provide metadata about M&S resources while "repository" refers to information systems that hold the actual resources.

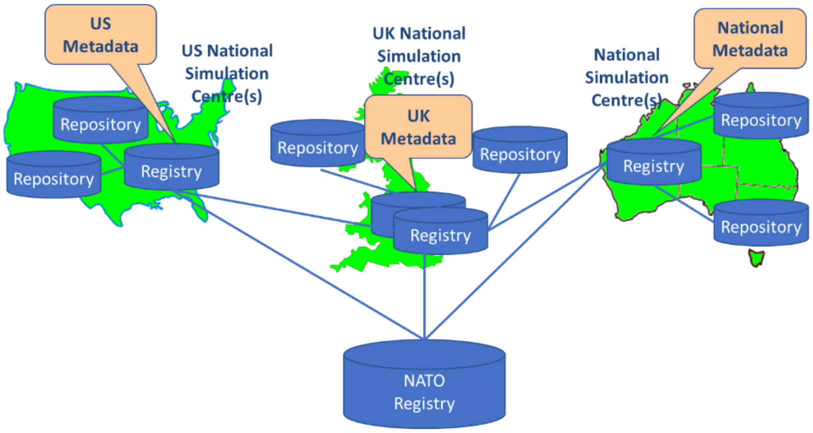

Figure 5: Example of Federated Registries. (NATO STO, 2019)

While Figure 5 illustrates the use case of federating registries (and repositories) of multiple nations, this use case could also apply to federating registries of multiple organizations within a single nation. The underlying setting and objective remain the same: Multiple stakeholders operate local registries (holding information about M&S resources that fall under their authority or ownership) and want to share some or all their information (about their M&S resources) with other stakeholders for the greater objective of achieving situational awareness about existing M&S resources in the wider community of interest.

This use case is specifically investigating the federation of registries, i.e., the exchange of metadata between stakeholders. Knowledge about existing M&S resources in the wider ecosystem is the fundamental requirement for potentially using those resources. Actual access to non-local M&S resources (i.e., those resources that are provided by another stakeholder) may require additional steps, like agreeing on licensing terms, usage conditions etc.

3.3 Information Exchange Requirements

Achieving the objective of creating a global situational awareness about existing M&S resources requires information exchange between the participating stakeholders respectively the federated (i.e., networked) registries. Figure 6 illustrates the two key elements required: a common data model (defining 'what' information is exchanged) and an agreed data exchange protocol (defining 'how' the information is exchanged).

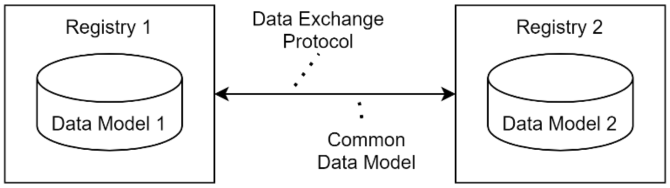

Figure 6: Common Data Model and Data Exchange Protocol are required.

As the underlying assumption of federated M&S registries is to connect registries from various stakeholders, proprietary data models that are not available to all stakeholders are obviously not a viable option. To enable wide adoption by many stakeholders, openness and free availability of the common data model is essential.

For the investigated domain (i.e., exchange of information about M&S resources), the "Modeling and Simulation (M&S) Community of Interest (COI) Discovery Metadata Specification" (MSC-DMS) provides a de facto open standard for describing M&S resources. (Department of Defense (DoD) Modeling and Simulation Coordination Office (MSCO), 12 July 2012) While the MSC-DMS specification is publicly available, the specification itself was developed in a rather closed process and makes references to a set of US-specific metadata specifications. To address lessons learned since developing MSC-DMS and to resolve issues with references to US-specific metadata standards, thereby making it more appealing to use for non-US stakeholders, the Simulation Interoperability Standards Organization (SISO) has established a Product Development Group (PDG) to develop the "Discovery Metadata Specification for M&S Resources" (DMS-MSR). (Simulation Interoperability Standards Organization (SISO), 7 Feb 2020) The DMS-MSR will be based on the MSC-DMS and is developed in an open fashion, and will be published as an open standard.

The NATO Consultation, Command and Control Board (C3B) is currently developing the "NATO Core Metadata Specification" (NCMS) as a metadata standard for a wide variety of resources. The NCMS might potentially be a good candidate for exchanging information about M&S resources, but is, as of yet, not published as an open standard.

Currently, the MSC-DMS is the best option for describing M&S resources in terms of maturity of the standard and availability. Once development of DMS-MSR is completed and the standard is published, the authors expect that DMS-MSR quickly takes over this role.

It is emphasized (again) that this use case is about exchanging metadata about M&S resources, but not about the exchange of actual M&S resources like models, simulations, databases, etc. The metadata would nevertheless include information how to access the actual resource. In the easiest form, this might simply be a pointer (hyperlink) to the actual resource. Depending on the resource (size, accessibility, constraints etc.), other information may be required.

The second element needed to exchange information between two (or more) registries is the data exchange protocol that defines how the information is exchanged. To date, no standards exist addressing this issue. The MSaaS working group of the NMSG is currently developing a "Resource Discovery Specification" to specifically address this issue. Initial tests have been successful, but completion of development and publication will take some time and is not expected before late 2022.

3.4 Challenges And Constraints

Key challenges for establishing a federation of M&S registries are:

- Multiple stakeholders with various trust relations: By definition, multiple stakeholders are participating in a federated M&S catalogue. While some stakeholders may have established and intimate trust relations, this cannot be assumed for all stakeholders and all relations. Therefore, any approach to information sharing has to take this into account.
- Local authority versus global situational awareness: Each stakeholder is responsible for the content in its own registry and usually reserves its rights to modify and manage the respective content. Similarly, for a variety of reasons, stakeholders may not be in a position to share all information, but need to be able to decide to share a portion of their local content with the wider community. Despite local control over the content, the overall objective is to achieve a global situational awareness.
- Credibility of information: Depending on the importance and criticality of the shared M&S resources (like models or simulations), the credibility and validity of shared information must be assured. This includes integrity checks to ensure that information was not modified, changed, or corrupted due to technical failures, operating error or deliberate actions.
- Key constraints for the design of a federated registry are: The variety of stakeholders and the absence of a single, centralized node that could act as a trustworthy master disallows the use of a centralized registry that could act as a hub for all information exchanges.
- The exchanged data volume is comparatively low as only metadata is exchanged. Individual M&S resources, even when represented in MSC-DMS compliant XML, are usually only a few Kilobytes large. Most registries observed today hold between a few hundred to a few thousand M&S resources. Thus, even full duplication of information between two registries is not posing substantial challenges in terms of data volume to be transferred, processed and stored.
- The requirements regarding latency of the information exchange are moderate. Update cycles for M&S resources are usually measured in days to weeks (depending on development speed of a resource and associated required changes to the metadata), thus synchronization of federated registries is not time critical.

3.5 Use Of Blockchains

The key idea is to use a blockchain as decentralized database to store the metadata about all M&S resources published by the participating stakeholders (see Figure 7).

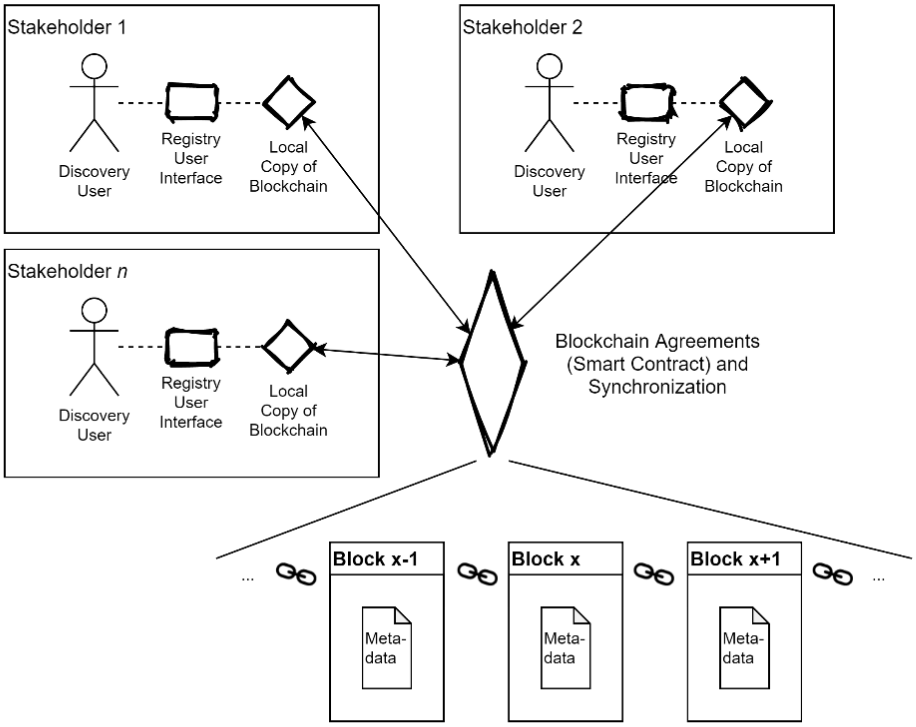

Figure 7: Blockchain for secure data exchange between M&S registries.

The individual blocks of the blockchain would contain transactions related to the publication of metadata for a specific M&S resource. Figure 8 shows a simplified Smart Contract defining allowed transactions. Essentially, the participating stakeholders can publish M&S resources (or, more accurately, metadata about M&S resources), edit/modify M&S resources and invalidate M&S resources.

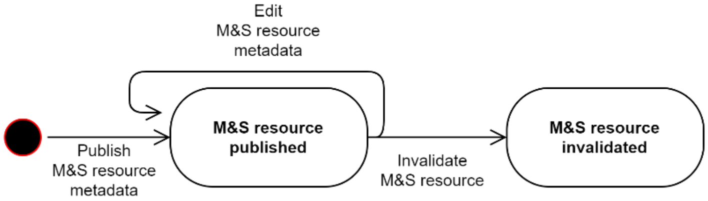

Figure 8: Simplified Smart Contract (State Diagram) for M&S resources metadata.

The transaction content (see Figure 7) contains the state transition plus the actual metadata, e.g., provided as MSC-DMS compliant XML snippet.

3.6 Anticipated Benefits And Open Issues

The proposed approach of using a blockchain as decentralized database for realizing a federated M&S registry provides the following benefits:

1. Stakeholders can share information about individual M&S resources, thereby establishing a global situational awareness about existing M&S resources.
2. Each participating stakeholder can verify the integrity of the entire dataset (blockchain). No central element of trust (like a centralized database, or trusted partner) is required.
3. All modifications of metadata are stored in a traceable and auditable way.
4. The Smart Contract (i.e., the allowed blockchain modifications) is the binding element between all stakeholders. Individually each stakeholder may use its own registry implementation, metadata format, client software, user interface etc. Thereby, this approach gives participating stakeholders a large degree of freedom

with regards to their individual registry implementations. Given the diversity of stakeholders and, in general, the lack of an overarching authority that could mandate a specific implementation, this aspect is critical to enable wide-spread adoption.

By default, each stakeholder is allowed to access (i.e., view) the entire blockchain. In situations where information sharing between different communities is required (e.g., certain resources are shared nationally, while other are shared with NATO partners, or with Five Eyes nations) a single blockchain as illustrated here may not suffice. An obvious approach would be the use of multiple blockchains for individual communities. Also, specific blockchain implementations offer options for satisfying this requirement. For example, Hyperledger Fabric allows to segregate the blockchain into so-called ‘channels’, where each channel represents a subset of participants that are authorized to see the data for the Smart Contract that is deployed to that channel. (Hyperledger, 2021)

4. USE CASE 2: BLOCKCHAIN FOR AVIONICS MISSION DATA PROVISIONING

4.1 Background Information

The provisioning of mission data for aerial platforms is a second use case for which we investigate the benefits of using blockchain technology. The authors are participating in a research effort that is investigating options to improve the existing mission planning processes of the German Air Force, specifically improving the data flow end-to-end, from data sources to the mission planning systems.

Current processes are characterized by a high degree of manual labor to verify data and transfer data between multiple, disconnected information systems and networks. The objective is to accelerate the data flow, while providing higher data quality and better usability for end users. The scope of the research includes identifying quick wins for current challenges in the mission planning process and providing long-term considerations related to upcoming technologies and features in future weapon systems, including inter-linked airborne platforms, and the accompanying new requirements for mission data provisioning and the mission planning process.

4.2 Mission Data Provisioning

Gathering data from many different data sources and organizations is a complex and time-consuming task. The process of mission data gathering needs to unite many different processes and different ways of handling data. Seen from the perspective of each individual data source, each process and the dedicated tools used are often well-aligned with the individual domain’s needs and purposes. However, the overall mission planning process is broad and comprehensive. From this superordinate view, the mission planner (i.e., the actual user of the mission data) looks at a vast number of mostly unaligned processes and tools. For an appropriate level of comprehensive data awareness, the organization needs to unite all kinds of different domains and their processes under one common roof without trying to squeeze them into just one generic framework and ignoring all the specific restrictions and necessities.

Usually, mission data must be available within a short timeframe. Long preparation times for mission data are increasing the overall preparation time to warfighter readiness. The length of this overall preparation time is a crucial part in today’s fast evolving operational environments. A highly sophisticated and expensive weapon platform like today’s modern fighter jets is of no use as long as the jets are grounded during the preparation process for a mission. Reducing preparation times is essential for a timely response to new situations and threats.

The challenge at hand is to produce a quick and comprehensive overview of available mission data. A smooth process to request and receive needed, but not yet available, mission data is essential for the reduction of lead times between mission statement and mission execution.

Data provisioning is currently done by several different source systems, using different processes due to requirements from the different domains. The processes are often strictly separated from each other and not aligned at all (see Figure 9). The involved parties are, depending on the domain, most of the times not aware of the other processes of data provisioning ongoing at the same time. There is a high potential to leverage benefits in these separate processes when making data sharing easier and more transparent. Sharing data is a necessity in mission planning, while the different security levels must be adhered. Additionally, awareness of already available data can prevent stakeholders from doing double work.

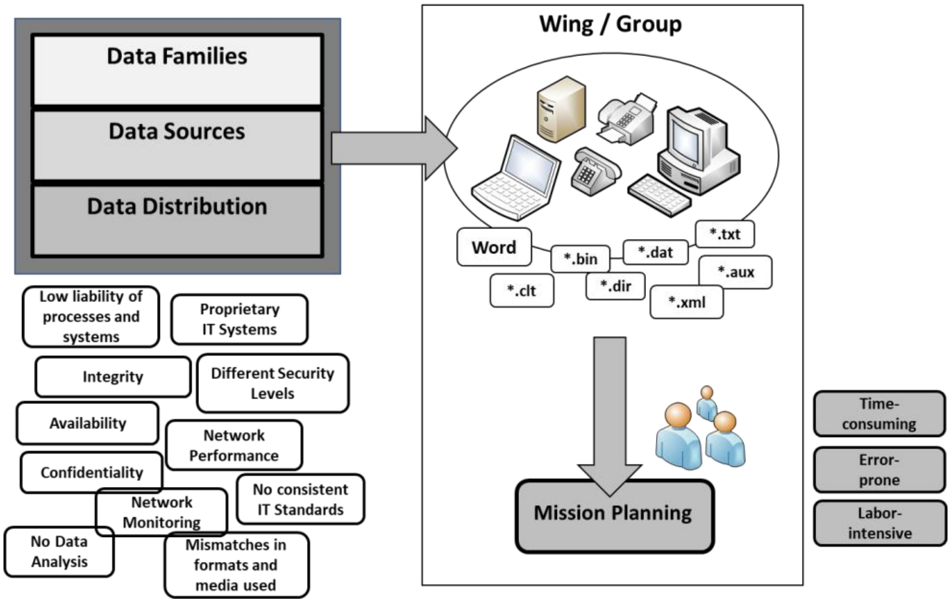

Figure 9: Challenges within current mission data provisioning processes.

Overall, a solution that supports and even rewards data sharing and thus enables data awareness throughout the organization is needed. The solution must adhere to the organization's security standards and sustain principles such as "need-to-know", meaning only the basic information of data should be accessible for a first evaluation of the data. Also, the solution must be easy to handle and allow speedy and (semi-) automated data processing and provisioning.

4.3 Information Exchange Requirements

The exchange of information of the available data should be based on open standards. Also, the solution should be scalable, and should be usable with partners e.g., jointly with other national or international partners. The NATO Consultation, Command and Control Board (C3B) is currently developing the "NATO Core Metadata Specification" (NCMS) as a metadata standard for a wide variety of resources. The NCMS is, as of yet, not available as an open standard. Nevertheless, NCMS is a technically mature standards and is considered to be a good candidate for exchanging information about mission data (and related resources).

In the context of the requirements described above, we decided to base a prototype system on NCMS, despite the fact of its pre-release status. NCMS provides a well-organized and comprehensive set of metadata fields. Additionally, it is flexible and can be extended by Community of Interest (CoI) specific metadata fields.

4.4 Use Of Blockchain In A Prototype Implementation

The current research effort with the German Air Force (GAF) includes the development of a prototype system for testing different approaches and technologies. The prototype system is called "Data Fusion System GAF" (DFSLw). To address the described challenges and requirements, the prototype uses a blockchain-based solution where the blockchain contains the metadata information about the actual mission data. The mission data itself can remain on the source system's storage (while the metadata is published to the blockchain) or may also be stored in DFSLw.

A permission-based blockchain is used to provide the metadata information to the participants of the blockchain network. We are currently using Hyperledger Fabric from the Linux Foundation (Hyperledger, 2021). The prototype system has been tested and presented during an experimentation week together with the customer, focusing on metadata sharing and blockchain security. As the research effort continues, we will go on experimenting with security features, but also with the authentication and authorization features of Hyperledger Fabric.

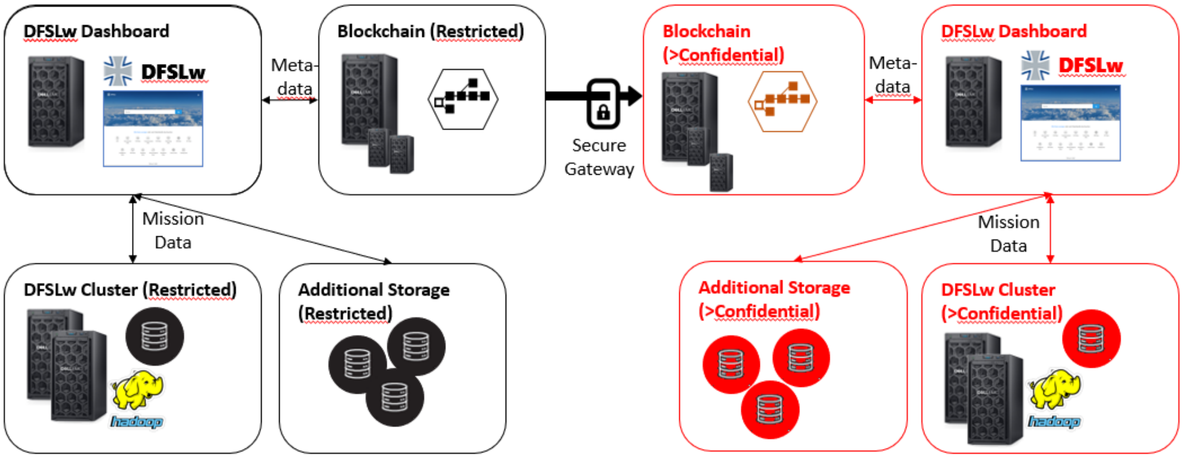
DFSLw Dashboard
DFSLw
Meta-data
Blockchain (Restricted)
Secure Gateway
Blockchain (>Confidential)
Meta-data
DFSLw Dashboard
DFSLw
Mission Data
Mission Data
DFSLw Cluster (Restricted)
hadoop
Additional Storage (Restricted)
Additional Storage (>Confidential)
DFSLw Cluster (>Confidential)
hadoop

Figure 10: Overview of a multi-domain, distributed, blockchain-based data sharing solution.

Figure 10 provides a high-level overview of the current system design. To satisfy multi-domain security requirements (due to different classification levels of mission data), the system is replicated in multiple domains and uses accredited Secure Exchange Gateways (SEGs) to control information flow between separate security domains.

4.5 Lessons Learned

Preliminary lessons learned from using blockchains for secure data sharing and data awareness are:

- A blockchain-based solution for sharing mission data overcomes some of the blocking points or obstacles of cross-organizational data sharing, e.g., the lack of trust.
- Metadata sharing via blockchain enables organizations to enhance overall data awareness and to leverage blockchain technology for achieving data integrity, while being able to limit and control the transfer of the data itself.
- A decentralized approach allows each participant to share a limited amount of information and to keep the own IT infrastructure mostly separate from other partners.
- Applying an established blockchain framework is helpful for faster deployment and reliability.
- Finding a suitable blockchain framework requires analysis and a good understanding of the requirements.
- Blockchain security features (including authentication and authorization) are important for future accreditation of any blockchain-based solution by cyber security authorities.
- The usage of blockchain to store data is not a feature that provides direct benefit to the everyday user. Hence, the integration of blockchain must happen seamless without causing considerable extra effort or loss of performance for the user.

4.6 Open Issues And Next Steps

The research and the development of the prototype system for sharing avionics mission data is still ongoing. These are the blockchain-related topics next in scope:

- Setting up a multi-domain blockchain solution based on multiple blockchain networks to support sharing of differently classified data in compliance with all relevant cyber security regulations (see Figure 10).
- Further research on blockchain security, e.g., resilience against network failures, and blockchain authentication and authorization in the military context.
- Re-evaluation of Hyperledger Fabric and alternative DLT frameworks based on (new) upcoming requirements.

5. CONCLUSIONS AND RECOMMENDATIONS

The usage of a blockchain-based approach to securely store and manage metadata of mission data or simulation resources is promising if at least one of the following challenges need to be addressed:

- Resources are owned by different services, organizations, or nations, and resource metadata is to be shared between them.
- Several partners are contributing to a comprehensive (meta-)information overview where each partner wants to only share a limited amount of information and prefers to keep the own IT infrastructure mostly separate from other partners.
- High requirements on metadata integrity and availability.

In a distributed and decentralized environment, blockchain technology can provide the trust that is needed to nourish transparency amongst the involved partners.

Blockchain technology is often described as a "trustless system". This does not mean that blockchain technology eliminates trust, but it distributes trust over a network of several partners. These partners of the blockchain network come to consensus on what the actual truth is. This consensus constitutes the trust. Additionally, each partner can verify the truth individually, based on the local copy of the blockchain. This minimizes the trust that each partner has to put into the other partners. The trust (and hence the power) is shared among the partners across the network. This contrasts with centralized systems where the truth is concentrated on one central institution while all network partners have to trust this centralized truth.

Further research and development are required to reinforce and validate the findings of this paper. A key topic for further investigation is the resilience of such blockchain-based solutions against network interruptions, data corruption, malicious behavior and other threats.

Additionally, it seems evident that the solution approach can be transferred to other application domains and industries with similar demanding requirements on data integrity, like logistics (e.g., of aircraft spare parts) or healthcare (e.g., traceability of pharmaceutical and medical samples). The described benefits of blockchain technology most likely also apply in these areas as well as in other areas where data integrity is a key requirement, e.g., for successful data sharing and data awareness.

ACKNOWLEDGEMENTS

The "Allied Framework for MSaaS" (including the concept of federated registry) is developed by the NATO Modelling and Simulation Group (NMSG), specifically the technical team MSG-164. The use of blockchains for improving mission data provisioning is part of a research study of the German Air Force. We cordially thank the MSG-164 team and the German Air Force for great discussions and their enduring support.

REFERENCES

Ali, A., & Afzal, D. M. (2018, January). Confidentiality in Blockchain. *International Journal of Engineering Science Invention (IJESI)*, 50-52.

Department of Defense (DoD) Modeling and Simulation Coordination Office (MSCO). (12 July 2012). *Modeling and Simulation (M&S) Communit of Interest (COI) Discovery Metadata Specification (MSC-DMS).*

European Commission. (2021). *Support Centre for Data Sharing*. (D. G. European Commission, Producer) Retrieved 06 15, 2021, from https://eudatasharing.eu/what-data-sharing

Foundation, H. (2020). *Hyperledger Fabric*. Retrieved from https://www.hyperledger.org/use/fabric

Hyperledger. (2021). *Hyperledger Architecture Reference*, *Channels*. Retrieved 06 14, 2021, from https://hyperledger-fabric.readthedocs.io/en/release-2.2/channels.html

Mulligan, C., Zhu Scott, J., Warren, S., & Rangaswami, J. (2018). *Blockchain Beyond The Hype - A Practical Framework For Business Leaders.* World Economic Forum. World Economic Forum. Retrieved 06 21, 2021, from http://www3.weforum.org/docs/48423_Whether_Blockchain_WP.pdf

NATO STO. (2019). *Operational Concept Document (OCD) for the Allied Framework for M&S as a Service.* NATO STO.

NATO STO. (2021). *MSaaS Technical Reference Architecture.* to be published: MSG-164.

NATO STO. (October 2018). *Modelling and Simulation as a Service (MSaaS) - Rapid Deployment of Interoperable and Credible Simulation Environments.*

Richardson, J., Moran, M., Logan, V., Edjlali, R., & Faria, M. (2017). *Survey Analysis: Third Gartner CDO Survey — How Chief Data Officers Are Driving Business Impact.* Gartner Research.

Roshdy, R., Fouad, M., & Aboul-Dahab, M. (2013, August). DESIGN AND IMPLEMENTATION A NEW SECURITY HASH ALGORITHM BASED ON MD5 AND SHA-256. *International Journal of Engineering Sciences & Emerging Technologies*, pp. 29-36.

Shah, N. (2019). *Blockchain for Business with Hyperledger Fabric.* Neu Delhi: BPB Publications.

Simulation Interoperability Standards Organization (SISO). (7 Feb 2020). *Discovery Metadata Specification for M&S Resources.* Product Nomination.

UK Government Office for Science. (2016). *Distributed Ledger Technology: beyond block chain.*