Compatibility Check of smartDSM with the IDSA Reference Architecture

Igor Koropiecki, Krzysztof Piotrowski
IHP - Leibniz Institut für innovative Mikroelektronik
Frankfurt (Oder), Germany
{koropiecki, piotrowski}@ihp-microelectronics.com

***Abstract*—The paper evaluates the compatibility of a middleware platform called smartDSM, with the IDSA Reference Architecture Model 4.0 for building secure data spaces, that facilitate trusted and secure data sharing across various organizations and sectors. Considered aspects include technical interoperability, security measures, data governance policies, and scalability. The findings identify both strengths and areas for improvement.**

***Index Terms*—smartDSM, IDSA, compatibility, interoperability, data security**

I. Introduction

The digital transformation and growing significance of data collection (with focus on privacy and sovereignty) show the necessity for robust and interoperable data systems, that leverage modern technology [1]. These systems must provide strong tooling for data sovereignty, governance and access control to ensure that the value of the data is fully realized and protected. While the systems offer the necessary tools and infrastructure, it is the data itself that holds the true value, enabling organizations to innovate and maintain competitive advantages. In order to maximize the benefits from the data, it is necessary to have a common language that allows to freely exchange metadata and the data itself.

Large, well-structured and addressable data-sets (using common ontology and vocabularies) allow stakeholders to solve problems (efficiency and cost savings), make informed decisions (business growth through strategic planning), improve processes (increased productivity), and adapt to customer behaviors (increased customer satisfaction).

Simultaneously, consumer awareness about data privacy is on the rise, driven by increased amount of online interactions and advancements in data collection and analysis technologies [2]. Research indicates that consumers are becoming more aware of how companies use their personal information, however many believe that they do not have control over their data. This can lead to justified privacy concerns [3].

Regulations like GDPR (General Data Protection Regulation) [4] legally ensure data privacy by establishing clear conditions on how data is managed and protected, which is crucial for building trust and maintaining customer relationships.

On the global level, the IDSA (International Data Spaces Association) [5] initiative has introduced the Reference Architecture Model (RAM) [6] and the concept of *data spaces* [7], [8], which are key initiatives attempting to create a trusted environment for data exchange among various stakeholders (public, private or personal).

The IDSA RAM is a framework for secure, decentralized and standardized data exchange with data linkage in a trusted ecosystem. The model leverages existing standards and technologies, along with well-accepted governance models, to ensure data sovereignty and trust among participants.

Data spaces are foreseen as the foundation for secure and standardized environments for data sharing across different domains. They enable integration and utilization of data from diverse sources, which is crucial for many sectors such as smart cities, healthcare or finance.

The smartDSM platform [9], [10] is a middleware solution that facilitates secure and distributed data exchange with processing capabilities in distributed measurement systems. Developed using Java and web technologies, smartDSM supports various operating systems and hardware configurations.

The platform can be used with another middleware platform, called tinyDSM [11] (for embedded systems, also based on variables) to streamline a distributed embedded system deployment on the low and the high level.

In this paper, we assess how well smartDSM features align with the guidelines and standards in IDSA RAM 4.0. The goal is to identify similarities, highlight differences, and suggest potential improvements.

II. Related Work

The concepts presented in the Reference Architecture Model served as a foundation for the development of DIN SPEC 27070 [12], a standard that focuses on the requirements, architecture, and cybersecurity measures for a security gateway in industrial data exchange.

On the European level, the Gaia-X [13], [14] initiative funded by the European Union (EU), extends the IDS RAM framework by focusing on creating a secure European data infrastructure. It emphasizes the governance and infrastructure to be primarily under European control (data residency within the region). It also ensures adherence to strict European data protection regulations like GDPR. Gaia-X adheres to high interoperability standards defined by the RAM, while having the ability to also incorporate additional European-specific protocols. Through legislation, funding, and strategic frameworks, the EU is promoting data spaces as the future of data-driven ecosystems in Europe. Fig. 1 presents a high-level overview of the IDSA, Gaia-X and their relation.

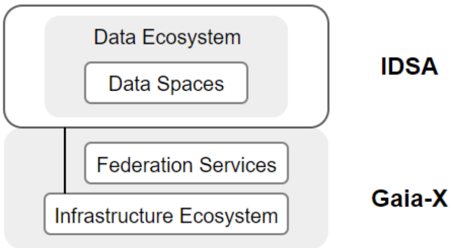

Fig. 1: Gaia-X in context of IDSA

The concept of data spaces (coined about 15 years ago [15]) is gaining adoption, with emerging platforms that implement (or align with) the paradigm. Examples include FIWARE [16], Eclipse Dataspace Components [17] and others.

III. Methodology

To evaluate how smartDSM aligns with the IDSA RAM 4.0 framework, a literature review was performed to extract and present the ideology of the IDSA RAM 4.0, outlining its key components such as layers, technologies, and approaches. The smartDSM platform was then compared to the identified components. The results were compiled into a table to visualize the compatibility level in each dimension.

IV. Compatibility Evaluation

The IDSA Reference Architecture Model (RAM) 4.0 [6] utilizes existing standards, such as *data formats* (e.g. JSON-LD, XML), *data models* (e.g. Resource Description Framework, RDF; Web Ontology Language, OWL), *security protocols* (e.g. TLS/SSL, OAuth) and *policy languages* (e.g. Open Digital Rights Language, ODRL; Business Process Model and Notation, BPMN). It also provides guidance (for concrete implementations) in terms of technologies, and governance models.

The IDSA RAM 4.0 contains five layers (which include the following tasks):

Business layer categorizes the roles of participants (core, intermediary, service provider, etc.) within the International Data Spaces (IDS), with their activities and interactions.

Functional layer defines the functional requirements (identity, validation, etc.) of the IDS, the features (data transmission, policy enforcement, etc.) and data sovereignty which allow the secure data exchange.

Information layer provides a model based on linked-data to describe the properties (metadata, usage policies, etc.) of the IDS components (connectors, brokers, apps, etc.).

Process layer details the interactions between the IDS components using the BPMN notation.

System layer describes the architecture of software components (data exchange is handling, brokers managing metadata, app delivery, deployment configurations, etc.).

There are three perspectives defined: *security* (measures to ensure trusted exchange and protect data sovereignty), *certification* (of IDS participants and components), and *governance* (framework, roles, and processes for operating the IDS ecosystem).

In smartDSM, a variable is the core data structure that encapsulates stakeholder data with its metadata. Each variable is a tuple-based structure maintaining previous values (similar to time-series) with attributes such as name, sub-fields, types or limits. This defines the metadata. The actual data, owned by stakeholders, includes values, ownership information, timestamps, and source identifiers. Stakeholders (individuals or organizations), own assets within the system and interact with variables through actions, such as read, write, update, or clear. Permissions are managed by the stakeholders, who can grant or revoke access to other stakeholders as needed. The distributed nature allows to build systems that store data centrally [18], on user premises [10], [19] or both. Communication security is ensured through the Public Key Infrastructure (PKI, using Java's JSSE), with each communicating entity using certificates for encryption and identification. The platform comprises of several core components like the server, data interface, proxy, discovery service, and web certificate manager. The components work together to manage data access, permissions, and communication between local and remote servers.

The RAM uses standardized frameworks suitable for large-scale, state-of-the-art implementations, while smartDSM is positioned as a more prototype-oriented solution and could be a smaller branch (or data space) within the ecosystem.

Table I contains the evaluation result. The first column includes the identified dimensions, the second column contains approaches of IDSA RAM 4.0. The next columns include evaluation of smartDSM, a symbol rating scale of the compatibility, and a comment with supporting information.

V. Conclusion and Further Steps

The smartDSM platform partially aligns with the principles of the IDSA RAM. Focus is put on data sovereignty, security and ownership, ensuring that data providers retain control (user consent and explicit permissions). Usage of PKI aligns with the security blueprint introduced in the IDSA RAM. The distributed architecture of nodes (in smartDSM) acting as local servers, aligns with the vision of a decentralized ecosystem.

The similarities are mostly on the technical level, though some disparities still exist. SmartDSM relies on variables (tuples), which limits the compatibility with data formats supported by the IDSA RAM, such as JSON-LD and XML. It can be addressed through creating adapters, or integrating the formats directly. The smartDSM platform lacks a governance model and provenance mechanisms, which are one of key components of IDSA RAM.

TABLE I: Evaluation of smartDSM compatibility with the IDSA Reference Architecture

	IDSA RAM 4.0	smartDSM	○	Comment for smartDSM
Data types	Standardized types, e.g. JSON-LD, XML	Variables (tuples); Time-series	◐	Custom data type restricts interoperability
Data models	RDF and OWL	None	○	No ability to describe data and relationships
Data governance	Clearing house; Data usage policies	Minimal	○	Only access control is possible
Data provenance	Tracking (clearing house)	Limited	◐	Logging could be implemented by creating a permission-respecting, listening service
Data ownership	Data usage policies; Access control	Access control	◐	Opt-in variable-level consent to user-premises data
Data sovereignty	Decentralization; Data usage control; Identity management	Decentralization; Data usage control; Identity management	●	Certification authority, web-based dashboard for stakeholders
Interoperability	Standardized interfaces and protocols (IDS Connector)	Data Interface (Java, Python)	◐	Partial compatibility in terms of core concepts
Trust, security & privacy	Digital certificates (X.509); Encryption	Digital certificates (X.509); Encryption; Opt-in sharing	●	Stakeholders must present a valid certificate to participate in data exchange
Scalability	Endorses distributed approach	Distributed approach	●	Nodes can be arranged in various hierarchies and managed through a container orchestration system
Modularity	Endorses modular approach	Modular approach	●	Microservices architecture, all data are structured as variables

○ – not compatbile ◐ – partially compatible ● – compatible

It is worth noting that the platform is not a standalone ready-to-go solution; rather, it is a set of building blocks that can be used to create a distributed system. Further development should address limitations in data types, models, interoperability, and governance to achieve full compatibility with the IDSA RAM. Other directions include analysis in terms of data quality, platform adaptability and usability.

References

[1] McKinsey & Company, What is Digital Transformation, Available online: https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-digital-transformation (accessed on 16 July 2024).

[2] F. M. Shipman and C. C. Marshall, Ownership, Privacy, and Control in the Wake of Cambridge Analytica: The Relationship between Attitudes and Awareness, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020.

[3] W. Presthus and H. Sørum, Consumer perspectives on information privacy following the implementation of the GDPR, International Journal of Information Systems and Project Management, 2021.

[4] European Parliament and Council of the European Union, General Data Protection Regulation (GDPR), Regulation 2016/679, 27 April 2016.

[5] International Data Spaces Association Website, Available online: https://internationaldataspaces.org/ (accessed on 16 July 2024).

[6] IDSA, Reference Architecture Model 4, Available online: https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4 (accessed on 16 July 2024).

[7] European Commission, Common European Data Spaces, Available online: https://digital-strategy.ec.europa.eu/en/policies/data-spaces (accessed on 16 July 2024).

[8] European Commission, Towards a common European data space, COM(2018) 232 final, April 2018.

[9] I. Koropiecki, K. Piotrowski and R. Szulim, SMARTDSM: Data Space Middleware for Distributed Measurement Systems, 14th Scientific Conference on Measurement Systems in Research and in Industry. Łagów, Poland, 2022.

[10] I. Koropiecki and K. Piotrowski, SmartDSM: Towards User-Centric IoT Middleware Platform for Privacy-Focused Smart Systems, Proc. of the 2023 IEEE Int. Conf. on IoT and Intelligence Systems, IoTaIS 2023.

[11] K. Piotrowski, P. Langendörfer and St. Peter, tinyDSM: A Highly Reliable Cooperative Data Storage for Wireless Sensor Networks, Proc. of the IEEE International Symposium on Collaborative Technologies and Systems (CTS), 2009.

[12] DIN SPEC 27070, Reference Architecture for a Security Gateway for Sharing Industry Data and Services. Available online: https://www.din-mitteilungen.de/de/din-spec-27070-und-kuenftige-din-spec-27072-standardisiertes-security-gateway-und-security-anforderungen-fuer-iot-geraete-im-small-business-home-umfeld-272902 (accessed on 16 July 2024).

[13] Gaia-X website, https://gaia-x.eu/ (accessed on 16 July 2024).

[14] Gaia-X Initiative, Gaia-X - Architecture Document - 22.04 Release. Available online: https://gaia-x.eu/wp-content/uploads/2022/06/Gaia-x-Architecture-Document-22.04-Release.pdf (accessed on 16 July 2024).

[15] B. Otto et al., Designing Data Spaces, Fraunhofer ISST. Dortmund, Germany, 2022.

[16] FIWARE Foundation website, https://www.fiware.org/ (accessed on 12 August 2024).

[17] Eclipse Dataspace Components, https://projects.eclipse.org/projects/technology.edc (accessed on 12 August 2024).

[18] I. Koropiecki and K. Piotrowski, Environment monitoring backend and dashboard, GI/ITG KuVS Fachgespräch Sensornetze, FGSN 2023.

[19] I. Koropiecki and K. Piotrowski, IoT and Data Spaces as a highly distributed measurement and control system, 15th Scientific Conference on Measurement Systems in Research and in Industry, SP 2024. Zielona Góra, 2024.