Decentralised semantics in distributed health data ecosystems: Harmonising quantitative data from medical devices for improved insights-driven applications[⋆]

Paul Knowles[1,*,†], Robert Mitwicki[1,†] and Michal Pietrus[2,†]

[1] *MeDDEa Solutions AG, Grosspeteranlage 29, Basel, 4052, Switzerland*
[2] *ArgonAUTHs sp. zoo, ul. Zamknieta 10, Krakow, 30-554, Poland*

Abstract
In this paper, we introduce *Decentralised semantics*, the segregation of task-specific, content-addressable objects, as a solution for harmonising quantitative data from medical devices within and across distributed health data ecosystems. Overlays Capture Architecture (OCA) is a semantic architecture that offers a scalable harmonisation solution between data models and data representation formats from different sources while providing a roadmap to resolve privacy-compliant data sharing between healthcare actors, including regulatory agencies. As an illustration, we introduce the concept of applying ”overlays” as a transformation mechanism to convert source measurements from medical devices to a standard system of measures and units for improved insights-driven applications.

Keywords
Decentralised semantics, Overlays Capture Architecture (OCA), data harmonisation, distributed health data ecosystems, Internet of Things (IoT), medical devices,

1. Introduction

Repeatable and reproducible measurements are crucial in the medical field, where regulatory agencies require process traceability. In addition, with the global medical devices market projected to grow from 495.46*billionin*2022*to*718.92 billion by 2029[1], an increasing amount of device-generated data is set to enter the digital health landscape over the next several years with its interpretation used for clinical decision-making.

The projected growth of the global medical devices market coincides with the current sectoral drive towards scalable, distributed health data ecosystems for improved digital sovereignty–each party's right and ability to control its digital data. Furthermore, data anomalies in training datasets compromise Artificial Intelligence (AI) and Machine Learning (ML) solutions, where

14th International SWAT4HCLS Conference, FEBRUARY 13–16, 2022, Basel, Switzerland
[⋆] Contribution as a short paper
[*] Corresponding author.
[†] These authors contributed equally.
paul.knowles@meddea.global (P. Knowles); robert.mitwicki@meddea.global (R. Mitwicki); michal.pietrus@ds-aht.eu (M. Pietrus)
0000-0003-4379-5037 (P. Knowles)

algorithms analyse data and make predictions based on that information. As such, there is an imminent need for an unintrusive solution for harmonising device-generated quantitative measurements at scale as a scientific base for patients, nurses, physicians, laboratory technical staff, and other external entities such as regulators, insurance companies, and healthcare organisations to generate meaningful analytics and insights.

2. Decentralised semantics

Decentralised semantics[2] describes the separation of task-specific metadata objects that, when combined, provide an accurate representation of a complex object. The following diagram illustrates Overlays Capture Architecture (OCA)[3] as a layered architecture of ”overlays” providing a utility for representing, in this case, a multi-language HL7® FHIR®[4]-compliant Continuous Glucose Monitoring (CGM) form.

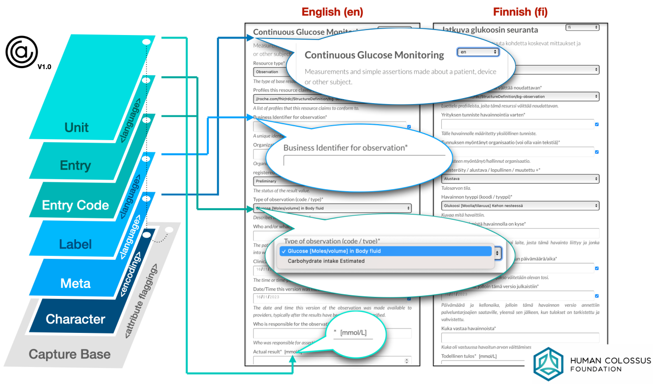

Figure 1: Overlays Capture Architecture (OCA). Separating task-specific overlays within a standard architecture brings an extensible, flexible, and interoperable solution for capturing the metadata necessary to render complex objects.[EUPLv1.2, Human Colossus Foundation]

The primary objective of decentralised semantics is to enable data and metadata harmonisation at scale, providing an opportunity to structure unstructured data, transform data from one data representation format to another, and offer a long-term solution for data language unification within and across distributed health data ecosystems.

Decentralised semantics is one of the foundational concepts of a Dynamic Data Economy (DDE)[5], a next-generation data-agile economy offering a new paradigm in digital living, interaction, and growth, with a vision of empowering people and businesses to make better-informed decisions based on insights from interpretable, harmonised data framed by sound data governance.

The separation of task-specific metadata objects allows the custodianship of granular data collection, validation, transformation, and presentation tasks under the same governance to reside with different actors, departments, or entities without compromising the overall integrity of those processes. This level of interoperability is essential in a data-agile economy where multiple actors from various institutions participate in complex use cases, supply chains, and data flows supported by multistakeholder data governance administrations and frameworks.

Decentralised semantics provides a new paradigm for domain-driven design (DDD)[6], enabling software models to match specific domains with a richer understanding of the processes and rules that apply.

3. Distributed health data ecosystems

The European Commission's earliest proposal of a distributed health data ecosystem is the European Health Data Space (EHDS)[6], an initiative to address health-specific challenges to electronic health data access and sharing. The EHDS aims to create a framework for health data ecosystems where natural humans can control their health data whilst enabling purpose-based, privacy-preserving data exchanges between all stakeholders, public or private. A substantial amount of quantitative data in those multi-stakeholder environments will come from purpose-driven medical devices.

he EU roadmap for EHDS will require all health data ecosystems to find innovative software, hardware, and governance solutions for managing, sharing, and discovering digital information, where ecosystem participants can uniquely identify entities and objects across diverse systems.

Internet of Things (IoT) application for medical devices: Converting quantitative measurements and units using overlays and conversion tables.

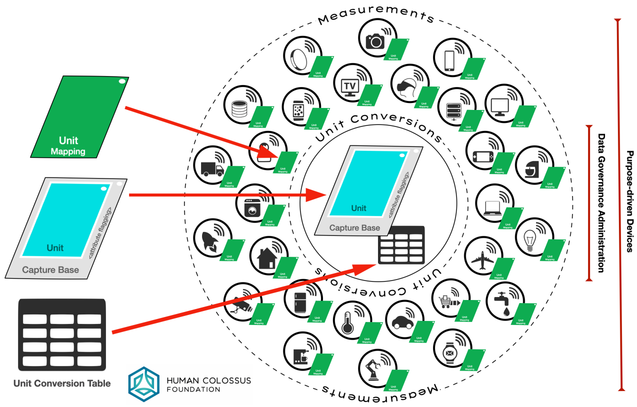

Figure 2: Purpose-driven medical devices collect and transmit quantitative measurements within a distributed health data ecosystem. Unit and unit-mapping overlays are cryptographically bound to a common capture base, with unit conversion tables providing recipes for accurate real-time measurement conversions.[EUPLv1.2, Human Colossus Foundation]

The Internet of Things (IoT) describes physical objects ("things") with sensors, processing ability, software, and other technologies that connect and exchange data with other devices and systems over any communications network. Often accompanied by units defined and adopted by convention or law, many IoT devices focus on collecting and transmitting quantitative measurements in real-time for continuous insights on changes in the measurand.

Differences in preferred measurement systems and regulatory differences between countries and jurisdictions bring additional complexity when working between medical device operating systems. For example, standard measurement systems implemented may include the International System of Units[7] or SI (the modern form of the metric system), the British imperial system[8], or the United States customary system[9]. As a result, measurement comparisons from different devices across jurisdictional healthcare systems and applications can take significant time and effort due to the added complexity of the conversion factors required for accurate quantitative data harmonisation.

Aligning measurement units is particularly important when converting measurements or sharing quantitative data. For example, independent IoT sensors may use different units to represent their measurements, in which case, standardised unit mappings provide mechanisms for accurate conversion.

Leveraging the Unified Code for Units of Measure (UCUM)[10], purpose-driven medical devices collecting quantitative data in one unit can use unit-mapping overlays under their issuance and control as a mechanism to facilitate unit conversions for transmitting harmonised measurements for IoT applications. Furthermore, adding unit-tagging overlays into the process would enable subject matter experts to annotate and tag ontology terms at a unit level according to, for example, the QUDT Ontologies[11] for deterministic unit relationships for improved analytics, querying, and data integration.

Unit-mapping overlays provide the necessary mappings to enable medical devices to convert quantitative data in real time according to metadata definitions defined by ecosystem data governance administrations (DGA). Capture bases and associated overlays issued and published by the DGA provide common targets for harmonising quantitative data with unit conversion tables providing recipes for accurate measurement conversions. Specifically, the unit-mapping and unit overlays contain the cryptographic identifier of the capture base, ensuring the integrity of those objectual relationships and facilitating a secure means for unit transformations.

In addition to measurement conversions, different overlay types provide transformation and presentation mechanisms for various use cases and harmonisation processes.

Some simple examples include using

- *Unit overlays* for temperature conversions;
- *Format overlays* for date format conversions;
- *Label overlay*s for data presentations in multiple languages.

Distributed health data ecosystems, such as the EHDS, require interoperable data modelling solutions capable of capturing extensible metadata on an ongoing basis. Strong bindings between overlays and capture bases, in conjunction with the cryptographic identifiers products of one-way hash functions[12] of the objects, provide objectual integrity throughout the data lineage[13] lifecycle, including data collection, validation, transformation, and presentation, to ensure the same results upon running the model under the same initial conditions.

4. Conclusions

Using Overlays Capture Architecture (OCA) as a represented architecture, we introduced the concept of applying overlays as transformation mechanisms to convert source measurements from medical devices to a standard system of measures and units for improved insights-driven applications. OCA continues to provide a comprehensive solution for defining deterministic metadata, structuring digital objects and harmonising data payloads, showing promising results in several use cases across the healthcare and life sciences sector. Decentralised semantics offers a solution for collaborative algorithmic modelling for innovative data and metadata cleaning to enhance the development of accurate and ethical AI while providing a solution for quantitative measurement harmonisation from medical devices.

Decentralised semantics offers an enhanced data and metadata harmonisation solution while ensuring objectual integrity throughout any data lifecycle. As a result, its application in IoT use cases would greatly simplify the technical assessment of hardware choices for medical devices producing quantitative data.

References

References are in order of appearance in the body of the text.

[1] Fortune Business Insights, Medical Devices Market Size, Share & COVID-19 Impact Analysis, By Type (Orthopedic Devices, Cardiovascular Devices, Diagnostic Imaging, In-vitro Diagnostics, Minimally Invasive Surgery, Wound Management, Diabetes Care, Ophthalmic Devices, Dental Devices, Nephrology, General Surgery, and Others), By End-user (Hospitals & ASCs, Clinics, and Others), and Regional Forecast, 2022-2029), June 2022. https://www.fortunebusinessinsights.com/industry-reports/medical-devices-market-100085

[2] P. Knowles et al., Decentralised semantics in distributed data ecosystems: Ensuring the structural, definitional, and contextual harmonisation and integrity of deterministic objects and objectual relationships, 2022. https://ceur-ws.org/Vol-3249/paper4-OSS.pdf

[3] P. Knowles, Overlays Capture Architecture (OCA)–Version 1.0, October 2022. https://static1.squarespace.com/static/5ead4c8660689c348c80958e/t/634d55c7cb2c9a43206480c1/1666012626086/HCF+Overlays+Capture+Architecture+%28OCA%29+v1.pdf

[4] HL7® FHIR® (Fast Healthcare Interoperability Resources), https://hl7.org/fhir/

[5] R. Mitwicki, Genesis of a Dynamic Data Economy, 2020. https://humancolossus.foundation/blog/dde-first-contact

[6] European Commission, European Health Data Space. https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

[7] Bureau International des Poids et Mesures (BIPM), The International System of Units (SI), https://www.bipm.org/en/measurement-units

[8] Wikipedia, British imperial system, https://en.wikipedia.org/wiki/Imperial_units

[9] Wikipedia, United States customary system, https://en.wikipedia.org/wiki/United_States_customary_units

[10] Unified Code for Units of Measure (UCUM),https://ucum.org/

[11] QUDT, https://www.qudt.org/

[12] M. Hawthorne, One-Way Hash Function, 2019. https://www.technipages.com/definition/one-way-hash-function

[13] Wikipedia, Data lineage, https://en.wikipedia.org/wiki/Data_lineage