Utilising the FOXS Stack for FAIR Architected Data Access

John MEREDITH\textsuperscript{a1}, Nik WHITEHEAD\textsuperscript{b} and Michael DACEY\textsuperscript{b}

\textsuperscript{a}Wales Institute for Digital Information, Cardiff UK
\textsuperscript{b}University of Wales Trinity Saint David, Swansea UK

Abstract. A FOXS stack assembles HL7 FHIR, openEHR, IHE XDS and SNOMED CT as an operational clinical data platform to build digital systems. This paper analyses its applicability for FAIR-enabled medical research based on a summary of key principles. It highlights the benefit of the blended approach to operational technology stacks for health systems, and a need for industry standard technologies to enable greater semantic coherence for primary/secondary data use.

Keywords. interoperability, HL7 FHIR, openEHR, IHE XDS, SNOMED CT, FAIR.

1. Introduction

There is now a paradigm shift in health data accessibility to the requirement for fully structured, semantically coherent data available via open APIs. The open architecture approach is key to the meaningful use of data for operational clinical tools as well as providing a foundation for secondary use for research purposes. The FAIR Principles to ensure data are \textit{Findable}, \textit{Accessible}, \textit{Interoperable} and \textit{Reusable} \cite{1} have been adopted by organisations such as Health Data Research UK, who consider embracing open standards as a necessity \cite{2}. However there has been little requirement for FAIR to be adopted by the UK NHS outside of research.

In this paper, we present an approach to supporting FAIR principles for health data stemming from the open clinical data interoperability platform; the FOXS stack. This employs four commonly used technology standards within the domain of digital health; HL7 FHIR, openEHR, IHE XDS and SNOMED CT. These are specifications and technologies to persist clinical data, bound by standardised terminologies, represented by syntax and metadata harmonised messaging and document structures. We present and discuss the high-level summary of each FOXS component and assess compatibility with FAIR principles.

1.1. The FOXS Stack

The FOXS stack is assembled from the following components:

\begin{itemize}
  \item FHIR: Fast Healthcare Interoperability Resources (FHIR) from HL7 are regarded as the emerging standard for technical and syntactic interoperability \cite{3},
\end{itemize}
- openEHR is a specification [4] that describes clinical models and the rules which govern them. It is constrained by a reference model to support the longitudinal record, and acts as the core data persistence layer of FOXS.
- XDS: Cross-Enterprise Document Sharing (XDS.b) from IHE is a standards-based specification [5] to support the sharing of documents and images between health organisations.
- SNOMED CT is a hierarchal clinical vocabulary for use with digital tooling and patient records used widely across the world [6], mappable to other code systems and bound within data structures used by the above.

Both FHIR and openEHR specifications are freely available under open-source licenses. However, it is essential that standards have been adopted for use by the healthcare providers. While it is possible to create standardised structures for documents in both openEHR and FHIR, neither has demonstrated implementation at scale that rivals the more mature standard of IHE XDS.b for interoperability across the health sector [7].

2. Methods

The applicability of utilising some FOXS components for FAIR research has been established. NIH has issued a RFI [8] on clinical research utilising FHIR indicating a desire to test its efficacy. Recent progress also includes work groups aiming to harmonise FHIR with the BRIDG reference model [9]. Compliance to FAIR principals has also been established for openEHR [10]. This notion is extended here to the generalised FOXS stack based upon previously published implementation considerations [11].

Certain principles were established for the definition of metadata within FOXS stack to delineate between provenance and knowledge-based data. Provenance assumes an inherent ownership and position within the clinical pathway and can be found in FHIR, openEHR and XDS.b specifications. Knowledge metadata is attributed to ontological aspects of openEHR archetypes and FHIR based resources, to differing levels of detail.

For example, the clinical model for blood pressure exists as an openEHR archetype but described within the contents of a FHIR observation resource. While openEHR classes are analogous to certain FHIR resources, openEHR offers richer published definitions to support knowledge metadata for research. Any given model may contain references to external SNOMED CT codes (e.g. a ‘record artifact’ based document type) as well as contained within the data itself. In this scenario, we consider openEHR, FHIR and SNOMED CT with capabilities to describe clinical content, as well as metadata elements. IHE XDS.b represents a standardised container used in healthcare for additional clinical content such as documents or images (e.g. using the DICOM standard).

3. Results

Each FAIR principle is summarised with compliance to FOXS components as a whole.

3.1. Findability

Unique and persistent identifiers are the basis for findability to ensure computability. FHIR, openEHR and XDS rely upon location-based URLs for resources, compositions
and documents respectively. OpenEHR facilitates metadata contained within a composition structure to be located via this identifier. FHIR presents optional metadata elements within individual resources. XDS utilises metadata structures as part of an affinity domain to locate records which may consist of a variety of different data standards for content (e.g. FHIR).

In terms of metadata richness, FHIR resources contain mandatory backbone elements and referenced Resources such as 'Patient' and 'Encounter' as well as optional metadata elements. OpenEHR attributes metadata at multiple levels to consider the provenance of the data as well as the models themselves, akin to FHIR but in more granular detail owing to the archetype modelling approach. The combination of FOXS components facilitates data descriptors identified by class of archetype or resource, document type, care context or clinical term. Computability is made possible with content due to the nature of archetypes, resources and terms being unique identifiers to data. (Meta)data may be stored separately in model repositories such as CKM [12].

3.2. Accessibility

FOXS stack components rely on W3C standard web-based protocols such as SOAP XML (XDS) and REST (FHIR, openEHR). FHIR offers a framework for search capabilities and within openEHR architecture, all data items attributed to a specific archetype may be queried independently with the archetype query language (AQL, [8]). This presents as an increasingly granular capability as query use cases move from the higher-level interoperability space (FHIR, XDS.b) to the data persistence layer (openEHR), with both perspectives augmented by SNOMED CT.

Role base access and authentication is supported with rich metamodels. Rules on persistence and retention may be subject to domain specific policy such as Caldicott data sharing principles in the UK[13]. Maintaining these structures to support findability also facilitates accessibility when record management policy has a consideration for the destruction of data when no longer needed [14], as this allows metadata to persist longer.

3.3. Interoperability

Utilising archetypes as the base model for FOXS facilitates semantic coherence and a common representation of knowledge. This may be demonstrated through a variety of widely used, machine-readable formats (e.g. XML, JSON, RDF). OpenEHR provides an internal terminology and data may be enhanced with the use of SNOMED CT to act as a vocabulary. In addition, FHIR has successfully been utilised to represent SNOMED CT reference data sets[15] which can in turn be embedded within openEHR models. The addition of XDS.b completes the interoperable capability for document bound data.

3.4. Reusability

The final principle concerns how (meta)data are reused and their applicability to clinical use cases. This requires implicit understanding of the use and misuse of data supported by robust and detailed metadata. A FOXS-based platform is able to make use of conformance archetypes that map directly to the interoperability layer (e.g. matching an XDS.b profile). Data usage may be encapsulated within a consent focussed archetype that describes specific scenarios such as research. Provenance-based (meta)data enables operational and secondary use filtering (e.g. only include observations recorded recently
or in specific clinical contexts). Minimal data standards may be reflected as cardinality within an archetype/resource, or at the conformance API layer (i.e. a FHIR profile).

4. Discussion

FOXS implementations view SNOMED CT as a component of the wider stack, rather than it being the ontology itself. Wider structures provided by the persistent archetype model provide each data point. The increased ontological context for the uniquely identifiable nature of clinical data via a cumulatively standards based approach enables the data to become FAIR [16]. While it is feasible to include multiple terminologies or classifications within openEHR archetypes and FHIR resources, SNOMED CT becomes an invaluable tool due to its capability to map to external terminologies such as ICD-10. This is essential for enabling FAIR enabled research [17].

While the interoperable aspects of openEHR, through the native API service layer, comply with FAIR principles, it could be argued that it lacks formal standardisation in terms of industry usage. This is because the openEHR specification enables standards for clinical data models to be created that reflect use cases at the data persistence layer. The commonality of shared archetypes does not imply that the assembly of said models will be standard across all implementations. The resulting templates and APIs will reflect the decisions taken by local implementers to support the specific use case at hand. This also applies to FHIR (and to some extent XDS and SNOMED CT) that all are subject to various levels of localised standardisation to reflect the heterogenous practice across organisation or geographical boundaries.

Efforts have been made to align FAIR principles to FHIR [18], however progress has been hampered due to existing gaps between health and research standards for basic model elements such as demographics [19]. FHIR is regarded as the messaging standard for health, and is seeing increased use in the UK through initiatives such as the UK Core [20]. This advocates a desire for compliance to accessibility standards consumed by operational systems, through a common syntax. Where these interoperable APIs share common data structures, they may be used by multiple systems or actors, but also customised to support local implementation requirements. This necessitates a degree of transformation between systems, producing non-standardisation that would seem to oppose some FAIR principles such as accessibility. Nevertheless, the ability to rely on a common, persistent baseline model in openEHR, supported by SNOMED CT terms as a common vocabulary enables a FAIR-enabled technology stack to reside within operational digital systems and not exist as a purely research-based endeavour. Additionally, openEHR acts as a proxy for common data elements, essential for FAIR data sharing [19]. By abstracting this model away from the requirements of messaging, openEHR compliments FHIR, providing the flexibility demanded by implementors while maintaining the semantic coherence that benefits FAIR.

We suggest that by utilising FHIR as the interoperability gateway, we begin to standardise FAIR principles alongside established industry practice. Relying on XDS profiles to support document and image archive-based paradigms, supported by a robust metadata model supports this view.
5. Summary

The assembly of FOXS components represents an act of domain-based convergence. It juxtaposes the detailed curation of clinical models for data persistence with interoperability, utilising standard syntax and protocols. While this generalises at the enterprise scale, the assembly attempts to enable FAIR-ness by way of facilitating data access through one or more routes within a FOXS platform. Future research will seek to develop a specification for a FAIR-enabled FOXS stack and assess how this aligns to currently available FAIR maturity models [21] to support clinical data research.

References

[1] Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016 31 [Internet]. 2016 Mar 15 [cited 2021 Jul 21];3(1):1–9. Available from: https://doi.org/10.1038/sdata.2016.18
[2] Health Data Research UK. Data Standards Principles. 2020.
[3] Health Level Seven (HL7). FHIR Overview [Internet]. [cited 2021 Jul 26]. Available from: shorturl.at/c594
[4] openEHR Foundation. What is openEHR? [Internet]. 2019 [cited 2021 Jun 6]. Available from: http://bit.ly/openehr
[5] IHE International. Cross-Enterprise Document Sharing (XDS.b) [Internet]. IHE ITI Technical Framework (Rev17 Vol1). [cited 2021 Jul 27]. Available from: https://www.nlm.nih.gov/healthit/snomed/snomed_overview.html
[6] Wettstein R, Merzweiler A, Klass M, et al. Using openEHR in XDS.b Environments – Opportunities and Challenges. Stud Health Technol Inform [Internet]. 2020 [cited 2021 Aug 4];272:300–3. Available from: https://doi.org/10.3233/shti200554
[7] National Institutes of Health. Use of the HL7 Fast Healthcare Interoperability Resources (FHIR) for Capturing and Sharing Clinical Data for Research Purposes [Internet]. 2019 [cited 2021 Jul 26]. Available from: https://grants.nih.gov/grants-guide/notice-files/NOT-OD-19-150.html
[8] BRIDG, HL7 & FHIR - Biomedical Research Integrated Domain Group [Internet]. National Cancer Institute. 2021 [cited 2021 Jul 26]. Available from: https://bridgmodel.nci.nih.gov/hl7-fhir
[9] Bönisch C. FAIRness of openEHR Archetypes and Templates. 2019;
[10] Jacobsen A, de Miranda Azevedo R, Juty N, et al. FAIR Principles: Interpretations and Implementation Considerations. Data Intell. 2020;2(1–2):10–29.
[11] OpenEHR Clinical Knowledge Manager (CKM) [Internet]. Ocean Informatics / Ocean Health Systems; 2019. Available from: https://ckm.openehr.org/ckm/
[12] The Eight Caldicott Principles [Internet]. UK Caldicott Guardian Council. 2020 [cited 2021 Jul 30]. Available from: https://www.ukcegc.uk/manual/principles
[13] Information Governance Alliance. Records Management Code of Practice for Health and Social Care [Internet]. 2016 p. 38, 65. Available from: shorturl.at/fpETY
[14] Metke-Jimenez A, Steel J, Hansen D, et al. Ontoserver: A syndicated terminology server. J Biomed Semantics. 2018;9(1):1–11.
[15] Frexia F, Mascia C, Lianas L, et al. openEHR Is FAIR-Enabling by Design. Stud Health Technol Inform [Internet]. 2021 May 27 [cited 2021 Jul 21];281:396–45. Available from: https://doi.org/10.1093/bib/bbb204
[16] Vesteghem C, Brondum RF, Sonderkaer M, et al. Implementing the FAIR Data Principles in precision oncology: review of supporting initiatives. Brief Bioinform [Internet]. 2020 May 18 [cited 2021 Jul 21];21(3):936–45. Available from: https://doi.org/10.1093/bib/bbb204
[17] IHE SOA Work Group. HL7 FHIR and FAIR principles (FAIR4FHIR) [Internet]. FHIR for FAIR Implementation Guide. 2021 [cited 2021 Sep 16]. Available from: shorturl.at/pVMo7
[18] Kush RD, Warzel D, Kush MA, et al. FAIR data sharing: The roles of common data elements and harmonization [Internet]. Vol. 107, Journal of Biomedical Informatics. J Biomed Inform; 2020 [cited 2021 Jul 25]. Available from: https://pubmed.ncbi.nlm.nih.gov/32407878/
[19] FHIR UK Core - NHS Digital [Internet]. NHS Digital. 2020 [cited 2020 Apr 8]. Available from: https://digital.nhs.uk/services/fhir-uk-core
[20] Willems M. FAIR Data Maturity Model: specification and guidelines [Internet]. 2020 [cited 2021 Sep 15]. Available from: shorturl.at/inqKO