Cookbook for Regional Interoperability Detailed Design Paper #031

Data Release Management

PRELIMINARY DRAFT

Version 1.2 – 26th October 2021

Abstract Interoperability Cookbook Anchor Points

Section	Title
3.1.2	FHIR Service Bus
4.1	FHIR Service Point

Table of Contents

Version Control

Version	Release Date	Released By	Reason for Release
1.0	03/01/2021	Robert Hickingbotham	Preliminary Draft
1.1	11/01/2021	Robert Hickingbotham	Incorporate Comments From TD
1.2	26/10/2021	T Davey	Add enhancements to Contextual Data Filters

Reviewers

Initials	Name	Role	Organisation
TD	Tim Davey	YHCR Architect	PA Consulting

1 Introduction

1.1 Purpose of this Document

This document is one of a series of design papers which underpin the Abstract of a Cookbook for Regional Interoperability (the Abstract Cookbook). These papers, in their totality, describe the technical components and the standards which form the YHCR System of Systems (SoS). They are intended as a basis for developing or procuring software and so are expressed at a level of precision which aims to avoid ambiguity but consequentially, they are focussed to technical readers.

Design papers are anchored to topics which are discussed in the Abstract Cookbook. They are elaborations of the concepts which were first introduced by the abstract and new content is further detail rather than variations of previously established core principles.

This document (design paper 031 - “Data Release Management”) specifies a design for components running in the regionally hosted FHIR Aggregator which simplify control over the release of data for data providers.

A principle of the SoS is that data providers have discretion (subject to adherence to regional standards and information governance framework) when releasing data to data consumers. A data consumer provides details of its reason for accessing regional data and the type of user making a request. This information is passed on in the content of a JWT bearer token (design paper 005 – "Identity and Access Management") to data providers who can use it to determine whether to release data and to tailor data content to usage.

These options and the discretion available to data providers are not altered by this paper: a data provider is still at liberty to implement appropriate measures to control data which flows from their boundary. Rather, this paper has been triggered by a recognition that many data providers will wish to operate similar controls over the release of data and the implementation of certain features centrally will simplify onboarding generally resulting in lower cost implementation and greater confidence that the SoS is being used transparently for purposes mutually agreed between its participants.

1.2 Why Restrict Access to Data?

The SoS has been designed for meeting a multitude of uses including:

- direct care;
- population health management;
- algorithmic analysis of condition development;
- care co-ordination;
- engagement of the patient in their own care;
- etc.

There are several classes of people who will use the SoS:

- clinicians and social care workers with a direct relationship with the patient;
- care administrators;
- researchers;
- patients;

- carers.

Thorough information governance dictates that data providers need be selective when releasing data, being cognisant of how it will be used, for what purpose and by whom. These factors affect the applicability of data sharing agreements, the safe interpretation of data, and the sensitivity of the data.

It should be noted that these factors relate to the context in which data is being requested. For instance, it may be appropriate to release data for the purpose of direct care for use by a clinician with a legitimate care relationship. It may not be appropriate to release data for other uses. The context is established by the data consumer in their claim to access the SoS (design paper 005). The control can be implemented as a relatively simple filter based on the properties of the claim.

There are other non-contextual reasons for filtering data. Several data providers are dependent on their system vendor for a connection to the SoS. Unlike data providers who are self-determinant in extracting data from core systems. These organisations are less in control of the data types offered to the SoS. Even with vendor provided selectivity over the resource types made available, there will be circumstances where a decision to turn on a data feed is undesirably binary.

Organisations wishing to review, and clean data will organise their work in different ways, potentially focusing on service lines, care teams or patient cohorts. Delaying the release of any data until all data is clean would set back contribution to the SoS and attainment of clinical benefits. A regional capability which could allow these organisations to specify rules which selectively controlled the release of data could overcome this obstacle. Examples of rules which may exclude data for release:

- which has been collected by particular healthcare services;
- where the patient is in the care of particular care teams;
- data which has been recently recorded.

1.3 Data Sensitivities and the Responsibilities of Data Consumers

Whilst the discussion above has focused on the responsibilities of Data Providers in controlling data release, there are also responsibilities for a Data Consumer. Data which is normally legitimate to share for direct care and other purposes may have sensitivities which should lead to special treatment by data consumers. Examples of sensitivities include:

- data whose subject is an employee or family member of an employee;
- a diagnosis which has not yet been shared with a patient;
- potentially stigmatizing data pertaining substance abuse, genetic diseases, psychiatric conditions etc.

The context information supplied to data providers is insufficient to determine whether sensitive data should be withheld – the decision may depend on subtle factors which are particular to the organisation retrieving the data. Withholding data is only one possible treatment, others include logging access, alerting a supervisor, displaying warnings alongside the data.

This paper proposes use of the HL7 FHIR meta data standard for tagging sensitive data to enable an appropriate response by data consumers.

1.4 Controlling Data Filters

Data filters must be under the direct control of the organisation acting as a data provider. This paper proposes that the onboarding suite hosts a user interface that allows the technical officers at a participating organisation set up and parameterise filters. This will be done on an environment-by-environment basis.

1.5 Relationship with Data Access and Consent Management

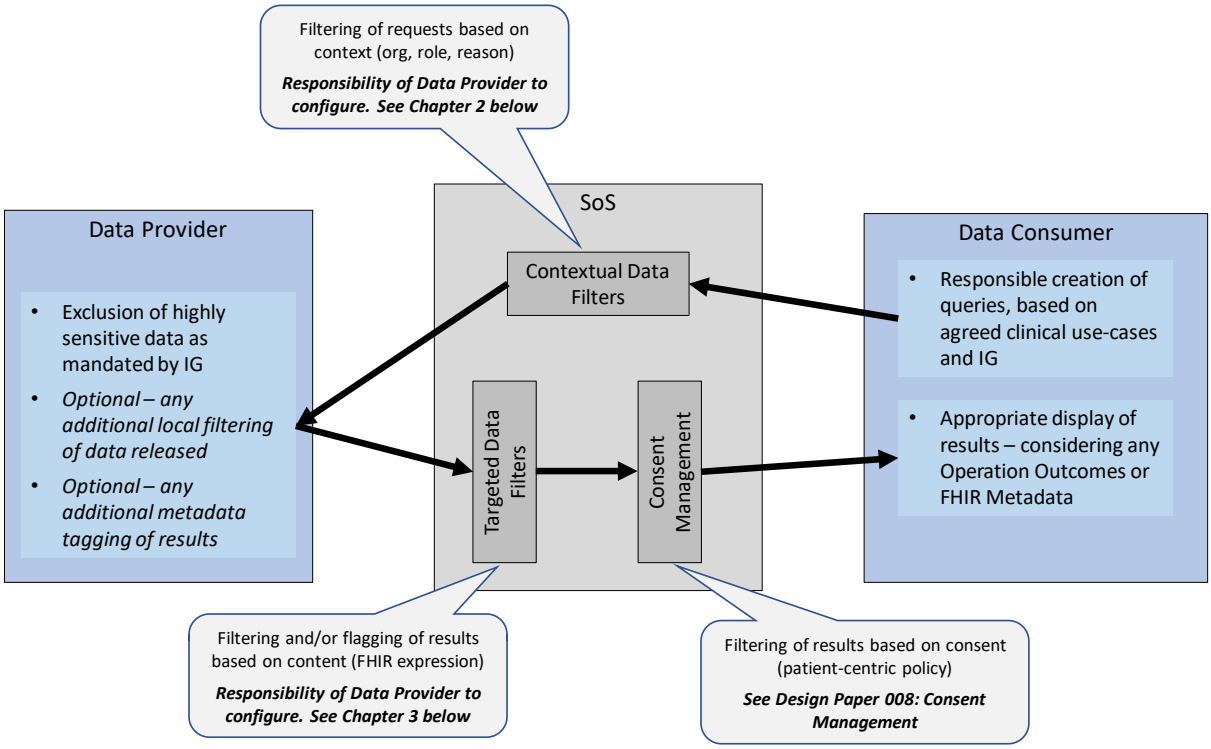
Design paper 008 describes regionally hosted components that allow data to filtered (or released with qualification) in accordance with machine interpretable policies which describe the context in which they apply and detail the attributes of data items that they cover.

These are principally the features that are required to support the requirements identified here and it is assumed that much of the implementation of this componentry will be reused for data filtering. Key differences between Data Access and Consent Management and the Data Filters described here are:

- data access policies are patient centric, a policy is applied to individual patients;
- data access policies require that all data is retrieved from a data and provider and then filtered whereas the implementation of some data filters will prevent queries being forwarded to a data provider;
- data access policies are not accounted for by the Data Availability Service (design paper 002) whereas some data filters will result in the presence of data being hidden from data consumers;
- data access policies are not controlled by data provider organisations.

1.6 Summary of Responsibilities and Technical Mechanisms

Taking into account all of the above, the diagram below summarises the end-to-end responsibilities and technical mechanisms available for controlling data release in the System of Systems:



1.7 Relationship of this Document with Other Standards

This document assumes familiarity with the HL7 FHIR STU3 Metadata standard:

- HL7 FHIR STU3 Resource Metadata standard.

1.8 Intended Users of the This Document

Developers of the SoS core components, data providers and data consumers.

2 Contextual Data Filters

Contextual data filters enable a decision to be made by the FHIR Aggregator (design paper 010) as to whether to issue a query to a data provider. The decision is made solely based on information included in data consumers claim to the Identity and Access Management Service (design paper 005) and carried in the JWT bearer token accompanying a request for data.

A contextual data filter will only prevent a transaction from being issued to a data provider which would have otherwise been forwarded. A data filter cannot result in transaction being issued to a data provider which would not otherwise have taken place.

Transactions include queries, direct resource retrieval, and data management actions. They apply to the following interaction patterns:

- synchronous;
- asynchronous;
- subscription creation.

They do not apply to:

- subscription notification;
- reliable messaging receipt or dispatch.

2.1 Configuration of Contextual Data Filters

Data filters are configured in a user interface within the onboarding suite for data providers. The configuration can be read and modified by a technical officer from a participant organisation or by administrators of the SoS.

Data filters are defined separately for Sandpit, Staging, System Test and Production environments.

The following filtering mechanisms are available:

1 Provider / Consumer Pairings

These allow a Data Provider to select the list of Data Consumers they will accept requests from. This might be used, for example to pilot with a smaller group before full rollout.

- The default is a blank list, which will allow ALL Data Consumers
- An “allow” list may be configured – in this case the Data Provider will ONLY accept requests from the listed Data Consumers
- Alternatively a “deny” list may be configured – in this case the Data Provider will accept requests from any Data Consumer EXCEPT those listed

2 Permitted Roles and Reasons

This allows a Data Provider to select the IAM Roles and/or Reasons which they will accept.

The reader should refer to design paper 005 for currently valid reason and role codes. At the time of writing the following codes are operating in the YHCR.

Code	Reason
1.1	Direct care (Emergency). Access is in the context of a patient;

1.2	Direct care (Non-emergency). Access is in the context of a patient.
2	Indirect care with the consent of the patient. Access is in the context of the patient.
3	Indirect care not in the context of a patient. (Not patient-centric).
4	Analytics with access restricted to pseudonymised data. (Not patient-centric).
5	Administration (Not patient-centric).
6	PDS Trace
7.1	Clinical Safety Testing (data)
7.2	Clinical Safety Testing (UI)

Code	Role
1	Clinical Professional.
2	Social Care Professional,
3	Citizen.
4	System or Robot.
5	Administrator (of YHCR systems).
6	Auditor.
7	Authorised Carer.

The default is to allow all Roles and Reasons

3 FHIR Resource Type Publication

This allows a Data Provider to configure the types of FHIR Resources which it will publish (eg Patient, Encounter, DocumentReference, etc). The FHIR Aggregator will optimise queries and only route to Data Providers who indicate via this list that a particular type of FHIR Resource is supported.

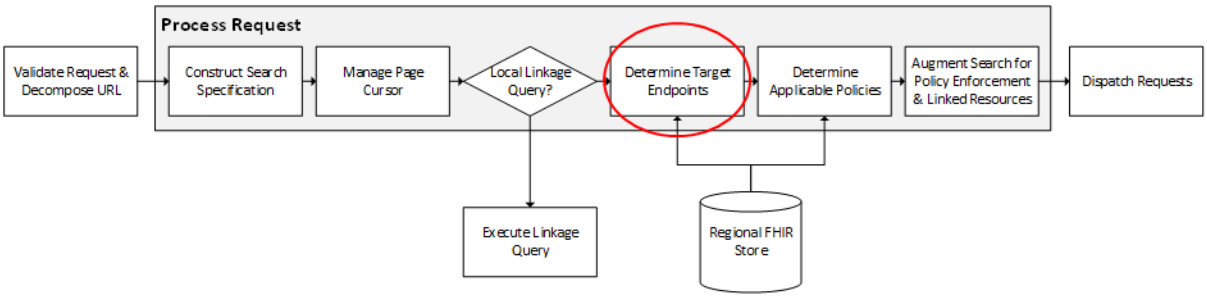
Each type of FHIR Resource in the list has a status:

- ***Public*** – this resource is generally available to all users
- ***Clinical Safety*** – this resource is restricted for pre-release clinical safety testing. It may only be accessed by users with the Reason for Access of "*7.1 Clinical Safety Testing (data)*"

Each type of FHIR Resource in the list also has the opportunity to enter a paragraph of free text description. This may be used to provide more information to users, eg including the extent of coverage, any known deficiencies, etc

2.2 Application by the FHIR Aggregator

The FHIR Aggregator design (design paper 010) is impacted in the following location:



This processing step determines a list of data providers to which a transaction or query will be dispatched. Data filters will require an enhancement to test the context of the interaction for each of the data providers against their configured filters. Filters should be processed in order. The first filter which matches the context of the request will determine whether the provider is retained or removed from the list.

Note that this enhancement will block interactions with providers which they would normally have received because:

- contact has been registered with PIX (design paper 004);
- the provider is explicitly referenced in the interaction through a resource identifier or meta data source tag;
- an allowable non-patient centric query is issued to all providers.

2.3 Implications for the Data Availability Service

Prior to this enhancement the Data Availability Service (design paper 002) was context insensitive.

Its role is to report whether data is available from the SoS for a patient and optionally to detail the data providers from which it is available.

This enhancement requires the service to take into account the context of the request and to modify the data providers included in results for operational contextual data filters.

This has the implication that different data consumers and users will have different impressions of what data is available for a patient from the YHCR.

3 Targeted Data Filters

Targeted data filters enable a data provider to specify data items that should:

- be excluded from search a request or direct resource retrieval request;
- be accompanied in search results by a data impairment report;
- be annotated as being sensitive, the definition being expanded on in section 4.

If resources are excluded from search requests, then optionally (and probably normally) a data impairment report will be added to explain the data omission.

This facility will allow data providers to selectively enable data to flow from a data feed to SoS where there are potential problems known about a subset of the data. This subset can be released from the source system but excluded or qualified in transit through the FHIR Aggregator.

The rules which exclude data are deterministic and are based on FHIR expressions which can be evaluated by the FHIR Aggregator against FHIR resources returned from a data provider. If one of the expressions matches, then the resource matches the filter and an action is taken.

3.1 Configuration of Targeted Data Filters

It is intended that targeted data filters be a tool which allow an organisation to progressively release data to the SoS during an onboarding and data maturing process.

Data filters are configured in a user interface accessible from the onboarding suite for data providers.

People reviewing and cleansing categories of data in local organisations will express categories of data in a clinical language which the onboarding suite must translate to machine interpretable expressions that can be evaluated against individual FHIR resources. For example, the business requirement might be to exclude data captured in the course of encounters with Psychiatric services. This will translate to several resource specific FHIR expressions in the form:

Observation: encounter.incomingreferral.service=Psychiatry

Administrators of the SoS will be responsible for defining the meaning of a business context in terms of FHIR expressions. Users of the onboarding suite in organisations will use the business context to define active filters. Design paper 008 – "Data Access and Consent Management" specifies in detail the use of FHIR expressions to filter data.

Active filters can be setup and maintained by technical officers from the participant organisation. The data filter configuration defines an action to be taken when a data item matches a targeted filter. Options comprise:

- exclude the data item from the result set silently;
- exclude the data item from the result set and add a data impairment report in the form of an *OperationOutcome* resource;
- include the data item in the result set and add a data impairment report in the form of an *OperationOutcome* resource.;
- include the data item in the result set add a metadata sensitivity code and label from the YHCR coding system (see section 4).

OpertionOutcome resources are selected from a standard bank maintained centrally.

Data filter configuration can be accessed in a read-only form by administrators of the SoS.

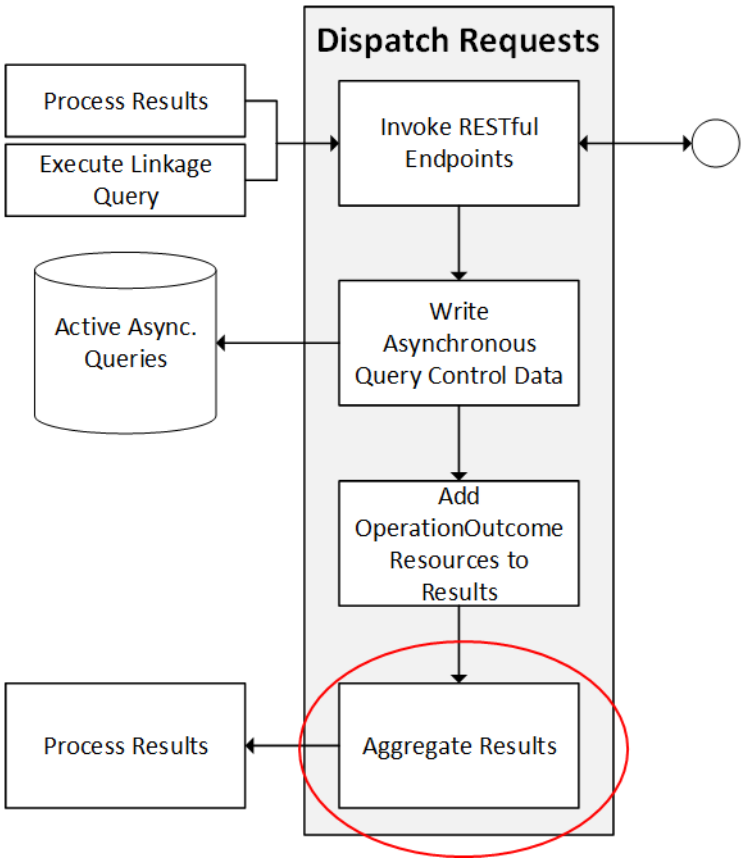
Data filters are defined separately for Sandpit, Staging, System Test and Production environments. The Onboarding Suite should make it simple to replicate filters between environments.

An unlimited number of filters may be defined for a data provider. A filter may optionally have a data range defining its period of applicability.

All changes to filters are logged.

3.2 Application by the FHIR Aggregator

The FHIR Aggregator design (design paper 010) is impacted in the following location:



When processing a search request this step builds a consolidated bundle of search results. The enhancement requirement required will be to execute FHIR expressions against every resource returned by a data provider for active targeted data filters.

Resources may be excluded from the bundle. Note that other resources in the bundle may reference excluded resources. These references will not be modified by the FHIR Aggregator.

OperationOutcomes may be inserted into the bundle. A single *OperationOutcome* will be inserted for all resources affected by a business rule. If data is excluded by the rule then the *OIperationOutcome* will explain why. If the data is included but qualified then the *OperationOutcome* will cross reference the affected resources using the issue.expression attribute. The use of OperationOutcomes is detailed in design paper 017 – "Data Impairments".

4 Publishing Sensitive Data

The definition of what constitutes sensitive data is broad, complex, and sometimes subjective and this subject is left for the Data Architecture Design Authority (DADA) to progress. This paper offers a technical mechanism for a data provider to annotate individual resources as being sensitive and to codify the reason why. The actions that should be taken by a data consumer when receiving notification of a sensitivity will be defined by DADA as regionally important sensitivities are defined.

The mechanism uses the FHIR meta data standard for applying security labels (includes sensitivity labels) to resources. The standard uses an extensible coding system defined here which allows data providers to inform consumers of sensitivity of individual resources and the use of the coding system is a responsibility of data providers when mapping data. As an extensible value set, the YHCR is able to define additional concepts and determine their interpretation by data consumers.

The YHCR extension coding system will be published on the YHCR Metadata server (design paper 030).

Appendix 1 – Maturity Matrix

Section	Narrative	Consultative	Draft	Normative
1 Introduction 1.1 Purpose of this Document	X			
1.2 Why Restrict Access to Data	X			
1.3 Data Sensitivities and Responsibilities of Data Consumers		X		
1.4 Controlling Data Filters		X		
1.5 Relationship of this Document with Data Access and Consent Management		X		
1.6 Relationship of this Document with Other Standards	X			
1.7 Intended Users of the This Document	X			
2 Contextual Data Filters 2.1 Configuration of Contextual Data Filters			X	
2.2 Application by the FHIR Aggregator			X	
2.3 Implications for the Data Availability Service			X	
3 Targeted Data Filters 3.1 Configuration of Targeted Data Filters			X	
3.2 Application by the FHIR Aggregator			X	
4 Publishing Sensitive Data			X	