



**INTERWEAVE**  
CONNECTING CARE

Cookbook for Regional Interoperability  
Detailed Design Paper #019

Document Management

PRELIMINARY DRAFT

Version 2.1 – 5<sup>th</sup> July 2021

**Abstract Interoperability Cookbook Anchor Points**

Section	Title
3.1.7	XDS/FHIR Document Registry

## Table of Contents

1	Introduction .....	4
1.1	Purpose of this Document .....	4
1.2	Relationship with IHE XDS .....	4
1.3	Medical Media and DICOM.....	5
1.4	FHIR Documents .....	5
1.4.1	Compatibility of Document Based FHIR provision and the YHCR.....	5
1.4.2	The YHCR as Mediator for Document Based Data Sources .....	6
1.5	Relationship of this Document with Other Standards.....	6
1.6	Intended Users of the This Document.....	6
2	Relationship with Other Components of the System-of-Systems.....	8
2.1	Requirements for Regional Infrastructure.....	8
2.2	Document Proxy Server .....	9
2.2.1	Security .....	9
2.2.2	Purging Document Locator Records .....	10
2.2.3	Auditing.....	10
2.3	Requirements for Data Providers .....	10
3	The YHCR <i>DocumentReference</i> Profile .....	11
4	FHIR Documents and Compositions in the YHCR .....	13
4.1	Use of FHIR Compositions as a History of Care .....	13
4.2	FHIR Documents and Transfer of Care .....	13
4.3	FHIR Documents as an Interfacing Mechanism.....	13
4.4	Decomposition of Vendor Supplied Bundles .....	14
5	YHCR Document and Composition Production .....	16
5.1	Composition Template .....	16
5.2	Composition Generator .....	17
	Appendix 1 – The YHCR Document Architecture, XDS and MHD .....	19
	Appendix 2 – A Primer On FHIR Documents and Compositions.....	20
	Appendix 3 – Additional Guidance on document content .....	21
	Appendix 4 – Maturity Matrix .....	25



## 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 to inform the development or procurement of software and so are expressed at a level of precision which aims to avoid ambiguity but with the consequence that they are targeted at 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 019 - "Document Management") establishes an architecture for documents in the YHCR. It is expressed in two parts.

Part 1 concerns itself with the dissemination and consumption of what might be considered traditional documents: reports, discharge summaries, clinic letters etc. These documents provide a narrative of medical history and are produced in volume by clinicians who communicate pertinent facts to other clinicians involved a continuity of care.

Part 2 focuses on structured documents and FHIR compositions. These are a collection of coded data items which are packaged alongside human readable narrative. FHIR documents and compositions are the basis of several national standards which include standards for transfers of care from acute inpatient episodes, emergency department encounters and mental health encounters. The YHCR is also modelling its work on the development of standards from transfers of care from an ambulance on the national transfer of care initiative and it will be proposing a standard for a FHIR *Composition* for national adoption.

The aim of Part 2 is to show how the YHCR's FHIR based infrastructure can be used by its participants to simplify the production of FHIR *Compositions* and documents and to propose a regional standard for expressing the structure of compositions in a manner that enables them to be automatically generated.

### 1.2 Relationship with IHE XDS

The body, Integrating the Healthcare Enterprise (IHE), publishes a profile or standard for cross-enterprise document sharing (XDS). Whilst the standard is not prevalent in the YHCR there are instances of XDS compliant technology (a document exchange in South Yorkshire and document management systems such as Kainos and WinDIP). XDS is also an important technology for the NHS in other regions. The YHCR document architecture should be cognisant of the desirability for compatibility with the standard without being constrained by it.

The IHE has put work into reconciling XDS with the FHIR standard and has published the Mobile access to Health Documents (MHD) standard for translating an XDS implementation into a FHIR based presentation of XDS data structures. This design is consistent with MHD. It would allow an MHD compliant client to consume documents from the YHCR. It would also be possible to bridge the

YHCR document architecture to allow it to be represented as an XDS document registry and repository, although this would entail further development which is outside of the scope of what is envisaged by this design.

Further details of the relationship between the YHCR document architecture and XDS are provided in Appendix A.

### **1.3 Medical Media and DICOM**

A long-term goal of the YHCR is likely to include improving access to medical images and videos. Images and videos can be considered as traditional documents and as such, subject to file size, are candidates for sharing using the document architecture described in Part 1. However, there are constraints inherent in the architecture's transport protocols which will limit the size of file which can be transported to a few 100MBs. A hard limit will be published in the YHCR Operations Guide.

If the YHCR is to become a regional transport layer for use by Picture Archiving and Communication Systems (PACS) then further work is required to define the application of the Digital Imaging and Communications in Medicine (DICOM) standard within the YHCR. This work is outside the scope of this paper.

### **1.4 FHIR Documents**

This paper assumes that its readers are familiar with the concept of FHIR documents and a related *Compositions*. A primer is offered in Appendix 2.

A new breed of document is emerging which is based on FHIR standards. These documents are constructed of FHIR resources and contain structured data packaged alongside textual narrative to aid human interpretation.

FHIR documents are important to the YHCR because:

- national standards for transfer of care, GP-Connect and maternity systems dataset rely on them;
- the YHCR is using FHIR documents to meet some early-day use cases including transfer of care from ambulances;
- some vendors' and national systems' FHIR based integration capabilities rely on producing a FHIR document (or similar FHIR structure) rather than making atomic resources available for on-demand query.

This last point requires some elaboration as it leads to a potential incompatibility between systems which legitimately claim FHIR compliance and the YHCR.

#### **1.4.1 Compatibility of Document Based FHIR provision and the YHCR**

For the purpose of direct care, most data which is accessed through the YHCR will be acquired through direct query. A user interface displaying, for instance, a list of allergy intolerances known by care settings for a patient will issue a query for *AllergyIntolerance* resources. A system, whose sole interface mechanism is the publishing of FHIR documents will not be able to respond to the query. It may have information about allergy intolerances and the FHIR document that it produces for the patient may include an *AllergyIntolerance* resource in its content, but the resource is invisible to querying systems.

A data consumer might elect to solve the problem by also issuing queries for FHIR documents (technically a query for a FHIR Bundle of type document) and extracting any *AllergyIntolerance* resources from the documents it receives. This approach is computationally expensive for the consumer, potentially requires large amounts of irrelevant data to be sourced and represents a clinical safety risk.

Extracting individual resources from a document and presenting them to a user as if they had been acquired as independent atomic resources is risky because context implied by the publisher is lost in the process. Consider for example, a document containing both a *CarePlan* and a *Condition*. The *Condition* identifies that the patient is suffering from anxiety. The document, when consumed in its entirety, would imply that the *CarePlan* is related to the assessment of anxiety. Context is lost if the *CarePlan* is extracted and presented independently of the *Condition*.

The *CarePlan* could have been coded to explicitly reference the assessment of anxiety. The resource definition allows this by providing an "addresses" attribute which could link the *CarePlan* to *Condition*. However, the link is optional and the author of a document whose structure makes the association explicit could well consider this an unnecessary duplication of information.

#### **1.4.2 The YHCR as Mediator for Document Based Data Sources**

In principle the YHCR could offer an adapter that extracted resources from documents and offered them in atomic form to the YHCR. The adapter would allow consumers to 'drill' into FHIR documents and could be constructed in a generic fashion to ingest data from any document. However, the adapter must do this in a clinically safe way and only extract resources which can be consumed as self-contained units.

The adapter will augment FHIR resources contained in documents to enable them to be consumed safely where feasible. Applying the example given above, the adapter would augment the *CarePlan* resource to reference the *Condition* in the "addresses" attribute.

Regardless of the mediation provided by the YHCR, data consumers will be able to query and retrieve document and other bundles via the SoS from the document source.

### **1.5 Relationship of this Document with Other Standards**

This design paper reference the following standards

- [HL7 FHIR STU3](#)
- [Mobile access to Health Documents \(MHD\)](#)
- [Digital Imaging and Communications in Medicine \(DICOM\)](#)
- [HTTP/2](#)
- [Extensible Stylesheet Language Transformations \(XSLT\)](#)

### **1.6 Intended Users of the This Document**

Developers of the System of Systems, data providers and data consumers.

## Part 1 – The YHCR Document Architecture

## 2 Relationship with Other Components of the System-of-Systems

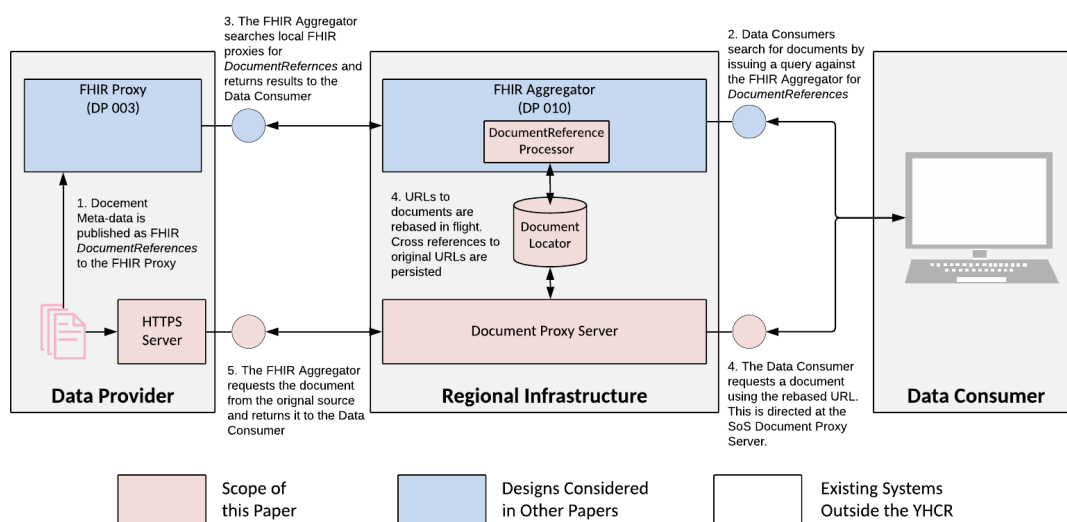
The YHCR document architecture is, in the main, an application of the generic YHCR architecture for data sharing. It uses FHIR resources to describe the documents which are accessible through the YHCR. The FHIR resources are provisioned by data providers who have documents to share and are accessed by data consumers who use the regional FHIR infrastructure as a FHIR aggregator. The process uses the same security infrastructure as all other data sharing in the YHCR and FHIR resources describing documents can be exchanged between data providers and data consumers using the standard interaction patterns being synchronous query, asynchronous query, subscriptions, and reliable messaging.

The *DocumentReferences* and FHIR resource is used to describe documents. Standard search terms allow documents to be located by subject (usually the client or patient), the author, the type of document and several other elements. By querying *DocumentReference* objects a consumer can assemble a catalogue or timeline of documents for a patient or client. Given the historically document centric nature of medical records this alone is a valuable capability to support direct care.

A *DocumentReference* may embed the document that it describes and if it has been created to support a specific interaction with a known participant then this may be an appropriate usage. However, for discovery type use cases where consumers need to understand what information is available before accessing the detail then it is appropriate to separate the *DocumentReference* from the document itself. The *DocumentReference* should reference the document through its content attachment URL property. Please refer to Appendix 3 for further details and requirements.

In most cases the URL will be served from an endpoint on the data provider's network. Regional infrastructure can improve security and simplify firewall requirements by providing a proxy over which the document can be retrieved. The proxy's endpoint is secured to the YHCR standard. The YHCR FHIR aggregator replaces URLs with a URL served from the proxy endpoint which brokers access to the document on behalf of the data consumer.

The document architecture and its relationships to the System of Systems is illustrated below.



### 2.1 Requirements for Regional Infrastructure



The regional FHIR Aggregator provides enhanced processing for all *DocumentReference* resources which transit through its synchronous query, asynchronous query or subscription pathways.

For all document references with a non-null content attachment URL, the FHIR aggregator replaces the URL with a URL directed at a regional Document Proxy Server. The rebased URL embeds an identifier which can be used by the Document Proxy Server to uniquely determine the original URL on which the document is available.

The relationship between the identifier and the original URL are persisted in a Document Locator store. The Document Locator also persists (for security reasons discussed below) details from the original Javascript Web Token (JWT) which was used to gain access to the YHCR and used to initiate the data flow. The original JWT is:

- for a synchronous query, the JWT used in HTTP GET request;
- for an asynchronous query, the JWT used to place the query;
- for subscriptions, the JWT used to place the subscription which gave rise to the notification.

Note that subscription notifications may be distributed to more than one subscriber (see design paper 007 – "Subscriptions Infrastructure"). This has the consequence that a single *DocumentReference* transiting in a subscription notification may result in multiple Document Locators and multiple resultant *DocumentReferences* – one for each subscriber.

## 2.2 Document Proxy Server

The Document Proxy Server is an HTTPS server which supports HTTP 1.1 (and specifically chunked transfer encoding) and is secured according to the principles established by the design paper 005 – "Identify and Access management". Chunked transfer encoding is used to synchronise receipt of data from a data provider with its dispatch to the data consumer.

HTTP/2 has been considered and discounted because of the additional implementation cost. Whilst there would be performance benefits from using HTTP/2 binary frames (approximately 15%), there is unlikely to be uniform support for HTTP/2 from all data consumers and so this would require the regional infrastructure to support both HTTP 1.1 and HTTP/2 to be universally accessible.

The service implements the HTTP GET verb and on performs the following processing.

1. Extracts a document identifier from the request URL.
2. Retrieves the persisted Document Locator and in the event that the URL is an invalid format, or the Document Locator does not exist, returns with an HTTP error code.
3. Validates that the JWT details recorded on the Document Locator matches the relevant JWT attributes presented in the request, if necessary, returns an HTTP error code.
4. Connects and issues HTTP headers to the endpoint identified in the Document Locator and if the connection fails return an HTTP error code to the data consumer.
5. Sends to the data consumer in turn each transfer encoded chunk received from the data provider.

### 2.2.1 Security

Connections from the document proxy server to a data provider's HTTPS endpoint are secured using TLS1.2 using mutual authentication of certificates. Certificates installed at the data provider are signed by the YHCR certifying authority.

Attributes from the JWT presented on document retrieval are validated against attributes of the original JWT of the request which initiated the retrieval. *DocumentReference* resources is performed to secure the integrity of any data access control policies operating at the boundary of the data provider.

By releasing a *DocumentReference* a data provider declares that the context under which the *DocumentReference* complies with any data access policies in force and that BOTH the metadata and document itself may be released to the data consumer.

The context of the request is defined in the JWT by the following attributes:

JWT Attribute	Explanation
iss	The application which issues the claim.
ods	The ODS code of the organisation that issued the claim.
rsn	The reason for the access request.
usr.rol	The user's role.
usr.org	The ODS code of the organisation who employs the user in the capacity for which they are accessing the YHCR.
usr.rel	The user's relationship with the patient.

The model FHIR Proxy (design paper 003 – "Conceptual Design for a FHIR Proxy Server") allows a data access policy to be enforced when serving resources. It is possible that documents will be served from a relatively simple HTTPS server and that there will not be equivalent access control mechanisms. The YHCR enables the data provider to operate a simple document architecture by ensuring that the context in which a *DocumentReference* is the same as that used in a subsequent request for the referenced document.

### 2.2.2 Purging Document Locator Records

Document Locator records are periodically purged. The periodicity is specified in the YHCR Operations Guide.

### 2.2.3 Auditing

Audit records should be written by the document proxy server in compliance with design paper 009 – "Auditing". Records should be written for:

- failed attempts to retrieve documents;
- documents successfully retrieved.

## 2.3 Requirements for Data Providers

A data provider who operates a FHIR Proxy server has the technical capability to serve FHIR resources including *DocumentReferences* in compliance with regional standards subject to their ability to prepare the data. See Appendix 3 for further details and requirements.

The document architecture additionally requires the provider to implement a mechanism for serving files over HTTPS. There are many product options available and the choice of technology is left to the data provider.

### 3 The YHCR *DocumentReference* Profile

The YHCR Maturity Model (design paper 023 – "The YHCR Maturity Model") provides a framework for evolving data quality over maturity levels. At the highest level of maturity the model sets a goal of all FHIR resources to comply with Care Connect profiles, YHCR refined versions of Care Connect profiles, or YHCR defined profiles where Care Connect does not provide a profile.

The objective for the YHCR *DocumentReference* is consistent with this approach and is a refinement of the Care Connect [DocumentReference](#).

At this stage the YHCR refinements are:

Element	Resource Cardinality	Profile Cardinality	Implementation Notes
type	1..1	1..1	Binding (mandatory): SNOMED CT UK document type ( <a href="https://fhir.hl7.org.uk/STU3/ValueSet/CareConnect-DocumentType-1">https://fhir.hl7.org.uk/STU3/ValueSet/CareConnect-DocumentType-1</a> )  (The Care Connect binding is preferred).

Additional refinements may be introduced by the Data Architecture Design Authority as groundwork in modelling and consuming document meta data is undertaken.

The Care Connect profile constrains all resource references from the *DocumentReference* to be to resources which themselves comply with other Care Connect profiles. This will not necessarily be feasible at lower levels of maturity and this requirement will be relaxed by the maturity model.

## Part 2 –FHIR Documents and Compositions

## 4 FHIR Documents and Compositions in the YHCR

This section requires an understanding of FHIR Compositions and documents. Readers who are not familiar with these concepts are guided to Appendix 2 which provides a brief refresher,

FHIR *Compositions* and documents provide a snapshot of data at a point in time and are useful devices for collating a set of data at key points in a care pathway. The immutability of the FHIR document provides for an attested audit trail of information on which care professionals can agree as representing the facts as known at the time that the dataset was produced. The national standards for [datasets to be exchanged on handover of care](#) depend on FHIR documents for this reason.

### 4.1 Use of FHIR Compositions as a History of Care

A FHIR *Composition* which is assembled from references to specific FHIR resource versions and has been attested and finalised by a clinician acts as bookmark in a spell of care. It collates all the relevant information known at a point in time, and through the mechanism of FHIR versioning, is immutable and therefore is a reliable source of evidence of the care provided. A chain of *Compositions* which are generated during a spell of care establish causality and visibility of the care pathway which has been followed. It is thought that *Compositions* will be a tool which the Data Design Authority will wish to use in maturing data across the region. Because *Compositions* reference resources rather than package them up for transfer to other locations, they are entirely compatible with the design principles of the YHCR which seek to federate data control and enable consumers to acquire data on demand.

### 4.2 FHIR Documents and Transfer of Care

FHIR documents are a convenient mechanism for packaging data in a handover of care from one care setting to another: the data requirements can be anticipated, the parties to the handover are known, and the document will usually be delivered to the receiving party using a messaging service which means that all relevant information is included in a single package rather than requiring the receiving technology to issue follow-up queries.

A precedent has also been established by national transfer of care initiatives which have defined FHIR documents as the means for transferring care to GPs.

There is a potential consequence of moving FHIR resources in documents from one care setting to another. It may suggest that it is legitimate to take a copy of the data. A core principle of the YHCR is that it aims to reduce duplication of data by making data available on-demand from its original source as it is required by whoever consumes it. Sending a FHIR resource packaged in a document invites a consumer to persist it and if done so leads to data duplication and conflicts in determining the provenance of data.

### 4.3 FHIR Documents as an Interfacing Mechanism

Less compatible with YHCR design principles is the use of documents and other FHIR *Bundle* types to package the 'current' medical record for consumption on-demand. Bundling of data has consequences:

- the provenance of FHIR resources is often lost with the implication that it is difficult to merge with other 'current' medical records;
- it constrains data consumers in how they acquire data and presumes how the data will be used;
- it implies that the context of the data is defined by the dataset as a whole and so it may be clinically unsafe to extract individual FHIR resources and use them independently.

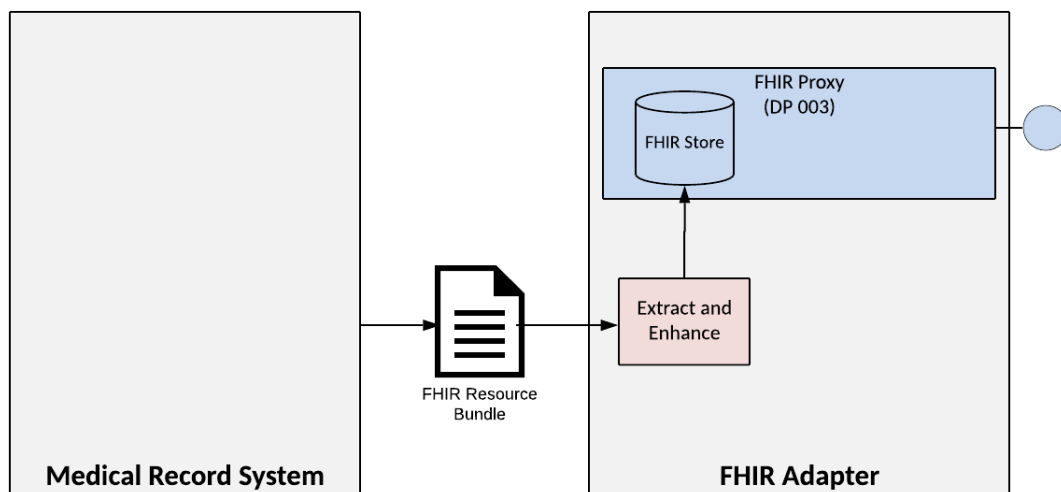
However, the YHCR will encounter systems that use FHIR in this way and it has a responsibility to reconcile the vendor's method of presentation with its core approach to exchanging FHIR resources as atomic units.

#### 4.4 Decomposition of Vendor Supplied Bundles

A data consumer of the YHCR expects to issue a FHIR search for specific FHIR resource types and to receive an aggregation of resources which meet its criteria from all data providers. A system which presents the 'current' medical record as a bundle will not be represented in the search results. Without prior knowledge the consumer might assume that an absence of data means that no data has been collected which could present a risk to clinical safety.

It is possible for a consumer to target a specific system and request its bundle of resources. However, such point-to-point relationships introduce complexity and are not in the ethos of the YHCR.

It is hoped that over time vendors will have an incentive to support the YHCR approach natively. But in the meantime, the YHCR will almost certainly need to adapt the capability of FHIR-bundle-offering system to allow the data to be safely decomposed.



The function of the Adapter is to service search requests and resource retrieval requests for resource types that are embedded in *Bundles* produced by the source system.

The Adapter consumes *Bundles* which are emitted by the medical record system. There are 3 main consumption strategies:

- a record is emitted when a key event is registered in the source system, such as a baby being born, or a patient's attendance at a clinic;

- a record is produced periodically by the source system;
- a record is requested by the Adapter when a search is received from the regional Aggregator.

Under the last strategy the contents of the *Bundle* may be cached by the adapter to reduce the frequency of requests made against the source system,

Resource data in the *Bundle* may require enhancement to allow it to be interpreted atomically, This could be as simple as introducing a common *Encounter* resource which other resources reference or it may include complexities such as de-duplicating data or merging resources with data obtained elsewhere. The requirements will be established on a case-by-case basis. As a minimum the FHIR Adapter will have the option of releasing an *OperationOutcome* which announces to consumers that there is an impairment in the data available from this provider (see design paper 017 – " Data Impairments - Reporting and Interpretation").

Going forward the YHCR should offer a boiler-plate statement for inclusion in new contracts with vendors which details what is meant by FHIR support.

## 5 YHCR Document and Composition Production

The YHCR, through its FHIR based infrastructure has an opportunity to minimise the burden on its participating organisations in relation to the production of compositions and FHIR documents.

The goals of this tooling are

1. To enable the structure of compositions and documents to be standardised across the region.
2. To enable them to be automatically assembled from FHIR resources which are available through on-demand query.

The tools might operate in local data provider organisations to help them comply with national or regional requirements for FHIR documents such as compliance with national transfer of care initiatives. They may operate regionally to allow compositions to be assembled from a blend of FHIR resources obtained from several data providers in support of events which impact a number of care settings (such as the enrollment of a patient on a care pathway). A data consumer may use the tools to generate compositions in response to events in the consumer organisation (such as a referral to a service operated by the consumer's organisation).

Implementation of the tools described here are not an immediate priority and the purpose of this section is to illustrate how the core capabilities can be used to support FHIR compositions and documents and ensure that the architecture is compatible with national initiatives that depend on these structures.

The approach is based on the concepts of:

- a composition template which describes the structure of a composition and provides instructions for assembling its content;
- a composition generator which interprets the template follows the instruction to build its content.

These are novel concepts which have been envisaged specifically for the YHCR and not based on another standard.

### 5.1 Composition Template

Compositions are related to a context. The context of a discharge summary is the encounter at the care setting from which a patient is being discharged. The context a bed usage report is a hospital ward etc.

A composition template describes how a composition could be constructed given a context. The context is the 'input' from which a composition can be generated, and a template is collection of rules which allow the content for a composition to be acquired using this input.

The context of a composition is a FHIR resource. In the above examples, the discharge summary can be constructed with an *Encounter* resource as input to the rules. The rules describe how the patient demographics medications, conditions, test results etc. can be assembled given an *Encounter* resource as context.

The content of the composition is made up of FHIR resource references and the mechanisms to assemble the relevant FHIR resources are FHIR searches.



A composition template lists the FHIR searches that could be executed given the context. So, the rule that enables prescribed medications to be assembled for the discharge summary might be expressed as:

```
/MedicationRequest&status=active&context=%Encounter%
```

where %Encounter% is a reference to the *Encounter* resource supplied as context.

The composition template is structured in a similar way as to a composition itself. It lists the sections which will make up the generated *Composition* resource and for each section lists the rules that can be used to acquire resources for the section content.

It also describes how narrative can be generated. Sections in a composition can include human readable narrative in HTML form which is intended to be consistent with the structured FHIR resource data by the composition. It should be possible to generate the narrative from the content of these FHIR resources. A standard such as Extensible Stylesheet Language Transformations (XSLT) could be applied.

The composition template might also provide values for other attributes of the *Composition* resource such as title, confidentiality etc.

A standard set of regional composition templates ensures that compositions of standard types are generated consistently by all participants in the region. A composition template, whilst not a FHIR resource, could be structured in a similar way to FHIR resources and could be served from a regional FHIR Store for use in local composition generation.

## 5.2 Composition Generator

The composition generator is software that uses composition templates to produce *Composition* resources on demand. A composition generator takes as input the FHIR resource which defines context, executes the rules (FHIR searches), collects FHIR resources, generates narrative by applying XSLT to the data which has been collected and constructs the *Composition* resource.

Composition generation will normally be triggered by an event such as a patient discharge. In the YHCR events can be subscribed to using the Subscription Infrastructure (design paper 007) and the composition generator could be the target of a subscription notification. The FHIR resource which is the content of the notification in most cases will be the context required by the composition template. However, an interim step to assemble context might be needed.

Composition generation could be triggered by other mechanisms such as an end user making a request. In this case context would need to be provided by the end user.

The following functionality might be envisaged for a composition generator

Function	Processing Performed
Register composition type	Issues a subscription which defines the event on which composition generation will occur.
Receive subscription notification	Locates the composition type for which the subscription was registered. If subscription resource type is not the required context then obtain the context through a FHIR query. Invoke the composition generation.

---

Generate composition	Accepting a FHIR resource as context, retrieve the appropriate composition template from the regional FHIR store and execute, overriding regionally defined attributes with local variants as necessary.
Attest and finalise composition.	Accepting a Patient, Practitioner or Organisation resource, record attestation and mark the composition status as final.
Generate FHIR document	Collect all resources references in a composition and build a document bundle.

Composition templates and software used for composition generation by a data provider could be used anywhere in YHCR with the only configuration difference being the FHIR endpoint against which searches are executed. A data provider can use the same software and templates directed at a local FHIR store as the region or consumers can use when directed at the regional FHIR Aggregator. The difference being the scope of the content of the generated composition or document.

## Appendix 1 – The YHCR Document Architecture, XDS and MHD

XDS defines the concepts of a document registry (which contains meta-data about documents) and a document repository (to which documents can be published and retrieved). The document registry contains meta-data about the following:

- a document (its author, type of document, subject etc.)
- a submission set in which a group of documents were published.
- a folder which groups related documents for same subject.

### FHIR as an Overlay to XDS

The FHIR document architecture has equivalent concepts to the XDS document and submission set metadata in the *DocumentReference* and *DocumentManifest* resources and IHE publishes the Mobile access to Health Documents (MHD) standard for translating an XDS implementation into a FHIR based presentation of XDS data structures.

MHD defines 4 actors: Document Source, Document Recipient, Document Responder and a Document Consumer.

A Document source publishes a document to a Document Recipient. A Data Consumer queries meta-data and retrieves documents from a Document Responder. In most implementations the roles of Document Recipient and Document Responder are assumed by a document management system.

Part 1 of this paper is only concerned with the relationship between Document Consumers and Document Responders. The means by which a participant to the YHCR lodges a document with a document management system or otherwise collates documents for public dissemination is not the concern of the YHCR.

It concentrates on standards for the *DocumentReference* resource. Whilst the YHCR does not preclude data providers from publishing details of collections of related documents using *DocumentManifest* resources, this is considered an alignment with XDS which is unnecessary given the existing use of documents in the region.

## Appendix 2 – A Primer On FHIR Documents and Compositions

### The FHIR Composition Resource

A FHIR *Composition* is a collection of references to other FHIR resources which are all related to a subject. Typically, the subject is a patient or client but can be any FHIR resource.

A *Composition* embeds textual narrative alongside resource references. The intention being that the narrative provides a human readable interpretation of the content which is consistent with coded data in the FHIR resources.

Sections provide structure to a *Composition*. The definition of a section is left to the designer of the *Composition* but given that narrative is provided at a section level, they are a useful construct for binding narrative more closely to individual resources.

A *Composition* can exist within a workflow and its elements can be used to represent stages in an information gathering process such as issuance, attestation and finalisation.

### FHIR Documents

A *Composition* is foundation of a FHIR Document.

A *Composition* is an assembly of FHIR resource references rather than the resources themselves. The FHIR document is a *Bundle* which contains the *Composition* resource alongside the resources that it references.

A FHIR document is immutable – it must not change once it has been published. If one of the underlying resources changes that renders the document obsolete, then it must be republished as a different document.

Compositions and related resources can be packaged together in other types of *Bundle* and some systems use these to package a snapshot of a medical record in a similar manner to a document. For instance, GP Connect uses a *Bundle* type of search set. For the purpose of this paper any use of a *Bundle* to package a composition with related resources has a similar implication – the information is intended for consumption in its entirety and represents a snapshot of a medical record at a point in time.

## Appendix 3 – Additional Guidance on document content

The key FHIR Resource for handling documents is the DocumentReference <http://hl7.org/fhir/stu3/documentreference.html>. However this offers several different ways to transmit the document content and it is important to understand the details of this. The sections below cover each option in turn.

### 1) Embedded document (only supported for messaging NOT online query)

This is the “simplest” approach, as the document content is embedded directly into the DocumentReference. Note that the contentType (mime-type) is provided so that the consumer is able to judge whether it has the ability to display this format of document eg PDF, etc. (This is the same for all three approaches).

For example the “content” section of a DocumentReference would look like:

```
"content": [{
  "attachment": {
    "contentType": "text/html",
    "data":
    "PGRpdiB4123ucz0iaHR0cDovL123dy53My5vcmcvMTk5OS94aHRtbCI+PGgx123W5pY2FsIElOZW
    1zPC9oMT48ZGI2IGNsYXNzPSJkY123WJhbm5lciI+PHA+QWxsiHJlbGV2YW50IG123W1zPC9wPjwv
    ZGI2Pjx0YWJsZSBpZD0123xpLXRhYil+PHRoZWZkPjx0cj48dGg+RGF0ZTwvdGg+PHRoPkVudHJ5PC9
    0aD48dGg+RGV0YWlsczwvdGg+PC90cj48L3RoZWZkPjx0Ym9keT48dHI+PHRkIGNsYXNzPSJkYXRIL
    WNvbHVtbil+MDItSnVuLTE5NjM8L3RkPjx0ZD5NYW1tb2dyYW08L3RkPjx0ZD5Ob3JtYWw8L3RkPj
    wvdHI+PC90Ym9keT48L3RhYmxlPjwvZGI2Pg=="
  }
}]
```

This embedded approach is often suitable for messaging, where a relevant document is being specifically sent (eg Ambulance Transfer of Care). However for online query it has the significant disadvantage of making the DocumentReference resources very large - as the whole document is transferred even when just viewing the list of documents available. Typically there may be many documents about a patient, but only one or two which are relevant for the user to actually view. It is very inefficient (and slow) to transfer the entirety of the documents when a user is only perusing the list.

*(NB: Despite this limitation, GP Connect is an important use-case which returns embedded document content. However GP Connect documents are not returned from ordinary online queries and instead must be specifically requested from a screen developed to meet NHSD accreditation requirements. The reason for embedding is that, having requested a patient’s GP Record, it is almost always going to be desired to view the document content).*

### 2) Linked document (recommended)

This replaces the content in the DocumentReference with a link. For example:

```
"content": [{
  "attachment": {
    "contentType": "text/html",
    "url": "https://my.server.com/documents/21f51e78-a46d-402c-aa22-dd43e0fec530"
  }
}]
```

---

This is clearly much more efficient. This list of DocumentReferences contains only metadata about the documents, and if the user wants to open the document then it is a case of following the link to download it. This performs a simple binary download of the document bytes – as would be provided by hosting the document on any standard web server.

*Note that InterWeave Connect offers a proxy service to avoid the need for point-to-point networking for document downloads. Instead the urls in the DocumentReference are automatically rewritten to point at a central hub which in turn connects to the Data Provider and streams document downloads.*

### 3) FHIR Binary Resource linkage (only recommended if implemented comprehensively)

Yet another possible approach is to replace the content in the DocumentReference with a link to a FHIR Binary Resource <http://hl7.org/fhir/stu3/binary.html>. This is similar in principle to the Linked document approach, but is more complicated for the Data Provider to implement (and for that reason not recommended). Note: that the contentType of the document reference must be the mime type of the document encoded by the binary resource (see GET behaviour below) The “content” section of a DocumentReference with a linked Binary Resource would now look like this:

```
"content": [{
  "attachment": {
    "contentType": "text/html",
    "url": "https://my.server.com/fhirpath/Binary/21f51e78-a46d-402c-aa22-dd43e0fec530"
  }
}]
```

When the link is followed then it again returns the binary content – as would be provided by hosting the document on any standard web server.

*(This is because of a feature of the FHIR specification that treats Binary resources differently from any other resource. When a FHIR server receives a GET request for a Binary resource then it MUST return the binary document content rather than the Binary resource. However the FHIR specification also requires that this behaviour can be overridden by a consumer by inserting the HTTP Accept header with value “application/fhir+json” (or even “application/fhir+xml”). This then returns a more complete FHIR data structure with the document content embedded – although for document viewing purposes it is difficult to envisage why a consumer would desire this)*

### 4) FHIR Composition Resource linkage (only where relevant)

In addition to unstructured human-readable documents such as PDFs, some Data Providers may offer semi-structured documents in the form of a FHIR Composition. Whilst these are valid in-and-of-themselves as FHIR Resources, it is recommended to also “advertise” them via a DocumentReference. Thus providing Data Consumers with a complete list of all documents whether unstructured or semi-structured.

For example:

```
"content": [{
```

---

---

```
"attachment": {
  "contentType": "application/fhir+json",
  "url": "https://my.server.com/fhirpath/Composition/21f51e78-a46d-402c-aa22-
dd43e0fec530"
}
}
```

When the link is followed then it will return the FHIR Composition data structure. (Whilst all consumers will thus be able to show the existence of the composition, evidently not all will have the ability to follow the link and render the data for any given type of composition)

## Summary of Requirements

**Data Consumers MUST support viewing of documents via a DocumentReference resource which contains a link for downloading the actual document content**

- This covers both of (2) and (3) above as, in practice, both contain a link to download the document content

**Data Providers SHOULD use the “linked document” mechanism when publishing documents for on-demand retrieval**

- This is (2) above, which is recommended as a simple and efficient approach.
- Data Providers MUST NOT publish embedded documents for on-demand retrieval, as this is inefficient
- Data Providers MAY publish documents as FHIR Binary Resources. This is equivalent from a Data Consumer perspective as it provides a link to follow to download the document. However it is not recommended due to the additional complexity for a Data Provider. If this option is chosen then the Data Provider MUST fully support the FHIR specification.

**Data Providers and Data Consumers SHOULD support the following contentTypes:**

- application/pdf
- text/html
- image/tiff
- image/jpeg

(This list may be added to by request and discussion)

Other contentTypes should only be used with caution, as not all Data Consumers may be able to view them. The contentType MUST be advised by the Data Provider in the DocumentReference resource, so that consumers can be aware and process accordingly.

**Documents MUST be classified according to the agreed regional code-list**

- This ensures consistency for consumers. The following SNOMED code list MUST be used for the document “type”: **(UK) EHR composition types – 24501000000105** (See [termbrowser.nhs.uk](http://termbrowser.nhs.uk) for details)

**Data Consumers MAY leave the HTTP “Accept” header blank, and SHOULD NOT populate it with anything other than the mime-type advertised by the Data Provider in the DocumentReference**

- There is no expectation of Data Providers being able to offer translation capabilities between mime-types

**Data Providers who offer FHIR Compositions SHOULD also publish a DocumentReference of content type “application/fhir+json” with a link to the composition**

- This is (4) above. It allows all Data Consumers to be aware of the existence of a composition and to provide a more complete list of all documents about a patient – including both unstructured and semi-structured “documents”

**Data Consumers MAY have the ability to retrieve and render FHIR Composition links**

- This is the choice of the data consumer. Evidently specific renderers would need to be created to display each type of composition to a user
- A data consumer MAY display the list of DocumentReferences but have no ability to retrieve and render any particular type of composition. However it MUST handle this possibility elegantly.



## Appendix 4 – Maturity Matrix

Section	Narrative	Consultative	Draft	Normative
<b>1 Introduction</b>	X			
1.1 Purpose of this Document				
1.2 Relationship with IHE XDS	X			
1.3 Medical Media and DICOM		X		
1.4 FHIR Documents	X			
1.4.1 Compatibility of Document Based FHIR provision and the YHCR				
1.4.2 The YHCR as Mediator for Document Based Data Sources	X			
1.5 Relationship of this Document with Other Standards	X			
1.6 Intended Users of the This Document	X			
<b>2 Relationship with Other Components of the System-of-Systems</b>			X	
2.1 Requirements for Regional Infrastructure				
2.2 Document Proxy Server			X	
2.2.1 Security				
2.2.2 Purging Document Locator Records			X	
2.2.3 Auditing			X	
2.3 Requirements for Data Providers			X	
<b>3. The YHCR DocumentReference Profile</b>		X		
<b>4. FHIR Documents and Compositions in the YHCR</b>	X			
4.1 Us of FHIR Compositions as a History of Care				
4.2 FHIR Documents and Transfer of Care		X		
4.3 FHIR Documents as an Interfacing Mechanism		X		
4.4 Decomposition of Vendor Supplied Bundles		X		
<b>5. YHCR Document and Composition Production</b>		X		
5.1 Composition Template				
5.2 Composition Generator		X		