INTERWEAVE

CONNECTING CARE

Cookbook for Regional Interoperability Detailed Design Paper #021

Onboarding for Data Consumers

PRELIMINARY DRAFT

Version 1.1 – 26th October 2021

Abstract Interoperability Cookbook Anchor Points

Section	Title
4	Requirements for Data Providers

Table of Contents

Version Control

Version	Release Date	Released By	Reason for Release
1.0	17/12/2019	R Hickingbotham	Preliminary draft
1.1	26/10/2021	T Davey	Updated with Portal environments

Reviewers

Initials	Name	Role	Organisation

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 021 - “Onboarding for Data Consumers”) defines the process for bringing a new data consumer onto the System-Of-Systems. It covers the steps which are followed from when a potential data consumer first approaches the YHCR up to the point when the consumer can make a connection to the live SoS environment and acquire data from data providers.

The paper is written from the perspective of the operators of the System-of-Systems. The technical, governance, or transformational activities undertaken by the data consumer are crucial to prepare an organisation for participation, but other design papers deal with these considerations such as 015 – “Governance for Data Consumers”. This paper focuses on the actions which will be taken by operators to:

- establish the identity of the data consumer and its responsible officers;
- enable and secure TCP/IP connections from the consumer;
- provide access to the consumer to a series of environments with increasing requirements for compliance with technical standards;
- test the consumer’s adherence with mandatory governance requirements;
- record the scope of consumer’s data requirements.

1.2 Relationship with this Paper and Design Paper 020 – Onboarding for Data Providers

This paper is a companion to design paper 020 – “Onboarding of for Data Providers”. Much of the onboarding process is the same for data consumers as for data providers and this paper avoids duplication by focussing on the differences. Design paper 020 also details the configuration of SoS environments and the interdependencies between change control of the software which makes up the SoS and the requirements of SoS participants as they progress through sandpit, staging and live environments. This discussion is equally relevant to data consumers and readers of this document are encouraged to also read design paper 020 as the material is not repeated here.

1.3 Onboarding and Risk Management

Onboarding is a key risk activity for the YHCR. There is the potential to compromise both security and data integrity. The SoS will ultimately be hosted in the open internet and its security is wholly dependent on the integrity of the public key infrastructure and consistency of the configuration of firewalls, DNS, and the participant registry. Risks are mitigated by automating the onboarding process and minimising the need for manual configuration of components.

Design paper 020 introduces a design for an Onboarding Suite: a software product which automates the workflow underpinning the onboarding process and manages configuration of the resources responsible for security. This paper elaborates on the requirements for this product from the perspective of onboarding data consumers.

1.4 The Participant Registry

The participant registry is an important source of configuration data for the SoS. It identifies the data providers and data consumers which have access to each of the SoS environments and contains all the details needed to manage interactions with the participants. This paper specifies the data elements which are managed for data consumers which is hence termed the consumer registry.

1.5 Relationship of this Document with Other Standards

This paper does not require knowledge of any particular standard.

1.6 Intended Users of the This Document

Operators of the YHCR, developers of the regional onboarding suite. Those seeking to onboard data consumers.

.

2 Environments

2.1 System of Systems Environments

As noted in “Onboarding for Data Providers”, the SoS comprises several environments which are used to separate live operations from software development, testing and participant onboarding.

Non-live environments can be considered as being arranged in two discrete sets: a set of environments which are dedicated to software development and testing, and a set of environments provided for the purpose of participant onboarding. Design paper 020 – “Onboarding for Data Providers” provides much detail on the structure and use of these environments which is not repeated here.

For the purpose of onboarding data consumers, it is sufficient to understand the role of three environments:

1. **Sandpit** – an environment which is connected to “sandpit” test systems at data providers and provides consuming software with access to work-in-progress test datasets. The sandpit is used by software still under development to test their functionalities and ensure compliance with regional standards.

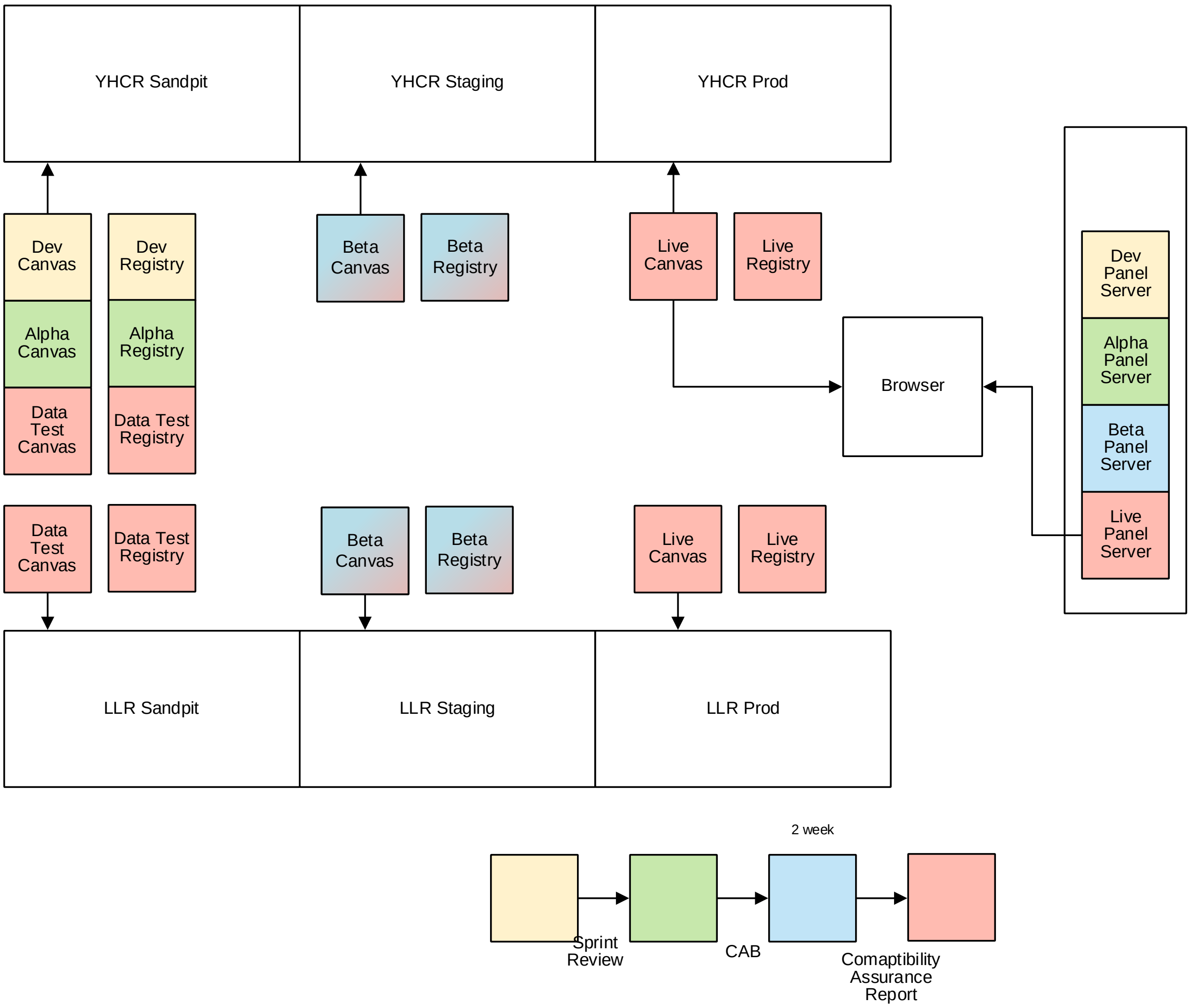
2. **Staging** – an environment which is connected to “staging” test systems at data providers and which is as representative as possible of the live environment. It is in staging that formal assurance of compliance with regional governance and clinical safety compliance occurs.

3. **Live** – which provides access to live data providers.

The function of the onboarding process is to migrate a product through these environments in sequence whilst ensuring that the correct assurance has been received at each stage.

2.2 Portal Environments

The Interweave Portal is an example of a Data Consumer - and has its own environments which are used both for testing of new Portal functionality and also as a tool for reviewing new data offered by Data Providers. The diagram below illustrates the configuration of these Data Consumer environments.



Key points are as follows:

- New Portal code progresses through a cycle of Dev (yellow), Alpha (green) and Beta (blue) before reaching Production (red)
  - Dev and Alpha are connected to the Sandpit data, and are used for initial testing in an environment with few restrictions. They are for initial developer and internal quality testing.
  - Beta is connected to Staging data, for final testing and User Acceptance Testing against more stable datasets

- The Portal's secondary purpose for viewing and testing Data Provider data is also covered:
  - The "Data Test" Portal environment uses the fully proven production Portal code for viewing Data Provider datasets in Sandpit
  - The "Beta" Portal environment is used for viewing and testing Data Provider datasets in Staging. (Depending on the point in the release cycle, this may be the production code or a beta release)

  - The “Production” Portal environment is used, obviously, for viewing Production datasets

- Other points shown by the diagram include:
  - The YHCR environment is, for historical reasons, the region used for initial development – and therefore the Dev and Alpha environments for internal testing exist only against the YHCR Sandpit
  - LLR is shown as an example of an additional region – with the environments needed for data testing and Portal UAT. This would be replicated for any other regions in a similar way.
  - The diagram also shows how any given environment is composed of three components: Canvas, Registry, and Panel Server. The Panel Server is essentially the “app store” of Panels, and the diagram shows how this is shared in common across multiple environments and regions.

3 The Onboarding Workflow

Data consumers follow the same onboarding workflow as for data providers and the process is controlled by the same software product, the Onboarding Suite, which is described in design paper 020. Requirements for this product are broadly the same as for data providers and these are not repeated here. The following constitute key areas of difference between providers and consumers which have consequences for the onboarding process:

	Data Providers	Data Consumers
Testing	Can be largely automated with the SoS providing a test harness, test scripts and visualisation tooling. The gate controlling progression from the Sandpit to Staging can be automatically controlled through measurement of successful test coverage.	Testing of consumer functionality can only be performed by the suppliers of consuming software. The SoS can support this by offering rich test data covering all maturity levels but it depends on supplier self-assurance to control the gate between Sandpit and Staging.
Compliance with Standards	Assurance can be mainly automated. The SoS validates data content and adherence to FHIR profiles. Compliance of endpoints and audit data is automatically tested during onboarding.	Assurance is manual and relies on statements of compliance by suppliers.
Security	Security focuses on trusted relationships enforced through a public key infrastructure. Inaccurate identity management risks a rogue provider compromising data quality.	Security focuses on trusted relationships enforced through a public key infrastructure. Inaccurate identity management risks a rogue consumer having unauthorised access to patient data. The reward for compromising security is greater with a correspondingly higher risk.
Controlled release of live data	Prior to generally releasing live data, a data provider must undertake a clinical safety assurance review. This needs to be controlled by the SoS.	Release of consumer software to clinical staff is under control of the consuming organisation. Control by the SoS team in allowing access to data is not needed.
Ongoing Compliance Testing	Key features of compliance can be tested automatically.	Forensic analysis tooling (design paper 016 – "Securing the YHCR") can be used to detect errant consumer behaviour but rigorous compliance assurance is a manual process.
Presence in the System Test Environment	Data Providers contribute to the System Test environment by providing servers which are populated with data which is used	Data Consumers have no presence in the System Test environment. Testing is performed using model consumer software. Compliance by

	for system testing new YHCR functionality	real consumer software with new YHCR features is proven in the Sandpit/Staging environments.

3.1 Anatomy of a Data Consumer

A data consumer connects to the SoS from a pre-defined static IP address and from which it issues claims against IAM and interacts with regional data through the FHIR aggregator.

A consumer also optionally implements one or two endpoints to which the System of Systems connects to deliver:

- subscription notifications;
- transactional messages.

3.2 Onboarding Steps

The following table largely replicates the steps for onboarding data providers. Cells coloured grey are identical for data providers and consumers and do not need to be considered separately.

Identify Participant	1	Register prospective participants to the YHCR in the Onboarding Suite.
	2	Obtain and record details of senior officers, contact methods and domain name. Corroborate details from an independent source and record method of corroboration. All email addresses must be at the registered domain name for the organisation.
	3	Identify senior officer responsible for participation in the YHCR (SRO).
	4	Receive notification of interest to participate in the YHCR.
		Record details of the product or service to be onboarded.
	5	Obtain and record contact details of technical personnel responsible for onboarding. Corroborate participation and technical contacts through an exchange of email with the SRO.
	6	Generate online accounts on the onboarding suite for technical officials.
	7	Generate a synonym for the consumer which will be used as a uniform identifier by the SoS.
	8	Provide access to Onboarding suite portal to technical officials. Capture password to be use in certificate signing requests.
Development and Sandpit	9	Capture IP addresses and port numbers of the endpoint of the participant's FHIR development server. Separate IP addresses may be provided for: • consumer client endpoint (IP address only); • reliable messaging delivery point; • subscription notification delivery point.

		Capture minimum maturity level of paired data providers.
	10	Obtain electronic signature verifying IP addresses from technical official.
	11	Corroborate veracity of IP addresses with SRO
	12	In the Sandpit environment: • generate DNS address records for the consumers IP addresses; • configure outbound firewall rules to permit access to server addresses/ports and an inbound rule to permit access to the consumer client; • register the data consumer in the Sandpit participant registry as an inactive participant.
	13	Prove IP connectivity between the Sandpit environment and data consumer endpoints.
	14	Accept server and client certificate signing requests.
	15	Confirm receipt of certificates with SRO.
	16	Apply certificates to Sandpit Participant Registry.
	17	N/a
	18	N/a
	19	Activate participant in Participant Registry.
	20	Determine pairings with data providers
	21	Participant confirms testing is complete.
	22	Participant completes governance checklist for data consumers.
	23	Confirm progression to Staging with SRO.
Staging	24	Capture IP addresses and port numbers of the endpoint of the participant’s FHIR development server. Separate IP addresses may be provided for: • consumer client endpoint (IP address only); • reliable messaging delivery point; • subscription notification delivery point. Note that the participant may elect to migrate the same servers from development to staging in which case these IP addresses may be the same as in step 8.
	25	Obtain electronic signature verifying IP addresses from technical officer.
	26	Corroborate veracity of IP addresses with SRO
	27	In the Staging environment: • generate DNS address records for the consumers IP addresses; • configure outbound firewall rules to permit access to server addresses/ports and an inbound rule to permit access to the consumer client; • register the data consumer in the Staging participant registry as an inactive participant.

	28	If the same servers are being used in Staging as in the Sandpit, then deactivate the participant in the Sandpit environment.
	29	Prove IP connectivity between the Sandpit environment and data provider endpoints.
	30	Accept server and client certificate signing requests.
	31	Confirm receipt of certificates with the SRO.
	32	Apply certificates to the Sandpit Participant Registry.
	33	N/a
	34	Activate participant in the Sandpit Participant Registry.
	35	Execute pipe cleaning tests aimed at proving functional connectivity.
	36	N/a
	37	Schedule volume testing.
	38	Schedule assurance testing.
	39	Execute volume tests.
	40	Execute assurance tests.
	41	Confirm progression to Live with the SRO.
System Test Representation	42	N/a
	43	N/a
	44	N/a
	45	N/a
	46	N/a
	47	N/a
	48	N/a
	49	N/a
	50	N/a
	51	N/a
	52	N/a
	53	N/a
	54	N/a
	55	N/a
	56	N/a
	57	N/a

Migrate to Live	58	Capture IP addresses and port numbers of the endpoint of the participant's FHIR development server. Separate IP addresses may be provided for: • consumer client endpoint (IP address only); • reliable messaging delivery point; • subscription notification delivery point. Note that the Sandpit servers cannot be repurposed for live operation and must remain accessible from the Sandpit environment.
	59	Obtain electronic signature verifying IP addresses from technical officer.
	60	Corroborate veracity of IP addresses with SRO
	61	In the Live environment: • generate DNS address records for the consumers IP addresses; • configure outbound firewall rules to permit access to server addresses/ports and an inbound rule to permit access to the consumer client; • register the data consumer in the Staging participant registry as an inactive participant.
	62	N/a
	63	Accept server and client certificate signing requests for Live.
	64	Confirm receipt of certificates with the SRO.
	65	Apply certificates to the Live Participant Registry.
	66	Activate participant in the Live Participant Registry.
	67	Execute pipe cleaning tests aimed at proving functional connectivity.
	68	Suspend all outgoing connection points to the data provider (subscriptions will be generated as historic patient contact is loaded and these should queue in the SoS).
	69	N/a
	70	N/a
	71	N/a
	72	N/a
	73	Schedule go-live date.
	74	Confirm data compliance review results and go-live date with SRO.
	75	N/a
	76	N/a
	77	Open connection points.

4 The Participant Registry for Data Consumers

The participant registry covers both data consumers and data providers. Requirements here relate only to data consumers; a parallel definition is provided by design paper 020 for data providers.

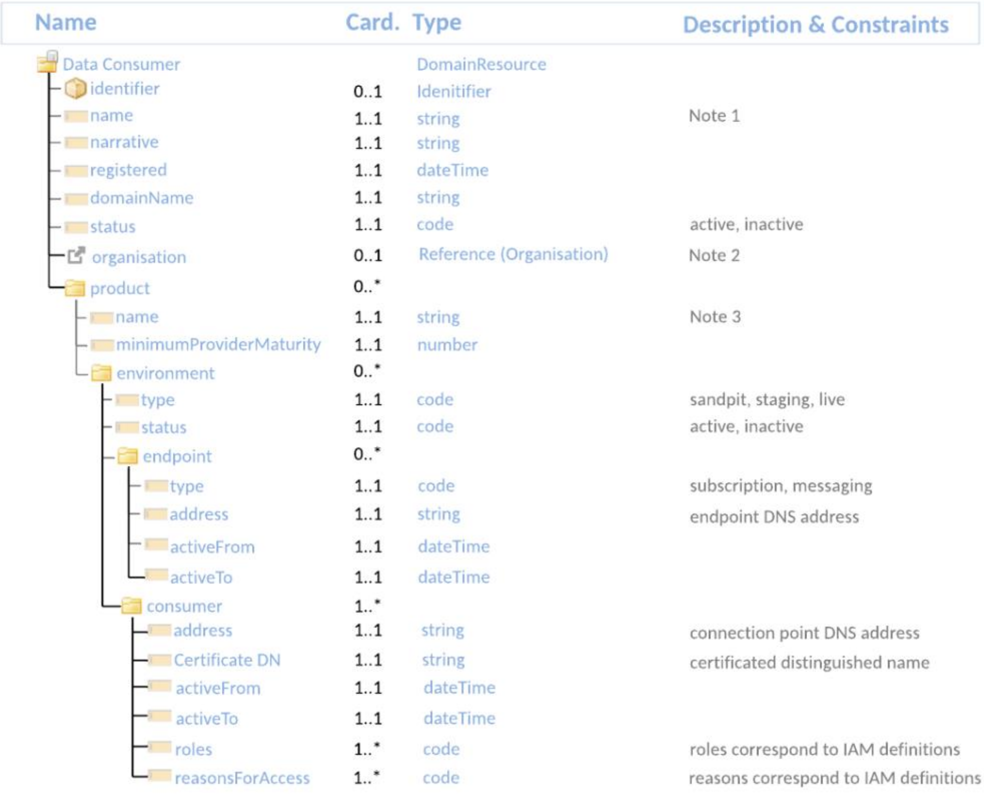
4.1 The *DataConsumer* Resource

The data model described below is used by components of the SoS to control the relationships between data providers and data consumers. It can also be accessed by consumers of the YHCR that interested in determining the registration details of other participants.

The data is not a FHIR resource but it uses the notation and inheritance of model of FHIR and therefore can be served from FHIR server which is configured with an appropriate schema.

The data structure represents a data consumer which may be associated with an Organisation which is also known to the YHCR as a health & care provider. Normally only one data consumer will be associated with an organisation. The data consumer can operate one or more products that each have a presence in one or more environments. In each environment the product must have a consumer (the client application which queries the YHCR) and may have endpoints for subscription notification and message delivery. The details of consumers, and endpoints may change over time and can be superseded.

The participant registry for data consumers is defined as follow:



Name	Card.	Type	Description & Constraints
Data Consumer		DomainResource	
identifier	0..1	Idenitifier	
name	1..1	string	Note 1
narrative	1..1	string	
registered	1..1	dateTime	
domainName	1..1	string	
status	1..1	code	active, inactive
organisation	0..1	Reference (Organisation)	Note 2
product	0..*		
name	1..1	string	Note 3
minimumProviderMaturity	1..1	number	
environment	0..*		
type	1..1	code	sandpit, staging, live
status	1..1	code	active, inactive
endpoint	0..*		
type	1..1	code	subscription, messaging
address	1..1	string	endpoint DNS address
activeFrom	1..1	dateTime	
activeTo	1..1	dateTime	
consumer	1..*		
address	1..1	string	connection point DNS address
Certificate DN	1..1	string	certificated distinguished name
activeFrom	1..1	dateTime	
activeTo	1..1	dateTime	
roles	1..*	code	roles correspond to IAM definitions
reasonsForAccess	1..*	code	reasons correspond to IAM definitions

Notes:

(1) The name of the data consumer is unique among *DataConsumers*. An ODS code is preferred but not mandated.

(2) The association to an *Organisation* is optional. Whilst most of the organisations operating a data consumer might also be expected to be providers of health and care services there may be exceptions. If it is provided then it must to be a reference to an Organisation resource held by the YHCR (reference design paper 013 – "Interfaces with the Organisational Data Service (SDS)". The onboarding suite may choose to use details of a FHIR resource to store organisational details but it is not mandated and the potential operational benefits of decoupling the definition of an organisation and its officers from the perspective of onboarding to that from the perspective of a health & care provider.

(3) The name of the product is unique within the *DataConsumer*. In combination with the consumer name, the product name uniquely identifies the consumer and is used by the consumer when making a claim against IAM as the *iss* attribute of the JWS and HTTP header clientId. These are in the format:

[consumer name]-[product name] ie: RFR-Sepia

Search parameters comprise:

Name	Type	Description	Expression
name	string	The name of the data consumer	DataConsumer.name
identifier	token	The unique id for a particular data consumer	DataConsumer.identifier
status	code	The status of a data consumer	DataConsumer.status
organisation	reference	The organisation that operates the data consumer	DataConsumer.organisation (Organisation)
productName	string	The name of a product operated by a consumer	DataConsumer.product.name
environment	code	The environment in which the product is operating	DataConsumer.product.environment.type
connectionType	code	The status of a product in an environment	DataConsumer.product.environment. endpoint.type | DataConsumer.product.environment. cosumer.type

4.2 APIs for Interacting with *DataConsumers*

DataConsumer resources will be persisted in a regional FHIR Store and will be accessible from a URL published in the Operations Guide by regional components and any consumer of the YHCR using standard FHIR STU3 search syntax.

Resources will be managed by the Onboarding Suite. Management APIs will not be available to any other component or participant.

Note that the content of a DataConsumer resource should be based on the environment from which it was obtained:

- references to *Organisations* must resolve to the organisation within the same environment as the *DataConsumer;*
- only one environment branch should be populated – that which corresponds to the environment from which the *DataConsumer* was obtained.

4.3 Use of the *DataConsumer* Resource by YHCR Components

All YHCR components manage data structures which involve data consumers by reference to a DataConsumer resource or to the combination of a product name and consumer name.

4.3.1 Identity and Access Manager

IAM validates that the JWS claim made by a client is associated with an active product for a registered data consumer. Specifically the Client ID of the HTTP header in which the claim is made is verified to be in the form [consumer name]-[product name] and that both the consumer and product are active in environment in which the claim is made.

IAM validates that Distinguished Name in the certificate used to sign the claim is the same as that which is active for the consumer part of the product.

IAM validates that the DNS entry registered which is active for the consumer part of the product resolves to the claimant's IP address.

4.3.2 FHIR Aggregator

If the product's minimum provider maturity level is specified, then the FHIR Aggregator restricts data interactions to the those data providers at or above the specified maturity level.

The FHIR Aggregator also restricts access to those data providers who have an inclusive pairing with the *DataConsumer* and those data providers which have an exclusive pairing but don't exclude relationships with this consumer (see design paper 020 for more detail).

4.3.3 Subscription Manager

The Subscription manager delivers subscription notifications to the active subscription endpoint address recorded for the product that created the subscription.

A subscription cannot be created by a product which does not have an active subscription endpoint.

4.3.4 Message Delivery

Messages sent over the reliable messaging channel (design paper 006) are delivered to the active messaging endpoint address recorded for the product that created the subscription.

A message will not be accepted for products which do not have an active messaging endpoint.

Note that the *MessageHeader* specifies an *Organisation* resource as a receiver. By implication to be deliverable the *Organisation* must be associated with exactly one active *DataConsumer* which has exactly one active product with an active messaging endpoint.

4.3.5 Consent Manager

DataConsumers may be referenced by *Policy* resources so as to restrict release of data to particular consumers. By implication all products operated by the consumer are covered by this policy

5 Compliance Assurance

For data providers a significant part of the assurance of compliance with standards can be automated and gates between environments opened by the onboarding suite with little manual involvement. Data consumers are different in that assurance processes are manual and the YHCR must corroborate that the correct assurance has taken place at each stage in the onboarding process.

Key determinants for the entry into an environment are as follows:

Sandpit

- Evidence that client software has been developed with an understanding of the architecture and technical standards of the YHCR.
- Governance responsibilities are understood.

Staging

- Evidence of testing in the sandpit environment using appropriate resource types at the required maturity levels.
- Evidence of testing using the required interaction patterns.
- Evidence of testing in situations that there are known data impairments.
- Successful execution, without errors being reported by any component of the YHCR, of a scripted integration test.

Live

- Signed-off clinical safety audit which asserts that YHCR data is presented or otherwise used in a clinically safe manner for each use case supported by the consumer. The audit should specifically cover the treatment of statements of data impairment and access restrictions.
- Signed-off governance report that asserts:
  - the identity management system employed ensures that that the identity of the user of the YHCR is known to the consuming organisation;
  - the user is assigned a role and the system enforces appropriate restrictions to data content for the role;
  - the user has a legitimate relationship with all patients for which data is accessed using the YHCR;
  - users' access to data is logged, that the log is associated with access-token provided by IAM, and the access token is searchable;
  - the claim made by the consumer in gaining an access token accurately states the user's regional role and reason for access.
- Signed-off security assurance from a YHCR security expert.

Appendix 1 – Maturity Matrix

Section	Narrative	Consultative	Draft	Normative
1 Introduction 1.1 Purpose of this Document	X			
1.2 Relationship with this paper and Design Paper 020	X			
1.3 Onboarding and Risk Management	X			
1.3 The Participant Registry	X			
1.4 Relationship of this Document with Other Standards	X			
1.5 Intended Users of the This Document	X			
2 Environments of the System of Systems			X	
3 The Onboarding Workflow 3.1 Anatomy of a Data Consumer			X	
3.2 Onboarding Steps			X	
4. The Participant Registry for Data Consumers 4.1 The *DataConsumer* Resource			X	
4.2 APIs for Interacting with *DataConsumers*		X		
4.3 Use of the *DataConsumer* Resource by YHCR Components 4.3.1 Identity and Access Manager			X	
4.3.2 FHIR Aggregator			X	
4.3.3 Subscription Manager			X	
4.3.4 Message Delivery			X	
4.3.5 Consent Manager			X	
5. Compliance Assurance		X		