Cookbook for Regional Interoperability Detailed Design Paper #009

Auditing

PRELIMINARY DRAFT

Version 1.3 – 8th December 2022

Abstract Interoperability Cookbook Anchor Points

Section	Title
3.1.4	Audit Service
4.6	Governance for Data Providers

Table of Contents

1 Introduction 4
1.1 Purpose of this Document 4
1.2 Topics of Interest 4
1.3 FHIR and Auditing 5
1.4 Relationship of this Document with Other Standards 5
1.5 Intended Users of the This Document 5
2 General Requirements for FHIR Endpoints 6
2.1 Persistence of *AuditEvent* Resources 6
2.2 Considerations for the use of AuditEvents 6
2.3 Event Identifiers 7
2.4 Outcome Codes 10
2.5 The YHCR *AuditEvent* Profile 11
2.6 Agent Contents 14
2.6.1 Data Consumer 14
2.6.2 IAM 16
2.6.3 Regional Aggregator 16
2.6.4 Data Provider 17
2.7 Entity Contents 18
2.8 NHS Number Entity 19
2.9 Additional Auditing Requirements for IAM 19
2.10 Retention Period for Audit Records 20
3 Minimal Auditing Requirements for Data Consumers and Data Providers 21
3.1 Key Principles 21
3.2 Requirements for Data Consumers 21
3.3 Requirements for Data Providers 22
4 Additional Requirements for Regional Components 23
5 Security of Audit Records 24
Appendix 1 – Approach to Audit Investigations 25
Appendix 2 – Maturity Matrix 26

Version Control

Version	Release Date	Released By	Reason for Release
1.0	28/04/2019	R Hickingbotham	Preliminary draft
1.1	05/07/2021	T Davey	Clarification of minimum requirements for Data Consumers
1.2	18/08/2021	T Davey	Clarification of audit resource usage and investigations approach
1.3	08/12/2022	T Davey	Addition of audit codes for merge and deletion

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 Interweave Exchange. They are intended as a basis for developing or procuring software and so are expressed at a level of precision which is intended to avoid ambiguity but with a consequence that 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 009 - "Auditing") focuses on requirements for auditing use of the YHCR. Auditing is a collective responsibility for data consumers, data providers and the YHCR regional components. The fundamental requirements for auditing are to:

- support a definitive statement of who has accessed what patient data;
- track the provenance of data created or updated through the YHCR;
- highlight potential security threats;
- enable analysis of user behaviour and identification of potentially inappropriate usage.

1.2 Topics of Interest

At a basic level, auditing is the recording of an event related to the person that caused the event with identification of the data involved in the event. However, the federated nature of the YHCR and the potentially vast transaction rates (current estimates of peak usage are in the order of 1,500 transactions per second) give rise to some very specific questions:

How can the individual interacting with data be identified? The components in the YHCR which service access to data are remote from the user interfaces which present data to end users. The user identity, as known to the data consumer, must be reconcilable to the user identity presented to data providers.

How is a consistent approach to auditing achieved? The YHCR is composed of many distributed components operated by a number of autonomous organisations. The audit record must be a uniform explanation of who has accessed or updated what data. This requires collaboration on what events are audited, what data is recorded for each event, and how audit records are aggregated across components.

How is access to audit records controlled? Audit records potentially contain revealing information about citizens and insecurities in the auditing system can invalidate the security benefits provided by auditing. Consideration must be given to the data content of audit records and to the segregation of duties around data access.

What use can be made of audit records in a massive transaction throughput environment? If the YHCR is to operate and anywhere near the projected peak transaction rates, then manual review of audit records will have no value. Tooling will be needed to extract information which is targeted to use cases.

How can the integrity of audit records be ensured? Alongside other mitigations, audit records provide assurance that records have not been accessed inappropriately. This control is compromised if those with access to infrastructure where audit records are held are able to tamper with audit records.

1.3 FHIR and Auditing

FHIR defines an *AuditEvent* resource and, given the level of dependency the YHCR has on FHIR, it is a natural candidate for use here. The YHCR FHIR infrastructure also focusses on aggregating FHIR resources from federated sources and so adopting it for auditing goes some way towards addressing the requirement for assembling records from distributed components. Both the regional components and data providers, through their implementation of a FHIR proxy (design paper 003 "FHIR Proxy Server") have FHIR stores in which *AuditEvent* resources can be recorded and this paper proposes that this mechanism is used by these parties.

Data consumers also have the option of using a FHIR Store and the *AuditEvent* resource to meet their audit responsibilities. However, many of the early adopters of the YHCR are likely to have established technology for viewing medical records with their own established mechanisms for logging. Rather than create an unnecessary barrier to entry, this paper accepts that all data consumers will not standardize on the use of FHIR, and focusses on establishing a capability that is required from data consumers rather than stipulating the technical mechanism through which it is supplied.

As ever, FHIR resources need constraining so ensure that they are consistently interpretable and much of this document is dedicated to specifying a profile for the *AuditEvent*.

1.4 Relationship of this Document with Other Standards

The following standards form the basis for this document:

- FHIR Release 3 (STU) – Messaging Using FHIR Resources;

1.5 Intended Users of the This Document

This document is a reference guide for data providers implementing a FHIR Proxy Server with consent management features, developers of regional components, and developers of user interfaces which provide end user access to the YHCR.

2 General Requirements for FHIR Endpoints

FHIR endpoints are software which permits interaction with FHIR resources. In the YHCR FHIR endpoints include:

- the boundary of a data provider, for which, often, the endpoint is implemented using a FHIR proxy server but may be offered natively by an Electronic Patient Record (EPR);
- the regional Identity Management Service;
- the regional FHIR Aggregator;
- the regional FHIR Store.

This section establishes a uniform requirement for all of these components.

2.1 Persistence of *AuditEvent* Resources

AuditEvent resources will be persisted in a FHIR Store. The regional FHIR Store (design paper 018) is available for this purpose, but, for performance reasons, it is strongly recommended that data providers establish a local capability for this purpose. The FHIR Store usually write *AuditEvents* for resources created, modified, or read from it. The exception is the *AuditEvent* resource itself for which management will be unaudited (subject to the considerations of section 5).

Locally persisted *AuditEvents* must be readable and searchable from the FHIR endpoint which caused the audit record to be written. Regionally recorded records are readable and searchable from the regional FHIR aggregator.

It must not be possible to modify or delete an *AuditEvent* resource from a FHIR endpoint.

It must not be possible for database administrators (DBA) to manipulate the data that represents an *AuditEvent* in a FHIR Store. Note that this will mean that access privileges to update and delete data must be restricted to a database role which is not assigned to database administrators. Use of this role must be logged by the database middleware to a directory on a file system to which access is restricted to the root account and the database software.

Depending on technology used the it may not be possible for a FHIR Store to operate different resource types. In this eventuality a separate, dedicated, FHIR store, which is inaccessible to DBAs, must be used for *AuditEvent* resources.

2.2 Considerations for the use of AuditEvents

An *AuditEvent* resources records:

- the actor(s) involved in the event, including system actors;
- the source or participant recording the audit event;
- the resource(s) involved in the event;
- meta data about the event including the event type, date/time, reason for access, and outcome.

<u>Observations about the *AuditEvent*</u>

1. Relationships or causality between audit events cannot be expressed using the FHIR model. As a possible requirement, the regional FHIR aggregator may wish to record events for a FHIR query received from i) a data consumer and ii) a query executed against a data

provider. ii) may be a consequence of i) and the relationship should be recorded in the *AuditEvent.* An extension property is proposed for this purpose.

2. Policies are tied to an event. The policy that enabled the event to occur (or prevented it from happening) is recorded at the actor level, being interpreted as, say: Policy XYZ prevented/enabled data being returned to ABC. In the YHCR, policies are applied more granularly. A single search may return some data enabled by policy ABC and withhold some data because of policy DEF. This characteristic means that different *AuditEvents* must be recorded for a search request and for the data which is withheld or released from the search results.

3. The *securityLabel* property of resources which are referenced by the *AuditRecord* is not searchable. This is a logical candidate for describing whether a resource was withheld, released, or released subject to restrictions, but is unusable because of the inability to search it. A query cannot be expressed which identifies, say, resources not seen by clinicians because of policy restrictions. This type of query thought to be a fundamental requirement and being able to meet it is a guiding factor in the design of audit event types.

4. Being able to search audit records by patient is essential. Therefore an additional entity which identifies each NHS Number involved is added to each event to aid searching. Whilst this aids with searching, additional detail is also needed – as the patient is not necessarily the subject to all resources referenced by the event. A search for, say, prescriptions for aspirin, will yield results for many patients. FHIR provides for 'details' of referenced resources and this design proposed that the YHCR resource profile uses this concept to identify the patient's NHS Number for all patient identifiable resources.

In summary, the STU3 resource definition requires interpretation if it is to be used for the YHCR. This design proposes a YHCR resource profile which should lead to consistent adoption among the YHCR participants.

2.3 Event Identifiers

An *AuditEvent* resource is recorded for an event type and an event sub-type. Together these imply the scope of auditing within the YHCR.

FHIR endpoints must record the following events. Event codes are defined in the value set: http://yhcr.nhs.net/fhir/valueset-audit-event-type. The FHIR value set http://hl7.org/fhir/valueset-audit-event-type.html has been discarded as being too tied to DICOM to be useful.

Event	**Responsibility**	**FHIR Event Code**
Authentication Request	Data Consumer	YHCR001
Use of *AuditEvents* by data consumers are optional, but if used, then this event represents an attempt by an end user to authenticate with the data consumer and the data consumer's attempt to gain a bearer-token from the regional IAM service		
Authorization Request	IAM	YHCR002
The event signifies that an authorization claim has been made to IAM. The AuditEvent identifies the user making the claim and, for successful requests, the access token issued in response.		
FHIR Operation	All	YHCR003

A generic event which is recorded for all FHIR operations (inbound or outbound). The AuditEvent identifies the operation type, the resources being operated on, and the query string used for queries and bulk operations.		
Content Released	Regional FHIR Aggregator and Data Providers implementing consent management.	YHCR004
Recorded by endpoints releasing query results. This event is related to a FHIR operation search event. Multiple content released events may be recorded for a search, each describing the policy that allowed the contents to be released. Policies, currently, will not be required for data to be released for the purpose of direct care and search conducted on this basis will be associated with a single "content released" event. The *AuditEvent* references the resources (and their version) released.		
Content Withheld	Regional FHIR Aggregator and Data Providers implementing consent management.	YHCR005
Recorded by endpoints withholding query results because data access management policies. The event is related to a FHIR operation search event and is recorded for all resources withheld by the endpoint because of a particular policy. The *AuditEvent* references resources (and their version) which were withheld and the policy which caused the action. If more than one policy results in resources being withheld, then multiple "Content Withheld" events are written.		
Restricted Content Released	Regional FHIR Aggregator and those Data Providers implementing consent management.	YHCR006
Recorded by endpoints releasing query results which implement data access management policies. The event is related to a FHIR operation search event and is recorded for all resources released with the status of restricted content by the endpoint. The *AuditEvent* references resources (and their version) which were released and the policy which caused the action. As noted above, the actions multiple policies may necessitate multiple events to be written for a single search.		
Asynchronous Event	Data Providers	YHCR007
A set of events which relate to the processing of asynchronous query requests. Events are related to the original "RESTful Operation" event and track execution of the query and delivery of result-set parts to the invoker. The event uses related "Content Released", "Content Withheld" and "Restricted Content Released" events to record resources subject to the query processing.		
Subscription Result	All	YHCR008
A set of events which record key stages in the execution of the subscriptions process. For each stage the *AuditEvent* either by embedding resource references within the *AuditEvent* or by using associated 'Content Released', 'Content Withheld' or 'Restricted Content Released' events.		

Event sub-type codes are defined in the value set: http://yhcr.nhs.net/fhir/valueset-audit-event-sub-type as follows:

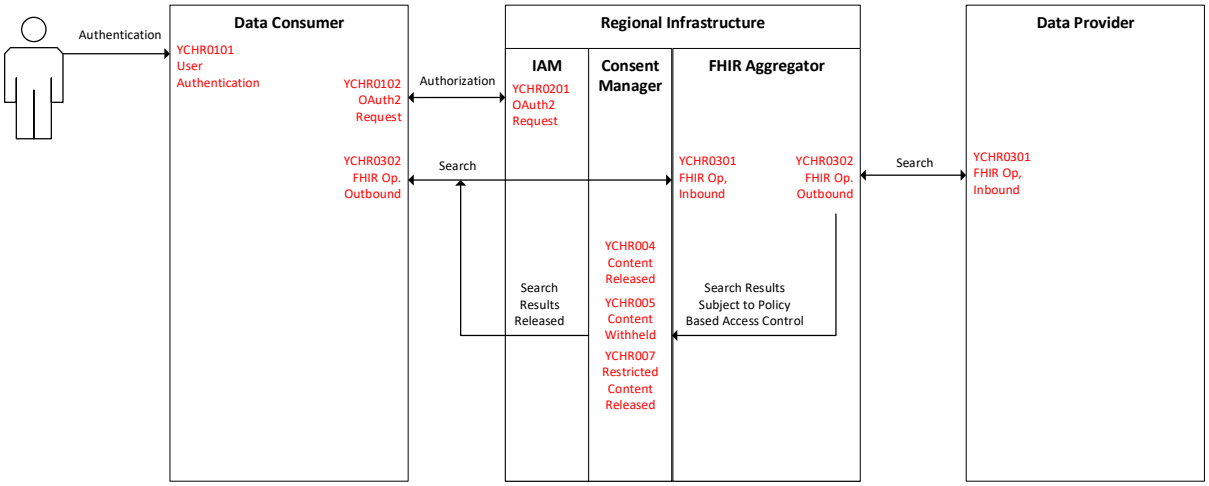
Event	Sub-type	FHIR Sub-type Code

Authentication	User Authentication	YHCR0101
	OAuth2 Request	YHCR0102
Authorization Request	OAuth2 Request	YHCR0201
FHIR Operation	Inbound	YHCR0301
	Outbound	YHCR0302
	Merge	YHCR0303
	Delete (soft delete)	YHCR0304
	Erase (hard delete)	YHCR0305
Content Released		
Content Withheld		
Restricted Content Released		
Asynchronous Event	Query Queued	YHCR0701
	Query Part Received	YHCR0702
	Query Part Released	YHCR0703
	Query Part Purged	YHCR0704
Subscription Result	Subscription Executed	YHCR0801
	Subscription Result Delivered	YHCR0802
	Subscription Result Received	YHCR0803

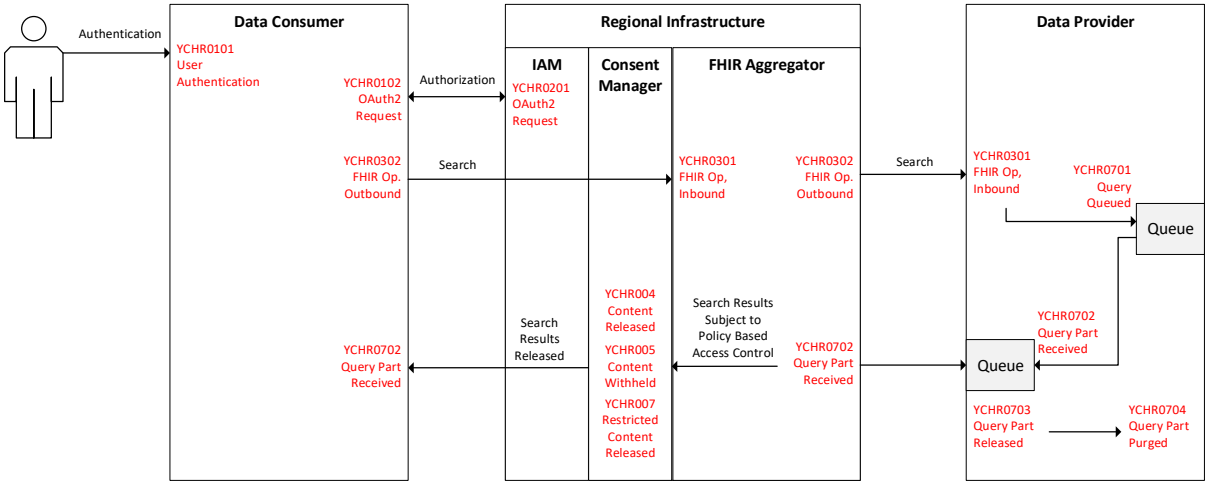
The *AuditEvent* resource carries details about the outcome of the event and so separate events are not required for negative occurrences such as Authorization Request Rejected.

The following diagrams illustrate how events are recorded by all actors in a transaction chain

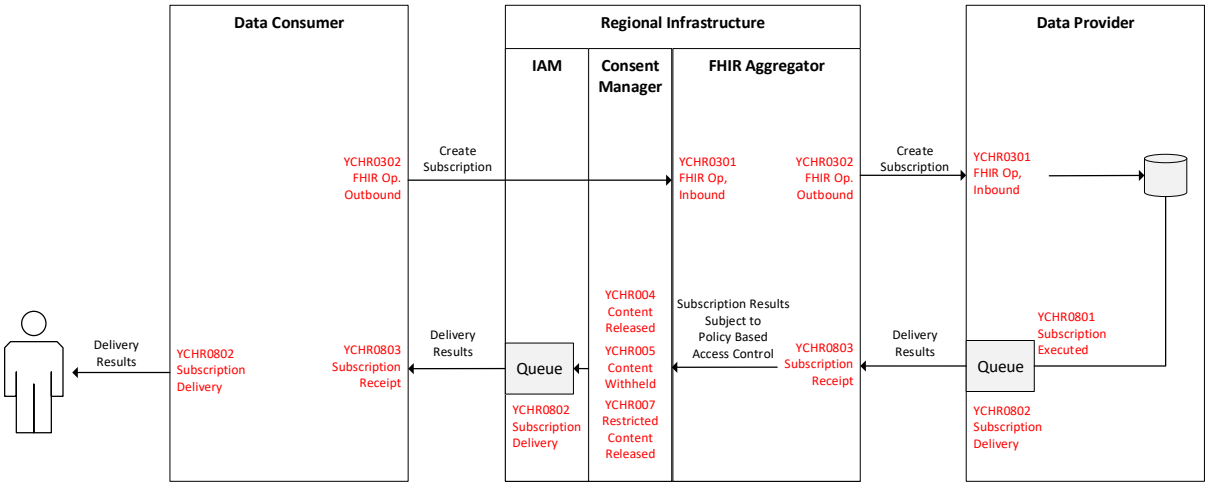
Events Recorded for Authentication, Authorization and Search



Events Recorded for Authentication, Authorization and Asynchronous Search



Events Recorded for Subscription Creation, Execution and Results Delivery



2.4 Outcome Codes

The FHIR STU3 standard defines event outcomes as follows:

Code	Display	Definition
0	Success	The operation completed successfully (whether with warnings or not).
4	Minor failure	The action was not successful due to some kind of catered for error (often equivalent to an HTTP 400 response).
8	Serious failure	The action was not successful due to some kind of unexpected error (often equivalent to an HTTP 500 response).
12	Major failure	An error of such magnitude occurred that the system is no longer available for use (i.e. the system died).

The YHCR coding system http://yhcr.nhs.net/fhir/valueset-audit-event-outcome adds.

Code	Display	Definition
99	Denied	The action was not undertaken for security reasons. For example, an authentication request could not be completed, au authorisation claim failed or JWT has expired.

2.5 The YHCR *AuditEvent* Profile

The YHCR *AuditEvent* is a constrained version of the STU3 standard resource. The profile adjusts the STU3 *AuditEvent* resource definition as described in the following table. A number of the fields are populated with data from the JWT bearer-token which enables access to services. See design paper 005 – "Identity and Access Management" for details.

Element	Resource Cardinality	Profile Cardinality	Search Index	Implementation Notes
type	1..1	1..1	type	Uses the YHCR coding system: http://yhcr.nhs.net/fhir/valueset-audit-event-type
subtype	0..1	0..1	subtype	Uses the YHCR coding system: http://yhcr.nhs.net/fhir/valueset-audit-event-sub-type
action	0..1	1..1	action	Type of action performed during the event. The FHIR AuditEventAction must be used. For the FHIR operation event (YHCR003) then an appropriate create, update or read actions is used. For authentication (YHCR001) and authorization (YHCR002) events then the execute action is used. For all other events the read action is used.
recorded	1..1	1..1	date	The time when the event occurred.
outcome	0..1	1..1	outcome	The YHCR profile makes an outcome mandatory and uses the http://yhcr.nhs.net/fhir/valueset-audit-event-outcome value set.
outcomeDesc	0..1	0..1		An appropriate reason must be recorded for all non-success outcomes.
purposeOfEvent	0..*	1..1		The 'Reason for Access' code from the JWT. The reason codes are maintained as a FHIR value set at: http://yhcr.nhs.net/fhir/valueset-audit-event-purpose-of-use
extension (related)	n/a	0..1		A reference to an *AuditEvent* being the event which caused this event to occur. This property is mandatory for event types YHCR004, YHCR005, YHCR006, YHCR007 where it must be a reference to an inbound FHIR operation event (YHCR0301). It is also used to relate an outbound FHIR operation (YHCR0302) to an inbound operation (YHCR0301).
agent	1..*	1..3		The FHIR standard allows multiple actors (agents) to be involved in an event. In most cases two actors will be recorded, to capture the "from" and "to" This will be always include: a) The "self" system[1] And then one of either: b) The human end user involved - for events *YHCR0101 User Authentication, YHCR0102*

[1] This "self" system agent is strictly speaking redundant, but provides a complete picture of "from" and "to" and may be useful to disambiguate if audit records from several different systems are extracted and combined together for analysis. It also provides a place to record the System Id. (In addition to the ODS code of the owning organisation which is recorded in the "Source"). For internal events it also allows the JTI to be captured.

				OAuth2 Request and *YHCR0803 Subscription Receipt*)[2] c) The "other" system involved in the interaction - for all communication-related events Whilst there will generally be two actors, there are exceptions as follows: • Some events are internal and therefore only involve one "self" actor (*YHCR0701, YHCR0702, YHCR0704, YHCR0801*) • *Event YHCR0102 OAuth2 Request* require three actors. It is executed by the Data Consumer, communicates with IAM, but must also link back to the session id established for the human user *See "Agent Contents" below for full details*
agent.role	0..*	1..1	agent-role	The role of the agent must be identified. • For human agents the role code from the JWT will be mapped to one of the standard FHIR role codes in SecurityRoleType (see section 2.6 for details) • For system agents extension role codes will be defined in https://yhcr.nhs.uk/Coding/audit-agent-role as follows: ○ data-consumer ○ data-provider ○ aggregator ○ iam
agent.reference	0..1	0..0	n/a	Agents are not referenced as FHIR Resources. Instead the JTI (see altId below) provides a link back to the JWT which contains further details
agent.userId	0..1	1..1	user	For human agents this is the user id. For system agents this will be either: • For Data Providers / Consumers - the Participant Id from the participant registry. (https://yhcr.nhs.uk/Id/participant-id) • For central infrastructure: values which represent the identity of IAM or the Aggregator (eg FQDN) *See "Agent Contents" below for full details*
agent.altId	0..1	1..1	altId	Must be recorded for all actors and be the jti (unique identifier) of the JWT. The jti binds all requests made under a single authorization together and acts here as cross system session identifier. For authorization events this must be the jti of the bearer-token issued by IAM.
agent.name	0..1	1..1	n/a	For human agents this is the user's name. For system agents this will be a system Display Name from the participant registry. *See "Agent Contents" below for full details*

[2] It is also intentional to not record details of the originating human user on audit events other than those directly involving the user and occurring within the Data Consumer system. This is to improve the anonymity of audit records. Instead the jti of any audit events requiring further investigation must be presented to the central team for reidentification of the user

agent.requestor	1..1	1..1		True if this agent is initiating the interaction. *See "Agent Contents" below for full details*
agent.location	0..1	0..0		Not needed for the YHCR. (Equivalent details may be retrieved either via the JWT for the user, or the Data Provider / Consumer registry for systems)
agent.policy	0..*	0..*	policy	Recorded where relevant for events FHIR004, FHIR005 and FHIR006 (Content Released, Content Withheld and Restricted Content Released). The property is a Reference to a Policy resource (design paper 008 – "Data Access and Consent Management".
agent.media	0..1	0..0		Not needed for the YHCR.
agent.network	0..1	0..1		This essential for events involving communication with another system, but not relevant for internal events.
agent.network.address	0..1	1..1	address	IP Address of other system. *See "Agent Contents" below for full details*
agent.network.type	0..1	1..1		2: IP Address
agent.purposeOfUse	0..1	0..0		Not used. See instead "purposeOfEvent" above
source	1..1	1..1		The organisation writing the audit event.
source.site	0..1	0..0	n/a	The identifier is sufficient for uniqueness.
source.identifier	1..1	1..1	source	The ODS code of the organisation writing the audit event.
source.type	0..1	0..0		Not needed for the YHCR.
entity	0..*	0..*		Resource references (see section 2.6 for requirements).
entity.identifier	0..1	0..1	entity-id	Identity references must be used not names or identifiers. *However the NHS Number entity is an exception – see section 2.8 below*
entity.reference	0..1	0..1	entity	Identity references must be used not names or identifiers. *However the NHS Number entity is an exception – see section 2.8 below*
entity.type	0..1	1..1	entity-type	Whilst references imply the type, explicitly recording the type facilitates searching and forensic pattern analysis. *An extension to the standard code list of "nhs-no" will be defined – see section 2.8 below*
entity.role	0..1	0..0	n/a	Derivable from the entity type.
entity.lifecycle	0..1	0..0		Information not available from the context of requests.
entity.securityLabel	0..*	0..0		Not searchable and of no additional value above the event type.
entity.name	0..1	0..0	n/a	Identity references must be used not names or identifiers.
entity.description	0..1	0..0		As little identifiable data should be recorded in the audit record to improve anonymity.
entity.query	0..1	0..1		The FHIR search string for searches and conditional operations. (NB: base64 encoded)

entity.detail	0..*	0..*		All patient identifiable resources (see design paper 005 – "Identity and Access Management" for a list) must identify thee NHS number of the resource's subject. In addition used to record Operation Outcome details only where relevant (see below)
entity.detail.type	1..1	1..1		"NHS"
entity.detail.value	1..1	1..1		NHS number of the resources subject. (NB: base64 encoded)
entity.detail.type	1..1	1..1		"OPERATIONOUTCOME" (only where relevant)
entity.detail.value	1..1	1..1		Contents of Operation Outcome (NB: base64 encoded) (only where relevant) This is for non-fatal Operation Outcomes ("Data Impairments") which may be returned as part of the results bundle. These may contain additional warnings that are important from a clinical safety perspective. They are transient and so unlike other FHIR Resources cannot be reconstructed from a reference. Therefore the details must be stored here.

Search indexes are shown as-per the FHIR specification for the Audit resource. Where a field is searchable in FHIR but not used by YHCR then the index is marked as "n/a".

2.6 Agent Contents

The contents of the Agent fields will vary depending on the type and subtype of audit event. In further elaboration of the Profile above, some fields will vary and will be populated as follows:

2.6.1 Data Consumer

Self Actor

This is recorded for all data consumer events

agent.role	data-consumer
agent.userId	The Participant Id of the Data Consumer (As-per the Participant registry, retrieved from configuration)
agent.altId	For YHCR0101 = "Unknown" For all other events = the JTI from the JWT.
agent.name	The Data Consumer display name (As-per the Participant registry, retrieved from configuration)
agent.requestor	The inverse of the value for the other actor involved in the event
agent.network.address	n/a

In addition the following event specific actors are recorded:

Interactions With Human User

- User Authentication: YHCR0101
- Subscription Receipt: YHCR0803

agent.role	The user's role, mapped from the IAM role code used in the JWT as follows:

Consumer Role	FHIR Role
Clinical Professional	AUTM
Social Care Professional	AUTM

	Citizen: PAT Authorised Carer: AUCG System or Robot: AULR
agent.userId	The human user id as known to the Data Consumer system, and as per the "subject" of the JWT
agent.altId	The internal session id established as the user logs on (NB: This is not the JTI as it may not be known at this stage)
agent.name	The human user's name
agent.requestor	YHCR0101 = True YHCR0803 = False
agent.network.address	The IP address of the user's computer

Interactions With IAM

- OAuth2 Request: YHCR0102

agent.role	The user's role, mapped from the IAM role code used in the JWT as-per the table above
agent.userId	The human user id as known to the Data Consumer system, and as per the the JWT "subject"
agent.altId	The internal session id established as the user logs on
agent.name	The human user's name
agent.requestor	False
agent.network.address	n/a

agent.role	iam
agent.userId	IAM system id, as used in the "issuer" field of the JWT
agent.altId	The JTI from the JWT.
agent.name	"IAM"
agent.requestor	False
agent.network.address	IP address of IAM endpoint

NB: This event is unique in having three Agents. It provides the link between the "internal" session id (established when the user first logs on), and the "external" JTI session id with the regional infrastructure (established when authorisation with IAM occurs)

Interactions With Regional Aggregator

- FHIR Op Outbound: YHCR0302
- Query Part Received: YHCR0702
- Subscription Result Received: YHCR0803

agent.role	aggregator
agent.userId	Aggregator system id ie FQDN of the environment, eg "fhir.sandpit.yhcr.nhs.uk" (from configuration)
agent.altId	The JTI from the JWT.
agent.name	"Aggregator"
agent.requestor	YHCR0302 = True YHCR0702 and YHCR0803 = False
agent.network.address	IP address of Aggregator endpoint

NB: This assumes that all communications go via the Aggregator, as is currently the case. In theory a Data Consumer might attempt to contact a Data Provider directly – in which case it would need to instead populate appropriate "Data Provider" details into this event.

2.6.2 IAM

Self Actor

This is recorded for all IAM events

agent.role	iam
agent.userId	IAM system id, as used in the "issuer" field of the JWT
agent.altId	The JTI from the JWT.
agent.name	"IAM"
agent.requestor	The inverse of the value for the other actor involved in the event
agent.network.address	n/a

In addition the following event specific actors are recorded:

Interactions With Data Consumer

- OAuth2 Request: YHCR0201

agent.role	data-consumer
agent.userId	The Participant Id of the Data Consumer, as per the JWT "issuer"
agent.altId	The JTI from the JWT.
agent.name	The Data Consumer display name, as looked up from the System Id in the Data Consumer registry
agent.requestor	True
agent.network.address	Incoming IP address of Data Consumer

2.6.3 Regional Aggregator

Self Actor

This is recorded for all Aggregator events

agent.role	aggregator
agent.userId	Aggregator system id ie FQDN of the environment, eg "fhir.sandpit.yhcr.nhs.uk" (from configuration)
agent.altId	The JTI from the JWT.
agent.name	"Aggregator"
agent.requestor	The inverse of the value for the other actor involved in the event
agent.network.address	n/a

Interactions With Data Consumer

- FHIR Operation Inbound: YHCR0301
- Subscription Result Delivered: YHCR0802
- Content Released: YHCR004
- Content Withheld: YHCR005
- Restricted Content Released: YHCR 006

agent.role	data-consumer
agent.userId	The Participant Id of the Data Consumer, as per the JWT “issuer”
agent.altId	The JTI from the JWT.
agent.name	The Data Consumer display name, as looked up from the Participant Id in the Participant registry
agent.requestor	For YHCR0301 = True For YHCR0802, 004, 005, 006 = False
agent.network.address	IP address of Data Consumer For YHCR0301 = Incoming IP address For YHCR0802, 004, 005, 006 = intended delivery IP address

Interactions With Provider

- FHIR Op Outbound: YHCR0302
- Query Part Received: YHCR0702
- Subscription Result Received: YHCR0803

agent.role	data-provider
agent.userId	The Participant Id of the Data Provider
agent.altId	The JTI from the JWT.
agent.name	The Data Provider display name, as looked up from the Participant Id in the Participant registry
agent.requestor	For YHCR0302 = False For YHCR0702, YHCR0803 = True
agent.network.address	IP address of Data Provider For YHCR0301 = Intended delivery IP address For YHCR0802, 004, 005, 006 = Incoming IP address

2.6.4 Data Provider

Self Actor

This is recorded for all data provider events

agent.role	data-provider
agent.userId	The Participant Id of the Data Provider (from configuration)
agent.altId	The JTI from the JWT.
agent.name	The Data Provider display name (from configuration)
agent.requestor	The inverse of the value for the other actor involved in the event
agent.network.address	n/a

In addition the following event specific actors are recorded:

Interactions With Regional Aggregator

- FHIR Operation Inbound: YHCR0301
- Query Part Released: YHCR0703
- Subscription Result Delivered: YHCR0802

agent.role	aggregator
agent.userId	Aggregator system id ie FQDN of the environment, eg “fhir.sandpit.yhcr.nhs.uk” (from configuration)
agent.altId	The JTI from the JWT.
agent.name	“Aggregator”

agent.requestor	For YHCR0302 = True For YHCR0703 and YHCR0802 = False
agent.network.address	For YHCR0301 = Incoming IP address For YHCR0802, 004, 005, 006 = Outgoing IP address of Aggregator endpoint

NB1: This assumes that all communications go via the Aggregator, as is currently the case. In theory a Data Provider might allow Data Consumers to contact it directly – in which case it would need to instead populate appropriate "Data Consumer" details into this event. This may be non-trivial as this information will be available only at the TLS termination point. This may very likely be on an separate proxy server, which would thus need to extract details and pass through (eg as custom HTTP headers).

NB2: By a similar argument, the incoming IP address to the Data Provider may in fact be logged as that of a proxy server. The proxy server logs would need to be consulted for further details of actual external connections. This is not seen to be a serious issue where there is a single Aggregator coming from a well-known endpoint.

Internal

- Query Queued: YHCR0701
- Query Part Received: YHCR0702
- Query Part Purged: YHCR0704
- Subscription Executed: YHCR0801

No additional agent, other than the "self" agent, is relevant

2.7 Entity Contents

The content of the entity section of the *AuditEvent* depends on the event type being logged. Requirements are summarised in the following table.

Event	Action	Entity Contents
Authentication	E	None
Authorization Request	E	None
RESTful Operation	C	A reference to the created resource if successful, otherwise none.
	R	References to resources successfully read or searched.
	U	References to resources successfully updated.
	D	References to resources successfully deleted
Content Released	R	References to resources successfully read or searched and released after policy-based access controls have been applied.
Content Withheld	R	References to resources successfully read or searched and withheld after policy-based access controls have been applied.

Restricted Content Released	R	References to resources successfully read or searched and released with restricted status after policy-based access controls have been applied.
Asynchronous Event	R	References to resources successfully searched.
Subscription Result	R	Reference to a resource prepared for dispatch or dispatched as a result of the subscription.

2.8 NHS Number Entity

Being able to search audit events by patient (ie NHS Number) is an important capability that is, however, not obviously supported by the FHIR STU3 audit event. Two mitigations are already in place:

1) Capturing the NHS Number in the entity.detail field – this is useful to provide precise linkage to the entities for each patient, however it is not indexed for ease of searching
2) Searching the IAM history – this contains details of the patient (includng NHS Number) and is fully indexed - albeit only for Direct Care interactions which reference a patient in the JWT token

To further strengthen the ability to search audit records directly by NHS Number in all cases, additional "NHS Number" entity entry(s) will be added to each audit event as follows:

entity.identifier	The NHS Number
entity.reference	Not populated
entity.type	"nhs-no". *(An extension to the standard code list)*
All other entity fields	Not populated

An additional NHS Number entity such as this will be added to the audit event for each distinct NHS Number that is involved. The entity.identifier field is indexed, thus allowing for ease of searching.

2.9 Additional Auditing Requirements for IAM

Rich information about the originating user, organisation, and system are captured in the JWT which was used to gain access to data. Furthermore, when the token is for direct-care then information about the patient is also captured. JWT and FHIR are complementary standards, and it is not possible to capture and index the full richness of information in a JWT claim within the fields available in a FHIR Audit event. These are gaps in the audit record in must be filled by IAM.

IAM must persist:

- all signed claims;
- a reason for rejecting a claim;
- the JWT issued in response to a claim;
- the time the claim was received;
- a reference to local persona of the user as maintained by IAM;
- the IP address of the network device from which the socket connection was made to IAM.

Successful claims must be keyed by the JTI of the issued JWT and indexed by all fields in the claim presented to IAM.

- The JTI acts as a "session identifier" and is referenced in all related FHIR Audit records. When reviewing FHIR Audit records then the JTI provides a mechanism to link back from the FHIR Audit records to the full information contained in the original JWT
- This fully indexed history of IAM claims provides an alternative starting point for audit investigations - and may often be the best place to begin a search. Having identified relevant IAM tokens (eg based on searching by user, patient, timestamp, issuer, etc) then the JTIs can be used as a key into the FHIR Audit records to extract further details.

2.10 Retention Period for Audit Records

Retention applies to both:

1. The FHIR Audit Records
2. The IAM History of JWT tokens issued (IAM Fire Store)

The retention period for online audit records is configurable, and be set to the same value in both cases. This will be tuned based on storage and performance considerations. Initially data will be kept online indefinitely, but in future records may be archived to offline storage as performance and sizing needs dictate.

3 Minimal Auditing Requirements for Data Consumers and Data Providers

3.1 Key Principles

In an ideal world then all Data Consumer and Data Provider systems would implement a FHIR-compliant audit approach as described in this design paper. This would include:

- Keeping an audit based on FHIR Audit events, as defined above
- Providing a FHIR endpoint to allow suitably authorised "auditor" users to retrieve audit records automatically

This is the approach being taken with the new systems being created by the YHCR – ie the Portal (Data Consumer) and the FHIR Appliance (Data Provider).

It is recognised however that there will be situations where the YHCR needs to interact with pre-existing Data Provider and Data Consumer systems. These are likely to have existing and established audit approaches – of equal merit, but not based around FHIR Audit events as described in this paper. In many cases it may be unrealistic to expect these existing systems to implement a new audit approach. However to enable a robust end-to-end audit capability there are key principles which all participating systems must uphold:

1) The ability to provide, on request, an extract of relevant audit records to the YHCR if needed for an investigation.
    - This paper envisages an automated approach, based on the ability to call a FHIR Audit API. However as a minimum there must be a manual process to request an audit extract (eg via a Service Request)

2) The ability to correlate the JTI Session Ids used by YHCR with existing internal audit trails
    - This is likely involve implementing a "link" mechanism between existing internal audit identifiers and the external YHCR JTI ids

3.2 Requirements for Data Consumers

A data consumer MUST be able, on request, to provide correlation between the JTI of the token used for access to the YHCR and the local session which provides the context of their activities. (For example any local IG controls and other related data access and activities within the session)

A mechanism to achieve this would be:

- audit every authentication attempt with sufficient information to identify the authenticated individual;
- audit every claim made to IAM (including renewal of expired JWTs) and for successful claims record the JTI of the returned JWT;
- associate audit records of claims made to IAM with authentication audit records;
- provide search facilities to identify the authenticated individual from a JTI.

Another possible approach would be:

- Audit every interaction with YHCR, including both the JTI from the returned YHCR token and the local session id

- Provide search facilities to review this audit by both JTI and local session id

3.3 Requirements for Data Providers

A data provider MUST be able, on request, to provide an extract of activity based on a JTI provided by the YHCR

A data provider MUST be able, on request (and subject to suitable Information Governance controls) to provide an extract of data based on FHIR Resource references to assist in an investigation

4 Additional Requirements for Regional Components

A FHIR based approach to auditing provides a uniform foundation for transactional auditing. FHIR resources are indexed and can be queried in a variety of manners which supports investigative interrogation of audit data. Various questions can be asked of the audit data such as:

- Who has seen what data for a particular patient?
- What data has a particular user accessed?
- How did a particular resource come to be in the state that it is?

The audit data is transactional, and the transactions can be assembled when a question needs to be answered, However, the format of the data is not well suited to more forensic type analysis, analysis which prompts the investigative questions. Forensic analysis might involve questions such as:

- Which user has accessed the most patients' records this week?
- Have any users accessed patient records over an unusual geography?
- Do any users have unusual temporal patterns of access?
- Have any users made bulk updates to data?

These questions have potential to reveal improper use of the YHCR and are arguably the most important product of audit data. However, they need data in a non-transactional format: multi-dimensional indexes over key properties of the audit record or OLAP cubes.

Requirements for forensic analysis will evolve over time. However, at this stage it is proposed that focus be restricted to FHIR operations passing through the FHIR Aggregator and the following dimensions of analysis:

- the end-user requesting the operation;
- the type of operation (read, create or update);
- the data provider hosting the FHIR resource operated on;
- the time of the operation;
- the patient who is the subject of the FHIR resource operated on.

5 Security of Audit Records

From version 1.1, design paper 005 – “Identity and Access Management” defines an Auditor role. This role has read only access to *AuditEvent* resources and no other type other resource. No other role has access to *AuditEvent* resources.

All access to *AuditEvent* resources must be logged by FHIR Stores in FHIR Proxy Servers at data providers and the regional FHIR Store. The log must be written to a file in a directory on a file system where write access is restricted to the root account and the account running the FHIR Store software. The log must record:

- the time of access of the AuditEvent;
- the user accessing the AuditEvent;
- the identifier of the AuditEvent resource.

Appendix 1 – Approach to Audit Investigations

All audit investigations will be coordinated centrally. Where concerns about improper use exist then these should be raised in the first instance with the central Information Governance team. This team will then raise any Service Requests necessary for supporting technical activities.

It is envisaged that the starting point for most audit investigations will often be the central IAM History. This is a rich source of information, fully indexed and searchable, and including:

- User – UserId, name, role, organisation (ODS code) and other identifiers (eg staff number)
- Patient (for Patient Centric, direct care) – NHS Number, name, DoB
- Consumer System – System Id, and owning organisation (ODS code)
- Reason for access
- Timestamp

See ***YHCR Design Paper 005 - Identity and Access Management*** for further details of the JWT token contents.

- This IAM History will allow an initial assessment of activity by patient and/or user to be established.
- It can also provide a basis for analysis of other more complex requests based on other criteria such as system, organisation or reason for access. (Noting that very complex requests may require an initial "rough cut" of data to be extracted into other analytical tools for further refinement)

Based on this initial assessment, the unique JTI token identifiers may be used to add detail as required. For example:

- Using the JTIs to linking with the central Audit FHIR Store to elaborate the specific actions undertaken (eg queries executed, success or failure of the request, identifiers of entities returned)
- Using the JTIs and/or entity identifiers to link with a Data Provider Audit and Resource FHIR Stores - thus following the entity references and see the actual data items returned
- Using the JTIs to link with a Data Consumer and discover more details of the user's session (eg usage statements agreed with, screens actually viewed, other related activity of the user)

Additional notes:

- The IAM History will only contain NHS Number information for Patient-Centric (direct care) interactions. A more general search of the central Audit FHIR Store based on the NHS Number may therefore reveal additional information for non-direct-care requests. This additional search of the Audit FHIR Store by NHS Number can be done using the additional entities described in Section 2.8. In this case the JTI may be followed in the other direction - back from the Audit FHIR Store to the IAM History to discover further details of the user and other token-based information.
- For investigations which pre-date the online retention period then data will need to be restored prior to the investigation beginning – this may add extra time to the request. A decision will need to be made at that point whether it is most effective to reload the data into the original system or into a separate analytical area.

Appendix 2 – Maturity Matrix

Section	Narrative	Consultative	Draft	Normative
1 Introduction 1.1 Purpose of this Document	X			
1.2 Topics of Interest	X			
1.3 FHIR and Auditing	X			
1.4 Relationship of this Document with Other Standards	X			
1.5 Intended Users of the This Document	X			
2 General Requirements for FHIR Endpoints 2.1 Persistence of AuditEvent Resources			X	
2.2 Considerations for the use of AuditEvents		X		
2.3 Event Identifiers		X		
2.4 Outcome Codes		X		
2.5 The YHCR AuditEvent Profile		X		
2.6 Agent Contents		X		
2.7 Entity Contents		X		
2.8 NHS Number Entity		X		
2.7 Additional Auditing Requirements for IAM		X		
2.8 Retention Period for Audit Records	X			
3 Minimal Auditing Requirements for Data Consumers		X		
4 Additional Requirements for Regional Component		X		
5 Security of Audit Records		X		