Management and maintenance policies for EHR interoperability resources

Authors:
Dipak Kalra, University College London, UK
Gerard Freriks, TNO, NL
François Mennerat, ProRec France, FR
Jos Devlies, ProRec Belgium & EuroRec, BE
Archana Tapuria, University College London, UK
Geert Thienpont, RAMIT, BE

Table of Contents

1 Introduction

This report, produced as part of the Q-REC project, describes the quality management and maintenance landscape for four kinds of resource that support the interoperability of electronic health records:

- Clinical archetypes;
- Open source components and XML Schemas;
- Legislative and industry standards;
- Coding schemes and terminology systems.

This report is to be read in conjunction with two others published by the EuroRec institute as part of its series of Q-REC project reports:

- Deliverable 3.1: Inventory of resources supporting EHR interoperability
- Deliverable 3.2: Repository and web portal providing managed access to resources

Deliverable 3.1 focuses on inventories of standards publications and on coding schemes, for each of which it is presently practical to enumerate and document a set of resources that contribute to the design and implementation of interoperable EHR systems. Deliverable 3.2 focuses on the design of and maintenance polices for the EuroRec repository of EHR quality criteria that have been developed through the Q-REC project.

This report, Deliverable 3.3, focuses on archetypes, for which a definitive inventory is not feasible at this stage in their evolution but for which quality criteria and a pathway towards certification are realistic to propose. This report also focuses on open source components (including but not limited to XML Schemas): this report proposes that it is also not feasible to inventorize these and also that formal quality criteria are difficult to assert for this class of artefact. This report also briefly outlines how the inventories of standards and coding schemes will be maintained and managed by EuroRec.

2 The quality labelling and certification of archetypes

2.1 Introduction

Clinical archetypes are specifications of the knowledge data and their inter-relationships that play an important role in determining how clinical information is represented and organised inside EHRs when they are interoperably communicated between systems. Archetypes will often also influence the way in which clinical data are managed within individual EHR systems, how users enter data and how data are presented. They form an absolutely essential part of the semantic interoperability underpinning of e-Health. Archetypes therefore need to be quality managed and quality assured.

The experience and evidence base for the large scale design and use of archetypes is not yet mature enough for a robust certification process to be applied to them, but consensus good practice quality criteria are emerging. This report summarises the origins and state of the art of clinical archetypes, and their trajectory for adoption internationally. It proposes a set of quality criteria by which archetypes may be assessed, and which will evolve to become formal quality labelling criteria.

2.2 Background to Archetypes

The origin of clinical archetypes
The requirement for clinical teams to share patient record information to support longitudinal continuing care and to follow multi-professional care pathways is well recognised. Delivering shared regional or national Electronic Health Records is now central to every e-Health programme. It is also recognised that the support of shared care through records that are only human readable is not sufficient: patient safety management and the pursuit of evidence based care require computable information that can be linked to and queried by alerting components, decision support and clinical pathway systems. The efficient management of health services, and the support of public health and clinical research through audits and population analyses also require semantically processable EHRs. These computable purposes of use ideally also require that the clinical findings within EHRs are represented and organised consistently across vendor products and communities of use: semantic interoperability.

International research over the past sixteen years has catalogued the clinical, ethical and technical requirements that need to be met in order to realise interoperable EHRs, for example published by the Good European Health Record project, the EHCR Support Action and the Synapses project. These have formed the basis of architecture formalisms and standards to represent and communicate personal health data comprehensively, in a manner that is medico-legally rigorous and preserves the clinical meaning intended by each original author. Concerns about protecting the confidentiality of sensitive personal information have also to be addressed if consumer confidence is to be maintained when EHRs are widely accessible. Much of this work is now embodied within international standards for EHR architecture requirements (ISO 18308), EHR interoperability (ISO/EN 13606) and through a major open source initiative: the *open*EHR Foundation.

The approach pioneered by the Synapses project and the *open*EHR Foundation, and mirrored in EHR interoperability standards, is to define a Reference Model as a generic high-level representation of the EHR and its governance and medico-legal properties. These properties include:

- dates and times of when observations occurred, health events took place and when information was recorded;
- persons who provided, composed, entered or authorised (signed) particular entries, or who played particular roles in a health care process;
- version management information, including who changed any of the entries, when and why;
- the degree of sensitivity of the information and who should be allowed to access it;
- a clinical label for each point in the record hierarchy, to name each folder, document, heading and the parts of each detailed entry;
- who the information is about, if not the patient (e.g. if it about a family member, or a third party);
- a standard way of representing coded clinical terms, measured quantities, dates, times and various kinds of multimedia data;
- a linkage between the component entities documented in an EHR and formalised knowledge representations, now known as archetypes, that have directed how each clinical entity is structured within the EHR.

Provided that the Reference Model for the EHR is standardised across sending and receiving information systems, any health record extract exchanged between them will contain all of the structure, names and medico-legal information required for it to be interpreted faithfully on receipt even if the nature of the clinical content have not been "agreed" in advance. This is sometimes termed structural or syntactic interoperability.

However the wide-scale sharing of health records, and their meaningful analysis across distributed sites, for example to enable decision support, also requires that a consistent approach is used for the clinical (semantic) data structures that will be communicated via the Reference Model, so that equivalent clinical information is represented consistently. This is necessary in order for clinical applications and analysis tools safely to request and process EHR data that have come from heterogeneous sources. This is termed semantic interoperability.

Clinical archetypes are a formal, rigorous and standardised (interoperable) specification for an agreed consensus or best practice representation of clinical data structures (within an electronic health record). They provide a standardised way of specifying EHR clinical data hierarchies and the kinds of data values that may be stored within each kind of entry. An archetype defines (or constrains) relationships between data values within an EHR data structure, expressed as algorithms, formulae or rules. An archetype may logically include other archetypes, and may be a specialisation of another archetype. In order for it to be managed and used appropriately, its metadata needs to define its core concept, purpose and use, evidence basis, authorship, versioning and maintenance information.

Archetypes offer a tractable way of binding generic EHR models to compositional terminology. They provide target knowledge representations for use by guideline and care pathway systems, and so support knowledge level interoperability: systems may interoperate not only at the data level, but also at the concept level. EHR entries identify the archetypes used when the data were created, and/or to which they map, which aids future interpretation, analysis, querying.

Clinical archetypes originate from over ten years of research and clinical demonstrators in Europe and Australia; much of this work has in recent years been pioneered within the *open*EHR Foundation. ISO/EN 13606 Part 2 incorporates the *open*EHR archetype approach as a standard knowledge model, and an exchange representation, for the communication of clinical archetypes. This approach therefore has significant international endorsement.

Archetype specifications and tools

Clinical archetypes are a knowledge representation that define the way in which the EHR Reference Model is to be applied to represent particular clinical entities (i.e. particular kinds of finding, assessment, hypothesis, plan or intervention). An archetype defines a data structure, including optionality and multiplicity, data value constraints, and relevant bindings to natural language and terminology systems

To support semantic interoperability, these archetypes need to be shared and used consistently by EHR system vendors and their users, so that the EHR data they create is consistently organised. Archetypes therefore need to be shared and managed as a common knowledge asset, and incorporated into the design of clinical applications, rather like a terminology system. Many of the formalisms and tools needed for archetypes to be a global resource are now in place.

A formal set of technical requirements for the representation of an archetype are included as a normative section in ISO/EN 13606 Part 2.

An object model for the representation of archetypes (in UML) has been published by the *open*EHR Foundation and also included as a normative specification in ISO/EN 13606 Part 2.

One particular serialised format for representing an archetype, Archetype Definition Language (ADL), that conforms to the object model, is published and maintained by the *open*EHR Foundation, and has been included in ISO/EN 13606 Part 2. Other serialised formats such as XML are in development.

Several tools have been published for the authorship and technical validation of an archetype: by Ocean Informatics and by the University of Linköping. Others tools are being developed that link archetype specifications to legacy/feeder systems. These are now being used by many different groups and projects across the globe, and are being constantly refined to reflect the requirements that arise from these user communities.

Ongoing research and development activities within the *open*EHR community are currently focussed on: repository services to store and distribute archetypes, knowledge management services to support searching for and comparisons of different archetypes for specific purposes, template tools to profile and combine archetypes for specific clinical workflows, and the definition of business rules and design guidance for the binding of archetype nodes to SNOMED CT. University College London and Ocean Informatics are each also exploring the inclusion of application presentation and workflow instructions within archetypes that have been operationalised for internal EHR system use, towards the realisation of fully knowledge-driven systems.

One notable challenge in designing libraries of archetypes to meet broad areas of clinical practice, for example to cover the complete clinical information needs of a

speciality or professional discipline, is to ensure that archetypes are evidence based or meet *de facto* agreed clinical needs (e.g. established by consensus, or reflecting existing practice) . Given that many archetypes may be needed to cover a given domain, it is also important for them to be mutually consistent and bind to terminology systems in appropriate and consistent ways. This is necessary in order to minimise the diversity of ways in which a given kind of EHR data might be represented. This consistency is needed by clinical applications, decision support and other analytic software that need to retrieve or filter EHR data, or assist users with selective navigation through a large EHR or across populations of EHRs. In order for them to be accepted and adopted widely, archetypes also have to be of demonstrable good quality. This is an area of ongoing learning through communities that have begun to build up libraries of archetypes for their clinical domains.

2.3 The benefit of using archetypes

For semantic interoperability to be possible communicating partners need to achieve agreements at many levels, from technology up to coding systems. Complex IT-systems can be described using 5 independent layers[1]. Each level needs its own set of agreements and standards:

- Enterprise Viewpoint
  - Purpose, scope, and policies (not necessarily means "real enterprise")
  - Often represented through a requirements specification
- Information Viewpoint
  - Meaning and processing of information
  - Information model and schema
- Computational Viewpoint
  - Functional decomposition (close to application objects)
  - Often represented through an interface object model or component models
- Engineering Viewpoint
  - Infrastructure to enable interaction between computational objects
  - Could be considered as infrastructure/middleware for distributed processing systems
- Technology Viewpoint
  - Choice of hardware and software, and conformance

Two methods to support semantic interoperability for electronic health records are available today: messages and archetypes/templates.

Messages
An older way is the use of messages. It is a characteristic of messages (EDIFACT, DICOM, HL7v2, HL7v3) that in one message specification (message standard) several viewpoints are defined rather than just one:

- Enterprise viewpoint will contain the use case, i.e. the standardised work process,
- Information viewpoint contains the Message Information Model

---

[1] ISO/IEC 10746: 1996 Open Distributed Processing - Reference Model Parts 1-3

- Computational viewpoint is about the choreography of messages in the Interaction schema's
- Engineering viewpoint is the level where the XML schema is defined.

In any message specification changes can occur at any or all layers. Work processes change, new data elements need to be stored or exchanged, new interfaces are needed, etc. Even the smallest change will lead to a new version of the message. Since the implementation of messages in all EHR-systems in an uniform way (e.g. via the IHE process) is time and money consuming, it is clear that messages do not facilitate innovation because the flexibility and adaptability of this technology is poor.

Archetypes/Templates based on ISO EN13606 and *open*EHR
In healthcare archetypes and templates express the requirements from the Enterprise viewpoint level as constraints on the Reference Model. The Reference Model[2] is a very generic model of any health record or document. The resulting collection of defined archetypes and templates constitute the Information Viewpoint.

The European standard EN13606 defines how archetypes and templates are produced in a standardised way. Therefore the European EHR-standard is operative on the Information Viewpoint level only.
*open*EHR[3] has extended the European EHR-standard to the Computational Viewpoint so that EN13606 conformant EHR-systems become possible. (Other standards will govern the other ODP layers.)

Archetypes and templates play a key role in semantic interoperability. Archetypes define what is maximally documented in the world about a specific health record entity. Templates define what in a specific context at a specific point in time, will be stored, retrieved, presented, exchanged and archived. In part, clinical meaning within an EHR will be expressed through the structure of the archetype/template, and in part the meaning will be expressed through codes from coding systems. A way to view this metaphorically is:

- codes are the words in a dictionary;
- the structure of the archetype/template is the grammar;
- with both codes and archetypes sentences can be formed that make or do not make sense;
- but archetypes define what makes sense.
- and templates define what makes sense in a specific context.

In the case of EN13606/*open*EHR archetypes provide a lot of flexibility and adaptability. Using archetypes healthcare providers can define and re-define at any moment templates that are needed in their work process at that point in time. Systems based on archetypes and templates support easy customisation and localisation, and rapid evolution to meet new clinical requirements.

---

[2] Note: the Reference Model of the EN13606/*open*EHR is not the same as the Reference Information Model of HL7.
[3] A not-for-profit organisation Please see www.openEHR.org

2.4 Archetype quality criteria

Clinical archetypes need to be quality assured, since they will direct the ways in which clinical data is captured, processed and communicated. It is important that the design of individual archetypes is an accurate and faithful reflection of good practice for the clinical disciplines in which each of them might be used. They need to be optimally designed for their purpose, and considered trustworthy within their intended communities of use. This requires not only sound methodologies for designing each archetype in accordance with, for example, published clinical guidelines or peer consensus, but rigorous and robust processes for validating any given archetype against its clinical evidence base and in the context of other archetypes alongside which it might be used. Pan-European (or international) applicability will be an increasingly-important requirement for good quality archetypes.

If a clinical team is to trust and adopt a set of archetypes, the following set of questions probably all need to be addressed through archetype quality assurance and quality labelling:

- is it clear what clinical situations it is to be used for?
- how inclusive is it of the kinds of patients we treat?
- is it flexible enough for our needs?
- what kinds of patients is it intended for? (children?, elderly?)
- has it been designed with suitable multi-professional input and domain experts?
- what clinical evidence and guidelines does it follow?
- or, is its model based on an existing well-accepted system?
- by whom has the archetype been peer reviewed?
- has it been endorsed by one or more professional bodies?
- has it been quality labelled by a body that I trust?

For a regional care manager, the questions might be:

- what clinical use cases has it been designed for?
- will it be used consistently and safely across care teams?
- does it align with other archetypes we use: it is clear how they fit together?
- has it been approved by my national health service?
- what national data sets does it conform to?
- what terminologies (and versions) does it bind to?
- will it align with national audit and governance reporting?
- how up to date is it?
- when and who will review and maintain it? how frequently?
- has it been quality labelled by a body that I trust?

And how could a CTO or vendor know if an archetype is safe to implement?

- which use cases and users should have access to it?
- does it clash with any other archetypes we already implement?
- does it conform to a technical standard?
- does it align with data standards that I also have to report on?
- has it been tested?
- can I verify the authenticity of the copy I have?
- can I verify its currency (is it the latest version)?
- how will I be notified of updates?

- how are terminology bindings maintained and disseminated?
- it is published by a certified repository?
- has it been quality labelled by a body that I trust?

If record-sharing communities are to construct safe EHR instances in accordance with archetypes, and to trust EHR data conforming to archetypes, a formal process of verification and certification is needed for archetypes that provides assurance of their suitability and safety. The EuroRec Institute is partnering the *open*EHR Foundation in developing governance practices for archetype development, and the quality criteria and editorial policies by which certified libraries of archetypes can be recognised.

Since the development of large libraries of archetypes by clinical communities is still relatively new, the experience and evidence base for the quality assurance and quality labelling of archetypes is not yet strong enough to support a formal certification process. However, there is a growing consensus on the kinds of quality criteria that good archetypes should meet, which are described in this section.

Business requirements

- An archetype shall define a formal representation for one or more discrete kinds of clinical (health or health care) entity within an electronic health record.

- An archetype shall define the structural organisation and kinds of permitted data content for representing one or more clinical entities as a use pattern (i.e. a constraint pattern) for a specified electronic health record information model.

- An archetype shall specify the use pattern in sufficient detail and with sufficient precision that different conforming clinical data instances drawn from different EHR systems and communities of practice can be represented consistently when using the same (specified) electronic health record information model.

- An archetype shall include or reference information about its intended usage sufficiently that potential or current technical or clinical adopters can unambiguously determine the clinical scenarios and kinds of EHR data to which it applies.

- An archetype shall include or reference information that enables a potential or current user to determine its evidence basis, quality and currency.

- An archetype shall include or reference information that informs a current or potential user about the certifications, approvals and uses of it, globally.

Clinical requirements

Clinical usage requirements

- An archetype shall specify the precise nature of the clinical entity (or set of entities) for which it defines a use pattern.

- An archetype shall specify any particular clinical scenarios or workflows for which it is particularly intended.

- An archetype shall specify any particular sub-populations of citizens for whose health or health care it particularly applies.

- An archetype shall specify any particular speciality, discipline or professional groups for whose use it is primarily intended.

- It shall be possible for an archetype to include specific usage guidance, such as a restriction to certain sub-populations or scenarios, that apply to individual EHR nodes and/or constraints within it (i.e. that apply to individual parts of the archetype rather than its whole).

Clinical domain coverage

- An archetype shall include or reference one or more concepts from an internationally registered terminology system to which it corresponds most closely, in order to permit its clinical scope to be widely understood, and compared with other archetypes.

- The clinical scope of an archetype shall be sufficiently precise that EHR instances conforming to the archetype may meaningfully be interpreted and analysed collectively (i.e. that their data values are comparable - e.g. it would be meaningful to list the values in a table or plot them on a graph).

- An archetype shall include or reference sufficient information to permit areas of clinical scope overlap between archetypes to be identified, for example by mapping individual nodes within it to internationally registered terminology concepts.

- An archetype shall be able to include part or all of another pre-existing archetype if part of the entity it represents has already been defined in a way that meets the requirements of its use cases and users; such re-used archetype fragments shall be identifiable as being identical across the various archetypes that use them.

- An archetype shall be able to be a constrained (specialised) version of a pre-existing archetype, for example to narrow its applicability to a sub-set of the use cases of the original archetype; a specialised archetype shall be uniquely identified independently of the archetype it specialises, but shall identify the archetype it specialises.

- It should be possible to identify parts of two or more archetypes that have the same scope (i.e. if they define constraints to represent the same portions of a clinical entity), so that differences or similarities between them can be recognised.

- An archetype's use pattern should be inclusive of all of the minor variations in clinical entity representation across its use cases, users and scenarios; i.e. it should be a superset of the sub-components of the various required representations of the clinical entity.

- The representation of sub-components (data items) of an entity within an archetype should be optional unless those data items are accepted to be mandatory across all of its intended use cases, users and scenarios.

- Notwithstanding the above, an archetype's design should avoid meeting an over-inclusive set of use cases and including so many optional properties that it results in very diverse kinds of conforming EHR instances.

- An archetype's scope should be focussed enough that the likelihood of overlap with other archetypes in the same domain is minimised.

- An archetype should reflect the extent of consensus and degree of alignment of requirements across the relevant user communities; multiple archetypes

should be considered in areas where consensus is limited or sound reasons exist for fostering diversity of representations (such as an area of active research or innovation, or to comply with differing jurisdictional policies).

- It should be possible for a community of practice to identify the set of archetypes that is relevant to its domain, and to identify the extent of domain coverage (including gaps and overlaps of coverage).

Evidential basis

- An archetype shall be able to include references to one or more kinds of published knowledge that have informed its overall design, and/or to which it conforms; (examples of relevant knowledge include clinical guidelines, care pathways, standard data sets, professional policies, reporting templates).

- An archetype shall be able to include references to one or more kinds of published knowledge or policy to which any individual node or nodes within it conform.

- An archetype shall enable any reference to published knowledge or policy to include a textual reference to it, a description of it, an executable link such as a URL, and any notes provided by the author to specify the extent of conformance, or reasons why conformance has not been considered appropriate or feasible.

- An archetype shall enable any reference to published knowledge or policy to include a date when that knowledge is due to be reviewed (and therefore when the archetype itself might also need to be reviewed).

- An archetype shall be able to include information about *de facto* specifications (such as existing clinical information systems) that have been its primary design basis.

- An archetype shall be able to include information about the set of clinical and non-clinical stakeholder communities that have provided requirements that it meets.

- An archetype shall be able to include information about the set of clinical and non-clinical stakeholder communities that have verified its correctness via peer review.

- The author of an archetype should first ensure that appropriate effort has been made to identify relevant evidence, consult relevant stakeholders and examine existing systems in use.

- The author of an archetype should first ensure that any existing published archetypes are examined for potential duplication or overlap, and should aim to re-use relevant existing specifications.

- An archetype should specify if its draft versions have been through an open consultation or social computing form of peer review (e.g. published on a wiki site for public comment).

Communities of use

- An archetype shall be able to include or reference multiple instances of information relating to its scope, purpose, usage, stakeholders and evidential basis so that different communities of use can include such information as is relevant to their own jurisdictions.

- An archetype shall be able to include multiple natural language translations of any or all of its textual content, and be able to distinguish pure translations

from alternative wording for a different community of practice or jurisdiction.

Technical requirements

Conformance to standards

- An archetype shall conform to the requirements specified in Section 6 of ISO/EN 13606 Part 2.

- The information in an archetype shall be capable of being represented using the information model specified in Section 7 of ISO/EN 13606 Part 2 or to any more recent model version published by the *open*EHR Foundation (on www.openEHR.org).

Modelling requirements

- An archetype shall specify the EHR information model for which it is a use pattern.

- An archetype shall specify the class within the EHR information model that is the root for EHR instances that conform to the archetype's constraints.

- Every node in the archetype shall specify the class within the EHR information model that is the corresponding node for EHR instances that conform to the archetype's constraints.

- The identifier of an archetype, and of each of its nodes, shall be globally unique and replicated consistently whenever it is communicated.

- The clinical label for each node (its name) shall be drawn from a published controlled vocabulary, and preferably from a published international terminology.

- The definition of each node shall permit the unambiguous and consistent mapping of appropriate original EHR data and EHR system data items to it.

- Any node in the archetype shall be capable of being mapped to additional terms that offer an equivalent meaning to its name, and to natural language translations of the name.

- The existence and multiplicity (cardinality) of each node will reflect the most inclusive requirements from across its use cases and users (i.e. it will specify optional in preference to mandatory constraints unless there is a consensus otherwise).

- An archetype hierarchy shall avoid redundant, duplicate or near-duplicate nodes unless there are clear requirements for these and their definitions make clear how each is to be used.

- Data value constraints (such as value ranges and term value lists) shall cater adequately for the diversity of anticipated values from its defined patient populations.

- If an archetype constraint permits a null data value, it shall have been verified that corresponding EHR instances are consistent with the requirements of the user communities and do not introduce the risk of ambiguous or unsafe meaning to the clinical entity being represented.

- The term value list associated with an archetype node that has a coded or enumeration data value shall be demonstrably consistent semantically with

the name of that node.

- The units of measurement (and corresponding value ranges) should permit the use of any relevant units that are in appropriate use internationally.

- Language translations for values within a term value list shall always be complete and correspond to the original terms on a specified one-to-one basis.

- An archetype node shall constrain the values of any relevant properties of the corresponding class of the EHR information model to exclude values that might otherwise contradict or conflict with the consistent representation of the clinical entity corresponding to the archetype as a whole.

- References to term value lists by means of a pattern or query specification for a given terminology shall specify the terminology system and version for which it has been validated.

- References to other archetypes and/or archetype fragments to be included within an archetype shall be specific to the version of each archetype.

NOTE: research is presently in progress to define business rules for the appropriate binding of archetype nodes to SNOMED-CT: this work is expected to give rise to additional quality criteria relevant to this sub-section.

Information governance requirements

Authorship

- An archetype shall always include information about the person and/or organisation that has taken primary responsibility for its creation.

- An archetype shall always include information about the person and/or organisation that has taken primary responsibility for its design basis.

- The person and/or organisation details may include professional or academic qualifications, organisational accreditation or other credentials.

- An archetype shall include the data and time and location (jurisdiction) of its creation.

- An archetype shall include the data and time when it must either be reviewed (to verify its clinical validity and evidence basis) or deprecated.

- An archetype shall specify the party or organisation that is primarily responsible for its quality maintenance.

Version management

- Any modification to an archetype shall result in a revised version that references the former version.

- Archetype version management shall distinguish modifications: (a) that extend its descriptive or quality management data but do not alter its current use or the constraints that determines conformant EHR instances; (b) that enlarge or reduce or alter the ways in which it might be used but do not alter the constraints that determines conformant EHR instances; (c) that alter the constraints and extend the domain of conformant EHR instances (i.e. the change is backward compatible).

- No revision to an archetype may render non-conformant any instance of EHR

data that conformed to a previous version: in such circumstances a totally new archetype shall be created and the existing archetype shall, if appropriate, be deprecated from further use.

- All modifications shall specify the person and organisation responsible for the change, the date and time of the change, a description of what has been changed and the reasons for making the change.

Access and licensing

- An archetype shall include a clear statement of any copyright or usage restrictions that apply to it.

- An archetype that has restrictions on its use shall include license information and details of how any relevant permissions may be obtained.

- An archetype shall include a clear indication if it is a draft version (and liable to change), or if it is deemed complete but has not yet been endorsed by its authoring organisation.

Endorsement, quality labelling, certification

- An archetype shall list and date stamp any approvals and endorsements for its use in different jurisdictions or by different communities of practice.

- An archetype shall include a time-stamped indication of its intended deprecation from future use by any jurisdiction, optionally with an explanation of the reason and optionally a reference to any successor archetype(s).

- An archetype shall list and include or reference any formal certification or quality labels that have been applied to it.

- Any quality label applied to an archetype shall include the date and time and details of the approving body.

- Any quality label applied to an archetype shall include the date and time when this approval is to expire (unless it is renewed).

Archetype repository requirements

- The controller of an archetype repository shall publish and implement a quality management plan that includes a quality assessment of any candidate archetype offered for storage; this might for example be undertaken by a scientific review board.

- This quality assessment shall include either the undertaking of a validation against the quality criteria listed here or any future more formal criteria, or by requiring evidence of this assessment having been undertaken by the archetype authors, or by ensuring that the archetype carries a quality label or certificate from a recognised issuing body.

- The controller of the repository shall ensure conformance to any relevant licences or restrictions for use of an archetype, and provide appropriate means for potential users of it to be informed of these.

- The repository shall index each contained archetype using terms and other mechanisms that enable users and software components to locate the set of archetypes that are relevant to a query or retrieval request.

- The repository shall enable archetypes to be identified by searching on any of its structured information properties.

- The repository shall support the provision only of archetypes that have been certified or quality labelled, or approved for use within a given jurisdiction, if this is a condition specified in the request.

- The repository shall be able to provide any if its archetypes in at least one format that conforms to a published international standard or specification.

- Where more than one format is supported, a user or requesting service shall be able, per request, to nominate one of these as the preferred retrieval format.

- The repository shall ensure that it can be notified of any modifications or updates to an archetype that it holds by its original authors, or other recognised authoring bodies, in a timely fashion.

- The repository and its services shall maintain a complete and audited version history for all of its archetypes.

- Requesters of obsolete versions of an archetype shall be provided with a notification that an update (or updates) exist and be able to nominate the version(s) to be returned.

- An archetype repository shall support a standardised set of interfaces and services once these are defined.

- A repository service should provide a notification service to its registered clients of relevant archetype updates and additions.

- A repository service should provide a service whereby registered clients may maintain and keep synchronised a local copy of the set of archetypes for a given domain.

2.5 The adoption of high quality archetypes

For semantic interoperability based on the European EHR-standard to become possible it is absolutely essential at least in Europe, but preferably in the rest of the world, that healthcare providers use the same sets of archetypes with their bindings to terminology systems.

Most public domain archetypes produced so far have been published via the *open*EHR web-site and by the NHS in England, both in a relatively non-quality-assured fashion. The European Institute for Health Records has agreed to play a central role in Europe in Archetype Governance[4].

Archetype/template production developments

Initial phase
Many of the archetypes and templates that have been produced so far have been exploratory learning exercises. *open*EHR has produced a initial collection. Others like TNO in the Netherlands, NHS-Scotland and the NHS-England have produced archetypes and templates to be used in diabetes care, nursing, maternity, emergency medicine and mental health. They have found this to be a successful

---

[4] This is through the Q-Rec project.

way to obtain clinical information requirements from healthcare professionals, but experience of their operational use is limited.

Start-up phase
With the official adoption of the European EHR standard and the *open*EHR implementable specification in countries and regions, the process will really take off when several places in Europe and the world start at scale to produce archetypes to be used for semantic interoperability. At this moment (May 2008) Australia, the UK, Denmark, Sweden, Latvia, Slovakia, Belgium and the Netherlands clinical groups are preparing to start the production of archetypes and templates, plus their mappings to terminology systems. It is expected that one or more European projects will surface within the next year where archetypes and templates play an important role. (possibly through the European SOS-LSP[5]).

Archetype production: key players
Key players in the production of archetypes and templates are:

- healthcare providers defining their needed templates
- health information experts
- healthcare scientific bodies providing input to the health content of archetypes
- Standardization bodies CEN/TC251, ISO/TC215, IDHTSO (SNOMED-CT)
- IT-industry
- education
- *The open*EHR Foundation
- National ProRec centers
- The EuroRec Institute
- Governments

Each of these involved players needs to play their respective natural role and collaborate with the others in order to achieve persistent high quality results.

Healthcare providers
It is the natural role of healthcare providers to express what data or information they have to store, retrieve, present, exchange or archive to support their local work processes. They need to be able to produce local templates, using quality assured archetypes from a national or European repository.

Health Information Experts
Healthcare providers are trained in the execution of care processes and not necessarily in the more generic science of information modelling, and so may not be proficient in the use of the archetype/template tools and in the definition of business rules. Healthcare providers will need the assistance of trained Health Information Experts who are knowledgeable about health, information, repositories, business rules and tools.

Scientific and professional bodies
At national, regional and international levels archetypes and templates will be produced that reflect the present state of the art. Therefore at all these levels scientific and professional bodies will play an important role in shaping and approving the common agreed sets of archetypes that reflect the state-of-the-art in their clinical fields or health sectors.

---

[5] Smart Open Services (S.O.S.): an open eHealth initiative for a European large scale pilot of patient summary and electronic prescription, a multi-national initiative funded by the European Commission from 2008-2011.

Standardization organisations
CEN/TC251 and ISO/TC215 have as their natural role the maintenance of ISO EN 13606 and other complementary EHR related standards. Since they are inevitably highly governed organizations with fixed procedures and business models, their strength is to maintain stable artefacts such as formal standards. They are not capable of maintaining an ever changing dynamic digital resource such as a knowledge artefact library.

IHTDSO, the Standards Development Organisation responsible for the Reference Terminology SNOMED-CT, does not behave in the same way as the statutory standards bodies like CEN and ISO: it maintains a very dynamic digital resource: SNOMED Clinical Terms.

Industry
It is the natural role of industry to adopt the ISO EN 13606/*open*EHR specifications in their systems. In addition, for their own profit and in the interests of their clients, they need a correct, complete and usable set of archetypes and templates. Based on their enormous historical expertise, industry could play a role in the production of archetypes and templates that are practical to implement and maintain.

Ocean Informatics has produced two open source archetype tools and made these available via *open*EHR. This year (2008) it will release the Archetype Knowledge Manager. This is a repository for archetypes and templates that makes their production, governance and publication possible at national, European and international scales.

It is expected that another worldwide operating IT-company[6] will start or support the formation of a European organisation responsible for the definition, translation and maintenance of health knowledge artefacts like coding systems, archetypes and templates. This will be based on its vision for IT in healthcare within next-generation IT-systems.

Education
The ISO EN 13606/*open*EHR based paradigm of archetypes and templates will be new for many actors in many sectors of society. A general awareness and education about the use and governance of data and information in the EHR is needed. At more detailed levels the Health Information Experts and industry need special training in the correct use of archetypes, templates and coding systems. Curricula need to be devised and existing educational resources need to be extended.

openEHR
The *open*EHR Foundation has played a major role as the custodian of an open implementable specification (which extends EN 13606) through its Architecture Review Board. This role is akin to that of the standardization organisations but in a synergistic pro-active flexible way for dynamic adaptation and extension of the fixed stable standards in the light of implementation experience. This *open*EHR implementable specification is a digital dynamic resource.

*open*EHR has started a Clinical Review Board that will be responsible for the quality assurance of archetypes, and will work in partnership with EuroRec on European certification. In addition, the Foundation is exploring long-term governance relationships with the International Health Terminology Standards Development Organisation (IHTSDO), since it is recognised that archetypes offer

---

[6] This announcement has not yet been made public.

a tractable way to quality assure a terminology against specific and well-specified use cases, and vice versa.

National and Regional ProRec Centers
Increasingly national and regional ProRec Centres have been formed and are active. It is their natural role to become a body where all key players mentioned in this document meet and work on local coding extensions and subsets, archetypes and templates, coding mapping and translations. Regions and Member States have their own identities, cultures and languages. At these levels ProRec centres will serve their constituents but have to work together at the European level within EuroRec.
An additional role is the quality labelling and certification of EHR-systems in cooperation with the European Institute of Health Records. Part of this is the governance of the quality of archetypes, templates and terminology bindings.

European Institute for Health Records (EuroRec)
The natural role for EuroRec is to promote and label patient-safe and privacy-respecting semantic interoperable EHR-systems in Europe that enable the provision of healthcare in Europe for all (mobile) citizens. For this purpose it takes responsibility to set up a European resource that national and regional responsible organizations like ProRec Centers can use.

As part of the quality labelling and certification of EHR-systems, it may take joint responsibility for the governance of archetypes and templates alongside the *open*EHR Foundation, since these artefacts play an extremely important role in semantic interoperability in Europe. EuroRec has started (or will start) discussions with bodies like the Commission, CEN/TC251, ISO/TC215, *open*EHR and IHTDSO in order to create a framework where all can become responsible for a defined aspect in their natural roles.

EuroRec has expressed its ambition to become a European Agency responsible for the quality (patient safety and privacy) of semantic interoperability in healthcare in Europe.

European Commission and Member States.
Making rules and regulations (Directives, Laws) are their natural roles. Since only European standards can play a defined role in legislation they can refer to relevant National and European standards in order to create the free movement in Europe of people, goods, money and services. They provide the (legal) framework and resources for persistent organisations like Standardization Organizations, Eurorec and ProRec centers to continue to provide a joint common high quality resource with high societal impact, needed for semantic interoperability in Europe.

In June 2008 the European Commission has announced the Recommendation[7] on semantic interoperability in Europe as a step towards European legislation.

2.6 Conclusions

There is widespread and world-wide recognition that a formalised and scalable means of defining and sharing clinical data structures is needed to achieve the

---

[7] European Commission Recommendation on cross border interoperability of Electronic Health Records, (in press)

value of investment in e-Health. Clinical archetypes are gaining acceptance as the best of breed and best supported approach for defining these structures, reflected in its international standardisation and the activities connected with the *open*EHR Foundation and its global community.

Large and comprehensive sets of archetypes are needed that cover whole clinical domains in a systematic and inclusive way, catering for the inevitable diversity of use cases and users but helping to foster consensus and best practice. For these to be endorsed by health systems, implemented by vendors and trusted by end users, these archetypes need to be quality assured, and to be published and maintained by reliable certified sources.

Through *open*EHR and EuroRec there is a momentum to establish the means to assess, quality label and certify archetypes and archetype repositories. This process will include organisational governance, artefact governance and quality assured processes, and will build on the approach and criteria presented in this deliverable.

3 Open-Source Components and XML Schemas for EHR systems

3.1 Introduction

Open source software, sometimes known as FOSS (Free and Open Source Software) is a paradigm of software development and licensing that enables different teams to collaborate on the development of software, offers transparency and peer review of code development, and includes a license that permits free or low cost terms for some or all kinds of use. Open source projects and software exists across all sectors of industry, including health care, to varying degrees of penetration. There are some health open source products that are setting-specific EHR systems, some that are middleware components supporting the capture, storage and communication of EHR data, and some that are tools that can be used to design knowledge artefacts used within an EHR deployment environment.

The quality assurance of open source software may be considered on three levels:

- processes inherent in the open source development processes themselves;
- quality assurance processes undertaken by each specific software development community;
- accreditation processes and quality labelling (including certification) of the resulting products that are identical to those applicable to commercial software.

This section of the report provides an overview of open source development practice, and summarises some example projects in the health sector. It considers each of these three levels of quality assurance, and discusses the potential roles for an external quality labelling approach to open source software.

3.2 Overview of open source software development

Open source software is not new nor radical, although sometimes stereotyped as such. It offers a different approach to the provision of information systems and design tools from conventional commercial software development. In summary, open source is a development method for software that utilises distributed peer review and transparency of process with the aims of producing software of better quality, higher reliability, more flexibility and lower cost than is produced by the traditional closed model, and an avoidance of vendor lock-in. The term "open source" is sometimes thought to be synonymous with "public domain." However, open source has a formal definition which must be reflected in the way in which the software is licensed (source: Open Source Initiative[8]):

> **1. Free Redistribution**
> The license shall not restrict any party from selling or giving away the software as a component of an aggregate software distribution containing programs from several different sources. The license shall not require a royalty or other fee for such sale.
> **2. Source Code**

---

[8] Please see: http://www.opensource.org/docs/osd

The program must include source code, and must allow distribution in source code as well as compiled form. Where some form of a product is not distributed with source code, there must be a well-publicized means of obtaining the source code for no more than a reasonable reproduction cost preferably, downloading via the Internet without charge. The source code must be the preferred form in which a programmer would modify the program. Deliberately obfuscated source code is not allowed. Intermediate forms such as the output of a preprocessor or translator are not allowed.

3. Derived Works
The license must allow modifications and derived works, and must allow them to be distributed under the same terms as the license of the original software.

4. Integrity of The Author's Source Code
The license may restrict source-code from being distributed in modified form only if the license allows the distribution of "patch files" with the source code for the purpose of modifying the program at build time. The license must explicitly permit distribution of software built from modified source code. The license may require derived works to carry a different name or version number from the original software.

5. No Discrimination Against Persons or Groups
The license must not discriminate against any person or group of persons.

6. No Discrimination Against Fields of Endeavor
The license must not restrict anyone from making use of the program in a specific field of endeavor. For example, it may not restrict the program from being used in a business, or from being used for genetic research.

7. Distribution of License
The rights attached to the program must apply to all to whom the program is redistributed without the need for execution of an additional license by those parties.

8. License Must Not Be Specific to a Product
The rights attached to the program must not depend on the program's being part of a particular software distribution. If the program is extracted from that distribution and used or distributed within the terms of the program's license, all parties to whom the program is redistributed should have the same rights as those that are granted in conjunction with the original software distribution.

9. License Must Not Restrict Other Software
The license must not place restrictions on other software that is distributed along with the licensed software. For example, the license must not insist that all other programs distributed on the same medium must be open-source software.

10. License Must Be Technology-Neutral
No provision of the license may be predicated on any individual technology or style of interface.

There is a variety of licenses that are used within the open source community, but they almost always involve these principles:

- Free Redistribution: There is no restriction on the distribution of the software in a royalty-free manner.
- Source Code: The source code must be included in the distribution or readily made available in some form.
- Derived Works: The license must allow for modifications and derived works to be distributed under the same terms as the original software.
- Integrity of the Author's Source Code: There must be some provision to protect the integrity of the original authors' code.

Examples of commonly used licences are listed below.

- The GNU General Public License (GPL) stipulates that anyone in possession of the software may use, modify, and redistribute it without restriction as long as the source code is made available with any distributions and the recipients of any distributions also use, modify, and redistribute the software under the terms of the GPL.

- The GNU Lesser General Public License (LGPL) is similar to the GPL, but, unlike the GPL, it allows LGPL software libraries to be linked to proprietary software systems without the latter becoming derivative works (and thereby no longer proprietary).

- The Mozilla Public License (MPL) allows source code and derivative works covered under its terms to be directly integrated into proprietary software systems (not just linked to such systems), although the license also requires that any modifications made to MPL software libraries be distributed with source code.

- Apache Software license, which allows for commercial products to be derived from an open source base without requiring giving back that software to the community.

There are a number of recognised variants of these two kinds of licence, each suitable for subtly different intentions.

Benefits of open source software
As reflected by the definition of open source, there are two important characteristics of open source software that underpin the approach.

First, the source code for the software is accessible:

- It is available for scrutiny, for peer review by the wider development community, providing a kind of quality assurance other than the black-box testing of software, which is the norm commercially.
- It can be held by any user site as a form of protection against the original team folding: the user is empowered to take over the software's evolution or to commission this from another source.
- Vendor lock-in is practically impossible.
- Other component developers can understand in detail how they might best interface with it or incorporate it into other components and systems.

Second, the compiled, executable software is usually available at no cost or at a nominal cost:

- Potential user sites can pilot the software quite extensively before making a final decision on its adoption.
- It can be adopted for a lower capital cost and probably maintained at a lower ongoing cost of ownership.
- The costs are based on recurring services rather than on recovering the full development costs from each client.
- Since open source often makes use of and interoperates well with other open source software, the cost benefits of the approach tend to propagate across the overall enterprise information technology solutions.

Most importantly, the open source software is often nurtured within an active community of end users in a wide range of settings. Clearly, if the user community providing input on the requirements and feedback on the components is extensive, the component will evolve functional breadth and adaptability, and more comprehensive testing, making it suited to a diversity of settings.

It is recognised internationally that good (that is, well-liked and well-used) clinical systems often arise through iterative cycles of development and clinical piloting in small-scale settings, and are often driven by locally-recognised needs and championed both by the clinical and by the development teams. Open source developments are ideally suited to supporting such iterative development processes.

However, there are some very large open source projects. For example, the U.S. Veterans Administration has historically adopted an open source approach to the systems used throughout its US hospitals and clinics (the software is known as VistA, and is discussed further below).

Support of interoperability
Open source software is frequently written initially within a small-scale and close working partnership of health care domain experts and systems developers. It is therefore very likely to be well tailored to the functional requirements of its specifying community. Because these usually are domain-specific specialised software components, they will have been developed on the assumption that they will in practice be used alongside other components that have not been developed by that group, and in multiple and diverse settings. Interoperability and standards compliance are critical to the success of the open source software.

In contrast, the traditional procurement of single supplier large-scale systems fails to encourage the adoption of interoperability standards:

- Large vendors do not need them, but can instead use private interfaces internally and leverage their size and position to dictate the interfaces they choose to offer to others.
- Standards compliance is sometimes claimed but in practice often partial, and rarely verified by the purchaser; DICOM conformance has proved to be a good example of this problem internationally.
- There is limited financial incentive for vendors to adopt standards, since the customer requirements for interoperability within a procurement specification are usually quite weak and are not key determinants within the selection process.
- The whole financial model leans toward vendor ownership of the software, and of the strategic direction of its evolution; the private finance initiative (PFI) approach adopted by several countries is perhaps an extreme example.

Open source projects behave in this regard much like European SME's: the support and adoption of standards is considered a fundamental business interest to enable mutual collaboration, and thereby compete with large vendor integrated systems.

3.3 Example open source projects in health care

This section summarises several key projects that focus on the development of tools and services to support electronic health record systems. It is by no means an exhaustive list: these examples are included to illustrate the breadth and diversity of these initiatives.

Open Health Tools (OHT)

Open Health Tools is an open source community with a vision to enable a ubiquitous ecosystem where members of the Health and IT professions can collaborate to build interoperable systems that enable patients and care providers to access vital and reliable medical information at the time and place it is needed. Its goal is to create a common health interoperability framework, exemplary tools and reference applications to support health information interoperability.

Open Health Tools aims to involve software product and service companies, medical equipment companies, health care providers, insurance companies, government health service agencies, and standards organisations.

OHT has based its governance, legal and intellectual property policies, development processes, marketing, and business models on the well-known Eclipse Foundation. Many of those working to create Open Health Tools were directly involved with creating Eclipse.

The Eclipse Open Health Framework (OHF) and other Eclipse software will be used as the basis for the OHT Framework. OHT will use the Eclipse Public License (EPL): a commercially-friendly license that allows organisations to include EPL-licensed software in their commercial products, while at the same time requiring those who modify derivative works of EPL code to contribute the modifications (but not the derivative works) back to the community.

From a governance perspective, OHT has four councils which are responsible for planning and managing the developments. The Clinical Council which is chaired by the Chief Clinical Officer and is the primary domain expertise resource. The Requirements Council which is co-chaired by the CCO and the Chief Technical Officer and determines set of requirements. The Architectural Council which is chaired by CTO and creates and manages high level architecture and road map of Open Health Technology. The Planning Council which is chaired by the CTO and is responsible for seeing that the road map is implemented.

Several projects have already been launched, of which a few examples are summarised below.

1. The OHT HL7 Tooling Project, led by the NHS (England), will provide second generation tools to support the HL7 version 3 message modelling methodology. The toolset will be based on the Eclipse Platform and form part of a wider suite of tools covering conformance/testing, clinical modelling and terminology maintenance.

2. The Open Health Tools SNOMED CT Tooling Project, led by the IHTSDO, will provide second generation tools which can be used to develop, maintain, promote and enable the uptake and correct use of SNOMED CT in health systems around the world.

3. A project led by Canada Health Infoway to support the implementation of standards through the provision of conformance testing of artefacts and tooling.

It is too early in the life of OHT, which was formally announced only in April 2008, to critique its intended governance and conformance testing processes, but given its links with the Eclipse Foundation that has a proven track record of supporting a rich pluralistic environment of inter-working components, it is likely that OHT will impose strict governance policies on its portfolio of projects.

The openEHR Foundation

The *open*EHR Foundation was formed in 2000 jointly by University College London and Ocean Informatics (Australia) as a not-for-profit company to consolidate and build on the two founding organisations' research and development in electronic health records, and to foster an international community of interest in progressing the design, implementation and evaluation experience of EHR systems in diverse clinical and health service settings. The company is registered in the UK with five directors, from UCL and Ocean Informatics. The Foundation has three boards to effect and govern its strategy: a Foundation Board, a Clinical Review Board and an Architecture Review Board.

The intellectual property vested in *open*EHR represents well over a hundred person years of effort, accumulated through around £50 million of accumulated research activity over fifteen years. The *open*EHR specifications build on a strong research pedigree in the requirements for EHR interoperability, the design of information architectures to meet these, and proof-of-concept implementations of EHR systems in small-scale demonstrator sites. The Foundation firmly believes that active engagement in implementation and clinical evaluation is essential to validate the integrity of the approaches it proposes, and to continue learning cycles to evolve the understanding of requirements and to refine specifications. The Foundation has published around 30 different specification documents, each evolving with strict version control, that collectively define international best practice in the design of various aspects of an EHR system. These specifications are now finding favour amongst other organisations: for example the *archetype* approach for specifying clinical data structures is now being used by the NHS in England, in Denmark, the Netherlands and Australia. It is being considered by a new international multi-agency forum known as the Detailed Clinical Models Group.

The *open*EHR architectural approach has also had a strong impact on international standards. Its core EHR Reference Model has, in a simplified form, been the basis of the ISO/EN 13606 Part 1 EHR communications model, the archetype approach has now also been standardised by CEN as EN 13606 Part 2. Its data types have influenced the draft International Standard for data types in ISO, and Foundation members are contributing to the new EHR Requirements draft standard in ISO.

Members of the Foundation are presently collaborating on open source reference implementations of these specifications:

- in Eiffel and Microsoft .net;
- in Java;
- in Python.

The implementation activities broadly cover two kinds of software: components that represent and manage EHR data instances (and so would form part of an EHR system), and tools used to manage knowledge artefacts (such as archetypes). Open issues, new requirements and matters arising from the implementation are considered through discussion lists, and the membership of *open*EHR now stands at approximately 800 members from nearly 80 countries,and its members include representatives from many large and small health software vendors, as well as senior figures in a number of health departments internationally.

Governance of *open*EHR specifications is managed by the Architecture Review Board and the Clinical Review Board. A more detailed discussion of the governance of archetypes is discussed elsewhere in this report. Its technical implementation projects are supported by discussion lists, and there are plans to develop a formal conformance testing procedure and tools, but this work has not yet commenced. However, there is a reference implementation that would permit the ready development of a conformance testing capability.

VistA
Probably the most successful and widespread open source effort in health care began in the U.S. Department of Veterans Affairs (VA) more than 20 years ago. The goal was to connect and automate common processes across its nationally distributed medical centres. Named the Veterans Health Information Systems and Technology Architecture (VistA), it is a rich, automated environment that supports day-to-day operations at Department of Veterans Affairs (VA) health care facilities, and versions of it have been adopted by the Indian Health Service and the Department of Defence. The software is in the public domain. VistA is used by over 90% of VA physicians, probably contains the largest volume of clinical patient information in a single open source system, and provides the VA with a rich resource to enhance the quality of care it provides and an evidence basis for strategic planning and research.

VistA includes links that allow commercial off-the-shelf software and products to be integrated. However, given its size and complexity it has not to date been ported in full and re-used to any other health setting, in the US or elsewhere. WorldVistA EHR is based on a public domain and smaller scale version of VistA. It has been piloted as a physician office system and achieved CCHIT certification as an ambulatory care EHR system in 2007.

OpenEMed
OpenEMed is a set of healthcare information service components built around the Object Management Group's distributed object specifications and HL7 (and other) data standards; it is written in Java for platform portability. It comprises a set of components that can be assembled and configured, alongside other components, to meet a variety of tasks rather than serve as a complete system in itself. OpenEMed includes reference implementations of OMG's Person Identification Service, Clinical Observation Access Service, Resource Access Decision, and Terminology Query Service.

OpenEMed started as an example of the US National Information Infrastructure in 1993 as part of the Sunrise project at Los Alamos, California, to demonstrate a common infrastructure that would support the use and value of distributed applications to a number of disciplines. The system is designed around services that were designed to be very flexible in the support of patient identification, terminology discovery, clinical observations, and distributed access control.

OpenEMed is the core architecture of the FIRST project at the City of Hope which is seeking to create a distributed collaborative system for managing clinical protocols. It also is a core component in an effort to provide a comprehensive integration strategy for bio-surveillance within the United States.

The software is distributed under an open source license from Berkley Software Design (BSD). Most participation beyond Los Alamos has occurred from Europe, most notably from the University of Maastricht. The project continues to invite interested participants both from the commercial and research sectors to participate.

openEMR
OpenEMR is a free medical practice management application. OpenEMR is licensed under the General GNU Public License (General GPL). The system offers clinical documentation, problem lists, medication lists, lab ordering, lab result reporting, and document management. The core EMR supports problem lists, allergies, medications management, immunisations, surgical and dental notes. The system also includes electronic billing, medical claims and accounts, appointments. Other functions are available through the purchase of commercial add-on modules, such as drug-interaction checking and decision support for coding. There is support for voice recognition software, electronic or scanned digital document management for records, and support for HL7 messages. OpenEMR has a large pool of developers from multiple companies participating in its maintenance and extension.

FreeMed
FreeMED has a combination of practice management and medical-record functionality. It supports clinical documentation, problem lists, lab reporting, document management, and some prescribing. Its other modules include: accounts receivable, administration, billing, claims management and a scheduling calendar. The system uses standardised data coding where possible, and provides options for data interfacing. Electronic ordering is not yet available, and limited decision-support capabilities are present.

The FreeMED project is managed by a non-profit organisation, FreeMED Software Foundation Inc, that is committed to preserving the open-source status of the software. The Foundation has both a Board of Directors and a Community Clinical Advisory Board.

ClearHealth
ClearHealth is a recently marketed open source product designed for physician offices and small to medium sized US health centres. Its core electronic medical record comprises a wide range of functions: health status alerts, patient dashboard, encounters, allergies, social history, problem list, self management goals, medical history, SOAP notes, drawings and images, clinical summary, CCR records, real-time lab results, electronic lab ordering, LOINC and SNOMED code support, patient pictures, barcode scanning and workflows, decision support, ID card support, speciality configurations, mobile device support, patient portal. Available speciality modules include: obstetrics and gynaecology, chiropractic, urology, oncology, home health, mental health. There is little readily available information on its governance processes or about the open source community who have underpinned its development. It is being marketed commercially on a contractual service and support basis even through the core product is FOSS.

Misys Open Source Solutions

This organisation, a division of Misys Healthcare, was formed in late 2007 to develop open source software solutions. Its main focus will be to deliver an open source infrastructure to connect electronic health information systems. Misys has chosen the Eclipse Foundation platform for development and deployment. It is an example of the less common scenario in which a commercial company with many proprietary health IT products has chosen to place part of its portfolio into an open source environment.

GNUmed

GNUmed is a German based open source project that has developed a multi-professional general practice system. It was developed by a small number of doctors and developers working internationally. From an EHR point of view it manages clinical documents, observation measurements, problems and health issues, medication, SOAP notes and the generation of correspondence. Administratively it manages patient registers, reports and audit analyses.

Tolven Healthcare Innovations

Tolven is a commercial firm that has developed a document repository that provides both PHR and EHR functionality through advanced clinical coding and a sophisticated knowledge base. It works closely with HL7 and is also investigating ways in which *open*EHR archetypes might inter-work with Tolven's clinical knowledge base.

IndivoHealth

This project is developing a Web-based personally controlled health record system that combines features of personal health records (PHRs) and Health Information Exchanges (HIEs). Indivo emerged from an academic setting and is still in its pilot phase with few active user sites.

openMRS

OpenMRS is an open source medical record system which is focused on developing countries. It is Web based, written in Java, and is an under active development. Most of the core developers are from the Regenstrief Institute and Partners in Health.

3.4 The role for certification of open source EHR systems and tools

Health IT has always been a complex and heterogeneous industry. The larger organisations, principally hospitals, have tended to be served by the larger vendors offering integrated (sometimes called monolithic) and relatively comprehensive solutions, and with limited need to support interface standards and limited co-operation with SME's. These large-scale systems have usually been procured by those managing the organisation, and as a consequence the current generation of hospital information systems are not well regarded for their support of clinical care except in data intensive areas such as radiology, laboratory and intensive care. (In contrast, SME's have done better at gaining penetration in niche areas such as the support of detailed clinical requirements for specific domains and disciplines.)

The open source movement in health care has been stimulated by many drivers over the past fifteen years, including: a way to meet the needs of small practices and clinical teams that were below the business radar of the large vendors; in recognition that health IT budgets are generally low and that Free and Open Source Software (FOSS) might be the only affordable option for many settings; as a practical way to innovate in exploratory areas of health IT by sharing costs (and pooling knowledge) across a community that would be prohibitive for one organisation to cover.

The growth in the EHR system marketplace has been accelerated by the many national e-Health programmes that have triggered significant new public investment in health ICT across the globe. These programmes have radically altered the health ICT landscape over the past three years, with an impact on the drivers outlined above. Health IT is no longer a low-cost "Cinderella" sector in many countries and indeed recruitment needs have outstripped workforce capacity in some. The need to join up regions and nations means that small primary care settings, which hold the most longitudinal and holistic patient information, are no longer below the business radar but are an integral part of the new programmes. Contracting arrangements have tended now to favour organisations that do not just deliver software but can support and maintain a large-scale deployment setting, which has always been the Achilles Heel of FOSS. As a consequence, the growth of open source EHR systems that was anticipated only a few years ago has not occurred. Small groups continue to survive but have limited penetration within large scale e-Health programmes. Mainstream vendors have largely not adopted an open source approach. Announcements are occasionally made about solutions becoming open source (such as Misys), but as yet we have not seen communities of developers growing around these in the original spirit of open source.

In contrast to this situation for EHR systems, for EHR resource and knowledge management tools, open source is gaining ground internationally, and has recently been strongly stimulated by the Open Health Tools (OHT) collaborative. As an example, the NHS in England intends that all new contracts for tools development will include a requirement for the adoption of the Eclipse Framework and publication via OHT. Admittedly the present set of completed tools is small, and for some time will be incomplete, but the trend is for open source to move into the tooling rather than live system direction.

Quality management of open source projects
The quality assurance of open source software may be considered on three levels:

- processes inherent in the open source development processes themselves;
- quality assurance processes undertaken by each specific software development community;
- accreditation processes and quality labelling (including certification) of the resulting products that are identical to those applicable to commercial software.

Open source projects receive code contributions from multiple and often widely distributed developers, and so need to have systems in place for code checking and an editorial policy that reviews candidate enhancements for their suitability and fitness for purpose. The code review process might be undertaken by the hosting team, or by peer review across the development community, but in any event is transparent and usually of good quality. This "micro-governance" is part of what keeps an open source project alive, and its community willing to

contribute. Editorial and strategic policies, "macro-governance", are not always transparent, and the various open source projects reviewed earlier varied in the extent to which these kinds of governance were documented. (It may be that such details are only available to members, and so the absence of public macro-governance information cannot be taken to mean that this is not well undertaken.) However, it must be recognised that neither kinds of governance are usually made transparent in the development of proprietary software.

For software development, there are established quality processes that are defined by the software industry, and Integrated Development Environments (IDEs) that help to enforce those. Similarly, there are established practices for software safety testing. More recently, health informatics standards have been defined that specify how health software should be risk assessed and quality assured [for example, ISO/TR 29322:2007]. It would be reasonable for EuroRec to develop quality criteria that require adherence to such standards, but not to target only the open source community for such criteria: when appropriate they should apply to all EHR systems.

There are no absolute quality criteria for the editorial and strategic governance of open source software. It might be argued that these are evaluated by the market through the success of a product to meet user needs.

By the very nature of the spirit of open source, projects often have active user communities as well as developer communities. These users often provide the input requirements for a product's evolution, and offer critique of the quality and functionality of the various software releases. In this regard they are no different to any other system user community or user group, and most health IT vendors support such groups to ensure product relevance and to nurture reference sites. It would be reasonable for EuroRec to consider future quality criteria around the support of user engagement, the influence of contributions that are made by users individually or via a user group, and the availability of reference sites for scrutiny as a complement to classical *in vitro* software testing as undertaken, for example, by CCHIT.

Most importantly, open source software must deliver to the same functional and safety quality standards as proprietary software. In other words, any product be it open or closed source needs to meet the quality labelling criteria that have been collated and systematised by EuroRec in its repository, according to profiles defined by purchasers. There is no case for adopting a special or different approach to the quality labelling or certification of open source software.
It had at one point been considered feasible for EuroRec to maintain an inventory of open source software. (A repository was not considered since this would require open source projects to donate a copy of the software to the EuroRec site, with consequent copyright and version management implications, and no clear advantage.) However, a list of open source projects would be volatile, and include some organisations that do make a profit from support services etc. (rather like as occurs with Linux). It might therefore be considered inequitable for EuroRec to list open source projects and not, for example, academic closed source projects or commercial products. It is therefore not recommended that this option be pursued at present.

3.5 XML Schemas

XML Schemas are used throughout industry to represent data formats used within components and products and for interchange between them. They are only one

technical way of representing a logical information model, but one that is gaining popularity as a generic Implementable Technical Specification (ITS).

Since an XML Schema is simply a representation format, there is nothing inherently open or closed about it except through any copyright and/or usage rights associated with it. Interest in XML Schemas for interoperability has grown as a format by which interoperability standards may be represented. HL7 is a standards body that has published an XML Schema for each of its logical (information model) standards, and its developer communities use this as a way of validating the model. Both CEN TC/251 and ISO TC/215 (Health Informatics) have produced a small number of these, and plan to extent the practice of developing and supporting an XML ITS for relevant standards. However, for all of these bodies, the XML is tightly coupled to the published standard (sometimes included in the same document) and therefore subject to the same copyright restrictions i.e. it is not public domain. Since conformance testing to industry standards (such as DICOM) is a commercial activity, XML Schemas to these are also usually restricted.

Some national e-Health programmes have developed libraries of XML Schemas to define their messages for communication, for example the NHS in England. These are restricted to the contracted vendor community (and in England's case to members of HL7 UK who provide guidance to the NHS on them).

At this stage, therefore, the majority of XML Schemas of greatest interest to the EHR and EHR system development community are either copyright protected and need to be purchased or are issued formally to organisations if they are part of a national e-Health programme. There does not seem to a case for EuroRec to seek to maintain an inventory of these as part of its EHR system quality labelling and certification programme. EuroRec and its customer base of EHR system purchasers may, however, require that conformance to one or more particular XML Schemas has been demonstrated as part of the quality assessment of individual EHR systems.

3.6 Conclusion

The marketplace for open source EHR systems is modest and its potential growth appears for the time being to have been largely overshadowed by the high-budget and contractually-driven e-Health programmes. Whilst open source projects have several reasons for having good quality processes, it is not equitable for open source projects to be assessed on these when commercial closed source products are not. However, an insistence on demonstrated software quality processes including testing might become part of future quality labelling criteria for EHR systems. Open source EHR systems should be expected to demonstrate conformance to the same kinds of quality labelling criteria as closed source systems.

There is no present case for EuroRec to maintain an inventory of open source systems. However, EuroRec may in the future wish to extend its present set of quality criteria to include a requirement for the adoption of industry good practice in software development and safety testing. It may also consider future EHR system vendor requirements on the engagement of users, the influence of user groups, and the availability of reference sites.

3.7 Further reading

Kalra D, Forslund D. Open Source Health Systems. In Demetriades JE, Kolodner RM, Christoperson GA (eds.). Person-Centred Health Records: Toward Health-ePeople, 2005, p169-185. Springer, NewYork, USA. ISBN 0-387-23282-6

Kalra D, Beale T, Heard S. The openEHR Foundation. In Saranummi N, Piggott D, Katehakis DG, Tsiknakis M, Bernstein K (Eds). Regional Health Economies and ICT Services: The PICNIC Experience. Stud Health Technol Inform. 2005; 115: 153-173. IOS Press. ISBN:1-58603-538-X PMID: 16160223. Open-Source EHR Systems for Ambulatory Care: A Market Assessment. Sujansky & Associates

California HealthCare Foundation. January 2008. Available from: http://www.chcf.org/topics/view.cfm?itemID=133551

ISO/TR 29322:2007. Guidance on the management of clinical risk relating to the deployment and use of health software systems. ISO, Geneva, 2007.

Specific open source projects:
ClearHealth: http://www.clear-health.com

IndivoHealth: http://www.indivohealth.org

FreeMed: http://www.freemedsoftware.org

GNUmed: http://www.gnumed.org

openEHR Foundation: http://www.openehr.org

OpeneMR: http://www.openemr.net

Open Health Tools: http://www.openhealthtools.org

Tolven Healthcare Innovations: http://www.tolven.org

WorldVistA eHR: http://www.worldvista.org/World_VistA_eHR

4   Legislative and industry standards

A significant body of standards published by CEN, ISO and HL7 have been listed and summarised in Q-REC Deliverable 3.1. This set of resources has been published as a searchable resource by EuroRec:
(available at http://www.eurorec.org/standards/standards.cfm?actief=services).

The standards listed and described within this inventory are those considered to be of most relevance to the developers, purchasers and users of EHR systems. (Some of them are referred to in the EuroRec inventory of EHR requirements statements.)

Among the reasons for applying standards (formal or informal) to EHR systems (EHR-Ss), two of them, related to interoperability issues, are outstanding:
–   the need to communicate between systems in such a way that data originating from any EHR-S should be usable straight away by any receiving systems;
–   the need for durability and long-term availability and processability —possibly lifelong and beyond— of personal health data, primarily for the sake of continuity of personal health care, secondarily for public health and other similar collective purposes.

A significant body of standards published by CEN, ISO and HL7 have been listed and summarised in Q-REC Deliverable 3.1. This set of resources has been published as a searchable resource by EuroRec, available at:
http://www.eurorec.org/standards/standards.cfm?actief=services

The standards listed and described within this inventory are those considered to be of most relevance to the developers, purchasers and users of EHR-Ss. (Some of them are referred to in the EuroRec inventory of EHR requirements statements.)

Standards are evolving along time. New topics, related to particular facets of the use of EHR-Ss lead to the need to devise new specifications. Existing standards are reviewed, based on implementation experience or on the appearance of new requirements, some arising from the need to harmonise with other newer standards. Therefore their scope may evolve, or more simply some of their provisions must be adapted and amended.

A list of standards such as the one set up by Q-REC, is bound to be frequently updated, both to include new references and to track the amendments made to the ones of existing documents. This ongoing activity will, beyond the termination of the Q-REC project, need a long-lasting flow of resources to be allocated to it.

5 Coding schemes and terminology systems

The EuroRec institute offers a web-based mechanism whereby the developers and publishers of health terminology systems and coding schemes can formally register in compliance with European standard EN 1068:2005 "Health Informatics - Registration of Coding Systems". The European Institute for Health Records has been designated by CEN as the Registration Authority for EN 1068:2005, with the task of implementing that comprehensive register of health coding systems used in Europe, and making it available to all those parties who may benefit from the information it contains.

Through the work of the Q-REC project this register and inventory has been enhanced and enriched, and is available at:
http://www.eurorec.org/services/inventories/coding/index.cfm?actief=services

This site may be used for the registration of new coding schemes, updating an existing registration and to review the existing inventory of registered schemes. The web-based inventory is documented Q-REC Deliverable 3.1.

The EuroRec institute offers a web-based mechanism whereby the developers and publishers of health coding systems and associated terminological systems can formally apply for registration in compliance with the European standard EN 1068:2005 "Health Informatics - Registration of Coding Systems".

The European Institute for Health Records has been designated by the CEN Management Centre as the Registration Authority for EN 1068:2005, with the task of implementing that comprehensive register of health coding systems used in Europe, and making it available to all those parties who may benefit from the information it contains.

There are many coding systems in use in health, and it will most probably remains so for yet a very long time, since the worldwide adoption of a unique terminological system for each particular field of clinical information, for example:
- symptoms, findings, diagnoses, procedures, which would only need to be translated in any language, looks impracticable for an unpredictable time.
- the use of coding systems will keep inevitable for an unpredictable time.

The only current alternative is the use of coding systems to represent the informative content both unambiguously and in a manner that allows it to be easily handled and processed by computers. But the communication of coded data raises an immediate practical issue: any receiver of a code value needs to relate it to its corresponding code meaning. For this she or he needs to identify unambiguously the coding system that has been used by the sender.

The goal of the EuroRec register of health coding systems is to provide health coding systems users with a list of those systems, not to have direct access to the full tables of code values and code meanings each of these is made of. The task of implementing the European register of Health Coding Systems has been performed as part of the Q-REC project.

The register site, available at
http://www.eurorec.org/services/inventories/coding/index.cfm?actief=services
may be used:
- for the initial application to register a coding system, and to submit the update of a previously registered one
- as well as, of course, to access the content of the register.

Beyond the primary implementation of the register, there is an ongoing need to:
- identify the Responsible Organisations for health coding systems and suggest their application to register their product with EuroRec;
- monitor, usually on an annual basis, the revisions of those coding systems that had been registered or track their new versions, and, again, suggest their Responsible Organisations to submit these updates.

Whatever its business model, this ongoing activity will, beyond the termination of the Q-REC project, need a long-lasting flow of resources, to be allocated to it.

The web-based inventory is documented Q-REC Deliverable 3.1.