

# City Research Online

## City, University of London Institutional Repository

**Citation**: O'Sullivan, D. and Plastiras, P. (2017). Combining ontologies and open standards to derive a middle layer information model for interoperability of personal and electronic health records. Journal of Medical Systems, 41(195), doi: 10.1007/s10916-017-0838-9

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: http://openaccess.city.ac.uk/18466/

**Link to published version**: http://dx.doi.org/10.1007/s10916-017-0838-9

**Copyright and reuse:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

City Research Online:

http://openaccess.city.ac.uk/

publications@city.ac.uk

# Title Page

### Title

Combining Ontologies and Open Standards to Derive a Middle Layer Information Model for Interoperability of Personal and Electronic Health Records

#### **Abstract**

**Objectives**: To enable better interoperability between Personal Health Record (PHR) and Electronic Health Record (EHR) systems to allow exchange of data from patients to providers and vice versa in order to encourage PHR use and patient self-management.

**Methods**: A non-binding middleware based on open technologies and standards that resides between a PHR and EHR system has been developed. Specifically, the middleware consists of an ontology-driven information model based on the HL7 Reference Information Model (RIM) and a set of transformation rules that work in conjunction with the information model to process data exported from a PHR or EHR system and prepare it according to constraints imposed by the receiving system.

**Results**: The information model was evaluated by executing a set of use case scenarios containing data exported from a PHR system, transformed according to the transformation rules, transferred to an EHR system and vice versa (EHR to PHR). This allowed various challenges to emerge as well as revealed gaps in current standards in use.

**Conclusions**: The proposed middleware information model offers a number of advantages. When modifications are made to either a PHR or EHR system, they can be incorporated by altering only the instantiation of the information model. The model uses classes and attributes based on HL7 RIM to define how data is captured which allows greater flexibility in how data can be manipulated by receiving systems. The solution is applicable to existing PHR systems, or could be used as a blueprint to develop new PHR applications.

**Keywords**: Medical Records Systems, Computerized; Personal Health Record: Electronic Health Record; Health Information Exchange; Interoperability.

## Background and Significance

PHR systems capture health data entered by individuals and provide information related to the care of those individuals. By contrast EHR systems serve the information needs of health care professionals and are managed by those professionals and/or a healthcare institution. The value of PHR systems is increased greatly when they are integrated with EHR systems so that health information about an individual can flow seamlessly among systems used by health professionals and the patient [1]. Public interest and the current availability of PHR systems is high [2] and although the evidence about the effect of PHR systems on health outcomes for patients is limited [3, 4], a majority of people see benefits of accessing their own data for health management and improved communication with healthcare providers [5]. From an organizational point of view, robust PHR systems allow greater knowledge about the patient's lifestyle, allow access to data in unusual circumstances such as emergencies [5], and open up new opportunities for telemedicine and home monitoring initiatives [6, 7]. However, in spite of this, adoption of PHR systems remains relatively low with estimates that between 7 and 10% of Americans use a PHR [8-10].

There are two prevailing models of PHR systems - "untethered" standalone systems which are entirely under the control of the patient who must enter their own information or arrange for it to be transferred from another system; and "tethered" systems, often referred to as patient portals which are sponsored by an organization and where the record is automatically populated without the patient needing to enter information [11]. The majority of existing and emerging PHR systems are untethered despite the growing trend of sponsored patient portals [12]; therefore, the success of these systems is determined by a person's willingness to maintain their PHR information or on their health provider's willingness to share and transfer data from an EHR. The tethered model places fewer burdens on the patient and has shown the highest adoption rates, for example, 5.2 million of 9 million members of Kaiser Permanente had registered to use My Health Manager [13] as of the third quarter of 2015, and approximately one-fifth of veterans' report using the US Department of Veterans Affairs PHR, My HealtheVet [14]. If EHR and PHR systems have not been originally designed to follow the tethered model, the adaption of these systems to allow sharing of institutional and personal data usually presents challenges for healthcare providers who must 'retrofit' proprietary EHR systems for purposes they were not originally intended, a costly and time consuming process [8, 15-17]. As a result, many tethered PHR systems focus on providing simpler data to patients, for example, hospital visits or prescription drugs dispensed in a read only format [11], rather than clinical data which is more useful if patients are to be encouraged to self-manage in a meaningful way.

According to reviews on PHR adoption barriers, the most significant factors for non-adoption are privacy and security, usability, and interoperability and integration [5, 8, 9, 15, 18-21]. User attitudes to PHR security and privacy are complicated. Some e.g. [8] report it as the most significant barrier deterring users whereas others e.g. [19], have found that privacy and security are no longer pertinent concerns given better familiarity among consumers with online services such as e-banking. Poor usability has arisen because of a lack of understanding and poor assessment of lay users' workflow as well as their health literacy, which has led to the development of systems which are challenging to use and content that is difficult to comprehend [20].

In this paper we focus on the interoperability and integration barrier. Some of the challenges to interoperable PHR are the fact that ideal information in a PHR is lifelong and cross institutional, and thus covers data from various providers but also citizens own entries as well as emerging data from sensors and integrated devices. Further, citizens are the primary managers of the data but can also possibly want to share it with others – family, employers, care givers and healthcare providers [22].

Another complicating factor is that dispersed care is the general rule and coordinated care the exception. Compounding this is the increase in frequency of co-morbid chronic conditions, further dispersing patient care across multiple providers and EHR systems [23]. This poses major challenges due to the resultant provider-to provider misunderstandings, information gaps, and identification of responsibility.

Slow adoption of standards for interoperability has delayed the rollout of PHR systems. In particular, standards for data interchange, messaging, content encoding, 'lay' representation of encoded data have not been widely adopted [1]. The two most commonly used standards for PHR are the Continuity of Care Record (CCR) [24] and the Continuity of Care Document (CCD) [25]. CCR was developed by ASTM International and CCD is a joint effort by HL7 and ASTM to link CCR with HL7's Clinical Documentation Architecture (CDA) [26] which underpins EHR systems. The CCR is a core data set of the most relevant administrative, demographic, and clinical information facts about a patient's healthcare, covering one or more healthcare encounters and its primary use case is to provide a snapshot of pertinent clinical, demographic, and administrative data. To ensure interchangeability of CCRs, the specification specifies an eXtensible Mark-up Language (XML) coding that is required when the CCR is created. The CCD represents a complete implementation of CCR, combining HL7 technologies with the CCR's clinical data representation. CCD is an XML-based standard that specifies the structure and encoding of a patient summary clinical document. However real-world implementation of these standards is still at very early stages [27].

Previous work on PHR systems has tackled interoperability and integration issues in various ways. Ved et al [28] and Ming et al [29] propose a cloud-based architecture allowing different PHR systems to access patient data from various EHR systems where the emphasis is on security, scalability, and efficiency. Others have focused on the use of open standards to develop interoperable PHR systems. For example, Xie et al [30] use XML at every level of development from data storage as XML documents to rendering final data to the screen, while Li Hui et al [31] use Cross-Enterprise Document Sharing from Integrating the Healthcare Enterprise (IHE-XDS) and the Web Ontology Language (OWL) to create a shareable and integrated PHR system. Puustjärvi and Puustjärvi [32], also use an ontology and open standards to develop a PHR system interoperable with relational databases. Lähteenmäki et al [33] attempt to integrate non-clinical information using an architecture based on HL7 CDA messages and Simple Object Access Protocol (SOAP). Yuksel and Dogac [34] and Brut et al [35] customized the HL7 RIM [36] in order to support data acquired from other sources. By mapping to HL7 RIM, Yuksel and Dogac derived a Refined Message Information Model (RMIM) which allows transforming data to HL7 CDA which could subsequently be used by both EHRs and PHRs. Brut et al created an alteration of RIM utilizing classes to generate a domain specific RMIM. Both solutions refined the RIM to produce messages, whereas our work focuses on the development of a new information model that allows for the exchange of full medical documents.

## Objectives

The aim of our work is to enable better interoperability between PHR and EHR systems to allow meaningful exchange of clinical data from providers to patients and vice versa. We have developed an ontology-driven information model based on previous analysis on different PHRs [37], which stands as a middleware layer between PHR and EHR systems and a set of transformation rules for effectively transferring data in a standardized way between PHR and EHR systems. The middleware solution is equally applicable to tethered and untethered PHR systems that have basic import/export

capabilities as it abstracts away from the specific PHR and EHR system. We demonstrate our methods with a number of case study scenarios.

#### Methods

The process of designing the information model involved a number of discrete stages. We began by analysing different PHRs in order to identify common entities and used this analysis to develop a new ontology-based representation for PHR data and their relations. Classes from our information model were then mapped to relevant HL7 RIM classes. Finally, we implemented a set of transformation rules for PHR-EHR and EHR-PHR exchange using a custom PHP transformation engine. The transformation of data to a syntactically correct format was aided by predefined XML templates.

#### High Level Architectural Model

The architectural model consists of four structural components. The first component is the conceptual layer, namely the definition of each concept to be used and how concepts are related. The second component is the semantic layer which provides a formal representation of the entities defined in the conceptual layer. The third component is the syntactic layer which aims to ensure that represented data will fully conform to pre-defined structure and syntactic rules. Finally, the data layer manipulates data exported from either a PHR or EHR system in an appropriate way. Figure 1 shows how data exported from a PHR system will be transferred to an EHR system. The ontology instantiation reflects the data and associated attributes a given PHR (e.g. Microsoft Health Vault) can hold. Moreover, the ontology is referenced from both the conceptual layer and the semantic layer to ensure the exported data are understood correctly and transformed to the appropriate format ready for import. The transformation rule engine implements the transformation rules to manipulate the data and lies between the data, the semantic and the syntactic layer. Finally, the exported HL7 document produced after the transformation is represented in the data layer.

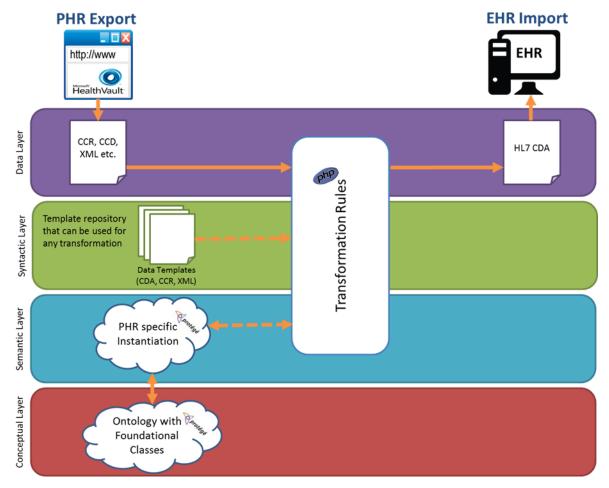


Figure 1 - High Level Architectural Model

#### Representing PHR Data and Relationships

Previously we have analysed data structures and functionality of six different PHRs (Health Vault, Telemedica, NoMoreClipboard, Health Spek and Health Companion) in order to identify similarities and differences and to produce a new high level information model to cover all existing data [000337]. The resulting information model can be seen in Figure 2.

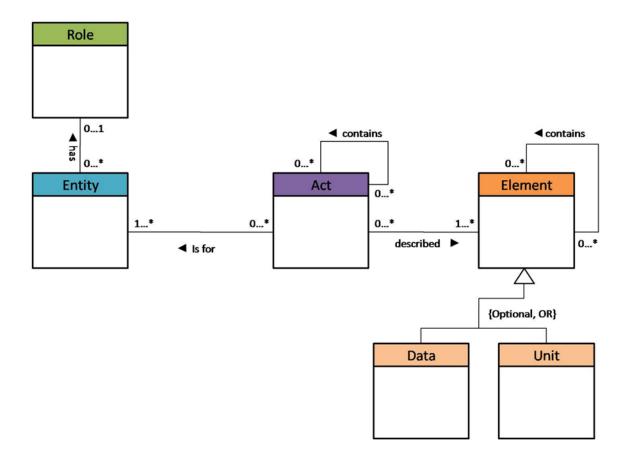


Figure 2 - Information model - baseline expressed in UML notation

Using HL7 RIM foundation classes as a basis, our information model consists of four basic classes: Role, Entity, Act and Element (later we describe in detail how classes and their relations are different in our information model and in RIM). In our information model classes are defined as follows:

- Role: participants to one's personal health record
- Entity: roles are played by an Entity
- Act: any event in a PHR, e.g. 'monitoring' vital signs or 'recording' an allergy,
- Element: any data corresponding to Acts, e.g. an allergy will be an element itself and then it can contain some sub-elements such as allergy name, effective date etc.

Moreover, the class Element has two subclasses named "Data" and "Unit" to manipulate represented data. These subclasses characterize data input or saved by a user as part of an act. An Element may contain many other Elements of type Data or Unit and hence it can be used as a container. Sub-classing data into its constituent elements allows for finer-grained representation of patient data thus allowing the information model to capture variations among data stored by various PHR systems, as well as to adequately capture the greater amounts of data and data types stored by PHR systems when compared with EHR systems. The above classes can be considered a baseline; both classes and attributes can be expanded in order to cover the requirements of a specific PHR system. Use cases can be used to demonstrate relevant concepts and relations of the information model. Figure 3 demonstrates a patient use case for monitoring vital signs and describes a person (Entity) who is the patient (Role), monitoring (Act) their vital signs (Element). Monitor is a composite act that involves the measurement of different elements. Moreover, a simple element

may consist of data elements or/and unit elements. Figure 4 shows how Act and Element classes may be instantiated to accommodate such a case.

#### Patient monitors vitals.

Patient measures Height/Weight/BMI
Patient measures Blood Pressure

Patient measures Systolic Blood Pressure
Patient measures Diastolic Blood Pressure

Patient measures Glucose Monitoring
Patient measures Triglycerides Monitoring
Patient measures Cholesterol Monitoring

Figure 3 - Use case scenario to monitor vital signs

#### **Blood Pressure**

Diastolic BP (value, unit) Systolic BP (value, unit)

Simple Element: Blood Pressure

Simple Element: Diastolic BP

Value Element: Diastolic BP (i.e. 120)

Unit Element: BP (i.e. mmHg)

Simple Element: Systolic BP

Value Element: Systolic BP (i.e. 80)

Unit Element: BP (i.e. mmHg)

Figure 4 - Use case scenario to capture data and unit elements

The use of each of the outlined classes and their relations differs from RIM in the following ways. In RIM the class Role is related to the class Act through another class named Participation, and to the class Entity. In our information model, class Role is related to class Entity and the latter is then related directly to class Act. This is because roles in PHR systems (e.g. doctor, patient, care given, relative etc.) are more limited than in EHR systems (e.g. access role, employee, licensed entity etc.) and thus entities participate directly in acts. Further, in contrast with RIM, the attributes of each class and the classes themselves are flexible so that the four classes can accommodate all relevant information from PHR systems. This flexibility can overcome the reported recently RIM shortcoming of limited status codes (i.e. for care plans) [38], lack of basic clinical information (i.e. normal range for lab results) [38], incapability to cover specialized domains [39] and conceptual misunderstanding of capturing information about an action versus the action itself [39]. An example of a possible set of attributes for each class is shown in Figure 5. The information model has been instantiated using Protégé [40] and OWL [41].

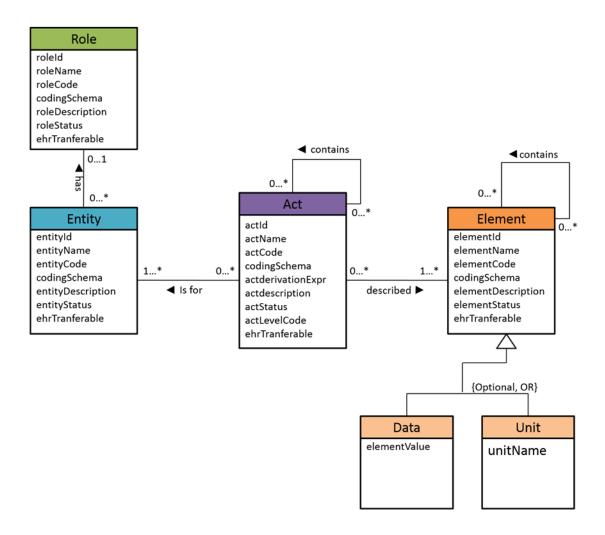


Figure 5 - Information model - instantiated with a possible set of class attributes

Despite other emerging standards it has been decided to map our information model to RIM classes as this allows the production of a CDA document. In addition HL7 v3 messages ordocuments can be transformed to other standards such as HL7 v2 if necessary. This decision reduces complexity towards achieving integration at an early stage given the current most commonly used standards for PHRs are CDA-based. Then by altering the transformation mapping rules (described later), other standards such as OpenEHR [42] and FHIR [43] can be also be incorporated.

#### Mapping Classes from the Information Model to HL7 RIM Relevant Classes

Entity and Role classes from our information model can be matched to the respective RIM classes, however Act as it has been defined in the new information model cannot be directly mapped either to the Act class of RIM or to any of Act's subclasses. The solution is to map the proposed Act class to a combination of the classes Act, Participation and Act Relation available in RIM. The latter two classes are used in RIM to complement the usage of the initial class Act and the way it interacts with the class Role. The class Act Relationship expresses a direct association between a source and target act. The class Participation is an association between an act and a role with an entity which plays that role. The Element class from the information model can be replicated utilizing the Observation class from RIM which is a sub-class of Act. An observation often involves measurements or other methods of investigation. In addition, an observation may also be a simple statement. Thus, an

Observation class has also been used to replicate the subclasses Unit and Data Element (subclasses of the Element class).

The most important class used in the information model is Act. It can be used to capture varied information depending on personal preferences in contrast with the use of the Act class in RIM. In order to map the information model, the structure shown in Figure 6 has been used.

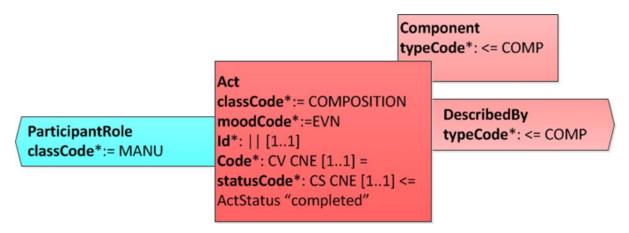


Figure 6 - Act Class Mapped to RIM Notation

At the centre sits a RIM Act class with a classCode COMPOSITION, namely grouped patient information represented in an EHR. The Act class has a 'one-to-many' connection to a Participation class named 'ParticipantRole' and a 'one-to-many' connection to an Act Relationship class named 'DescribedBy'. In order for the recursive relationship that has been used in the information model to be replicated, a different Act Relationship class named 'Component' has been used. The Act class is connected to a role via a ParticipantRole. A role is always related to an entity. Both Role and Entity classes have been used with no alterations and they can encapsulate any classCode available in the RIM. The Act class from RIM has been reused to represent the Element class as shown in Figure 7.

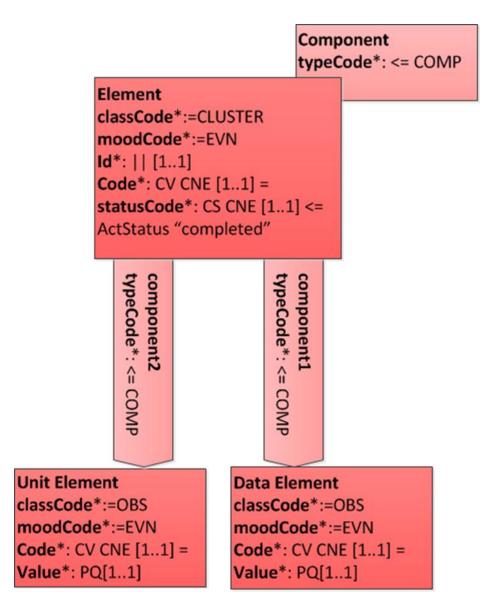


Figure 7 - Class Element with subclasses and recursive relationship in RIM Notation

The Observation class has been renamed Element with a 'classCode' of CLUSTER. A cluster is used to describe a group of entries within a composition, topic or category that have a logical association with one another. This class contains two more Observation classes for Unit Elements and Data Elements respectively. Both these classes are connected to the Element class utilizing two different Act Relationship classes named 'Component1' and 'Component2' respectively. As these two classes are simple Observations, the 'classCode' for both is OBS. In the case of class Element, the recursive relationship from our information model has been replicated the same way as described in Act class above, namely by using another Act Participation class. The final information model can be seen in Figure 8.

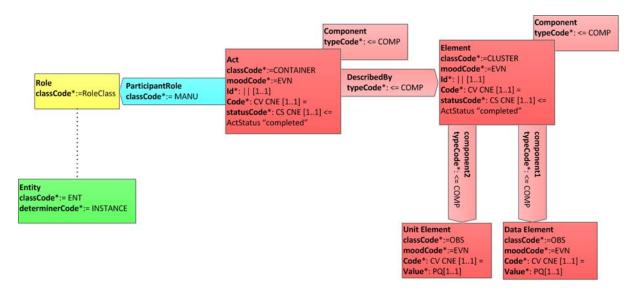


Figure 8 - Information model reconstructed utilizing RIM classes

#### Rules for transforming PHR-EHR and EHR-PHR data

In the previous section we described a high level mapping of the information model to customized RIM classes. In practice such smooth mapping can rarely be achieved due to the nuances and complexity of data from PHR systems and initial data exported from either a PHR or EHR system needs to be prepared for transfer. To accommodate this preparation, we have defined a set of transformation rules. The rules have been implemented in PHP; however, they may be implemented in any programming language of choice.

Rule 1 – Identify participating roles: Identify the entities and their roles that may participate in an act. In particular, the demographic characteristics for the patient from an exported document need to be identified. Accordingly, other participant entities such as author, custodian and legal authenticator of the document may also be included on the header and thus need to be identified. The location of the participaantrole can in the exported data is different depending on the export file. For instance, if the exported file is CCD then the information will be in the header. In case of custom export formats (e.g. custom XML schema) the information will be extracted from a different location.

<u>Rule 2 – Identify elements that have been exported</u>: During this step the exported file is parsed to identify high level elements (i.e. data items) that have been exported. Such elements might be allergies, medications, past medical history etc. With current PHR systems it is not common for the user to have a choice of what to export, however as systems evolve this may be desired.

Rule 3 – Identify the expected attributes for the elements highlighted from rule 2: This rule parses the PHR specific ontology instantiation (i.e. the OWL file), to identify the expected attributes that need to be identified in the exported document. In case of composite elements, after each main element has been identified, the algorithm will search for any sub-elements before it moves to the next main element. For instance, for the element 'Allergy', the sub-elements 'Allergy Name', 'First Observed', 'Reaction' and 'Treatment' may be identified.

Rule 4 – Collect the expected attributes: During this step the exported file is parsed in order to extract the expected elements and sub-elements which were identified during the previous step. For example, considering the example from the previous step we will need to extract the actual allergy name (e.g.. Pollen).

Rule 5 – Decide how to process each value: Establish how each of the identified values from the previous step (rule 4) should be manipulated. For example, not all the information in a PHR always needs to be transferred into an EHR and therefore assigning data properties to elements (e.g. EHR\_Transferrable or NOT\_EHR\_Transferrable) allows constraining the data to be transferred. In a similar manner, attributes can be assigned to indicate special treatment of a specific piece of data. For instance, if a piece of information should be mapped to the unstructured body of the final CDA document this could also be specified with a relevant attribute.

Rule 6 – Apply specific transformation to the extracted data: Transform the extracted data according to specific syntax rules of the receiving system (e.g. CDA). Transformation is achieved using predefined XML templates which are sets of mark-up tags representing data for different scenarios. For instance, an allergy element represented in a PHR system as shown in Figure 9 will have a corresponding template as shown in Figure 10. For transformation, template data such as [-AllergyName-] should to be replaced with extracted data values.

```
<ObjectPropertyAssertion>
   <ObjectProperty IRI="#contains"/>
    <NamedIndividual IRI="#Allergy"/>
    <NamedIndividual IRI="#Allergy First Observed"/>
</ObjectPropertyAssertion>
<ObjectPropertvAssertion>
   <ObjectProperty IRI="#contains"/>
    <NamedIndividual IRI="#Allergy"/>
    <NamedIndividual IRI="#Allergy_Type"/>
</ObjectPropertyAssertion>
<ObjectPropertvAssertion>
   <ObjectProperty IRI="#contains"/>
    <NamedIndividual IRI="#Allergy"/>
    <NamedIndividual IRI="#Allergy_Reaction"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
   <ObjectProperty IRI="#contains"/>
    <NamedIndividual IRI="#Allergy"/>
    <NamedIndividual IRI="#Allergy_Notes"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
    <ObjectProperty IRI="#contains"/>
    <NamedIndividual IRI="#Allergy"/>
    <NamedIndividual IRI="#Name"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
    <ObjectProperty IRI="#contains"/>
    <NamedIndividual IRI="#Allergy"/>
    <NamedIndividual IRI="#Notes"/>
</ObjectPropertyAssertion>
```

Figure 9 - Microsoft Health Vault ontology showing object property allocations for element 'allergy'

```
<act classCode="ACT" moodCode="EVN">
    <code code="10153-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="Allergy"/>
    <statusCode code="completed"/>
    <effectiveTime value="1234"/>
    centryRelationship typeCode="REFR">
    <observation classCode="OBS" moodCode="EVN";</pre>
        <code code="10153-2" codeSvstem="2.16.840.1.113883.6.1" codeSvstemName="LOINC" displayName="Allergy" />
        <value xsi:type="CE" code="55561003" displayName="[-AllergyName-]" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT" />
    </entryRelationship>
    <entryRelationship typeCode="REFR">
    <observation classCode="OBS" moodCode="EVN";</pre>
        <templateId root="2.16.840.1.113883.10.20.1.47" />
           de code="10153-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="Reaction" />
        <statusCode code="completed" />
        <value xsi:type="CB" code="55561003" displayName="[-Reaction-]" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT" />
    </observation>
    </entryRelationship)</pre>
    <entryRelationship typeCode="REFR">
        <templateId root="2.16.840.1.113883.10.20.1.47" />
        <code code="10153-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="Type" />
<statusCode code="completed" />
        <value xsi:type="CE" code="55561003" displayName="[-AllergyType-]" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT" />
    </entryRelationship>
    <observation classCode="OBS" moodCode="EVN">
        <templateId root="2.16.840.1.113883.10.20.1.47" />
        <code code="10153-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="First Observed" />
        <statusCode code="completed" />
        <value xsi:type="CE" code="55561003" displayName="[-FirstObserved-]" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT" />
    <entryRelationship typeCode="REFR">
       <templateId root="2.16.840.1.113883.10.20.1.47" />
              code="10153-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="Notes" />
        <statusCode code="completed" />
        <value xsi:type="CB" code="55561003" displayName="[-Notes-]" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT" />
    </observation>
    </entryRelationship>
```

Figure 10 - Pre-defined XML template for transforming element 'allergy'

Such templates can be constructed at various levels of granularity depending on the required flexibility and variability between systems. Moreover, the HL7 Interoperability Toolkit [44] contains similar predefined templates for integration between EHR systems. By restructuring and repeating these templates, appropriately formatted documents can be produced.

<u>Rule 7 – Attach the header to the exported document</u>: Use the data identified by rule 1 and place it in the final document as a header. This will be facilitated using data templates in a similar manner as in the previous step.

#### Results

Our proposition has been evaluated by executing two sample scenarios in which selected information exported from a PHR in CCD format wastransformed to a CDA document according to the rules outlined in the previous section and accordingly integrated into an her and in which information exported from an EHR in CDA format was transformed to CCD and integrated into a PHR system. Microsoft Health Vault [45] has been selected as the PHR system as it is the most commonly used PHR as well as its range of import/export capabilities (e.g. ability to import/export in CCD, CCR format as well as in custom MS Health Vault specific XML or CSV format). Regarding the EHR, an open source solution CityEHR [46] has been selected due to its ability to work with XML database and directly save data in CDA format. By selecting an open source solution, vendor dependencies were avoided. Finally, it is noteworthy that any other EHR which allows integration could be used instead. In this case further transformation of the CDA output to the supported format (e.g. HL7 v2 messages, FHIR etc.) might have been necessary.

The first step is to identify the participating entities and roles. In both scenarios the required information was available in the header of the exported document whether that was in CCD or in CDA format. Figure 11 below shows an example of CCD header where the relevant sections of Patient Role, Author (expanded sections at the top), custodian, legal authenticator and participants (collapsed sections at the bottom).

```
<id assigningAuthorityName="MS Health Vault" root="2.16.840.1.113883.19.5"/>
          <streetAddressLine/>
          <streetAddressLine/>
          <state/>
          <postalCode/3</pre>
               </addr>
       <telecom use="HP"/>
        <telecom use="WP"/>
       <telecom/>
       <telecom use="MC"/>
       <telecom value="mailto:pplastiras@smail.com"/>
       <patient>
             <given>Panagiotis</given>
             <family>Plastiras</family>
          <administrativeGenderCode code="M" codeSystem="2.16.840.1.113883.5.1" codeSystemName="HL7 AdministrativeGender" displayName="Male"/>
          <br/><br/>birthTime value="19870725"/>
          <languageCommunication>
             <languageCode/>
          </languageCommunication>
       </patient>
    </patientRole>
 </recordTarget>
 <author contextControlCode="OP" typeCode="AUT">
    <assignedAuthor classCode="ASSIGNED">
       <addr use=""/>
       <telecom use=""/>
       <assignedPerson classCode="PSN" determinerCode="INSTANCE">
          <name use=""/>
       <representedOrganization classCode="ORG" determinerCode="INSTANCE">
           <id root="2.16.840.1.113883.19.5"/>
          <name>MS Health Vault</name>
          <addr use=""/>
       </representedOrganization>
    </assignedAuthor>
 </author>
<legalAuthenticator>
Code="CP" typeCode="REF">
```

Figure 11 - Example of CCD header from MS Health Vault

The information has been extracted using a set of XPath [47] expressions. For instance, Table 1 below shows example XPath expressions for Patient and Author roles as well as different entities. Similar XPath expressions were used in order to retrieve all the necessary data from both the exported documents as well as the instantiation of the relevant ontology which was expressed in OWL format.

Table 1 - Examples of XPath Expressions for CCD Header

Participant Type	CCD XPath Expression
Patient	/ClinicalDocument/recordTarget/patientRole
Person	/ClinicalDocument/recordTarget/patientRole/patient
Author	/ClinicalDocument/author
Custodian	/ClinicalDocument/author
Legal Authenticator	/ClinicalDocument/legalAuthenticator

For the next step the transformation rules had to identify various elements that have been exported such as allergies, immunizations and medications. Both CCD and CDA documents contain each main

element within separate components tags in the XML structure. Child tags of components are sections, text and entry tags respectively. From the content of these tags the elements that have been exported can be identified immediately. An example of the Allergy section can be seen below in Figure 12 and 13 below.

```
<component>
  <section>
    <templateId assigningAuthorityName="HL7 CCD" root="2.16.840.1.113883.10.20.1.2"/>
    <code code="48765-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC"</pre>
        displayName="Allergy"/>
   <title>Allergy</title>
    <text>
      <thead>
           Type
             Substance
             Substance RxNorm code
             Reaction
             Date Identified
             Status
           </thead>
```

Figure 12 - Allergy Section in a CCD document

```
V<entry typeCode="DRIV">
V<entry typeCode="CRIV">
V<entrolled controlled controlled
                            </code>
</ple>

</ple>
</participantRole>
</participantRole>

<p
                                             </entryRelationship>
```

Figure 13 - Allergy Entries in a CCD document

Next the expected attributes for the elements highlighted from the previous step must be identified in the instantiation of the ontology for Microsoft Health Vault. Using as an example the allergy element, Figure 14 shows the instantiation of the ontology and the attributes that the PHR can support.

```
<ObjectPropertvAssertion>
    <ObjectProperty IRI="#contains"/>
   <NamedIndividual IRI="#Allergy"/>
   <NamedIndividual IRI="#Allergy_First_Observed"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
    <ObjectProperty IRI="#contains"/>
   <NamedIndividual IRI="#Allergy"/>
   <NamedIndividual IRI="#Allergy_Type"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
    <ObjectProperty IRI="#contains"/>
   <NamedIndividual IRI="#Allergy"/>
    <NamedIndividual IRI="#Allergy_Reaction"/>
</ObjectPropertyAssertion>
<ObjectPropertvAssertion>
    <ObjectProperty IRI="#contains"/>
   <NamedIndividual IRI="#Allergy"/>
    <NamedIndividual IRI="#Allergy_Notes"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
   <ObjectProperty IRI="#contains"/>
   <NamedIndividual IRI="#Allergy"/>
   <NamedIndividual IRI="#Name"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
    <ObjectProperty IRI="#contains"/>
    <NamedIndividual IRI="#Allergy"/>
   <NamedIndividual IRI="#Notes"/>
</ObjectPropertyAssertion>
```

Figure 14 - MS Health Vault Ontology Properties Assertion for Allergy

The values for those attributes must be gathered from the exported documents and a set of preasserted data attributes from the ontology instantiation file have been used. For example, the property 'hasEHRTransferable' is used to dictate whether an element is going to be transferred to an EHR. An example assertion in OWL can be seen in Figure 15. In this case the allergy element will be transferred to the EHR system.

Figure 15 - Data assertion for element 'allergy'

Once the data are extracted the relevant transformation must be applied. This transformation was carried out utilizing pre-defined XML CDA templates. An example template can be seen in Figure 10 on the previous section. The engine successfully identified all the required templates for this transformation and collated them to produce the final CDA document. An example of the final transformed document section for Allergies can be seen in Figure 16 below.

Figure 16 - Allergy section transformed to CDA according to our transformation rules

Finally, the information identified in rule 1 must be mapped to the header of the final CDA document. For this transformation predefined XML templates as per previous steps.

The final CDA/CCD document were validated to verify that were well-formed, namely following the syntax rules of XML and valid against the CDA content-specific rules. The validation in all cases was successful and the final document were integrated to the receiving system.

# Challenges encountered when executing the scenarios (make this subheading within results)

The prime challenge during our evaluation was how to treat elements exported from thesystem which were unsupported by the destination system. This was mainly observed with data exported from EHR to PHR. For example, elements such as the X-Ray section that was included under the physical examination section of the exported file as or elements which were included as vital signs on the EHR system (i.e. pain level and saturation) but were not supported by the selected PHR system. The solution that was applied in such scenarios was to map all the unsupported elements as text entries in the relevant 'Comment' fields of the most relevant section to avoid losing important clinical information. This mapping was identified during the transformation rule 3 but the actual transformation occurred during rule 4.

A second challenge was the difference between the attributes captured by a PHR and the attributes that are actually supported by a CCD document. For example, according to the instantiation of the ontology, for an immunization, the elements immunization type, date given, number of sequence, who administers the immunization, adverse effects, location administered on the body, how it was administered and the manufacturer of the substance, were expected in the exported document. However a set of fewer elements was identified in the exported document (immunization name, administration date, administrator and sequence. Therefore, data from the PHR system was not fully transferred to the EHR system. The example provides clear evidence of shortcomings of current standards, namely the CCD standard has predefined capacity for the number supported elements. In

order to overcome this our transformation rules ignores any elements captured in the PHR namely exist in the ontology instantiation but are not present in the exported file.

A third challenge involved the semantic interpretation of the expected data. For example, elements 'Medications' and 'Procedures' which appear as an outcome of rule 2 in one scenario , were not directly identified. According to the ontology file the expected elements were 'list of Procedures' and 'List of Medications'. Our information model can handle such scenarios by adding values to the relevant attributes, elements can be mapped to relevant coding schemas to resolve sematic irregularities. Figure 17 demonstrates the addition of different codes (in this example from coding schemes SNOMED-CT and LOINC), allowing to map the term 'List of Medication' from the PHR to the term 'Medications' in the CDA document.

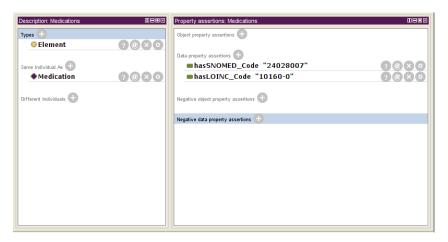


Figure 17 - Declaration of different codes for the same element in Protégé

Finally, as mentioned above a set of pre-asserted data attributes such as the 'hasEHRTransferable' property from the ontology instantiation file have been used to determine how the transformation rules should treat various elements. As a different example of how such attributes were used Figure 18 demonstrates data properties as they appear in Protégé for 'Dietary Intake' which has a value of 0 for property 'hasEHRTransferable' indicating that the exported elements should not be transferred to an EHR and hence 'Dietary Intake' will not be included in the next step. Another attribute 'hasNonXMLBody' indicates whether the respective element should be mapped to the structured or unstructured body of a CDA document (in the case where 'hasEHRTransferable' has a value of 1, thus allowing the information model to support both clinical and non-clinical data. The example demonstrates how data properties may allow for extra data manipulation and flexibility. This rule was completed as expected.

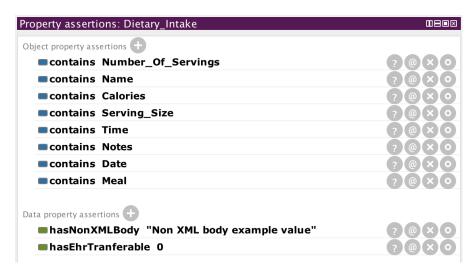


Figure 18 - Data properties in Protégé for element 'Dietary Intake'

#### Discussion and Conclusion

In this paper we have presented an information model that attempts to overcome inadequacies of existing standards for PHR data such as CCD. Our attempt utilizes a multilayer architectural model which combines ontologies in order to allow flexibility and potentially extend beyond and untie the solution from any specific PHR or EHR system. We have mapped our information model to HL7 RIM classes and proposed a set of transformation rules defining how to apply the information model to produce standardized PHR and EHR documents.

The evaluation of the information model allowed various challenges to emerge most of which were due to the semantics of data and it also revealed gaps in the current standards in use... Our solution, in contrast with existing standards is not based on using class attributes to capture values about specific aspects of data (i.e. the actual value or the unit of measurement). Rather our proposition uses classes to capture the data and attributes to define the way the data will be manipulated by various applications. This offers a more flexible approach to manipulate the captured data in the way a user requires. As well as being applicable to existing PHR systems (both tethered and nontethered), the findings could be used as a blueprint to develop new PHR applications.

As future work for our project, connectivity and compatibility with other emerging standards such as openEHR archetypes and FHIR should be examined. Moreover, a comprehensive evaluation of our proposition with the latest trending personal health and wellbeing management applications as well as consumer devices is a necessity. Finally, other data formats which extent beyond the strict clinically defined context such as Observations of Daily Living (ODL) must be examined in order to identify whether our proposition is able to insure seamless integration from these new sources.

#### References

- 1. Detmer D, Bloomrosen M, Raymond B, Tang, P. Integrated personal health records: Transformative tools for consumer-centric care. BMC Med Inform Decis Mak. 2008;8:45.
- 2. Tang PC, Ash, JS, Bates DW, Overhage JM, Sands DZ. Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption. J Am Med Inform Assoc. 2006;13:121-6.
- 3. Jilka SR, Callahan R, Sevdalis N, Mayer EK, Darzi A. "Nothing About Me Without Me": An Interpretative Review of Patient Accessible Electronic Health Records. J Med Internet Res. 2015;17: e161
- 4. Giardina TD, Menon S, Parrish D, Sittig D, Singh H. Patient access to medical records and healthcare outcomes: a systematic review. J Am Med Inform Assoc. 2014;21:737–41.
- 5. Kim K. Benefits of and barriers to the use of personal health records (PHR) for health management among adults. OJNI 2012 16:e3.
- 6. Dorr D, Bonner LM, Cohen AN, Shoai RS, Perrin R, Chaney E, Young AS. Informatics systems to promote improved care for chronic illness: A literature review. J Am Med Inform Assoc. 2007;14:156-163.
- 7. Amante DJ, Hogan TP, Pagoto SL, English TM. A systematic review of electronic portal usage among patients with diabetes. Diabetes Technol Ther. 2014;16:784–93
- 8. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. J Am Med Inform Assoc. 2011;18:515–522.
- 9. Undem T. Consumers and Health Information Technology: A National Survey. Oakland (CA): California HealthCare Foundation; 2010.
- 10. Markle Foundation. The Public and Doctors Largely Agree Patients Should Be Able To View, Download and Share Their Health Info. 2011.
- 11. Studeny J., Coustasse A. Personal Health Records: Is Rapid Adoption Hindering Interoperability?. Perspect Health Inf Manag. 2014 Summer; 11(Summer): 1e.
- 12. Vydra T. P., Cuaresma E., Kretovic M., Bose-Brill S. Diffusion and Use of Tethered Personal Health Records in Primary Care, Perspect Health Inf Manag. 2015 Spring; 12(Spring): 1c.
- 13. Kaiser Permanente. Kaiser Permanente News Center: 2015 Annual Report. 2015.
- 14. Tsai J, Rosenheck RA. Use of the internet and an online personal health record system by US veterans: comparison of Veterans Affairs mental health service users and other veterans nationally. J Am Med Inform Assoc. 2012;19(6):1089-1094.
- 15. Beard L, Schein R, Morra D, Wilson K, Keelan J. The challenges in making electronic health records accessible to patients. J Am Med Inform Assoc. 2012;19:116–120.
- 16. Kahn J S, Aulakh V, Bosworth A. What It Takes: Characteristics Of The Ideal Personal Health Record. Health Aff. 2009;28:369-376.
- 17. de Lusignan S, Mold F, Sheikh A, Majeed A, Wyatt JC, Quinn T, Cavill M, Gronlund TA, Franco C, Chauhan U, Blakey H, Kataria N, Barker F, Ellis B, KoczanP, Arvanitis TN, McCarthy M, Jones S, Rafi I. Patients' online access to their electronic health records and linked online services: a systematic interpretative review. BMJ Open. 2014;4:e006021.
- 18. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. BMJ 2010;341:c5814
- 19. Lafky DB, Horan TA. Prospective Personal Health Record Use Among Different User Groups: Results of a Multi-wave Study. HICSS 2008: Proceedings of the 41st Hawaii International Conference on System Sciences; 2008 Hawaii, USA. New York: IEEE Press; 2008. p. 233-241.

- 20. Liu LS, Shih PC, Hayes GR. Barriers to the Adoption and Use of Personal Health Record Systems. iConference 2011: Proceedings of the iConference; 2011 Feb 8-11, Seattle, WA, USA. New York: ACM; 2011; p. 363-370.
- 21. Lafky DB, Horan TA. Personal health records: Consumer attitudes toward privacy and security of their personal health information. Health Informatics J. 2011;17:63-71.
- 22. Wartena F, Muskens J, Schmitt L. Continua: The Impact of a Personal Telehealth Ecosystem. eTELEMED 2009: Proceedings of IEEE Computer Society, International Conference on eHealth, Telemedicine, and Social Medicine; 2009 Feb 1-7, Cancun, Mexico. Piscataway, NJ: IEEE; 2009;p.13–18.
- 23. Pham H, Schrag D, O'Malley AS, Wu B, Bach PB. Care Patterns in Medicare and Their Implications for Pay for Performance. N Engl J Med. 2007;356;1130-1139.
- 24. American Society of the International Association for Testing and Materials [Internet] Continuity of Care Record [cited 2016 May 31] Available from: http://www.astm.org/Standards/E2369.htm
- 25. Health Level 7 [Internet] Continuity of Care Document [cited 2016 May 31] Available from: http://www.hl7.org/implement/standards/product\_brief.cfm?product\_id=6
- 26. Health Level 7 [Internet] Clinical Document Architecture [cited 2016 May 31] Available from: http://www.hl7.org/implement/standards/product\_brief.cfm?product\_id=7
- 27. Binh Q. Tran and Pedro Gonzales. Standards and Guidelines for Personal Health Records in the United States: Finding Consensus in a Rapidly Evolving and Divided Environment. J Health Med Informat S6:001.
- 28. Ved V, Tyagi V, Agarwal A, Pandya AS. Personal Health Record System and Integration Techniques with Various Electronic Medical Record Systems. HASE 2011: Proceedings of 13th International Symposium on High-Assurance Systems Engineering; 2011 10-12 Nov, Boca Raton, FL, USA. IEEE; 2011. p.91-94.
- 29. Li M, Yu S, Zheng Y, Ren K, Lou W. Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption. IEEE Trans. Parallel Distrib. Syst. 2013;24: 131-143.
- 30. Xie L, Yu C, Liu L, Yao, Z. XML-based Personal Health Record System. BMEI 2010: Proceedings of 3rd International Conference on Biomedical Engineering and Informatics; 2010 16-18 Oct, Yantai, China. IEEE; 2010. p.2536-2540.
- 31. Lee LH, Chou YT, Huang EW, Liou DM. Design of a Personal Health Record and Health Knowledge Sharing System Using IHE-XDS and OWL. J Med Syst. 2013;37:9921-9924.
- 32. Puustjärvi J, Puustjärvi L. Designing and Implementing an Active Personal Health Record System. eTELEMED 2011: Proceedings of the 3rd International Conference on eHealth, Telemedicine, and Social Medicine; 2011 Feb 23-28; Gosier, Guadeloupe, France. Piscataway, NJ: IEEE;2011.p.102-107.
- 33. Lähteenmäki J, Leppänen J, Kaijanranta H. Interoperability of Personal Health Records. IEEE CBMS 2009: Proceedings of 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2009 Sep 3-6; Minneapolis, Minnesota, USA. Piscataway, NJ: IEEE;2009.p.1726-1729.
- 34. Yuksel M, Dogac A. Interoperability of Medical Device Information and the Clinical applications: An HL7 RMIM based on the ISO/IEEE 11073 DIM. IEEE Trans Inf Technol Biomed. 2011;15: 557-66.
- 35. Brut M, Al-Kukhun D, Péninou A, Canut MF, Sèdes F. APHR: Annotated Personal Health Record for Enabling Pervasive Healthcare. IEEE MDM 2011: Proceedings of 12th IEEE International Conference on Mobile Data Management; 2011 June 6-9; Luleå, Sweden. Piscataway, NJ: IEEE; 2011.p.73-79.

- 36. Health Level 7 [Internet] Reference Information Model [cited 2016 May 31] Available from: <a href="http://www.hl7.org/implement/standards/rim.cfm">http://www.hl7.org/implement/standards/rim.cfm</a>
- 37. Plastiras P.,O'Sullivan D., Weller P., An Ontology-Driven Information Model for Interoperability of Personal and Electronic Health Records, The Sixth International Conference on eHealth, Telemedicine, and Social Medicine, 24-03-2014 27-03-2014, Barcelona.
- 38. Bosca D., Moner D., Maldonado A., Robles M. Combining Archetypes with Fast Health Interoperability Resources in Future-proof Health Information Systems, 2015 European Federation for Medical Informatics (EFMI).
- 39. Smith B., Ceusters W., HL7 RIM: An Incoherent Standard, Studies in Health Technology and Informatics, vol. 124 (2006), 133–138
- 40. Protégé [Internet] [cited 2016 May 31] Available from: http://protege.stanford.edu/
- 41. Web Ontology Language [Internet] [cited 2016 May 31] Available from: http://www.w3.org/standards/techs/owl#w3c\_all
- 42. openEHR [Internet] [cited 2016 July 20] Available from: http://www.openehr.org/
- 43. Fast Healthcare Interoperability Resources (FHIR) [Internet] [cited 2016 July 20] Available from: https://www.hl7.org/fhir/
- 44. NHS Interoperability Toolkit [Internet] [cited 2016 May 31] Available from: <a href="http://www.hl7.org.uk/marketing/itk/itk.asp">http://www.hl7.org.uk/marketing/itk/itk.asp</a>
- 45. Microsoft Health Vault [Internet] [cited 2016 May 31] Available from: <a href="https://www.healthvault.com/gb/en">https://www.healthvault.com/gb/en</a>
- 46. CityEHR [Internet] [cited 2016 May 31] Available from: <a href="https://sourceforge.net/projects/cityehr/">https://sourceforge.net/projects/cityehr/</a>
- 47. XML Path Language (XPath) [Internet] [cited 2016 May 31] Available from: <a href="https://www.w3.org/TR/xpath/">https://www.w3.org/TR/xpath/</a>