Advancing Patient Record Safety and EHR Semantic Interoperability

Konstantinos Perakis[*], Thanasis Bouras[*], Dimitris Ntalaperas[*], Panagiotis Hasapis[†], Christos Georgousopoulos[†], Ratnesh Sahay[‡], Oya Deniz Beyan[‡], Cristi Potlog[ʃ], Daniela Usurelu[ʃ]

[*]UBITECH Ltd., Athens, Greece, {kperakis, bouras, dntalaperas}@ubitech.eu
[†]INTRASOFT International S.A., Luxembourg, Luxembourg, {panagiotis.hasapis, christos.georgousopoulos}@intrasoft-intl.com
[‡]Digital Enterprise Research Institute (DERI), NUI Galway, Galway, Ireland, {ratnesh.sahay, Oya.Beyan}@deri.org
[ʃ]SIVECO Romania S.A., Bucharest, Romania, {Cristi.Potlog, Daniela.Usurelu}@siveco.ro

Abstract—**Electronic Health Records (EHRs) contain an increasing wealth of medical information, which has the potential to significantly advance medical research and health policies formulation, providing society with additional benefits within a global health perspective. However, the European healthcare information space is fragmented due to the lack of legal and technical standards, cost effective platforms, and sustainable business models. Providing an interoperability infrastructure for EHRs is on the agenda of many regional, and international eHealth initiatives. The semantic interoperability of patient data between EHRs and medical research can transform today's process of drug discovery and development, enable faster access to effective new medications, provide improved patient outcomes, and provide a key foundation for targeted (personalized) medicines. The scope of the current paper is the description of the effort undertaken by the Linked2Safety consortium towards the development of an innovative interoperability framework, for the efficient, homogenized access to and the effective utilization of the increasing wealth of medical information contained in the EHR systems deployed and maintained at regional and/or national level across Europe.**

Keywords - EHR, interoperability, semantic, patient, safety, datacubes;

I. INTRODUCTION

Electronic Health Records (EHRs) contain an increasing wealth of medical information which has the potential to significantly advance medical research, personalized healthcare, as well as health policies formulation, providing society with additional benefits within a global health perspective, with applications ranging from disease prevention and genetics to surveillance and epidemiologic studies [1]. Healthcare Organizations and the pharmaceutical industry in Europe both share a common goal - to deliver the best possible personalized treatments and innovative medicines to improve patient outcomes. The industry believes that technological advances and broad implementation of EHRs in Europe can achieve this goal and accelerate clinical research.

However, the European healthcare information space is fragmented due to the lack of legal and technical standards, cost effective platforms, and sustainable business models. The potential gains in efficiency and effectiveness for primary care afforded by rapid and secure access to patient healthcare data in electronic form are widely recognized today across the EU.

Providing an interoperability infrastructure for EHRs is on the agenda of many regional, pan-European and international eHealth initiatives [2], while about half of the member states are currently working on national eHealth infrastructures[3]. Towards this end, common standards and interoperability will bring opportunities for a global approach for the benefit of patients, health systems and the market. Effective use of EHRs has the potential to positively influence both the quality and the cost of health care [4]. Consequently, sharing patient's EHRs is becoming a global priority in the healthcare information technology domain [5].

On top of the common standards, the semantic approaches to promote interoperability among standard-compliant information systems, e.g. reference ontologies and mediation, have proven to be able to have significant potential as regards the integration of information from distributed EHR databases [6, 7]. The semantic interoperability of patient data between EHRs and clinical research can transform today's process of drug discovery, development and commercialization, enable faster access for patients to effective treatments, provide improved patient outcomes, improve medication safety and adverse event signal detection, and provide a key foundation for personalized medical models.

The scope of the current paper is the description of the effort undertaken by the Linked2Safety consortium towards the development of an innovative interoperability framework, for the efficient, homogenized access to and the effective utilization of the increasing wealth of medical information contained in the EHRs and Electronic Data Capture (EDC) systems deployed and maintained at regional and/or national level across Europe [8].

The Linked2Safety project addresses research challenges in the fields of a) clinical terminology mapping and integration, b) querying and reasoning over heterogeneous and distributed clinical data stores, c) the secure and reliable access of sensitive clinical resources, and d) annotation and semantic interlinking of clinical resources. Only few projects such as Artemis[1], RIDE[2], ACGT[3], @neurIST, and W3C COI[4] use healthcare

---

[1] http://www.srdc.com.tr/metu-srdc/projects/artemis/
[2] http://www.srdc.com.tr/metu-srdc/projects/ride/
[3] http://www.eu-acgt.org/

standards in their architectural frameworks. The EHR models used with the Linked2Safety framework is based on the openEHR [5] standard. Healthcare and Life Sciences (HCLS) projects contribute to different aspects of EHR/EPR interoperability. For example, the Artemis project focuses on data and services level interoperability by semantically annotating the HL7 compliant clinical messages and service descriptions. CALBC[6] is more focused on semantic annotation and interlinking of scientific documents in the biomedical domain. Such an interlinking mechanism is also crucial in linking the patient records with clinical trial documents. In HCLS projects, mappings/alignments between terminologies, coding schemes, standard and proprietary vocabularies, and standard information models, is applied only at the schema level. Linked2Safety supports semantic interlinking of clinical resources used internally as well external links to the Linked Open Dada (LOD[7]) cloud. As regards the formal description of ontologies, a number of different languages have been introduced during the last decade. The most widely used ones are Web Ontology Language[8] (OWL) and RDFs[9]. The current version of OWL allows expressiveness via complex descriptions of classes and relationships. It offers many forms of syntaxes and expressiveness levels (separated in different sub-categories of languages (named OWL Lite, OWL DL and OWL Full) in order to address computational complexity issues. It is a widely accepted description language and is the basis of what has been used for the ontologies described in this paper.

II. METHODOLOGY

The main goal of the Linked2Safety project is to facilitate efficient and homogenized access to the increasing wealth of medical information contained in heterogeneous EHR and EDC systems deployed and maintained at regional and/or national level across Europe, by establishing the concepts and tools that will enable scalable, standardized, secure sharing and reuse of statistical medical data as a foundation for policy prediction, planning and adjustments.
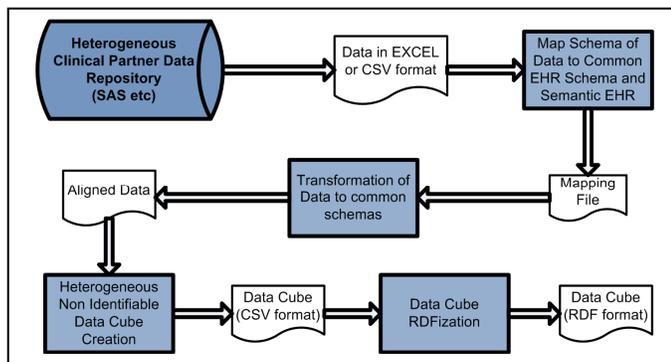

Figure 1.   Methodological pipeline

[4] http://www.w3.org/wiki/HCLS/ClinicalObservationsInteroperability

[5] http://www.openehr.org/

[6] http://www.ebi.ac.uk/Rebholz-srv/CALBC/project.html

[7] http://lod-cloud.net/

[8] http://www.w3.org/TR/owl-features/

[9] http://www.w3.org/RDF/

The Linked2Safety consortium has adopted the data cube generation, cell-suppression and perturbation techniques as described in [9] to address the ethical requirements of handling sensitive patient data, namely: respecting patient anonymity, data ownership and privacy, as well as compliance with legislative, regulatory and ethical requirements. The data cube approach also enables the use of only non-identifiable, small, truncated statistical healthcare data aligned for online computation and analysis. In this way, the healthcare data maintained by data providers remain in their secure, on-site, off-line physical environment, and only selected, homogenized and anonymized data cubes are allowed to leave the clinical data providers' premises and enter the Linked2Safety platform.

In order to realize these goals, the Linked2Safety consortium followed a step-wise methodology which comprises of the following steps depicted in Fig. 1.

This step-wise methodological approach serves the goal of developing an innovative interoperability framework, meeting the project goals which are:

- to facilitate the alignment, transformation and standardized bridging of existing EHR and EDC repositories to a common-reference EHR schema (Linked2Safety Common EHR Schema) - addressing syntactic and technical interoperability between EHR and EDC repositories. This is achieved by defining the Linked2Safety Common EHR Schema, which is outlined in section II.A.

- to facilitate the establishment of a common reference model for the semantic annotation, representation and alignment between EHR artifacts and standard medical vocabularies. This is achieved by defining the Linked2Safety Semantic EHR Schema, which is outlined in section II.B.

- to allow the (manual) alignment of the medical variables of data providers to the Linked2Safety Common EHR Schema, thus enabling extensive processing of health data coming from heterogeneous sources, regardless of their origin. This is achieved by developing the Linked2Safety Schema Mapping and Alignment Tool, which is outlined in section II.C.

- to allow the automatic transformation of EHRs and EDCs to the Linked2Safety Common EHR Schema based on the mapping file produced by the Linked2Safety Schema Mapping and Alignment Tool. This is achieved by developing the Transformation to the Linked2Safety Common EHR Schema Tool, which is outlined in section II.D.

- to address the ethical requirements of handling sensitive patient data, as well as the compliance with legislative, regulatory and ethical requirements, and to enable the use of only non-identifiable, small, truncated statistical healthcare data aligned to the Linked2Safety Common EHR Schema. This is achieved by defining the Linked2Safety Data Cube Creation Mechanism, which is outlined in section II.E.

- to allow the conversion of the aggregated data in the form of a data cube into an RDF version of the same data adhering to the RDF Data Cube Vocabulary. This is achieved by developing the Linked2Safety RDFisation Mechanism that is outlined in section II.F.

Each goal is attributed by specific processes and model that are shown in Fig. 1. In the next few sections we present the details of our method.

A. Linked2Safety Common EHR Schema

The Linked2Safety Common EHR Schema is an ontological reference model used in the Linked2Safety project for mapping all proprietary and non-proprietary protocol-based EHRs and EDCs of clinical data providers, and also by alignment engines that will facilitate the transformation of the EHR records and EDC databases to datacubes that can be used for creating analytic reports. In order to achieve these goals, the Linked2Safety Common EHR Schema Specification builds upon both the openEHR Reference Model[10] and the openEHR Archetype Model[11], translating basic openEHR concepts to OWL/RDF notation, thus enabling an easy conversion to typical openEHR format of most of the information received from different clinical data providers. The definition of the Linked2Safety Common EHR Schema includes also conceptual bridges and mappings among the openEHR, HL7-CDA[12] and ISO/CEN 13606[13] standards, which are laid out as semantic annotations to object properties and data properties derived from openEHR.

As illustrated in Fig. 2, the Reference Model consists of the following schema files in OWL2:

- BasicTypes.owl – types from rm.data_types and rm.support.identification packages;

- Structure.owl – types from rm.data_structures and rm.common.generic packages;

- Resource.owl – types from the rm.common.resource package;

- Content.owl – item, Section, Entry and all subtypes;

- Composition.owl – rm.composition;

- Version.owl – the Version classes from rm.common.change_control;

- Semantics.owl – annotation properties used semantic mappings of EHR concepts.

- The Archetype Model is composed of the following schema files in OWL2:

- Archetype.owl – types from the Archetype Object Model (AOM);

- Profile.owl – types from the openEHR Archetype Profile package;

- Clinical.owl – types used for representing Linked2Safety clinical trials data from medical partners.

The Linked2Safety Common EHR Schema is the first step towards interoperability and interconnection of healthcare resources based on:

- fostering "separation of concerns" to resolve heterogeneity issues (with data and schema) at a conceptual level;

- employing interlinking/mapping of cross-domain data/schema (e.g. EHR, clinical trial).
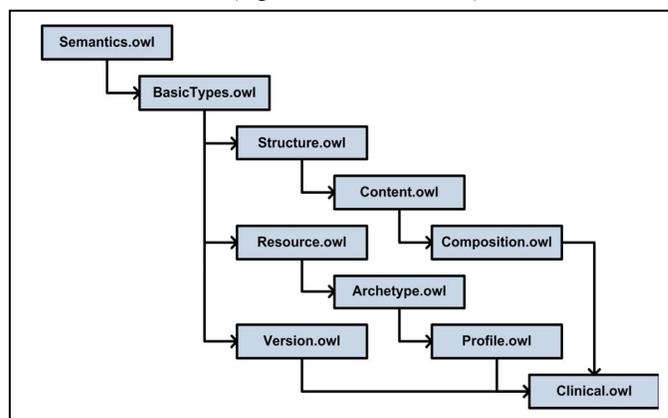

Figure 2. Common EHR Layering

B. Linked2Safety Semantic EHR Schema

The Linked2Safety Common EHR Schema incorporates the most widely-used attributes and best practises from multiple open standards, in the sense that although no particular fields of OpenEHR were utilised (due to the medical domains on which Linked2Safety focuses), it was considered for declaring the structure of the EHRs. By studying the data fields as they were provided by the project's clinical partners, the Linked2Safety Common EHR Schema encapsulates only the basic notions and data fields from the clinical partners (which are common to all of them).

The Linked2Safety Semantic EHR (SEHR) Model is a light-weight and extensible ontology that covers multiple sub-domains of Healthcare and Life Sciences (HCLS) through specialisation of the upper-level Basic Formal Ontology (BFO). The main difference from the Linked2Safety Common EHR Model is that the SEHR ontology contains domain specific clinical variables (e.g., mental disorder, diabetic, breast cancer) used by the clinical data providers, allowing expressiveness. Also, it goes a step further by combining entities that are similar together, either in concepts or in properties, along with a way of enabling an "interface" with other upper layer ontologies.

In order to query across diverse and disparate clinical repositories, the Linked2Safety SEHR provides formal and consistent definitions of clinical terms. The Linked2Safety SEHR Model is built using an alignment methodology, namely the Plug and Play Electronic Patient Record (PPEPR) Methodology, which ontologises the Health Level Seven (HL7) standard, and which had been proposed in our earlier work

---

[10] http://www.openehr.org/releases/1.0.2/architecture/rm/ehr_im.pdf
[11] http://www.openehr.org/releases/1.0.2/architecture/am/aom.pdf
[12] http://www.hl7.org/implement/standards/product_brief.cfm?productid=7
[13] http://www.en13606.org/the-ceniso-en13606-standard

[10]. One of the main features of the Linked2Safety SEHR is the layering of local clinical terminologies with the upper-level BFO. The layering task arranges ontologies into global (shared by all clinical users) and local (used within a particular institute) spaces.
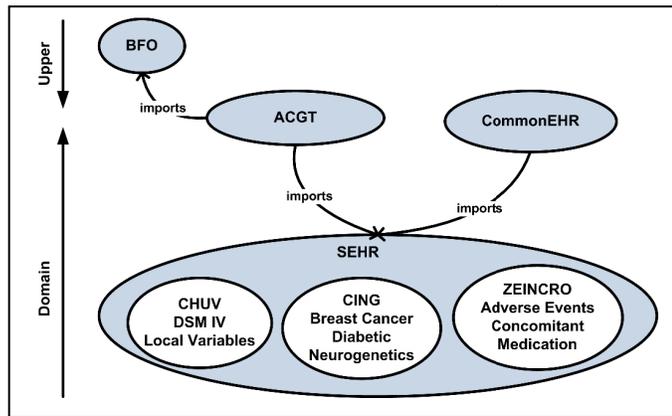

Figure 3.   Semantic EHR Layering

Figure 3 illustrates how global ontologies are arranged in a top-down fashion where ACGT extends the BFO. The ACGT Ontology was developed within the ACGT project, which focused on breast cancer and nephroblastoma (Wilms' Tumor) [11], with one of its aims being to create an ontology-driven clinical trial management system. Local ontologies are created from the clinical variables or terminologies and later aligned with the ACGT ontology. In Fig. 3 the circled cross symbol ($\times$) implies alignment of ontologies as two conceptual ambiguities exist between local ontologies: (i) semantically similar concepts are named differently; (ii) corresponding concepts are represented at different structural levels. These ambiguities arise because local systems have flexibility and different alternative choices in the design and use of clinical terminologies. To deal with both issues (i.e., different naming schemes and structural differences), one option is to provide directed alignments between local ontologies. However, directed set of local alignments (i.e., bi-directional alignments) would result in a quadratic size "n×(n-1)" alignments for "n" number of clinical applications. Considering the diversity of the HCLS domain, the use of local terminologies and corresponding alignments cannot be completely avoided. In other words, local clinical terminologies originating from various clinical sites cannot be predefined in a central conceptual model (or an ontology). Therefore, in the Linked2Safety SEHR a hybrid approach of aligning the majority of local terminologies with the global ontologies was employed; thus allowing only minimal set of bi-directional local alignments. In this way, the bilateral correspondences between clinical applications by delegating the majority of alignments with the upper-level (global) ontologies is introduced. Further, the meeting points of the "bottom-up" and "top-down" arrows in Fig. 3 (left hand side), require extensions within local and global ontologies to find suitable correspondences between them.

The PPEPR Methodology that builds the Linked2Safety SEHR Model is iterative; therefore, it allows further extension and adaptation of the Linked2Safety SEHR Model which might be required due to future changes and different applications scenarios. Furthermore, the evaluation of the Linked2Safety SEHR Model is conducted to ensure the correctness and completeness of concepts, properties, and axioms described as part of the Linked2Safety SEHR Model. The evaluation process includes clinical data providers, technical partners, and the domain experts. Our empirical evaluation shows an agreement of clinical experts confirming the Linked2Safety SEHR's usability in clinical trials.

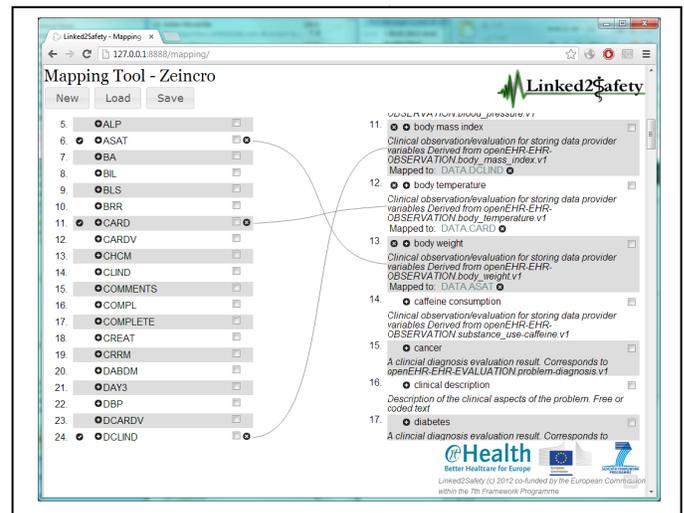*C. Linked2Safety Schema Mapping & Alignment Tool*

Figure 4.   Mapping Tool UI Design

The Linked2Safety Schema Mapping and Alignment Tool is the entry point of patients' data in the Linked2Safety platform, incorporating open-source mapping and alignment engines required for the automatic transformation of EHRs and EDCs to the Linked2Safety Common and Semantic EHR schema described in Section II.A, a baseline against which all medical data variables from different clinical data providers are aligned. The mapping process is performed by expert users from the clinical data providers, and allows them to visually align their medical data variables to the Linked2Safety Common and Semantic EHR Schema using this dedicated tool, thus enabling further extensive processing of health data coming from heterogeneous sources, regardless of the originating application, in the Linked2Safty platform. This component generates a mapping file that is later consumed by the transformation tool to automatically convert data to the Linked2Safety Common EHR schema reference representation. Figure 4 presents a design concept for the user interface of this tool.

D. Linked2Safety Transformation Tool

The Linked2Safety Transformation Tool performs the automatic transformation of EHRs and EDCs to the reference Linked2Safety Common EHR schema, by employing custom developed software tools to ease the task of transformation between metadata standards, thus bridging the gap between analysis and execution.This component uses input data from the clinical data providers in association with a given mapping file, and outputs data transformed and aligned to the Linked2Safety Common and Semantic EHR Schema according to the given mappings in a representation that can be stored in the Linked2Safety repository. Figure 5 presents a design concept for the user interface of this tool.
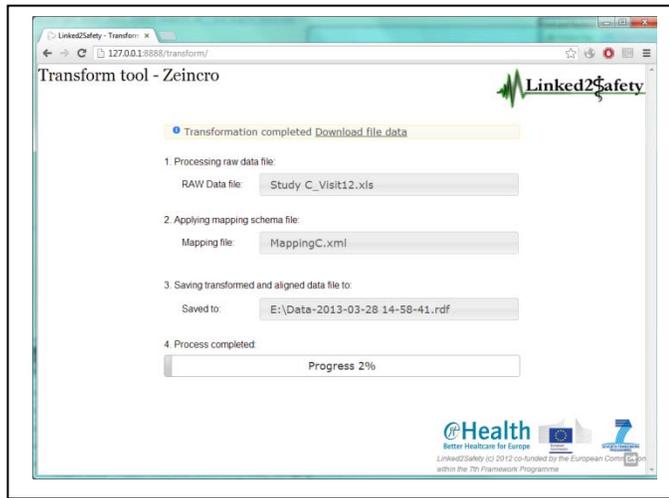

Figure 5.   Transformation Tool UI Design

E. Linked2Safety Data Cube Creation Mechanism

Following the transformation of EHRs and EDCs to the reference Linked2Safety Common EHR schema, and in order to address the ethical requirements of handling sensitive patient data, as well as to comply with the associated legislative, regulatory and ethical requirements, and to enable the use of only non-identifiable, small, truncated statistical healthcare data aligned to the Linked2Safety Common EHR schema, the Linked2Safety Data Cube Creation Mechanism is applied on the transformed data.

Data cubes are created from records that adhere to the Linked2Safety ontologies. A set of records along with a list of medical variables which are used to define the dimensions of the data cube is given as input to the component. If the values of an attribute are continuous, the user can define the ranges according to which the values of the attributes are to be categorized.

Then, a set of records which have already been converted to the Linked2Safety Common EHR schema by the Linked2Safety Transformation tool is gathered, while the variables, along with any categorization of their domain values, are also given as output of the Linked2Safety Transformation tool. Each variable is used for the formulation of one-dimensional slices of the data cube. The data for each slice derives from aggregating the counts of patients that have specific values for the corresponding values. The above process is repeated for each variable. An "NxM" dimensional cube is produced, with "N" being the number of attributes and "M" being the combination of each attribute's values.

In addition, the data cube is both i) filtered, so that data cube cells of smaller values are excluded, and ii) perturbed according to a predefined perturbation policy, so as to limit the possibility of reverse engineering the data cube in order to identify the patients from the aggregated data [12]. The user, taking into consideration the percentage of cells that was filtered out, decides whether the amount of data loss is acceptable. If not, the data cube is discarded and the process may be repeated with a different set of input parameters. If the data cube is of acceptable quality the user may add annotations to the resulting data cube, thus finishing the process.

The Data Cube Creation Algorithm utilized in the context of Linked2Safety is presented in Fig. 6. Initially, for each line, the algorithm first computes the unique key identifying a patient. Then, it obtains the values of the variables. Each variable is categorized according to the ranges present in a mapping file by calling *convertInRange()*. Thus, for each line, an array *value[]* is created which stores all the variables values after classification. From this array an index is computed by calling *indexOf()*. The index is uniquely determined by the combination of a specific selection of ranges from the variables list.

```
foreach line
        computeunique_key
        foreach cell in line_cells
            cell_field++
            value[cell_field] = valueOf(cell)
            value[cell_field] = convertInRange(value, ranges[cell_field])
        Cube[unique_key][indexOf(value[cell_field])]++
foreachunique_key
        for  i=1 to num_of_fields
        perturbe(Cube[unique_key][i])
        if Cube[unique_key][i] <= threshold
            discard (Cube[unique_key][i])
```

Figure 6.   Data cube Creation Pseudocode

F. Linked2Safety RDFisation Mechanism

The last step of the methodological pipeline described in the beginning of section II and illustrated in Fig. 1 concerns the conversion of a data cube into a format where it can later be semantically retrieved using an RDF query language (such as SPARQL). The Linked2Safety Data Cube Creation Mechanism produces the desired data cube(s) in CSV format, which are then provided as input to the Linked2Safety RDFisation mechanism. This mechanism is responsible for converting the aggregated clinical data which are in the form of a data cube, into an RDF version of itself adhering to the RDF Data Cube Vocabulary [13].

The RDF Data Cube Vocabulary introduces a reusable set of concepts and components based on SDMX (Statistical Data and Metadata Exchange). The SDMX standard includes a set of Content Oriented Guidelines (COG) which define a set of common statistical concepts and associated code lists that are intended to be reusable across data sets. In the context of the current work, RDF analogues to the COG are created.

The process of the RDFization of a data cube is straightforward. The core class element of an RDF Data Cube Vocabulary is the "Dataset". This class describes the structure of a data cube as a collection of the class "Observation". The latter class holds the description of an occurrence number, the dimensions coordination for that number along with the semantics of those dimensions.

```
Output imports of vocabularies
Output structure and URI's for the newly created data cube
For each line in data cube
  Create a new Observation
   In each observation assign the values & data types for each dimensions
```

Figure 7. Data cube RDFizing Pseudocode

The pseudo-code of the Data Cube RDFizing Algorithm is presented in Fig. 7. Before processing the data cubes lines that contain the data, the structure of the cube with a unique URL for that is created. Next, a new observation for each line that is being read by the data cube is created and it is outputted to a file. It should also be noted that the algorithm employs the Turtle format, a compact way to serialize RDF triples.

III. RESULTS

All of the components described within this paper were designed and developed in the context of the Linked2Safety project and were validated in real life scenarios at the premises of the consortium clinical data providers, namely ZEINCRO Hellas S.A., the Center for Psychiatric Epidemiology and Psychopathology of Lausanne, and the Cyprus Institute of Neurology and Genetics. At each step of the design and development of these components the clinical data providers were evaluating the process not only in terms of the final outcomes but also in intermediate stages, by providing feedback to the technical partners of the consortium by means of completed questionnaires or specific instructions based on the pre-existing knowledge of their own data. The final outcome of the project will be validated in real life scenarios, in which each of the data providers will be asked to execute predefined scenarios in order to ensure compliance of the desired and actual result. The results of the utilization of the Linked2Safety Common EHR Schema, the Linked2Safety Semantic EHR Schema, and of the application of the four components on datasets from the clinical partners highlighted the potential of the Linked2Safety platform.

More specifically, as a first step, the clinical piloting candidates indicated the data variables utilized in clinical practice, from which the two schema models were built. The second step included the provision of the software prototypes to the consortium clinical data providers so that they could perform mappings of their local schemas to the models created in the project, as well as create sample data cubes. This second step was split in two discrete phases; in the first phase, the software prototypes were provided to the consortium clinical data providers by whom they were evaluated. This evaluation produced a wave of change requests, that led to the second phase which included the incorporation of these requests in the second version of the software prototypes – thus adapting to the specific needs of the consortium clinical data partners. The procedure was highly connected to the provided data from each of the clinical data providers but not limited to that. The clinical partners, having the expertise in their field, evaluated the software prototypes taking into account the adjustments that were made to fit all of the clinical partners needs.

Then, the four components were applied with the ultimate results of producing RDFised data cubes from the specific datasets for purposes of performing statistical analysis and identifying potential adverse events. The statistical analysis can have a defined objective but given the multiple options that the Linked2Safety platform provides, a clinically significant result may be extracted even if not included in the initial objectives. With regards to the adverse events identification, a key point is the ability to relate different types of adverse events to different molecular fragments and thus the ability to detect the possibility of an adverse event occurring when specific conditions are in place. After an initial data analysis and model design phase, due to the difference in the nature of the data provided by the clinical data providers (genetic and non genetic data), the two models were further extended in order to include additional data variables and make the two models more holistic by meeting the clinical partners needs in order to cover a significant spectrum of the variables utilized in clinical practice.

IV. CONCLUSION

In this paper, the work that took place within the context of the Linked2Safety project regarding the development of an innovative interoperability framework, for the efficient, homogenized access to, and the effective utilization of the increasing wealth of medical information contained in the EHRs and EDC systems deployed and maintained at regional and/or national level across Europe was presented. In particular, emphasis was given to the various subsystems that comprise the lower level of the Linked2Safety platform architecture, which include: i) the Linked2Safety Common EHR Schema which allows for the alignment, transformation and standardized bridging of existing EHR and EDC repositories to a common-reference EHR schema, ii) the Linked2Safety Semantic EHR Schema which allows for the establishment of a common reference model for the semantic annotation, representation and alignment between EHR artifacts and standard medical vocabularies, iii) the Linked2Safety Schema Mapping and Alignment Tool which allows for the alignment of the medical variables of data providers to the Linked2Safety Common EHR Schema, iv) the Linked2Safety Transformation Tool which allows for the automatic transformation of EHRs and EDCs to the reference Linked2SafetyCommon EHR Schema, v) the Linked2Safety Data Cube Creation Mechanism which enables the transformation of data to non-identifiable, small, truncated statistical healthcare data aligned to the Linked2Safety Common EHR Schema, and vi) the Linked2Safety RDFisation Mechanism which allows for the conversion of the aggregated data from the form of a data cube into an RDF version of the same data adhering to the RDF Data Cube Vocabulary.

The initial results of the validation of the aforementioned models and components were very promising. The Linked2Safety Common EHR Schema and the Linked2Safety Semantic EHR Schema were designed and built based on the

data variables that were indicated by different clinical data providers. The application of the four components on these datasets highlighted the potential of utilizing the Linked2Safety platform for producing RDFised data cubes for the purposes of performing statistical analysis on them, and identifying potential adverse events.

As a future step, it is the intention of the Linked2Safety consortium to proceed with integrating the current project results for delivering a coherent environment with an intuitive front-end for realizing the next-generation, semantically-interlinked, secure medical and clinical information space in the enlarged Europe featuring a lightweight decision support system to facilitate both the early detection of potential patients safety issues, and the recruitment process of proper candidates for clinical trial executions. Then as a final step, the conduction of a fully scaled two-phase piloting procedure is envisaged to be conducted for validating and evaluating the Linked2Safety platform prior to its final release.

REFERENCES

[1] L. Kun, R. Beuscart, G. Coatrieux, C. Quantin, "Improving outcomes with interoperable EHRs and secure global health information infrastructure", 29th IEEE EMBS Int. Conf, pp. 6158 - 6159, August 2007

[2] O. Kilic, A. Dogac, M. Eichelberg, "Providing Interoperability of eHealth Communities Through Peer-to-Peer Networks", IEEE Transactions on Information Technology in Biomedicine, Volume: 14 Issue: 3, pp. 846 – 853, May 2010

[3] WHO, "Building Foundations eHealth in Europe", Available online at: http://www.who.int/goe/BFeuroFull.pdf, 2010

[4] B. Stewart, S. Fernandes, E. Rodriguez-Huertas, M. Landzberg, "A preliminary look at duplicate testing associated with lack of electronic health record interoperability for transferred patients", JAMIA 2010; 17:341-344 doi:10.1136/jamia.2009.001750

[5] O. Kilic, A. Dogac, M. Eichelberg, "Achieving Clinical Statement Interoperability Using R-MIM and Archetype-Based Semantic Transformations", IEEE Trans. Inf. Tech. Biomedicine, Volume: 13 Issue: 4, pp. 467 - 477, July 2009

[6] J. T. Fernandez-Breis, P. J. Vivancos-Vicente, M. Menarguez-Tortosa, D. Moner, J. A. Maldonado, R. Valencia-Garcia, T. G. Miranda-Mena, "Using semantic technologies to promote interoperability between electronic healthcare records' information models ", 28th IEEE EMBS Int. Conf, pp. 2614-2617, Sept. 2006

[7] L. Al-Safadi, "Semantic-Based Exchanger for Electronic Medical Record", 3rd Int. Conf. on Convergence and Hybrid Information Technology, 2008

[8] A. Antoniades et al., "Linked2Safety: A secure linked data medical information space for semantically-interconnecting EHRs advancing patients' safety in medical research", 12th Int. IEEE Conf. on Bioinformatics & Bioengineering (BIBE), IEEE, 2012

[9] N. Forgó, M. Góralczyk, and C. Graf von Rex. "Security issues in research projects with patient's medical data", 12th Int. IEEE Conf. on Bioinformatics & Bioengineering (BIBE), IEEE, 2012

[10] Sahay, R., Fox, R., Zimmermann, A., Polleres, A., and Hauswirth, M. (2011). A Methodological Approach for Ontologising and Aligning Health Level Seven (HL7) Applications. In ARES 2011 - Proceedings of the Availability, Reliability and Security for Business, Enterprise and Health Information Systems, Vienna, Austria, August 22-26, pages 102–117. Springer LNCS Series

[11] Brochhausen, M., Spear, A. D., Cocos, C., Weiler, G., Mart´ın, L., Anguita, A., Stenzhorn, H., Daskalaki, E., Schera, F., Schwarz, U., Sfakianakis, S., Kiefer, S., Doerr, M., Graf, N. M., and Tsiknakis, M. (2011). The ACGT Master Ontology and its applications - Towards an ontology-driven cancer research and management system. Journal of Biomedical Informatics, 44(1), 8–25

[12] A. Antoniades et. al., "The effects of applying cell-suppression and perturbation to aggregated genetic data", 12th Int. IEEE Conf. on Bioinformatics & Bioengineering (BIBE), IEEE, 2012.

[13] R. Cyganiak, D. Reynolds, J. Tennison. "The rdf data cube vocabulary." Accessed on November 2012, Available online at: *publishing-statisticaldata.googlecode.com/svn/trunk/specs/src/main/html/cube. html*