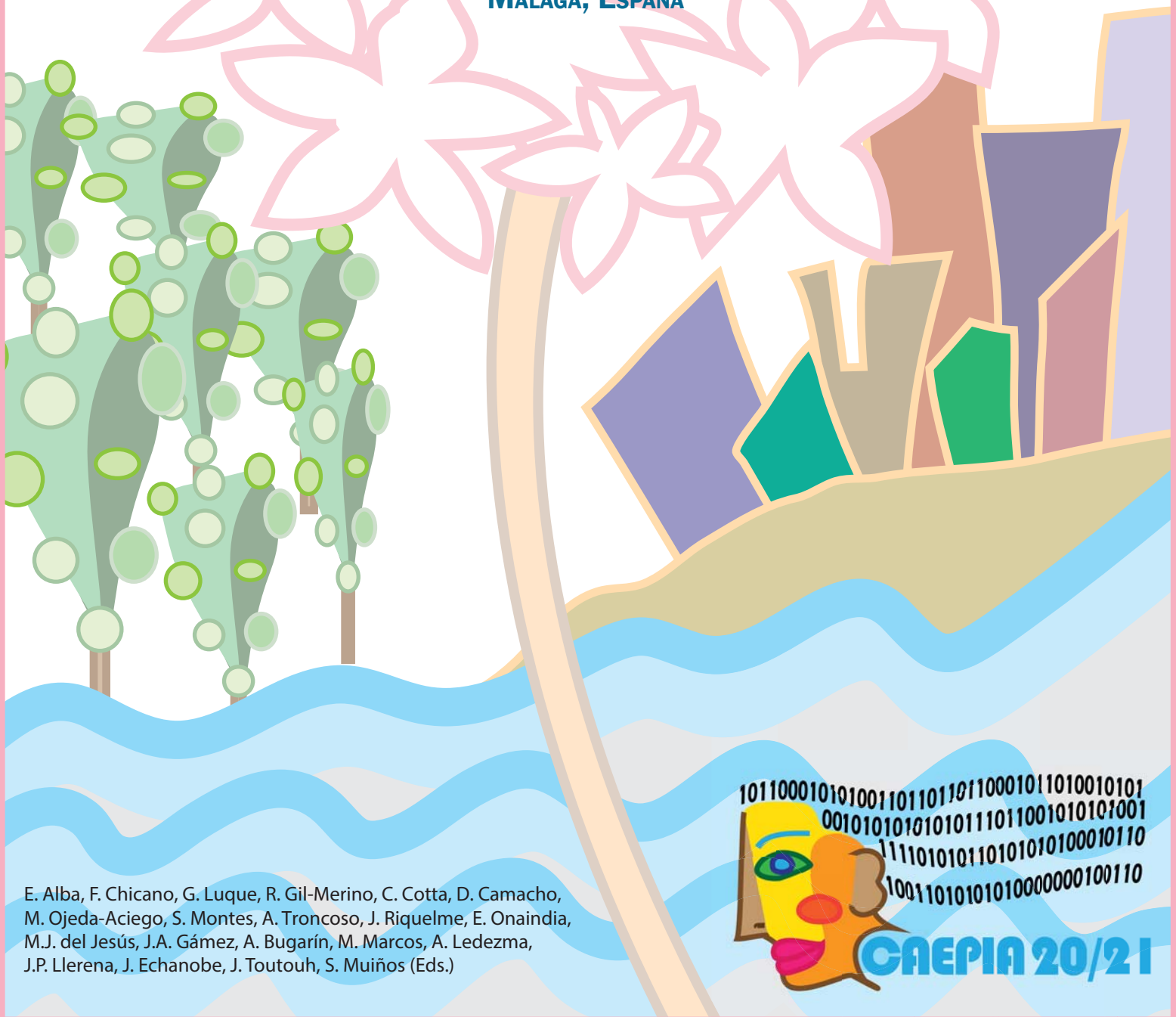
XIX Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 20/21)

Septiembre 22-24

Málaga, España



E. Alba, F. Chicano, G. Luque, R. Gil-Merino, C. Cotta, D. Camacho, M. Ojeda-Aciego, S. Montes, A. Troncoso, J. Riquelme, E. Onaindia, M.J. del Jesús, J.A. Gámez, A. Bugarín, M. Marcos, A. Ledezma, J.P. Llerena, J. Echanobe, J. Toutouh, S. Muiños (Eds.)

XIX Conferencia de la Asociación Española para la Inteligencia Artificial CAEPIA 20/21

22-24 de septiembre de 2021

Málaga, España

Enrique Alba, Francisco Chicano, Gabriel Luque, Rodrigo Gil-Merino, Carlos Cotta, David Camacho, Manuel Ojeda-Aciego, Susana Montes, Alicia Troncoso, José Riquelme, Eva Onaindia, María José del Jesús, José Antonio Gámez, Alberto Bugarín, Mar Marcos, Agapito Ledezma, Juan Pedro Llerena, Javier Echanobe, Jamal Toutouh, Santiago Muiños

Editores

Enrique Alba
Universidad de Málaga
Málaga, España

Francisco Chicano
Universidad de Málaga
Málaga, España

Gabriel Luque
Universidad de Málaga
Málaga, España

Rodrigo Gil-Merino
Universidad de Málaga
Málaga, España

Carlos Cotta
Universidad de Málaga
Málaga, España

David Camacho
Universidad Politécnica de Madrid
Madrid, España

Manuel Ojeda-Aciego
Universidad de Málaga
Málaga, España

Susana Montes
Universidad de Oviedo
Oviedo, España

Alicia Troncoso
Universidad Pablo de Olavide
Sevilla, España

José Riquelme
Universidad de Sevilla
Sevilla, España

Eva Onaindia
Universidad Politécnica de Valencia
Valencia, España

María José del Jesús
Universidad de Jaén
Jaén, España

José Antonio Gámez
Universidad de Castilla La Mancha
Albacete, España

Alberto Bugarín
Universidad de Santiago de Compostela
Santiago de Compostela, España

Mar Marcos
Universitat Jaume I
Castellón de la Plana, España

Agapito Ledezma
Universidad Carlos III de Madrid
Madrid, España

Juan Pedro Llerena
Universidad Carlos III de Madrid
Madrid, España

Javier Echanobe
Universidad del País Vasco
Bilbao, España

Jamal Toutouh
Universidad de Málaga
Málaga, España

Santiago Muiños
Centro Tecnológico AIMEN
Pontevedra, España

ISBN: 978-84-09-30514-8

Presentación de CAEPIA 20/21

Este volumen contiene un conjunto de artículos seleccionados y revisados por pares enviados a CAEPIA 20/21, la XIX Conferencia de la Asociación Española de Inteligencia Artificial, celebrada en Málaga, España, del 22 al 24 de septiembre de 2021. CAEPIA es un evento bienal español bien establecido sobre Inteligencia Artificial (IA) que comenzó en 1985. Ediciones anteriores tuvieron lugar en Alicante, Málaga, Murcia, Gijón, San Sebastián, Santiago de Compostela, Sevilla, La Laguna, Madrid, Albacete, Salamanca y Granada.

CAEPIA es un foro nacional abierto a investigadores de todo el mundo para presentar y discutir sus últimos avances científicos y tecnológicos en IA. Los autores podían optar por cinco tipos de contribuciones: trabajos inéditos de investigación para un volumen en la serie *Lecture Notes in Artificial Intelligence* de Springer, trabajos inéditos de investigación para estas actas, trabajos destacados ya publicados, proyectos de doctorado, desarrollos de aplicaciones móviles y vídeos divulgativos. La conferencia acogió tanto investigación teórica como metodológica, técnica y aplicada.

Dentro de CAEPIA se organizaron varios talleres y congresos federados relacionados con los temas más relevantes de la IA: XX Congreso Español Sobre Tecnologías y Lógica Fuzzy (ESTYLF); XIV Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB); X Simposio de Teoría y Aplicaciones de la Minería de Datos (TAMIDA); y seis talleres. También contamos con un Doctoral Consortium (DC). Este es un foro para que los estudiantes de doctorado interactúen con otros investigadores discutiendo sus planes de trabajo y avances en el doctorado. Como actividad adicional de IA, llevamos a cabo el 4º Concurso de Aplicaciones Móviles con Técnicas de IA, junto con una nueva edición del Concurso de Vídeos de Divulgación de IA.

Todas las actividades anteriores avalan la IA, y nos esforzamos por alcanzar una alta calidad en los artículos científicos, el DC y las competiciones. El programa científico de CAEPIA 20/21 también ofreció una vía para difundir trabajos destacados (Key Works: KW) publicados recientemente en revistas y foros de alto impacto científico. CAEPIA siempre ha tenido como objetivo ser reconocida como una conferencia insignia en IA y, por lo tanto, los artículos fueron revisados por pares. El número total de envíos a CAEPIA 20/21 fue de 186 (en este número no se incluyeron ni DC ni concursos ni presentaciones KW, que suman 83 contribuciones adicionales, y que pasaron por un proceso de evaluación diferente). Los revisores evaluaron la calidad general de los manuscritos presentados, junto con la calidad de la metodología empleada, la solidez de las conclusiones, la importancia del tema, la claridad de la redacción y su organización, entre otros criterios de evaluación. A partir de estas revisiones, los responsables de área, presidentes de congresos y organizadores de talleres y sesiones especiales propusieron un número final de artículos que fueron analizados y aprobados por los editores de este volumen.

CAEPIA 20/21 invitó a dos investigadores de renombre internacional a impartir una charla plenaria. Nuestros dos ponentes plenarios fueron Óscar Cordón (Inteligencia Artificial para Antropología Forense e Identificación Humana) y Yaochu Jin (Optimización Evolutiva Basada en Datos). Nuestra conferencia se celebró como un gran evento dentro de uno aún mayor: la Conferencia Española de Informática (CEDI), que también contó con charlas plenarias muy interesantes.

AEPIA y los organizadores de CAEPIA 20/21 reconocieron las mejores tesis doctorales y artículos originales en eventos federados escritos tanto por investigadores consolidados como por estudiantes. CAEPIA 20/21 también tuvo como objetivo promover la presencia de mujeres en la investigación de IA. Como en ediciones anteriores, el premio Frances Allen reconoció las dos mejores tesis doctorales defendidas por una mujer durante los dos últimos años.

Los editores de este volumen quieren agradecer a las numerosas personas que contribuyeron al éxito de CAEPIA 20/21: autores, miembros de los comités científicos y los comités de programa, ponentes invitados, organizadores de eventos, gestores de medios electrónicos, etc. También, agradecer el trabajo incansable del comité organizador, nuestros patrocinadores (como VRAIN en Valencia), el equipo de Springer y AEPIA por su apoyo.

Por último, pero no menos importante, en nombre de los participantes de CAEPIA 20/21, Enrique Alba (presidente) y Francisco Chicano (responsable de este volumen) dan las gracias a la organización de CEDI, la Universidad de Málaga (sede local de la conferencia) y a toda la comunidad española que trabaja en IA (y sus numerosos colaboradores extranjeros) por hacer de este evento un verdadero éxito.

Enrique Alba
Presidente de CAEPIA 20/21

A platform for the design and execution of clinical data transformation and reasoning workflows

José Alberto Maldonado
Veratech for Health SL
Valencia, Spain
jamaldo@veratech.es

Mar Marcos
Universitat Jaume I
Castellón, Spain
mar.marcos@uji.es

Jesualdo Tomás Fernández-Breis
Universidad de Murcia
Murcia, Spain
jfernand@um.es

Vicente Miguel Giménez-Solano
Veratech for Health SL
Valencia, Spain
vigiso@veratech.es

María del Carmen Legaz-García
IMIB-Arrixaca
Murcia, Spain
mcarmen.legaz@ffis.es

Begoña Martínez-Salvador
Universitat Jaume I
Castellón, Spain
begona.martinez@uji.es

***Abstract*—Effective sharing and reuse of Electronic Health Records (EHRs) requires a high level of semantic interoperability. We have developed a framework to support EHR interoperability based on workflows whose primary components are reusable mappings. The framework supports mappings which take advantage of the best features of EHR standards and semantic web technologies. Based on this framework, we have implemented CLIN-IK-LINKS, a web-based platform that enables users to create, modify and delete mappings, as well as to define and execute clinical data transformation and reasoning workflows. The platform has been applied in two use cases in which real data have been used. This *key work* is an abridged version of the paper published in the journal Comput Meth Prog Bio [1].**

***Index Terms*—Data Workflow, Electronic Health Records, Health Information Interoperability, Semantic Web**

I. Introduction

Electronic Health Record (EHR) systems promised a series of benefits including: savings from data capture and access; information connectivity for stakeholders; and improvement of efficiency, safety, and quality of healthcare through the use of decision support. Some studies have shown that such benefits are still far from being achieved due to the difficulties of sharing and reusing EHR content. Effective EHR sharing and reuse requires a high level of semantic interoperability of EHR data. To achieve this, leading international initiatives and projects have recommended the use of EHR standards and semantic resources.

In the last years we have explored how EHR standards and semantic web technologies could be combined to overcome data heterogeneity in a patient phenotyping scenario [2]. One lesson learned was the need to apply a number of data transformations of different types, and that the organisation of the transformations in a workflow could promote the reuse of the data models and the transformations. This led us to explore how web services could be used to implement such workflows. Another lesson learned was the need to precisely characterise the mappings underlying the data transformations to facilitate reuse.

Based on our previous work, in this paper we describe: (1) a conceptual framework for representing clinical data transformation and reasoning services, (2) the CLIN-IK-LINKS platform that we have implemented to support such framework, and (3) a series of EHR-related interoperability scenarios demonstrating its usefulness.

Work supported by the Spanish Ministry of Economy and Competitiveness and the ERDF programme through grants TIN2014-53749-C2-1-R, TIN2014-53749-C2-2-R, and TIN2017-85949-C2-1-R, and by the Spanish Ministry of Science, Innovation and Universities through grant PTQ2018-009924.

II. Material & Methods

CLIN-IK-LINKS is aimed at facilitating the composition and execution of clinical data transformation workflows to convert EHR data stored in a database into EHR and/or semantic web standards. While focused on data transformation workflows, the platform relies upon a collection of reusable *mappings* with associated *web services*. Mappings are declarative descriptions that specify the operations to be performed to obtain data conforming to a target data model from data stored according to a source data model, designed for specific transformation steps in a clinical scenario. Web services are the essential complement of mappings. Depending on the transformation task they perform, they can be categorised into: normalization, abstraction, semantic publishing and reasoning.

The platform is based on a conceptual framework allowing for the characterisation of mappings beyond the data/information artefacts they consume/produce, At the information model level, the framework distinguishes between the executable mappings (*Mapping*) and the entities that can be mapped with them (*MappableEntity*). The latter includes clinical information models or artefacts in various forms, e.g XML schemas. At this level, a mapping is related to one or more input artefacts and to one output artefact. At the concept level, conceptual mappings (*ConceptualMapping*) and concepts (*Concept*) are the counterpart of mappings and mappable entities, respectively. Using these elements, one can specify not only the connection between a mapping implementation and its definition (via the *IMPLEMENTS* relationship), but

also the relation between the mapping definition and its input and output concepts. Additionally, the connection between an information artefact and the clinical concept it represents can be explicitly stated (with the *REPRESENTS* relationship).

Using the above elements, the platform can accommodate a variety of cases, e.g. several mappings implementing the same conceptual mapping, based or not on the same artefacts, or even multiple conceptual mappings that serve as alternative methods to derive a particular concept.

III. Results

A. The platform

The CLIN-IK-LINKS platform has been implemented in J2EE and uses Neo4J (a graph database) for managing artefacts, concepts and mappings as well as the relationships among them. The overall architecture is diagrammed in Fig. 1. The workflow engine is implemented as a service that offers a REST API. Users interact with the platform by means of a web-based front-end application that allows them to create, modify and delete artefacts and their relationships, as well as to define and execute workflows.

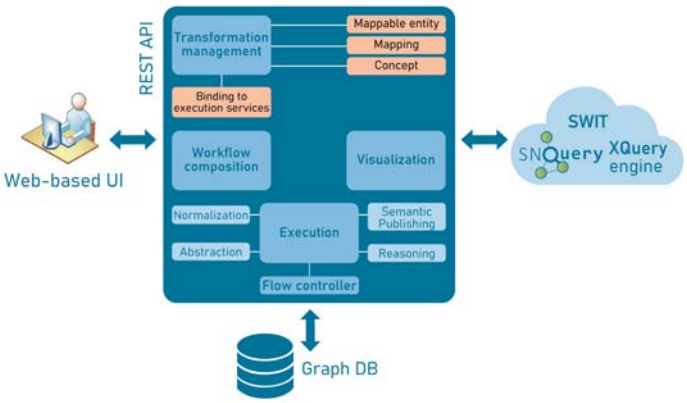


Fig. 1. Overall architecture of the CLIN-IK-LINKS platform.

The set of mappings, artefacts, concepts, and relations among them yields as a result a (disconnected directed acyclic) graph. In the platform, a workflow is defined through a query over this graph. Given either a source or a target artefact, or both, the platform offers a number of paths defining transformation workflows from the source to the target, made up of one or more mappings. For querying purposes, the visual interface presents the user with the set of paths between the source and target artefacts. Such visualisation has proven very useful in composing and understanding complex workflows. The set of paths (or possible transformation workflows) is calculated using the graph traversal capabilities of the Neo4J query language.

B. Use cases

We have implemented two different use cases, available at http://keg.act.uji.es/demoCLINIKLINKS/, which operate with real data. The first use case transforms clinical laboratory test results into an OWL representation in terms of a LOINC-based ontology, and the second one implements two different colorectal cancer screening protocols. In both use cases, we have used existing tools for the definition of mappings and the generation of data transformation programs, namely LinkEHR and SWIT. Note that any mapping tool providing an appropriate web service could have been used instead. In addition, both use cases required elements such as database schemas, clinical information models (e.g. archetypes), and ontologies. These items had to be developed specifically for the use case when they were not available.

IV. Discussion

CLIN-IK-LINKS provides an integrated representation, storage and exploitation of the different sorts of mappings created in the platform. This not only promotes reuse and collaboration among mapping designers, but also facilitates the discovery of mapping compositions/workflows not initially foreseen. In addition, the platform provides flexibility in the design of workflows, since they are composed from queries to the graph of mappings. This allows the user to analyse the different options, and to select the optimal one according to the actual needs.

Several widely used workflow systems are available in biomedical informatics, such as Taverna or Galaxy. However, according to our experience such systems do not provide the necessary features to model workflows with the specific requirements of the healthcare domain, where e.g. there is an increasing number of sophisticated terminological resources. Our platform includes a concept-level layer that not only allows you to describe the relationships between mapping definitions and their input and output concepts, but also to link concepts with appropriate terms from existing terminologies. On the other hand, state-of-the-art workflow systems are oriented towards the definition and exploitation of individual workflows, whereas CLIN-IK-LINKS leverages a graph of mappings which integrates content from all the workflows defined in the platform, thus facilitating sharing and reuse.

V. Conclusions

In this paper we have described CLIN-IK-LINKS, a platform for the composition and execution of clinical data transformation workflows to convert EHR data into EHR and/or semantic web standards. Although the platform was originally conceived to design EHR data transformation workflows for interoperability purposes, we have shown that it can be used to implement applications of a certain significance by themselves, e.g. normalisation and semantic publishing of EHR data.

References

[1] J A Maldonado, M Marcos, J T Fernández-Breis, et al. CLIN-IK-LINKS: A platform for the design and execution of clinical data transformation and reasoning workflows. *Comput Meth Prog Bio*, 197:105616, 2020.

[2] J T Fernández-Breis, J A Maldonado, M Marcos, et al. Leveraging electronic healthcare record standards and semantic web technologies for the identification of patient cohorts. *J Am Med Inform Assoc*, 20(e2):e288–e296, 2013.