

Cochrane PICO Annotator and PICOfinder: Using semantically-enabled evidence to find what works in health care

Chris Mavergames¹, Lorne Becker², Julian Everett³, Paul Wilton³, Anthony Hughes³

¹The Cochrane Collaboration, Head of Informatics & Knowledge Management, Freiburg, Germany

`CMavergames@cochrane.org`

²Cochrane Innovations, Tampa, USA

`lornebecker@gmail.com`

³DataLanguage (UK) Ltd, Leatherhead, United Kingdom

`{julian.everett,paul.wilton,anthony.hughes}@datalanguage.com`

Abstract. This software demo will show the functionality of Cochrane's PICO Annotator tool, which was developed for annotating evidence in clinical trial reports and systematic reviews, and the Cochrane PICOfinder, a discovery and visualisation tool for filtering evidence and underlying data by PICO criteria. Both tools are part of the Cochrane Collaboration's Linked Data Platform. Details about the underlying model, technology, and software design will be discussed together with plans for the future development and deployment of these tools across Cochrane and out into the domain of clinical guidelines, electronic and clinical decision support systems, and policy and guideline development. Lessons learned from the implementation of the OHDSI clinical terminology and ontology mappings (<http://www.ohdsi.org/>) of SNOMED CT, RxNorm, WHO ATC, MedDRA, and other vocabularies will also be briefly shared.

Keywords: evidence-based health care, RDF, OWL, PICO, Cochrane, SNOMED, Medra, RxNorm

1 Introduction

Cochrane is the world leader in the production of systematic reviews of healthcare interventions. Finding out what works (or harms) in health is a massive and time-consuming challenge of synthesising data from clinical trials and other sources of data. Once compiled and analysed, systematic reviews are published as lengthy, cumbersome PDF and HTML representations. The underlying trial data is buried in trial reports published in scientific journals and other grey literature formats. Historically, navigating what's available to answer a specific clinical question, usually framed as a PICO for the Population (age, gender, condition), the Intervention(s) in question, the Comparator (placebo, another intervention), and the Outcome(s) being measured, involved manually sifting through search results and then documents. To address this, over the last two years a suite of linked data UI components, REST APIs and data repositories have been developed by Cochrane in conjunction with their linked data technology partner DataLanguage, which deliver the ability to capture and query rich semantic micrographs against clinical evidence. This paper consists of application notes which summarise the technical and data architecture of the new platform.

2 Application Notes

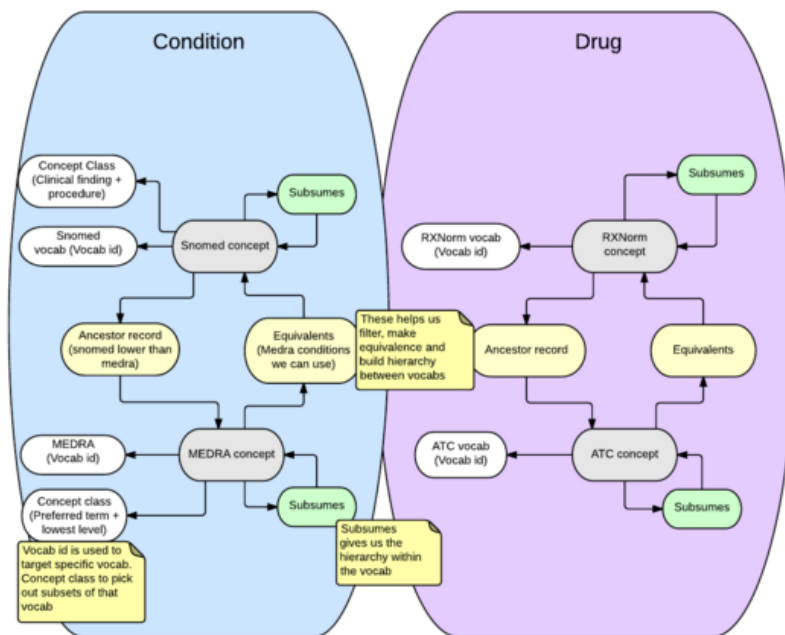
2.1 Technical Architecture

The PICO Annotator is a JQuery plugin which can be dropped into any web application container, and which simply takes a constructor parameter of the URI which identifies the content document or fragment being annotated. For the purposes of annotating systematic reviews within Cochrane, the PICO Annotator is currently being hosted within a review browser web container and will also shortly be integrated into the Central Register of Studies user interface. The Annotator plugin implements its own library of PICO user control JQuery plugins, which allow UI consumers to integrate at either the annotator or individual user control level. The controls

directly broker all calls to the underlying JSON-LD REST APIs, which are implemented as JSR-311 compliant Java web services hosted in Tomcat containers. The two primary APIs are concerned with exposing the knowledge graph (Concepts API) and reading/writing content annotations (Annotations API). The Annotations API reifies the submitted annotations and injects provenance information extracted from auth tokens generated via the OAuth2 security layer. The annotations are then persisted into an instance of Ontotext's GraphDB triple store, which is configured with connectors to materialise the relevant subgraphs into an Elastic Search cluster, to power discoverability and visualisation tools of which PICO finder is the first application instance.

2.2 Data Architecture

Previous work by the Cochrane Collaboration [1] confirmed the value to the organisation of describing clinical evidence using structure linked data. Building on that, domain models for PICO descriptors and Content Annotations have been iteratively developed and manifested as OWL ontologies. Reference data to instantiate the models was then generated by selecting and curating public domain datasets and/or dataset fragments which best fulfil the organisation's use cases for each dimension of the PICO graph. Following standard software architecture principles, this created a separation between public semantic interface (i.e. public domain identifier mappings) and internal implementation (i.e. Cochrane linked data identifiers) which has greatly aided data governance, versioning and deployment tasks. The vocabs used currently are as follows.



In addition, SNOMED is being used as a source for Procedures and Materials, MeSH for Population Age and Sex, and Outcomes are based on the work from the Comet Initiative.

References

1. Mavergames C, Becker L, Oliver, S. Systematic Reviews as an interface to the web of (trial) data: Using PICO as an ontology for knowledge synthesis in evidence-based healthcare research Proceedings of the European Semantic Web Conference, Montpellier, France. June 2013 <http://ceur-ws.org/Vol-994/paper-02.pdf>