Pharmacological Class Data Representation in the Web Ontology Language (OWL)

Qian Zhu[1]*, Cui Tao[2]
[1]Department of Information System, University of Maryland, Baltimore County, MD, USA
[2]School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA

INTRODUCTION

Dozens of drug terminologies and resources capture drug and/or drug class information; they range greatly in their coverage and their adequacy of representation. In this study, we generated a standardized Pharmacological Class Profile Ontology, named PCPO, which integrates multiple drug resources in the Web Ontology Language (OWL). PCPO will not only present a large volume of drug data in a well-organized formal form, OWL with possible inference capability, but also potentially support computational drug repurposing application development.

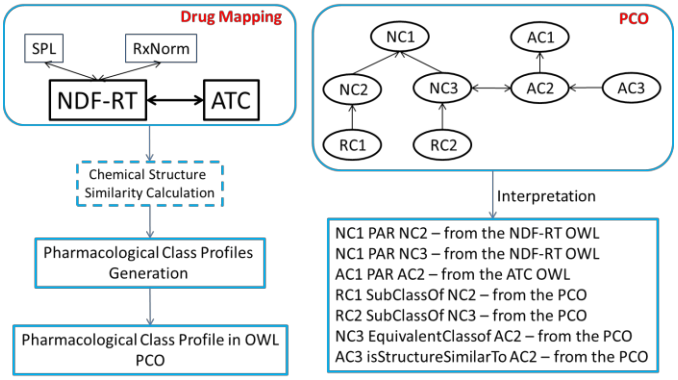

Fig. 1. Workflow for Pharmacological Class Profile Ontology (PCPO) Generation. A snapshot of the PCPO is given on the right side of this figure, along with interpretations for the relations presented. AC indicates ATC concept; NC, NDF-RT concept; OWL, Web Ontology Language; RC, RxNorm concept; PAR, isPatientOf.

METHODS

A. PCPO Development

We generated the PCPO by linking drug concepts from different drug resources in two layers, drug class layer and individual drug layer. For drug concept mappings among RxNorm, SPL, and NDF-RT, we directly extracted these mappings from 2 RxNorm files, RXNCONSO and RXNREL. Figure 1 shows the workflow of the PCPO generation, along with relationships expressed in the PCPO.

B. PCPO Representation in OWL

We applied an ontology-based method to represent the mappings among heterogeneous drug terms/concepts in the PCPO with the existing drug ontologies. For ATC, we adopted the ontology developed by Croset et al [1]; for NDF-RT, we used the ontology released by the NIH [2], and for RxNorm, we used the OWL/RDF file exported from the NCBO BioPortal [3]. There is no OWL representation of SPL at this time, so we have not represented the mappings to SPL semantically in the PCPO. We defined a new OWL class for each unique drug term in the mappings generated previously. For mappings between ATC and NDF-RT, we first created a new OWL class with a PCPO URI (Uniform Resource Identifier) for each unique ATC drug entity. This OWL class will then be defined as an equivalent class of the corresponding mapped NDF-RT OWL classes of this ATC class to indicate the mapping between ATC and NDF-RT. For NDF-RT to RxNorm mappings, we specified that an RxNorm concept (OWL class) is a subclass of a mapped NDF-RT concept (OWL class) accordingly. This way, the PCPO contains higher-level drug classifications derived from both ATC and NDF-RT and lower-level information about individual clinical drugs from RxNorm. We represented structural similarity between pairs of drugs in the PCPO. Because OWL only supports binary relationships, we followed the World Wide Web Consortium guidelines [4] and introduced a new class called "Similarity" to represent the target drug and the corresponding similarity score.

RESULTS

A total of 5,717 ATC drug entities are included in this study, which corresponds to 4,483 distinct ATC terms. That is, one drug can be categorized into multiple therapeutic classes. Of the 48,266 NDF-RT concepts, 34,011 concepts were used in this study, consisting of 15,857 VA products, 486 VA classes, 9,960 Chemical/Ingredients, 7,184 Generic Ingredient Combinations, and 524 EPCs.

The PCPO, which can be accessed at https://sourceforge.net/projects/PCPO/, currently contains 58,241 OWL classes, 98,677 subclass axioms, and 21,917 equivalent class axioms. It has defined 178,838 axioms with 120,594 logical axioms.

DISCUSSION AND CONCLUSIONS

We successfully integrated NDF-RT, ATC, RxNorm, and SPL and built PCPO for representing drug and drug class entities. In addition, the ontology was expanded from a chemical structure perspective by introducing chemical similarity calculation. PCPO supports automated reasoning, which can ultimately be applied for drug repositioning by identifying alternative drugs for a particular disease through drug-drug associations inference. To expand the coverage and usage of PCPO, other drug terminological resources and drug interaction information will be integrated in the future.

ACKNOWLEDGMENTS

We thank Pradip P. Kanjamalafor IT support. This work was partially supported by the Cancer Prevention & Research Institute of Texas (CPRIT R1307).

REFERENCES

[1] Croset S, Hoehndorf R, Rebholz-Schuhmann D. OWL representation of the anatomical therapeutic chemical classification system and mapping to DrugBank. Paper presented at: 4th Workshop on Ontologies in Biomedicine and Life Sciences; 2012 Sep 27-28; Dresden, Germany.
[2] Index of /ftp1/NDF-RT [Internet]. [cited 2013 Dec 23]. Available from: http://evs.nci.nih.gov/ftp1/NDF-RT/.
[3] Noy NF, Shah NH, Whetzel PL, Dai B, Dorf M, Griffith N, et al. BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res. 2009 Jul;37(Web Server issue):W170-3. Epub 2009 May 29.
[4] Noy N, Rector A. Defining N-ary relations on the semantic web [Internet]. W3C Working Group Note 12 April 2006. c2006 [cited 2013 Dec 23]. Available from: http://www.w3.org/TR/swbp-n-aryRelations/.

Pharmacological Class Data Representation in the Web Ontology Language (OWL)

Qian Zhu[1], Cui Tao[2]

[1]Department of Information System, University of Maryland, Baltimore County, MD, USA
[2]School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA

INTRODUCTION

Dozens of drug terminologies and resources capture drug and/or drug class information; they range greatly in their coverage and their adequacy of representation.

In this study, we generated a standardized Pharmacological Class Profile Ontology, named PCPO, which integrates multiple drug resources in the Web Ontology Language (OWL). PCPO will not only present a large volume of drug data in a well-organized formal form, OWL with possible inference capability, but also potentially support computational drug repurposing application development.

ACKNOWLEDGEMENTS

We thank Pradip P. Kanjamalafor IT support. This work was partially supported by the Cancer Prevention & Research Institute of Texas (CPRIT R1307).

REFERENCES

1.Croset S, Hoehndorf R, Rebholz-Schuhmann D. OWL representation of the anatomical therapeutic chemical classification system and mapping to DrugBank. 4th Workshop on Ontologies in Biomedicine and Life Sciences; 2012 Sep 27-28.
2. NDF-RT: http://evs.nci.nih.gov/ftp1/NDF-RT/.
3. Noy NF, Shah NH, Whetzel PL, Dai B, Dorf M, Griffith N, et al. BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res. 2009.
4. Noy N, Rector A. Defining N-ary relations on the semantic web [Internet]. W3C Working Group Note 12 April 2006. Available from: http://www.w3.org/TR/swbp-n-aryRelations/.

METHODS

A.PCPO Development We generated the PCPO by linking drug concepts from different drug resources in two layers, drug class layer and individual drug layer. For drug concept mappings among RxNorm, SPL, and NDF-RT, we directly extracted these mappings from 2 RxNorm files, RXNCONSO and RXNREL. Figure 1 shows the workflow of the PCPO generation, along with relationships expressed in the PCPO.

B. PCPO Representation in OWL

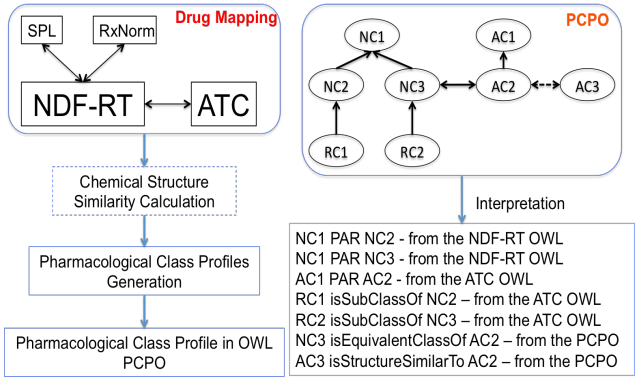

Fig. 1. Workflow for PCPO Construction

We applied an ontology-based method to represent the mappings among heterogeneous drug terms/concepts in the PCPO with the existing drug ontologies. For ATC, we adopted the ontology developed by Croset et al [1]; for NDF-RT, we used the ontology released by the NIH [2], and for RxNorm, we used the OWL/RDF file exported from the NCBO BioPortal [3].

We defined a new OWL class for each unique drug term in the mappings generated previously. For mappings between ATC and NDF-RT, we first created a new OWL class with a PCPO URI (Uniform Resource Identifier) for each unique ATC drug entity. This OWL class will then be defined as an equivalent class of the corresponding mapped NDF-RT OWL classes of this ATC class to indicate the mapping between ATC and NDF-RT. For NDF-RT to RxNorm mappings, we specified that an RxNorm concept (OWL class) is a subclass of a mapped NDF-RT concept (OWL class) accordingly. This way, the PCPO contains higher-level drug classifications derived from both ATC and NDF-RT and lower-level information about individual clinical drugs from RxNorm. We represented structural similarity between pairs of drugs in the PCPO. Because OWL only supports binary relationships, we followed the World Wide Web Consortium guidelines [4] and introduced a new class called “Similarity” to represent the target drug and the corresponding similarity score.

RESULTS

A. 5,717 ATC drug entities are included in this study, which corresponds to 4,483 distinct ATC terms. Of the 48,266 NDF-RT concepts, 34,011 concepts were used in this study, consisting of 15,857 VA products, 486 VA classes, 9,960 Chemical/Ingredients, 7,184 Generic Ingredient Combinations, and 524 EPCs.
B. The PCPO, which can be accessed at https://sourceforge.net/projects/PCPO/, currently contains 58,241 OWL classes, 98,677 subclass axioms, and 21,917 equivalent class axioms. It has defined 178,838 axioms with 120,594 logical axioms.

DISCUSSION AND CONCLUSION

- We successfully integrated NDF-RT, ATC, RxNorm, and SPL and built PCPO for representing drug and drug class entities.
- PCPO was expanded from chemical perspective by introducing chemical similarity calculation. PCPO supports automated reasoning, which can ultimately be applied for drug repositioning by identifying alternative drugs for a particular disease through drug-drug associations inference.
- To expand the coverage and usage of PCPO, other drug terminological resources and drug interaction information will e integrated in the future.