



Maestría en Ingeniería en Sistemas y Cómputo Inteligente

**A web-based interface to retrieve
information of Onto4AIR instances, an
ontology for institutional repositories**

*Ana Laura Lezama Sánchez
María Auxilio Medina Nieto
Jorge de la Calleja Mora
Antonio Benitez Ruiz*



A web-based interface to retrieve information of Onto4AIR instances, an ontology for institutional repositories

Ana Laura Lezama Sánchez, María Auxilio Medina Nieto,
 Jorge de la Calleja Mora, Antonio Benitez Ruiz
 Departamento de Posgrado. Maestría en Ingeniería en Sistemas y Cómputo Inteligente
 {ana.lezama4301, maria.medina, jorge.delacalleja, antonio.benitez}@uppuebla.edu.mx
 Tercer Carril del Ejido Serrano S/N, San Mateo Cuanalá, Juan C. Bonilla, Puebla, México

1. Introduction

An institutional repository (IR) is a centralized digital platform that stores academic and scientific production. In México, organizations such as the *Repositorio Nacional (NR)*, the *Consejo Nacional de Ciencia y Tecnología (CONACYT)* and the *Red Mexicana de Repositorios Institucionales (REMERI)*, promote open access in IRs as part of a national strategy to build an open science policy.

The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is an interoperability standard between repositories, that uses the Dublin Core metadata format (DC) to describe contents. The use of these metadata is proposed at a recommendation level, that beneficiates flexibility and adoption, however, causes ambiguity to retrieve shared data.

Onto4AIR is a formal machine-readable model designed to reduce that ambiguity, this represents domain and operational knowledge of IRs [1]. This document describes a web-based interface that allows users to retrieve information stored as instances of this ontology, that is, bibliographic data that belongs to 114 thesis from different IRs of the state of Puebla and of the Universidad Politécnica de Puebla.

2. Objectives

2.1. General objective

Represent the bibliographic data of thesis from RIs of the REMERI network as linked open data that can be queried and published in a web interface.

2.2. Specific objective

- Design an accessible interface via web that allows users to retrieve information about instances of the Onto4AIR ontology.

3. Methods

Figure 1 shows the steps to populate the Onto4AIR ontology. Figure 2 and 3 illustrates the tasks of the first and third steps.

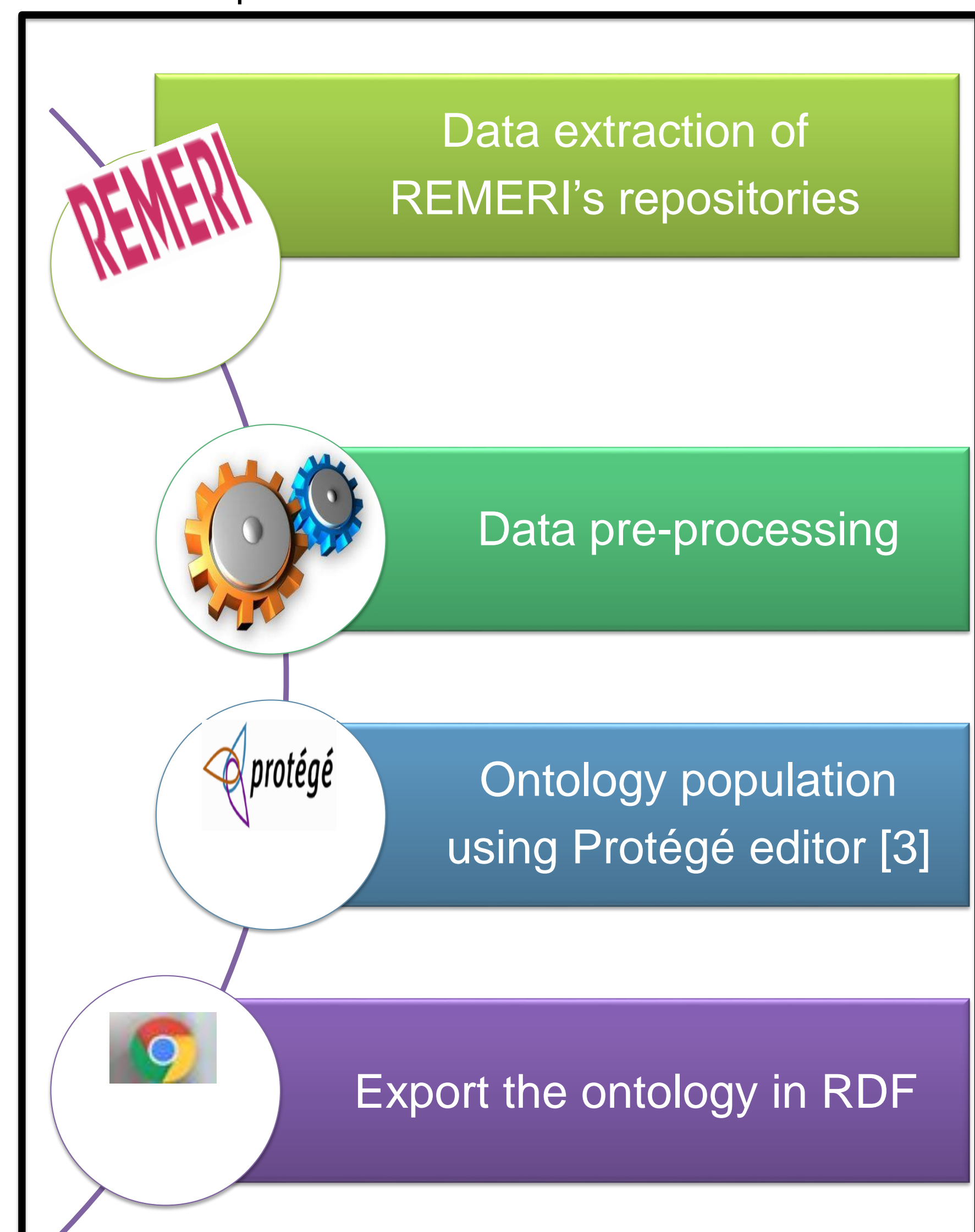


Figure 1. Methodology to populate the Onto4AIR ontology

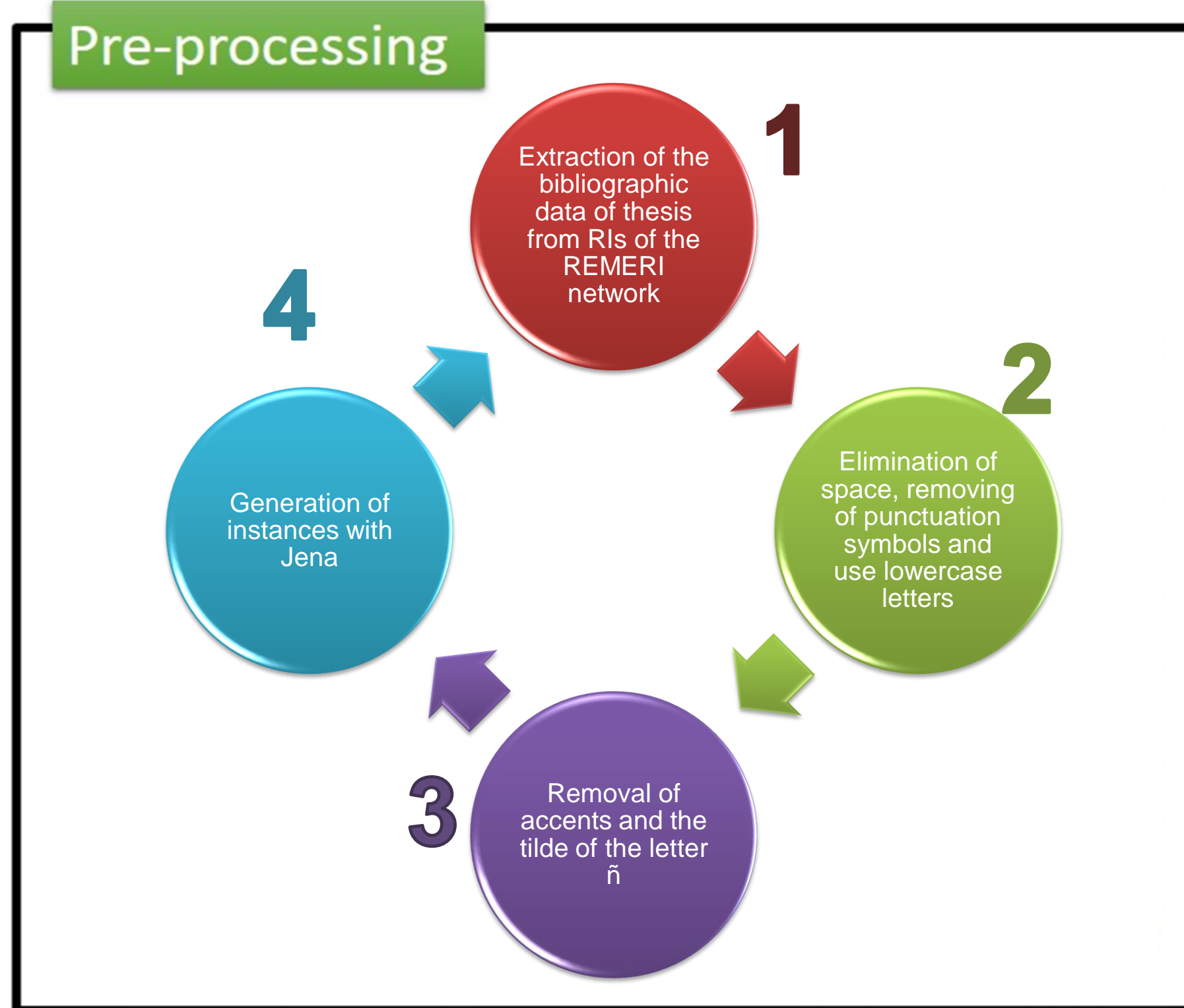


Figure 2. Tasks to process data from IRs.

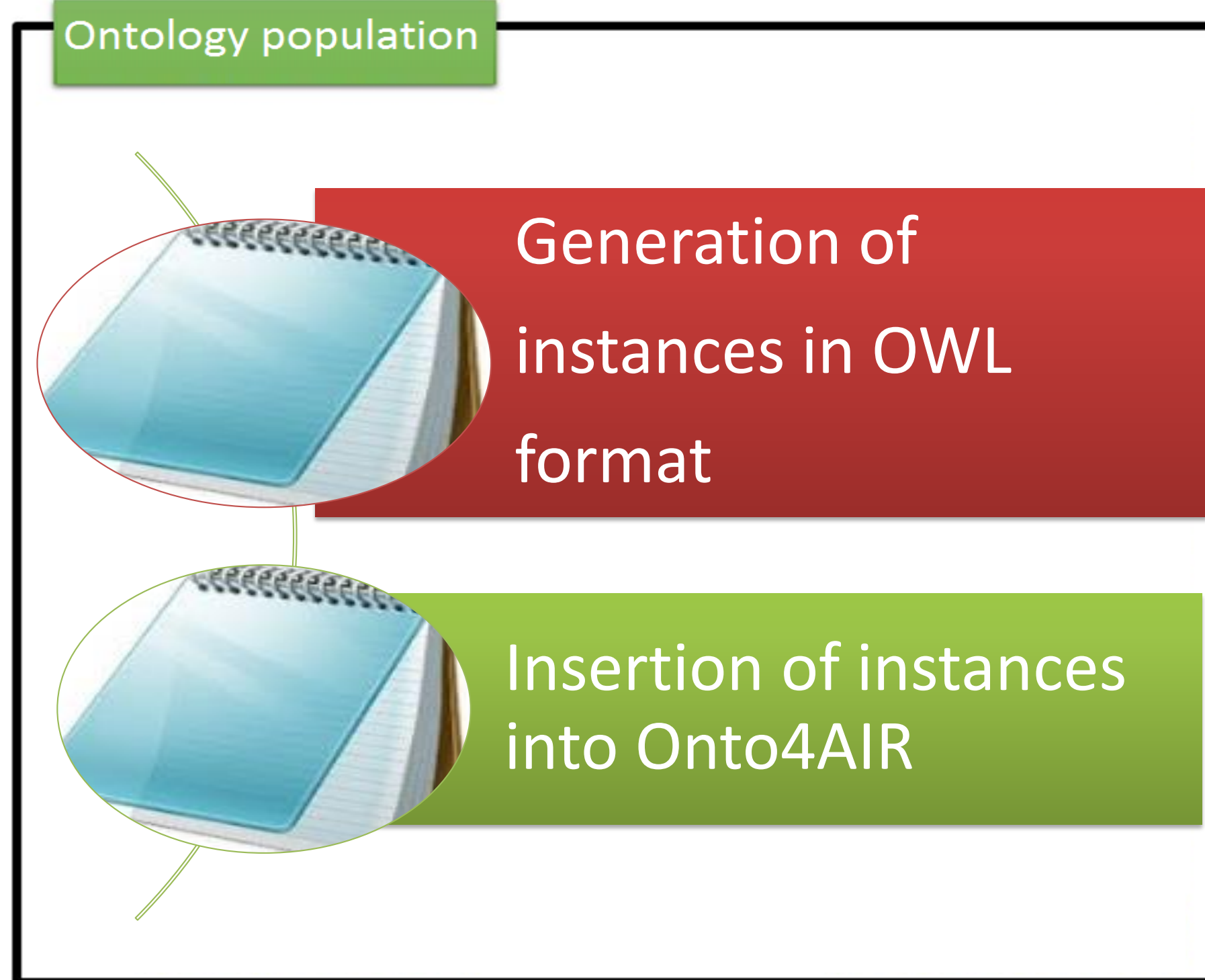


Figure 3. Tasks to populate the Onto4AIR ontology

The Protégé editor is used to check syntax and the execution of SPARQL queries that were previously considered as competence questions by potential users. Figure 4 shows the open-source software tools used to implement this interface.

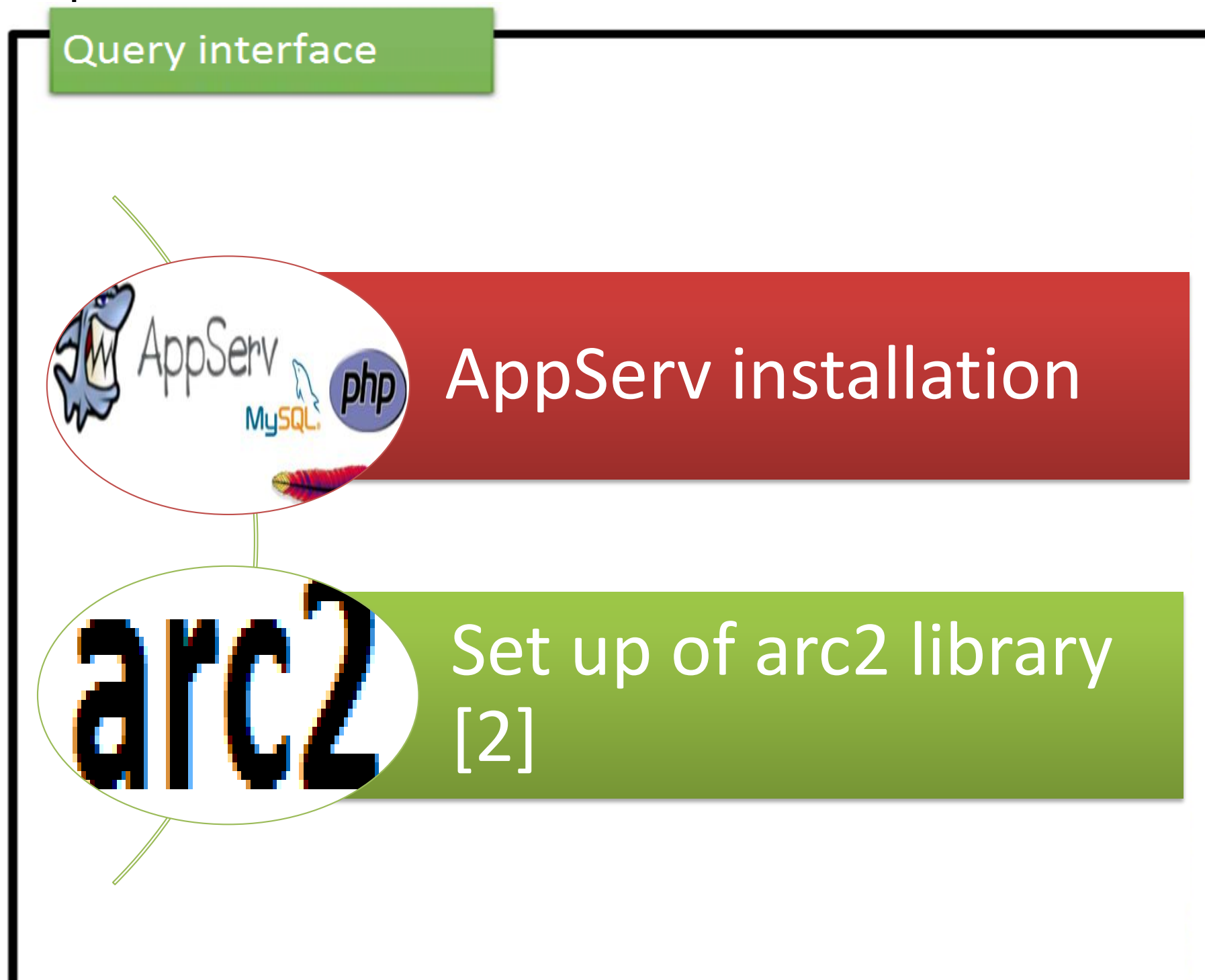


Figure 4. Software tools used to implement the web-based interface.

4. Results

Figures 5-8 show screenshots of the web-based interface. Note that queries are considered *pre-determined* if they are introduced by using forms and text fields or *free queries* when users introduce a complete SPARQL expression.

localhost/InterfazD/busqueda_autor.php

Search by author

Instructions: Write the first name and second name of the author

First name*: medina
 Second name: nieto
 Names: maria auxilio

Search Delete

*Required fields

Figure 5. Form to search thesis by author

localhost/InterfazD/resultados_consulta.php

Results by author

Author: medina nieto maria auxilio

Title	Subject	Institution
ontoair construction of lightweight ontologies to support information retrieval from multiple collections of documents	ciencias de la computacion	universidad de las americas puebla

Save results Regresar

Figure 6. Result of a query by author

Figure 7 shows a query to retrieve science thesis registered before 2017-12-25 whose author's last name is "medina".

```

  SPARQL query
  SELECT ?Titulo ?Autor ?Tema ?editor ?Fecha WHERE{
  ?x ni:Titulo ?Titulo.
  ?x ni:Autor ?Autor.
  ?x ni:editor ?editor.
  ?x ni:Tema ?Tema.
  ?x ni:Fecha ?Fecha
  FILTER regex(?Autor, "medina").
  FILTER regex(?Tema, "ciencias").
  FILTER (?Fecha < "2017-12-25"^^xsd:dateTime)}
  
```

Figure 7. Insertion of a free query

Results

Author	Title	Subject	Institution	Fecha
martinez medina lourdes angelica	query optimization using casebased reasoning	ciencias de la computacion	universidad de las americas puebla	2010-01-25T00:00:00Z
medina nieto maria auxilio	ontoair construction of lightweight ontologies to support information retrieval from multiple collections of documents	ciencias de la computacion	universidad de las americas puebla	2008-09-19T00:00:00Z

Saved results Return

Figure 8. Results of a free query

5. Conclusions

The implementation of the web-based interface allows that non-expert and expert users can access to data stored in the Onto4AIR ontology by using semantic technologies; these technologies enables data integration.

Between the benefits of semantic technologies is the possibility of guarantee logical consistency of data by means of reasoners and the capacity to share explicit and implicit knowledge stored in the ontology between users and machines. As future work, a usability assessment of the interface is proposed.

Acknowledgements

This research is supported by CONACYT Scholarship 788155. PNPC.

References

[1] Medina N. M. A., Sánchez J.A, Cervantes O., Medina R. R.C., De la Calleja, M. J., Benítez A. Representación semántica de conocimiento operativo y de dominio para repositorios institucionales. 2017. Registro público del derecho de autor. Número: 03-2017-042511235500-01. Fecha de concesión: 26 de Abril del 2017.

[2] Shekhar, M. 2013. Semantic Web Search based on Ontology Modeling using Protege Reasoner. *arXiv preprint arXiv:1305.5827*.

[3] Noy, N. F., M. Crubezy, et al. 2003. Protege-2000: an open-source ontology-development and knowledgeacquisition environment. *AMIA Annual Symposium Proceedings*: 953.



Este material se distribuye bajo los términos de la
Licencia 2.5. de Creative Commons

2018