



Data acquisition for analytical platforms: Automating scientific workflows and building an open database platform for chemical analysis metadata

Sana Tfaili, Diem Bui Thi, Karima Rafes, Ali Tfayli, Arlette Baillet-Guffroy, Cécile Germain, Pierre Chaminade

► To cite this version:

Sana Tfaili, Diem Bui Thi, Karima Rafes, Ali Tfayli, Arlette Baillet-Guffroy, et al.. Data acquisition for analytical platforms: Automating scientific workflows and building an open database platform for chemical analysis metadata. *Chimie X*, Jan 2016, Namur, Belgium. <hal-01423371>

HAL Id: hal-01423371

<https://hal.inria.fr/hal-01423371>

Submitted on 29 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DATA ACQUISITION FOR ANALYTICAL PLATFORM

AUTOMATING SCIENTIFIC WORKFLOWS AND BUILDING AN OPEN DATABASE PLATFORM FOR CHEMICAL ANALYSIS METADATA

Sana Tfaili¹, Diem Bui Thi¹, Karima Rafes², Ali Tfayli¹, Arlette Baillet-Guffroy¹, Cécile Germain^{3,4} and Pierre Chaminade¹

¹Lip(Sys)²- Chimie Analytique Pharmaceutique³, Univ. Paris-Sud, Université Paris-Saclay, F-92290 Châtenay-Malabry, France.

³(FKA EA4041 Groupe de Chimie Analytique de Paris-Sud)

²Bordercloud 245, quai de la Bataille de Stalingrad, F-92130 Issy-les-Moulineaux, France.

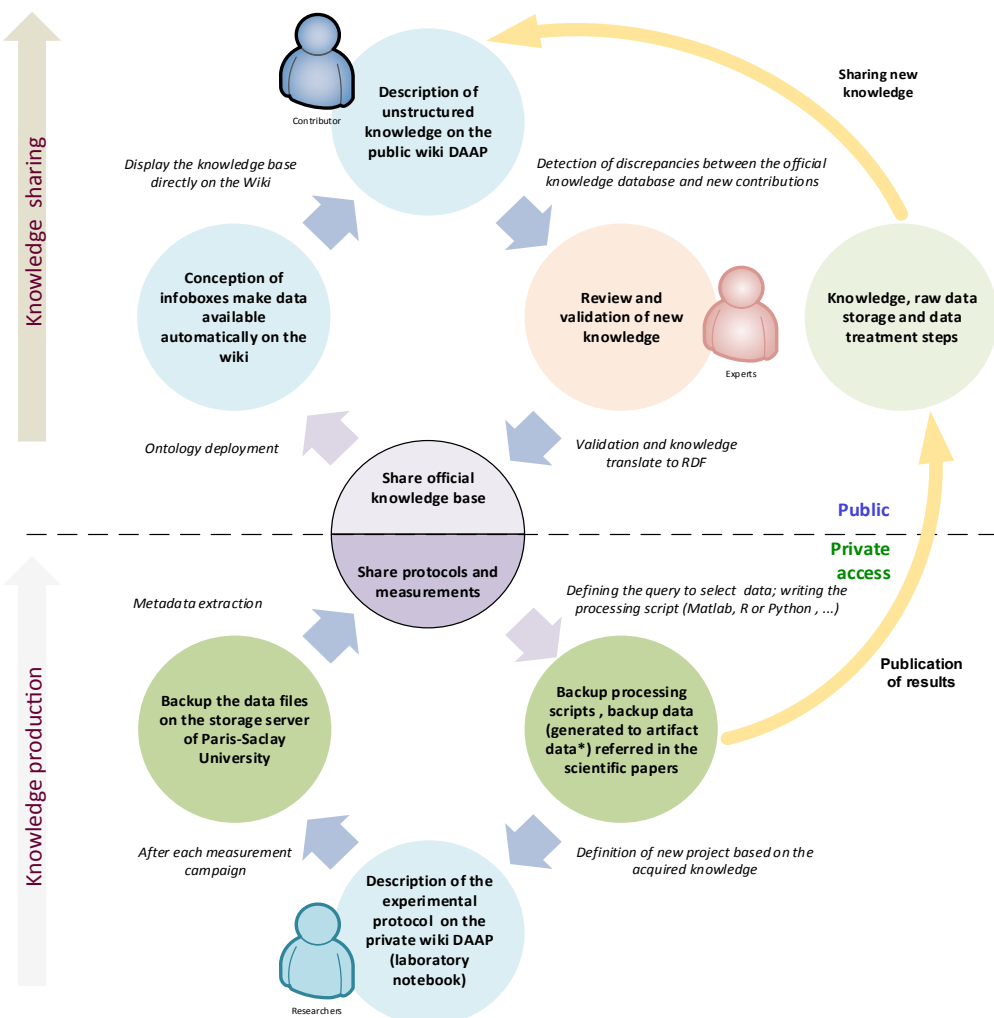
³LRI- Laboratoire de Recherche en Informatique, Bât 650, rue Noetzlin, F-91190 Gif-sur-Yvette, France.

⁴LAL- Laboratoire de l'Accélérateur Linéaire, Bât 200, voie de la faculté, F-91440 Bures-sur-Yvette, France.

TARGET: The project DAAP "Data Acquisition for Analytical Platform" engages the collaboration in the Analytical Chemistry research community. The platform allows shared data among scientists by accessing private Linked Database Platform (LDP) or public Linked Database Platform. The first step is to reference the resources available to researchers using a wiki, to define the ontology by the domain scientists and to benefit from Linked Data technologies.

PERSPECTIVE: open data access, relevant information extraction, data fusion

Efficient knowledge production with the scientific community



Step 1: Protocols are described on the private wiki DAAP by the researcher.

Step 2: Experimental data is stored within the University server.

Step 3: A plugin extracts the metadata from experimental data, imports the metadata into the RDF database and links metadata to the domain ontology.

Step 4: Researchers can simultaneously query the data in the database according to the domain ontology and find the related experimental data.

Step 5: After publishing a paper, data and knowledge are available on the public wiki DAAP.

Step 6: Experts organize a meeting to review and validate the shared new knowledge and the shared resources according to the domain ontology.

Step 7: Researchers can query the domain ontology using Linked Data technologies.

Step 8: The ontology is presented as an infobox within wiki DAAP pages. In case of disagreement, users can change it. The divergence in ontology will then be discussed in step 6.

Keys :

Using a MediaWiki with LinkedWiki extension

RDF database with shared access via SPARQL

Storage server in the public network (http)

SPARQL query

Using TopBraid software to write in RDF knowledge and SHACL constraints

RDF database with private access via SPARQL

Network-Attached Storage (NAS) with restricted access

*The artifact data: The data is computed from the experimental data.

lip(sys)²
Lip(Sys) Center for Data Science

BorderCloud

sana.tfaili@u-psud.fr

Contacts:

sana.tfaili@u-psud.fr

karima.rafes@bordercloud.com