

# SMART Protocols: SeMAntic RepresenTation for Experimental Protocols

Olga Giraldo<sup>1</sup>, Alexander García<sup>2</sup>, and Oscar Corcho<sup>1</sup>

<sup>1</sup> Ontology Engineering Group, Universidad Politécnica de Madrid, Spain  
{ogiraldo, ocorcho}@fi.upm.es

<sup>2</sup> Linkingdata I/O LLC, Fort Collins, Colorado, USA  
alexgarcia@gmail.com

**Abstract.** Two important characteristics of science are the “*reproducibility*” and “*clarity*”. By rigorous practices, scientists explore aspects of the world that they can reproduce under carefully controlled experimental conditions. The clarity, complementing reproducibility, provides unambiguous descriptions of results in a mechanical or mathematical form. Both pillars depend on well-structured and accurate descriptions of scientific practices, which are normally recorded in experimental protocols, scientific workflows, etc. Here we present SMART Protocols (SP), our ontology-based approach for representing experimental protocols and our contribution to clarity and reproducibility. SP delivers an unambiguous description of processes by means of which data is produced; by doing so, we argue, it facilitates reproducibility. Moreover, SP is thought to be part of e-science infrastructures. SP results from the analysis of 175 protocols; from this dataset, we extracted common elements. From our analysis, we identified document, workflow and domain-specific aspects in the representation of experimental protocols. The ontology is available at <http://purl.org/net/SMARTprotocol>

**Keywords:** experimental protocol, ontology, in vitro workflow, reproducibility.

## 1 Introduction

Scientific experiments often bring together several technologies at in vivo, in vitro and sometimes in silico levels. Moreover, the biomedical domain relies on complex processes, comprising hundreds of individual steps usually described in experimental protocols. An experimental protocol is a sequence of tasks and operations executed to perform experimental research. The protocols often include equipment, reagents, critical steps, troubleshooting, tips and all the information that facilitates reusability. Researchers write the protocols to standardize methods, to share these documents with colleagues and to facilitate the reproducibility of results.

Although reproducibility, central to research, depends on well-structured and accurately described protocols, scientific publications often lack sufficient information when describing the protocols that were used. For instance, there is

ambiguity in the terminology as well as poor descriptions embedded within a heterogeneous narrative. There is the need for a unified criterion with respect to the syntactic structure and the semantics for representing experimental protocols.

Here we present SMART Protocols (henceforth SP), our ontology-based approach for representing experimental protocols. SP aims to formalize the description of experimental protocols, which we understand as domain-specific workflows embedded within documents. SP delivers a structured workflow, document and domain knowledge representation written in OWL DL. For the representation of document aspects we are extending the Information Artifact Ontology (IAO).<sup>1</sup> The representation of executable aspects of a protocol is captured with concepts from P-Plan Ontology (P-Plan) [1]; we are also reusing EXPO [2], EXACT [3] and OBI [4]. For domain knowledge, we rely on existing biomedical ontologies. SP results from the analysis of 175 experimental protocols gathered from several sources. From this dataset, we extracted common elements and evaluated whether those protocols could be implemented.

Our main assumption is that *“experimental protocols are fundamental information structures that should support the description of the processes by means of which results are generated in experimental research”*. Hence our approach should allow answering questions such as: *Who is the author of the protocol? What is the application of the protocol? What are the reagents, equipment and/or supplies used? What is the estimated time to execute a protocol? Which samples have been tested in a protocol?* This paper is organized as follows: Section 2 presents related works, Section 3 describes the methodology stages to develop the SP ontology, section 4 shows the results and ontology evaluation. Finally Section 5 provides discussion and conclusions.

## Related Work

In an effort to address the problem of inadequate methodological reporting, the MIBBI<sup>2</sup> project brings under one umbrella most of these projects. The ISA-TAB also illustrates work in this area; it delivers metadata standards to facilitate data collection, management and reuse [5].

The Ontology for Biomedical Investigations (OBI)<sup>3</sup> aims to model the design of investigations, including the protocols, materials used and the data generated. OBI has key classes for the description of experiments, namely: `obi:investigator`, `obi:instrument`, `obi:biomaterial` entity. The generic ontology of scientific experiments (EXPO)<sup>4</sup> aims to formalize domain-independent knowledge about the planning, execution and analysis of scientific experiments. This ontology includes the class `expo:ExperimentalProtocol` and defines some of its properties: `expo:has_applicability`, `expo:has_goal`,

---

<sup>1</sup> <https://code.google.com/p/information-artifact-ontology/>

<sup>2</sup> <http://mibbi.sourceforge.net/portal.shtml>

<sup>3</sup> [http://obi-ontology.org/page/Main\\_Page](http://obi-ontology.org/page/Main_Page)

<sup>4</sup> <http://expo.sourceforge.net/>

`expo:has_plan`. EXACT suggests a meta-language for the description of experiment actions and their properties.

Recently, PLOS ONE in collaboration with Science Exchange and Figshare launched “The Reproducibility Initiative”.<sup>5</sup> This project aims to help scientists to validate their research findings. The Research Object initiative<sup>6</sup> aims to deliver a model to represent experimental resources; this model facilitates accessibility, reusability, reproducibility and also a better understanding of in silico experiments.

Publishers are also actively addressing the problem of experimental reproducibility; F1000Research,<sup>7</sup> an open science journal, suggests data preparation guidelines to capture the processes and procedures required to publish scientific dataset. The Force 11 initiative,<sup>8</sup> a community of researchers addressing issues in scholarly communication, has published a set of metadata standards for biomedical research. These standards focus on three recommendations: Gene accession numbers, organism identification and reagent identification. Vasilevsky et al., [6] recently published a study addressing the issue of material resource identification in biomedical literature. Interestingly, the results indicated that 54% of the resources are not uniquely identifiable in publications.

Unlike other approaches, the SP ontology provides a formalized representation of the domain that is not sufficiently covered by other ontologies. For instance, SP-document delivers a structured vocabulary representing a specific type of document, a protocol. This vocabulary includes rhetorical components (e.g. introduction, materials, and methods); it also has information like application of the protocol, advantages and limitations, list of reagents, critical steps. In addition, The formalization of instructions in the protocol, or steps, is covered in SP-workflow by the class `p-plan:Step`. The order in which these steps should be executed is captured by the property `bfo:isPrecededBy`. Inputs and outputs from each step are represented by the class `p-plan:Variable`.

## 2 Methodology

For designing SP, we followed practices recommended by the NeOn methodology [7]. Also, we carefully considered the experience reported by García [8]; for example, we used conceptual maps to better understand the correspondences, relations and feasible hierarchies in the knowledge we were representing. In addition, concept maps proved to be simpler for exchanging models with domain experts. The stages and activities we implemented throughout our ontology development process are illustrated in Fig 1.

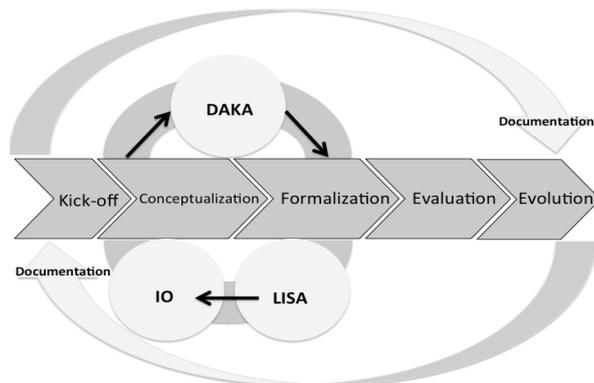
---

<sup>5</sup> <http://blogs.plos.org/everyone/2012/08/14/plos-one-launches-reproducibility-initiative/>

<sup>6</sup> <http://www.researchobject.org/>

<sup>7</sup> <http://f1000research.com/data-preparation>

<sup>8</sup> [https://www.force11.org/Resource\\_identification\\_initiative](https://www.force11.org/Resource_identification_initiative)



**Figure 1.** Methodology used to develop SMART Protocols.

## 2.1 Kick-off

In this stage we gathered motivating scenarios, competency questions, and requirements. We focused on the functional aspects we wanted the ontology to support. Competency questions were specified with domain experts, some of them are presented below: i) *Who is the author of the protocol?*, ii) *What is the application of the protocol?*, iii) *What is the provenance of the protocol?*, iv) *Who are the manufacturer and catalog number of reagents, equipment or supplies used?*, v) *What is the estimated time to execute a protocol?*, vi) *Which samples have been tested with a protocol?*, vii) *What are the critical steps, tips or troubleshooting of a protocol?* viii) *What are the basic steps of protocols in molecular biology?*

## 2.2 Conceptualization and formalization

In this stage we identified reusable terminology from other ontologies; supporting activities throughout this stage we used BioPortal.<sup>9</sup> We also looked into minimal information standards,<sup>10</sup> guidelines and vocabularies representing research activities [9-11]. Issues about axioms required to represent this domain were discussed and tested in Protégé v. 4.3; during the iterative ontology building, classes and properties were constantly changing. We identified three main activities throughout this stage, namely:

1. **Domain Analysis and Knowledge Acquisition, DAKA:** from the journals we worked with, protocols and guidelines for authors were analyzed; theory vs. practice was our main concern, *What information elements were required? Was there any relation between terminology from ontologies and these set of requirements from journals?* We also manually verified if published protocols were following the guidelines, if not, *What was missing?* Throughout this activity

<sup>9</sup> <http://bioportal.bioontology.org/>

<sup>10</sup> <https://www.force11.org/node/4145>

we were also analyzing existing ontologies and minimal information standards against published protocols. DAKA was facilitated because the knowledge engineer, namely Olga Giraldo, was also a domain expert with over ten years working in a laboratory of biotechnology. 35 domain experts were active participants in the development of SP; they were responding surveys, attending workshops, assisting in the definition of competency questions and scenarios of use. They were also validating the terminology and the relations.

We manually reviewed 175 published and non-published protocols from domains like biotechnology, virology, biochemistry and pathology. The non-published protocols (75 in total) were collected from four laboratories located at International Center for Tropical Agriculture (CIAT).<sup>11</sup> The published protocols (open access protocols in plant biology) were gathered from 9 repositories: Biotechniques,<sup>12</sup> Cold Spring Harbor Protocols (CSH Protocols),<sup>13</sup> Current Protocols (CP),<sup>14</sup> Genetics and Molecular Research (GMR),<sup>15</sup> Journal of Visualized Experiments (JoVE),<sup>16</sup> Protocol Exchange (PE),<sup>17</sup> Plant Methods (PM),<sup>18</sup> Plos One (PO)<sup>19</sup> and Springer Protocols (SP)<sup>20</sup> (Table 1).

**Table 1.** Repositories and number of protocols analyzed.<sup>21</sup>

Repository	CP	JoVE	PE	PM	CSH	Bio Tech.	GMR	PO	SP
No. of protocols	25	21	13	12	9	6	5	5	4
<b>Total</b>	<b>100</b>								

**2. Linguistic and Semantic Analysis, LISA:** this is the most complex activity throughout our development process. We identified linguistic structures that authors were using to represent actions; we needed to understand how instructions were organized. We were interested in understanding how verbs were representing actions, what additional information was there for indicating attributes for actions. By analyzing texts we were also identifying terminology and determining whether these terms were already available in existing ontologies. Minimal information standards were also considered; how could these be used when describing an experimental protocol?

<sup>11</sup> <http://ciat.cgiar.org/>

<sup>12</sup> <http://www.biotechniques.com/protocols/>

<sup>13</sup> <http://cshprotocols.cshlp.org/>

<sup>14</sup> <http://www.currentprotocols.com/WileyCDA/>

<sup>15</sup> <http://www.geneticsmr.com/>

<sup>16</sup> <http://www.jove.com/>

<sup>17</sup> <http://www.nature.com/protocolexchange/>

<sup>18</sup> <http://www.plantmethods.com/>

<sup>19</sup> <http://www.plosone.org/>

<sup>20</sup> <http://www.springerprotocols.com/>

<sup>21</sup> <http://goo.gl/MC4mR9>

From our dataset we extracted common elements and evaluated whether those protocols could be implemented. Initially, we focused our analysis on identifying necessary and sufficient information for reporting protocols. From our inspection, we determined workflow aspects in experimental protocols. The sequence of instructions had an implicit order, following the input output structure. Actions in the workflow of instructions were usually indicated by verbs; accurate information for implementing the action implicit in the verb was not always available. For instance, structures such as “*Mix thoroughly at room temperature*”, “*Briefly spin the racked tubes*” are common in our dataset. Due to the ambiguity and lack of detailed information for specifying actions in the instructions, it was difficult to understand how could these be implemented. Domain expertise was usually required in order to interpret some of the actions in our dataset.

In addition, we also isolated elements pertaining to domain knowledge as well as document related characteristics. We classified our protocols within 4 groups according to the purpose, namely: i) plant genetic transformation, ii) DNA/RNA extraction and purification, iii) PCR and their variants, iv) electrophoresis and sequencing. Within each group we identified basic steps (or common patterns), which we consider as necessary in the structure of the protocol. For example, we found that a cell disruption step is essential in DNA extraction protocols. We also identified that a digestion reaction (removing the lipid membrane, proteins and RNA) follows and that the DNA precipitation or purification comes at the end of this process.

Variables	Constants
Cell disruption (CD)	First= first step
Digestion reaction (DR)	Second= second step
DNA precipitation or purification (DNAP)	Third= third step

```
dna_extraction_protocol(CD, DR, DNAP):-
    CD= first, DR= second, DNAP= third
```

**3. Iterative ontology building and validation, IO:** as we were gathering information and learning about this domain, we started by building concept maps; these were rapidly mapped to parts of speech from the texts we were analyzing and also to existing ontologies. As concept maps were growing in complexity, number of concepts and relations, we then started to build draft ontologies –baseline ontologies representing specifics from parts of speech we identified. The knowledge engineer conducted the evaluation of the draft ontologies against competency questions. Models were also exchanged with domain experts; the process was iterative and, the models were constantly growing.

By building ontology models as well as by carefully analyzing the information we were gathering from LISA and DAKA activities, we were able to identify the modularity needed to represent experimental protocols. The module SP-document was designed to provide a structured vocabulary of concepts to represent information

for recording and reporting an experimental protocol. The module SP-workflow aims to provide a structured vocabulary of concepts to represent the execution of experimental protocols in life sciences.

### 2.3 Evaluation

The goal of the evaluation is to determine what the ontology defines, and how accurate these definitions are. Here we follow the activities proposed by Gómez-Pérez et al. [12] for terminology evaluation, which provide the following criteria:

1. **Consistency.** It is assumed that a given definition is consistent if, and only if, no contradictory knowledge may be inferred from other definitions and axioms in the ontology.
2. **Completeness.** It is assumed that ontologies are in principle incomplete [12, 13], however it should be possible to evaluate the completeness within the context in which the ontology will be used. An ontology is complete if and only if: o All that is supposed to be in the ontology is explicitly stated, or can be inferred.
3. **Conciseness.** An ontology is concise if it does not store unnecessary knowledge, and the redundancy in the set of definitions has been properly removed.

According to the criteria for evaluation proposed by Gomez-Perez [12], our ontologies were developed using the OWL-DL because of expressiveness and computational completeness.<sup>22</sup> The Protégé plugin OWLViz<sup>23</sup> was used to visualize and to correct syntactic inconsistencies. The Ontology Pitfall Scanner (OOPS),<sup>24</sup> was useful to detect and correct anomalies or pitfalls in our ontologies [14]. In relation to the evaluation of the terminology, we represented the 175 protocols using the SMART Protocol formalism, emphasizing on informative elements. For most of the cases, we had insufficient information from the protocols; domain expertise was therefore required in order to determine what was missing and how to best represent it. We also used surveys<sup>25</sup> in order to determine how complete our model was. As a result from our analysis we proposed a checklist<sup>26</sup> to report experimental protocols in plant biology; 35 domain experts validated this checklist.

### 2.4 Evolution

At the end of the cycle, new classes, properties and individuals are identified. These are then analyzed against the set to competency questions, existing ontologies, parts of speech and linguistic structures. The model evolves as new knowledge goes through the whole cycle. Although our ontology is a young ontology, we could observe how it evolved in its conceptualization as well as in the explicit specification.

---

<sup>22</sup> <http://www.w3.org/TR/2004/REC-owl-features-20040210/#s1.3>

<sup>23</sup> <http://protegewiki.stanford.edu/wiki/OWLViz>

<sup>24</sup> <http://oeg-lia3.dia.fi.upm.es/oops/index-content.jsp>

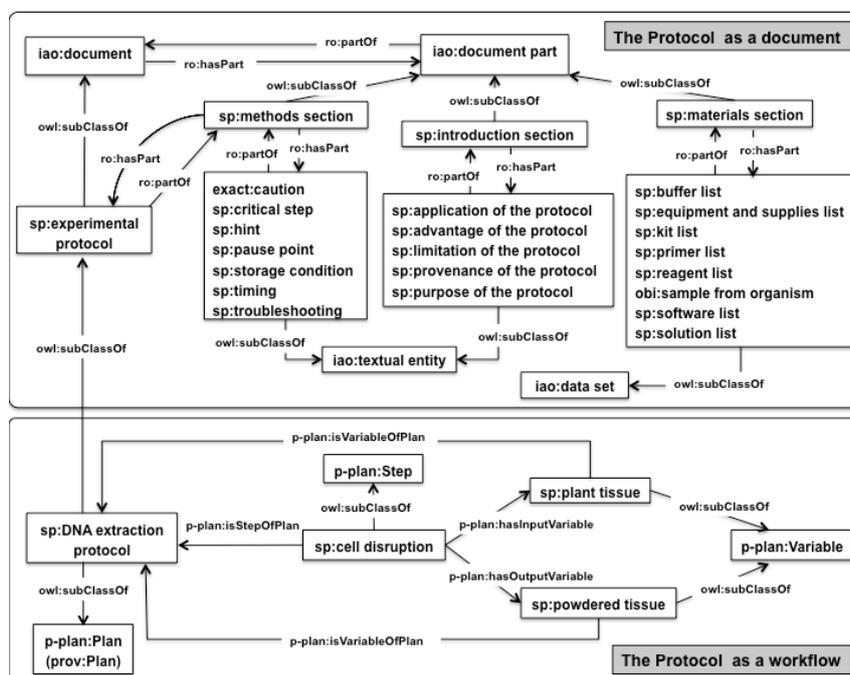
<sup>25</sup> [goo.gl/jBHPo](http://goo.gl/jBHPo)

<sup>26</sup> [goo.gl/gAVnn](http://goo.gl/gAVnn)

### 3 The SMART Protocols Ontology

The SMART Protocols approach follows the OBO Foundry principles [15]. Our modules reuse the Basic Formal Ontology (BFO).<sup>27</sup> Also, we reused the ontology of relations (RO) [16] to characterize concepts. In addition, each term from SP is represented by annotation properties imported from OBI Minimal metadata.<sup>28</sup>

An overview of the two modules comprising SP is illustrated in Figure 2. The classes, properties and individuals are represented by their respective labels to facilitate the readability. The prefix indicates the provenance of each term. The ontology describing the experimental protocol as a document is depicted at the top. The class `iao:information content entity` and its subclasses `iao:document`, `iao:document part`, `iao:textual entity` and `iao:data set` were imported from The Information Artifact Ontology (IAO) to represent the document aspects in the protocol. The ontology describing the experimental protocol as a workflow is depicted at the bottom. The representation of executable aspects of a protocol is modeled with the classes `p-plan:Plan`, `p-plan:Step` and `p-plan:Variable` from the P-Plan Ontology (P-Plan).



**Figure 2.** SMART Protocols as an extension of the ontologies IAO and P-Plan. The document aspects in a protocol are captured with IAO. The workflow aspects in a protocol are captured with P-Plan. The terms proposed in SMART Protocols use the `sp` prefix.

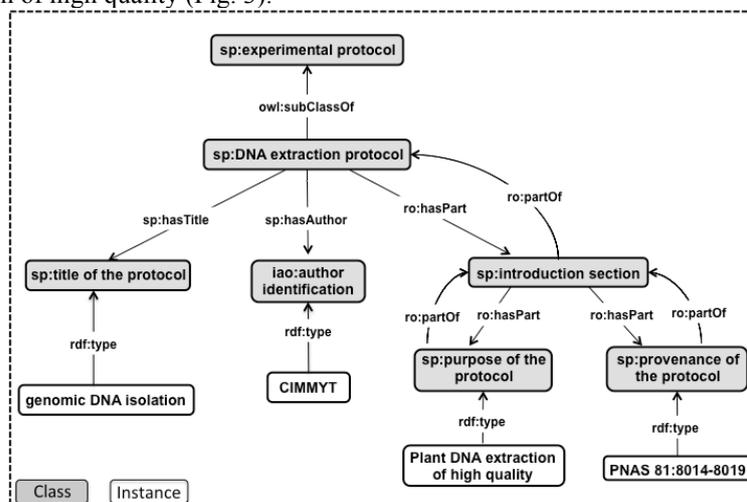
<sup>27</sup> <http://www.ifomis.org/bfo/>

<sup>28</sup> [http://obi-ontology.org/page/OBI\\_Minimal\\_metadata](http://obi-ontology.org/page/OBI_Minimal_metadata)

### 3.1 The protocol as a document

The document module of SMART Protocols reuses classes from CHEBI [17], EXACT, MGED [18], SO [19], OBI and SNPO.<sup>29</sup> Also, SMART Protocols-document (henceforth SP-document) extends the class `iao:information content entity` proposed by the Information Artifact Ontology (IAO) to represent the experimental protocol as an `iao:document` that has parts, `ro:has_part`, such as `iao:document part` (`iao:author list`, `sp:introduction section`, `sp:materials section` and `sp:methods section`). See the top of Figure 2 for details.

**Use Case.** SP-document represents information such as, the protocol type, `sp:DNA extraction protocol`; it has a title, identified by the property `sp:has title`, it is instantiated by genomic DNA isolation. Also, the author entry, `iao:author identification`, is instantiated by CIMMYT [20]. This protocol is derived, `sp:provenance of the protocol`, from the protocol published by [21] (`sp:PNAS 81:8014-8019`) and its purpose is instantiated by plant DNA extraction of high quality (Fig. 3).



**Figure 3.** The document aspects in a protocol are captured with IAO. The terms proposed in SMART Protocols use the `sp` prefix.

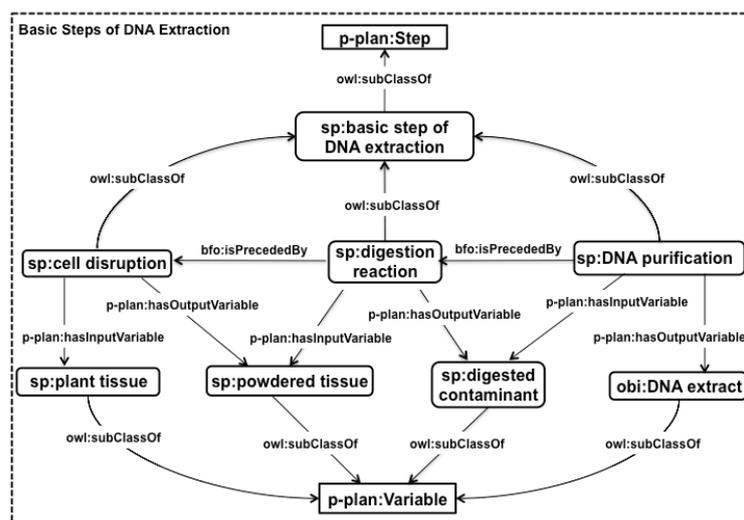
### 3.2 The protocol as a workflow

The workflow module extends the P-Plan Ontology (P-Plan). This ontology was developed to describe scientific processes as plans and link them to their previous executions. In the workflow module of SMART Protocols (henceforth SP-workflow), the experimental protocol, `p-plan:Plan`, is a description of a sequence of

<sup>29</sup> [http://www.loria.fr/~coulet/snponontology1.4\\_description.php](http://www.loria.fr/~coulet/snponontology1.4_description.php)

operations, `p-plan:Step`, that includes an input and an output `p-plan:Variable`. In this sense, a protocol is a type of workflow. See the bottom of Figure 2 for details. SP-workflow also reuses classes from CHEBI, MGED, SO, OBI and NPO [22].

**Use Case.** DNA extraction is a procedure frequently used to collect DNA for subsequent molecular or forensic analysis. DNA extraction includes 3 basic `p-plan:Steps`: i) cell disruption or cell lysis, ii) Digestion reaction (in this step, contaminants such as lipid membrane, proteins and RNA are removed from the DNA solution), and iii) DNA purification. Each one of these steps may include different protocols (or `p-plan:Plans`) to be executed. For example, the step `sp:cell disruption` or cell lysis may be achieved by chemical and physical methods - blending, grinding or sonicating the sample. Also, the ontology considers that each step is executed following a predetermined order. For instance, according to the protocol published by CIMMYT, the cell disruption by lyophilization and grinding has an input variable, `p-plan:hasInputVar`, as well as `sp:plant tissue`; it also has an output, `p-plan:hasOutputVar`, and `sp:powdered tissue`. The next step, `sp:digestion reaction`, has as input the output of the immediately previous step, `sp:powdered tissue`, and as output `sp:digested contaminant`. The last one, `sp:DNA purification` has as input `sp:digested contaminant`, and as output `obi:DNA extract` (Fig. 4).



**Figure 4.** Extending the P-plan ontology to represent experimental protocols in life sciences. The sp prefix indicates the terms proposed by the SMART Protocols ontology.

## 4 Discussion and Conclusions

Science has, among other, two important characteristics, reproducibility and clarity.<sup>30</sup> Clarity provides unambiguous descriptions for results in a mechanical or mathematical form. The lack of clarity about "*how to do* or *how to execute*" an experimental procedure hinders the reproducibility and impedes comparing results across related experiments. SMART Protocols addresses clarity by formalizing the objects that should go together with actions. Besides, SP reuses and extends minimal information standards, incorporating these structures within the representation of experimental protocols. By delivering a semantic and syntactic structure SP also facilitates reproducibility.

Our ontology-based representation for experimental protocols is composed of two modules, namely SP-document and SP-workflow. In this way, we represent the workflow, document and domain knowledge implicit in experimental protocols. Our work extends IAO and P-Plan ontology. Actions, as presented by [3] are important descriptors for biomedical protocols; however, in order for actions to be meaningful, attributes such as measurement units and material entities (e.g., sample, instrument, reagents, etc.) are also necessary.

Formalizing workflows has an extensive history in Computer Science; not only in planning but also in execution –as in Process Lifecycle Management and Computer Assisted Design/Computer Assisted Manufacturing. We have considered some of these principles for representing workflow aspects in protocols just as we have reused knowledge formalisms like the P-Plan Ontology. By formalizing the workflow implicit in protocols the execution can be ontologically represented in a sequential manner that is intelligible by humans and processed by machines.

Modularization, as it has been implemented in SP, facilitates managing the ontology. For instance, the workflow module can easily be specialized with more specific formalisms so that robots can process the flow of tasks to be executed. The document module facilitates archiving; the structure also allows to have fully identified reusable components. By combining both modules we are delivering a self-describing document.

## References

1. Garijo, D. and Y. Gil. *Augmenting PROV with Plans in P-PLAN: Scientific Processes as Linked Data*. in *2nd international Workshop on Linked Science 2012 - Tackling Big Data (LISC2012), in conjunction with 11th International Semantic Web Conference (ISWC2012)*. 2012. Boston, MA: Springer-Verlag.
2. Soldatova, L.N. and K.R. D., *An ontology of scientific experiments*. journal of the royal society interface, 2006. **3**(11): p. 795–803.

---

<sup>30</sup> <http://www.aaas.org/page/so-can-science-explain-everything>

3. Soldatova, L.N., et al., *The EXACT description of biomedical protocols*. Bioinformatics, 2008. **24**(13): p. i295-303.
4. Courtot, M., et al. *The OWL of Biomedical Investigations in OWLED workshop in the International Semantic Web Conference (ISWC)*. 2008. Karlsruhe, Germany.
5. Sansone, S.A., et al., *Toward interoperable bioscience data*. Nat Genet, 2012. **44**(2): p. 121-6.
6. Vasilevsky, N.A., et al., *On the reproducibility of science: unique identification of research resources in the biomedical literature*. PeerJ, 2013. **1**: p. e148.
7. Suárez-Figueroa, M.C., *Ontology engineering in a networked world*. 2012, Berlin ; New York: Springer. xii, 444 p.
8. Garcia-Castro, A., *Developing Ontologies in the Biological Domain in Institute for Molecular Bioscience*2007, University of Queensland: Queensland. p. 275.
9. Zimmermann, P., et al., *MIAME/Plant - adding value to plant microarray experiments*. Plant Methods, 2006. **2**: p. 1.
10. Bustin, S.A., et al., *The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments*. Clin Chem, 2009. **55**(4): p. 611-22.
11. Gibson, F., et al., *Guidelines for reporting the use of gel electrophoresis in proteomics*. Nat Biotech, 2008. **26**(8): p. 863-864.
12. Gomez-Perez, A., *Evaluation and assessment of knowledge sharing technology*, in *Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing*, N.J.I. Mars, Editor. 1995, IOS Press: Amsterdam, The Netherlands. p. 289-296. .
13. Gómez-Pérez, A., M. Fernández-López, and O. Corcho, *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the Semantic Web*. 2004: Springer.
14. Poveda-Villalón, M., M. Suárez-Figueroa, and A. Gómez-Pérez, *Validating Ontologies with OOPS!*, in *Knowledge Engineering and Knowledge Management*, A. ten Teije, et al., Editors. 2012, Springer Berlin Heidelberg. p. 267-281.
15. Smith, B., et al., *The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration*. Nature Biotechnology, 2007. **25**(11): p. 1251 - 1255.
16. Smith, B., et al., *Relations in biomedical ontologies*. Genome Biology, 2005. **6**(5): p. -.
17. de Matos, P., et al., *Chemical Entities of Biological Interest: an update*. Nucleic Acids Research, 2010. **38**: p. D249-D254.
18. Stoeckert Jr, C.J. and H. Parkinson, *The MGED ontology: a framework for describing functional genomics experiments*. Comparative and Functional Genomics, 2003. **4**: p. 127-132.
19. Mungall, C.J., C. Batchelor, and K. Eilbeck, *Evolution of the Sequence Ontology terms and relationships*. J Biomed Inform, 2011. **44**(1): p. 87-93.
20. CIMMYT, *Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory*, 2005, CIMMYT: Mexico, D.F. p. 102.
21. Saghai-Marouf, M.A., et al., *Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics*. Proc Natl Acad Sci U S A, 1984. **81**(24): p. 8014-8.
22. Thomas, D.G., R.V. Pappu, and N.A. Baker, *NanoParticle Ontology for cancer nanotechnology research*. J Biomed Inform, 2011. **44**(1): p. 59-74.