International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 9 (2018) pp. 6529-6535

Automatic Workflow Scientific Model for Lab Test in Civil Engineering

Ingrid-Durley Torres*, Jaime Guzman-Luna, Francisco Javier Moreno

Department of Computational Decision Science
School of Mines at Universidad Nacional de Colombia, hearquarters Medellin, Colombia.

*(*Corresponding author)*

Abstract

Lab tests are essential in the field of civil engineering. These procedures are conducted by technical personnel who supported by a set of guides conduct laboratory practices which work with multiple computational calculations, which are done manually. We have focused on providing a mechanism that may contribute to easing human tasks and which facilitates conducting Lab tests in the specific domain of civil engineering; therefore, we have implemented and discussed a workflow representation model working on a platform to execute computational calculation resources, presenting a transition from the manual recording of civil engineering Laboratory practices up to the automation process.

INTRODUCTION

Currently, civil works are designed, constructed and supervised using different materials; therefore, it is essential to know the physical and mechanical characteristics of those materials [3]. This is done through their study in the laboratory, as part of material quality control. In this context, this study contemplates Lab tests, which have standardized procedure guides to conduct laboratory practices [11]. These guides conserve a structure that allows technicians to describe the experimental process being conducted with an adequate level of detail. One of the items that is mandatory in the structure is the one represented by mathematical calculations, which describe the equations necessary to quantify an item's properties. In their vast majority, these calculations are conducted manually by lab technicians, who must record the significant value of input and output data of each operation on the tables written on paper or in digital files that are later transcribed into calculators, executed in pre-recorded excel formulas, or in the best of the cases, recorded on a software which has restrictions for permitted operations. If to the above, one adds calculation variability, which goes from simple operations up to complex activities like logarithms, mixed with iterative summations and other operations; the task becomes slow and humanly costly. To provide a tool, that in addition to making human tasks easier, conserves the coherence determined by the standardization of procedure guides. This proposal presents a lab test representation model, which supported by a service-oriented architecture, allows computational calculation automation, using as an input tests previously conducted. The advantages of implementing this proposal reflect a transition from manual documentation practices in the field of civil engineering to the automation of the process with advantages as: i) the representation of lab practices, in a procedural model comprehensible to machines (semantic) [1]; ii) automatic execution of computational calculations [14] and iii) recording, searching, recovering and conducting lab tests that favor reproducibility and the reuse of results [4, 11]. Although there are composition proposals for laboratory processes in other domains, this is the first to focus on lab practice descriptions, targeting the domain of civil engineering.

To provide a more detailed view of this work, this article is organized as follows. First, the article presents a description of a test and teaches the representation model. Then, the article provides the way resources are linked (computational calculation services). Then, the article describes the architecture, which is that presented in a case study; afterwards, the article shows a comparison with other works, and finally, the article incorporates implications and discussion concerning the subject matter.

LAB TEST

Commonly, a Lab test includes a concise description of ordered instructions to determine a property or characteristic of a material or product [3]. Instructions to conduct the test must include all the essential necessary details that must be implemented in the laboratory. Therefore, the structure of the test includes an introduction which includes the scope of the test, relevant definitions to understand the procedure and the importance and application of a test method (or methods). It also includes a list of materials and equipment which will be used on the test, followed by a detailed updated description of a corresponding lab procedure [16]. As an important part, the article details the mathematical formulas to achieve satisfactory precision and trends, and a format to record the data gathered during the test and the corresponding information obtained from the calculations conducted during the test.

Seen from this angle, these elements can be grouped into 2 more generic categories; one is documentary and the other is procedural (commonly referred to as workflow). The former includes textual description items as an introduction and materials. While in the latter, a workflow [1], all the steps are associated one by one described in the procedure including the corresponding calculation activity (when applicable, in that step). Identifying for each calculation its input and output data. Figure 1 shows an example of how these items can be grouped. It corresponds to the representation adopted in this study.

On the other hand, Figure 2 belongs to a detailed abstract model, of the description of the procedure (workflow) conducted in the laboratory [17]. This procedure was modeled on a semantic representation which makes it comprehensible to the machine. This representation is composed of set of steps connected via a dependence on order, expressed through the relation *IsPrecededBy*. This semantic model is inspired by P-Plan [9]. Each step represents a logical step of a lab test procedure; whose subject is to calculate a result. It is important to clear up that even

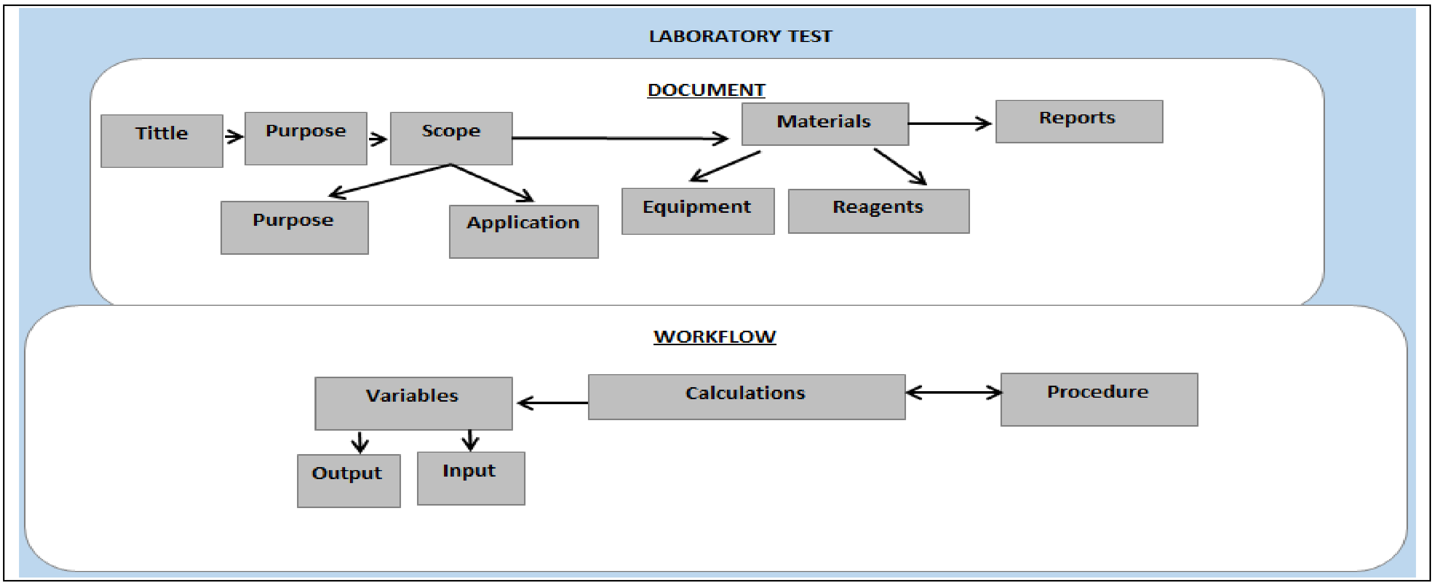

Figure 1. Structure of a lab test

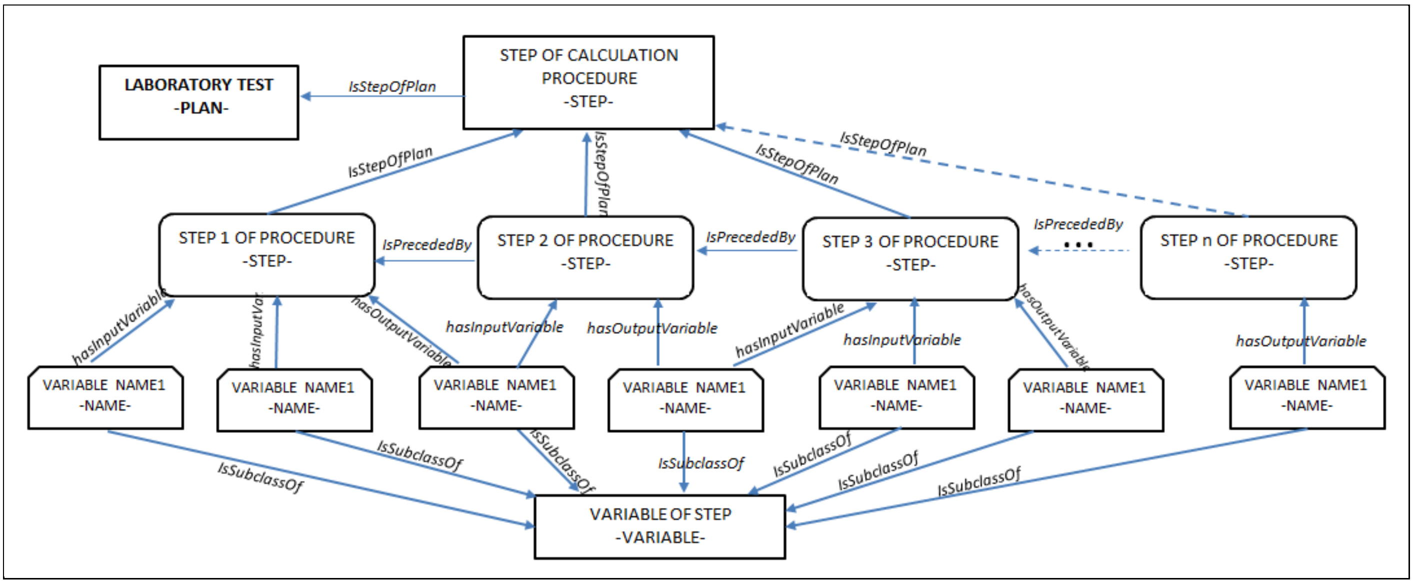

Figure 2. Description workflow ASTM2216 Standard Test Method for Particle Size Analysis of Soils

if the representation model perfectly allows the incorporation of tasks done by a human user (mixing, weighing, counting, etc.), as a step, and they are not taken into account because of the intention to automatize. Nevertheless, it is presumed that human tasks are adequately executed, and in this case, they are limited to providing a corresponding datum (for instance, sample weight). In other words, the model is in charge of the test representation in the lab steps computationally, in other words *in silico* [12], which is nothing more than the activities that can be conducted on computing machines.

Steps and their variables (workflow) are put into practice depending on the laboratory they are representing; hence, even though the model is abstract, it is unique (Figure 2). One or more concrete versions can be generated for each abstract laboratory. That will depend on the methods that the lab and the software resource diversity used to tie each calculation.

RESOURCE ALLOCATION

The abstract representation is not complete if it does not define how to tie software resources in charge of calculation tasks. In this case, to fulfill the task, the study generated a software application (known as web service [7]) for each calculation step. It is a software application that permits the interoperation of software which lab technicians traditionally use. This is done establishing a standard communication to conduct real-time data exchange through a request-response mechanism named web service. In that case, the study resorts to an abstract workflow representation and its specification in one of the platforms defined in the literature (See section 6). Concerning this, services which permit access to mathematical calculation processes are tied Figure 3 describes an in *silico* workflow. In this case, the workflow is represented and converted to a Tavern model [20], and through the model, researchers made a connection with a software resource which is in charge of realizing the corresponding equation and delivering the result in the next step until it terminates the complete execution described by the workflow plan. The graphic representation used to model workflow allows the incorporation of new services through simple means, and it links it to its corresponding step when conducting computational calculation. Thus, one by one, the steps are tied to a service which in turn, interacts with its respective software. This way, they ask for the input data one, or as a set, and access the service (or services) that conduct corresponding mathematical operation. This is achieved in real time and in accordance with the input data which was provided. The final result is returned as output data, to the following step in the procedure, which in turn, will perform the corresponding calculation. This process repeats successively until all the steps have been executed and the final output is delivered to the user who is in charge of conducting the Lab test.

One of the key aspects of this model is represented in releasing resources which must be accessible from their URLs. Nevertheless, that permissibility can be determined only for input and output specifications that the software requires for its execution. It is not necessary to know its internal composition, or to reveal more functionalities than the ones required for the computational operations for which that software has been bound. Therefore, these descriptions require a mental paradigm change of the people who dispose of the resource; in this case, a lab technician, who must express total agreement to release a resource. Since the model is accessible to several lab technicians, this service is highly distributed, and as such various users must be able to access it.

ARCHITECTURE OF THE MODEL

In order to understand the functionality of the model (named ROCE- Research Object [21] Civil Engineering), it is important to describe the environment in which the lab test workflows can be automated. The model assumes that the abstract workflow is represented by an expert engineer, and it is this engineer, according to the test description, who incorporates the procedural sequence of the computational calculations, adjusting to the model that was adopted (Figure 2).

Each generated workflow is stored in an abstract representation repository. It is named so because even if it describes the sequence of the steps for computational calculations, depending on the type of It neither includes the real value of those input and output data that it respectively generates. The lab technician is in charge of the following functional step. In this case, it is the lab technician who logs in (under a logging model), and seeks from a user-friendly interface the abstract

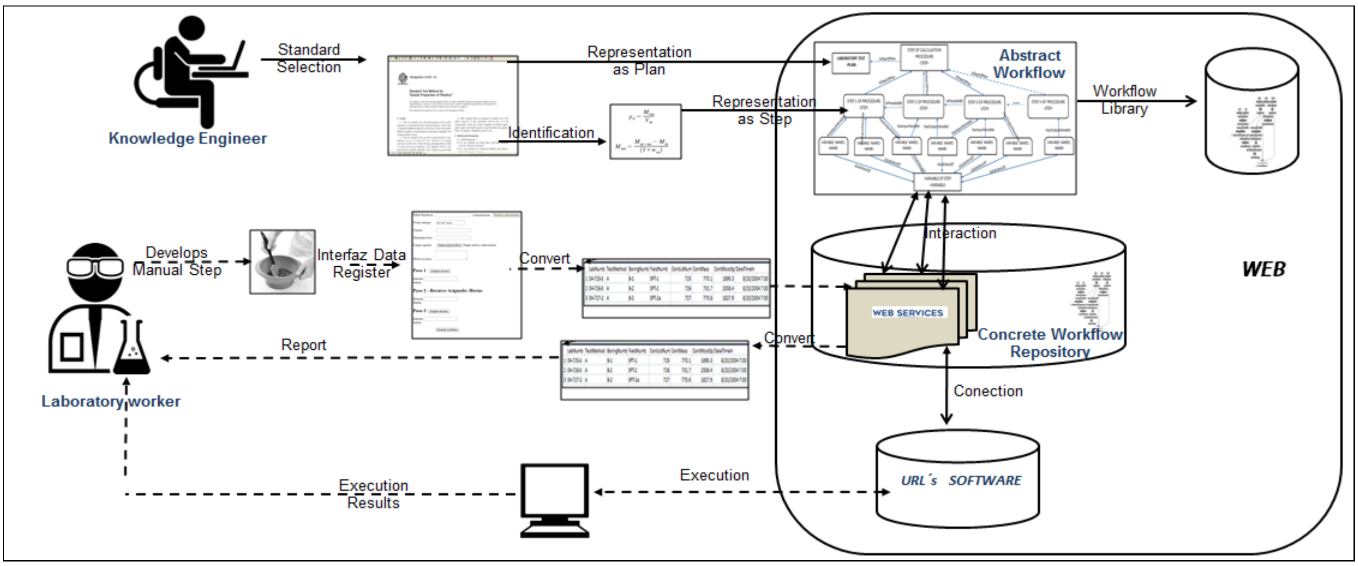

Figure 3. Architecture of the functional model

O workflow of the test that they want to conduct in the laboratory. Once identified, this study proceeded to generate, a procedural sequence defined as human activities to deliver the input data required for the test. In the following step, the lab technician must fill out the fields corresponding to the type of input data, already set up by the system providing the real value of that datum. In turn, the technician must find the software on which that technician wishes to execute calculations. Under this option, once you key in the software for each step, ROCE will enable some questions aiming to connect the service and these resources. The model warns that the connection will be enabled to be accessed by other lab technicians in the future. Once, this is accepted, ROCE not only executes the complete workflow, but also finishes the execution. It delivers a table with all the variables associated to this test. Finally, it stores the results in a repository named concrete workflow. Afterwards, other lab technicians can access these workflows to reproduce the test (the same data and the same conditions, the same results), or reuse the test (with their own data, and the same resources, to obtain their own results).

This way, ROCE executes all the computational workflow without the slightest human intervention. To do so, users do not need to download or install anything; they simply provide input data and receive a corresponding output.

CASE STUDY

Figure 4 presents the typical to type in input data and the software resource registration presented to the lab technician.

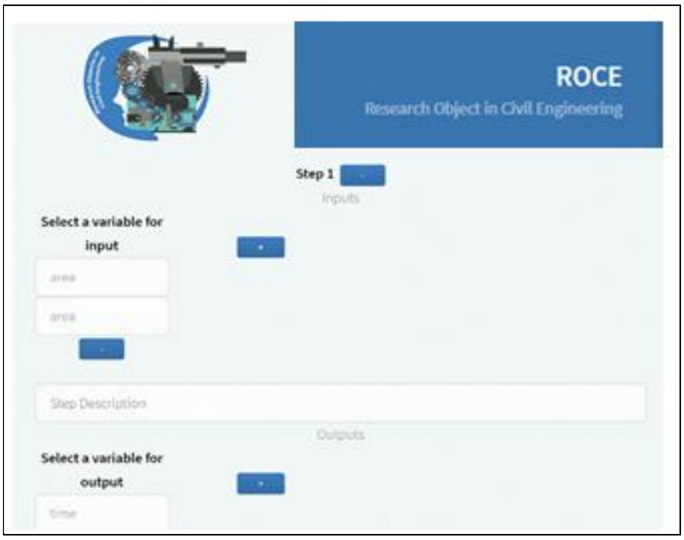

Figure 4. Portion. Concrete workflow interface

As it can be seen, the information is requested from the user without demanding any technical know-how different from the ones that the lab technicians already master. In this case, it is important to Define the type of datum with which the variables are going to be identified.

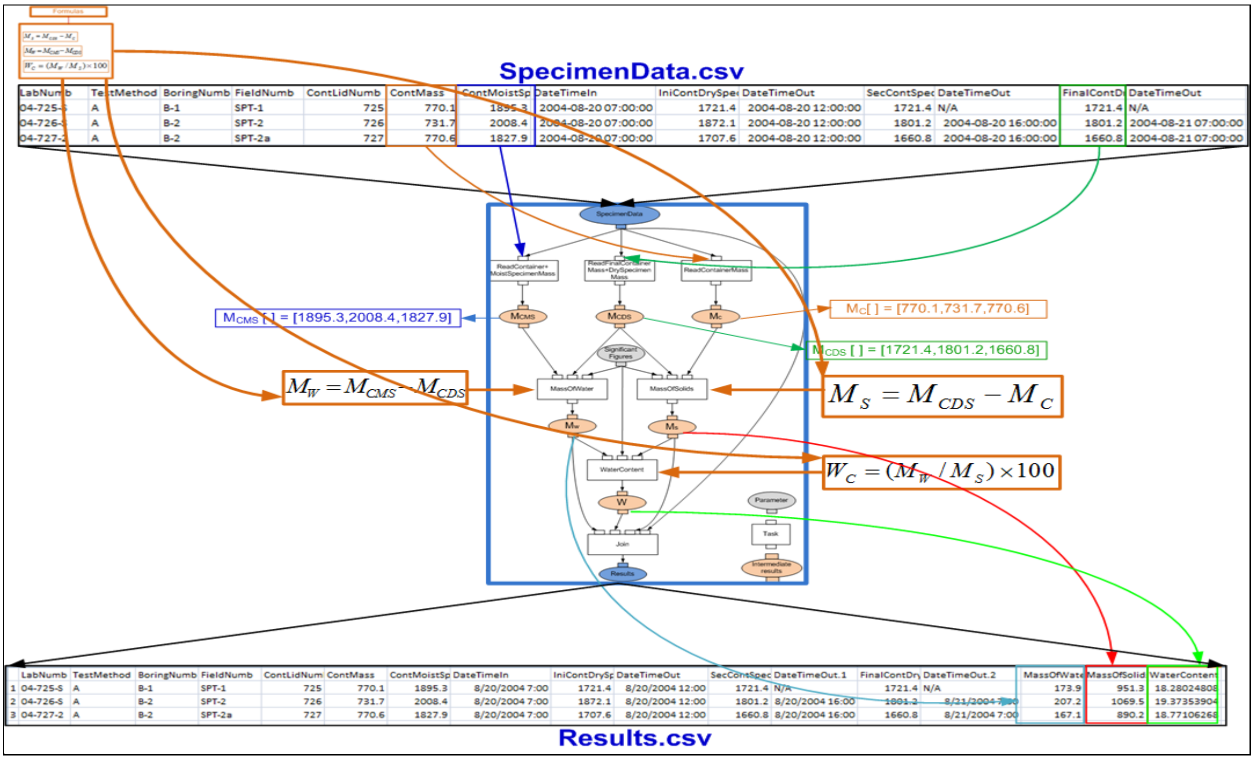

Figure 5. Portion. Concrete workflow interface

On the other hand, Figure 5 presents the concrete workflow just as ROCE understands it. In this specific case, it corresponds to an ASTMTest 2216 norm test, whose objective is *to determine the water (moisture) content of soil*, and which is made up of five steps: *calculation mass of specimen water*, *calculation mass of oven dry specimen*, and *calculation water*, , in which the first receives as input variables, the mass of thecontainer with a moist specimen (*MassOfContainerAndMoistSpecimen*) and the mass of the container with a dry specimen (*MassContainerandOvenDrySpecimen*) to provide as an output variable the mass of water (*MassOfWater*). Meanwhile, the second step is in charge of calculating the mass of the dry specimen (calculation mass of oven dry specimen), and it receives as input (*MassContainerandOvenDrySpecimen*) and the mass of the container (*MassContainer*) whose output variable corresponds to the mass of a dry specimen.

Finally, it calculates the water content *(calculation water)*, that it receives as inputs, the outputs of the previous two processes *MassOfWater* and *MassContainer* to deliver the user the *contentWater* variable.

The input data are delivered in a table, which is read and executed by the workflow, which communicates with the corresponding services (one for each step), to deliver the outputs to the following step. Finally, it documents those outputs completing the initial table; and then, it is finally delivered to the user.

RELATED WORK

There are similar works in other domains as: Bioinformatics [8, 20], astronomy [6], Oceanography [5], Biology [15], chemistry [19], but as it seemed, there is nothing to cite in terms of civil engineering. Table 1 summarizes part of these studies. Our proposal resorts to representation of DCG (Directed Cyclic graph) workflow which is contrary to others whose specification is more complex because it is DAG (Directed Acyclic graph), yet it is a start.

Table I. Related works

	DOMAIN	WORKFLOW REPRESENTATION	WORKFLOW EXECUTABLE	DEVELOPED BY	SEMANTIC REPRESENTATION
NEPTUNE [5]	Ocean buoys	Workflow management system (WMS).	✓	University of Washington	×
PEGASUS [6]	Bioinformatics astronomy, earthquake science, gravitational wave physics, and ocean science	DAG	✓	University of Southern California	✓
TAVERNA [20]	bioinformatics	DAG	✓	University of Manchester	×
TIRIANA [15]	Biology	DCG	✓	Cardiff University	×
KEPLER [2]	Biology, ecology, geology, chemistry, and astrophysics	DCG	✓	University of California	×
ASKALON [8]	Bioinformatics	DCG	✓	University of Innsbruck,	
Weka4WS [13]	Data mining systems	DCG	✓	Grid Lab of University of Calabria	×
GWES [10]	Genetic, Bioinformatics	Petri nets GridWorkflow Execution Service	✓		✓
CoG Kit [18]	Urban Water Distribution Systems.	DAGs	✓	Argonne National Laboratory	×
DIS3GNO [12]		DAG	✓		
SEDNA [19]	Chemistry	BPEL	✓	UK EPSRC	×
SWIFF [17]		DAG and DCG	✓		×

Another of the proposed contributions is to incorporate a semantic model which up to date allows storing, searching and recovering lab trial workflows in civil engineering. However, in turn, it empowers other interoperability processes with other domains.

DISCUSSION

The model this study proposed has permitted a migration from the traditional manual way to execute lab tests in civil engineering to an automatic execution model. it is a start to reproduce and reuse the results of the experiments more In other words, it contributes to *in silico* experimenting in the field of civil engineering.

That although, it only concentrates on computational activities, efficiently. This is always done relieving human labor. Up-to-date, 30 lab tests have been modeled abstractly in soils, hydraulics and pavements, and all the tests are ASTM-Test norm supported. 15 tests have been implemented after materialization with real resources (only soils) although the work in the immediate future is to represent materials tests and obviously, lab technicians will materialize them with real resources. Is important to highlight that considering semantic representation has enabled us to generate a test recovery system which searches closer to a user's language (currently being evaluated) and projecting a future inclusion of complex processes like the composition of lab tests incorporating, in addition to workflow, the documentary characteristics of reports not described in this study.

ACKNOWLEDGEMENTS

This study has been partially supported by a project named "a semantic recovery system for experimental protocols in civil engineering" Código Hermes 34685.

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] P. Alper, K.Belhajjame y C. Goble. "Small is Beautiful: Summarizing Scientific Workflows Using Semantic Annotations". IEEE International Congress on Big Data, Santa Clara, CA, USA, pp. 318-325, 2013

[2] I. Altintas, C. Berkley, E. Jaeger, M. Jones, M., Ludäscher, B., S. Mock, Kepler. "Towards a Grid-Enabled system for scientific workflows". The Workflow in Grid Systems Workshop in GGF10-The 10th Global Grid Forum, 2004, pp.1-5.

[3] ASTM. Form and Style for ASTM Standards. West Conshohocken, PA, USA. 19428-2959. January 2015. www.astm.org.

[4] R. Bergmann, Y., Gil, "Retrieval of semantic workflows with knowledge intensive similarity measures". 19th International Conference on Case-Based Reasoning Research and Development, pp. 17–31, 2011.

[5] W. Chen, R. Silva, E. Deelman, R. Sakellariou, "Balanced task clustering in scientific workflows". IEEE 9th International Conference on e-Science, pp. 188–195, 2013.

[6] E. Deelmana, K. Vahia, G. Juvea, M. Ryngea, S. Callaghanb, P. Maechlingb, R. Mayania, W. Chena, R. Ferreira, M. Livnyc, K. Wengerc. "Pegasus, a workflow management system for science automation". Future Generation Computer Systems, Vol. 46 pp. 17-35. May 2015.

[7] K. Douglas, D. Dick. Web Services, Service-Oriented Architectures, and Cloud Computing: The Savvy Manager's Guide (Second Edition). Morgan Kaufmann. 2013.

[8] T. Fahringer, R. Prodan, R. Duan, J. Hofer, F. Nadeem, F. Nerieri, S. Podlipnig, J. Qin, M. Siddiqui, H. Truong, A. Villazon, M. Wieczorek, M. "ASKALON. A development and grid computing environment for scientific workflows". Taylor I.J., E. Deelman, D.B. Gannon, M. Shields M. (eds) Workflows for e-Science. Springer, pp. 450–471, London, Springer 2007.

[9] D. Garrijo, Y. Gil Augmenting PROV with Plans in P-PLAN: Scientific Processes as Linked Data. Proceedings of the Second International Workshop on Linked Science 2012 – Tackling Big Data, in conjunction with 11th International Semantic Web Conference, Boston, MA, 2012, pp.25-33.

[10] A. Hoheisel, "User tools and languages for graph-based Grid workflows," Concurrency Computation Practice and Experience, vol. 18, no. 10, pp. 1101–1113, August 2006.

[11] A. de Oliveira, D. Oliveira, M. Mattoso. Clouds and Reproducibility: A Way to Go to Scientific Experiments?. In: Antonopoulos N., Gillam L. (eds) Cloud Computing. Computer Communications and Networks. Springer, Cham. 2017.

[12] D. Talia, "Workflow Systems for Science: Concepts and Tools," ISRN Software Engineering, vol. 2013, Hindawi. pp. 1-15. 2013.

[13] D. Talia, P. Trunfio, O. Verta, "The Weka4WS framework for distributed data mining in service-oriented Grids," Concurrency Computation Practice and Experience, vol. 20, no. 16, pp. 1933–1951, 2008.

[14]] D. Talia, P. TrunFo. "How distributed data mining tasks can thrive as knowledge services,". Communications of the ACM, vol. 53, no. 7, pp. 132–137, 2010.

[15] I. Taylor, M. Shields, I. Wang, A. Harrison. "The Triana workflow environment: Architecture and applications". Workflows for e-Science, pp. 320–

339. Springer, 2007.

[16] Universidad Nacional de Colombia. Laboratorios Universidad Nacional de Colombia, sede Medellín. Facultad de Minas. Retrieved Abril, 2017, from http://www.minas.medellin.unal.edu.co/dirlab/index.php/laboratorios

[17] W. Van Der Aalst, K. Hee. “Workflow Management: Models, Methods, and Systems (Information Systems)”. Paperback. June, 2004.

[18] G. Von, M. Laszewski, M. Hategan, and D. Kodeboyina, “Java CoG kit workflow”. Workflows for e-Science, pp. 143–166, Springer, New York, 2007.

[19] B. Wassermann, W. Emmerich, B. ButchartNick, C. Liang, C. Jignesh. “Sedna: A BPEL-Based Environment for Visual Scientific Workflow Modeling”. Workflows for e-Science pp 428-449, SPRINGER, 2007.

[20] K. Wolstencroft, R. Haines, D. Fellows, A. Williams, D. Withers, S. Owen, S. Soiland-Reyes, I. Dunlop, A. Nenadic, P. Fisher, J. Bhagat, K. Belhajjame, F. Bacall, A. Hardisty, A. Nieva de la Hidalga M. Balcazar, S. Sufi, C. Goble. “The Taverna workflow suite: designing and executing workflows of web services on the desktop, web or in the cloud”. Nucleic Acids Research. Vol. 41, pp. 557–561, July 2013.

[21] J. Zhao, G. Klyne, P. Holubowicz , R. Palma RO-Manager: A Tool for Creating and Manipulating Research Objects to Support Reproducibility and Reuse in Sciences: ceur-ws.org. (2012).