ISSN 2395-8618

Planning Techniques to Compose Experimental Protocols in Civil Engineering

Ingrid-Durley Torres, Juan Felipe Muñoz-Fernández, Jaime A. Guzmán-Luna

***Abstract*—This paper presents a study that shows how planning techniques in Artificial Intelligence can be used in experimental protocols composition. Also, it presents how it is necessary to represent the experimental protocol models under ASTM standard with the emerging technologies as ontologies. The main goal of this paper is to discuss the way composition process can be used to meet the needs of the users. Additionally, this paper also searches to define the way the Civil Engineering domain must normalize its experimental practices, if we want the results of experimental activities to be reproducible and reusable.**

***Index Terms*— planning techniques, experimental protocols composition, standard ASTM, ontologies, and Civil Engineering**

I. Introduction

THE experimental practices of Civil Engineering are standardized in documents that are defined by ASTM International (American Association for Testing and Materials) [2]. Some of these documents describe, as an experimental protocol [8, 10, 14], how individual (atomic) tests of laboratory must be done for the testing of specific materials, which are used in processes of the Civil Engineering domain. For ex-ample, materials such as soils, metals, among others are subject to these laboratory tests.

In this domain, at the same time, there exist more complex processes as the design of pavements [1], which requires a soil study in its preliminary phase. This soil study should be composed of the individual execution of the experimental protocols de-fined in documents ASTMD0422, ASTM4318, ASTM698, just to mention some of them. This is to say, each experimental protocol, seen in an isolated way and independent from the context of the study, is not very useful for the user of the domain of Civil Engineering [7], but as a whole, they can lead to structure more complex processes that cannot be assisted by only one protocol, but for the composition of some of them. Nevertheless, the task is not trivial, since the protocols that should be obliged to participate in the process of composition must be identified accurately, according to the desired process. Similarly, it is required to know the order of execution of every protocol because, in some cases, the result of the execution of a proto-col constitutes the input to execute the next one or a subsequent one (functional requirement). It is also necessary to identify the protocols that can be executed using the resources of the laboratory (non-functional requirements). For example, the protocols defined in documents ASTM2974 and ASTMD4643 measure the percentage of humidity of a soil sample. However, the first one describes the test using a drying oven, while the second one describes the test using a microwaves oven.

In this way, a process of civil engineering [12] as a soil study can be characterized in a protocol composed of several atomic protocols, which have been obtained as a subordinate sequence of elements described from their functional and non-functional requirements.

The composition is an emerging solution, which is successfully applied in other domains such as Web Services [11] and learning routes [12]. It corresponds to a process, which is attained on one of its most promising lines by implementing planning techniques, and the automatic interaction of pre-existing elements to generate new elements that suit the specific needs of a user, when those needs cannot be satisfied by the preexisting elements.

In this paper, the concept of composition of experimental protocols used in the domain of Civil Engineering is presented, using planning techniques in artificial intelligence. In this composition, to consider the functional and non-functional aspects can be useful to overcome the adaptation and contextualization problems of the users aim, formulated as a complex process, and becoming a significant contribution for this domain. As an emergent contribution, it is proposed to model the semantic representation of the documents ASTM as well as the hierarchy of the main composed processes of the civil engineering area. Nevertheless, as a study case for this work, the representation will be focused on presenting the

Manuscript received October 9, 2001. (Write the date on which you submitted your paper for review.) This work was supported in part by the U.S. Department of Commerce under Grant BS123456 (sponsor and financial support acknowledgment goes here). Paper titles should be written in uppercase and lowercase letters, not all uppercase. Avoid writing long formulas with subscripts in the title; short formulas that identify the elements are fine (e.g., "Nd–Fe–B"). Do not write "(Invited)" in the title. Full names of authors are preferred in the author field, but are not required. Put a space between authors' initials.

I-D T. Author, is MsC. and researcher professor at Institución Univeristaria Salazar y Herrera, currently a PhD student at the Universidad Nacional de Colombia, headquarter Medellin (corresponding author to provide phone: 57-4-4600700; Ext: 2218; e-mail: idtorresp@unal.edu.co).

J. F. M-F. Author, is Esp. and MsC. student author, at the Universidad Nacional de Colombia, headquarter Medellin (e-mail: jfmunozf@@unal.edu.co).

J. G-L. Author, is PhD. and MsC A, is professor associate at the Universidad Nacional de Colombia, headquarter Medellin; leader of the research group SINTELWEB (e-mail: jaguzman@unal.edu.co).

ISSN 2395-8618

preliminary phase of the design of pavements.

This paper is organized as follows: the second section describes the relative concept of ASTM standards and the Civil Engineering process. The following section shows the representative model used for describing the experimental protocols under ASTM standard, with ontologies. In the fourth section, we describe the composition process, including the functional and non-functional requirements of the users, and the conclusions are discussed in the fifth section.

II. Civil Engineering Domain

One of the key aspects for the users involved in the domain of Civil Engineering: engineers, laboratory workers, auditors, technicians, etc., is the knowledge that they must have when applying the ASTM standards. Such knowledge will enable them to read and understand the documents of a project, as well as prepare them; and perhaps most importantly, this knowledge will enable them to avoid mistakes in quality assurance and the control of the construction and manufacturing [7]; The ASTM standards can be classified in five big groups: testing methods, material specifications, recommended practices, nomenclature and guides. For the scientific interest, the first group of this work is focused on standards of testing methods, which deter-mine the properties and characteristics of the material; this is the reason why they correspond to experimental practices carried out in laboratories [13]. At present, these standards are applied, hand documented, and written in natural language; and only in some cases, they are partially transcribed to a digital file. It is the applied ASTM standard which ultimately makes it an experimental protocol, that is to say "a sequence of tasks and operations that are represented as recipes providing the description of the processes step by step, described in natural language. Such sequence, in experimental research, is considered a fundamental unit of knowledge" [3]. In daily life, the laboratory workers document every protocol with valuable information, which is related to the used resources, the data employed, the implemented activities, the obtained results, the generated conclusions, together with some other related an-notations. However, if a user wants to accompany a process, as in the case of pavement design. According to the profile, this user must analyze a specific aspect of the result of a sequence of protocols that represent the phase of the addressed study (as shown in Fig. 1). This task must be done manually, aiming to read and understand the documentation (written) involved in the project, and related in this phase. So, in the case of the engineer, he can concentrate on the results of the tests, while an auditor can, using the same information, identify the calculations and tests (just to mention some of them). Even if another general user wants to implement only those protocols that are used by certain equipment, this user must use the manual search. As quoted, each user has different specific interests, even using the same documentation. This is the reason why the normalization of the documented information is valuable to strengthen processes such as the indexation, search, recovery, and their composition, in order to ease the human labor.

III. Representation Model

The formal description of the experiments is necessary in order to obtain an efficient analysis of the investigation, and to allow the exchange of scientific results, turning into a fundamental part of the practice of science. Thus, representing the domain of the experimental protocols, of the field of civil engineering semantically, allows modeling the scientific investigation of the area. Such representation must be unambiguous, reach an agreement, and it must contain all the common and needed elements to reuse the experiments; no matters for whom or for what are they required.

A. Experimental Protocol Ontology

For the goal of proposing a semantic representation that supports the needs of the experimental protocols in civil engineering, it is proposed to implement an ontology.

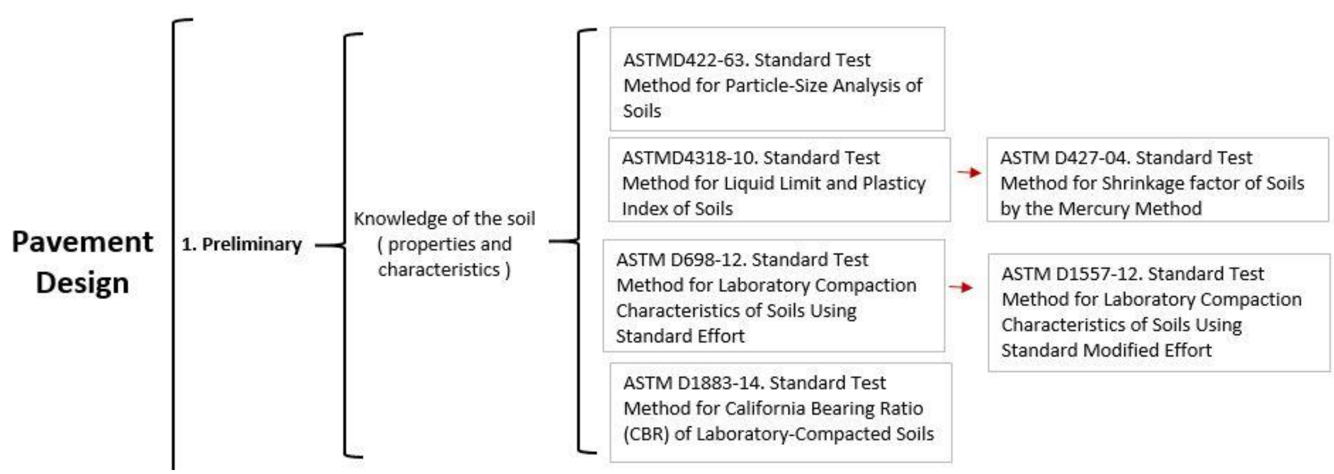

Fig. 1 Portion of the Process of Pavement Design (Preliminary Phase)

ISSN 2395-8618

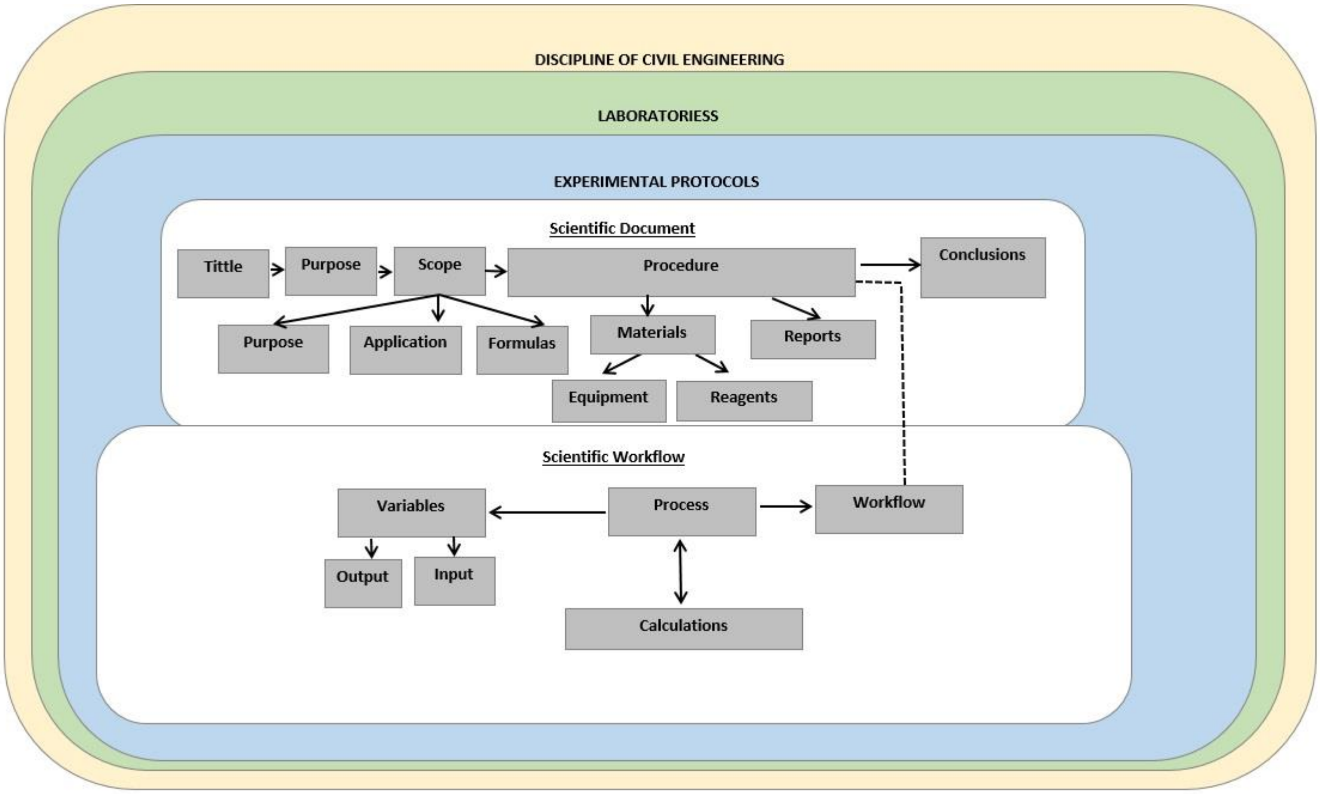

Fig. 2. Experimental protocol, as document and as scientific workflow, inspired [8]

In this stage, it is possible to notice that the experimental protocol is composed of two key elements: a documental structure that represents the steps of the scientific meth-od¸ and a description modeled through a flow of procedures (known as scientific workflow) [4, 5]. The previous description reflects the SMART Protocol [8], which is designed for the domain of the biotechnology of the vegetable plants. The goal is to model the experimental protocol in two parts (see Fig. 2), separating the static information model (called Document Scientific in Fig. 2) from the dynamic model (called Scientific Workflow in Fig 2, too).

The first being represented as a scientific document that registers the basic textual information of the methodology of investigation (modeled as ontology after), and with which it is possible to model non-functional requirements (tittle, objective, purpose, results and other metadata of documenting the experiment, from the steps of the scientific method). Whereas, the second represents the protocol as a flow of processes, commonly known as workflow, i. e. the set of step in an experiment. A Workflow can be defined as a directed graph in which the nodes are implemented by human ("mixed the sample") or software programs (calculations or simulations). This element is used for representing the functional requirements (the Fig. 4 show an example).

The intention of using the ontology was the clearest task so far: The classification and reuse of experimental knowledge generated in the laboratories of the Civil Engineering field. Thus, the most focused recommendation was about reusing the existing ontologies, since by the time of considering the reuse of standards of ontological type in this field, proposals under the scheme of modeling experiments were done. In order to enumerate the most important terms of the domain, the decision aims to reuse the SMART Protocol ontology, obviously extended and adapted to the needs of the ex-perimental civil engineering; thereby generating a PECE-Protocol Experimental Civil Engineering Ontology as document (see Fig. 3). PECE show a model ontology, which enumerates of the set of requirements to be considered, when an experimental proto-col in civil engineering is represented. In this case there are elements that are continuant and others are currents. Between the first are all entities that exists in full at any time in which it exists at all, persists through time, under this specification are all elements of scientific document (see Fig. 2). But in this case these elements are sections of the *class:documentpart;* in the second share, the occurrences entities, are all ele-ments that exist in a given moment of time, scientific workflow (see Fig. 2) *class:workflow*.

For space limits the Fig. 3 shows only a portion of subclasses *class:Entity* and *Class:Continuant* besides of the some elements *subclassOf:DocumentPart* development in Protégé software.

A. *Complex Processes Domain Ontology*

Retaking the idea of dealing with the semantic heterogeneity among users (scientist, engineer, laboratory workers, and auditor), the protocol composition model goes beyond simple ontological domain modeling, and it includes semantic formalization of a complex process.

ISSN 2395-8618

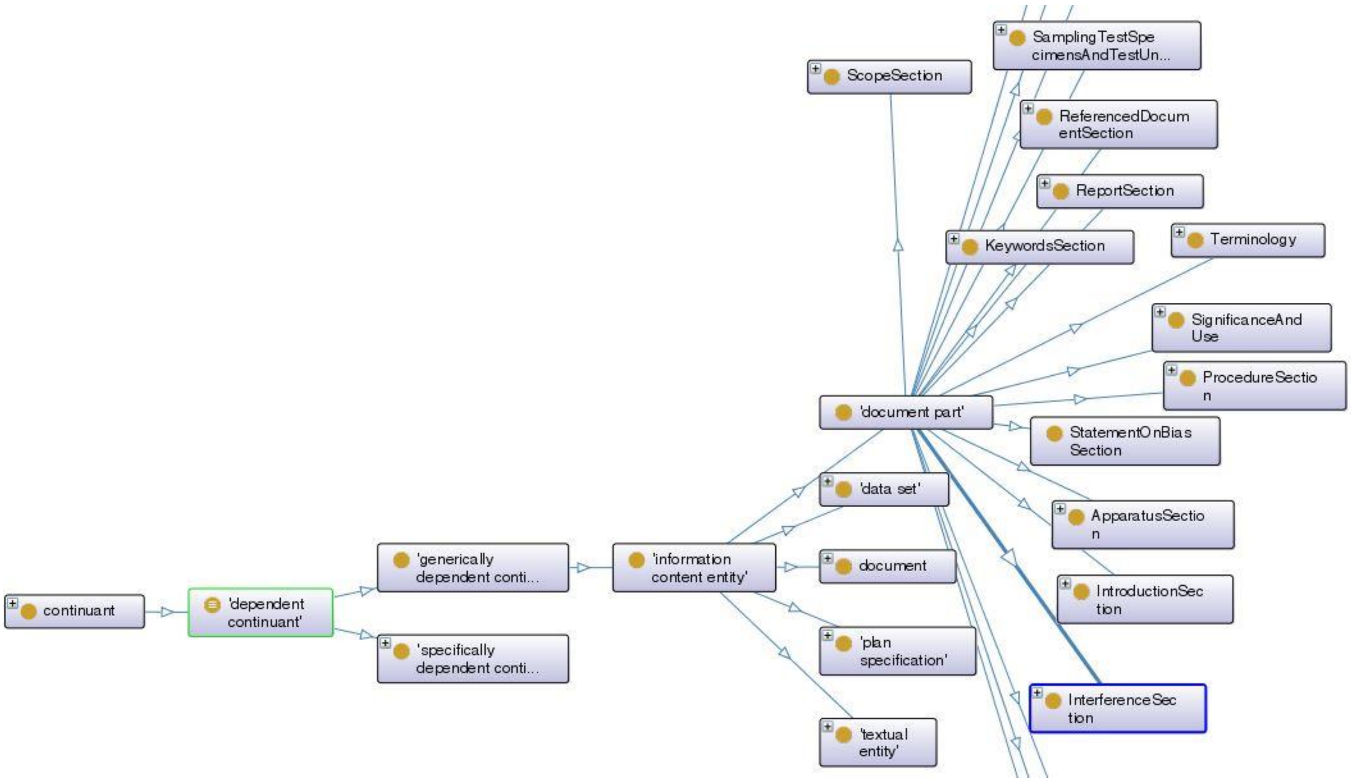

Fig. 3. Portion of Ontology PECE as Document

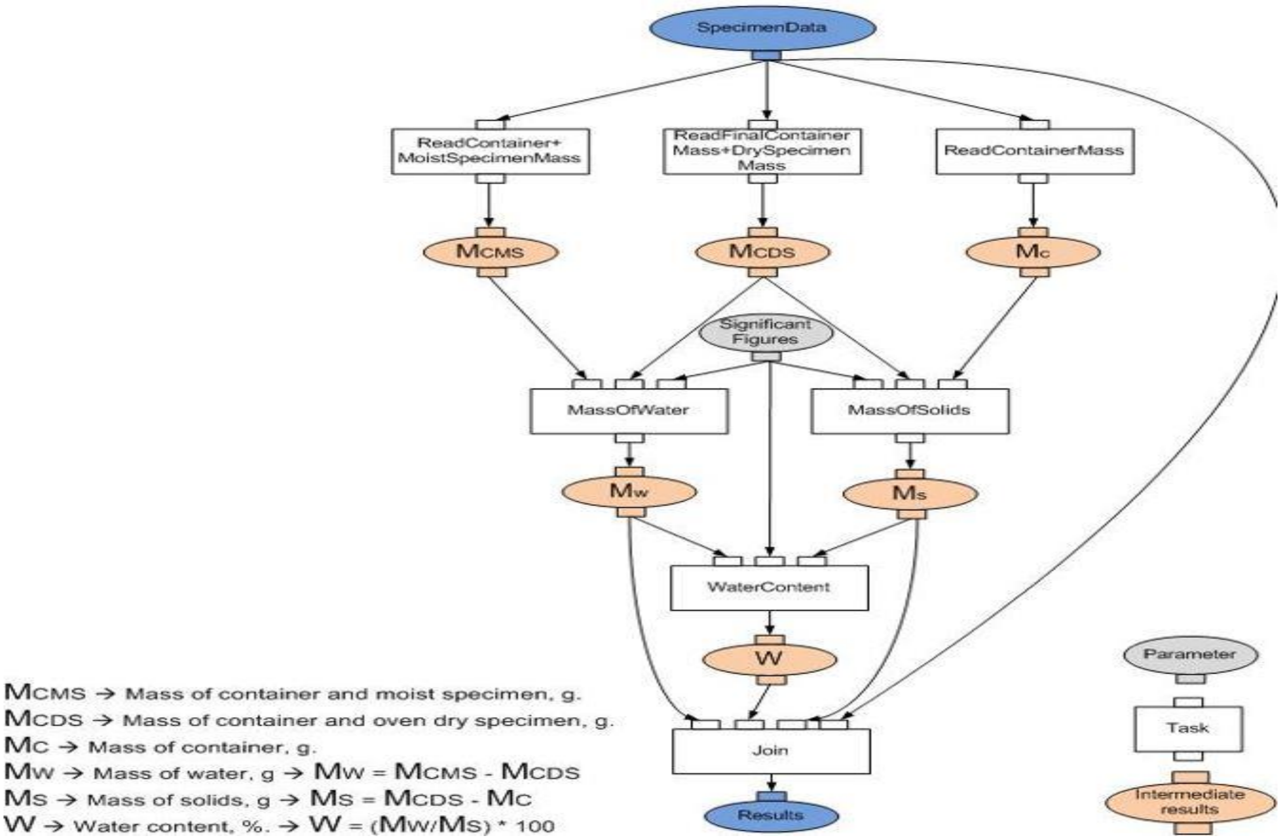

Fig. 4. ASTMD2216-5122 Protocol as Workflow

ISSN 2395-8618

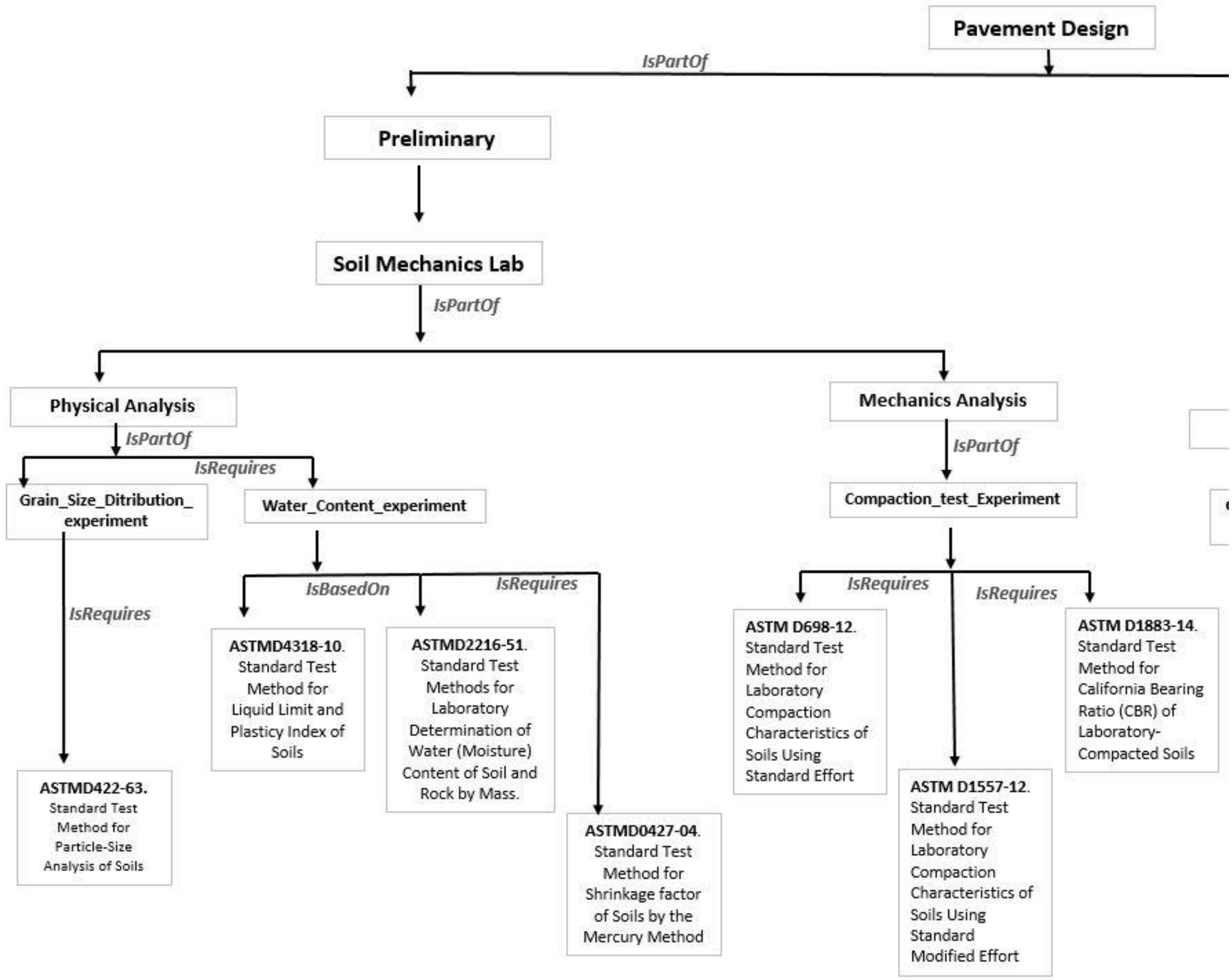

Fig. 5. Portion of Model Complex Process Domain Ontology

It defines a concept like *OWL:Concept* [15] which belongs to an *OWL:Class*. Consequently, it is possible to build a network of concepts depending on the cause-effect relation, which may take place among the various phases of the complex process (See Fig. 5). The idea is to model each complex process as a hierarchical network of tasks where the branches correspond to the last level phases mentioning experimental protocols to be achieved with the application of each ASTM. Order these values enable to define a linear order among the actions, or among the internal or external processes for the satisfaction of the objectives of the user. i) *IsBasedOn*, provides order relations among protocols (ASTM Instanced) and the processes. It can provides order relations among protocols (ASTM Instanced) and the processes; it can be defined as disjunctive "O", since a process can need to know the outcomes provided by a "Y" protocol, also obtained by a "Z" protocol in order to achieve the knowledge goals of the "Y or Z" specific process. This is, then, at least one of the "*OR"* type must have been completed in order to access the following protocol. ii) *Requires*, reports dependencies among type; this happens when a process needs the outputs from another process. This is a hard requirement, that is to say, it must be obtained before addressing the next one. Hierarchy: iii) *IspartOf:* this concept describes a hierarchical structure among the elements of the process. The aim is to model the domain of the complex processes of Civil Engineering, as shown in Fig. 1, through this ontology.

V. The Composition Process

The composition of an experimental protocol, from the point of view of a user's requirement, is directly related to the formal specification of a planning problem, which includes collections of actions with pre-requisites and results (preconditions and effects respectively). Inspired in [3, 11]. When an action supports another, this generates a causal link between them, meaning that the preceding action is finished before starting a succeeding action. With these actions, it is possible to establish a plan of the actions which will be transferred to users (Scientist, engineer, auditor, etc.) from an initial state of knowledge of a complex process in Civil Engineering, to a state in which formulated goals are achieved

ISSN 2395-8618

by users themselves. Expressed analogically, with a civil domain, users express their requirement depending on one or various concepts of domain specific knowledge; these concepts are part of a complex process (as shown in Fig.2). This requirement must be achieved from a state of a user's initial knowledge of that domain (which is even considered null, for the specific case in which an individual knows nothing of a domain); also, including at the same time considerations regarding preference of the non-functional requirements in this case expressed as metadata of ontology as document. Thus, the experimental protocol compound for this case constitutes a (plan) sequence of protocol experimental associated to knowledge domain (the specific Civil complex process, i.e. pavement design, review database buildings, beam analysis, to name a few), and customized for a user.

As an attempt to transfer both specifications, in one sole model (Planning Domain Definition Language PDDL [9]), was born, which was expressed by the (S, So, G, A, R) tuple, where S is a set of all the possible states, defined by the all the concepts of the complex process map. $So \in S$ denotes the initial state of the problem, and is related with one or more, state of the complex process map (concepts of the domain Ontology); $G \in S$ denotes the goal state, defined by the user´s specific complex process, that, he must be achieve. This process, are expressed by a set of concepts associated to the domain ontology of the Civil complex process (Fig. 5), in which the composition system will try to search. A

is a set of actions which represent the protocols which the composer must consider to change from a user state of knowledge, to another state of knowledge, and the translation relation $R \in S \times A \times S$ defines the precondition and effects to execute each action.

As it is deduced from the tuple, S states are defined depending on the concepts associated with the complex process model; this is why, this stage can only be achieved once that, at least one of the experimental protocols of the process modeling has been defined. This composition process, in this case is done in two phases: (i) in order to select the sequence of protocols, that allow to reach the goal (expressed as one or more concept of Domain Ontology). In this phase, the system works with the functional requirements, this is the workflow, expressed by, its inputs and/or outputs. (ii) in order to consider the best adapted assignments to a given user preferences, it is only possible with the non-functional requirements; this is the metadata of the protocol as document.

In the first case, each protocol (ASTM Instanced) is translated to a planning domain as an action, where preconditions will be defined as two types (i) some of knowledge directly associated to Complex Process Domain Ontology concepts and to the assignments; and (ii) some precedence associated to be represented by is required and *isBasisFor*. The effects themselves correspond to the statement that says that the concrete protocol is already known and that task specific has already been carried out (*task_done*). Another important aspect, apart from the order of the protocol on the concept map of the ontology of domain of complex processes, is the sequence of input and output variables associated to the workflow (seen as black box [3]) of each protocol. These are also part of the domain description, incorporating as pre-conditions and effects of knowledge. The above types represent a specification problem, already in PDDL. Fig. 6 represents a portion a domain of pavement. In this Fig., it is possible to appreciate the way, the ASTM are converted in the Domain File expressed as PDDL actions; while Fig. 7 represents the file problem, also in PDDL, in this case it includes all the objects, the initial state, and the goal user´s specification.

```
(define (domain ROPavimentos)
(:requirements :strips :typing)
(:types ro inputvar outputvar process language - object)
(:constants MCMS MCDS MC MW MT MD N PL A R G G1 V VO Mmtf Wtf Pc Mt1 Mmd V1
            K Wc Yw W1 W2 B MassAsreceivedSpecimen
            MassoftheOvenDriedSpecimen - inputvar
            W LL PI P D SL DrySoilMass WetSoilMass Mdtf Pf Pm Pd Yd Wsat
            Gs A1 MoistureContentAirDriedSample AshContent - outputvar
            en es - language
)
(:predicates
 (inputvar_known ?e - ro ?i - inputvar)
 (outputvar_known ?e - ro ?o - outputvar)
 (task_ASTM_D2216_done ?process - process)
 (task_ASTM_D4318_done ?process  - process)
 (task_ASTM_D0422_done ?process  - process)
 (task_ASTM_D0427_done ?process  - process)
 (task_ASTM_D0698_done ?process  - process)
 (task_ASTM_D136_done ?process  - process)
 (task_ASTM_D2974_done ?process  - process)
 (null ?process  - process)
)

(:action ASTM_D2974
     :parameters (?e - ro ?i - inputvar ?o - outputvar ?process - process)
     :precondition (and (inputvar_known ?e MassAsreceivedSpecimen)
                   (inputvar_known ?e MassoftheOvenDriedSpecimen) (null ?process))
     :effect (and (outputvar_known ?e W)(outputvar_known ?e MoistureContentAirDriedSample)
             (outputvar_known ?e AshContent)(task_ASTM_D2974_done ?process))
)
```

Fig. 6. Portion of PDDL domain (pavement)

In the second case, the convenience of the experimental protocol is selected according to the preferences of the user, and it is achieved by building an abstract plan (see Fig. 7), achieved in the previous phase. After the discovery of each protocol in the repository of the protocol, it is necessary to filter each one according to the preference of each user. This is achieved by a semantic matching, each preference with the corresponding metadata. In this case, the selection preferences metadata are shown to the user for him to select the ones of his interest. For instance, the user can only require to analyze the pavement protocols of the preliminary phase, carried out by Universidad Nacional de Colombia (*Metadata Author*), in July 2016 (*Metadata Date*); it may even request the system to report only the ones, whose materials use a microwaves. As it is possible to observe, the preferences of the user, based on interests, must be paired with the metadata of the protocol as a document. For this reason, once the plan is built, as pointed in Fig. 8, the system must search only those elements whose matching is precise to the one formulated by the user. Such specification is expressed as an "*and*", this is to say, each and every preference must be mapped in the repository of protocols. In addition, this matching exercise can show results of two types: i) that one does not exist (or any protocol), or ii) there exist more than one protocol matching the preferences.

ISSN 2395-8618

```
(define (problem P1)
    (:domain ROPavimentos)
    (:objects
        ASTM - ro
        pavimentos - process
    )
    (:init
        (null pavimentos)
        (inputvar_known ASTM MCMS)
        (inputvar_known ASTM MCDS)
        (inputvar_known ASTM MC)
        (inputvar_known ASTM MW)
        (inputvar_known ASTM MT)
        (inputvar_known ASTM MD)
        (inputvar_known ASTM N)
        (inputvar_known ASTM PL)
        (inputvar_known ASTM A)
        (inputvar_known ASTM R)
        (inputvar_known ASTM G)
        (inputvar_known ASTM G1)
        (inputvar_known ASTM V)
        (inputvar_known ASTM VO)
        (inputvar_known ASTM Mmtf)
        (inputvar_known ASTM Wtf)

    )
    (:goal
        (and (outputvar_known ASTM A1)(task_ASTM_D136_done pavimentos)
```

Fig. 7. Portion of PDDL problem

In the first case, the user must change his preferences in order to do a new selection, and if necessary, to build an abstract plan again (it can be considered alternative protocols defined under the relation *IsBasedOn*, being previously ignored). If in this case, the terms that make up the original search of the user do not coincide with the existing concepts in the context base, the agent returns a fault in the request. In the second case, there exists more than one protocol matching the preferences. Conversely, after finding these metadata, the system returns to an RDF document [9] with a list of the triplets related to information about all metadata.

```
2:(astm_d0698 astm ?i ?o pavement)
        6   -> (task_astm_d0427_done pavement)
        3   -> (task_astm_d4318_done pavement)
        0   -> (inputvar_known astm yw)
        0   -> (inputvar_known astm wc)
        0   -> (inputvar_known astm k)
        0   -> (inputvar_known astm v1)
        0   -> (inputvar_known astm mmd)
        0   -> (inputvar_known astm mt1)
        0   -> (inputvar_known astm pc)
        0   -> (inputvar_known astm wtf)
        0   -> (inputvar_known astm mmtf)
;

1:(astm_d136 astm ?i ?o pavement)
        2   -> (task_astm_d0698_done pavement)
        0   -> (inputvar_known astm b)
        0   -> (inputvar_known astm w2)
        0   -> (inputvar_known astm w1)
;

bindings =
{ ?e(3) ?e(5) ?e(6) ?e(4) ?e(2) ?e(1) } == astm
{ ?process(6) ?process(5) ?process(4) ?process(3) ?process(2) ?process(1) } == pavement
;
orderings = { 2<1 3<1 3<2 3<6 4<1 4<2 4<3 4<5 4<6 5<1 5<2 5<3 5<6 6<1 6<2 }
Time: 0
```

Fig. 8. Portion of a plan built

VI. *Discussion and Conclusions*

The technology of the experimental protocols semantically modeled is recent. It emerged from the application of the ideas coming from the Semantic Web field in the experimentation area. This technology is based on adding formal and comprehensible semantic contents, by computer systems, to the description of the capacities of the protocols, in such a way that software entities can process this information the same way that a human would. Although there are very well represented models, it has never been this necessary that each domain adjust its representation to its own requirements; if reusing automatically the results of the experimentation is being sought. Simultaneously, this paper has presented an approximation about the way in which the planning technique can be incorporated to the design of a protocols com-position in the domain of Civil Engineering; specifically when modeling the ASTM standard under functional and non-functional requirements.

The need for new alternatives that enable the reuse of documented results of an experimental research, better adapted to the needs of the role of a user, has motivated the application of planning techniques of Artificial Intelligence in this field. In this regard, all the efforts have been oriented to adapt the way to represent the domain of the protocols semantically, and to identify the best way to achieve eloquence from the existing planners, allowing to deal with these characteristics.

The described model is part of a representation, composition, and execution plat-form of experimental protocols, being the reason why we are working on the best way to adapt the execution process to the functional and non-functional requirements associated to the workflows, and to the semantic representations respectively. Nevertheless, the main effort is represented in the work demanded by the sensitization of the community, when sharing the results of the researches; if it is intended that e-science starts to permeate the scientific experimentation with its benefits in the do-main of Civil Engineering.

Acknowledgment

The research reported in this paper is supported from the investigation thesis “A Integrated Model of logical reasoning, and planning techniques in Artificial Intelligence for Automatic Composition of Research Object based in Experimental Protocols in the Area of Civil Engineering” and “A recommender model of experimental protocols based on user usage context”. Both funded under the Universidad Nacional de Colombia, headquarters Medellin.

References

[1] Asphalt Institute. A Basic Asphalt Emulsion Manual. Manual Series No. 19 (MS-19), (1979).

[2] ASTM Standard. American Society for Testing and Materials - ASTM International. Available: https://www.astm.org/ (1996).

[3] Barreriro D and Albers P.: Approaches of Web Services Composition Comparison Between BPEL4WS and OWL-S. In Proceedings of ICEIS'05(International Conference on Enterprise Information Systems), May 23-27,pp. 208-213, Miami-USA, (2005).

[4] Bechhofer S., Soiland-Reyes S. and Belhajjame K. Workflow Lifecycle Management Initial Requirements. , SWAT4LS, London, (2011)

[5] Beco S., Cantalupo B., Giammarino L., Matskanis N. OWL-WS: a workflow ontology for dynamic grid service composition. First International Conference on e-Science and Grid Computing, 1-1 July (2005).

[6] Corcho O. Research Semantic S-cience. http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/researchareas/3-semanticscience/index.html.

ISSN 2395-8618

[7] Delatte N. Las Normas y la Ingeniería en las Aulas Universitarias. Magazines & Newsletters / ASTM Standardization News. Entrevista Mayo/Junio (2009)

[8] Giraldo O, Garcia A, and Corcho O.: SMART Protocols: Semantic Representations for Experimental Protocols. 4th Workshop on Linked Science - Making Sense Out of Data (LISC2014). 19th or 20th October. Riva del Garda, Trentino, Italy (2014).

[9] McDermott, D., and the AIPS-98 Planning Competition Committee (1998). PDDL-the planning domain definition language. Tech. rep., www.cs.yale.edu/homes/dvm.

[10] Soldatova L. and King R.. An ontology of scientific experiments. Journal List. J. R. Soc Interface. Vol. 3 No. 11. doi: 10.1098/rsif.2006.0134. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1885356/. (2006).

[11] Torres I-D, Guzmán J..: Reactive Planning to Compose Learning Routes in Uncertain Environments. CISSE - University of Bridgeport. December 12-14, (2013).

[12] Torres I-D, Guzmán J.: Applying Case-Based Learning to Improve the Efficiency in the Web Service Compositions. International Journal of Engineering and Technology. vol. 6, no. 3, pp. 227-233, (2014)

[13] Universidad Nacional de Colombia. Laboratorios Universidad Nacional de Colombia, heardquare Medellín. Facultad de Minas. Retrieved Abril, 2014, from http://www.minas.medellin.unal.edu.co/dirlab/index.php/laboratorios. (2000).

[14] Verdes-Montenegro L.. E-Science for the SKA, WF4Ever: Supporting Reuse and Reproducibility in Experimental Science RadioNet Advanced Radio Astronomy, Commissioning Skills and Preparation for the SKA Manchester: November 15th (2012)

[15] OWL, Web Ontology Language, https://www.w3.org/OWL/.