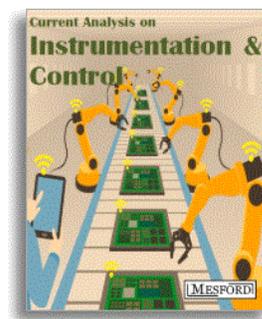


Automation of a Learning Process through Semantic-based Process Modelling and Reasoning Method

Kingsley Okoye*

School of Architecture Computing and Engineering, College of Arts Technologies and Innovation, University of East London, Docklands Campus, London, UK. E16 2RD.



Abstract:

Process modelling involves the extraction and synthesising of different kinds of data and its attributes that can be found in any given process domain. To ensure the usefulness of the captured datasets and informative value of the derived models, there is also need for technologies that have the capability of describing the relationships that hold between the different process instances or elements. For this purpose, the work in this paper proposes a semantic rule-based process mining approach that is directed towards discovering of meaningful patterns or models from any process events log and then respond by making decisions based on the resultant models or semantic information base. Practically, the work applies the method to the learning process settings in order to generalise and validate the proposed approach. Theoretically, the paper provides an effective way of using the semantic technologies (e.g ontological models) that is capable of automatically computing the various activities or patterns within the learning knowledge-base and to check the consistency of the defined object/data types and assertions. Thus, the method is grounded on inductive and deductive logic expressions (descriptions) that allow for the use of a reasoner to check that all the definitions in the resultant learning model are consistent, and can also determine (infer) which concepts that fit within each defined class. In fact, the inductive reasoning aptitude is applied in order to discover the various kinds concepts or relationships (e.g learner categories), while the deductive approach is used to verify and enrich the discovered patterns and/or rule expressions.

Publication History: Received: 09 November 2018 | Revised: 31 December 2018 | Accepted: 21 December 2018

Keywords:

Process Modelling, Process Mining, Learning Process, Ontology, Semantic Reasoning.

INTRODUCTION

The swift shift from the conventional big data to big data analysis is inundated by the ever-increasing volumes of data that are being recorded at an unprecedented rate in today's information systems. This spans the need for automatic computing or reasoning methods that can be applied to make sense of the said datasets (particularly as it concerns the data values). Perhaps, the semantic technologies are one of the scientifically proven techniques that are used to model different kinds and structure of the activities, events or processes as they happen in real-time. For instance, ontologies can be layered on top of existing information asset to provide a more formal expressiveness and/or enhancements to processes in real-time settings [1]. Indeed, the ontological concepts present to the data science community - the capability of using semantics to classify instances (process elements) in order to

explain the dependent variables in terms of independent ones. Moreover, the semantic annotations and reasoning aptitudes make it possible to match same ideas as well as use the coherence and structure itself to inform and answer questions about relationships the process instances share within an information knowledge base.

Equally, in terms of the semantic-based process mining and modelling approach - the various activities within a learning process can be related to exactly one case and assigned a case identifier [2]. The method results in the automatic creation of a workflow process for the individual activities [3] and can help to maintain the resulting hierarchy correctly. Indeed, the work in this paper shows that the automatic creation of the workflows can be achieved by using the ontology schema and/or annotations to represent the sets of various entities (classes, object/data properties, individuals, axioms or

*Address correspondence to this author School of Architecture Computing and Engineering, College of Arts Technologies and Innovation, University of East London, Docklands Campus, London, UK. E16 2RD; E-mail: k.okoye@uel.ac.uk

Mesford Publisher Inc

Office Address: Suite 2205, 350 Webb Drive, Mississauga, ON L5B3W4, Canada; T: +1 (647) 7109849 | E: caic@mesford.ca, contact@mesford.ca, <https://mesford.ca/journals/caic/>

relations) that can be found within the learning knowledge base, and then make use of the reasoner to classify and create inferences capable of providing new knowledge, or better still, a richer set of intelligence within the derived model. In other words, this work uses the process mining techniques in combination with semantic modelling (ontology schema/vocabularies) methods to model and discover sets of attributes and/or relationships that can be found within a process domain – using the case study of the learning process. Consequently, suitable learning paths were determined by means of the semantic reasoning aptitudes. Clearly, the method addresses the problem of extracting useful patterns from the captured datasets to the provision of useful and valuable knowledge (information) about the processes in view.

The rest of the paper is structured as follows: in Section 2, appropriate related works are analysed and discussed. Section 3, presents a description of the learning process modelling method and how we apply the representations to draw conclusions and make predictions based on the analysis of the available data. Section 4 illustrates the semantic process modelling technique - describing in detail the method for semantic representation and reasoning using the ontological schema. The prototype implementation and preliminary outcomes are discussed in Section 5. Finally, Section 6 concludes the paper and points out directions for future research.

2. RELATED WORKS

The effective use of semantic technologies can solve the problem of regulating the ever-changing, and yet, static measures of knowledge at both theoretical and technological levels [4]. The mechanism has been proven to improve and enhance the capability of process models by making inferences, retaining and applying what they have learned as well as the discovery of new information or processes.

The authors in [5] note that various process modelling and automation methods have been proposed in the literature which is directed towards obtaining more expressive models from knowledge bases [6]. Reference [5] argue that classification is a fundamental task for a lot of intelligent systems and that classifying through logical reasoning may be both too demanding and frail because of inherent incompleteness and complexity in the said knowledge-bases. However, the authors observe that these methods adopt the availability of an initial drawing of ontology that can be automatically enhanced by adding or refining concepts, and have been shown to effectively solve process modelling problems [1][3][6] using process description logics and/or queries [7] particularly those based on classification, clustering and ranking of individuals.

Likewise, the learning process modelling and analysis has been tackled over the years by adapting machine learning (ML) methods such as Instance Based Learning [8] and Support Vector Machine (SVM) [9] to Description Logics (DLs) queries [7] - which is recently the standard theoretical foundation upon which semantic technologies or languages such as the

Ontology Web Language (OWL) [10] and Semantic Web Rule Language (SWRL) [11] are built as well as used in the literature.

According to [12] and [13], Bayesian models have paved way for new machine learning algorithms with more powerful and more human-like capabilities. Perhaps, the semantic web technologies and its main application in real-time cannot be explained without mentioning the Bayesian theory of probability [14][15]. In fact, the Bayesian probabilistic theory has been proven to be one of the few mathematical interpretation of predictive concepts used for representing a state of knowledge. Thus, an extension of logic proposals that enables reasoning with the hypothesis whose true or false values is uncertain. Moreover, the Bayesian models are based on 3 vital probes: (i) what are the content of probabilistic theories? (ii) how can they be used to support reasoning? and (iii) how can they themselves be reasoned upon? Accordingly, the hypothesis is measured by computing the Bayes' rule; where the:

Probability, $P(x|h, T)$ - measures how well each argument predicts the data and the initial marking or likelihood.

Whereas, the $P(h|T)$ expresses the plausibility of the hypothesis given the users background knowledge, and

The posterior probability, $P(h|x, T)$ – which is proportional to the result of the two expressions representing the level of certainty in each of the hypothesis given both the constraints of the background theory T , and observed data x .

On the other hand, [16] notes that the challenge comes in specifying hypothesis and probability distributions that support Bayesian inference for a given task or domain. Interestingly, the authors argue that both structured knowledge and statistical inference are necessary to explain the nature, use, and acquisition of such human knowledge, and further introduced a theory-based Bayesian framework for modelling inductive learning and reasoning. Explicitly, the results in [16] and [17] shows that the problems of modelling learning processes can be solved by transforming the ontology population problem to a classification problem where - for each entity within the resolutely defined ontologies; the concepts (i.e. classes) to which the entities belong to have to be determined (i.e. classified) [5]. Generally, the approaches assume that there already exists a probabilistic and/or fuzzy knowledge-base upon which the proposed methods are able to predict the patterns/behaviour (e.g. classification or identification of newly but not previously observed patterns or behaviours).

Indeed, the inductive and deductive reasoning methods can be used as a building block towards the development of probabilistic and automated process knowledge-bases. The methods are achieved by learning the probability that an inclusion axiom or concept assertion holds between two objects. Besides, [5] argues that in presence of noisy and inconsistent knowledge-bases that could be highly probable in a distributed environment such as the world wide web, that deductive reasoning is no more applicable since it requires correct (i.e. true) premises. Although, if all premises are true and the rules of the deductive logic are followed, then the

conclusion reached is necessarily true. On the other hand, inductive reasoning which is grounded on the generalisation of specific process instances and assertions rather than correct premises - allows the formulation of conclusions even when inconsistent or noisy knowledge bases are being considered. Reference [18] is even more specific about concepts generalisation capabilities of the inductive approach. According to the authors [18] – the aim of the inductive learning methods is to infer general ideologies and/or values from specific facts (axioms) or instances (process elements) through the consideration of some kinds of background information or knowledge. Therefore, unlike deductive reasoning, inductive reasoning allows for the possibility that the conclusion is false, even if all of the premises are true, and does not rely on universal restrictions over a closed axiom to draw conclusions. Scientifically, inductive reasoning is the main practice for logical reasoning (e.g. ontologies, description logic queries and classifiers, SWRL syntax or format etc.) obtaining conclusions that are believed by the scientific community to be the most probable explanation of observed phenomena.

In general, reasoning on ontological knowledge plays an important role in the semantic representation of processes (e.g. the learning process). The method is important because the semantic reasoning aptitudes allow for the extraction and conversion of explicit information into some implicit information. For instance, the intersection or union of classes, the definition of axioms and entities relationship, and concepts/role assertions. Moreover, [19] describes such logical intersection or relations between the process instances as Workflow Activity Patterns (WAPS) [2] - which are common structures involving the interaction between individual entities and the control-flow constructs used to model the semantics of the activities being performed. Perhaps, the workflow systems assume that a process can be divided into small, unitary actions called activities [3]. To perform a given process, one must perform the set (or perhaps a subset) of the activities that comprise it. Hence, an activity is an action that is a semantic unit at some level, which can be thought of as a function that modifies the state of the process in terms of the semantics of the patterns and can be discovered automatically by means of semantic reasoning [1].

3. ONTOLOGICAL DESCRIPTION AND MODELLING OF DOMAIN PROCESSES

The ontological description and modelling of any given domain process is technically based on computer logic programming [20] and has been related to the natural process of human thinking. The work in [21] notes that inductive intelligent is made of the process of reasoning from the particular to the general through observation of particular events or data logs. The method associates new contents with prior knowledge which can lead to unrelated data being discovered, examined, and further grouped or labelled in order to draw conclusions as well as make predictions based on the analysis of the data.

According to [22], the ability to analyse information and create concepts is fundamental to the ontological reasoning process

and can be applied toward the automation of any given process domain (e.g. the learning processes)

Practically, the following steps/procedures are applied to support the conceptual analysis process and method of this paper, and can be applied to any given process domain independent the analysis questions or metrics as follows:

Step 1: examine the process knowledge base to determine unrelated entities.

Step 2: group entities with common attributes and ascertain/provide descriptive labels for the objects or datatype properties.

Step 3: identify relationships (taxonomies or class hierarchies) in order to generalise, predict and extract patterns from the existing properties within the knowledge-base.

Step 4: apply discovered patterns to a new and/or different context to demonstrate understanding (model validation)

Step 5: check that all facts (axioms) within the discovered classes or model is true and at least falls within the universal restriction of validity by definition and that there is no inconsistency of data or repeatable contradicting discovery.

Indeed, the purpose of the semantic-based method is to utilize the basic concept of ontological modelling to define and understand the process (in view) reality based on the discovered knowledge against the historical data. In other words, the ability to provide a link between the learning objects (properties, classes, individuals) for instance, and the data used in the discovery of the model.

3.1. Concept Matching and Definition of Variables

Association Rule Learning [2] is one of the other data mining techniques that aims at finding rules that can be used to predict the value of some response variables that has been identified as important just like decision systems [23] but without focusing on a particular response variable. The method aims at creating rules of the form.

IF X THEN Y

Where X is often called the antecedent and Y the consequent [2][24] Hence, $X \Rightarrow Y$

Moreover, the rule is similar and can be related to the semantic web rule language [11] used to provide a more improved class expression or ontological description to the discovered process models.

In terms of syntax, the SWRL rules are represented in the following form:

atom ^ atom (antecedent)..... \rightarrow atom ^ atom (consequent).

In theory, the association rule learning strongly supports the use of metrics frequently expressed in the form of support and confidence. These expressions help in measurement of the strength of the association (relations) between the learning objects.

On one hand, Support determines how often a rule is applicable to a given data set. Thus, the fraction of instances for which both antecedent and consequent hold. Perhaps, a rule with high support is more useful than a rule with low support. For instance, a rule that has low support may occur simply by chance and is likely to be irrelevant from a learning process perspective because it may not be profitable to monitor, recommend and/or promote learning activities or learning patterns. However, Support can be used to evaluate the learning process models and its execution;

Where:

N_x is the number of instances for which, x , learning activity holds.

N_y the number of instances for which learning activity y holds, and

$N_{x,y}$ is the number of instances for which activity x and y holds.

Consequently, support for the rule $X \Rightarrow Y$ is described as

Support, $s(X \square Y) = N_{x,y}/N$:: where N is the total number of instances.

On the other hand, Confidence is used to measure the reliability of the inference made by a rule over a given process in question. Thus, for a given rule of the form, $X \Rightarrow Y$, the higher the confidence, the more likely it is for the consequent Y (learning pattern extension) to be discovered within the learning process that contains X (learning patterns). In other words, confidence measures the conditional probability that the extension Y will happen given X .

Hence, Confidence, $c(X \square Y) = N_{x,y}/N_x$

In general, inferences made using association rule learning technique could suggest co-occurrence of relationships between items in the antecedent (X) and consequent (Y) of any rule. Therefore, for every given set of activities or item set, there exist rules having support $\geq \text{minSup}$ and/or confidence $\geq \text{minConf}$.

Where: minSup and minConf are the corresponding support and confidence thresholds, respectively.

Likewise, with the learning process models, these metrics can be used to effectively reduce the exploration or drilling down of space (simple models) when constructing the set of frequent activity logs. The simple requisite is that X and Y are non-empty and any variable appears at most once in X or Y . For instance, the following rule can be discovered in order to provide a more formal definition and/or improve the abstract analysis of the model.

IF Learner(X) AND hasLearning_Activities THEN hasPartLearning_Process(Y)

Thus, Learner(X), hasLearning_Activities(X , Activity) \rightarrow hasPartLearning_Process(Y)

Technically, the approach has been used to provide process specification and expressive language formats that are logical and fundamental to knowledge representation. For instance,

the Knowledge Interchange Format (KIF) [25] which makes it possible to understand the meaning of class expressions through declarative semantics. For example, with the KIF format, it can be expressed that “Every Learner has a Learning_Activity”. Thus, the expression:

```
( forall ( ?X )
  ( => ( Learner ?X )
    ( exists ( ?Y )
      ( and ( someActivity ?Y )
        ( Learning_Activity ?X ?Y ) ) ) ) ) )
```

Consequently, Every Learning_Activity is part of a Learning_Process and must have some kind of a Learner. Furthermore, the expression;

```
( forall ( ?X ?Y )
  ( => ( Learning_Process ?X ?Y )
    ( and ( someLearner ?X )
      ( someLearning_Activity ?X ?Y ) ) ) ) )
```

In fact, the aforementioned rule expressions suggest that a strong relationship exist between the Learning_Process and the Learner. This is because Learner(X) has_Activities described as a Learning_Activity, and Learning_Activity has been described as PartOfLearning_Process.

In turn, designers of knowledge base systems can use this type of rule expressions to help identify new opportunities especially for enhancement of the process models. Besides, the association rule learning is currently now being used in application domains such as the web mining and big data analysis. This is owing to the fact that the association of patterns (or similar attributes) helps in revealing interesting connection (relationships) among domain entities (especially the individual classes, and object/data types) to provide a better understanding of how the different elements within process knowledge-base relate and interact with each other.

Over the next section, the work describes and implement the steps for the ontological modelling and reasoning of the learning process activities capable of deducing inferences based on such design rule-base, thus, the semantic approach to automated learning. The focus is on defining the learning objects properties (restrictions) used for implementation of the different classes and relationships within the learning knowledge-base.

4. METHOD FOR ONTOLOGICAL MODELLING AND AUTOMATION OF THE LEARNING PROCESS (INDUCTIVE AND DEDUCTIVE REASONING APPROACH)

In terms of logic and ontological vocabularies - the inductive reasoning methods have as input data type from which a possible someValuesFrom or believable generalisation is computed. The technique is considered to be an existential restriction, which describes a set of process instances (individuals) that have at least one specific kind of relationship

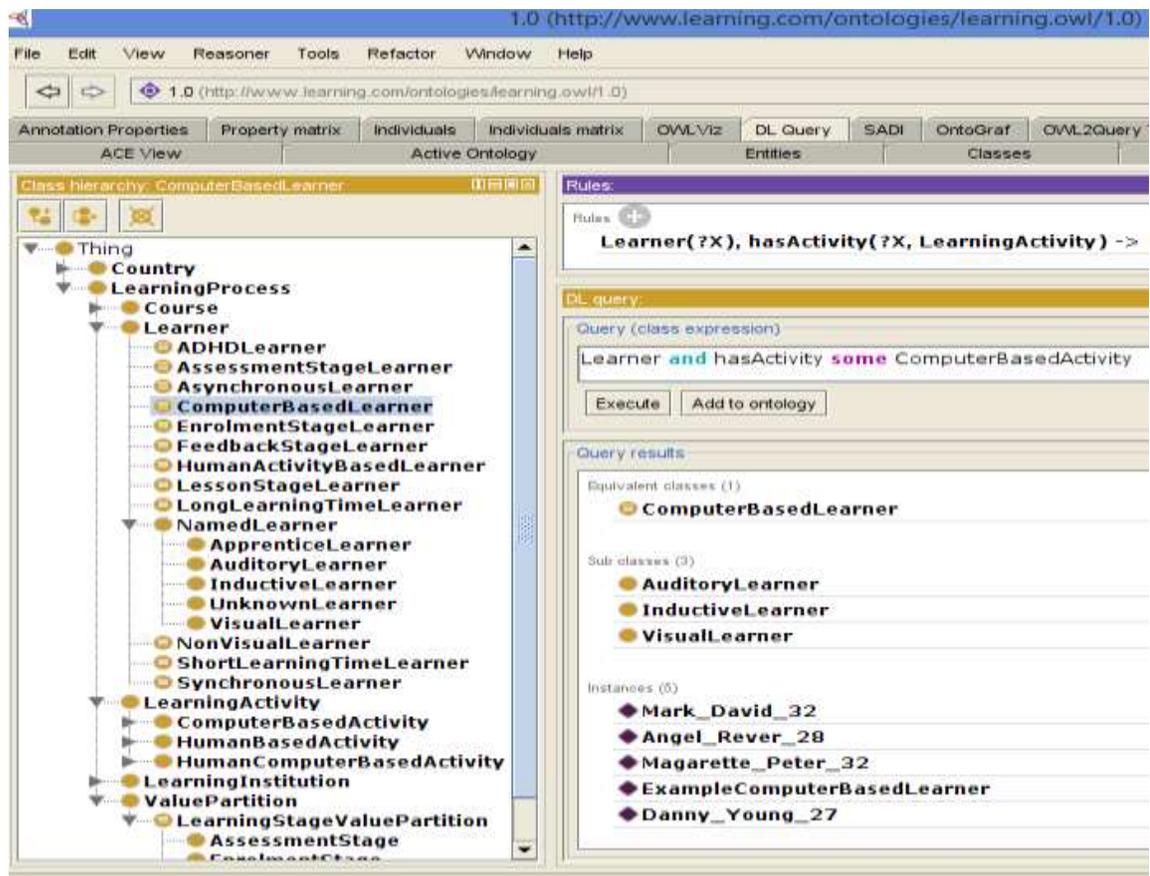


Fig. (1). The Learning Model in Protégé 4.3 Editor with SWRL Rule and DL Query.

to other instances that are a member of a specific class. In other words, it is a relationship that exists between two individuals (i.e. concept assertion that hold between two objects). On the other hand, deductive reasoning which has been generally adopted in the semantic web context assumes an allValuesFrom restriction whereby given a set of general axioms - precise and definite conclusions are drawn through the use of a formal proof (axioms). Moreover, the technique can also be referred to as universal restriction, which describes the set of instances (individuals) where for a given object property/criteria only have specified relationships to individuals that are members of a specific class.

Fig. (1) is an OWL version 2 model for the Learning model ontology used for the purpose of the work in this paper. The model has been developed and implemented in Protégé 4.3 and reasoned using Pellet 2. Moreover, protégé OWL editor [26] supports Description Logic (DL) Queries [7] and the SWRL rules [11] described in this paper

As shown in the figure (Fig. 1) the work uses the protégé Editor to construct a learning process ontology that expresses the functionality of the learning model in terms of the individual learning characteristics (activities or events). The Cases (learning categories) within the model were defined as sub-class of the main class LearningProcess. The class expression (taxonomies) is based on the OWL syntax primarily focused on collecting all information about a particular class or individual into a single construct, called a frame. Furthermore,

the DL Query provides the platform for searching the classified ontology to infer the learning activities of any named individual. The result of the logic expression and reasoning is what we use to show the process model and automated discovery of learning patterns. Indeed, the tactics aim at discovering rules similar to the association rule learning [27], but then without focusing on a particular variable to discover user interaction patterns and/or response by making decisions based on the expressive rules that are centred on the captured user profiles. Clearly, the goal is to discover and create rules of the form IF-THEN as noted in [28]. Thus;

$$X \Rightarrow Y \text{ (IF } X \text{ THEN } Y)$$

where X = Learning pattern (Antecedent) and Y= Learning pattern extension (Consequent)

e.g Learner (?X) , hasActivity (?X, LearningActivity) -> haspartLearningProcess (?X)

Learner (?X) , hasLearningActivity (?X, ComputerBasedActivity) -> isComputerBasedLearner (?X)

For example, driven by the variables as defined in the learning model ontology (Fig. 1), the resulting rules expressions as shown in the following Fig. 2 were derived to improve the conceptual analysis and/or semantic knowledge about the learning process.

Fig. (3) is an example of OWL 2 XML file format for the defined learning process model and ontologies.

ACE Snippets	ACE Word Usage	ACE Q&A	ACE Lexicon	ACE Text
ACE Snippets: 100 snippets (all shown)				
Find snippet by: <input checked="" type="radio"/> Highlight <input type="radio"/> Filter				
Snippet				
No ComputerBasedActivity is a HumanBasedActivity.				
If X hasActivities Y then X hasUnits Y.				
If X hasInstitutions Y then X hasUnits Y.				
If X isUnitOfs Y then Y hasUnits X. If X hasUnits Y then Y isUnitOfs X.				
If X hasInstitutions Y then Y isInstitutionOfs X. If X isInstitutionOfs Y then Y hasInst				
If X isInstitutionOfs Y then X isUnitOfs Y.				
If X isActivityOfs Y then X isUnitOfs Y.				
If X hasActivities Y then Y isActivityOfs X. If X isActivityOfs Y then Y hasActivities X.				
If X hasUnits something that hasUnits Y then X hasUnits Y.				
If X isUnitOfs something that isUnitOfs Y then X isUnitOfs Y.				
Everything hasInstitutions at most 1 thing.				
Everything isInstitutionOfs at most 1 thing.				
Everything that is hasActivated by something is a Learning_Activity.				
Everything that hasActivities something is a Learner.				
Everything that isActivityOfs something is a Learning_Activity.				
Everything that is isActivityOfed by something is a Learner.				
Everything that hasInstitutions something is a Learner.				
Everything that is hasInstitutioned by something is a LearningInstitution.				
Everything that isInstitutionOfs something is a LearningInstitution.				
Everything that is isInstitutionOfed by something is a Learner.				
Every Learner hasInstitutions a LearningInstitution.				

Fig. (2). Example of Rule expressions for learning activities as defined in the Learning Model.

```

<?xml version="1.0"?>
<ontology xmlns="http://www.w3.org/2002/07/owl#"
  xmlns:base="http://www.semanticweb.org/kingsley/ontologies/2014/11/learning-ontology-25"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  ontology:IRI="http://www.semanticweb.org/kingsley/ontologies/2014/11/learning-ontology-25"
  version:IRI="http://www.semanticweb.org/kingsley/ontologies/2014/11/learning-ontology-25">
  <annotation>
    <annotationProperty abbreviatedIRI="rdfs:comment"/>
    <literal datatypeIRI="http://www.w3.org/2000/01/rdf-schema#PlainLiteral">Learning Ontology that classifies the various Units
of Learning Process based on the Learning Activities</literal>
  </annotation>
  <EquivalentClasses>
    <class IRI="http://www.learning.com/ontologies/learning.owl#AsynchronousLearner"/>
    <objectIntersectionOf>
      <class IRI="http://www.learning.com/ontologies/learning.owl#Learner"/>
      <objectAllValuesFrom>
        <objectProperty IRI="http://www.learning.com/ontologies/learning.owl#hasActivity"/>
        <class IRI="http://www.learning.com/ontologies/learning.owl#AsynchronousActivity"/>
      </objectAllValuesFrom>
    </objectIntersectionOf>
  </EquivalentClasses>
  <EquivalentClasses>
    <class IRI="http://www.learning.com/ontologies/learning.owl#ComputerBasedLearner"/>
    <objectIntersectionOf>
      <class IRI="http://www.learning.com/ontologies/learning.owl#Learner"/>
      <objectSomeValuesFrom>
        <objectProperty IRI="http://www.learning.com/ontologies/learning.owl#hasActivity"/>
        <class IRI="http://www.learning.com/ontologies/learning.owl#ComputerBasedActivity"/>
      </objectSomeValuesFrom>
    </objectIntersectionOf>
  </EquivalentClasses>
</ontology>

```

Fig. (3). A fragment of the Learning Ontology OWL 2 XML file with Protégé.

5. DISCUSSION

Process automation or modelling of any given real-time process (e.g. the learning process) involves the visualization or mapping of the flow of activities within the learning knowledge-base - technically described as a workflow. Moreover, being able to use the reasoner to automatically compute the different class hierarchy (taxonomies) of activities within the underlying knowledge-base or model is one of the major benefits of building ontologies using the semantic technologies such as the OWL, SWRL and DL Queries.

Practically, the semantic-based method for annotation and properties assertions are used to add useful information (i.e. Metadata – data about data) to the resultant model. In other

words, the proposed method for developing the learning process model and ontologies in this paper - allows the meaning of object and data properties to be enhanced through the use of property descriptions and classification of discoverable entities.

In fact, the work makes use of the main function offered by the reasoner to ensure and check for consistency in the model. For instance, to test whether or not a class is a subclass of another class, or checking whether or not it is possible for a class to have any instances. Perhaps, a class is said to be inconsistent if it does not have any instances. Moreover, in addition to performing the model consistency test (i.e. classification), it also becomes possible for us to compute the inferred activity hierarchies to determine useful patterns within

the model. For example, inferring the object property assertions (for a given class) may mean that there can be at least one individual that is related to the class by means of the restriction. Besides, with OWL models, property restrictions are used to describe for instance - specific class of individuals or process instances based on the relationship the members of the class participate in.

Typically, in the learning model that was developed for the purpose of the work in this paper - we describe the class Learner to be a subclass of the LearningProcess. The necessary condition is: if something is a Learner, it is necessary for it to be a participant of the class LearningProcess and necessary for it to have a kind of sufficiently defined condition and relationship with other classes e.g. LearningActivity, LearningInstitution, Course, LearningStageValuePartition etc.

For instance, as gathered in Fig. (1) and the rule expressions in section 4 - we show that:

- ComputerBasedLearner is a subclass of, amongst other NamedLearners, a Learner, and
- also a subclass of the LearningActivity class that have at least one Activity that is ComputerBased.

Obviously, the assertions are achieved through the restriction property. In other words, the properties restriction is used to infer anonymous classes (Unnamed classes) that contains all of the individuals that satisfy the restriction. In essence, all of the individuals that have the relationship required to be a member of the specified class. Perhaps, the necessary and sufficient Condition makes it possible to implement and check for consistency in the model - which means that it is necessary to fulfil the condition of Object/Data Property Restriction for any individual to become a member of a class.

6. CONCLUSION AND FUTURE WORK

The work in this paper makes use of the ontological schema or vocabularies to develop and propose a semantic rule-based process mining or modelling method that leads to automated computing of different patterns within a learning process knowledge-base. In short, the approach makes use of the three main building blocks - annotated events logs/models, ontologies, and semantic reasoning to propose a method that is used to address the problem of determining the presence of different patterns within a process base (using the case study of the learning process domain). The results of the experiments show that any pattern or learning behaviour can be discovered as a consequence or condition of a rule. In essence, the ontology provides us with benefits in discovery, flexible access, and information integration due to the inherent connectedness (inference), concept matching and reasoning capabilities. Indeed, such characteristic is the ability to match same idea as well as the use of the coherence and structure itself to inform and answer questions about relationships the learning objects (process instances) share amongst themselves within the learning knowledge-base.

Future work will focus on applying the approach described in this paper to a different process domain or case study in order

to generalise and provide more validation to the proposed method. The aim is to cover the whole spectrum of the approach presented in this paper to help provide a more effective method for big data analysis in literature.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

- [1]. Okoye K, Naeem U, and Islam S. Semantic Fuzzy Mining: Enhancement of process models and event logs analysis from Syntactic to Conceptual Level. *International Journal of Hybrid Intelligent System (IJHIS)* 2017; 14 (1-2): 67-98.
- [2]. Van der Aalst. W.M.P. *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Springer 2011.
- [3]. Ferreira DL, Thom LH. A Semantic Approach to the Discovery of Workflow Activity Patterns in Event Logs. *Journal of Business Process Integration and Management* 2012; 6(1)
- [4]. Barbieri D, Braga D, Ceri S, Valle ED, Yi Huang, Tresp V, Rettinger A, Wermser H. Deductive and Inductive Stream Reasoning for Semantic Social Media Analytics. *Intelligent Systems, IEEE* 2010; 25(6): 32-41.
- [5]. d'Amato C, Fanizzi N, Esposito F. Query answering and ontology population: An inductive approach. In Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M., eds. *Proceedings of the 5th European Semantic Web Conference, ESWC 2008; 5021 of LNCS*, Springer: 288-302.
- [6]. Lehmann J, Hitzler P. Concept learning in description logics using refinement operators. *Machine Learning* 2010; 78: 203-250
- [7]. Baader F, Calvanese D, McGuinness DL, Nardi D, Patel-Schneider PF.: editors. *Description Logic Handbook*. Cambridge University Press (2003)
- [8]. Mitchell T. *Machine Learning* McGraw Hill (1997)
- [9]. Shawe-Taylor J, Cristianini N. editors. *Support Vector Machines & other kernel-based learning methods*. Cambridge Uni. Press (2000)
- [10]. Horrocks I, Patel-Schneider PF, McGuinness DL, Welty CA. Owl: a description logic based ontology language for the semantic web. In: F. Baader, et al. eds. *The Description Logic Handbook: Theory, Implementation, and Applications*. 2 ed. New York: Cambridge University Press 2007; 458-486.
- [11]. Horrocks I, Patel-Schneider PF, Boley H, Tabet S, Groszof B, Dean, M. SWRL A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission [Internet] 2004. Available from: <http://www.w3.org/Submission/SWRL/> Accessed [November 2018]
- [12]. Kemp CC, Griffiths TL, Stromsten S, Tenenbaum JB. Semi-supervised learning with trees. *Advances in Neural Information Processing Systems* 2004; 16:257-264.
- [13]. Griffiths TL, Chater N, Kemp C, Perfors A, Tenenbaum JB. Probabilistic models of cognition: exploring representations & inductive biases. *Trends in Cognitive Sciences* 2010; 14(8): 357-364.
- [14]. McGrayne SB.: *The Theory That Would Not Die*, at Google Books 2011; 10.
- [15]. Fienberg SE.: When did Bayesian Inference become "Bayesian"? *Bayesian Analysis* 2006;1(1): 1-40
- [16]. Tenenbaum JB, Griffiths TL, Kemp C.: Theory-based Bayesian models of inductive learning and reasoning. In: *Trends in Cognitive Sciences* 2006; 10(7): 309-318.
- [17]. Okoye K, Tawil ARH, Naeem U, Islam S, Lamine E. Semantic-based Model Analysis towards Enhancing Information Values of Process Mining: Case Study of Learning Process Domain. In: A. Abraham, A. K. Cherukuri, A. M. Madureira & A. K. Munda, editors. *Advances in Intelligent Systems and Computing (AISC) vol. 614. Proceedings of SoCPaR 2016*. Springer AG. pp. 622-633 (2018)
- [18]. Bühmann L, Lehmann J, Westphal P. DL-Learner-A framework for inductive learning on the Semantic Web. *Web Semantics* 2016; 39:15-24.
- [19]. Thom LH, Reichert M, Iochpe C.: Activity patterns in process-aware information systems: basic concepts and empirical evidence. *Journal of Business Process Integration & Management* 2009; 4(2): 93-110.
- [20]. Nguyen TH, Vo BQ, Lumpe M, Grundy J. KBRE: a framework for knowledge-based requirements engineering. *Journal of Software Quality* 2014; 22(1): 87-119.
- [21]. Lumpe A. *Forms of Models and Induction*. Seattle, WA: Seattle Pacific University, Lecture Notes (2007)

- [22]. Jepbarova N. Survey of Instructional Strategies Lecture Notes EDU 6526, March 7 (2007)
- [23]. Yarandi M, Jahankhani H, Tawil ARH. Towards Adaptive E-Learning using Decision Support Systems. *International Journal of Emerging Technologies in Learning* 2013; 8:44-51.
- [24]. Okoye K, Tawil ARH, Naeem U, Bashroush R, Lamine E. A Semantic Rule-Based Approach Supported by Process Mining for Personalised Adaptive Learning. In: *Procedia Computer Science* 2014: 37;203-210.
- [25]. Knowledge Interchange Format. (Online) Available at: <http://cl.tamu.edu/discuss/kif-100101.pdf> Accessed [April 2018]
- [26]. Protégé Editor. (Online) Available at: <http://protege.cim3.net> Accessed [March 2018]
- [27]. Han J, Kamber M, Pei J. *Data Mining: Concepts and Techniques*. The Morgan Kaufmann Series in Data Management Systems, [3rd Ed] Morgan Kaufmann Publishers (2011).