The Ontological Group Profiling for Virtual Communities of Practice Framework

Zeti Darleena Eri, Rusli Abdullah, Marzanah A. Jabar, and Masrah Azrifah A. Murad

Abstract—This paper presents a study that seeks to investigate the current state of Virtual Communities of Practice (VCoP). It aims to provide a strategy to manage research communities’ group profiling using ontology in the field of Knowledge Management (KM). Within the research communities of academics researchers, they consist of groups of individuals that have their own competencies and experiences. Managing these groups of people is not an easy task since most of the profiles are kept both static and individually. Many software development projects ignore the fact that people can be in the form of groups that share practices, such as CoP, with participation and interaction playing an active role in generating group profiles. The objective of the paper is to define and validate the ontoGP-VCoP framework. Finally, the framework that the study proposes will be used as both a guidance and example of a potential similar development in any institution of higher learning in Malaysia.

Index Terms—Group profiling, VCoP, ontology, KM strategy, research communities.

I. INTRODUCTION

Knowledge Management (KM) has been applied in public of higher learning institutions (IHLs) in Malaysia over the past years to provide strategies, processes and technologies to manage organizational knowledge [1]. A strategy that has become widely-used in many KM practices is the Communities of Practice (CoP) [2], [3], a practice-based theory of learning perspective known as “legitimate peripheral participation” (LPP). LPP is a concept which describes how newcomers become members in a group of “old timers” (experienced members) in a CoP or collaborative project by participating; subsequently newcomers will become old timers [4].

Previously, [5] focused on CoP as a knowledge perspective on organizational studies that provides a view of learning as being connection between working and innovating. Later works by Wenger has dropped the term of LPP and put more focus on online communities together with [6].

According to [7], CoP consists of groups of people who share a concern or a passion for something they do and continually aspire to better performance. They interact regularly and [8] identified CoP as being “groups within or maybe across organizations who share a common set of information needs or problem”.

Due to the phenomena of the internet and globalization respectively, the numbers of knowledge workers have rapidly increased, as well as somehow changing the conventional CoP into Virtual CoP (VCoP). The meaning of a VCoP according to [9] refers to a CoP having fully internet-based support interactions.

The impact of a VCoP of KM in IHLs is considerable, particularly concerning these two items, namely knowledge and learning. Knowledge is known to have a key competitive advantage, as well as being the core commodity. Learning from the perspective of VCoP will take into account the competencies of the members in the organization; hence, the creation of the VCoP is necessary in developing these competencies. From another perspective, the importance of VCoPs has been highly recognized in this current decade, particularly with regard to the strategic application of KM and Informal Learning [10]. With a combination of proper internet infrastructure, social networking and the capitalization of know-how respectively, the concept of CoP has been extended to include the concept of virtual CoP[9], [11]-[13]. Therefore, this study finds that the VCoP is a suitable strategy in KM by which to apply at many IHLs.

In KM, finding the right people to collaborate and the right knowledge domain is crucial [14] particularly in the terms of supporting new comer and managing current members of the communities. Many software developments ignores the sociability factors such as participation and interaction particularly in generation groups profile [15], [16]. In an effort to better understand and describe VCoP, the ontology will be used to represent the shared conceptual model and provide a common understanding between communities.

II. PROBLEM AND OBJECTIVES

In current the situation, information concerning groups of researchers (i.e. research group profile, research grants, publications, members’ profiles) is kept in a static form and individual profiles are regularly listed as a group’s members collectively are displayed most websites.

The research related information in a member’s profile is used to present an overview of general or specific research interest and is used for browsing. In order to strengthen knowledge sharing between groups and to manage these groups, it is important to establish a standardized and formal representation of interests and knowledge of the research communities [17].

Managing groups of people is still a potential problem and it has been put into focus by the following researchers [13], [18]-[21]. The newcomer or current member will be able to navigate the websites with relevant pages so as to
join the research groups. However, a problem exists in that the newcomers and current members of the communities will browse by looking at individual profiles, which can be simplified by understanding the VCoP of research communities in order to generate group profiling.

Knowledge is the asset and commodity traded within IHLs that creates a supply and demand scenario. Thus, finding the right knowledge and people at the right time to collaborate is also a critical aspect of KM which many IHLs should address [22]. The objectives of this study are to define and validate the ontoGP-VCoP framework for IHLs in Malaysia.

III. LITERATURE REVIEW

The literature review focuses on these components of research area, namely: VCoP, VCoP Model, ontology, group profiling and its related work in the KM field.

A. Virtual Communities of Practice and CoP

The setup of CoP is informal and the participation of its members operates by sharing common interests with three main characteristics, namely: domain, community and practice. Domain defines the setting or platform of those who are involved and their competency that is valued by other members. Community is defined as the social structure that creates the relationship of shared information and engaged activities in which members can learn from each other. Practice is the shared knowledge collectives that address the problem or task at hand.

VCoP was initiated from the concept of CoP, whose members utilize various internet technologies to support communication and participation. A CoP is called ‘virtual’ when CoP members are supported by internet technologies as part of their communication [23], as well as being supported by face to face meetings to enrich interaction. Further, one researcher defines VCoP as the solution to solve the dispersion of the members of the CoP by the following elements, specifically: geographic distribution, time and space that require internet technologies to communicate [24].

B. VCoP Model

The VCoP Model was constructed by [9] based on Wenger’s CoP attributes of mutual engagement, shared repertoire, joint enterprise, community and learning or identity acquisition. The concept of VCoP was only introduced relatively recently by several researchers including [9], [23], [25], [26].

However, only Murillo produced the VCoP model based on Wenger’s CoP constructs. Murillo’s work was focused on a particular area of social media named Usenet, which is a discussion network and also the largest discussion area of the internet. His work focuses on identifying a search strategy by which to detect online collectives from a high CoP attribute of online sharing and participation. He has drawn out two traits in his model, and these are known as Essential Traits and Exemplary Traits. The Essential Traits were taken based on Wenger’s CoP attributes as mentioned previously. The Exemplary Traits are attributes aiming to enable online searches to be more visible and measurable, as described in Table I.

VCoP can be in the form of several modes, specifically: 1) distributed mode [9] and 2) co-located mode [23]. The VCoP of distributed mode is separated by geographical factor with no physical existence beyond the support of ICT [23], [25]. Within a real situation, groups of VCoP are also dealing with practices using ICT. Having a mix of with and without face-to-face meetings can result in a lack of richness in interaction of the VCoP [23]. We use Murillo’s Murillo VCoP model to capture, identify and model research communities in our study.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Attributes Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essential</td>
<td>Mutual Engagement</td>
</tr>
<tr>
<td></td>
<td>(Practice) Members’ interactions to collaboratively solve problems, produce artifacts and discuss relevant issues</td>
</tr>
<tr>
<td>Exemplary</td>
<td>Core-periphery interaction structure</td>
</tr>
<tr>
<td></td>
<td>High volume of online interactions</td>
</tr>
<tr>
<td></td>
<td>Active interactions between members</td>
</tr>
<tr>
<td>Essential</td>
<td>Joint Enterprise</td>
</tr>
<tr>
<td></td>
<td>Provide an identity and define key issues or problem that members need to address</td>
</tr>
<tr>
<td>Exemplary</td>
<td>Focused on a profession</td>
</tr>
<tr>
<td></td>
<td>Highly-focused discussions</td>
</tr>
<tr>
<td></td>
<td>Quality of interactions with identifiable domain of knowledge</td>
</tr>
<tr>
<td>Essential</td>
<td>Shared Repertoire</td>
</tr>
<tr>
<td></td>
<td>Knowledge resources such as tools, procedures, symbols or artifacts that the community has produced</td>
</tr>
<tr>
<td>Exemplary</td>
<td>High-quality institutional documents</td>
</tr>
<tr>
<td></td>
<td>Useful resources produced by the members and to avoid repetition of routine question or practice</td>
</tr>
<tr>
<td>Essential</td>
<td>Community</td>
</tr>
<tr>
<td></td>
<td>A stable group with social interrelationships developed through sustained mutual engagement</td>
</tr>
<tr>
<td>Exemplary</td>
<td>Low level of conflict</td>
</tr>
<tr>
<td></td>
<td>Manifested by subject headers and subject tone on issues discussed to avoid online misinterpretation or “flame war”</td>
</tr>
<tr>
<td></td>
<td>Mid-sized online group</td>
</tr>
<tr>
<td></td>
<td>Each member knows and is able to engage with each other through online or face-to-face interactions</td>
</tr>
<tr>
<td>Essential</td>
<td>Learning/Identity Acquisition</td>
</tr>
<tr>
<td></td>
<td>Members’ participation resulted other members’ learning. Learning is treated as equal to acquisition to a new identity; becoming an insider</td>
</tr>
</tbody>
</table>

C. Ontology

Ontology provides a simplified and explicit specification on what we want to represent [27]. Ontologies are useful because they define and model each component or process systems as much detail as possible. Ontology can define knowledge in any formal definition using a combination of taxonomy, axiom, hierarchy tree or diagrams.

Ontology-based modeling is useful to capture and classify knowledge model of VC profiles and is considered to be a comprehensive knowledge model.

D. Group Profiling

Group is defined as a set of actors who interact with others within a community rather than with than those who are not. Similarly, group profiling refers to the process of extracting descriptive attributes for a group of people [28].
The context of profiling is divided into data subject, subject and the data controller respectively [29]. The data subject is the main point to which a profile refers (i.e. human, non-human, individual or group), the subject refers to the person or an organization and the data controller refers to the data user. Data user is defined as the subject that determines the purpose of the processing of the data and the application.

Profiling is the process of building data profiles. The profiles contain a list of related data that is used to represent a person, group or category. Personalized profiling refers to the process of building a list of related data that represents a particular individual. According to [29], [30], group profiling denotes the process of extracting descriptive attributes for a group of people that form a community. These attributes are then collected, aggregated, stored and processed in order to find shared features. Group profiles can be in the form of a distributional profile (i.e. all members shared all attributes) or a non-distributional profile (i.e. not all members shared all attributes).

E. Ontological Group Profiling VCoP

The most important role of ontology in the VCoP is to provide a shared understanding in order to support knowledge-sharing. Ontology provides the standard vocabularies by putting out concepts and relationships that come together with its definitions [31].

In managing groups of a VCoP in KM field using ontology, several related works are compared and summarized as shown in Table II. These related works include: the Palette project [32], KEPSNet project [13], framework of ontology-based KMS [33], framework for ontology-based KM [34] and the CIK-NET project [35].

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual Identification</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Group Identification</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Individual Classification</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Group Classification</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>VCoP</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Yes</td>
</tr>
<tr>
<td>Group profiling</td>
<td>Yes</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ontology</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Collaboration</td>
<td>Yes</td>
<td>No</td>
<td>NA</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Application</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Functionality/KM process</td>
<td>Yes</td>
<td>NA</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

IV. PROPOSED FRAMEWORK

Based on the literature analysis set out in Section III, we propose the use of the onto GP-VCoP framework as shown in Fig. 1. In reference to Table II, not many studies focus on describing VCoP, group profiling and group classification as part of the KM service process. Thus, it is crucial to identify and specify people as being a community comprising one important component in KM service strategy. Hence, we are keen to look into this area in our study by incorporating VCoP and group profiling into KM.

Basically, the main components of the framework consist of people, process and knowledge infrastructure. The people’s component contains the VCoP (i.e. research communities) that uses group profiling to provide its functionalities. The process components then contain the KM process activities. The third component is the knowledge infrastructure component that binds all of the components together.

V. METHODOLOGY

We construct the onto GP-VCoP framework with these following steps:

A. Step 1: Performing a Review of the Existing Literature

The process of reviewing the VCoP, VCoP Model, ontology and KM is conducted at this step. The previous and related works of other researchers are also included at this stage. The outcome of this step is the summary of the literature review knowledge gap as shown in Table II.

B. Step 2: Formulate the Proposed Framework Components

The next process continues with the construction of the proposed framework. This process involves identifying and selecting elements from the previous step. The activity then continues with the questionnaire construction for piloting.
C. Step 3: Conduct the Preliminary Survey

A preliminary survey is conducted in step 3. The survey questionnaire design is based on the identified components and elements detailed in step 2. The main reason for conducting the preliminary is to validate the components of the proposed framework.

D. Step 4: Preliminary Survey Analysis

The analysis process uses the Rasch Analysis Model with reliability and validity tests, data fitness tests as well as a the distribution on person-item with Wright map using Winstep.

E. Step 5: Translate the Form of the Framework into Software Architecture and Functionalities

This step translates the framework into a more workable state of software architecture and its functionalities.

VI. FINDINGS AND DISCUSSIONS

The findings in this section are divided into two parts consisting of a preliminary survey and the proposed framework suggesting software architecture and functionalities.

A. Preliminary Survey

A preliminary survey was conducted to validate the proposed framework. The survey is distributed using self-administered questionnaires. The target respondents are the academicians of Computer Science and IT-based faculties of public institutes of higher learning in the Klang Valley area. The returned samples are 42.5% only. The data are analyzed using the Winstep tool for Rasch Analysis Model.

Fig. 2 presents a moderately high Cronbach Alpha reliability coefficient of .89. This indicates that the design of the questionnaire is good. Both the person and item reliability index are moderately high with values of .87 and .86 respectively. These are considered as a good index for both person and item. Hence, the data can be used for further analysis.

Next, the analysis of infit and outfit tests with standards outfit MNSQ is 0.5<x<1.5 and outfit ZSTD is -2.0<x<2.0. The mean infit and outfit for person and item MNSQ are expected to be 1.00. In regard to this data, they are all close to 1.00. The mean infit and outfit ZSTD for a person are -.5 and -.7 respectively. The mean infit and outfit ZSTD for items are .0 and .2 respectively. These statistics indicate that the data are inbound, although the data fit model could be somewhat better due to redundant items.

The separation index for a person is 2.58, comprising a moderately good spread of items and persons along a continuum. In relation to item separation index, it also shows a moderately good spread of item index of 2.50.

Fig. 3 is the Wright map that presents the person-item distribution. The majority of items are easily endorsed by most respondents. This indicates that all respondents involved in the content validity test tend to agree with each of the components and their assigned criteria of the survey. The most agreeable items include: no. 9 (group profiles help in searching for the right member to participate in a research project), no. 18 (respondent has difficulty in searching for the right member to participate in my new project), no. 22 (I have difficulty in finding new members) and no. 8 (the research group profile from the faculty’s website helps in searching for the right member to participate in a research project).

B. Software Architecture and Functionalities

The ontoGP-VCoP framework with Murillo’s VCoP model is then translated into software architecture and its functionalities as an application are shown in Fig.4. Any member of the VCoP can deposit or store (1) knowledge in her data profile. She can also access and share (1) the ontoGP-VCoP application in order to request knowledge from the repositories. The ontoGP-VCoP will then display...
the output together with its description (2).

The knowledge and ontologies representation is develop using Protégé 4.0.2 and Table III shows the outline of how the ontologies will be displayed in the ontoGP-VCoP interface. The ontoGP-VCoP application will then be able to generate the ontologies and their description when a request is deposited to the application. The levels indicate the depth of the application in order to generate a request.

For example, let us consider scenario one: a member asks who the experts in the KM field are? He then deposits the question to ontoGP-VCoP with three categories selected as part of the request. He can be more precise as to which exact branch of the category he is interested in. As an output, the interface of the application will show the knowledge representation together with its descriptions.

As for scenario two: a new member submits a request to join a research group that has similar interest to her own field that he is interested. As an output, the ontoGP-VCoP will display the knowledge infrastructure. The ontoGP-VCoP framework will complement the KM area as part of a strategy offering a holistic platform of people, process and knowledge infrastructure.

As for future work, the development of ontologies for ontoGP-VCoP will take place. The ontology will also need to be validated. The application of ontoGP-VCoP is suggested so as to be an easy access tool for use by the research communities.

VII. CONCLUSION AND FUTURE WORK

In this paper, we introduced the ontoGP-VCoP framework and the process by which to validate the framework. The framework will contribute to providing the ontology on how to build a VCoP with KM process, in addition to an application to generate VCoP and group profiling for research communities. The ontoGP-VCoP framework will complement the KM area as part of a strategy offering a holistic platform of people, process and knowledge infrastructure.

As for future work, the development of ontologies for ontoGP-VCoP will take place. The ontology will also need to be validated. The application of ontoGP-VCoP is suggested so as to be an easy access tool for use by the research communities.

### REFERENCES


[28] R. Bin Haji Abdullah is an associate PROFESSOR in the Information System Department, Faculty of Computer Science and Information Technology of Universiti Putra Malaysia. He holds a B.Sc in Computer Science from Universiti Putra Malaysia (1988), M.Sc in Computer Science from Universiti Putra Malaysia (1996), and PhD in Knowledge Management field at Faculty of Computer Science and Information System at UniversitiTeknologi Malaysia (2005).

He has more than 12 years of teaching experience and as well as 8 years of system development experience as a system analyst at a higher learning institution. His research interests include: Knowledge Management and Software Engineering, Computer Supported Collaborative of Work, and Workflow Management.


Zeti Darleena Eri is a PhD student in the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia. She holds a B.Sc in Computer Science, (1999) and M.Sc. Computer Science (2002) from UniversitiTeknologi Malaysia. Currently, she is a lecturer at UniversitiTeknologi MARA (UiTM) Shah Alam, Malaysia. She has been an IACSIT member since 2009.

Her study involves developing aontoGP-VCoP framework that focuses on ontological group profiling for research communities in Malaysian institutes of higher learning. Research areas studied include: knowledge management, virtual communities of practice, group profiling and ontology.

[36] Marzanah A. Jabar is an associate professor at the Information System Department, Faculty of Computer Science and Information Technology of Universiti Putra Malaysia (UPM) Malaysia. She obtained her PhD in 2008 with the area of studies are Management Information Systems and Knowledge Management from UPM. She has vast experiences in developing and maintaining information systems when she works as Senior System Analysts for UPM before she embark her career as an academician. Currently, her area of research interest are software engineering and IS including Data warehouse, Decision Support System, ERP and knowledge sharing.

Masrah Azrifah Azmi Murad is an associate professor at the Department of Information System, University Putra Malaysia. She received her Bachelor of Management Information System (1997) from Drexel University, Philadelphia, USA; her Master of Computer Science (1999) from University Kebangsaan Malaysia and her PhD (2005) from University of Bristol, UK.

She worked as a tutor in University Putra Malaysia from 2000 to 2005 and became a lecturer in 2005 until present. Her areas of specialization are text mining, information retrieval and artificial intelligence. She is a committee member of IADIS International E-Commerce Conference (2007), Second International Conference on Informatics (2007) and Malaysian Software Engineering Conference (2007).