

---

## Editorial

---

### Lorna Uden

FCET,  
Staffordshire University,  
The Octagon, Beaconside,  
Stafford, ST18 0AD, UK  
Email: L.uden@staffs.ac.uk

**Biographical notes:** Lorna Uden is Emeritus Professor of IT Systems in the Faculty of Computing, Engineering and Technology at Staffordshire University. Her research interests include technology learning, HCI, big data, mobile learning, activity theory, knowledge management, web engineering, multimedia, e-business, service science and innovation, semantic web, internet of things, software as a service (SaaS), internet of things and problem-based learning.

---

Welcome to issue V9N4 of *IJWET*. This issue consists of three papers. The first paper is 'EXCLS: enhanced XML clustering by level structure accuracy' by Rehab Desoki and Ahmed Elfatraty. In this paper, they introduce a framework for clustering XML documents by structure. The framework consists of two phases. In the first phase, each XML document is represented by an enhanced level structure which includes the hierarchical relationships between nodes without summarisation. The proposed representation is based on hash table data structure to improve processing time. Based on the proposed representation, they presented a distance metric to evaluate the distance between two documents. Then, a hierarchical clustering algorithm has been applied to a sample dataset to evaluate the threshold. In the second phase, the XCLS incremental clustering algorithm has been applied to all datasets with a predefined threshold that resulted from the first phase.

The proposed method was implemented and evaluated using homogeneous and heterogeneous datasets. The authors argue that the experimental results show a significant improvement in clustering accuracy, especially in homogeneous XML documents without a significant impact on processing time. However, one of problems is the need to specify the clustering threshold in the XCLS incremental clustering algorithm. Another problem is the specification of the number of clusters by users in k-mean partition-based clustering used by XEdge which affects clustering results. In addition to enhancing the performance of clustering without summarising the structure, hash tables have been used to reduce the processing time. Further work will have to be carried out to address these problems.

The second paper is 'Towards a model-driven approach for treating quality attributes of service-based business processes' by Fabrício Teles, Fernando Lins and Nelson Rosa. The authors in this paper we present the 'model-driven quality attributes for business processes' (MOQUA) solution, which aims to provide an agnostic and systematic way for treating with quality attributes of business processes through the BPM lifecycle. MOQUA proposes a forward model-driven engineering process in which abstract

business processes models and their quality attributes are sequentially refined through model transformations until the production of executable and quality-aware business processes.

To illustrate the proposed approach, a supply chain management use case is used. In order to evaluate the model-driven approach, an evaluation plan based on goal/question/metric (GQM) is defined, providing the rationale to define and to apply evaluation metrics of interest. The MDE techniques that underlie MOQUA were evaluated against the proposals of the main related work, providing a comparative view of our approach and the ones defined in related work. The GQM measurement plan defined to evaluate the related approaches can be reused and extended to perform other goal-driven evaluations of MDE solutions for quality attributes. Although the authors argue that there are benefits of using the model, it is not possible to see the evidence in use. To validate the model, it is important that this is put to actual use.

The final paper is 'Latent semantic analysis for business protocol discovery using log files' by Abdelkader Moudjari, Salim Chikhi and Hamamache Kheddouci. In this paper, the authors present a new business protocol discovery algorithm based on latent semantic analysis (LSA) using log files. The proposed approach exploits the information present in a set of log files instead of only one. An incremental approach based on the notions of micro and macro-succession between messages has been adopted. The proposed approach uses synthetically generated log files, with the presence of noise and incomplete information.

To validate the usefulness of the proposed approach different configurations have been tested. These configurations contain critical cases, such as cycles, multiple final states and the belongingness of the initial states to cycles. The input test business protocols have been successfully discovered with only the use of the information present in the log files. The use of the LSA permitted overcoming the problem of noise presence and incomplete information. It also supports the case of a large number of messages. These authors argue that the results obtained from applying the proposed approach have proved that it is effective for business protocol discovery.