Editorial

Lorna Uden

Faculty of Computing, Engineering and Technology,
Staffordshire University,
The Octagon, Beaconside,
Stafford, ST18 0AD, UK
E-mail: L.uden@staffs.ac.uk

Biographical notes: Lorna Uden is an Emeritus Professor of IT Systems in the Faculty of Computing, Engineering and Technology at Staffordshire University. Her research interests include technology learning, HCI, activity theory, knowledge management, web engineering, multimedia, e-business, service science and innovation, semantic web, and problem-based learning.

Welcome to V6 N3 issue of *IJWET*. There are four papers in this issue covering from visualisation of geo-tagged pictures in the web to dimensionality reduction for blog tag mining.

The first paper is by Carboni, Marotto and Zanarini. This paper is about 'Visualisation of geo-tagged pictures on the web'. These authors argue that the release of Google Maps API has fostered a large number of geographical mashups based on interactive maps and user generated content. Google competitors have accepted the challenge and brought geographical information systems to the masses. This paper summarises a study of three different user interfaces for browsing geo-referenced pictures on web sites. The applications subjects of this study are two widely used systems – Panoramio and Flickr. The result of this analysis is a new application developed by the authors called the 'fractal view'. The study has been conducted by means of users' interviews for a qualitative evaluation and by means of eye tracking for a quantitative assessment. Users are required to browse pictures and to select ten images that they consider the most interesting in a sort of collection of favourites. The authors chose this type of task because the kind of applications analysed are not professional GIS. People often browse pictures on the web for fun and to search points of interest without a precise destination, even discovering new things unexpectedly. It would be good to have more analysis on this.

The second paper is by Gonçalves, de Souza Gimenes, Fantinato, Travassos, and de Toledo, 'Experimental studies of the e-contract establishment in the PL4BPM context.' These authors argue that electronic contract (e-contract) establishment is a complex activity due to both sundry involved information and required parameters. Product line for business process management (PL4BPM) was developed to reduce the complexity of e-contracts based on web services (WS-contract). It promotes business process reuse based on software product line (PL) concepts. Their paper presents results of two experimental studies, aimed at investigating benefits, usability and feasibility of PL4BPM, concerning its WS-contract establishment process. Participants from both academia and industry were involved. Two studies were performed, one with academic

participants in the scope of *in vitro* studies, and the other as a survey with participants who work in software development companies.

The results of the studies show strong evidence that the PL4BPM approach has advantages over *ad hoc* approaches. These advantages include: mechanisms that can save time and effort that would be spent learning a specification language syntax, in case of the establishment of WS-contracts by *ad hoc* procedures. Concerning the usability assessment, PL4BPM was considered easy to use, due to its well defined stages. In addition, the use of feature modelling helps the roles involved to understand the e-contract establishment process. Overall, PL4BPM was considered a feasible approach. The results and knowledge gained in the experimental process serve as a basis for proposing a set of activities to support experimental studies in the BPM domain. Although a significant number of data and aspects have been taken into account in these experimental studies, further studies and analyses, from different points of view and using additional criteria, would be useful.

The third paper is entitled, 'A classification scheme for open government data: towards linking decentralised data', by Kalampokis, Tambouris and Tarabanis. In their paper, these authors proposed a classification scheme for open government data (OGD). OGD refers to making public sector information freely available in open formats and ways that enable public access and facilitate exploitation. The proposed classification scheme includes two dimensions. The first dimension cares for the technological aspect of OGD initiatives, which is an important driver. This includes two main categories:

a   making data available on the web as downloadable files in well-known formats such as PDF, Excel, CSV, KML, XML, JSON, etc.

b   making data available on the web as linked data through RESTful APIs and/or SPARQL search interfaces.

The second dimension refers to the organisational approach followed for providing governmental data. Again, this includes two categories:

a   direct data provision, where data belonging to various public agencies is published by the one-stop government data portal

b   indirect data provision, where data belonging to various public agencies is published in a decentralised manner by these agencies (usually on their website) while the portal provides some kind of linking mechanism and/or metadata for the identification of the actual dataset.

Based on the proposed classification scheme, 24 official OGD initiatives were reviewed. These authors also present an architecture and prototype implementation for the most advanced OGD class in their scheme, which enables linking decentralised data. This exercise showed the approach is technically possible albeit immature (e.g., it requires significant manual handling). Further work is required to investigate the potential and limitations of indirect provision of linked data in the public sector from an organisational and technological perspective.

The last paper is by Tsai, 'Dimensionality reduction for blog tag mining'. In this paper, Tsai argues that when analysing large data of multiple tags such as blogs, it may be necessary to perform dimensionality reduction or projection techniques to transform the data into a smaller, more manageable set. Dimensionality reduction is the search for a small set of features to describe a large set of observed dimensions. By performing

dimensionality reduction, the author hopes to uncover hidden structure that aids in the understanding of the data. In this paper, the author utilises dimensionality reduction techniques to reduce the inherent noise in blog tags. A tag-topic model is combined with dimensionality reduction, and then evaluated on real-world blog data. By employing dimensionality reduction techniques to reduce the document-tag space, better classification results were achieved. According to the author, this indicates that the noise in tags can be effectively reduced by representing the original set of tags with a smaller number of latent tags, which can lead to more accurate real-time categorisation of blog documents. The results indicate that the classification results of the tag-topic model combined with autoencoders achieved higher precision, recall, and F-score than the baseline tag-topic model. This shows that the noise in tags can be effectively reduced by representing the original set of tags with a smaller number of latent tags, thus improving the classification of blog documents. It would be useful to apply other dimensionality reduction techniques to a larger set of blog tags, and evaluate the feasibility of real-time deployment of the techniques for automatic classification of blog categories.