

Use of Semantics to Build an Academic Writing Community Environment

Sirisha BAJANKI^a, Kathrin KAUFHOLD^b, Alex LE BEK^c, Vania DIMITROVA^a,
Lydia LAU^a, Rebecca O'ROURKE^b, Aisha WALKER^b

^a*School of Computing, University of Leeds, UK*

^b*School of Education, University of Leeds, UK*

^c*Web Design: Simple solutions from advanced technology, Leeds, UK*

Abstract. Writing a dissertation is a critical aspect of the learning experience for most university students in all disciplines. It is often accompanied by anxiety and uncertainty, which supervisors often struggle to understand and address. A community-driven approach to support dissertation writing based on semantic social scaffolding is presented here. The paper describes how a semantic wiki was tailored to develop a social writing environment to provide holistic support throughout the whole dissertation process. Based on initial evaluation studies, we discuss the benefits and pitfalls of semantics. The work contributes to a recent research strand that examines how to exploit new social computing technologies to develop effective learning environments.

Keywords. Community learning environments, academic writing, semantic markup, semantic wikis

1. Introduction

Writing a dissertation is a major challenge faced by most students across the higher education spectrum. This is seen as a highly individual process, a 'lonely journey', which often embeds many of the anxieties associated with academic writing [1,2]. A pedagogical solution known to be effective is *social scaffolding*. It is well-established that inexperienced writers who can support each other and work in collaboration with explicit models and processes show characteristics of more mature writers [3]. Social models and community experiences are at the heart of the new wave of social computing technologies which is the natural platform to adopt when developing technological solutions for social scaffolding.

However, the educational community is sceptical; pointing out that social computing provides busy spaces but not necessarily effective learning spaces [4], e.g. too much time may be spent for information search not for learning, the knowledge is poorly structured, tacit knowledge is buried in large textual resources where key nuggets are difficult to find and articulate. We, however, believe that this pedagogical issue presents new opportunities for AIED research to embark on analysing challenges and exploiting possibilities brought by social computing. Initial architectures and approaches for collective content creation, community annotation, and knowledge sharing [5,6,7,8,9,10] provide encouraging support for a new AIED research strand that exploits intelligent technologies to develop social, community-driven learning

environments. This strand is still in an embryonic stage – prospective domains are not fully explored, examples of embedding the new environments in real practices are lacking, the vastly expanding palette of social technologies and architectures is yet to be leveraged in AIED [5,10].

The paper makes a unique contribution to AIED research in social computing from two perspectives. First, we argue that academic writing in general, and the dissertation in particular, is a promising domain for social AIED architectures due to its significance and the existence of underpinning social pedagogical models. Although intelligent technological solutions for writing development have been built, they focus mainly on discrete aspects of the dissertation process, for example argumentation or research methods [11,12]. In contrast, we show that social computing can be exploited to provide the necessary holistic support throughout the whole dissertation process.

The second basis of our contribution is based on the intelligent techniques we exploit which follow recent advances in social semantic web [13]. Specifically, the paper examines *how semantics can be used to shape a community learning environment and what benefits and obstacles this may bring*.

The paper will first present a semantic-based social environment, called AWESOME¹ Dissertation (Section 2), where students and tutors share dissertation practices, comment on writing examples, and become aware of the specific characteristics of the dissertation processes. Based on initial evaluation studies outlined in Section 3, we will discuss potential and challenges brought by semantics (Sections 4 and 5). Finally, we will conclude and outline our future plans (Section 6).

2. The AWESOME Dissertation Environment (ADE)²

The AWESOME Dissertation Environment (ADE) is an evolving interactive learning environment in which both academic staff and students contribute knowledge and experience to dynamically create content. It is intended for use by UK undergraduate and taught postgraduate students during their dissertation writing process. The platform is designed to augment academic supervision with peer support, and to actively engage students with relevant content appropriate for the stage of development in a student's dissertation writing process. ADE is based on MediaWiki³ and uses its extension Semantic Media Wiki [14] to create and query semantic content. As shown in Fig. 1, ADE has three layers.

The **interaction layer** implements the user interface. A user can *contribute content* in the form of wiki pages. In addition, ADE provides semantic-enabled functionality for structured *community browsing* (Section 4), *personal bookmarking* to indicate community content that is relevant to him/her (e.g. a user can bookmark an interesting

¹ AWESOME (Academic Writing Empowered by Social Online Mediated Environments) is an interdisciplinary project funded by the UK JISC programme. The partners are the departments of Education, Computing, and Staff Development at the University of Leeds, the Centre for Academic Writing at Coventry University, and the Department of Education and Lifelong Learning at Bangor University.

² ADE can be accessed from <http://awesome.leeds.ac.uk>. Only the public instance enables browsing of the community content, public access to the other instances is disabled to comply with data protection regulations.

³ <http://www.mediawiki.org>

topic idea or a question related to their dissertation), and a *personal dissertation space* where a user accesses his/her collection of bookmarks.

The **content and semantics management layer**

makes extensive use of semantic features such as *semantic forms* (structured forms for adding and editing semantic markup, these are used as the main means for producing semantically tagged content) and *factboxes* (each shows a summary of the semantic markup entered in a wiki page). *Semantic queries* are used to mine the semantic markup to enable ‘mashing’ of content related to the main dissertation categories.

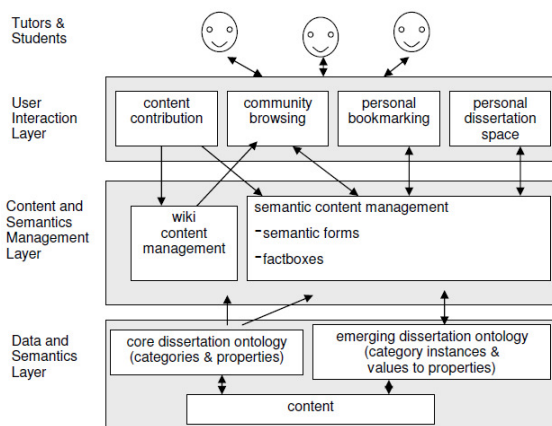


Figure 1. The ADE architecture

The **data and semantics layer** maintains the *content* and its links to a *core* and an *emerging dissertation ontology*. The core ontology includes pre-defined categories and properties, while the emerging ontology is being built during content creation.

Categories capture the main aspects of the dissertation process, such as getting an overview of dissertation `[[category::dissertation]]`, choosing a topic `[[category::topic]]`, adopting an appropriate research methodology `[[category::methodology]]`, building the body of literature `[[category::literature]]`, dissertation writing `[[category::writeup]]`, and management of the dissertation project `[[category::project management]]`. In addition, each category can have subcategories, e.g. `[[category::dissertationFAQ]]` is a sub-category of `[[category::dissertation]]`. When content is contributed in ADE, it is linked to a category either automatically, if semantic forms are used, or by a user adding simple markup. In this way, an instance of the category, in the form of a wiki page with content, is added to the emerging dissertation ontology.

Properties are the basic way of augmenting a page with semantic data. They express binary relationships between one semantic entity (a wiki page or a category) and another data entity or data value. In ADE, properties are mostly predefined as part of the core ontology⁴. ADE exploits properties to enable semantic social scaffolding. There are three types of properties: (1) characteristics of the main categories, e.g. *has research question* and *is inappropriate* are properties associated with the category *topic*; (2) general social scaffolding properties, e.g. *is top tip*, *is good writing*, *has feeling*; and (3) additional information about the contributions made, e.g. *has user*, *has question*, *has answer*.

The core dissertation ontology includes the predefined properties. Using semantic forms or direct semantic annotation, users can add semantic markup in the form of:

`[[property::value]]` which adds a triple `<content, property, value>`

⁴ Note that the users can add new properties too, which ADE does not disallow but we have not seen this happening in the current usage of the environment, see Section 4.

to the emerging ontology (content is the wiki page and value usually corresponds to some important part or characteristics related to the content in that page). For example, the sample markup below is taken from the public ADE instance:

```
[[is toptip::write your introduction at the end]]
[[is good writing::synthesise ideas and relate them to the topic]]
[[is bad writing::images should not be added just for decoration]]
```

The main ADE categories and properties have been derived via focus groups with academic writing experts within and outside the AWESOME project. The core ontology in the public instance of ADE includes 26 categories and 64 properties. In the remaining part of the paper, we will discuss benefits and challenges of using semantics for shaping a dissertation community by showing examples from current ADE instances and following feedback from user trials.

3. ADE Instances and User Trials

ADE instances have been created for user trials which involve different user groups and usage contexts, as indicated in Table 1.

Table 1. Summary of the user trials with ADE. AWESOME continues to be used after the user trials.

Institution	Context	Tutors	Students	Usage	Duration
School of Education, Leeds University	Master distance learning course in ICT in education	[1 tutor] Course coordinator, dissertation supervisor	[5 students] Mature, part-time, both home and international	Pilot study End of the dissertation period (writing up), intensive use, active tutor involvement	1 week in July 2008
School of Fashion Design, Leeds University	Undergraduate course with a very practical focus and little formal writing	[3 tutors] Course coordinators, dissertation supervisors	[10 students] Young undergraduates	Students had just returned from a year on work experience. Preparation for dissertation, additional support during the dissertation process	July 2008 – April 2009 Ongoing
		[1 tutor] Academic skills tutor	[19 students] Young undergraduates	Students in their second year preparing for their dissertation to follow in their final year	1 day session, October 2008
Centre for Academic Writing at Coventry University	Volunteers from the students who visit the Centre	[2 tutors] Academic writing tutors	[2X10 students] Diverse years, courses, and background	An online session facilitated by CAW tutors; feedback gathered immediately	2x 1 day sessions, Nov 2008 and January 2009
College of Education and Lifelong Learning	Mature part time students and postgraduate students on PG certificate teaching course	[2 tutors]	[12 students] Mature adult learners, limited IT experience	Supplementary support environment for face to face teaching, new dissertation regulations	Dec 2008 Ongoing
UK Higher Education Centre in Philosophy and Religious Studies	Tutors from UK universities – Leeds, Manchester Metropolitan, Sheffield, York	[6 tutors]	[17 students] undergraduate and master level	Developing a community repository for sharing dissertation practices	Nov 2008 – April 2009 Ongoing

ADE was used as an addition to the traditional face-to-face dissertation help provided. No specific instructions on how to use the ADE were given after an initial orientation to the environment and with one of the later trials even this was dispensed with. We wanted to replicate the exploratory nature of social software. Through the trials, we collected data comprising: ADE content, ontologies, and usage history; student questionnaires, individual student interviews; a focus group with students; and a focus group with tutors. A report on the evaluation trials is available from the project web site. In the following sections we will discuss only the main issues identified regarding benefits and obstacles brought by semantics.

4. Benefits of Using Semantics

Based on the AWESOME experiences in shaping and using ADE, as well as feedback gathered from users, we can identify several benefits of using semantics.

Capturing generality and discipline specificity of dissertation practices. The core and emerging dissertation ontologies in ADE make the dissertation practices explicit, which allows identification of common and discipline-specific aspects. The core ontologies in all instances comprise the same dissertation categories, with the exception of Philosophy and Religious Studies where an additional category for supervision is added because establishing a proper partnership between a supervisor and a student was identified as crucial. This will feed back into the core ontology, enabling other disciplines to consider the role and nature of supervision. Differences in the properties used across instances have been noted, e.g. discrepancies in different research methods between Education and Fashion Design and unique characteristics of argument structure in Philosophy. Clearly articulated differences come from the values linked to content via properties. The differences point at discipline-specific discourse and epistemology, whose importance has already been highlighted by researchers [15].

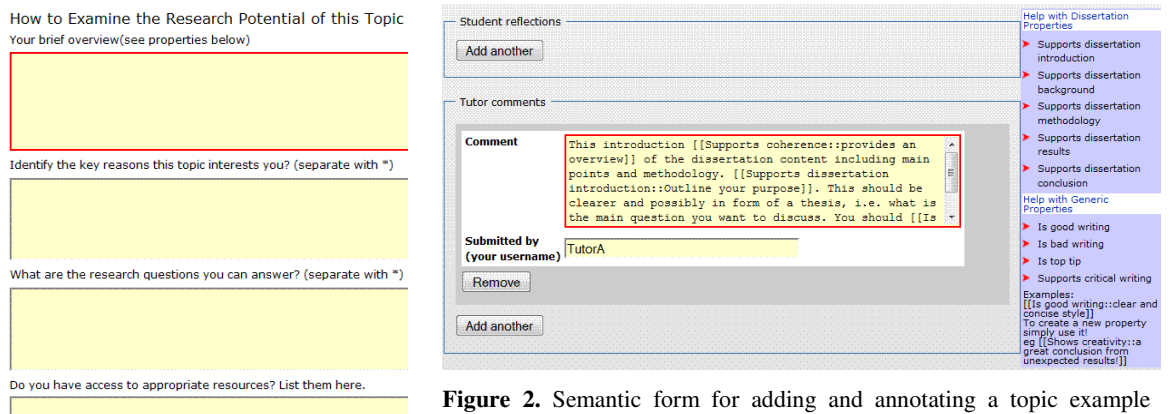


Figure 2. Semantic form for adding and annotating a topic example (left); semantic tagging of a comment added by the user (right).

Knowledge articulation through semantic tagging. The semantic markup enables connecting content to the ontology, thus identifying important knowledge chunks. In all trials, semantic forms have been used as the main means for content contribution. The forms assist the articulation of the key characteristics of dissertation

processes, as illustrated in Fig. 2, left. In a few occasions, content was contributed via direct semantic markup (Fig. 2, right) mostly by a community moderator.

Sample Introduction 1

Bookmark this...

Reference

Title	Sample Introduction
Submitted by	TutorA
Link	Ex Introduction.pdf
Podcast	

Comment by TutorA

This introduction provides an overview of the dissertation content including main points and methodology. Outline your purpose. This should be clearer and possibly in form of a thesis, i.e. what is the main question you want to discuss. You should write your introduction at the end, after you have written all the other chapters.

Facts about Sample Introduction 1

- Has title: Sample Introduction
- Has wikilink: Ex Introduction.pdf
- Is top tip: write your introduction at the end
- Submitted by: TutorA
- Supports coherence: Provides an overview
- Supports dissertation introduction: Outline your purpose

Figure 3. Example introduction with a factbox which gives a summary of the semantic properties related to the example (see the area at the bottom).

Student interviews indicate that semantic access may develop greater awareness of the dissertation processes. The support for the use of semantics to facilitate knowledge discovery is in line with recent ontology-driven approaches for browsing through educational content [16].

Improved community awareness. Students commented that they felt less isolated and relieved to see that others were experiencing similar problems. In general, this can be attributed to the openness of social environments. However, we also noted that semantic elements, such as FAQs summaries, AWESOME Press, AWESOME top tips (see Fig. 4), aimed to promote overall awareness, were extensively used and were identified by both tutors and students as positive features in ADE.

Top Tips

Practice Top Tips

Dissertation	Top Tip...
How to find a good topic	Originality isn't everything but it is a great start
Sample Introduction 1	write your introduction at the end

AWESOME Culture

AWESOME Contributor	Top Tip...	Feeling...
Catherine	Chocolate	Happy
Harriet	Weight isn't everything	Still a bit anxious. Will I ever get to the end?
Peter	BE COMFY Be fundamental, comfortable and always online	TEPID
Scsbsa	No Tip	Excited

Figure 4. Overview of recent top tips and the AWESOME culture presented on the main page. The tables are built dynamically by using semantic queries to mine semantic markup.

5. Challenges and Future Directions to Using Semantics

This section outlines challenges brought by semantics, points out how they are being addressed in the next version of ADE, and draws some wider research directions.

Adding semantic markup is difficult. The trials showed that the potential benefit of semantics for knowledge articulation could only be exploited via using semantic forms. However, users often struggled to understand the technical jargon of ‘semantic markup’ which put them off from contributing semantic content. The new version of ADE makes extensive use of semantic forms. The disadvantage of using forms is that the content is necessarily more prescriptive and part of the openness is lost. However, the advantage is that users contribute content without knowing how to semantically tag, as semantics is connected to the form fields and automatically asserted.

Semantics increases the complexity and may hinder navigation. Almost all users commented that they found it difficult, especially initially, to navigate and to understand exactly what was going on. This was partly due to having too many possibilities for discovering content and the lack of an apparent system backbone to follow. We believe that the potential of semantics should be explored further to assist the navigation through the information space. This is addressed in the next ADE version by enabling the site navigation from the ontology in a transparent way through exposing the categories of the core ontology as a tree. Hence, the ontology provides a single navigation point and a general overview of the overall environment.

Helping students exploit the benefits of social scaffolding. Many students pointed out that it was hard to follow what was there for them without going through the whole content. Although this browsing improves the overall awareness of all dissertation aspects, there are clearly parts which can be unrelated and confusing. We believe that semantics can help address this need for personalisation. Currently, ADE provides personal bookmarking enabling users to extract content chunks and store them in their personal spaces. The bookmarks can provide an individual starting point for exploring the environment based on the user’s past visits. Then, community modelling algorithms and social bookmarking approaches can be exploited to offer social recommendations based on similarities between users.

Helping tutors see what is going on in the environment. Although all tutors were generally positive about the environment, they felt that there was no single point which could show them an overview of what was going on in the environment. Currently, tutors have to use special pages which is challenging for people coming from non IT-intensive disciplines. We believe that semantics can be further exploited to support tutors. For instance, the most extensively populated parts of the ontology can be highlighted or active community members can be identified.

Showing pedagogical value of using semantics. A crucial challenge faced in ADE, and in semantic-driven social learning environments in general, is to show that there is a pedagogical value brought by the adding of a semantic layer to the traditional social settings. The discussion in Section 4 points at benefits from semantics in facilitating content creation, annotation, and sharing. A step further would be to connect these benefits to expected pedagogical outcomes. Longer studies are required to examine whether dissertation problems (e.g. students’ unfamiliarity with the dissertation as a genre and inability to effectively engage with the processes associated with dissertation writing, delay between information delivery and the time when students actually face the complexities of dissertation processes [1,2]) can be reduced when deploying social environments like ADE.

6. Conclusions

A prospective research direction which examines how to leverage the new social computing to empower learning is emerging. The work presented here contributes to this research direction by (a) showing a fruitful domain, (b) examining how semantics can be used to provide opportunities for community content creation and sharing, and (c) examining benefits and pitfalls of semantics.

The AWESOME project studies how through technological innovations the social component and academic components of dissertation writing can be captured in a unified and natural way to enhance the learning experience. Presented in this paper is the first stage of the project which focuses on the shaping of a community environment, called ADE, by tailoring and extending a semantic wiki platform. The ADE enables community-driven content creation and sharing of dissertation practices amongst students and tutors. We are currently in the second stage where parallel user trials are conducted and a volume of qualitative data is being collected. Initial findings, discussed in this paper, point at potential benefits of semantics to discover cross-disciplinary consensus and discrepancies in dissertation practices, to aid knowledge elicitation, to improve knowledge structuring, and to develop overall community awareness. The next ADE version addresses obstacles identified in user trials and lays the foundation for wider deployment of the AWESOME approach. This will enable larger user trials to systematically examine the pedagogical value of ADE.

References

- [1] L. Ganobcsik-Williams, (ed.), *Teaching Academic Writing in UK Higher Education*, Palgrave, 2006.
- [2] M. Lea, & B. Stierer, (eds.), *Student Writing in Higher Education: New Contexts*, OU Press, 2000.
- [3] T. Lillis & J. Turner, Student Writing in Higher Education: Contemporary Confusion, Traditional Concerns, *Teaching in Higher Education* **6**(1) (2001), 57-68.
- [4] P. Anderson, What is Web 2.0? Ideas, technologies and implications for education, *JISC Technology and Standards Watch*, February 2007.
- [5] J. Vassileva, Social Learning Environments: New Challenges for AI in Education, keynote, ITS2008, abstract in Woolf et al. (eds.), *Proceedings of ITS2008*, LNCS 5091, 8-8, 2008.
- [6] T. Alehmad, V. Aleven, R. Kraut, Open Community Authoring of Targeted Worked Example Problems, Woolf et al. (eds.), *Proceedings of ITS2008*, LNCS 5091, 216-227, 2008.
- [7] S. Bateman, C. Brooks, G. McCalla, J. Greer, Learning Object Metadata: From the Socially Prescribed to the Socially Derived, *LORNET WORKSHOP @ ITS2008*, June, 2008.
- [8] S. Bateman, R. Farzan, P. Brusilovsky, G. McCalla, OATS: The Open Annotation and Tagging System, *Proceedings of 3rd Annual Int. Scientific Conf. on Object Repository Research Network*, 2006.
- [9] S. Kleanthous & V. Dimitrova, Modelling Semantic Relationships and Centrality to Facilitate Community Knowledge Sharing. *Proceedings of AH2008*, Springer, LNCS, 2008.
- [10] J. Jovanovic, C. Torniani, D. Gasevic, S. Bateman, M. Hatala, Leveraging the Social Semantic web in Intelligent Tutoring Systems, Woolf et al. (eds.), *Proceedings of ITS2008*, LNCS 5091, 563-572, 2008.
- [11] B.M. McLaren, O. Scheuer, M. De Laat, R. Hever, R. De Groot, C.P. Rose, Using Machine Learning Techniques to Analyse and Support Mediation of Student E-discussions, *Proceedings of AIED07*, IOS Press, 331-338, 2007.
- [12] P. Wiemer-Hastings & K-I. Malatesta, Introducing RMT: A Dialog-based Tutor for Research Methods, *Proceedings of AIED2001 Workshop on Dialogue-based Tutoring*, 2001.
- [13] S. Schaffert, Semantic Social Software: Semantically Enabled Social Software or Socially Enabled Semantic Web?, *Proceedings of SEMANTICS 2006*, 99-112, 2006.
- [14] M. Krötzsch, D. Vrandečić, M. Völkel, H. Haller, R. Studer, Semantic Wikipedia. *Journal of Web Semantics*, **5** (2007), 251-261, Elsevier.
- [15] U. Wingate, Doing away with study skills, *Teaching in Higher Education* **11**(4) (2006), 457-469.
- [16] C. Dichev, D. Dicheva, J. Fischer, Identity: How to Name It, How to Find It. *Proceedings of WWW07*, ACM, 2007.