What is the Real Problem: Using Corpus Data to Tailor a Community Environment for Dissertation Writing

Lydia Lau, Royce Neagle, Sirisha Bajanki, Vania Dimitrova, and Roger Boyle
School of Computing, University of Leeds, UK

{ L.M.S.Lau, R.J.Neagle, S.Bajanki, V.G.Dimitrova, R.D.Boyle }@leeds.ac.uk

Abstract. Training in soft skills is becoming paramount in today's educational and societal climate, and receives increasing attention in the area of intelligent learning environments for ill-defined domains. We present a study that analyses written feedback given to undergraduate students by tutors at a key stage of dissertation preparation. This allows us to identify key problems students are facing and to understand how these problems are articulated and addressed by tutors. The results of the study are applied to tailor an existing social semantic web environment (AWESOME Dissertation) to address the needs of a particular community for dissertation writing in Computing.

Keywords: dissertation writing, semantic wikis, social computing, scaffolding.

1 Introduction

Dissertation writing, which is a major challenge faced by most students in higher education, is an example of soft skill training as the process requires the learners to explore, interpret, communicate, and manage their own work and progress during a sustained period of time. A fundamental step in developing such intelligent learning environments for ill-defined domains is to articulate *what problems learners are facing and how to shape the learning environment to effectively address these problems.*

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 [1,2]. An earlier attempt has been made in developing the AWESOME Dissertation Environment (ADE) which exploits social computing to provide holistic support throughout the dissertation process [3]. The intention was to make ADE a generic platform for any students writing a dissertation.

The issue of generic versus subject specific issues quickly emerged in the pilot testing of the ADE and, subsequently, instances for different disciplines were developed for further trials [4]. Experiences from these trials led us (i) to question the adequacy of high level (generic) views of the dissertation writing issues for supporting students who face individual and discipline-specific problems; and (ii) to drive for an environment which can be 'tailored' and 'evolve' with usage in the

community where understanding and interpretation of domain-specific vocabulary and concepts can be shared.

This paper presents a domain-specific study on the use of ADE for Computing. On one hand, we analysed how some dissertation writing problems were handled in the current practice of tutors giving written feedback to final year students at a key stage of dissertation preparation. In parallel, we tailored an initial AWESOME-Computing instance by (a) integrating examples of previous dissertations; and (b) seeding the environment with content that corresponds to some typical problems and tutor feedback. We developed example scenarios of students and tutors interacting with the AWESOME Computing environment to simulate the process of social scaffolding which enables further 'tailoring' as the ADE evolves with use.

2 The AWESOME project

The platform for the study is a novel community environment 'AWESOME Dissertation Environment (ADE)' which uses semantic wikis to implement the pedagogical approach of 'social scaffolding'. It uses MediaWiki[1] and its extension Semantic Media Wiki (Krötzsch et al., 2007) which provides more user-friendly interface to create and query semantic content. The ADE was developed within an interdisciplinary UK research project called AWESOME (Academic Writing Empowered by Social Online Mediated Environments) which involved the universities of Leeds, Coventry and Bangor[2]. The environment was instantiated and trialled in several domains: Education, Fashion and Design, Philosophy and Religious Studies, and an Academic Writing Centre.

The ADE architecture consists of a core ontology which supports semantic mark-ups for

- a scaffold in the form of main stages of dissertation writing process; for examples: getting an overview of dissertation, choosing a topic, adopting an appropriate research methodology, literature `review`, writing up, and project management;
- some common issues associated to each of these stages; for example: for choosing a topic, students need to think about if the topic "has a research question" or "is appropriate for the discipline";
- personal contributions in terms of related top tips, or examples of good writings.

As part of the tailoring process, a separate emerging ontology is built during content creation for additional community driven scaffolds. Features are also provided for users to link content between community and personal spaces. Readers are referred to [4] for a more detailed description of the ADE architecture.

Following both the encouraging feedback from the trial instantiations and the challenges faced in deploying the environment in practice in earlier studies, we conducted a systematic approach in understanding the disparity between generic and

[1] http://www.mediawiki.org

[2] See the AWESOME web site http://awesome.leeds.ac.uk/ for more information.

domain specific vocabulary when adapting the ADE to dissertation writing in Computing.

3 Dissertation Writing Problems Faced by Computing Students

The final year project (or dissertation) is a hallmark of most Computing and Informatics programmes worldwide [5]. Common to many other programmes is the difficulty experienced by Computing students in recording their work: the write-up represents a challenge they have often not encountered earlier in their studies and usually represents the primary (or sole) artifact that is used for assessment.

In our School of Computing, there is an established practice for every dissertation student to prepare a mid-project report under the guidance of the tutor. It is typically 10 pages long, containing background research and progress to date. This mid-project report will be commented on by another academic member of staff (i.e. 'assessor') to provide early written feedback to the student, with no marks given. This collection of assessor feedback forms is potentially a good resource for us to identify common problems and the feedback given to the students as a comparison with the scaffold/core ontology to be provided by the tailored ADE.

A systematic analysis was conducted on a set of 63 authentic feedback forms from 2008 with the aim to identify any common early problem indicators flagged up by the assessors to the students, the suggestions they made and the language used.

3.1 Procedure

1) Content analysis on twenty mid-project report feedback forms was conducted independently by four staff members (as coders), three of whom had considerable experience in assessing and supervising dissertations. Initially, each coder chose his/her own way to annotate the categories of the problem and the associated issues, with the broad understanding of the need to relate the annotation to the final dissertation marking scheme as the student/tutor may use it to judge the impact of the feedback on the work. The marking scheme consisted of: 'Understanding the problem', 'Produce a solution', 'Evaluate the solution', 'Write up' and 'Reflection'.

2) For each issue extracted from the assessor's feedback, the following were recorded: a) the problem as cited, b) the solution as cited, c) annotation to capture the general category that represented the problem and solution (e.g. write-up), and d) annotation to capture the issue category associated to the general category (e.g. scientific style, referencing).

3) A joint review of the annotations used for the first twenty feedback forms was conducted by the four coders with the aim of arriving at a taxonomy of annotated issues.

4) The updated list of annotations (or taxonomy) was then used by the coders to analyse the remaining feedback forms as in step (2) above.

3.2 Results and Analysis

A total of 250 issues were identified from the 63 feedback forms. Some feedback were positive encouragements reinforcing what were done correctly; but most were constructive feedback for further improvements. Following is an example of a feedback which was classified as "evaluation" for generic category and "criteria" for issue category:

- issue cited: "Your evaluation criteria need work";
- problem cited associated to this issue: "what were put in the report were subjective .. and unconvincing";
- solution cited associated to the issue: "focus groups won't help unless the users have a real task they are trying to achieve".

Table 1 summarises the taxonomy emerged from the analysis and the frequency of each being raised as an issue. It is clear that 'write-up' was most problematic as it was highlighted 86 times and with a wide range of issues being commented on. 'Methodology' came second by being mentioned 40 times.

For a specific issue category, such as 'scientific style', a range of comments can be found. For examples: "*no evidence of three prototypes claimed to be produced ... no pointer to the corpus generated*", "*it is important to clearly identify what you have created/developed yourself, and what you have 'inherited' from others*", and "*there should be more x-referencing*". Quite often, these comments represent a range of 'take-for-granted' common knowledge by academics experienced in scientific writing, but the concept of which has not been grasped properly by the students concern.

Here is another example of a common feedback which students are often at lost on remedial actions: "*there is a lack of critical analysis on the literature read*". Tutors found themselves needing to stress on this issue repeatedly, despite classes were run to discuss 'literature review' every year.

3.3 Implications

Our study revealed that the problem for Computing students in recording their work persists: the write-up represents a challenge they have often not encountered earlier in their studies and the dissertation usually represents the primary (or sole) artifact that is used for assessment for this piece of independent study. Previous studies suggested that common dissertation problems are due to dissertation 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 [6,7].

Table 1. Taxonomy for Dissertation Writing Issues for Computing Students

Generic category	Sub-total	Issue category	Frequency
evaluation	29	criteria	14
		depth	10
		missing	5
literature review	35	criticality	11
		depth	22
		web dependence	2
methodology	40	aims	6
		justification	10
		methods	16
		missing	3
		requirements	5
project definition	12	complexity unclear	7
		requirements specification	5
project management	24	milestones	6
		schedule unclear or delayed	18
topic selection	24	novelty	3
		problem clarity	19
		suitability	2
write-up	86	acronyms	4
		code	1
		explanation unclear	7
		formatting (diagrams/maths)	10
		presentation	1
		referencing	13
		scientific style	32
		specific content	4
		structure	2
		use of english	12

Perhaps, the solution lies not only in the prevention of problems but also in the provision of support when issues arise. A deeper understanding of how feedback is being acted on by the students and tutors is needed. The current practice in Computing is for a tutor to discuss the feedback face to face with his/her student. Experience shows that most students need assistance from their tutors to interpret the feedback which often followed by further assimilation in order to understand how to address the issues. This process could take between a week to months, with the exact path of inquiry taken by an individual student unpredictable. A combination of learning-by-example, social and individual learning processes is expected for an individual student to proceed with the rest of the dissertation journey. Pedagogically, it would be beneficial to provide some kind of structure or ‘scaffold’ [8] for

channeling and focusing these activities and a 'public platform' for the sharing of experience.

4 AWESOME-Computing

We suggest that "learning by example" is of great value, but more than that, open discussion by peers and staff of favourable and unfavourable examples *as the issues arise* would be of most help. Thus, a semantic wiki framework in which earlier dissertations exhibiting fragments of annotated good practice, which allowed the student to add her own annotations or questions may be productive.

4.1 Tuning and Seeding

The first step to tailor the generic ADE for Computing was by adding links to previous dissertations. Although these dissertations were available online on a website, there was no facility to comment on specific good practice or examples. AWESOME-computing enabled this by allowing students or tutors to add comments to a dissertation in a wiki fashion (see Fig. 1). Additional semantic markups enabled the example to be pulled into other appropriate wiki pages on specific issues.

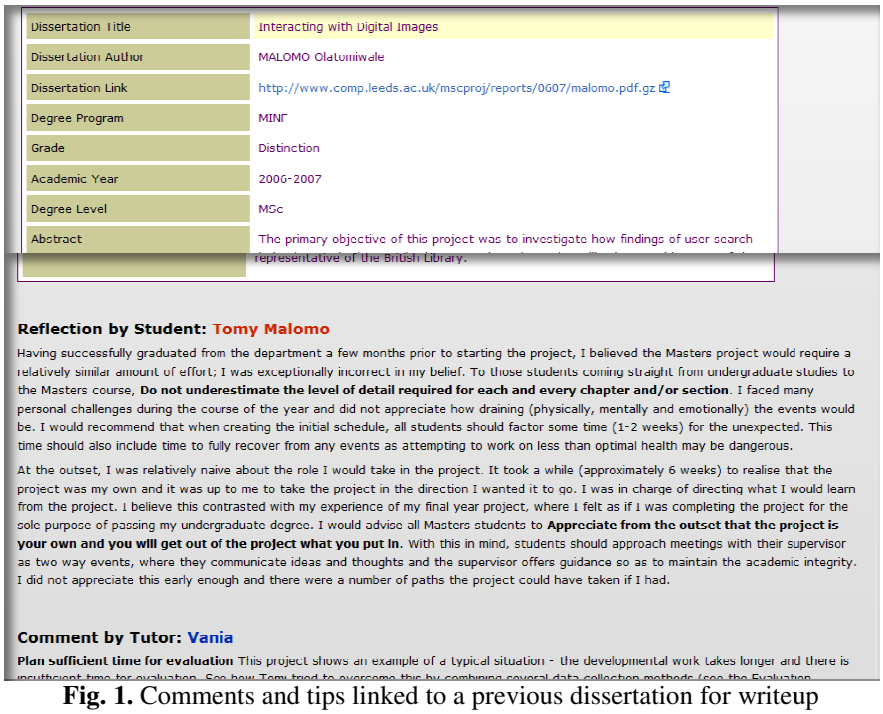

Fig. 1. Comments and tips linked to a previous dissertation for writeup

Secondly, we scaffolded and seeded the environment with content for some anticipated problems and tutor feedback (see Fig. 2 for an example).

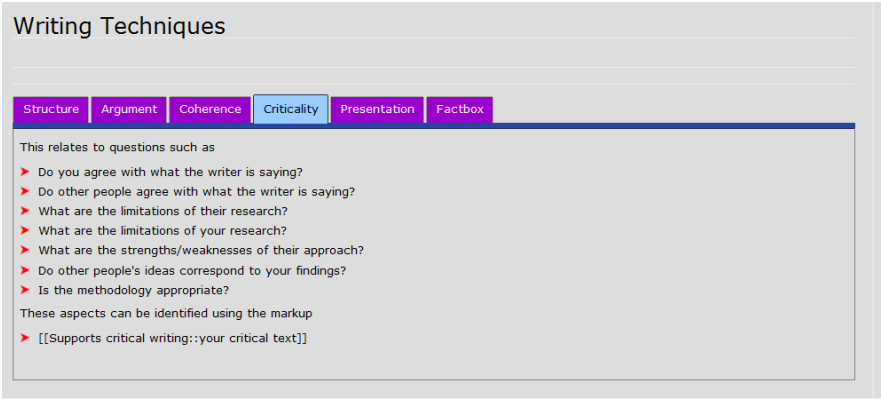

Fig. 2. Having problems relating to 'critical writing'?

4.2 Scenarios for Further Seeding of Content

To provide some useful initial content, a number of scenarios were developed based on real experiences by some tutors. These scenarios were then 'walked through' to populate the ADE with authentic content. Fig. 3 showed an example of the end result of this seeding process, which appears on the home page.

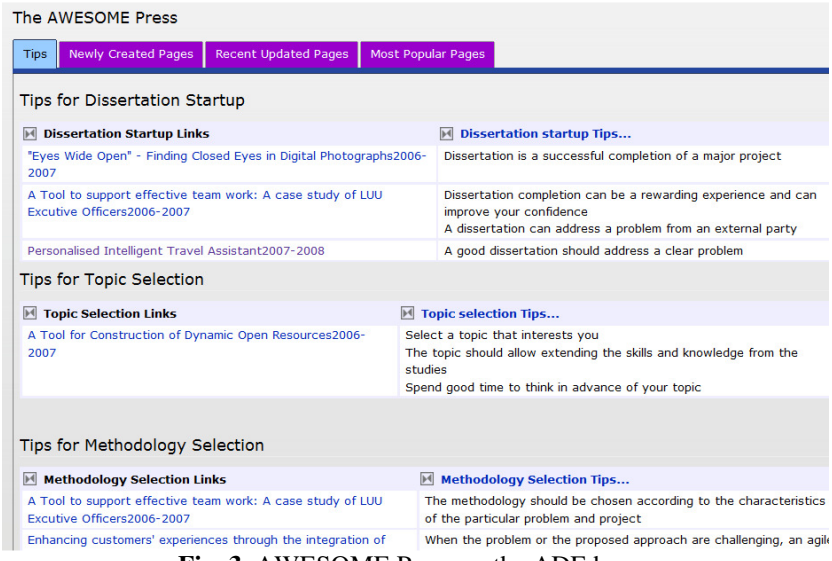

Fig. 3. AWESOME Press on the ADE home page

5 Conclusion

This paper presented an empirical study into the problems faced by a group of Computing students in their dissertation writing by analysing a set of 63 feedback forms on their mid-project reports. The top three problems were: (i) writing not in the expected scientific style, (ii) lack of depth in literature review and (iii) lacking problem clarity. Although classes were held every year to prepare students in tackling these issues, experience showed that many students still struggled to fully understand their relevance when taught.

AWESOME-computing, based on semantic wiki technology, was proposed as a solution to provide complementary support for students to get further assistance in a social context when the issues arisen. We believe in the pedagogical approaches of 'scaffolding' and 'learning by example'. We learned from previous trials that a well-seeded environment is vital for the success of the system launch. Hence, specific scenarios were design to seed a scaffolded environment with real examples and feedback.

References

1. McLaren, B.M., Scheuer, O., De Laat, M., Hever, R., De Groot, R., Rose, C.P. : Using Machine Learning Techniques to Analyse and Support Mediation of Student E-discussions. In:Proceedings of AIED07, pp.331--338, IOS Press (2007)
2. O'Rourke, S.T., Calvo, R.A.: Analysing Semantic Flow in Academic Writing. In: Frontiers in Artificial Intelligence and Applications; Vol. 200; Proceeding of the 2009 conference on Artificial Intelligence in Education, pp.173--180, IOS Press (2009)
3. Dimitrova, V., Lau, L., Le Bek, A.: Sharing of Community Practice through Semantics: A Case Study in Academic Writing. In: Sixth Int'l Workshop on Ontologies and Semantic Web for E-Learning, in conjunction with Int. Conf. on Intelligent Tutoring Systems - ITS'08, pp.30--39 (2008)
4. Bajanki, S., Kaufhold, K., Le Bek, A., Dimitrova, V., Lau, L., O'Rourke, R., Walker, A.: Use of Semantics to Build an Academic Writing Community Environment. In: Proceedings of AIED09, pp.357—364, IOS Press (2009)
5. Fincher, S., Petrie, M., Clark, M.: Computer Science Project Work: Principles and Pragmatics, Springer (2001)
6. Ganobcsik-Williams, L. (ed.): Teaching Academic Writing in UK Higher Education, Palgrave (2006)
7. Lea, M., Stierer, B. (eds.) : Student Writing in Higher Education: New Contexts, OU Press (2000)
8. Pea, R.D.: The Social and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. In: The Journal of the Learning Sciences, 13(3), pp.423--451 (2004)