CROWDSOURCING THE CLASSROOM: INTERACTIVE APPLICATIONS IN HIGHER LEARNING

Richard Littauer, Antonia Scheidel, Marc Schulder, Sibel Ciddi
Universität des Saarlandes (Germany)

{ littauer, ascheidel, mschulder, sciddi }@coli.uni-saarland.de

Abstract

The traditional university classroom paradigm frequently leads to fragmented knowledge dispersal; each student creates, collates, and curates their own notes individually. Here, we present a more holistic approach to student participation, based on experience from a student-initiated experiment using interactive applications in higher education. We explore several methods to help university students work collaboratively: Simultaneous note-taking using an online, interactive notepad; centralised information storage using a dedicated wiki; student-run centralised storage in the cloud; and communication management using a dedicated email list. In this paper, we elaborate on these points, and draw upon relevant previous literature, as well as our experience with implementing this system in several classes with diverse subjects in a graduate level Computational Linguistics degree. We also suggest future avenues to explore, such as slide-and-note integration.

Keywords: digital collaboration, interactive learning, teaching, crowdsourcing, instructional technology

1 INTRODUCTION

The internet is fundamentally changing the course of modern education. Whereas previously students were taught solely from static books, through experience, or in specific classes in their geographic location, today any student with a computer can take online courses and communicate and collaborate with peers across the globe. As the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, a milestone in the open access movement, notes: "The Internet has fundamentally changed the practical and economic realities of distributing scientific knowledge and cultural heritage…In order to realize the vision of a global and accessible representation of knowledge, the future Web has to be sustainable, interactive, and transparent. Content and software tools must be openly accessible and compatible."[1] This aim has led to the creation of new ways of disseminating knowledge, including Web 2.0 tools, collaborative learning environments, and an emphasis on open resources. However, little of this change has become realised within the sphere of higher education, and within the traditional classrooms themselves. But within the classroom, successful collaboration and research methods can gain the most traction; the classroom is where academic habits are first learned, and where students become used to the idea of collaboration.

In this paper, we discuss possible ways to bridge the gap between desired research practices and actual student work. The authors are all students who participated in the same Masters courses for Computational Linguistics at Saarland University. At the beginning of the course, we, along with the other students, began to work together to find the most efficient way to share notes, readings, and material from the lectures, which would maximise learning while minimising administerial work. Here, we present

[1] http://oa.mpg.de/berlin-prozess/berliner-erklarung/

four possible technologies that can be used by students to interactively work together in the classroom, which we found worked best for us and for our coursemates. These are: simultaneous note-taking, central storage, wikis, and email lists. For each of these topics, we give a brief discussion of our methodology in employing these tools. We go on to discuss their relative uses. This information is pertinent for not only classroom planners and education theorists, but also to students and teachers who can apply the methods used here to their own classrooms.

This work follows in the tradition of previous research on computer-supported collaborative learning [1]. Regarding the integration of technology and digital devices into the education arena, with the widespread use of personal computers, laptops and the availability of reliable internet connections, collaborative note-taking and note-sharing has become easier than it had been in the past. For the case study and the methods we describe, it is necessary for each student to have their own personal computer. Provided this is possible, no additional devices are needed (although some system administration is necessary for some approaches).

2 USE CASE STUDY

Over the course of one semester (15 weeks of classes), the authors collaboratively took notes in a number of courses. We discuss and evaluate the process and the results of joint note-taking for the most extensive and most representative class. Although students in each class used at least two of the four collaboration methods discussed in this paper, the range of utilisation varied. This may be due to several reasons: the content of the classes varied, from higher math to programming, linguistics to theoretical modelling and research; the teaching time for each class differed, as did the amount of teachers for each internal module; the subject matter was easier or harder to understand, or, in the case of statistics and formal logic, notation made taking notes electronically more difficult or not.

By far the most prolific class, regarding collaborative note-taking, was the Foundations of Language Science and Technology (FLST) course, a mandatory course for both the Erasmus Mundus Language and Communication Technology Masters and for the Language Science Technology Masters programs. The data gathered (either empirically and anecdotally) from this course was the most useful for showcasing the technologies discussed, and their uses. For using a wiki, the Computational Linguistics for Low Resource Languages (CL4LRL) seminar is the most suitable example, as the wiki was used often and was the core website for the course. For the discussion of centralised storage and dedicated email lists, we will use information gathered from all of the classes together, with only a shallow qualitative analysis in this section.

2.1 Collaborative Note-Taking

While collaborative notes were taken in all of the classes, FLST had the highest amount of notes over the course of the semester. This has been due to several reasons: it is an obligatory course in the programs attended by the authors, it covers a wide range of topics, and the students generally have different backgrounds, so that each student learned at least some novel material. In the course, weekly or bi-weekly homeworks were assigned and discussed in designated tutorial sessions. FLST is a lecture series meant to give an overview of the field and thus ranging from linguistic theories over cognitive foundations to statistical natural language processing. Class met three times a week for 90 minutes each. Lecture slides and exercise sheets were uploaded to a designated course website after the lectures (as with all of the courses), and in some cases also sample solutions to the exercises. Background reading was sometimes uploaded, but mostly only given in form of references in the slides.

Depending on the stability of the respective servers, the authors used etherpad services provided by the Etherpad Foundation[2], MoPad[3] or Riseup[4]. An etherpad is a collaborative online notepad that allows multiple users to edit a document simultaneously, and has a chat bar, basic formatting functionality,

[2] http://etherpad.org/

[3] https://etherpad.mozilla.org/

[4] https://pad.riseup.net/

and stores data for 30 days from the last edit. It was first developed by Appjet Inc., but is now owned by Google - although the technology remains open source.[5] A central etherpad for the entire semester was used to store links to individual etherpads for each class. At the beginning of a class the first contributor to reach the central pad would create a new pad and link to it, specifying the date and topic of the current class. The 30 day limit (when the server would remove the note) meant that the notes were also periodically manually stored in text files, so that the information would not be lost.

There was no explicit division of labour, but aspects of both the technology and the classroom setting facilitated smooth and efficient interaction. As the etherpad shows if and where a collaborator is typing, it was easy to see what a colleague was taking care of and which point still needed to be recorded. E.g., if the lecturer gave a list of three points and one collaborator was transcribing Topic A, another collaborator could move to Topic C to save time and avoid conflicts. Given the fact that notes were taken in small groups (as classes were never attended by more than a total of 30 students) and in seminar rooms rather than large lecture halls, it was also possible to check if an important point was being transcribed simply by glancing over to a fellow note-taker.

Notes were generally taken in a linear fashion, following the lecture. Most of the time, notes were ordered strictly chronologically: If the lecturer spoke about Topic A, diverged to Topic B and then made some more remarks on A, the remark would, too, be added after B and not to a designated section in the notes pertaining to Topic A. This was mostly to keep each individual's contributions visible to all collaborators, since the etherpads we used could not point out that someone was making additions in a document section scrolled out of sight. Editing the document afterwards, for instance during examination revision, often removed such discrepancies. Editing, in some cases, was required: it is significantly more difficult to take notes for statistics or logic classes than for a seminar paper, where the symbols require more detailed attention. In such cases, quick shorthand notes would be taken to keep up with the lecture, and the notepad would be cleaned up later during lulls in conversation or homework.

For the FLST lecture, notes were created in 25 out of 40 classes. Fig. 1 shows the number of words for each set of shared notes, and also the distribution between lectures and tutorial sessions. One of the first things to observe about the distribution is that notes were not usually taken in tutorial sessions - in fact, while notes were taken in 69% of lectures, only 36% of tutorials were transcribed. This can be attributed to several reasons. In some cases, tutorials were held in different locations where students would not necessarily have access to their laptops as freely. Most tutorials, as opposed to lectures, also required active participation in the discussions and did not leave enough time for reflection during the class. Another reason may have been that note-taking was not considered necessary because students corrected their solutions on their own submissions and/or did not have to correct them. Finally, the pedagogical aim of tutorials is different from lectures - it is more about practical experience than dissemination of information.

The discrepancy in coverage between lectures and tutorials will need to be discussed for future efforts. In a largely theoretical course as FLST, tutorials may not be as relevant for exam preparation as the lectures because they do not introduce new material. However, the exercises discussed there are the only time in the semester where students are required to produce answers to questions similar to those asked in the exam. Especially for those struggling with lecture content, tutorials provide a chance to ask questions directly to the teacher and to see the theory put into practice. Because tutorials rarely come with slides of their own, it is a lot harder to review their content than slide-based lectures.

[5] http://etherpad.com/

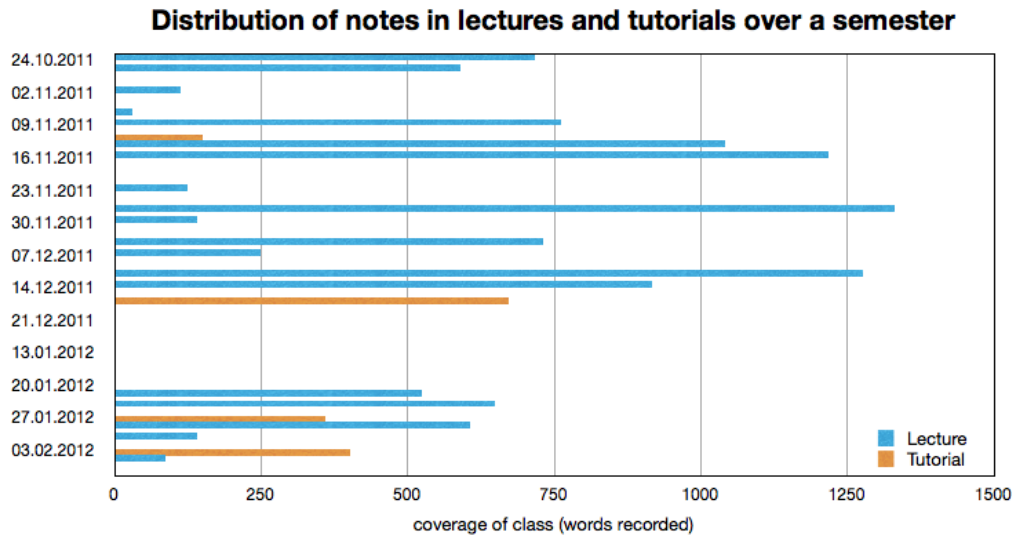

Fig. 1: Words recorded in each lecture or tutorial

Due to these reasons, it may be just as important to transcribe tutorials as lectures - if not even more so. Because of the difficulties mentioned above, note-taking in tutorials will need to be more carefully planned than in lectures: maybe by preparing questions to be asked so only the answers will need to be filled in, or taking notes manually and assembling them later. Most likely, duties will have to be assigned to individual students for portions of the tutorial, with one note-taker at a time focusing only on transcription for a short amount of time, thus making the inevitable neglect of the discussion less pronounced.

For the evaluation of this part of the case study, we analyzed the contents of six of the most extensive sets of notes (between 750 and 1250 words), cf. Fig. 2. For each set, we grouped notes into one of five categories: Slide transcription, links to external sources, text copied from the black board, comments on the lecture, and personal comments. In “comments on the lecture” we summarize subcategories like transcriptions of questions and answers that arose in class, notes on what the teacher pointed out as particularly important, but also exchanges between the students which are pertinent to the lecture. In particular, summaries of slide or lecture content belong to “comments on the lecture”, while “slide transcription” contains all contributions either transcribed or copied from slides which still have the same wording as the slides. The “personal comments” category comprises everything not pertaining to the lecture at all. (It needs to be noted that as the semester progressed, the designated chat window was used for that purpose a lot more. However, we have not included the chat logs in our analysis.)

Fig. 2 shows analyses of notes covering four distinct sub-topics in FLST: Statistical Natural Language Processing, Cognitive Foundations (two lectures each), Semantics and Pragmatics and Finite State Automata. It is visible from the charts above that the same topic can, but does not have to, result in the same note structure, as Statistical NLP and Cognitive Foundations show. Influencing factors were the teaching style (Did the teacher follow their slides very closely or did they resort to the blackboard?), availability of slides (Had they been uploaded before the class? Did they have slide numbers that could easily be referred to?), number of collaborators and also the familiarity of collaborators with the topic.

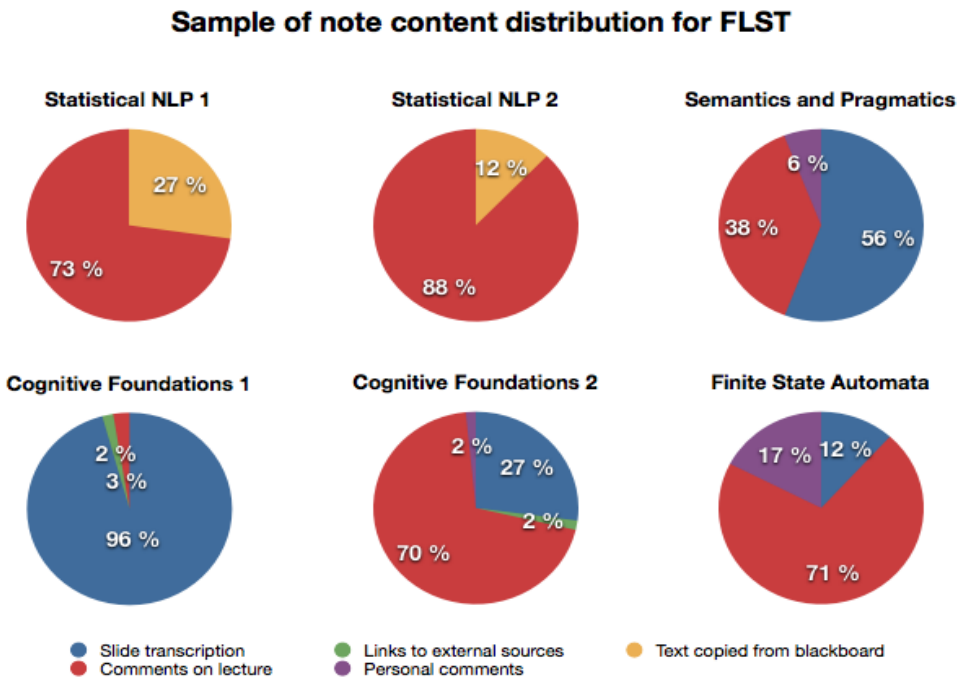

Fig. 2: Distribution of content across a set of shared notes from the FLST lecture

As the lecture series covered a wide range of topics and the authors all have different professional or educational backgrounds, it was not always the case that one collaborator could be called an expert in a given field. For the linguistic foundations, which we were all well familiar with, the decisions which parts of the lecture to comment on seem to have been made mostly based on personal preference of the topic, resulting in a less-than-even coverage of the lecture, albeit with high-quality comments summarizing the points that did make it into the notes. On the other hand, when the lecture covered areas we were all novices in, coverage was broader but shallower (because it was harder to tell which points were important), with a number of unanswered questions remaining in the notes. The most salient notes (not much overlap with slides and a high number of simply worded explanations of complex ideas) were created when one of the collaborators was at an advanced and the others at novice level. This resulted in the “expert” taking concise notes of the most important points and answering questions asked by the novices on the etherpad.

With regard to the quality and helpfulness of the notes, the slide transcription ratio in Fig. 2 is worth inspecting: A very high amount of slide contents can have two disadvantages. On the one hand, genuine new or explanatory content can easily get lost amid the familiar slide content when revising. Notes should be seen as an addition to the slides in exam preparation, not a substitute (As opposed to the plain text of the notes, slides will have an easier to read format - also, even notes with a high slide transcription percentage do not necessarily cover a large portion of the slides.) On the other hand, it suggests that for this particular lecture, an annotation tool (e.g., the comment function in a pdf viewer) would possibly have been more useful than a plain text based etherpad (See Section 3.1).

Out of the notes presented in the above pie charts, the ones on Statistical NLP were probably the most helpful, since their blackboard transcriptions recorded calculations not otherwise available and their comments preserved explanations which the lecturer considered too obvious to be put on slides.

2.2 A Dedicated Wiki

The course Computational Linguistics for Low Resource Languages was a seminar, composed of around

twenty students, where each week two students would present on a relevant assigned paper. There was no final written exam, although there was an oral one, as well as final essays on a topic of the student's choosing. For this reason, there was almost no collaborative note-taking - there would be nothing to study as intensely for, especially considering that the papers were read by the entire class before the presentations. However, it was useful to have a course website with a reading list, as well as a place to archive relevant papers, some extra assignments, and information about the course and the topic. A dedicated, open-source wiki, by MediaWiki[6], was set up for this purpose. The wiki was only open to participants of the course, for privacy and legal concerns. Below, the most frequented pages' view counts are shown in Fig. 3.

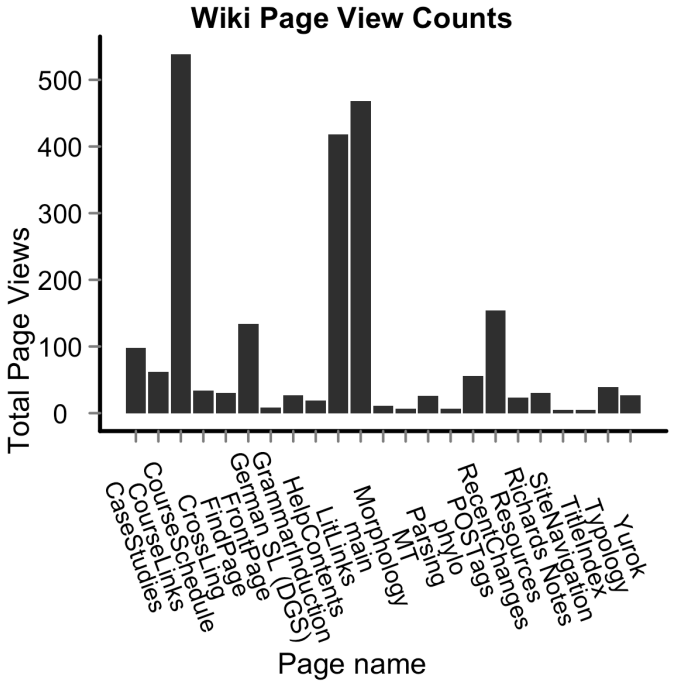

Fig. 3: Wiki Pages View Counts

As can be seen above, the most frequent page was the course schedule, followed by the main page and the literature links. This was in line with the purpose of the wiki - a place to store information about the course. However, the other pages also had a non-trivial amount of use.

The CL4LRL class wiki had roughly 25 edits over the course's length. These ranged from page creation, to uploading presentations, but did not include small edits, such as fixing spelling errors, as quality assurance on an editing level was not necessary. Roughly a dozen users (differentiated by both their login name and semi-anonymous IPs) edited the wiki. In total, there were 18 pages. Four of those pages were orphans, and linked to nothing - in each case, this was because they were created for a specific presentation given by a student, which would be often minimally useful to other students and would not be called upon in the final examination. As this was not a course that overly encouraged the use of a wiki for shared information, 18 pages is not necessarily an exceptionally low number.

2.3 Centralised Storage and Dedicated Email Lists

For each course in the program, there was a centralised email list that could be used to talk among other students. Usage patterns differed - for some courses, the lists were barely used. For others, there were often emails several times a week. The topics of conversation varied widely: administration details, homework assignments and questions, requests for papers that were held behind paywalls, questions

[6] http://www.mediawiki.org/wiki/MediaWiki

about specific lectures, mentions of relevant books or research outside of the syllabus, and so on. For the entire program, there was also a centralised list that was used for larger concerns, such as examination procedures.

In response to early inefficient practices of repeatedly forwarding multiple copies of papers and homework assignments, a centralised Dropbox[7] was set up to gather all of the information in one place. Dropbox is a file hosting service operated by Dropbox, Inc., offering limited storage for free. As of the date of publication, the Dropbox held 814 files, and was 1.07 GB in size. This is not trivial - for many courses, there were dozens of relevant articles, answer sheets, past examinations, and notes. Dropbox was chosen because it was free, easy to use, easily synced to the master storage unit in the cloud, easy to moderate (both to allow members and to ensure quality), and because it had version control (for instance, when someone accidentally deleted a file, it could be recalled.) The majority of the students in the department had access to the Dropbox, except a few who had minor security issues, or did not wish to share files with others.

3 DISCUSSION AND RECOMMENDATIONS

In this section, we discuss the overarching use of our technological methods, and how and why they could be utilised in the classroom.

3.1 Collaborative Note-Taking

We found that using etherpads offers a straightforward and versatile solution to the subtasks of note-taking, i.e., transcribing slide content, summarizing the vocal instructions and additional comments of the teacher, recording external references, and adding personal comments and questions. We found out that our method allowed coverage of all four subtasks, even without a pre-arranged division of labor. When a question was asked about confusing points by one note-taker, awareness of that concern was raised among the others as well, thus making it not just an individual's problem. This simple exchange of questions and reactions to them helped students bring up better questions for the whole class discussion. Therefore, it led to more interactive lecture sessions. It could even become a future strategy to designate the note-taker with the highest expertise in a topic to moderate and encourage others to ask additional questions that might then either be answered in the etherpad or brought up in the lecture.

In the best case, our joint notes resulted in a single, comprehensive resource which could be used for later reference. However, as discussed above, there were a number of limitations. Analysis of our notes showed that tutorials had been largely neglected in the case study of the FLST lecture. As outlined in section 2.1, we believe that tutorials may require a different approach in collaborative (or general) note-taking, to help find a balance between the need for each student (including the note-takers) to follow and benefit from the tutorial, and preserving the insights gained from it for later revision. A clear division of labour among a group of note-takers could ensure that only one student at a time needs to shift their attention from participating in the tutorial to transcribing its content.

The authors felt that taking notes was worthwhile in itself, by helping them focus better on the lecture and improving recall for visual learners. However, another impression was that the notes could become much better revision material by spending some more time on them after the lectures. In a postprocessing step, the linear nature of the notes can be broken up to rearrange items based on which topic they really belong to rather than the order in which they were mentioned. Notes should also be formatted to better reflect the structure, references to newly uploaded slides could be integrated. In short, redundant and confusing parts should be edited, with increased readability and compactness as the main goal.

While the lightweight, OS-agnostic etherpads were easy and fun to use, experiences outside the case study showed that taking notes directly on a saved copy of the slides (in, e.g., Adobe Acrobat or OSX Preview) had a lot of advantages over our method: The visual structure of the slides stays intact, and the note-taker does not have to specify in many words which line it is that they are referring to, as most applications allow users to add comments anywhere on the page. Downsides to this approach include

[7] http://www.dropbox.com

reliance on the availability of uploaded lecture slides, the limited space available for comments (because one slide can only take so many comment boxes) and the fact that these (free) applications do not usually have real-time, collaborative functionality. If they did, as the commercial software Microsoft OneNote does, the note-taking would not involve copying the slides so much as it would following the lecturer. This is definitely a possibility for future work. In the meantime, an option would be to transfer comments to the points on the slides they refer to in the postprocessing step mentioned above. Crucially, commenting on the slides requires uploading of the slides before the class - something which, in our experience, rarely happened.

3.2 Student-run Centralised Storage in the Cloud

Each student essentially needs access to the same material. This includes not only personal notes, but also assigned readings, assigned homework assignments (along with their personal answers and often the corrected answers from the teachers), relevant programs needed for assignments, and so on. Often, students gather notes via email from students who attended the course in previous years, and so the material that being accessed can extend beyond the immediately relevant course work. In our case, we found that this was especially important when the access to assigned readings required paid subscriptions to various journals or article databases. When the course materials were stored for future reference, this eliminated one of the potential obstacles for students.

Rather than fragmenting this storage, collation and curation can be done efficiently using a cloud-based repository, with access for each student. Collaborative cloud storage means that all documents can be accessed not only in a short-term scenario, but also by future classes. Access to the cloud storage can be achieved through a web browser. In addition, many storage services offer software for various operating systems to seamlessly integrate the storage into the computer's file manager, allowing users to manipulate files directly, without having to up- and download them manually. Access to the cloud can be moderated depending on privacy concerns.

We found that having multiple users curating the same archive for material meant that there was a much broader coverage for each class than would exist otherwise. Often, a single student would go through and make sure that all of the lecture notes were there. Further, many students freely uploaded additional materials which they thought were relevant to a particular topic or homework, which results in a broader database than would have existed if only a single student compiled materials. The archiving of past examination questions - including, in some cases, crowd-sourced documents with answers remembered after a test - was particularly useful, as such are not always available for each course. In some cases, the Dropbox was the only folder with stored information on a student's computer; when this was not the case, it was often because a student would have a private folder for personal exercises. Even then, students were actually motivated to share and research more information, due to the exemplary effect of the cloud collaboration.

3.3 Using a Dedicated Wiki

As there is already a broad literature looking at wikis in higher education [2], we present only brief comments here. As described in 2.2, a dedicated wiki as opposed to a static web page run by the school or individual teachers allows for greater flexibility of formatting, for immediate uploading of files and linking of diverse content, and for easy student participation, while establishing a logical content structure. A wiki encourages the creation of a knowledge-base whose focus can be on topics instead of courses. While the scope of the wiki presented here was small, pertaining only to one course which did not require much note taking, there is a greater possibility for more interaction and storage using a wiki. As a permanent and more structured resource combining central storage and collaborative authorship, it is possible to archive data previously acquired through collaborative note-taking and cloud storage more easily. As moderating a wiki takes more thought and effort than these other activities, this can be treated as a long-term task to be undertaken during less intense periods in the academic year.

3.4 A Dedicated Email List

Students often have questions outside of the classroom, for both the teacher and for other students. By being able to crowdsource answers from the rest of the class and the teacher simultaneously, repetitive and unnecessary work can be avoided for all parties. An email list is the most suitable tool for this sort of questioning: it does not require that the students sign in through a complicated content management structure, their university's firewall, or a third party social media source. All queries go directly to the user's inbox. Thus, questions can be resolved quickly using the collective hive-mind of the other students.

Apart from mailing lists for particular courses, we found that a general email list for all students of a year is also very useful in establishing a social and academic group. The list can be used to deal with overarching academic concerns, such as degree regulations, sharing of information not particular to a course, how to deal with administration problems or whether to arrange a study group; but also to plan social outings and promote extracurricular events. Using a mailing list for such activities also has the advantage that no individual students are isolated from the group by not receiving messages.

The idea of a dedicated email list for a university course is not new. However, combined with the other suggestions for electronic student collaboration, we find that an email list was very useful and instrumental in setting up, managing, and using the wikis. Often, a student could answer a question by providing a link to a relevant article in the Dropbox, or page on the wiki. It was also the case that, instead of forwarding an email to that participant, the relevant file that might have helped a questioning student was uploaded to the Dropbox, therefore allowing access for all students.

4 RECOMMENDATIONS

From our experiences, we have several overarching recommendations for future students and teachers regarding how to maximise the use of the four methods described above:

- Spend time after lectures, or during lulls in conversation, to make the notes readable and more structured. This leads to a better understanding of the material both during the lecture and for later reference.
- Communicate with lecturers about the use of slides being uploaded before a class. We found that does help significantly with following along. It also helps with note-taking, exporting to different documents and formats, and provides an impetus to prepare for lectures so that extra information can be garnished from the instructor instead of merely following the slides at the front of a class. Page numbers, sign-posting, and overview slides are especially helpful.
- Short breaks between sections during a class can be incredibly helpful, even if they seem redundant and counter-productive. They allow students to mull over new thoughts, and to structure their notes and their understanding.
- Make sure that someone archives and collates the notes together in a central repository, so that collaborative work is not lost (for instance, due to the 30-day use limit to the Etherpads). crowdsourcing quality assurance and checking the repository is also much easier than having one individual doing this alone.
- Keep systems simple and straightforward to encourage people to contribute regularly. Complicated processes and interfaces inhibit students, by causing them to spend the extra effort in an already intense study environment. The more seamlessly collaboration can be integrated into daily routine, the better.

We hope that these recommendations help with future classes. Another aspect, not mentioned so far in this paper, that makes collaborative work suitable for classrooms is that it is very engaging, and in some cases, *fun* - more so than following a single lecturer without interacting with fellow students.

5 FUTURE WORK

Part of the problem in collaborative note-taking and knowledge archiving is that the original slides and

pedagogical material themselves are seen as static. Ideally, the course slides could be edited from student feedback, or adjusted to meet the pace of the learners. If the teacher were to keep track of the students' collaborative notes, this could be possible, and would be an exciting experiment for a university class. Another future avenue to explore would be course-relevant blogging, which encourages the students to learn the material well enough to disseminate it correctly, or using simultaneous notes in other settings such as group meetings, seminars, conferences, and paper writing.

6 ACKNOWLEDGEMENTS

This work was sponsored by the Erasmus Mundus Language and Communication Technologies Masters program, in which two of the authors, Richard Littauer and Sibel Ciddi, are students.

7 REFERENCES

[1] Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), Cambridge handbook of the learning sciences, pp. 409-426. Cambridge, UK: Cambridge University Press.

[2] Ebner, M., Kickmeier-Rust, M. & Holzinger, A. (2008). Utilizing wiki-systems in higher education classes: A chance for universal access? Universal Access in the Information Society 7(4), pp. 199-207.