

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

Another Suitcase, Another Student Hall – Where Are We Going To? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing

The thirteenth Joint International Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing, held at New York University in Greenwich Village, New York City, between June 13th – June 17th 2001, showcased the wide spectrum of research being undertaken within the field of “Digital Media and Humanities Research” at present. This was the largest ever ACH/ALLC (the oldest established meeting of scholars working at the intersection of advanced information technologies and the humanities), providing a meeting point for the expanding Humanities Computing community, whilst giving an indication of the type and quality of work being pursued in the arena.

The conference was well attended, with over 324 participants registered, and a further 50 or so witnessing the webcasts of the opening and closing sessions (a first for the ACH/ALLC). Whilst over half of these participants came from academic institutions within the USA, the international academic community was also well represented¹. There was an increased presence of attendees from projects out with academic institutions, and industry. The community of ACH/ALLC, whilst being very westernised, is international, predominantly academic, but with necessary and welcome links to the practical application of its techniques in the heritage, teaching, and commercial sectors. This segue between the research based and the practical, industry and academia, demonstrates the rich paradox of the meeting of the computing and humanities fields.

Johanna Trucker, the Robertson Professor in Media Studies at the University of Virginia, in her opening plenary “Reality Check: Problems and Prospects in Digital

¹ There were sizeable numbers from the United Kingdom, Canada, Norway, and Germany. There were also participants from institutions in Italy, Ireland, Austria, Spain, Czech Republic, Romania, Poland, Holland, Belgium, Sweden, Finland, Mexico, Puerto Rico, Japan, Hong Kong, South Africa, and New Zealand.

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

Humanities” explored the meeting of these two different academic disciplines thoroughly. It is hard to know enough to work in the field of Humanities Computing – after all, it is a discipline which embraces the technical subjects but depends on a long background of humanities research. From the perspective of the humanist this is not a happy synthesis: there are un-reconcilable differences between the humanities and computing, but it is Trucker’s belief that keeping those differences alive is the most productive way forward for research in the domain. (A theme revisited in Alan Liu’s closing Keynote “The Tribe of Cool: Information Culture and History”, where he stressed that Humanities Computing has always been about collaboration, and that the cusp between the disciplines offers the most rewarding prospects for research.) Trucker pointed out that there is a malaise surrounding the condition of the humanities, and traditional humanities research, in contemporary culture – but the digerati are happy. In a society where the dominant destructive force of “disneyfication” depends on a sense of historical amnesia, the humanist’s job is to create historical memory. Although it would be romanticising to suggest that technology could ever be our only saviour, those working in digital media and the humanities are optimistic and excited about their research and their place in the more traditional disciplines; we are now in the situation where research is not pushing but being pushed by Humanities Computing. There may indeed be an inherent contradiction in the coupling of the humanities and computing as an academic subject, but Trucker stressed that instead of trying to wholly embrace Engineering or Computing Science working practises, or worrying about whose territory we are poaching on, we should embrace the sense of not really knowing where Humanities Computing is going, and celebrate the “mutually destabilising otherness” of the field, adding that this “fraying and fragmentation is essential to critical thought”.

But if we don’t know where the field is going, can this conference give some indication of where it is at present, or at least where it has been recently? Of course, the papers presented are not the only worthwhile work that is being done in the area of computing and the humanities at the moment, but they do reflect a wide international academic audience and authorship. The conference comprised of 67 individual papers and sessions, split into 10 parallel sessions, with Harold Short commenting that “the key intellectual problem at this conference will be in choosing which of the three

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

sessions to attend". A survey of these sessions, papers, and posters should present a snapshot of the kind of work being undertaken. This conference indicates that work is being done across the whole spectrum of humanities research, with dominant themes emerging in literature and linguistics, the development of digital resources, the evaluation of these resources, and also in the creation of policy, strategy and standards to maintain, manage and preserve such research. The question of how to utilise such resources in teaching was repeatedly raised, as was how to deliver Humanities Computing as a taught course itself. Furthermore, a large proportion of papers at the conference dealt with the development of software and systems to aid in humanities teaching and research, indicating that Humanities Computing scholars are increasingly able and willing to embrace the technically complex, whilst forging interdisciplinary partnerships and associations which further research in the field.

Literary and Linguistic Computing

There were a number of papers addressing the more "traditional" aspects of literary and linguistic computing, such as the analysis of corpora to indicate language change and community development, the stylistic and structural analysis of individual literary texts, and the interrogation of texts to answer questions of authorship attribution. These more traditional applications in the field showed the development of new techniques and software, and were applied to both earlier and modern sources.

Linguistics

Corpus studies was represented in a number of papers. Juola's "The Time Course of Language Change" describes an experiment in using techniques to determine rate of language change in a source that spans decades, to look at how small corpora can be analysed to give evidence of change in the type and structure of language used across a time frame. Horobin, in "The Evolution of Standard Written English: A Corpus Approach" described and demonstrated a corpus of Middle English to assess the development and influence of standard written English, suggesting important factors

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

in the design and use of historical corpora to allow the data to be interrogated in a number of intra- and extra-linguistic ways. A more specific study on the analysis of one feature of language change was illustrated in Blatna and Koprivova's poster, "Czech Approaching English-Verbal Forms With Personal Pronouns Across Styles in the Czech National Corpus" which looked at the increasing frequency of personal pronouns, showing how Czech is becoming closer to English on the lexical but also the grammatical level. The analysis of style was demonstrated in Bolshakova's poster, "Phraseological Database Extended by Educational Material for Learning Scientific Style", which presented an analysis of functional style on scientific and technical prose, examined the phraseology of scientific texts. A modern body of text was analysed in Giordano's "The Genre of Electronic Communication: A Virtual Barbecue Revisited" which argued that empirical linguistic analysis should be an alternative and fruitful way to understand the emergence and structure of a digital community, examining the linguistic structure of chat rooms and discussion groups.

Literary Studies

Techniques used to try and ascertain the order an author intended sections of a text to be published in were discussed in Spencer et al, "Reconstructing the stemma of a textual tradition from the order of sections in manuscripts", and also Bordalejo et al, "The Order of the Canterbury Tales: Praxis of Computer Analysis". The first appropriated analogous technique from evolutionary biology for the restructuring of family trees, describing two techniques where they can calculate the "distance" between a pair of manuscripts through the number of insertions, deletions and transpositions, and how it is possible to reconstruct a stemma from the matrix of pairwise distances among manuscripts. The second presented the results and implications of the use of such computational techniques to produce stemmata based on the tale order of the Canterbury Tales, in order to show a relationship between the textual tradition and the order of the tales. The analysis of more contemporary literature was presented in Gardener, "Versions of Interactivity: Meta-interpretive Response in Hypertext Fiction" which described techniques to identify and interpret the way in which feedback and non-trivial decision making influence or determine a reader's choices when confronted by hypertext as opposed to traditional texts. These

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

recordings, and subsequent analysis of interactions, can be used to show how a text is being interpreted by a reader, whilst paying attention to the distinctive aspects of hypertext fiction.

Elements of style, rather than narrative structure, were considered in Robey, "Rhythm and meter in Italian Renaissance narrative verse", which presented a systematic representation of Dante's *Divine Comedy* by creating an electronic text marked up in terms of accents and syllable divisions, indicating some quite substantial divergences in accentual structure between this and other Renaissance texts. Pawlowski et al, in "Time series modelling in the analysis of Greek metrics" detailed the controversy regarding the rhythmical organisation of Greek texts, and how this can be studied to gain understanding of the underlying rhythms in Greek oral literature.

Problems of authorship attribution were addressed in a number of papers. Hoover, in "Vocabulary Richness and Authorship Reconsidered", demonstrated the usefulness of vocabulary richness for attributing authorship in the stylistic analysis of literary texts. Rudman's, "The DNA Authorship Attribution Model" described a method of representation of an author's style based on the detection of their "intellectual DNA" – that is identifying the number and category of features which identify their work, using statistical analysis. A project specific analysis was covered in Holmes, "A Widow and Her Soldier: The Case of the Picket Letters", which detailed the problems in determining the authorship of a volume of letters, presenting a discussion of the sampling techniques, textual preparation, and the stylometric analysis used to raise questions of subterfuge and questionable authorship. Hoover's poster, "Experiments in Multi-variate Analysis and Authorship Attribution", reported on a project that is re-examining and comparing the statistical techniques used in stylistics. Burrows, in the Busa Presentation Plenary "Questions of Authorship; attribution and beyond" discussed the development and application of statistical and computing tools for the analysis of literary texts, detailing problems in computational stylistics and authorship attribution.

It is obvious from these papers that the role of computational techniques in answering problems in literary and linguistic computing is as pertinent as ever, and such analyses

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

are providing statistical representations of texts that would be impossible without the use of such tools. The use and development of such techniques continues to provide a different viewpoint, and a new way of interrogating and understanding texts for linguistic or literary purposes.

Digital Resources

Textual Encoding and Editing

A large proportion of papers and posters presented at the conference related to the creation, editing, and publishing of digital materials. Although there were a few projects regarding the development of multimedia resources, and problems in evaluation of digital resources, most of these papers focussed particularly on the implementation of TEI standards for encoding and display of textual matter (indicating the bias towards the textual in humanities computing as a discipline). Many of the projects had developed comprehensive, flexible, sets of specifications regarding digitisation, markup, presentation, and delivery when the TEI guidelines were not sufficient in these areas, and discussed the problems in developing encoding guidelines when there were often considerable challenges to design and implementation. This was often coupled with considerations to the extent to which such markup processes could be automated, and the development of suitable tools to view, manipulate and process the encoded data.

A large spectrum of types of source material was covered, for example the process of producing digital books (Gibson and Ruotolo, "Beyond the Web: TEI and the Ebook Revolution" and Bia, "Technical Aspects of the Production Process of Digital Books Using XML-TEI at the Miguel de Cervantes Digital Library"), and the digitisation, markup, presentation, and delivery of letters, (Vanhoutte, "Dancing with DALF: Towards a Digital Archive of Letters Written by Flemish Authors and Composers in the 19th and 20th Century" and Eide, "Putting the Dialog Back Together: Re-creating structure in letter publishing"). Problems with the encoding and display of poetry were covered, showing both the difficulties encountered when dealing with variations

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

between manuscripts (Price et al, "What's Interesting About Whitman's Poetry Manuscripts?") and the development of a standard for encoding verse within the framework of the TEI (Ore et al, "TEI for better or verse" and Schreibman and Chua, "Revisiting Revisions: Employing XML and XSL to Display Deeply Encoded, Multi-Versioned Text"). Using the TEI to produce academic electronic journals was covered in two papers (Unsworth, "Publishing originally digital scholarship at the University of Virginia", and Hannon et al, "Building Belphegor: A Multilingual Electronic Journal using TEI") and the application of the TEI to web pages, a less obvious than usual domain, was considered in Rahtz, "Using the TEI to author web sites".

The encoding of text and the development of suitable tools for use specifically in corpus studies was discussed in Biber et al, "AAC- Digital Resources in Textual Studies", and attention was also drawn towards the need to nurture the involvement of humanities scholars in the process of corpus building and annotation in O'Donnell et al, "OpenText.org: An Experiment in Internet-based Collaborative Humanities Scholarship". Using encoding to aid with the display and interrogation of linguistic systems was considered in Tu, "The Adaption and Breakthrough of Chinese Documents Encoding V: A Case Study of CBETA Digital Triptaka and TEI", and also Bia and Quero "Building Spell Check Facilities for Ancient Spanish". Problems in the markup of different typographical conventions were considered in Russom and Bauman, "Typographic Regularisation in the WWP Textbase" and also the poster presented by Anderson "Markup vs. Character Encoding: The Quandary of Handling the Epigraphical/Papyrological "Underdot" in Computer Representation".

On a more general note, problems with the encoding and integration of different types and formats of historical texts were considered in Fuchs et al, "Digitizing the Difference: the challenge of heterogeneity in the sources of early modern science", whilst a poster by Anderson and Crawford, "Form or Function: Considerations in Presenting Historical Documents on the Web", surveyed the difficulties in representing and encoding marginalia in historical documents. Robertson's, "The Historical Event Mark-up and Linking Project" looked at the need for a common historical markup language through which historical documents can be published.

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

Markup was considered on a more conceptual level in Brown et al, "Intertextual Encoding in the Writing of Women's Literary History" which presented the attempt to devise an encoding scheme for a broad-based, integrating literary history to emphasis intertextual relationships between texts, and this use of TEI in the encoding of an analytical framework was also considered in Wittern, "TEI and Topic Maps" which considered the possibility of encoding not only the concrete features of a text but also aspects of the world as presented through the encoded texts. McCarty, in "The DIY Commentary; Or, What the Reference and the Link Told each other" discussed what the most fully developed traditional examples of the commentary genre teach us about designing and implementing better scholarly information systems, and how it may be possible to imagine an interface to multiple commentaries that would better represent the plural text. The act of Text Encoding itself was considered in Caton, "Towards a Politics of Text Encoding" which undertook a critical evaluation of the process, exploring the notion that it is of a neutral, a-political nature.

These papers show a broad range of focus, from the project specific, to the more outward looking, from those that have taken a very practical stance towards problems of encoding and delivery, to those that are more concerned about the theoretical implications of their actions. However, they indicate the breadth to which the TEI is being implemented throughout the domain as a whole, and the critical processes that have to be developed to solve remaining problems with markup, presentation, and delivery of digital resources. As a representation of the type of textual encoding, editing, and, publishing that is happening in the field of humanities computing, these papers indicate the creativity necessary, and indeed present, when working within the guidelines of the TEI, and the need to embrace these guidelines to further the quality of humanities based digital resources.

Multimedia

However, although these text based resources represent the majority of papers presented regarding the creation of digital resources, this conference presented an

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

increased interest in the visual, and in the creation, dissemination, and evaluation of multimedia resources. Fraistat and Jones, in "Immersive Textuality, the Editing of Virtual Spaces" showed how exploiting the theatrical possibilities of digital environments allowed users to explore a literary text through its virtual representation in digital media, expanding the work into a game, or a space in which to travel through to explore its meaning. Stoicheff and Deshayes, in "The Visual Display of Literary Complexity in a Hypertext Critical Edition of William Faulkner's *The Sound and the Fury*" examined the visual display of textual information in the digital environment, indicating how representing texts visually can give new insights into their structure, and provide resources to aid in teaching. Interest in the physicality of digital texts and presentation was shown with Robertson's quirky, "An e|mediated rhetoric of visibility" exploring the role of typography as an information source within electronic resources, and the role of visual design as a cultural and literary shaping force was demonstrated in Walton's poster, "Cultures and Literacies: South African Students and Western Visual Design on the World Wide Web. The growing interest in the use of Virtual Reality within humanities computing was represented by Beacham and Denard's "The Pompey Project: Digital Research and Virtual Reconstruction of Rome's First Theatre", which detailed the practical and theoretical problems in putting together a Virtual Reality model of the theatre, and an exploration of the problems of integrating historical and archaeological data into a high-end VR space. The presence of these papers at the conference demonstrates a willingness to include new technologies, and their applications, in the evolving humanities computing pantheon, and as such to encourage further research in these non textual areas, which are as pertinent to the use of computing in humanities disciplines as the need for textual resources.

Resource Evaluation

There also seems to be a growing interest in the need for evaluation of these resources, in order to be able to evaluate digital resources in the areas of research, teaching, and in the provision of tools to search and retrieve information, such as in digital library projects. Smith et al, in their session, "Towards a Generative Evaluation Toolbox: a Roundtable" probed the present state of evaluation tools in the

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

areas of digital research projects, digital pedagogical projects, and digital library projects, questioning how you judge project quality, and project success. Mactavish, in "More than Words: Astonishment and Special Effect in Multimedia" called for a better way to evaluate how visual and aural elements function within interactive environments, as humanities computing expands to deal with more non-linguistic elements. Individual projects were also concerned with evaluation issues, such as Roz et al, "The Decameron Web. How does encoding help pedagogy?" which detailed evaluation by users as to the usefulness of the resource whilst indicating how the encoding of a primary text and the related retrieval system contribute to the teaching and learning experience in the digital environment, and many of the textual encoding and editing papers also showed such concern into evaluation issues. The need for evaluation of resources in image retrieval and management was covered in Chen, "Image Retrieval Knowledge and Art History Curriculum in the Digital Age" questioning the effectiveness of the current tools available to search for images from digital archives, and Kraus, "Mimetic Metadata: Linguistic Representations of Visual Objects in Image-Based Electronic Projects" criticising the subjective nature of image descriptions which are the basis for image retrieval systems, and how this effects usability of the resources.

These papers can be taken to indicate the state of "Digital Media and Humanities Research" in the area of digital resources, in that they show the bias towards the textual, although there is a growing interest in other media. Also, that there is a great deal of innovative and creative work being done, both in textual encoding and editing, and the development of other multimedia resources. Those involved are increasingly considering the development process, and how it can be made easier by the development of software and tools for creation and presentation. And ultimately, after exploring thoroughly how to create such resources, the field is learning how to evaluate and utilise such research, to enable the design and implementation of resources which meet the needs of researches, students, and institutions.

Policy, Strategy, and Standards

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

A growing number of papers at this conference discussed the need for the design and implementation of a framework for creating and managing digital resources, detailing the research needed in information storage, standards for preservation and access, and integrated and sophisticated search mechanisms to put in place dynamic architectures for digital scholarship. There was also concern for long term preservation of such resources, discussing reprocessing, and archival standards. Such papers also covered the strategic issues involved in implementing an integrated digital environment for the humanities, detailing the collaboration necessary between various institutions (universities, libraries, archives, museums, and media centres) and individuals to ensure the success and applicability of such endeavours, whilst examining the role of the humanist in the creation of new computer technologies.

Henry et al, in their session "International Strategic and Policy Issues in Networking Digital Resources in the Humanities" discussed the strategic issues regarding implementing an integrated digital environment for the humanities, and how the extensibility of models, strategies and practices between projects and institutions across the sector could aid in the integration and management of resources across a broad community. Staples et al, in their session "Progress of the Supporting Digital Scholarship Project" focussed on the creation, analysis, and reprocessing of digital resources, discussing the progress made regarding collecting and preserving existing digital resources of scholarly research for long-term use and preservation.

Rehberger et al's session "Digitizing the Human: Humanizing the Digital" addressed problems of organisation and implementation of large digital corpora in new media, querying how search engines interfaces influence users' access to the warehouses of information. On a more site-specific note, Ore and Eide, in their poster "The Norwegian Museum Project" discussed the steps taken to develop a common database system for the management of collections for the Norwegian University Museums.

McIver et al, in "A New Framework for Web-based Contributory Encyclopaedias" detailed the motivations, design, and implementation of a framework for creating and managing Web-based contributory encyclopaedias, exploring the continuing need for some centralising mechanism for knowledge organisation such as that in traditional encyclopaedias. Renear et al, in "The W3C Consortium and Standards" provided an

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

introduction into the development and implementation of standards by the W3C, particularly those regarding the TEI, and XML.

These papers indicate a concern over the management of digital resources – now that the humanities scholar is becoming more adept at producing such resources, it is important that these resources are managed, maintained, and preserved, and the implementation of frameworks to do so is logically the next step in the dissemination and preservation of such scholarship, although the development and adoption of such frameworks will demand large investments of time, effort, and finances to provide adequate infrastructures for the humanities.

Teaching

This conference also indicated a growing concern regarding teaching humanities computing as an academic subject, partly to fill the immediate cultural need of trained professionals who understand both the humanities and information technology, and also how to transform the enthusiasm for the subject into undergraduate programmes to enable a wider audience to develop the skills necessary to understand and contribute to the growing arena, but also how to incorporate digital resources into more traditional academic subjects. Hockey et al, in the session “MA Programmes for Humanities and Digital Media” discussed the increasing demand for Humanities Computing as a taught course, comparing and contrasting different MA programmes; one established course and one that has just been developed. The comparison covered both what is expected of the institution, regarding investment, teaching, assessment, and technical provision, and what is expected of the student, regarding course work, theoretical, and practical studies. The presentation showed an emphasis on evaluation and assessment of the programmes themselves, a focus that was mirrored in Unsworth’s session “A Masters Degree in Digital Humanities at the University of Virginia”, which detailed the coursework involved, topics covered, and structure of this particular course, whilst addressing how successful such a course is in providing a scholar with both the necessary computing and humanities based skills needed to have a true understanding of the Digital Humanities field. On a more theoretical level,

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

Liu's Closing Keynote, "The Tribe of Cool: Information Culture and History" questioned the distance between "cool" and the educational system, and how to encourage students to learn about the humanities, when such dominant knowledge structures were not fashionable in today's society, whilst considering the importance of teaching humanities computing as a singular subject.

There were a few papers that considered the role of the humanities computing department or unit in traditional academic institutions, and how they could aid in the integration of technology and teaching in more traditional humanities subjects. Burnard et al's session "Symbiosis or serfdom?... Are you/we/they being Served?" considered the relationship between the users and providers of humanities computing services, and suggested that the key to understanding humanities computing is a focus on the needs and requirements of clientele. The session also stressed the need to stimulate demand by demonstrating the benefits of such technology to a sometimes sceptical audience. Condon et al's poster "Sharing Expertise in the Use of Information and Communication Technology to Enhance Teaching and Learning in the Humanities" presented a collaborative project between institutions to explore how new technologies can support staff and students to make better use of small-group teaching. Stroupe's "Writing Against the Curriculum: Technology, Writing and Reconciling Disciplinary with Social Consciousness" indicated how the tools and techniques of distance, off campus, education can provide an infrastructure for communication across faculties and departments in the humanities and sciences.

Consideration of the use of digital resources for teaching in particular subject areas was also addressed. Duguld's "Digital Pedagogy in Film and Media Studies" discussed the use of digital moving images in the classroom, stressing how using such images provides the ability to develop rich multimedia resources, and eases their dissemination by utilising digital delivery mechanisms. Hamilton's poster "Remote Interactive Animated Projection" also addressed some of these issues whilst focussing on the development of a video streaming facility to aid in humanities research and teaching in an arts faculty, utilising animation because of its flexible nature.

The Development of Software and Tools

A great number of papers at this conference dealt with the development of software and computational tools to aid in humanities research, and how such software can be utilised and tested. Again, there was a bias towards the textual, with many papers detailing systems to aid in the markup of texts and their presentation. Other systems were developed regarding the organisation and analysis of texts, particularly hypertexts, and, again, there were some papers regarding the development of software and systems for non-textual elements of humanities research, including the visual, and music.

Textual Tools

Many of the papers presented tools to aid in the markup of texts. Akhtar et al, in "Automating XML Mark-Up" demonstrated a novel two-stage automatic XML markup system, to automatically extract and apply markup rules to documents by using self-organisation and adaptive automatic markup; learning from its own errors to increase accuracy. Bia and Carrasco, in "Automatic DTD Simplification by Examples", described a method for the automatic generation of simplified DTD from a source DTD and a set of sample marked up files, in order to create a minimum DTD. Sperberg-McQueen and Huitfeldt, in "Practical extraction of meaning from markup using XSLT", documented the development of software to aid in combating the problem of providing a clear, explicit account of the meaning and interpretation of markup by providing a notation for expressing the meaning of constructs in a markup language and the use of that notation to define elements and attributes. Burnard and Fix's, "Introducing Phelix: An Open XML Database System" demonstrated the development of a general purpose TEI-compatible XML database system, configured to support a complex XML DTD, tested on collaboratively designed database of several hundred manuscripts. Porter et al, in "Building flexible language-learning systems: Perl and HTML vs. XML and XSL", discussed two new systems to aid and support language learning, comparing the different technologies and different target

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

audiences, with the common requirements, and the need for build in user friendly systems administration tools.

A large proportion of the papers regarding software and systems development discussed techniques for the organisation, analysis, and querying of electronic texts. Smith, in "Linking and Gathering: Automatic Hypertext in the Perseus Digital Library" described a systems which aims to augment electronic documents by automatically generating hypertexts from the contents of a digital library, to aid in the creation of composite documents for richer contextualisation. Wong and Webster's "Linguistic Description and Exploration using RDF" proposed an approach to the information search and retrieval of linguistic data from texts based on the identification of linguistic information about the rhetorical structure of the text. White et al, in "Co-Cited Author Maps as Real-Time Interfaces for Web-Based Document Retrieval in the Humanities" presented an account of a web based information retrieval interface that aids in the mapping of scholarly literatures by creating maps of inter-related author names, so providing an aid to humanities research which tends to centre work around named persons.

A few of these papers dealt with the categorisation of texts, and how the use of techniques from the information retrieval field, for example numerical classification and categorisation strategies, can replace the manual categorisation process. de Pasquale and Meunier's "Categorisation techniques in computer assisted reading and analysis texts (CARAT) in the humanities" presented the development of tools to categorise texts, asking if these text classification techniques can be applied successfully to the reading and analysis of texts in the humanities and social sciences. Forest et al, "From mathematical classification to thematic analysis of philosophical texts" explored methods of classification and categorisation for the automatic reading and analysis of humanities texts. Leopold and Kindermann's poster "What can Hyperplane-Classifiers tell us about Texts" reported on a project which uses Support Vector Machines for Text Classification, utilising semantic spaces and latent semantic indexing to classify texts.

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

The development of tools to aid in the construction of dictionaries was demonstrated in Tufis and Barbu "Extracting multilingual lexicons from parallel corpora", showing a method for automatic extraction of translation equivalents from parallel corpora to be able to automatically produce bilingual dictionaries, and also Silberztein et al, "Large coverage dictionaries and grammars for text processing: the INTEX system". This session focussed on the use of a corpus processing system for the textual analysis of lexical resources and grammars, explaining advanced methods for information extraction, demonstrating how the tools may be used in research, and how applications may be built based on the technology. The INTEX system was utilised again in Dougherty et al, "INTEX solves Pronunciation and Intonation Problems in Text to Speech Reading Machines" A discussion on how to train text to speech machines to correctly intone sentences so that they retain their meaning, and what could be the optimal structure of lexical entries in order to account for lexical ambiguities in sentences, demonstrating how markup is a practical problem which comes to the fore when developing tools to do a specific task.

The management and analysis of text as part of the internet was addressed in Rockwell et al, "Tracking Culture on the Web: An Experiment" which demonstrated the use of the WWW as a tool to track ideas and culture by the development of a system for tracking selected items, and how such tools can indicate cultural change. Zarri's "The Euforbia Project: A Semantic Approach to the Filtering of Illegal and Harmful Content on the Internet" presented an experiment about the filtering of internet documents according to an unbiased and semantic-rich approach. The English bias of the internet and related technologies was pointed out in Golumbia's "The Computational Object: A Poststructuralist Approach" which demonstrated the focus on standard average European languages in today's computer programming infrastructure, querying how this effects the development of the code, and the accessibility to such systems.

Multimedia Tools

The papers dealing with the development of tools and systems for non-textual humanities research covered a broad scope. Kirschenbaum et al, in "The Virtual

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

Lightbox: The Potential of Peer-to-Peer Humanities Computing” demonstrated the development of an image based software tool which functions as an image based whiteboard for the web, allowing images to be juxtaposed for comparison, discussing the prospects the development of this kind of tool has for humanities computing. Brown and Seales, in “3D Imaging and Processing of Damaged Texts” showed how 3D imaging can be used as a means of creating, manipulating, restoring and carefully measuring features on digital facsimiles of manuscripts, applying new restoration techniques such as flattening to aid historians in their studying of manuscripts and other texts. Terras, in “Reading the Papyrologist: Building Systems to Aid the Humanities Expert” discussed the process involved in working with humanities experts in order to identify what type of computer tools will help them carry out their task, and the construction of a system to aid papyrologists in reading ancient texts. The analysis of music scores was presented by Ng’s poster, “Optical Music Recognition: Stroke Tracing and Reconstruction of Hand-written Manuscripts”, which documented an automatic and efficient method to transform paper-based music scores into a machine representation. Bod’s, “Using Natural Language Processing Techniques for Musical Parsing” presented an investigation into whether it is possible to use probabilistic parsing techniques from Natural Language Processing to parse Music into groups and phrases which can be represented in a tree structure. The paper presented the development of a new parser which combines techniques from probabilistic heuristics to solve ambiguity in order for it to parse music accurately. Ng, in “Music via Motion: Interactive Multimedia Performances” demonstrated a motion and colour detection system which uses a video camera to survey a live scene and track visual changes.

Again, these papers show the breadth of vision humanities computing is developing: the inclusion of papers dealing with non-textual elements show how the field is evolving and embracing a much wider remit, rather than just retaining a textual focus. These papers also indicate that some of the technical problems faced by scholars in the humanities are being solved by such scholars themselves, as they develop new techniques, tools, and software, and implement and test these new systems on a variety of sources. A number of the papers show that there is great inter-disciplinary collaboration between scholars to aid in the development of such tools, and that, as

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

Trucker intoned in her opening session, the meeting of the technical and the humanities based expert can often prove to be a rich collaboration.

The Social Aspect...

Socially, the conference was calmer than it has been in previous years. Being in the heart of Greenwich Village delegates were mostly left to fend for themselves, and the social schedule was kept to a minimum with a few drinks receptions, although there was still adequate opportunity to socialise between and after sessions. The conference culminated in a banquet at the 5th Avenue Ballroom, but unfortunately there was no dancing (no dancing?), with the hardiest having to go onto to a blues club in the Village. 9.30 session on Saturday morning, anyone?

Conclusions

The remit of such a conference has moved away from the merely literary and linguistic side of computing, towards computing and the humanities in its broadest sense, encompassing not only the development of digital resources, but their management, evaluation, and preservation, the development of computational tools to aid the humanities scholar in both textual and multimedia systems, and the development of teaching programmes and techniques to increase the presence of humanities computing in academic establishments. This conference has shown the interdisciplinary nature of such research, and how such collaborations can be fruitful. However, there still remains a lot to be done in all the areas mentioned, presenting great opportunities for the scholars involved. Humanities Computing is still a very young, developing field, indicated by the fact that so many young researchers were attending and presenting at the conference, alongside more established academics in the arena. There is a feeling of approachability – and although the conference is getting larger every year, there still remains a palpable sense of community, what Allan Reneer, the president of ALLC, called the “social work” of humanities computing; the building of connections between people and projects to aid work and advance research, and how such a conference is increasingly about the relationships

Terras, M. (2002). "Another Suitcase, Another Student Hall, Where are we going to? What ACH/ALLC 2001 can tell us about the current direction of Humanities Computing" *Literary and Linguistic Computing*, Volume 16, Issue 3.

fostered and connections made at such meetings. The quality of discussion after the papers was firm evidence of this, and was often commented upon. Compare this to conferences in more established traditional fields in the arts or sciences; anyone having attending such a conference will attest to the difference in tone. The phrase “warm and fuzzy” was used repeatedly to describe the ALLC/ACH Community at this conference – and we presume this had nothing to do with the NYC humidity.

Humanities Computing appears to be an academic field in transit: having not long started its journey, this conference attests to the momentum that is building, and the quality and disparate nature of the research presented indicates that it has far to travel whilst technologies are explored, infrastructures put in place, and resources developed, evaluated, and utilised. So where is it all going? In many directions. And Tübingen, 2002...