digital humanities 2012

Conference Abstracts

University of Hamburg, July 16–22

Hamburg University Press

The Alliance of Digital Humanities Organizations
The Association for Literary and Linguistic Computing
The Association for Computers and the Humanities
The Australasian Association for Digital Humanities
centerNet
The Society for Digital Humanities – Société pour l'étude des médias interactifs

Digital Humanities 2012

Conference Abstracts

University of Hamburg, Germany
July 16–22, 2012

Hamburg University Press
Publishing House of the Hamburg State and University Library
Carl von Ossietzky

Printed with the Support of the German Research Foundation

Editor
Jan Christoph Meister

Editorial Assistant
Katrin Schönert

Editorial Staff
Bastian Lomsché
Wilhelm Schernus
Lena Schüch
Meike Stegkemper

Technical Support
Benjamin W. Bohl
Daniel Röwenstrunk

Bibliographic information published by the *Deutsche Nationalbibliothek* (German National Library).
The *Deutsche Nationalbibliothek* lists this publication in the *Deutsche Nationalbibliografie*; detailed bibliographic data are available on the Internet at https://portal.dnb.de/
The *Deutsche Nationalbibliothek* stores this online publication on its Archive Server. The Archive Server is part of the deposit system for long-term availability of digital publications.
Also available open access on the Internet at:
Hamburg University Press – http://hup.sub.uni-hamburg.de
PURL: http://hup.sub.uni-hamburg.de/HamburgUP/DH2012_Book_of_Abstracts
For an elaborated version of this Book of Abstracts with color photos and zoom function, please visit: www.dh2012.uni-hamburg.de
ISBN 978-3-937816-99-9 (printed version)
© 2012 Hamburg University Press, publishing house of the Hamburg State and University Library Carl von Ossietzky, Germany
Cover design: Turan Usuk
Cover illustration: Dagmar Schwelle/laif
Printing house: Elbepartner, BuK! Breitschuh & Kock GmbH, Hamburg, Germany

The 24th Joint International Conference of
the Association for Literary and Linguistic Computing and
the Association for Computers and the Humanities

and

The 5th Joint International Conference of
the Association for Literary and Linguistic Computing,
the Association for Computers and the Humanities,
the Society for Digital Humanities – Société pour l'étude des médias interactifs,
for the first time organized together with
the Australasian Association for Digital Humanities and
centerNet

Table of Contents

List of Reviewers 1

Plenary Sessions

Dynamics and Diversity: Exploring European and Transnational Perspectives on Digital Humanities Research Infrastructures
Moulin, Claudine 9

Embracing a Distant View of the Digital Humanities
Shimoda, Masahiro 11

Pre-conference Workshops

Digital Methods in Manuscript Studies
Brockmann, Christian; Wangchuk, Dorji 15

Introduction to Stylomatic Analysis using R
Eder, Maciej; Rybicki, Jan 16

NeDiMAH workshop on ontology based annotation
Eide, Øyvind; Ore, Christian-Emil; Rahtz, Sebastian 18

Service-oriented Architectures (SOAs) for the Humanities: Solutions and Impacts
Hinrichs, Erhard; Neuroth, Heike; Wittenburg, Peter 20

Here and There, Then and Now – Modelling Space and Time in the Humanities
Isaksen, Leif; Day, Shawn; Andresen, Jens; Hyvönen, Eero; Mäkelä, Eetu 22

Crowdsourcing meaning: a hands-on introduction to CLÉA, the Collaborative Literature Éxploration and Annotation Environment
Petris, Marco; Gius, Evelyn; Schüch, Lena; Meister, Jan Christoph 24

Learning to play like a programmer: web mash-ups and scripting for beginners
Ridge, Mia 25

Introduction to Distant Reading Techniques with Voyant Tools, Multilingual Edition
Sinclair, Stéfan; Rockwell, Geoffrey 26

Towards a reference curriculum for the Digital Humanities
Thaller, Manfred 27

Free your metadata: a practical approach towards metadata cleaning and vocabulary reconciliation
van Hooland, Seth; Verborgh, Ruben; De Wilde, Max 28

Panels

Text Analysis Meets Text Encoding
Bauman, Syd; Hoover, David; van Dalen-Oskam, Karina; Piez, Wendell 33

Designing Interactive Reading Environments for the Online Scholarly Edition
Blandford, Ann; Brown, Susan; Dobson, Teresa; Faisal, Sarah; Fiorentino, Carlos; Frizzera, Luciano; Giacometti, Alejandro; Heller, Brooke; Ilovan, Mihaela; Michura, Piotr; Nelson, Brent; Radzikowska, Milena; Rockwell, Geoffrey; Ruecker, Stan; Sinclair, Stéfan; Sondheim, Daniel; Warwick, Claire; Windsor, Jennifer 35

Developing the spatial humanities: Geo-spatial technologies as a platform for cross-disciplinary scholarship
Bodenhamer, David; Gregory, Ian; Ell, Paul; Hallam, Julia; Harris, Trevor; Schwartz, Robert 41

Prosopographical Databases, Text-Mining, GIS and System Interoperability for Chinese History and Literature
Bol, Peter Kees; Hsiang, Jieh; Fong, Grace 43

Future Developments for TEI ODD
Cummings, James; Rahtz, Sebastian; Burnard, Lou; Bauman, Syd; Gaiffe, Bertrand; Romary, Laurent; Bański, Piotr 52

Compiling large historical reference corpora of German: Quality Assurance, Interoperability and Collaboration in the Process of Publication of Digitized Historical Prints
Geyken, Alexander; Gloning, Thomas; Stäcker, Thomas 54

Computational models of narrative structure
Löwe, Benedikt; Físseni, Bernhard; León, Carlos; Bod, Rens 57

Approaches to the Treatment of Primary Materials in Digital Lexicons: Examples of the New Generation of Digital Lexicons for Buddhist Studies
Nagasaki, Kiyonori; Tomabechi, Toru; Wangchuk, Dorji; Takahashi, Koichi; Wallman, Jeff; Muller, A. Charles 61

Topic Modeling the Past
Nelson, Robert K.; Mimno, David; Brown, Travis 64

Facilitating Research through Social-Document Networks
Pitti, Daniel; Simon, Agnès; Vitali, Stefano; Arnold, Kerstin 70

Digital Humanities as a university degree: The status quo and beyond
Thaller, Manfred; Sahle, Patrick; Clavaud, Florence; Clement, Tanya; Fiormonte, Domenico; Pierazzo, Elena; Rehbein, Malte; Rockwell, Geoffrey; Schreibman, Susan; Sinclair, Stéfan 72

Papers

Exploring Originality in User-Generated Content with Network and Image Analysis Tools
Akdag Salah, Alkim Almila; Salah, Albert Ali; Douglass, Jeremy; Manovich, Lev 79

Patchworks and Field-Boundaries: Visualizing the History of English
Alexander, Marc 82

Developing Transcultural Competence in the Study of World Literatures: Golden Age Literature Glossary Online (GALGO)
Alonso Garcia, Nuria; Caplan, Alison 84

Trees of Texts – Models and methods for an updated theory of medieval text stemmatology
Andrews, Tara Lee; Macé, Caroline 85

Mapping the Information Science Domain
Arazy, Ofer; Ruecker, Stan; Rodriguez, Omar; Giacometti, Alejandro; Zhang, Lu; Chun, Su 88

Words made Image. Towards a Language-Based Segmentation of Digitized Art Collections
Armaselu, Florentina 91

HisDoc: Historical Document Analysis, Recognition, and Retrieval
Baechler, Micheal; Fischer, Andreas; Naji, Nada; Ingold, Rolf; Bunke, Horst; Savoy, Jacques 94

Research infrastructures for Digital Humanities: The local perspective
Bärenfänger, Maja; Binder, Frank 97

Pelagios: An Information Superhighway for the Ancient World
Barker, Elton; Simon, Rainer; Isaksen, Leif 99

Putting TEI Tite to use – generating a database resource from a printed dictionary or reference type publication
Barner-Rasmussen, Michael 102

Digital Humanities in the Classroom: Introducing a New Editing Platform for Source Documents in Classics
Beaulieu, Marie-Claire; Almas, Bridget 105

DiaView: Visualise Cultural Change in Diachronic Corpora
Beavan, David 107

Catch + Release: Research and Creation of a Digital New Media Exhibition in the Context of a Cultural and Heritage Museum
Beer, Ruth 109

Opportunity and accountability in the 'eResearch push'
Bellamy, Craig 111

Connecting European Women Writers. The Selma Lagerlöf Archive and Women Writers Database
Bergenmar, Jenny; Olsson, Leif-Jöran 113

Stylometric Analysis of Chinese Buddhist texts: Do different Chinese translations of the 'Gandhavyūha' reflect stylistic features that are typical for their age?
Bingenheimer, Marcus; Hung, Jen-Jou; Hsieh, Cheng-en 115

Information Extraction on Noisy Texts for Historical Research
Blanke, Tobias; Bryant, Michael; Speck, Reto; Kristel, Conny 117

Modeling Gender: The 'Rise and Rise' of the Australian Woman Novelist
Bode, Katherine 119

Contextual factors in literary quality judgments: A quantitative analysis of an online writing community
Boot, Peter 121

Violence and the Digital Humanities Text as Pharmakon
Bradley, Adam James 123

Towards a bibliographic model of illustrations in the early modern illustrated book
Bradley, John; Pigney, Stephen 124

Automatic Mining of Valence Compounds for German: A Corpus-Based Approach
Brock, Anne; Henrich, Verena; Hinrichs, Erhard; Versley, Yannick 126

Networks of networks: a critical review of formal network methods in archaeology through citation network analysis and close reading
Brughmans, Tom 129

On the dual nature of written texts and its implications for the encoding of genetic manuscripts
Brüning, Gerrit; Henzel, Katrin; Pravida, Dietmar 131

Automatic recognition of speech, thought and writing representation in German narrative texts
Brunner, Annelen 135

Bringing Modern Spell Checking Approaches to Ancient Texts – Automated Suggestions for Incomplete Words
Büchler, Marco; Kruse, Sebastian; Eckart, Thomas 137

Designing a national 'Virtual Laboratory' for the humanities: the Australian HuNI project
Burrows, Toby Nicolas 139

Beyond Embedded Markup
Buzzetti, Dino; Thaller, Manfred 142

Myopia: A Visualization Tool in Support of Close Reading
Chaturvedi, Manish; Gannod, Gerald; Mandell, Laura; Armstrong, Helen; Hodgson, Eric 148

Translation Arrays: Exploring Cultural Heritage Texts Across Languages
Cheesman, Tom; Thiel, Stephan; Flanagan, Kevin; Zhao, Geng; Ehrmann, Alison; Laramee, Robert S.; Hope, Jonathan; Berry, David M. 151

Constructing a Chinese as Second Language Learner Corpus for Language Learning and Research
Chen, Howard 154

Social Curation of large multimedia collections on the cloud
Chong, Dazhi; Coppage, Samuel; Gu, Xiangyi; Maly, Kurt; Wu, Harris; Zubair, Mohammad 155

Sounding for Meaning: Analyzing Aural Patterns Across Large Digital Collections
Clement, Tanya; Auvil, Loretta; Tcheng, David; Capitanu, Boris; Monroe, Megan; Goel, Ankita 158

The Programming Historian 2: A Participatory Textbook
Crymble, Adam H.; MacEachern, Alan; Turkel, William J. 162

Multilingual and Semantic Extension of Folk Tale Catalogues
Declerck, Thierry; Lendvai, Piroska; Darányi, Sándor 163

Digital Language Archives and Less-Networked Speaker Communities
Dobrin, Lise M. 167

Language Documentation and Digital Humanities: The (DoBeS) Language Archive
Drude, Sebastian; Trilsbeek, Paul; Broeder, Daan 169

The potential of using crowd-sourced data to re-explore the demography of Victorian Britain
Duke-Williams, Oliver William 173

Sharing Ancient Wisdoms: developing structures for tracking cultural dynamics by linking moral and philosophical anthologies with their source and recipient texts
Dunn, Stuart; Hedges, Mark; Jordanous, Anna; Lawrence, Faith; Roueche, Charlotte; Tupman, Charlotte; Wakelnig, Elvira 176

Recovering the Recovered Text: Diversity, Canon Building, and Digital Studies
Earhart, Amy 179

Mind your corpus: systematic errors in authorship attribution
Eder, Maciej 181

Underspecified, Ambiguous or Formal. Problems in Creating Maps Based on Texts
Eide, Øyvind 185

A Frequency Dictionary of Modern Written and Oral Media Arabic
Elmaz, Orhan 188

Texts in Motion – Rethinking Reader Annotations in Online Literary Texts
Fendt, Kurt E.; Kelley, Wyn; Zhang, Jia; Della Costa, Dave 190

May Humanists Learn from Artists a New Way to Interact with Digital Technology?
Franchi, Stefano 192

A flexible model for the collaborative annotation of digitized literary works
Gayoso-Cabada, Joaquin; Ruiz, Cesar; Pablo-Nuñez, Luis; Sarasa-Cabezuelo, Antonio; Goicoechea-de-Jorge, Maria; Sanz-Cabrerizo, Amelia; Sierra-Rodriguez, Jose-Luis 195

HyperMachiavel: a translation comparison tool
Gedzelman, Séverine; Zancarini, Jean-Claude 198

Discrimination sémantique par la traduction automatique, expériences sur le dictionnaire français de Littré
Glorieux, Frédéric; Jolivet, Vincent 202

The Myth of the New: Mass Digitization, Distant Reading and the Future of the Book
Gooding, Paul Matthew; Warwick, Claire; Terras, Melissa 204

Designing Navigation Tools for an Environmental Humanities Portal: Considerations and Critical Assessments
Graf von Hardenberg, Wilko; Coulter, Kimberly 206

Processing Email Archives in Special Collections
Hangal, Sudheendra; Chan, Peter; Lam, Monica S.; Heer, Jeffrey 208

The Stylometry of Collaborative Translation
Heydel, Magda; Rybicki, Jan 212

Focus on Users in the Open Development of the National Digital Library of Finland
Hirvonen, Ville; Kautonen, Heli Johanna 215

The Rarer They Are, the More There Are, the Less They Matter
Hoover, David 218

Experiments in Digital Philosophy – Putting new paradigms to the test in the Agora project
Hrachovec, Herbert; Carusi, Annamaria; Huentelmann, Raphael; Pichler, Alois; Antonio, Lamarra; Cristina, Marras; Alessio, Piccioli; Lou, Burnard 221

Information Discovery in the Chinese Recorder Index
Hsiang, Jieh; Kong, Jung-Wei; Sung, Allan 224

Complex Network Perspective on Graphic Form System of Hanzi
Hu, Jiajia; Wang, Ning 228

A Computer-Based Approach for Predicting the Translation Time Period of Early Chinese Buddhism Translation
Hung, Jen-Jou; Bingenheimer, Marcus; Kwok, Jieli 230

Bridging Multicultural Communities: Developing a Framework for a European Network of Museum, Libraries and Public Cultural Institutions
Innocenti, Perla; Richards, John; Wieber, Sabine 232

Ptolemy's Geography and the Birth of GIS
Isaksen, Leif 236

Tracing the history of Noh texts by mathematical methods. Validitating the application of phylogenetic methods to Noh texts
Iwata, Yoshimi 239

Computing and Visualizing the 19th-Century Literary Genome
Jockers, Matthew 242

Using the Google Ngram Corpus to Measure Cultural Complexity
Juola, Patrick 245

'All Rights Worth Recombination': Post-Hacker Culture and ASCII Literature (1983-1993)
Katelnikoff, Joel 247

Evaluating Unmasking for Cross-Genre Authorship Verification
Kestemont, Mike; Luyckx, Kim; Daelemans, Walter; Crombez, Thomas 249

Literary Wikis: Crowd-sourcing the Analysis and Annotation of Pynchon, Eco and Others
Ketzan, Erik 252

Social Network Analysis and Visualization in 'The Papers of Thomas Jefferson'
Klein, Lauren Frederica 254

VariaLog: how to locate words in a French Renaissance Virtual Library
Lay, Marie Hélène 256

DeRiK: A German Reference Corpus of Computer-Mediated Communication
Lemnitzer, Lothar; Beißwenger, Michael; Ermakova, Maria; Geyken, Alexander; Storrer, Angelika 259

Estimating the Distinctiveness of Graphemes and Allographs in Palaeographic Classification
Levy, Noga; Wolf, Lior; Dershowitz, Nachum; Stokes, Peter 264

Academic Research in the Blogosphere: Adapting to New Opportunities and Risks on the Internet
Littauer, Richard; Winters, James; Roberts, Sean; Little, Hannah; Pleyer, Michael; Benzon, Bill 268

Feeling the View: Reading Affective Orientation of Tagged Images
Liu, Jyi-Shane; Peng, Sheng-Yang 270

Characterizing Authorship Style Using Linguistic Features
Lucic, Ana; Blake, Catherine 273

Investigating the genealogical relatedness of the endagered Dagon languages
Moran, Steven; Prokic, Jelena 276

Landscapes, languages and data structures: Issues in building the Placenames Database of Ireland
Měchura, Michal Boleslav 278

Interoperability of Language Documentation Tools and Materials for Local Communities
Nakhimovsky, Alexander; Good, Jeff; Myers, Tom 280

Content Creation by Domain Experts in a Semantic GIS System
Nakhimovsky, Alexander; Myers, Tom 283

From Preserving Language Resources to Serving Language Speakers: New Prospects for Endangered Languages Archives
Nathan, David John 286

Retrieving Writing Patterns From Historical Manuscripts Using Local Descriptors
Neumann, Bernd; Herzog, Rainer; Solth, Arved; Bestmann, Oliver; Scheel, Julian 288

Distractorless Authorship Verification
Noecker Jr., John; Ryan, Michael 292

Cataloguing linguistic diversity: Glottolog/Langdoc
Nordhoff, Sebastian; Hammarström, Harald 296

Geo-Temporal Interpretation of Archival Collections Using Neatline
Nowviskie, Bethany; Graham, Wayne; McClure, David; Boggs, Jeremy; Rochester, Eric 299

Enriching Digital Libraries Contents with SemLib Semantic Annotation System
Nucci, Michele; Grassi, Marco; Morbidoni, Christian; Piazza, Francesco 303

The VL3: A Project at the Crossroads between Linguistics and Computer Science
Nuñez, Camelia Gianina; Mavillard, Antonio Jiménez 306

'Eric, you do not humble well': The Image of the Modern Vampire in Text and on Screen
Opas-Hänninen, Lisa Lena; Hettel, Jacqueline; Toljamo, Tuomo; Seppänen, Tapio 308

Electronic Deconstruction of an argument using corpus linguistic analysis of its on-line discussion forum supplement
O'Halloran, Kieran Anthony 310

Citygram One: Visualizing Urban Acoustic Ecology
Park, Tae Hong; Miller, Ben; Shrestha, Ayush; Lee, Sangmi; Turner, Jonathan; Marse, Alex 313

Towards Wittgenstein on the Semantic Web
Pichler, Alois; Zöllner-Weber, Amélie 318

Uncovering lost histories through GeoStoryteller: A digital GeoHumanities project
Rabina, Debbie L.; Cocciolo, Anthony 322

Workflows as Structured Surfaces
Radzikowska, Milena; Ruecker, Stan; Rockwell, Geoffrey; Brown, Susan; Frizzera, Luciano; INKE Research Group 324

Code-Generation Techniques for XML Collections Interoperability
Ramsay, Stephen; Pytlik-Zillig, Brian 327

Uncertain Date, Uncertain Place: Interpreting the History of Jewish Communities in the Byzantine Empire using GIS
Rees, Gethin Powell 329

Code sprints and Infrastructure
Reside, Doug; Fraistat, Neil; Vershbow, Ben; van Zundert, Joris Job 331

Digital Genetic Criticism of RENT
Reside, Doug 333

On the Internet, nobody knows you're a historian: exploring resistance to crowdsourced resources among historians
Ridge, Mia 335

Formal Semantic Modeling for Human and Machine-based Decoding of Medieval Manuscripts
Ritsema van Eck, Marianne Petra; Schomaker, Lambert 336

The Swallow Flies Swiftly Through: An Analysis of Humanist
Rockwell, Geoffrey; Sinclair, Stéfan 339

The Digital Mellini Project: Exploring New Tools & Methods for Art-historical Research & Publication
Rodríguez, Nuria; Baca, Murtha; Albrezzi, Francesca; Longaker, Rachel 342

Intertextuality and Influence in the Age of Enlightenment: Sequence Alignment Applications for Humanities Research
Roe, Glenn H.; The ARTFL Project 345

Engaging the Museum Space: Mobilising Visitor Engagement with Digital Content Creation
Ross, Claire Stephanie; Gray, Steven; Warwick, Claire; Hudson Smith, Andrew; Terras, Melissa 348

Aiding the Interpretation of Ancient Documents
Roued-Cunliffe, Henriette 351

The Twelve Disputed 'Federalist' Papers: A Case for Collaboration
Rudman, Joseph 353

Writing with Sound: Composing Multimodal, Long-Form Scholarship
Sayers, Jentery 357

Intra-linking the Research Corpus: Using Semantic MediaWiki as a lightweight Virtual Research Environment
Schindler, Christoph; Ell, Basil; Rittberger, Marc 359

Corpus Coranicum: A digital landscape for the study of the Qu'ran
Schnöpf, Markus 362

The MayaArch3D Project: A 3D GIS Web System for Querying Ancient Architecture and Landscapes
Schwerin, Jennifer von; Richards-Rissetto, Heather; Agugiaro, Giorgio; Remondino, Fabio; Girardi, Gabrio 365

Multi-dimensional audio-visual technology: Evidence from the endangered language documentation
Sharma, Narayan P. 368

Contours of the Past: Computationally Exploring Civil Rights Histories
Shaw, Ryan Benjamin 370

Notes from the Collaboratory: An Informal Study of an Academic DH Lab in Transition
Siemens, Lynne; Siemens, Raymond 373

XML-Print: an Ergonomic Typesetting System for Complex Text Structures
Sievers, Martin; Burch, Thomas; Küster, Marc W.; Moulin, Claudine; Rapp, Andrea; Schwarz, Roland; Gan, Yu 375

Federated Digital Archives and Disaster Recovery: The Role of the Digital Humanities in Post-earthquake Christchurch
Smithies, James Dakin 380

Modeling Medieval Handwriting: A New Approach to Digital Palaeography
Stokes, Peter 382

A Digital Geography of Hispanic Baroque Art
Suárez, Juan-Luis; Sancho-Caparrini, Fernando 385

Approaching Dickens' Style through Random Forests
Tabata, Tomoji 388

Interfacing Diachrony: Visualizing Linguistic Change on the Basis of Digital Editions of Serbian 18th-Century Texts
Tasovac, Toma; Ermolaev, Natalia 392

Promise and Practice of Enhanced Publications to Complement Conventionally-Published Scholarly Monographs
Tatum, Clifford; Jankowski, Nicholas; Scharnhorst, Andrea 394

Culpeper's legacy: How title pages sold books in the 17th century
Tyrkkö, Jukka Jyrki Juhani; Suhr, Carla Maria; Marttila, Ville 396

The Differentiation of Genres in Eighteenth- and Nineteenth-Century English Literature
Underwood, Ted; Sellers, Jordan; Auvil, Loretta; Capitanu, Boris 397

Digital editions with eLaborate: from practice to theory
van Dalen-Oskam, Karina; van Zundert, Joris Job 400

Delta in 3D: Copyists Distinction by Scaling Burrows's Delta
van Zundert, Joris Job; van Dalen-Oskam, Karina 402

Wiki Technologies for Semantic Publication of Old Russian Charters
Varfolomeyev, Aleksey; Ivanovs, Aleksandrs 405

L'histoire de l'art à l'ère numérique – Pour une historiographie médiologique
Welger-Barboza, Corinne 407

Benefits of tools and applications for a digitized analysis of Chinese Buddhist inscriptions
Wenzel, Claudia 411

The ARTeFACT Movement Thesaurus: toward an open-source tool to mine movement-derived data
Wiesner, Susan L.; Bennett, Bradford; Stalnaker, Rommie L. 413

The electronic 'Oxford English Dictionary', poetry, and intertextuality
Williams, David-Antoine 415

Reasoning about Genesis or The Mechanical Philologist
Wissenbach, Moritz; Pravida, Dietmar; Middell, Gregor 418

The Digital Daozang Jiyao – How to get the edition into the Scholar's labs
Wittern, Christian 422

Posters

A Digital Approach to Sound Symbolism in English: Evidence from the Historical Thesaurus
Alexander, Marc; Kay, Christian 427

Collaborative Video and Image Annotation
Arnold, Matthias; Knab, Cornelia; Decker, Eric 429

Le Système modulaire de gestion de l'information historique (SyMoGIH): une plateforme collaborative et cumulative de stockage et d'exploitation de l'information géo-historique
Beretta, Francesco; Vernus, Pierre; Hours, Bernard 431

Realigning Digital Humanities Training: The Praxis Program at the Scholars' Lab
Boggs, Jeremy; Nowviskie, Bethany; Gil, Alexander; Johnson, Eric; Lestock, Brooke; Storti, Sarah; Swafford, Joanna; Praxis Program Collaborators 433

Supporting the emerging community of MEI: the current landscape of tools for note entry and digital editing
Bohl, Benjamin W.; Röwenstrunk, Daniel; Viglianti, Raffaele 435

'The Past Is Never Dead. It's Not Even Past': The Challenge of Data Provenance in the e-Humanities
Clark, Ashley M.; Holloway, Steven W. 438

The Social Edition: Scholarly Editing Across Communities
Crompton, Constance; Siemens, Raymond; The Devonshire MS Editorial Group 441

Courting 'The World's Wife': Original Digital Humanities Research in the Undergraduate Classroom
Croxall, Brian 443

The Academy's Digital Store of Knowledge
Czmiel, Alexander; Jürgens, Marco 445

Building a TEI Archiving, Publishing, and Access Service: The TAPAS Project
Flanders, Julia; Hamlin, Scott; Alvarado, Rafael; Mylonas, Elli 448

Author Consolidation across European National Bibliographies
Freire, Nuno 450

Historical Events Versus Information Contents – A Preliminary Analysis of the National Geographic Magazine
Fujimoto, Yu 453

'Tejiendo la Red HD' – A case study of building a DH network in Mexico
Galina, Isabel; Priani, Ernesto; López, José; Rivera, Eduardo; Cruz, Alejandro 456

Adaptive Automatic Gesture Stroke Detection
Gebre, Binyam Gebrekidan; Wittenburg, Peter 458

Towards a Transnational Multilingual Caribbean Digital Humanities Lab
Gil, Alexander 462

NUScholar: Digital Methods for Educating New Humanities Scholars
Graff, Ann-Barbara; Lucas, Kristin; Blustein, James; Gibson, Robin; Woods, Sharon 463

Latent Semantic Analysis Tools Available for All Digital Humanities Projects in Project Bamboo
Hooper, Wallace Edd; Cowan, Will; Jiao, David; Walsh, John A. 465

Machine Learning for Automatic Annotation of References in DH scholarly papers
Kim, Young-Min; Bellot, Patrice; Faath, Elodie; Dacos, Marin 467

An Ontology-Based Iterative Text Processing Strategy for Detecting and Recognizing Characters in Folktales
Koleva, Nikolina; Declerck, Thierry; Krieger, Hans-Ulrich 470

Integrated multilingual access to diverse Japanese humanities digital archives by dynamically linking data
Kuyama, Takeo; Batjargal, Biligsaikhan; Kimura, Fuminori; Maeda, Akira 473

Linguistic concepts described with Media Query Language for automated annotation
Lenkiewicz, Anna; Lis, Magdalena; Lenkiewicz, Przemyslaw 477

Virtual Reproduction of Gion Festival Yamahoko Parade
Li, Liang; Choi, Woong; Nishiura, Takanobu; Yano, Keiji; Hachimura, Kozaburo 480

Complex entity management through EATS: the case of the Gascon Rolls Project
Litta Modignani Picozzi, Eleonora; Norrish, Jamie; Monteiro Vieira, Jose Miguel 483

TextGrid Repository – Supporting the Data Curation Needs of Humanities Researchers
Lohmeier, Felix; Veentjer, Ubbo; Smith, Kathleen M.; Söring, Sibylle 486

RIgeo.net – A Lab for Spatial Exploration of Historical Data
Loos, Lukas; Zipf, Alexander 488

Automatic Topic Hierarchy Generation Using Wordnet
Monteiro Vieira, Jose Miguel; Brey, Gerhard † 491

Hypotheses.org, une infrastructure pour les Digital Humanities
Muscinesi, Frédérique 494

TXSTEP – an integrated XML-based scripting language for scholarly text data processing
Ott, Wilhelm; Ott, Tobias; Gasperlin, Oliver 497

Exploring Prosopographical Resources Through Novel Tools and Visualizations: a Preliminary Investigation
Pasin, Michele 499

Heterogeneity and Multilingualism vs. Usability – Challenges of the Database User Interface 'Archiv-Editor'
Plutte, Christoph 502

Medievalists' Use of Digital Resources, 2002 and 2012
Porter, Dot 505

Cross-cultural Approaches to Digital Humanities – Funding and Implementation
Rhody, Jason; Kümmel, Christoph; Effinger, Maria; Freedman, Richard; Magier, David; Förtsch, Reinhard 506

CWRC-Writer: An In-Browser XML Editor
Rockwell, Geoffrey; Brown, Susan; Chartrand, James; Hesemeier, Susan 508

The Musici Database
Roeder, Torsten; Plutte, Christoph 511

The TEICHI Framework: Bringing TEI Lite to Drupal
Schöch, Christof; Achler, Stefan 514

What Has Digital Curation Got to Do With Digital Humanities?
Schreibman, Susan; McCadden, Katiet Theresa; Coyle, Barry 516

Orbis Latinus Online (OLO)
Schultes, Kilian Peter; Geissler, Stefan 518

Semantically connecting text fragments – Text-Text-Link-Editor
Selig, Thomas; Küster, Marc W.; Conner, Eric Sean 520

The Melesina Trench Project: Markup Vocabularies, Poetics, and Undergraduate Pedagogy
Singer, Kate 522

Digital Edition of Carl Maria von Weber's Collected Works
Stadler, Peter 525

Data sharing, virtual collaboration, and textual analysis: Working on 'Women Writers In History'
van Dijk, Suzan; Hoogenboom, Hilde; Sanz, Amelia; Bergenmar, Jenny; Olsson, Leif-Jöran 527

Storage Infrastructure of the Virtual Scriptorium St. Matthias
Vanscheidt, Philipp; Rapp, Andrea; Tonne, Danah 529

Digital Emblematics – Enabling Humanities Research of a Popular Early Modern Genre
Wade, Mara R.; Stäcker, Thomas; Stein, Regine; Brandhorst, Hans; Graham, David 532

DTAQ – Quality Assurance in a Large Corpus of Historical Texts
Wiegand, Frank 535

The Digital Averroes Research Environment – Semantic Relations in the Editorial Sciences
Willems, Florian; Gärtner, Mattias 537

A flexible model for the collaborative annotation of digitized literary works

Gayoso-Cabada, Joaquin
gayoxo@gmail.com
Universidad Complutense de Madrid, Spain

Ruiz, Cesar
cruiz85@gmail.com
Universidad Complutense de Madrid, Spain

Pablo-Nuñez, Luis
lpnunez@filol.ucm.es
Universidad Complutense de Madrid, Spain

Sarasa-Cabezuelo, Antonio
asarasa@fdi.ucm.es
Universidad Complutense de Madrid, Spain

Goicoechea-de-Jorge, Maria
mgoico@filol.ucm.es
Universidad Complutense de Madrid, Spain

Sanz-Cabrerizo, Amelia
amsanz@filol.ucm.es
Universidad Complutense de Madrid, Spain

Sierra-Rodriguez, Jose-Luis
jlsierra@fdi.ucm.es
Universidad Complutense de Madrid, Spain

1. Introduction

The Complutense University has been one of the first European universities that has collaborated with Google's project[1] by putting on the Web 100,000 volumes from its ancient fund. However scholars notice that these digitized texts are often of no much use to professors-researchers-students in literature unless additional tools are provided, to enhance the educational and research value of this material. In particular, the ability of making annotations on these texts has been largely recognized as a basic mean of adding value to this kind of digitized resources (Rios da Rocha et al. 2009). In this paper we present the annotation model used in @Note 1.0, a system developed at UCM funded by the Google's 2010 Digital Humanities Award program.

@Note 1.0 allows us to retrieve digitized works from Google Books collection and add annotations to enrich the texts with research and learning purposes: critical editions, reading activities, e-learning tasks, etc. One of the main features of @Note annotation model, which distinguishes it from similar approaches (Azouaou & Desmoulins 2006; Bechhofer at al. 2002; Koivunen 2005; Rios da Rocha et al. 2009; Schroeter et al. 2006; Tazi et al. 2003), is to promote the collaborative creation of annotation schemas by communities of researchers, teachers and students, and the use of these schemas in the definition of annotation activities on literary works. It results in a very flexible and adaptive model, able to be used by many different communities of experts in literature defending different critical literary theories and for different annotation tasks. In this paper we present this annotation model.

2. The @Note Annotation Model

2.1. Structure of the model

The structure of the @Note annotation model is summarized in the UML class diagram (Booch et al. 2005) of Fig. 1. In this model:

- *Annotation management communities* are groups of *annotation managers*, experts in literature (teachers, researchers, etc) who act as administrators to create activities, to select works and to organize activity groups.
- *Annotation communities*, in their turn, are groups of *annotators*, students / pupils interested in literature who perform proposed annotation activities.
- Each a*nnotation activity* comprises (i) a *digitized work*, (ii) a *metalevel-oriented annotation schema*, (iii) a *work-oriented annotation schema*.
- In this context, the *works* are the literary texts that can be annotated during the annotation activities. *Annotations*, in their turn, are characterized by: (i) an *annotation anchor* (the region of the work to which the annotation refers), (ii) an *annotation content* (a free rich-text piece that actually configure the annotation), (iii) a set of *annotation types* (semantic qualifiers for annotations) chosen from the annotation schemas attached to the annotation activity (at least one from the metalevel-oriented annotation schema).

Figure 1: @Note information model

- The *annotation schemas* are explicit formalization of the types of annotations that can be carried out on works. In @Note, annotation schemas are hierarchies formed by annotation types and *annotation categories* (sets of annotation types and/or others, more specifically, annotation categories). In their turn, they can be *metalevel-oriented annotation schemas* (schemas which usually comprise concepts concerning particular literary theories around which the annotation activities are articulated), or *work-oriented annotation schemas* (schemas that capture aspects relative to the relationships between annotations and their anchors). While schemas of the first type are created by annotation managers, schemas of the second type are usually created by annotators.

In the context of the annotation management community, an annotation schema can be *public* or *private*. A private schema is only accessible for the annotation manager who created it. On the contrary, a public schema is accessible for all the annotation managers. Annotation managers have unlimited privileges on all the schemas to which they can access (i.e., they can create new annotation types and categories, they can blend two different types/categories in a single one, they have renaming and erasing privileges, etc), with the exception of modifying the *public/private* character (it only can be done by the schema's creator). In addition, when annotation managers create annotation activities, they only can choose those schemas to which they have access grants. Concerning annotators, they can add new types and categories to the book-oriented annotation schema, but they can't perform any other modification.

Figure 2: Example of rules governing the annotation process (informally described using natural language)

2.2. The annotation process

The @Note annotation process governs how to create the different types of information elements envisioned in the annotation model. For this purpose, @Note introduces a set of rules governing aspects like information visibility, creation and modification privileges of annotations and annotation schemas, etc. Although, by lack of space, these rules will not be detailed here, in Fig. 2 we include an example concerning an informal description of some of the rules governing the management of annotation schemas.

2.3. Annotation browsing and recovering

Annotation schemas in @Note are seen as T-*boxes* of description logic theories (Brachman & Levesque 2004). For instance, Fig 3a shows, edited in @Note, a fragment of the annotation schema used at UCM in an English Literature introductory course, while Fig 3b despites the description logic's counterpart. This simple interpretation is still powerful-enough to enable powerful *annotation browsing* and *annotation recovering* behavior. Indeed:

Figure 3: (a) A fragment of annotation schema (b) description logic counterpart

- Annotations can be browsed using annotation schemas, in a similar way to a folder explorer in a filesystem. In each step, there is a category or a type selected, and the user will *see* all the annotations entailed by such a selected element.
- Queries consist of arbitrary Boolean formulae involving annotation types and categories, being the outcomes the annotations entailed by such formulae.

In both cases, since entailment will be performed according to the description logic principles, the process will take into account the *is-a* relationship made explicit by the annotation schema.

2.4. Some technical details

The system has been entirely developed using Google technologies for the development of Rich-Internet Applications (RIAs) (Fraternali et al. 2010): GWT in the client side and the Google App Engine's facilities in the server side (Unruh 2010). Fig. 4 shows some snapshots of the system. The current version runs on the fully free-access books integrated in Google Books, and, in particular, on the UCM-Google collection. The works retrieval is achieved by the use of the Google Books API through REST (Richardson & Ruby 2007), and then presented to clients in an asynchronous way to keep them responsive to their events.

Figure 4: Some snapshots of @Note

3. Conclusions and Future Works

@Note promotes a fully collaborative annotation process, in which not only literary works are collaboratively annotated, but also annotation schemas are collaboratively created. The @Note system has been evaluated at UCM by several researchers and students in literature. They highlighted the flexibility of the annotation model, and, in particular, the ability to create and share annotation schemas tailored according to different critical perspectives and annotation activities. Additionally, they appreciate a sufficient expressive power from a browsing and recovering point of view. They also remarked the educational potential of the tool, although some advanced features could add some conceptual difficulties for students.

Currently we are working to adapt the annotation tool in order to facilitate its connection to a repository of learning objects, so as to allow the storage of literary texts' annotations as learning objects (Polsani 2003), and to make possible the recovery of those annotations and move about them according to the associated metadata. Thus, we are developing a communal working space for the creation of written compositions in different traditions and languages. We are also experimenting with the students' capability for developing their own catalogues, annotating the literary texts according to them and reusing their annotations in the production of critical essays. Additionally, we are working on connecting our system with other digital libraries (in particular, with Hathi Trust[2]). Finally, we are planning to address interoperability issues, in order to enable the interchange of annotations according to some of the emerging standards proposed by the digital humanities community (e.g., OAC[3]).

Acknowledgements

This work has been funded by Google with a grant of the Google's 2010 Digital Humanities award program entitled *Collaborative annotation of digitalized literary texts*. Additionally, this work has been performed in the context of the project grants of the Spanish Ministry for Research and Innovation (FF12008-06924-C02-01 and TIN2010-21288-C02-01), UCM (PIMCDs 2010/177 and 2011/313) and Santander-UCM (GR 42/10 - 962022).

References

Azouaou, F., and C. Desmoulins (2006). MemoNote, a context-aware annotation tool for teachers. *Proceedings of the 7th International Conference on Information Technology Based Higher Education and Training* ITHET'06, Sidney, Australia, July 2006.

Bechhofer, S., L. Carr, C. Goble, S. Kampa, and T. Miles-Board (2002). The Semantics of Semantic Annotation. *Proceedings of the First International Conference on Ontologies, DataBases, and Applications of Semantics for Large Scale Information Systems* ODBASE'02, Irvine, CA, USA, October 2002.

Booch, G., J. Rumbaugh, and I. Jacobson (2005). *The Unified Modeling Language User Guide* (*2nd Edition*). Boston: Adisson-Wesley.

Brachman, R., and H. Levesque (2004). *Knowledge Representation and Reasoning*. Amsterdam: Morgan-Kauffmann.

Fraternali, P., R. Gustavo, and F. Sánchez-Figueroa (2010). Rich Internet Applications. IEEE *Internet Computing* 14(3): 9-12.

Koivunen, M.-R. (2005). Annotea and Semantic Web Supported Collaboration. *Proceedings of the UserSWeb Workshop – 2nd European Semantic Web Conference*, Heraklion, Grece, June 2005.

Polsani, P. (2003). Use and abuse of reusable learning objects. *Journal of Digital Information* 3(4).

Richardson, L., and S. Ruby (2007). *Restful web services*. Beijing: O'Reilly.

Rios da Rocha, T., R. Willrich, R. Fileto, and S. Tazi (2009). Supporting Collaborative Learning Activities with a Digital Library and Annotations. *Proceedings of the 9th IFIP World Conference on Computers in Education* WCCE'09, Bento Gonçalves, Brazil, July 2009.

Schroeter, R., J. Hunter, J. Guerin, I. Khan, I., and M. A. Henderson (2006). Synchronous Multimedia Annotation System for

Secure Collaboratories. *Proceedings of the Second IEEE International Conference on e-Science and Grid Computing* e-Science '06, Amsterdam, Netherlands, December 2006.

Tazi, S., Y. Al-Tawki, and K. Drira K. (2003). Editing pedagogical intentions for document reuse. *Proceedings of the 4th IEEE Technology Based Higher Education and Training* ITHET'03, Marrakesh, Morocco, July 7-9, 2003.

Unruh, A. (2010). *Google App Engine Java and GWT Application Development*. Birmingham: Packt Publishing.

Notes

1. http://www.ucm.es/BUCM/atencion/25403.php
2. http://www.hathitrust.org/
3. http://www.openannotation.org/

HyperMachiavel: a translation comparison tool

Gedzelman, Séverine
severine.gedzelman@ens-lyon.fr
ENS de Lyon, France

Zancarini, Jean-Claude
jean-claude.zancarini@ens-lyon.fr
ENS de Lyon, France

1. Introduction

The HyperMachiavel project started with the idea of a tool that would aid research communities comparing several editions of one text and in particular comparing translations.

The Italian studies department (Triangle laboratory) at ENS de Lyon has been working for many years on fundamental texts, from Machiavelli, Guicciardini and other contemporary followers, that put forward new political concepts throughout Europe in the 16th century. The question addressed in the project was mainly about the transfer of these concepts from one language to another, and especially their reception in France. The first aligned corpora tested in our tool gathers different editions of Machiavelli's *Il Principe*, the *princeps edito* from Blado in 1532 and the first four French translations of the 16th century.

Inspired by machine translation and lexicographic domains, the system presented in this paper proposes an annotation environment dedicated to the edition of lexical correspondences and offers different views to assist humanities researchers in their interpretations of the quality and the specificities of translator's work.

2. Viewing and Searching in Aligned Corpora

2.1. Synoptic View

To be able to identify lexical correspondences, machine translation tools usually propose a frame of two panels, one for the source text and the other for the target text. The visualized interface is meant for annotators to easily revise the results obtained from automatic word alignment. In general it only considers a pair of texts at a time.

In the world of digital editions, text comparison has always been of great interest and the request to view diplomatic vs normalized transcriptions, or simply