

Digitale Paläographie

Veranstalter: Lehrstuhl für Computerphilologie, Universität Würzburg

Datum, Ort: 20.07.2011–22.07.2011, Würzburg

Bericht von: Simone Haag, Würzburg

Der internationale Workshop „Digital Palaeography“, der an der Julius-Maximilians-Universität Würzburg mit finanzieller Unterstützung der European Science Foundation (ESF) durchgeführt wurde, beschäftigte sich mit zentralen Fragestellungen und Herausforderungen der Paläographie im digitalen Zeitalter¹. Auf Einladung von Malte Rehbein (Würzburg) fanden sich 24 Wissenschaftler/innen verschiedener Disziplinen aus Europa, Nordamerika und Israel für drei Tage in Würzburg ein, um ihre aktuellen Projekte vorzustellen und neue Entwicklungen der Disziplin zu diskutieren. Mit dem Workshop wurde an die 2009 auf der Münchner Fachtagung „Kodikologie und Paläographie im digitalen Zeitalter“ formulierte Agenda angeknüpft, die die enge Zusammenarbeit und regelmäßige Treffen der wissenschaftlichen Gemeinschaft vorsieht.

Die Tagung eröffnete EEF OVERGAAUW (Berlin), der angesichts des wissenschaftlichen Nachwuchses und Interesses an noch nicht geklärten paläographischen Fragen keine Anzeichen für eine ‚Krise‘ dieser Disziplin sah. Er konstatierte im Gegenteil eine positive Weiterentwicklung der Paläographie in den letzten 50 Jahren, zu der auch die neuen Technologien beigetragen hätten; etwa auf dem Gebiet der quantitativen Kodikologie. Mit Hilfe der neuen Technologien könnten zwar seiner Ansicht nach keine neuen Fragestellungen entworfen werden, sie dienten aber dazu, befriedigende Antworten auf bekannte Fragen und Probleme zu finden, wie zum Beispiel dem Wunsch nach verlässlichen Kriterien für die Datierung und Lokalisierung eines Eintrages. Die neuen Möglichkeiten in der Paläographie (statistische Paläographie, digitale Bildanalyse und die Bildung von großen Datenbanken) hätten auch eine Veränderung der Forschungspraxis zur Folge. Dennoch sei nach wie vor die Gelehrtheit (*scholarship*), verstanden als wissenschaftlich fun-

dierte und durch Erfahrung ergänzte Ausbildung, und vor allem das Talent (*connoisseurship*) von zentraler Bedeutung, die zwar durch neue Bildanalysetechniken geübt und ergänzt, nicht aber ersetzt werden könnten. Obwohl Overgaauw den erleichterten Zugang zu digital verfügbar gemachten Manuskripten grundsätzlich begrüßte, erinnerte er eindringlich an die Pflicht eines jeden Forschers, nicht nur die neuen Medien und Methoden zu nutzen, sondern auch die Manuskripte selbst in die Hand zu nehmen, da zum einen das Original ein besseres Verständnis ermöglichte und zum anderen noch nicht jedes Manuskript digitalisiert und Gegenstand der Forschung sei.

In den thematischen Sitzungen stellten die Teilnehmer/innen daraufhin ihre Projekte vor, die Anlass zu Diskussionen über konkrete Probleme aus der Praxis sowie eine Weiterentwicklung der von den Teilnehmer/innen genutzten innovativen Methoden gaben.

Vier Präsentationen beschäftigten sich mit dem Umgang mit Texten und Bildern im digitalen Zeitalter. WENDY SCASE (Birmingham) stellte in ihrem Vortrag „New Methodologies for effective exploitation of Digital Manuscripts Corpora“ zwei ihrer Projekte vor, die sich mit Manuskripten aus den West Midlands und dem umfangreichen ‚Vernon Manuscript‘ befassen. Scase versucht durch die Digitalisierung dieses Werkes dieses einer breiteren Masse zugänglich zu machen, um dadurch weitere Forschungen zu initiieren. Die Bereitstellung von unbearbeiteten XML-Dateien im Netz der Universität Birmingham soll es Forschern zukünftig ermöglichen, durch allographische Analyse neue Datensätze zu erzeugen beziehungsweise Datensätze neu zu bestimmen und diese weiter zu entwickeln. Als langfristiges Ziel solle die Vernetzung zwischen abrufbaren und digitalisierten Ressourcen ermöglicht werden, allerdings stehe dem vorerst noch das Problem der Lizenzierung beziehungsweise Wiederverwendung des Datenmaterials im Wege.

THORSTEN SCHAßAN (Wolfenbüttel) griff das Thema „OCR for Manuscripts and early prints“ auf und stellte die Ergebnisse

¹ Konferenz-Website: http://www.zde.uni-wuerzburg.de/veranstaltungen/digital_palaeography/

der vom Deutschen Bibliotheksverband initiierten Arbeitsgruppe zur Optical Character Recognition (OCR) vor. Schaßan hob dabei zwei Programme, Abby Finereader und BIT Alpha, hervor und fasste deren Vor- und Nachteile in der Anwendung zusammen. Problematisch bleiben bei der Arbeit mit beiden Programmen die Fehlerbestimmung und die dafür nötige Parametrierung, sowie die Messung der Qualität der OCR-Analyse. Um dieses Verfahren zitierbar zu machen, müssen sowohl die Dokumentation des Entstehungsprozesses und die induzierten Unsicherheiten, die Beherrschung der Granularität und das Nebeneinander von Zitierungen, vor allem für die Ergebnisse der OCR in ALTO (Analyzed Layout and Text Object), einem standardisierten XML-Format zur Speicherung von Informationen zum Inhalt und Layout, berücksichtigt werden.

Im dritten Beitrag präsentierten STEWART BROOKES und PETER STOKES (London) das vom European Research Council finanzierte Projekt „Digital Resource for Palaeography, Manuscripts and Diplomatic“. Ihr Ziel ist es, eine Ontologie des lateinischen Alphabets und seiner paläographisch exakten Beschreibung, basierend auf angelsächsischen Manuskripten des 11. Jahrhunderts, aufzubauen und dieses Datenmaterial zum Studium der Schrift digital zur Verfügung zu stellen. Durch Such- und Sortierfunktionen sowie durch die Möglichkeit, Ähnlichkeiten verschiedener Exemplare innerhalb dieser Ontologie aufzufinden, wird das Forschen an Manuskripten dadurch erleichtert.

Der Vortrag „Spatial Exploration tools in the Graphem Project“ von MATTHIEU EXBRAYAT (Orléans) beschäftigte sich mit dem graphemischen Raum und den Möglichkeiten der automatisierten Klassifikation von mittelalterlichen Schriften. Das GRAPHEM-Projekt hat dabei zum Ziel, ‚character recognition‘, also die automatisierte Texterkennung innerhalb von Texten und Bildern, zu ermöglichen. Die Handhabung der Datenbank werde durch die Suchfunktion sowie der Möglichkeit, eigene Vergleichsbeispiele mit den in der Datenbank hinterlegten Schriften online zu vergleichen, erleichtert. Er demonstrierte daraufhin das Programm, wodurch deutlich wurde, dass die Visualisierung von Ergebnissen an-

gesichts der Komplexität und Masse an Daten ein hilfreiches heuristisches Instrument ist, um zu guten Ergebnissen zu kommen.

Zwei Beiträge aus dem Bereich der Kodikologie befassten sich mit dem Thema Pergament. PAOLA ERRANI (Cesena) stellte in ihrem Vortrag „Parchment and Scribes in the Malatestian Scriptorium“ die Methodik und die Ergebnisse ihrer Untersuchungen an 343 Manuskripten aus Ziegenpergament aus dem 15. Jahrhundert vor, die aus der Bibliothek der Familie Malatesta stammen. Nach fast 8.000 Messungen an je sechs Stellen eines Bifolios konnte sie feststellen, dass die Homogenität mit Zunahme der Größe des Manuskripts abnimmt und dass die für Malatesta hergestellten Pergamentseiten in der Regel dünner und weniger homogen sind, als die in der gleichen Zeit von den gleichen Schreibern für andere Auftraggeber hergestellten. Dabei liefern die Bereitstellung von vergleichbaren Messprotokollen und von unbearbeiteten Informationen in dauerhaften Datenbanken neue Ansatzpunkte für weitere Forschungen.

TIMOTHY STINSON (Raleigh) eröffnete in seinem Beitrag „DNA-Analysis and the Study of Medieval Parchment Books“ neue Perspektiven im Umgang mit Pergamenten. Die Analyse der mitochondrialen, die auch nach Jahrhunderten noch aus kleinsten Proben mittelalterlicher Pergamente gewonnen werden kann, könne dazu beitragen, Pergamente besser zu lokalisieren und zu datieren. Hierdurch könnten einzelne Häute bestimmten räumlich lokalisierbaren Herden zugewiesen und der mit ihnen betriebene überregionale Handel untersucht werden. Zwar seien die Daten dafür bis jetzt noch zu lückenhaft, sie ermöglichen aber jetzt schon, die materielle Struktur der Manuskripte zu untersuchen und eine mögliche Zugehörigkeit von Fragmenten zu einer gleichen Haut festzustellen.

Die Wiederherstellung von beschädigten Dokumenten war Gegenstand zweier weiterer Beiträge. LIOR WOLF (Tel-Aviv) zeigte in seinem Vortrag „Identifying Join Candidates in the Cairo Genizah“ die Problematik im Umgang mit bruchstückhaften Handschriften auf. Die Fragmente der Kairoer Genizah liegen darüber hinaus an unterschiedlichen Orten. Ziel des Projektes sei es daher, herauszufinden, welche Fragmente da-

von zusammenpassen, und in einem weiteren Schritt, diese wieder virtuell zusammenzufügen. Auf der Grundlage von bereits zugeordneten Fragmenten werde die Suchmaschine durch entsprechende Parametrisierung trainiert und zur Identifikation von anderen Bruchstücken herangezogen. Des Weiteren werden graphische Zeichen in graphemische Gruppen aufgeteilt, wodurch der Vergleich von verschiedenen Merkmalen ermöglicht werde. Hierdurch werde die automatisierte Klassifikation, die sich stark an den traditionellen Kriterien anlehnt, erleichtert.

Der zweite Beitrag „Investigation of Historic Documents with Focus on Automatic Layout and Character“ von MELANIE GAU und ROBERT SABLATNIG (Wien) befasste sich mit der Dokumentation, Analyse und Bearbeitung glagolitischer Handschriften aus dem 11. bis 13. Jahrhundert. Diese Handschriften wurden 1975 in einem Kloster auf dem Berg Sinai entdeckt; sie sind aus linguistischer und kultureller Sicht von Bedeutung. Aufgrund des schlechten Erhaltungszustands der Manuskripte, ihres Palimpsest-Charakters und der chemischen Umwandlung der Tinte stellt ihre Erforschung eine besondere Herausforderung dar. Jede Seite wurde daher zwölfmal in verschiedenen Lichtspektren fotografiert. Durch die digitale Löschung von hinzugefügten Texten und der Rekonstruktion von versteckten Merkmalen können Palimpsest-Texte so rekonstruiert werden. Dafür musste zunächst die Beschreibung der einzelnen glagolitischen Zeichen formalisiert werden, um es dem Computer zu ermöglichen, eine Transkription der Texte, in denen Buchstaben noch fragmentarisch sichtbar sind, vorzuschlagen. Da die ersten Ergebnisse vielversprechend seien, werde an der Weiterentwicklung des Programms gearbeitet.

Der Beitrag von SÉGOÛÈNE TARTE (Oxford) betrachtete die Arbeit mit den neuen Technologien von einer reflexiv-philosophischen Ebene aus. In ihrer Präsentation „Interpreting Ancient Documents: Of Avatars, Uncertainty, and Knowledge Creation“ hob sie die Bedeutung der Dokumentation von Arbeits- und Vorgehensweisen hervor. Bei der Erstellung digitaler Editionen werde auch immer die Realität durch interpretative Verfahren verändert. Daher sei das

digitale Objekt keine einfache Reproduktion, sondern eine zweidimensionale Interpretation des ursprünglich dreidimensionalen Ausgangsobjektes. Die Lektüre selbst sei ein doppelter Wahrnehmungsprozess, nämlich der Form und der Bedeutung. Die komplexen Prozesse, die zu einer Entscheidung führen, müssten dadurch durch Protokolle und Dokumentation nachvollziehbar sein, sodass im Falle von Neubewertungen der Ereignisse anhand von Entscheidungsbäumen eine Änderung des Modells möglich sei. Tarte sprach sich dabei für eine angemessene und selbst-reflexive, in Datenmodellen formalisierte Subjektivität aus.

Den Abschluss der Tagung bildete eine lebhafte Diskussion über die Probleme und Perspektiven der Digitalen Paläographie. Für die Teilnehmer/innen war klar, dass die schon vorhandenen Standards wie TEI und ALTO pragmatisch umgesetzt werden müssen, indem die Werkzeuge und Methoden, um Daten einfach und konsistent aufzuzeichnen und deren Ergebnisse zu bearbeiten, definiert und zugänglich gemacht werden. Vor allem für die Erscheinungsformen einzelner Schriftelemente („characters“) wurden allgemeine beschreibende Standards und die Entwicklung einer Ontologie gewünscht. Die Verwendung gemeinsamer Standards und Werkzeugen würde im Rahmen der Best Practice die Forschung erleichtern. Auch für die Werkzeuge gilt, dass, wie von Nataša Golob (Ljubljana) angeregt, ein hohes Maß an Benutzerfreundlichkeit und Zugänglichkeit der Daten beziehungsweise der Datenbanken erstrebenswert sei. Digitale Methoden müssen weiter vertieft werden, um durch heuristische Verfahren zu validierbaren Ergebnissen kommen zu können. Sobald diese Verfahren und Methoden etabliert seien, könnten auch die seit langem offenen paläographischen Fragen erneut behandelt und auf gemeinsamen Plattformen diskutiert werden.

Dass die Tagung mit dem gemeinsamen Wunsch nach enger Zusammenarbeit und dem weiteren Aufbau der wissenschaftlichen Forschungsinfrastruktur, wie NEDIMAH² und DARIAH³ endete, zeigt das Bedürfnis nach weiteren ähnlichen Veranstal-

²<http://www.nedimah.eu>

³<http://www.dariah.eu>

tungen und das Potential, das die Digitale Paläographie für die Forschung birgt.

Konferenzübersicht:

Malte Rehbein (Universität Würzburg): Begrüßung

Keynote

Eef Overgaauw (Staatsbibliothek, Berlin): Palaeography today. Old questions and new answers

Claudine Moulin (Universität Trier): Presentation of the European Science Foundation (ESF)

Enhancing Palaeography

Stewart Brookes (King's College, London): Digital Resource for Palaeography, Manuscripts and Diplomatic

Wendy Scase (University of Birmingham, Birmingham): New Methodologies for effective exploitation of Digital Manuscript Corpora

Paola Errani (Biblioteca Malatestina, Cesana): Parchment and Scribes in the Malatestian Scriptorium

Crossing the Disciplines

Moderation: Peter Stokes (King's College, London)

Timothy Stinson (North Carolina State University, Raleigh): DNA Analysis and the Study of Medieval Parchment Books

Thorsten Schaßan: OCR for manuscripts and early prints

Lior Wolf (Tel-Aviv University, Tel-Aviv): Identifying Join Candidates in the Cairo Genizah

Malte Rehbein (Würzburg): Case Study

The Digital World

Matthieu Exbrayat (Université d'Orléans, Orléans): Spatial exploration tools in the GRAPH-EM Project

Ségolène Tarte (University of Oxford, Oxford): Interpreting Ancient Documents: Of Avatars, Uncertainty, and Knowledge Creation

Melanie Gau, Robert Sablatnig (TU Wien): Investigation of Historic Documents with Focus

on Automatic Layout and Character Analysis

Conclusion

Perspectives

Moderation: Malte Rehbein (Würzburg)

Tagungsbericht *Digitale Paläographie*. 20.07.2011–22.07.2011, Würzburg, in: H-Soz-u-Kult 21.12.2011.