

„Das Projekt tranScriptorium ist eine tolle und spannende Sache, und wir haben viele positive Rückmeldungen dafür bekommen.“

Günter Mühlberger



2000

Laufmeter Akten und Dokumente lagern in den Räumen des Bozner Stadtarchivs. Die Ratsprotokolle der Jahre 1470 bis 1684 wurden bereits digitalisiert und online gestellt.

## Schätze aus den Archiven zutage fördern

WISSENSCHAFT: Stadtarchiv Bozen arbeitet mit einem EU-Projekt zur Umwandlung handgeschriebener historischer Quellen in digitalisierte Texte

VON HATTO SCHMIDT

**BOZEN/INNSBRUCK.** Im Bozner Stadtarchiv lagern riesige Schätze an historischen Quellen vom Mittelalter bis heute – schwer vorstellbare 2000 Laufmeter sind es. Ein Teil der Dokumente (die meisten Ratsprotokolle) ist zwar online einsehbar, der Großteil aber kann nur im Archiv konsultiert werden. Doch nur Experten können die alten Schriften entziffern. Das Stadtarchiv arbeitet nun an dem wissenschaftlichen Projekt „tranScriptorium“ mit, in dessen Rahmen versucht werden soll, handschriftliche Quellen digital zu erfassen und in editierbaren Text umzuwandeln.

Stadtarchivdirektor Hannes Obermair ist Feuer und Flamme für dieses von der EU finanzierte Projekt. Deshalb hat er sofort die Teilnahme zugesagt. Das Stadtarchiv hat einiges zu bieten, denn es ist ein Vorreiter in Sachen Digitalisierung von Handschriften: **Die im Archiv vorhandenen Ratsprotokolle aus den Jahren 1470 bis 1800 – insgesamt 30.000 Seiten – wurden bereits digitalisiert.**

### Großteil der Bozner Ratsprotokolle ist online

Online einsehbar sind seit einiger Zeit auf der Plattform BOhisto („Bozen-Bolzano's History Online“) die Protokolle von 1470 bis 1684. Sie halten polizeiliche, wissenschaftliche und kulturelle Themen fest. Bis 1900 wurden sie von Hand geschrieben.

Der Zugang zu diesen Protokollen ist frei, denn Obermair ist ein Verfechter des „open access“ (freier Zugang) gemäß der Berliner Deklaration 2003: Die Nutzer müssen nichts zahlen, sich nicht registrieren, die Dokumente tragen kein Wasserzeichen und sie können ohne Einschränkung heruntergeladen werden – einzig der Verfasser, oder in diesem Fall das Archiv müssen als Urheber bzw. Besitzer erwähnt werden.

Die Digitalisierung wurde von der Sozialgenossenschaft Kairos in Brixen durchgeführt; die Verarbeitung und Speicherung erledigt die Bozner Firma Alpin.

BOhisto ist ein Riesenerfolg. Vom Start weg erfolgten 10.000 Einzelzugriffe in nur sieben Mo-



Stadtarchivdirektor Hannes Obermair hält einen Originalband der Bozner Ratsprotokolle in der Hand; am Monitor ist eine der bereits digitalisierten und im Internet einzusehenden Seiten zu erkennen.

„Das Projekt bedeutet für das Stadtarchiv eine Aufwertung und Imagepflege und ich freue mich, dass wir an dieser Arbeit einen kleinen Anteil haben.“

Hannes Obermair

naten; 1000 Besucher griffen in diesem Zeitraum auf die online gestellten Protokolle zu.

Obermair ist stolz auf die Resonanz im Netz: „Ich hätte nie erwartet, dass zum Beispiel Leute aus den USA zugreifen und auf der Homepage auch verweilen, nicht nur einmal kurz hineinschauen.“ Das Projekt habe in der wissenschaftlichen Welt große Aufmerksamkeit und Anerkennung gefunden.

Eine ungeplante Folge von BOhisto war die Anfrage von Günter Mühlberger von der Universität Innsbruck, ob das Stadtarchiv sich nicht an dem Projekt „tranScriptorium“ beteiligen wolle.

### Anfrage vom Projekt „tranScriptorium“

Das Stadtarchiv übergab die digitalisierten Dokumente zweier

Bände seiner Ratsprotokolle samt Transkriptionen (500 transkribierte Seiten existieren bereits). Weitere „tranScriptorium“-Partner, die digitalisierte und transkribierte Dokumente für „tranScriptorium“ zur Verfügung stellen, sind die Universität London, die Spanische Nationalbibliothek, die Stadt Zwettl (Niederösterreich) mit ihren Ratsprotokollen und die Rechtshistorische Fakultät der Universität Tübingen.

„tranScriptorium“ ist auf solche Zuarbeit angewiesen, weil sonst der Aufwand viel zu groß wäre. Es werden digitalisierte

Dokumente benötigt mit größeren Abschnitten derselben Handschrift, von denen es bereits Transkriptionen gibt, denn die Ergebnisse der Digitalisierung („handwritten text recognition“, Erfassung handgeschriebener Texte) müssen anhand der von Experten besorgten Übertragungen der Handschriften überprüft werden.

### Die Texterfassung ist eine riesige Herausforderung

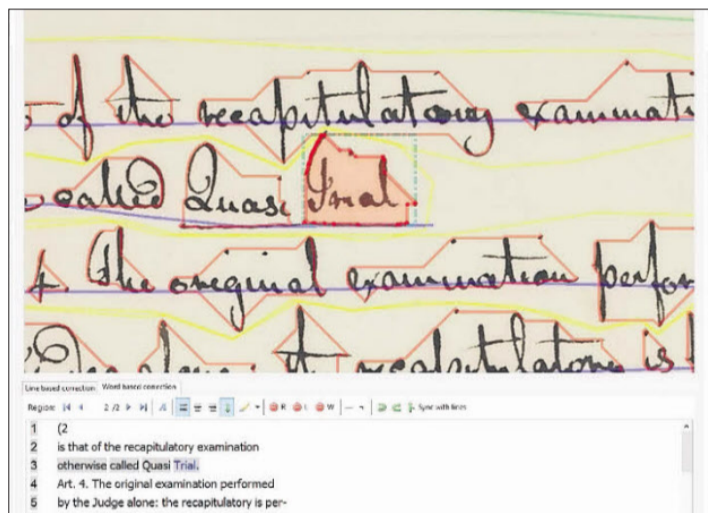
Ob es wirklich gelingen wird, die oft auch von Experten nur mit

Mühe entzifferbaren Texte zu digitalisieren, ist noch offen: **Die Handschriften sind oft verschlungen, Ober- und Unterzüge reichen mitunter über mehrere Zeilen, die Blätter weisen Flecken auf und hin und wieder sind die Textabschnitte nach nicht erkennbaren Kriterien in Blöcken über die Seiten verstreut.**

Verloren gegangene Buchstaben wie das lange -s-, Abkürzungen, regionale Schreibweisen und unbekannte Wörter stellen die Wissenschaftler vor weitere Probleme. Trotzdem konnten in ersten Testläufen mit englischen Schriften des 18. Jahrhunderts sieben bis acht von zehn Wörtern richtig erkannt werden.

### Sieben bis acht von zehn Wörtern richtig erkannt

Doch wer sich erinnert, wie ungenau noch vor zehn Jahren die Erkennung natürlicher Sprache oder die Gesichtserkennung waren und bedenkt, dass heutzutage etwa bei der Passkontrolle die eingesetzte Software inzwischen bessere Ergebnisse erzielt als der dafür zuständige Passbeamte, der wird kaum Zweifel haben, dass auch in diesem Bereich in wenigen Jahren ein Quantensprung in der Qualität erreicht werden kann.



Ein Beispiel für die Erfassung einer englischen Handschrift. Uni Innsbruck

Um diese Entwicklung voranzutreiben und sie für Archive und Geisteswissenschaftler möglichst transparent und zugänglich zu halten, arbeitet „tranScriptorium“ an einer zentralen Plattform: Die digitalisierten Handschriften werden in ein zentrales System eingespeist, dann von Geisteswissenschaftlern oder ehrenamtlichen Mitarbeitern transkribiert – und damit wird gleichzeitig Trainingsmaterial für die maschinenlernenden Verfahren geschaffen, die im Hintergrund der Plattform für die automatisierte Erkennung der Schriften sorgen.

### Maschinenlernende Verfahren

Sobald genügend Trainingsmaterial vorhanden ist, können dann auch alle 30.000 digitalisierten Seiten in einem Schwung – unter Zuhilfenahme eines Hochleistungsrechners der Universität Innsbruck – verarbeitet werden. Die Seiten stehen dann für eine Volltextsuche zur Verfügung.

„Das Projekt tranScriptorium ist eine tolle und spannende Sache, und wir haben viele positive Rückmeldungen dafür bekommen“, sagt Günter Mühlberger. Natürlich müsse man erst schauen, wie es wirklich läuft, „man darf die Erwartungen auch nicht zu hoch schrauben“. Der promovierte Germanist arbeitet seit 20 Jahren im Bereich Softwareentwicklung und ist mittlerweile vorwiegend in „digital humanities“ (digitalen Geisteswissenschaften) tätig.

Eines zu betonen ist Mühlberger wichtig: „Hannes Obermair nimmt wirklich eine vorbildliche Haltung ein“. Dass ein Archiv sich so offen zeige, sei nicht selbstverständlich. Das Bozner Stadtarchiv leiste einen wichtigen Beitrag für den Fortschritt des Projekts „tranScriptorium“, indem es die eingescannten Ratsprotokolle und die Transkriptionen zur Verfügung stelle.

Doch auch das Bozner Stadtarchiv profitiert: „Das Projekt bedeutet für das Archiv eine Aufwertung und Imagepflege“, sagt Hannes Obermair und freut sich, „dass wir an dieser Arbeit einen kleinen Anteil haben“.

© Alle Rechte vorbehalten

@ <http://stadtarchiv-archivistorico.gemeinde.bozen.it/>  
<http://transcriptorium.eu/>

**RENOVIERTEN SIE IHRE WOHNUNG MIT UNS...**

IHRER RENOVIERUNGSSPEZIALISTEN MIT

**idea CasaBagno**

ÜBER 40 JAHREN ERFAHRUNG, DIE ZU IHREN DIENSTEN STEHEN

0471 400879

BOZEN Drususstr. 51 • Meran Petrarcastr. 59

**EIN NÜTZLICHES GESCHENK**

**ALLES FÜR DIE PERFEKTE RASUR**

**... seit 1950 Ihre Fachwerkstatt für alle möglichen Schleifarbeiten**

**BOZEN - BINDERGASSE 28 - GOETHESTR. 36 - TEL. 0471/974706**

Katholisches **Sonntagsblatt**

Kirchenzeitung der Diözese Bozen-Südtirol

... jetzt auch auf **facebook**

Besuchen Sie uns im Internet unter:  
[www.facebook.com/sonntagsblatt.it](http://www.facebook.com/sonntagsblatt.it)