

Digitale Erschließung einer Sammlung von Volksliedern aus dem deutschsprachigen Raum

Manuel Burghardt (manuel.burghardt@ur.de)

Sebastian Spanner (sebastian.spanner@stud.uni-regensburg.de)

Thomas Schmidt (thomas.schmidt@stud.uni-regensburg.de)

Florian Fuchs (florian.fuchs@stud.uni-regensburg.de)

Katia Buchhop (katia.buchhop@stud.uni-regensburg.de)

Miriam Nickl (miriam.nickl@stud.uni-regensburg.de)

Christian Wolff (christian.wolff@ur.de)

Lehrstuhl Medieninformatik, Universität Regensburg,

1. Projektkontext

Dieser Beitrag beschreibt ein laufendes Projekt¹ zur digitalen Erschließung einer großen Sammlung von Volksliedern aus dem deutschsprachigen Raum, mit dem Ziel diese später über ein öffentliches Informationssystem verfügbar zu machen. Mithilfe dieses Informationssystems soll neben der üblichen Exploration gescannter Faksimiles der Originalliedblätter zusätzlich ein quantitativer Zugang zu den Daten ermöglicht werden, der diese anhand unterschiedlicher Parameter durchsuchbar und analysierbar macht. Ziel des Projekts ist also nicht nur, einen in dieser Form einzigartigen Bestand an Liedblättern nachhaltig digital zu erschließen und zugänglich zu machen, sondern darüber hinaus computergestützt nach Auffälligkeiten in Form wiederkehrender Phrasen und Themen oder melodischen Universalien zu suchen, die für verschiedene Regionen oder Zeitabschnitte charakteristisch sind.

2. Datenbasis

Die Datengrundlage des Projekts stellen umfangreichen Quellen zur Volksmusikforschung dar, die seit einigen Jahren von der Universitätsbibliothek Regensburg verwaltet werden. Die Regensburger Liedblattsammlung umfasst etwa 140.000 Blätter mündlich oder handschriftlich tradiertes Volkslieder aus dem gesamten deutschsprachigen Raum, und ist, was Abdeckung und Umfang angeht, in dieser Form einzigartig (Krüger, 2013). Die losen

¹ Anmerkung: Erste Vorarbeiten zu den hier beschriebenen Vorhaben erfolgten im Rahmen des DFG-Projekts „Erschließung von Quellen der Volksmusikforschung, Zugänglichmachung durch Digitalisierung sowie virtuelle Wiederherstellung zerstreuter Bestände“, vgl. <http://rvp.ur.de>.

Einzelblätter enthalten einerseits handschriftliche, monophone Melodien und andererseits Liedtexte, welche zumeist mit Schreibmaschine verfasst wurden (vgl. Abb. 1).

(Abschrift)

Klana Mann wollt' e grouß Fraa hou. Hä juch hä!

Grouß Fraa wollt' e klana Mann hou. Nadi nadi nadi

bums vallerä hop-sa sa sa

2. Grouß Fraa zum Tanzborrem ging, klana Mann wollt' a' mitgeh'n.

3. Klana Mann, du bleibst zu Haus, kehrst all' die Stowe aus.

4. Grouß Fraa vum Tanzborrem kam, klana Mann hinterm Ofen spann.

5. Mann, was host du gesponne? Dreimal hab ich abgesponnen.

6. Frau nahm den Rockenstock, schlug den Mann mit auf den Kopp.

7. Mann kroch ins Boddafass. Guckschte 'raus krieggte was.

8. Klana Mann lief zum Haus hinaus, lief schnell ins Nochbers Haus.

9. Nochber, soagt nor'mol oe, mich hat mei grouß Fraa geschloe.

10. Nochber, nor goar nett gekloagt, meini hott mer's grod so gemoacht.

Hähnlein, kr. Bensheim, - bl. 36
HESS.ARCHIV

A 23

Abbildung 1: Ausschnitt aus dem Liedblatt Nr. A23: „Klana Mann wollt' e grouß Fraa hou“.

Zu den Liedblättern existieren darüber hinaus Metadaten wie *Titel*, *Text-incipit*, *Sangesort* und *Jahr*, die ursprünglich in einem umfangreichen Zettelkastensystem vorlagen, mittlerweile jedoch in eine Datenbank (*Augias*) übertragen wurden. In Zusammenarbeit mit der Universitätsbibliothek Regensburg werden zunächst Scans der Liedblätter erstellt und mit den bereits vorhandenen digitalen Metadaten verknüpft. Daraufhin werden die Scans inhaltlich erfasst und in ein maschinenlesbares Format gebracht, das erlaubt, die Daten computergestützt zu durchsuchen und zu analysieren. Dieser Beitrag beschreibt Herausforderungen und Lösungsansätze bei der digitalen Erschließung der Liedblätter hinsichtlich ihrer Texte und Melodien.

3. Digitale Erschließung der Liedblätter

Für die Transkription der Texte und Melodien wurden Tools für die automatische Erfassung evaluiert. Neben automatischer Texterkennung (OCR, *Optical Character Recognition*), wurde auch die automatische Notenerkennung (OMR, *Optical Music Recognition*) untersucht (vgl. Bainbridge & Bell, 2001; 2006; Raphael & Wang, 2011; Rebelo, Capela, & Cardoso, 2010).

3.1. Erschließung der Liedtexte über OCR mit manueller Nachkorrektur

Die Evaluation der Eignung bestehender OCR-Tools für den Kontext der Regensburger Liedblattsammlung lehnt sich an Kanungo, Marton und Bulbul (1999) an. Das Testkorpus umfasst 102 Liedblätter, die möglichst viele unterschiedliche typographische und orthographische Phänomene abdecken, etwa Druckschrift (mit unterschiedlich starkem Kontrast), Frakturschrift, aufgeklebte Korrekturen, Sonderzeichen, etc. Für die Evaluation wurde die Textzone unterhalb der Notenzeilen ausgewählt, da die Noten als unbekannte Sonderzeichen das Texterkennungsergebnis negativ verfälschen würden. Für jene Textzonen wurde eine manuelle Transkription erstellt, die in der weiteren Evaluation als *ground truth* dient. Evaluiert wurden die folgenden drei OCR-Tools:

- *Abbyy Fine Reader* (<http://www.abbyy.de/>)
- *Omnipage Professional* (<http://www.nuance.de/for-individuals/by-product/omnipage/index.htm>)
- *Adobe Acrobat X Pro* (<https://helpx.adobe.com/de/acrobat/kb/acrobat-downloads.html>)

Mithilfe des OCR-Evaluationstools *ocrevalUAtion* (Carassco, 2014) wurde jeweils der Output der drei getesteten OCR-Tools mit den *ground truth*-Daten verglichen. Abb. 2 zeigt für jedes OCR-Tool die Anzahl korrekt erkannter Zeichen (*correct*), die Anzahl falsch erkannte Zeichen (*confused*), die Anzahl nicht erkannte Zeichen (*lost*) sowie die Anzahl überflüssiger Zeichen (*spurious*) als gestapeltes Balkendiagramm.

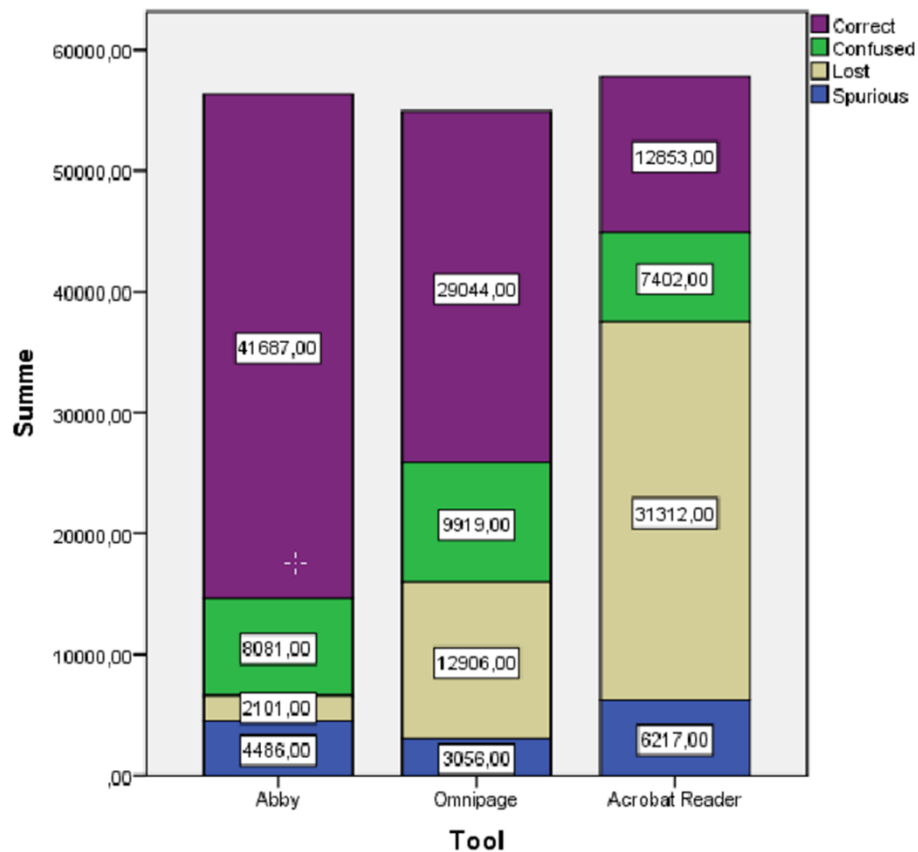


Abbildung 2: OCR-Evaluationsergebnisse für die getesteten Tools hinsichtlich der korrekt erkannten, der falsch erkannten, der gar nicht erkannten sowie der überflüssigerweise erkannten Zeichen.

Anhand dieser Parameter lassen sich Kennzahlen für die Tools berechnen, etwa die *precision* oder auch die *global error rate*. Bezüglich der korrekten Erkennung in Prozent wird deutlich, dass *Abby* mit einer Erkennungsrate von 80% (*Omnipage*: 56%, *Adobe*: 26%) und einer vergleichsweise geringen Streuung am besten in der Evaluation abschneidet (vgl. Abb. 3).

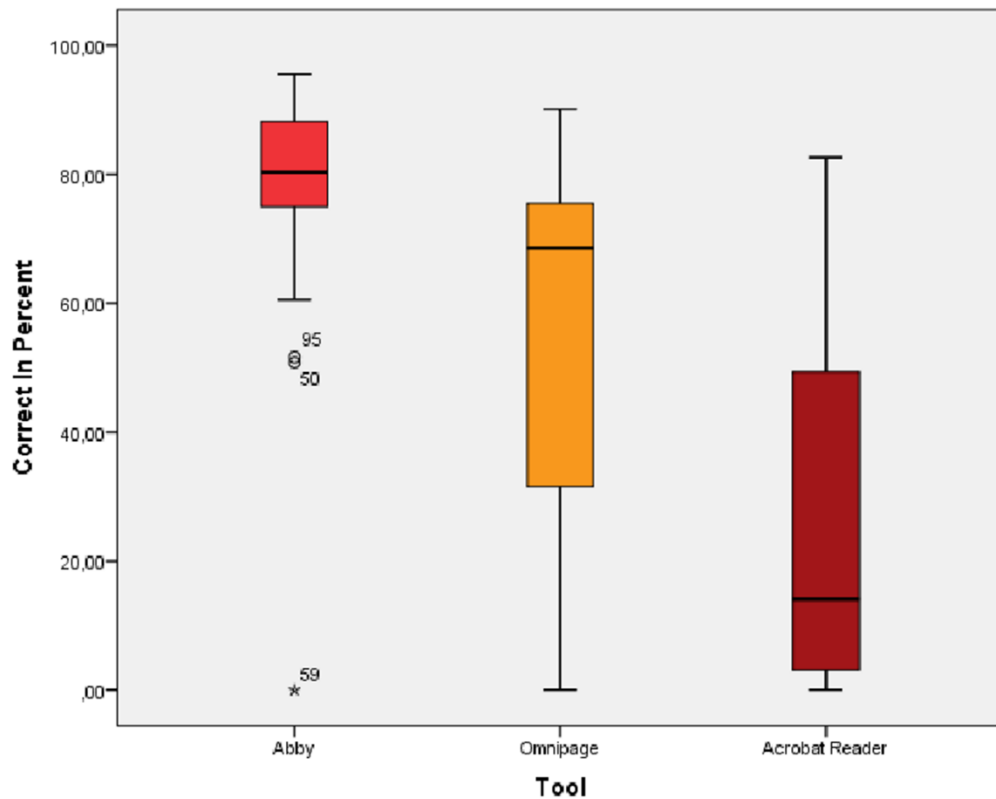


Abbildung 3: Boxplot zur Erkennungsgenauigkeit der einzelnen OCR-Tools.

Dass *Abby*-Tool liefert die besten Evaluationsergebnisse und wurde somit als OCR-Tool für die Liedblattsammlung ausgewählt. Die 80%-Erkennungsrate erlaubt erste explorative Analysen der Liedblätter anhand bestimmter Schlüsselwörter. Für die sukzessive Korrektur der Texte wurde ein Tool entwickelt, das die manuelle Korrektur des OCR-Outputs für jedes Liedblatt erlaubt. Um die Texte der insgesamt 140.000 Liedblätter möglichst effizient zu transkribieren, sind zudem weitere Evaluationsexperimente mit anderen OCR-Tools geplant. Zudem soll versucht werden, das *Abby*-Tool anhand der Liedblätter zu trainieren, um so die Erkennungsrate weiter zu verbessern.

3.2. Erschließung der Melodien über ein Crowdsourcing-Webtool

In Anlehnung an eine OMR-Evaluationsstudie (Bellini, Bruno & Nesi, 2007) wurden drei der am weitesten verbreiteten OMR-Tools hinsichtlich ihrer Eignung für die Liedblattsammlung evaluiert:

- *Photoscore* (<http://www.sibelius.com/products/photoscore/ultimate.html>)
- *SharpEye* (<https://www.columbussoft.de/SharpEye.php>)
- *CapellaScan* (<http://www.capella.de/de/index.cfm/produkte/capella-scan/info-capella-scan/>)

Anders als bei der OCR-Evaluation ist die Erstellung eines automatisch abgleichbaren *ground truth*-Datensatzes nicht ohne weiteres möglich, da die Erfassung musikalischer Notation

wesentlich komplexer ist als reine Textzeichenerkennung. Der Abgleich des jeweiligen OMR-Outputs mit dem entsprechenden Originalliedblatt erfolgte deshalb manuell. Insgesamt wurden auf diese Weise 20 Liedblätter ausgewählt, welche eine möglichst hohe Bandbreite unterschiedlicher Merkmalsausprägungen abdecken. Zu den Merkmalen zählen Zeichenabstand, Einheitlichkeit der Zeichen, allgemeiner Kontrast, Kontrast der Notenlinien, Größe der Notenköpfe, Länge der Notenhäse und das Vorkommen von Fremdzeichen.

Bei der Berechnung der Erkennungsgenauigkeit wurden dieselben Parameter verwendet wie schon bei der OCR-Evaluation (vgl. Abb. 2). Die Ergebnisse der OMR-Evaluation zeigen, dass hinsichtlich der durchschnittlichen Erkennungsgenauigkeit mit 36% bei *Photoscore*, 8% bei *CapellaScan* und 4% *SharpEye* keines der Tools auch nur ansatzweise für den produktiven Einsatz in Frage kommt (vgl. Abb. 4). Dabei ist selbst beim am besten evaluierten Tool *Photoscore* eine enorme Streuung zu beobachten, die bei 5 von 20 Blättern auf 0% kommt, und nur ein einziges Mal als beste Erkennungsrate 80% bei einem Liedblatt erreicht.

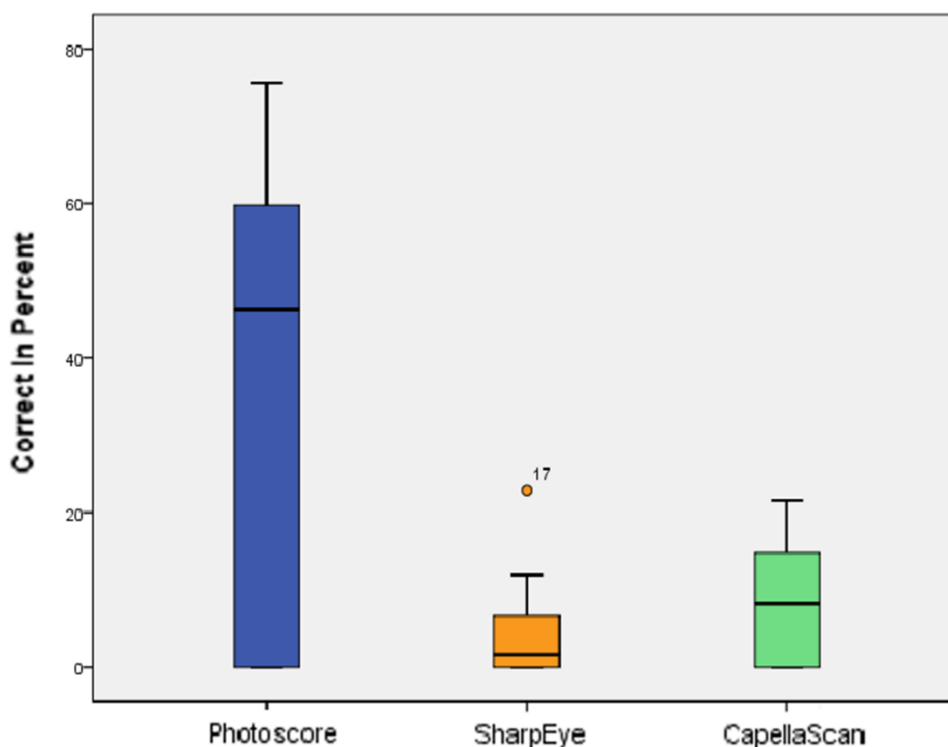


Abbildung 4: Boxplot zur Erkennungsgenauigkeit der einzelnen OMR-Tools.

Als alternative Erschließungsstrategie wurde ein Transkriptionstool namens *Allegro* entwickelt, welches aufgrund der erheblichen Datenmenge von mehreren tausend Liedblättern auf einen Crowdsourcing-Ansatz (Dunn & Hedges, 2013; Oomen & Aroyo, 2011) zurückgreifen soll. Erfolgreiche Beispiele für solche Ansätze im Bereich der Digital Humanities finden sich etwa beim Sammeln und Dokumentieren von urbaner Kunst (Burghardt, Schneider, Bogatzki, & Wolff, 2015), bei der Transkription von Manuskripten

(Causser & Wallace, 2012), bei der Verschlagwortung von Kunstwerken (Commare, 2011) und auch im Bereich der Transkription von Musikstücken, wie beim Projekt „What’s the Score?“².

Bei der Umsetzung des Tools für die Transkription der Regensburger Liedblätter wurde besonderes Augenmerk auf die einfache Bedienbarkeit durch iteratives *usability testing* während des Entwicklungsprozesses gelegt (vgl. ISO 13407:1999). Die Benutzeroberfläche wurde dabei so konzipiert, dass auch Personen, die keine Noten lesen können, in der Lage sind, die Noten zu transkribieren, indem sie diese auf ein virtuelles Notenblatt übertragen und das Original im Wesentlichen nachbauen (vgl. Meier et al., 2015). Die zusätzliche Möglichkeit der Transkription über ein Midi-Instrument soll später über einen speziell anzuwählenden Expertenmodus optional verfügbar gemacht werden.

Als erster Schritt wird in *Allegro* zunächst das Notenblatt manuell in einzelne Takte segmentiert (Abb. 5):

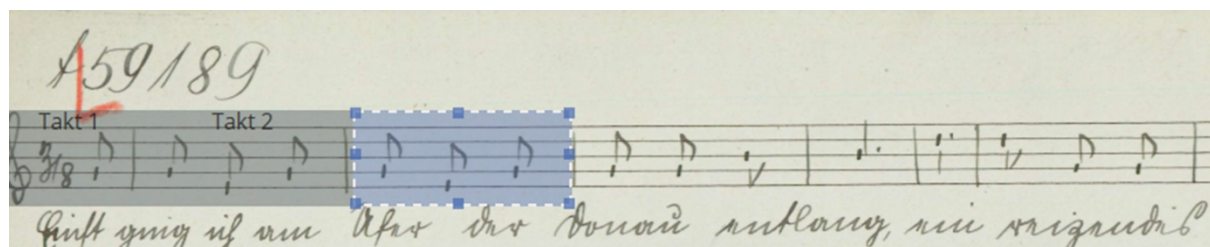


Abbildung 5: Taktweise Segmentierung der Liedblätter mit dem Allegro.

Nach Angabe der Liedblattnummer sowie der Auswahl von Taktart und Tonart gelangt man in den eigentlichen Transkriptionsmodus, bei dem Takt für Takt auf einer interaktiven Notenzeile mit Maus und Tastatur (Shortcuts) transkribiert wird (vgl. Abb. 6). Jeder einzelne Takt kann im Browser abgespielt werden, um so ggf. auf auditiver Ebene schnell Transkriptionsfehler zu erkennen.

² Projekt „What’s the Score?“ online: <https://www.bodleian.ox.ac.uk/weston/our-work/projects/whats-the-score>



Abbildung 6: Taktweise Transkription der Liedblätter mit dem Allegro-Tool.

Im Hintergrund werden die Eingaben auf das virtuelle Notenblatt schließlich in ein maschinenlesbares Format (*JSON*) übersetzt, das mithilfe einer *Converter-Toolbox* in beliebige andere Formate wie etwa *MusicXML* transformiert werden kann. Da die Transkription durch Laien eine erhöhte Gefahr für Transkriptionsfehler mit sich bringt, wird jedes Liedblatt doppelt übersetzt (vgl. das *double keying*-Konzept bei Texttranskriptionen). Liedblätter, bei denen die Transkriptionen nicht übereinstimmen, werden auf redaktioneller Ebene final geprüft. Um den Anreiz zur Beteiligung an der Transkription zu erhöhen, ist es den Teilnehmern möglich die selbst transkribierten Texte und Melodien in einer privaten Sammlung zu speichern und bei Bedarf als PDF bzw. als MP3 herunterzuladen.

Das Transkriptionstool befindet sich aktuell in der offenen Beta-Testphase und findet guten Zuspruch bei den Anwendern:

- *Allegro*: <http://138.68.106.29/>

4. Zusammenfassung

Dieser Beitrag gibt einen Einblick in ein laufendes Projekt zur digitalen Erschließung einer großen Sammlung von Liedblättern. Während OCR-Tools für die automatische Erfassung der Liedtexte annehmbare Ergebnisse mit einer Erkennungsrate von bis zu 80% liefern, so liegt die Erkennungsgenauigkeit bestehender OMR-Tools für die handschriftlichen Notensätze bei lediglich maximal 36%. Im Falle der Notenerkennung wurde von Grund auf

ein neues, intuitiv bedienbares Transkriptionstool entwickelt, welches über einen Crowdsourcing-Ansatz die sukzessive Erschließung der Notensätze sicherstellen soll.

5. Ausblick

Aktuell liegt der Projektfokus auf der Erschließung der Liedblätter. Parallel entstehen zudem erste Prototypen (vgl. Burghardt et al., 2016) für das angedachte Informationssystem, das die Analyse der Liedblätter anhand der verfügbaren Metadaten, der Liedtexte sowie anhand verschiedener melodischer Parameter (vgl. Mongeau & Sankoff, 1990; Orio & Rodá, 2009; Typke, 2007) erlaubt. Im Rahmen des weiteren Projektverlaufs sollen anhand der digital erschlossenen Liedblätter u.a. die folgenden Fragestellungen untersucht werden:

- Welche sind die häufigsten Wörter in den Texten deutscher Volkslieder, und welche Wörter treten besonders häufig zusammen auf (Kollokationen)? Lassen sich daraus Rückschlüsse auf wiederkehrende Themen ziehen, einerseits für das gesamte Liedblattkorpus, andererseits aus einer regionalen und diachronen Perspektive?
- Gibt es melodische Universalien, die typisch für deutsche Volkslieder sind, einerseits für das gesamte Liedblattkorpus, andererseits aus einer regionalen und diachronen Perspektive?
- Lassen sich musikalisch-linguistische Kollokationen identifizieren, kommen also bestimmte Melodien oder einzelne Rhythmen oder Intervalle besonders häufig in Texten mit auffälligen Schlüsselwörtern vor?

Bibliographie

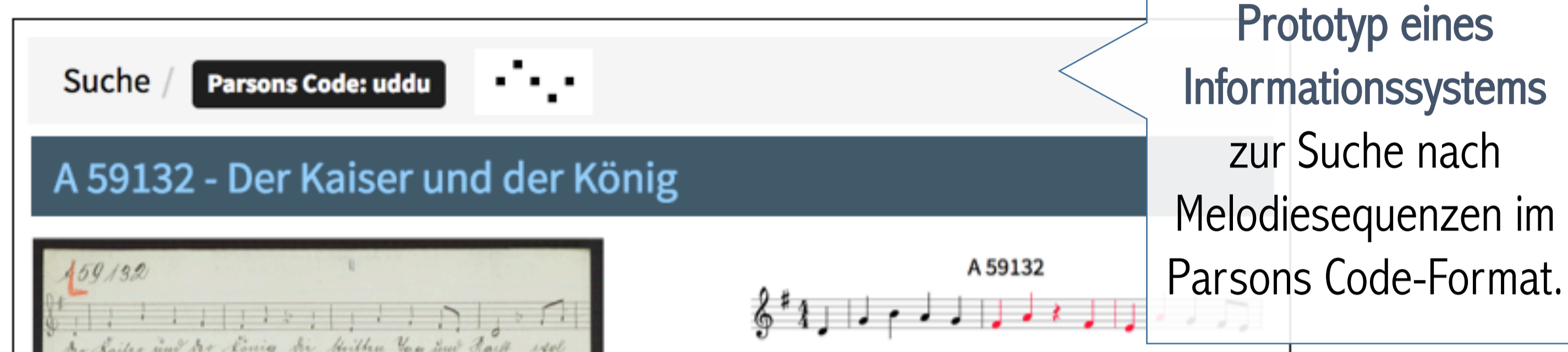
1. **Bainbridge, D. / Bell, T.** (2001). The challenge of optical music recognition. *Computers and the Humanities*, 35, 95–121.
2. **Bellini, P. / Bruno, I / Nesi, P.** (2007). Assessing Optical Music Recognition Tools. *Computer Music Journal*, 31(1), 68-93.
3. **Burghardt, M. / Lamm, L. / Lechler, D. / Schneider, M. / Semmelmann, T.** (2016). Tool-based Identification of Melodic Patterns in MusicXML Documents. In *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 440-442.
4. **Burghardt, M. / Schneider, P. / Bogatzki, C. / & Wolff, C.** (2015). StreetartFinder – Eine Datenbank zur Dokumentation von Kunst im urbanen Raum. In *Book of Abstracts, DHd 2015*.
5. **Carrasco, R. C.** (2014). An open-source OCR evaluation tool. In: *DATeCH 2014*. New York: ACM Press.
6. **Causer, T., / Wallace, V.** (2012). Building A Volunteer Community: Results and Findings from Transcribe Bentham. *DHQ202f: Digital Humanities Quarterly*, 6(2).
7. **Commare, L.** (2011). Social Tagging als Methode zur Optimierung Kunsthistorischer Bilddatenbanken – Eine empirische Analyse des Artigo-Projekts. In: *Kunstgeschichte. Open Peer Reviewed Journal*, 2011 (urn:nbn:de:bvb:355-kuge-160-9)
8. **Dunn, S., / Hedges, M.** (2013). *Crowd-sourcing as a Component of Humanities Research Infrastructures*. In *International Journal of Humanities and Arts Computing*, 7(1-2), 147-169.

9. **Kanungo, T. / Marton, G. A. / Bulbul, O.** (1999, January). Performance evaluation of two Arabic OCR products. In *The 27th AIPR workshop: Advances in computer-assisted recognition* (pp. 76-83). International Society for Optics and Photonics.
10. **Krüger, G.** (2013). Das „Regensburger Volksmusik-Portal“ der Universitätsbibliothek Regensburg. Bestände – Problematiken – Perspektiven. Zwischenbericht aus einem Erschließungsprojekt. In E. R. Mohrmann (Ed.), *Audioarchive – Tondokumente digitalisieren, erschließen und auswerten* (S. 119–131). Münster et al.: Waxmann Verlag.
11. **Meier, F. / Bazo, A. / Burghardt, M. / Wolff, C.** (2015). A Crowdsourced Encoding Approach for Handwritten Sheet Music. In J. Roland, Perry; Kepper (Ed.), *Music Encoding Conference Proceedings 2013 and 2014* (pp. 127–130).
12. **Mongeau, M. / Sankoff, D.** (1990). Comparison of Musical Sequences. *Computers and the Humanities*, 24, 161–175.
13. **Oomen, J. / Aroyo, L.** (2011). Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges. *C&T '11 Proceedings of the 5th International Conference on Communities and Technologies* (S. 138–149).
14. **Orio, N. / Rodá, A.** (2009). A Measure of Melodic Similarity Based on a Graph Representation of the Music Structure. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)* (S. 543–548).
15. **Raphael, C. / Wang, J.** (2011). New Approaches to Optical Music Recognition. In *12th International Society for Music Information Retrieval Conference (ISMIR)* (S. 305–310).
16. **Rebelo, A. / Capela, G. / Cardoso, J. S.** (2010). Optical recognition of music symbols. *International Journal on Document Analysis and Recognition*, 13, 19–31.
17. **Typke, R.** (2007). Music Retrieval based on Melodic Similarity. *Ph.D Thesis, Utrecht University*.

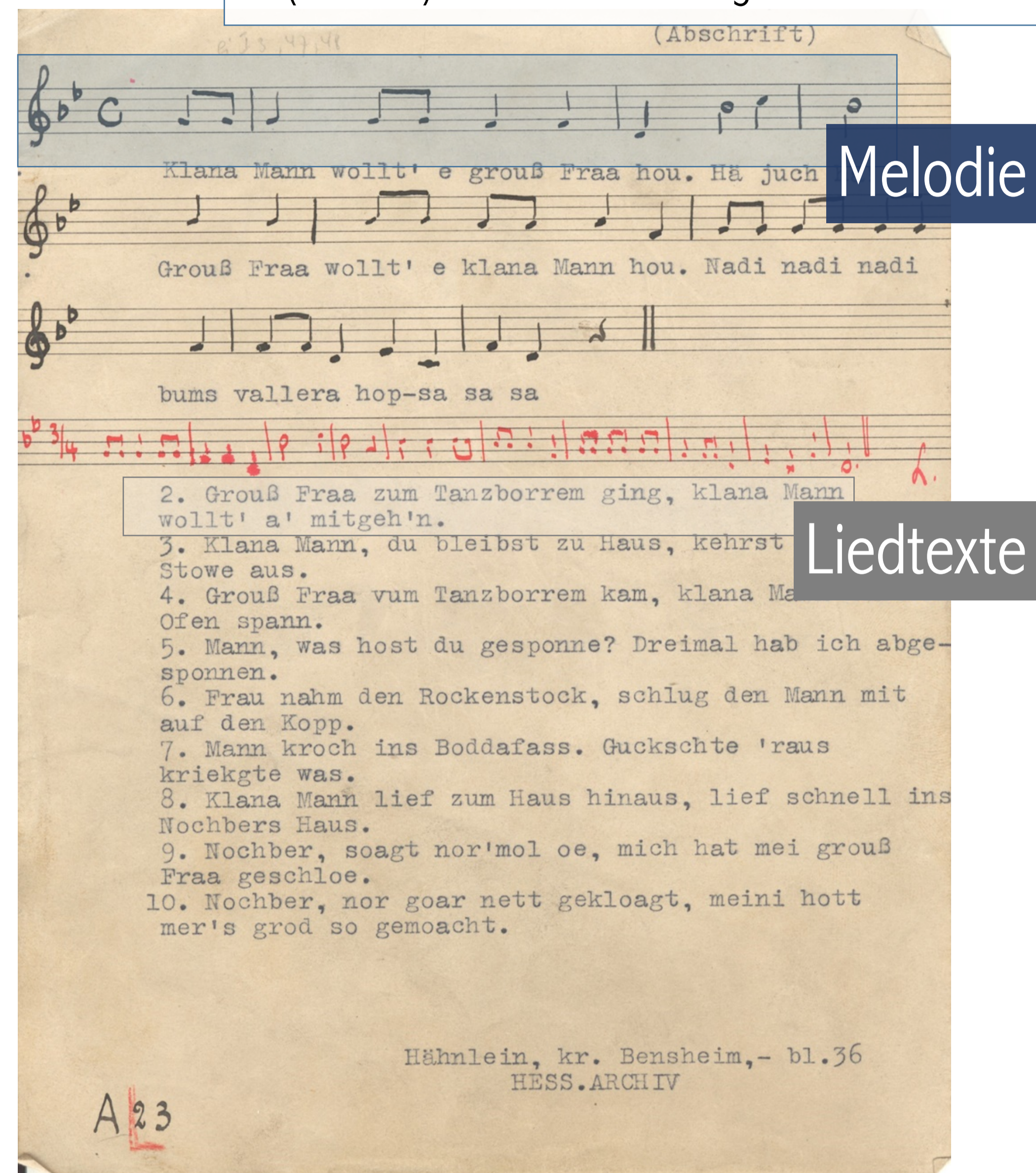
DIGITALE ERSCHLIEßUNG EINER SAMMLUNG VON VOLKSLIEDERN AUS DEM DEUTSCHSPRACHIGEN RAUM

I) PROJEKTKONTEXT UND ZIELE

- Digitale Erschließung einer großen Sammlung von Volksliedern aus dem deutschsprachigen Raum
- Langfristiges Ziel: Erstellung eines Informationssystems zur quantitativen Analyse der Liedblätter bzgl. Liedtexten und Melodien (Suche nach konkreten Melodiefragmenten und melodischen Ähnlichkeiten)



Beispielhaftes Liedblatt mit handschriftlichen Melodien und (zumeist) schreibmaschinengeschriebenen Liedtexten.



II) EVALUATION VON OCR-/OMR-TOOLS

OCR-STUDIE – DURCHSCHNITTliche ERKENNUNGSGENAUIGKEIT:

- Abby Fine Reader (80%) 😊
- Omnipage Professional (56%)
- Adobe Acrobat X Pro (26%)

Evaluationsdesign angelehnt an Kanungo, Marton & Bulbul (1999); OCR-Evaluationstool: ocrevalUAtion (Carasco, 2014)

OMR-STUDIE – DURCHSCHNITTliche ERKENNUNGSGENAUIGKEIT:

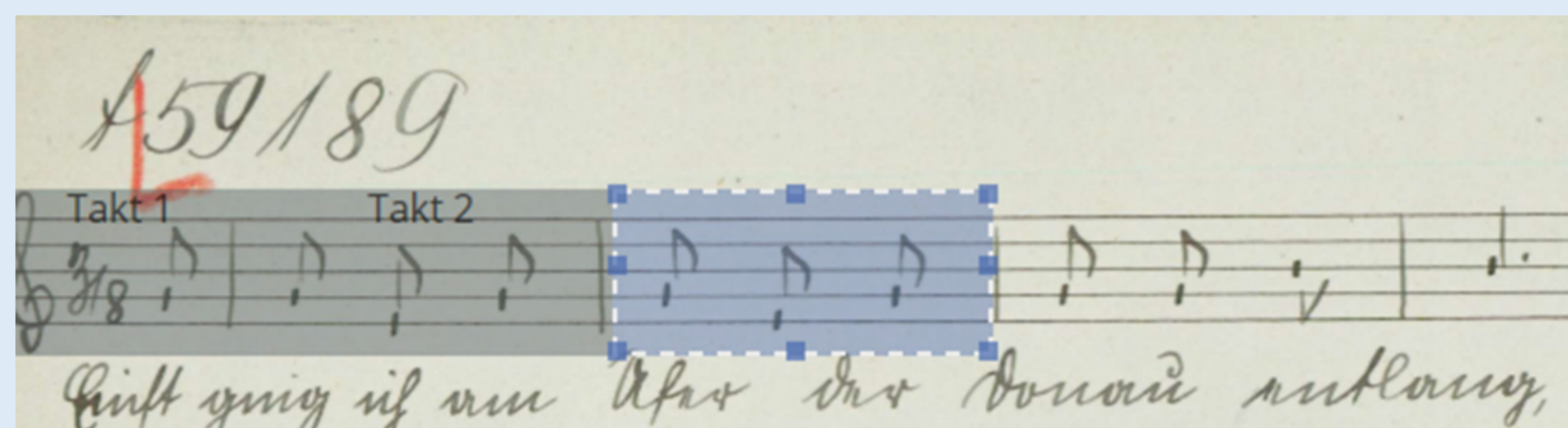
- Photoscore (36%) 😞
- CapellaScan (8%)
- SharpEye (4%)

Evaluationsdesign angelehnt an Bellini, Bruno & Nesi (2007)

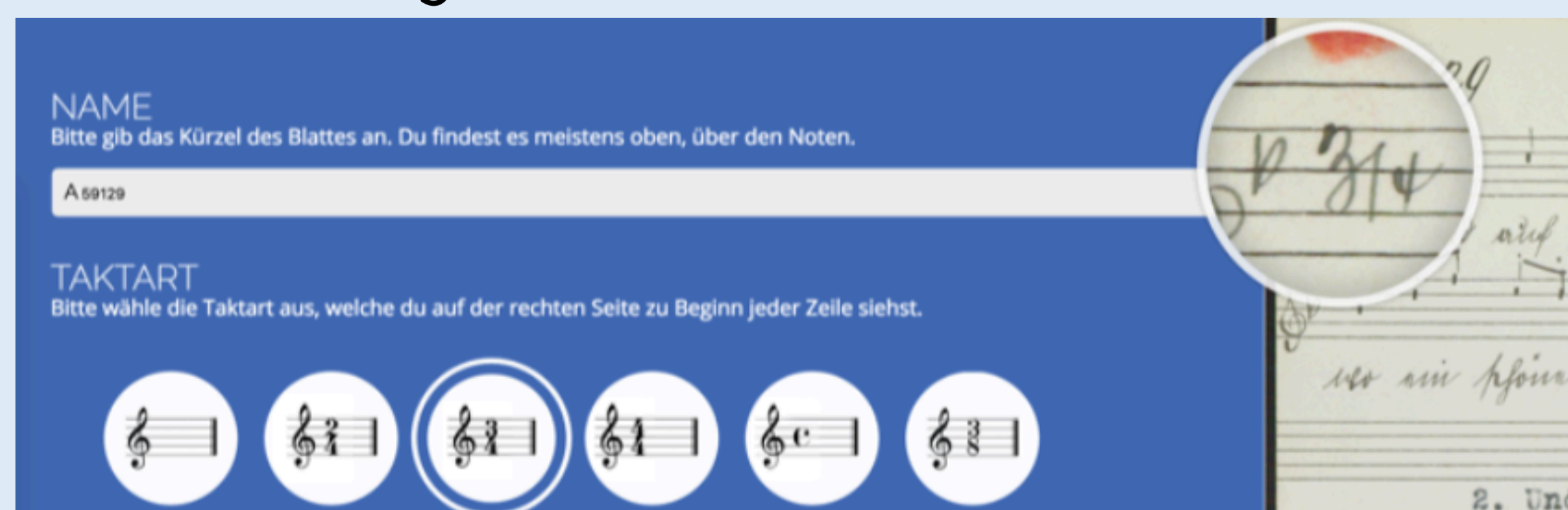
„OMR for handwritten scores as a major unresolved problem.“ (Müller, 2007, p. 20)

III) TRANSKRIPTION DER MELODIEN ÜBER EIN CROWDSOURCING-WEBTOOL

Schritt 1: Segmentierung der Takte



Schritt 2: Angabe von Taktart und Tonart



Schritt 3: Transkription der Takte



BIBLIOGRAFIE

Carrasco, R. C. (2014). An open-source OCR evaluation tool. In: *DATECH 2014*. New York: ACM Press.

Bellini, P. / Bruno, I. / Nesi, P. (2007). Assessing Optical Music Recognition Tools. *Computer Music Journal*, 31(1), 68-93.

Kanungo, T. / Marton, G. A. / Bulbul, O. (1999). Performance evaluation of two Arabic OCR products. In *The 27th AIPR workshop: Advances in computer-assisted recognition* (pp. 76-83). International Society for Optics and Photonics.

Müller, M. (2007). *Information Retrieval for Music and Motion*. Springer: Berlin.

► Mehr Informationen zum Projekt unter www.dhregensburg.de

Manuel Burghardt (manuel.burghardt@ur.de), Sebastian Spanner (sebastian.spanner@stud.uni-regensburg.de), Thomas Schmidt (thomas.schmidt@stud.uni-regensburg.de), Florian Fuchs (florian.fuchs@stud.uni-regensburg.de), Katia Buchhop (katia.buchhop@stud.uni-regensburg.de), Miriam Nickl (miriam.nickl@stud.uni-regensburg.de), Christian Wolff (christian.wolff@ur.de) – Arbeitsgruppe Digital Humanities, Lehrstuhl für Medieninformatik, Universität Regensburg