*Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, pages 885–890
Marseille, 11–16 May 2020


# LiViTo: linguistic and visual features tool for assisted analysis of historic manuscripts

**Klaus Müller[1], Aleksej Tikhonov[2], Roland Meyer[3]**
[1]Musterfabrik Berlin/Fraunhofer IPK, Brunnenstr. 156, 10115 Berlin, Germany
[2,3]Department of West-Slavic Languages, Humboldt University Berlin, Unter den Linden 6, 10099 Berlin, Germany
klaus.mueller@musterfabrik-fabrik.de, tikhonal@hu-berlin.de, roland.meyer@hu-berlin.de

**Abstract**

We propose a mixed methods approach to the identification of scribes and authors in handwritten documents, and present LiViTo, a software tool which combines linguistic insights and computer vision techniques in order to assist researchers in the analysis of handwritten historical documents. Our research shows that it is feasible to train neural networks for the automatic transcription of handwritten documents and to use these transcriptions as input for further learning processes. Hypotheses about scribes can be tested effectively by extracting visual handwriting features and clustering them appropriately. Methods from linguistics and from computer vision research integrate into a mixed methods system, with benefits on both sides. LiViTo was trained with historical Czech texts by 18th century immigrants to Berlin, a total of 564 pages from a corpus of about 5000 handwritten pages without indication of author or scribe. We provide an overview of the three-year development of LiViTo and an introduction into its methodology and its functions. We then present our findings concerning the corpus of Berlin Czech manuscripts and discuss possible further usage scenarios.

**Keywords:** mixed methods, digital humanities, machine learning, linguistics, manuscripts, authorship, Czech, Slavic, Slavonic, handwriting

## 1. Introduction

Historical manuscripts are an essential source for all humanities. In line with the apparently rising interest in everyday culture (Meyer, 2007), new research stimuli can be expected from hitherto unknown texts in small private or parochial archives. They may contain reports by personal witnesses, new information on everyday culture, and samples of everyday language. For a long time, many communities have been cherishing a purely handwritten culture that can be fully understood only by assessing personal manuscripts. The present project is devoted to exploring handwritten sources of a small, but linguistically, culturally and historically extraordinary community of migrants to 18th c. Berlin: the "*Czech brethren*" (aka *Moravian Church*, *Herrnhuter Brüdergemeine*), who fled from anti-Protestant persecution in the Czech lands to Saxony and Prussia. Research questions that arise in this context are: 1) Are the manuscripts originals or handwritten copies? 2) Are the originator/author and the scribe the same person? 3) How many authors and scribes worked on the manuscripts? 4) Can these authors and scribes be found in other manuscripts? 5) Are there revisions in the manuscripts and where are they?

For the identification of scribes and authors in manuscripts, methods from classical linguistic analysis are combined with modern computer vision approaches, such as neural networks. This approach enhances the knowledge discovery process and knowledge representation. The idea of automatically discovering knowledge from manuscripts is a very attractive and challenging task. The main motivation for applying neural networks to knowledge extraction and clustering tasks is that they require mainly data to make the algorithm work. Hence they do not depend on a specific language the manuscripts are written in and can be easily retrained for other languages. To get information about authors and scribes statistical analysis of linguistic features is used in order to make quantitative statements. Analysing the output generated by neural networks with the algorithm and the Euclidean or other distances allows to make more precise statements about the relationship between texts, authors and scribes even for rather short documents of less than 1000 tokens. Text length is a major issue for alternative, stylometric approaches to our research question, such as the one implemented in *stylo* (Eder, 2016). Unfortunately, stylo is suboptimal for our case due to the strong dependence of the text length and the accuracy of the results.

The potential target audience for LiViTo are generally researchers and students from different disciplines - humanities, social studies, law and medicine. In general, LiViTo is intended to be an assistance system for analysing, comparing and clustering of handwritten (historic) data. Research questions in law and medicine could be the origin or the linguistic and visual interdependence of handwritten legal documents such as birth certificates and clinical records, testaments or agreements. In that regard, the questions of the humanities and social sciences are often equal to law and medicine. Such tasks can occur in history, theology, literature, cultural studies, linguistics, cultural anthropology, sociology, political sciences and many other research subjects. Obviously there is a high amount of fields in the humanities and social studies which could reach the goals of their research questions easier and faster. Meticulous and close reading, understanding and analysis of handwritten texts is an inalienable part of such a research process. Such a qualitative research approach can be combined with the (half-)automatic methods of LiViTo focusing quantitative data, obtained by statistical research methods, resulting in a mixed methods research approach. It should be clear to the user that LiViTo is a data driven assistance system which provides results that can lead the user to both kinds of results - quantitative and qualitative. LiViTo is designed in such a way that it requires the user to have only minimal programming knowledge.

The intent of this software design is to enable the user to get first insights into the manuscripts and iterate faster through research questions rather than spending time learning a complex tool. This is why it is relevant for archivists, curators, museum employees, genealogists and people who are in general interested in history and historical manuscripts. The users are provided with a step by step manual which should help in the beginning working with LiViTo.

## 2. Related Work

The related works can be divided into three fields: linguistic and visual author attribution, stylometric research and computer aided keyword analysis/search on digital documents. LiViTo combines methods from all three fields.

Burrows (1987, 1989) designed the approach of analysing word frequencies to visualize the distance between two or more texts in terms of authorship. By projecting word frequency vectors in a two-dimensional coordinate system, the similarity of two or more texts could be deduced. For two texts, in practice it would mean that the farther apart two vectors are, the less similar they are, and the closer they are to each other, the greater is their similarity. Two comparable methods are analysis of variance (ANOVA) and Kullback-Leibler divergence (KL divergence). According to Koppel et al. (2009) KL divergence is of greater importance because it is not based solely on the relationships between individual word frequencies, but on the stochastic Markov chain and the probability distance (Koppel et al., 2009, p. 10). One of the central roles of function words in multivariate analysis is also inherited in machine learning approaches where text categorization is based on neural networks. This method of author assignment has been widely used in various disciplines since 1993 (Koppel et al., 2009, p. 11). All three methods of text investigation have found diverse applications over the past 30 years. Jonathan Hope (1994) studied the authorship of Shakespeare's plays. His sociolinguistic analysis explored the connections between John Fletcher, Thomas Middleton, and the English classic. The results allowed an objective view of Shakespeare's collaborative writing practices.

Much of such research is ultimately based on Baayen's (2008) seminal work on analyzing linguistic data in R (see Dévière 2009, Krajewski & Matthews 2009). The R-package *stylo* developed by Eder, Rybicki & Kestemont (see Eder 2010, Eder & Rybicki 2011, Eder, Rybicki & Kestemont 2016) is a tool for statistical analysis of the style of one or more texts. In recent years, stylometric techniques in combination with *stylo* have become popular among language, literary and cultural scholars who are concerned with the question of authorship of texts and language statistics. Zyl & Botha (2016) examine the direct speech of the television series *The Big Bang Theory* and prove that the main protagonist - physicist Sheldon Cooper - linguistically distinct from other protagonists or stands out in spoken language very strong to written academic style tends. In addition, they find that the scriptwriters had created clear linguistic concepts for all the characters, which could be measured stylometrically. Likewise, other current works show the wide range of (semi-)automatic stylometric methods by means of which anonymous blog and forum entries can be compared and cloned with e-mails and other forms of electronically transmitted messages. For example, the authorship of texts (see Afroz et al., 2014) or plagiarism (Oakes 2014).

The usage of tools for authorship analysis needs digital input data, but as most historical documents have not been transcribed into a digital format and the manual transcription process itself is very time consuming, there has been research on automatic optical character recognition dating back as early as 1913 (d'Albe & Lodge, 1914). A first system which was able to transcribe more than just single well separated characters was the omni-font software developed by Kurzweil Computer Products in 1974 (Haugher, 1995). A prominent free open source tool for OCR, which can transcribe various languages and styles is Tesseract. It was developed in the 1980s and became open source in 2005 (Kay, 2007). Recent advancements made in machine learning led to first research results on machine-recognized handwritten texts, which are on human level accuracy (Graves & Schmidhuber, 2009). Inspired by these technological improvements Transkribus, a service platform for computer-aided transcription, was developed in 2017 (Kahle et al, 2017).

## 3. Use Case and Data

During the period of cruel counter-reformation after the victory of the Catholic Union in the Battle of *Bílá hora* (1620), many Hussite protestants had to flee the Czech lands. A group of them belonging to the "Czech brethren" / "Brüdergemeine" community was granted religious liberty in Prussia and settled down in Berlin-Rixdorf in the 1730s.

The Brüdergemeine maintained its handwritten culture in individualized documents such as CVs and sermons up to the 1820's (Mettele, 2009). The object of the present study are previously scientifically unedited Czech documents from the *Archiv im Böhmischen Dorf* - the parish archive of the Brüdergemeine in Berlin, namely, CVs and so-called Chorreden (i.e., sermons to sub communities in the parish) – rare documents, valuable for Czech language history; but since neither translators nor Czech authors of the sermons are known, it is difficult to assess the apparent German language influence on them – is it due to long-term contact, widespread bilingualism, independent language change, or merely to "translationese"? Since some of the handwritings in the sermons may be identical with those of the autobiographies, it is important to uncover textual interrelations with efficient methods for analysing up to 5 000 handwritten pages. The part of the Rixdorf (RIX) dataset, analysed for this survey, consists of two books containing overall 183 CVs, some of which considerably vary in length. An example of the handwriting style and content can be seen in Figure 1:

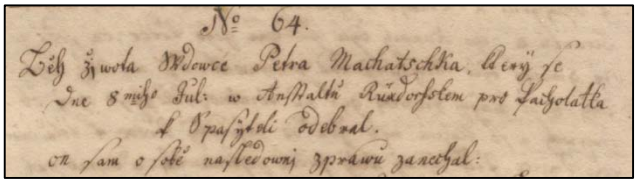

Figure 1. Example of the handwriting style

The shortest CV is two lines and the longest contains 23 pages. Therefore it is not possible to make a statistically sound assertion about all CVs. Due to the fact that the data set is monothematic the variance in the statistical analysis gets reduced, as the language corpus size is in terms of vocabulary limited to one topic. The CVs are partly constructed according to patterns commonly used in the parish. These patterns are not artistic, but rather specify obligatory elements of a CV, e.g. the moment of spiritual enlightenment. The CVs usually deal with the life story of common parish members, i. e. craftsmen who are in most cases also the originators of the texts. The function of the CV is close to that of the sermons - a teaching doctrine for other members (Lost, 2007). The CVs were introduced in 1747 by the bishop of the parish Nikolaus Ludwig Imperial Count von Zinzendorf und Pottendorf (1700-1760) and at first were obligatory self-written (ibid, p. 10), but from the end of the 1750's this rule was not followed strictly anymore. From that time most CVs were not or only partly self-written by the persons they are dealing with (Böß, 2016).

# 4. Methods

## 4.1 Preconditions

Since LiViTo processes manuscripts it needs image files. It can process various file formats like tif, png and jpg. It has to be taken into account that for scribe detection the analysis should include at least five pages per potential scribe as well as at least two potential scribes. In other words the data set should include not less than ten pages. In order to use the keyword detectors functionality transcripts of the manuscripts for the training of the neural network need to be provided by the user. The following applies in general for all neural networks: The more data is provided, the better neural network's results will be. To decrease the training time for neural networks it is advisable to use computer with a strong dedicated graphical processing unit (e.g. NVIDIA GeForce GTX 1060).

## 4.2 Software structure

LiViTos functionality shall be explained on three use cases. Figure 2 gives an overview of LiViTos module dependency and data flow. The main module, which is underlying the main functionalities is the pre-processor module. It needs to be executed before the other modules can be used. The pre-processor takes all input images and extracts text lines from the manuscripts, which will be needed for the other modules. The pre-processor also creates a data structure which will be built upon by the other modules.

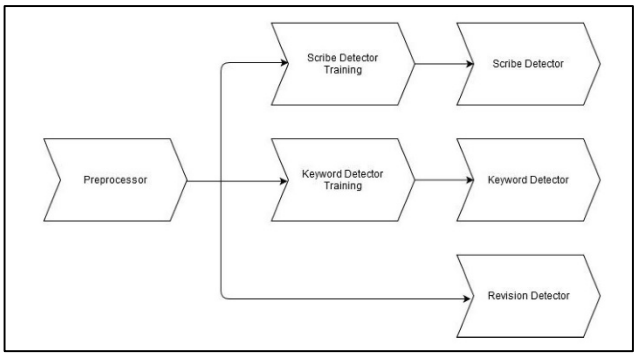


Figure 2. LiViTo's data flow and module dependency

### 4.2.1 Module 1: Scribe Detector

The interaction with this module is fairly simple. The user needs to provide a folder of about 100 text line segmentations per scribe to the training system, which were generated in the pre-processing step. Next the neural network will try to differentiate the scribes from each other. The grade to which the hypothetical scribes are distinguishable will be shown in a graph. The scribe detector has two use cases. First it can be used as a tool for hypothesis testing for identifying probable scribes in

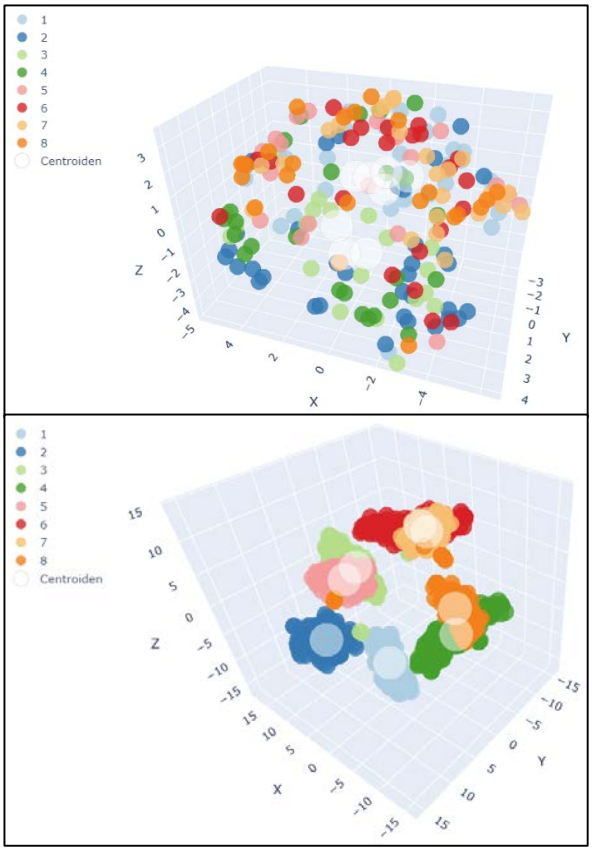


Figure 3. Output of Module 1: Scribe Detector

manuscripts. Figure 3 shows the users hypothesis/training results for two training processes. The upper graph in Figure 3 shows that the user made a hypothesis, which the network could not verify, as no mono colour clusters can be formed. Each data point represents a single text line segmentation, whereas the colour stands for the respective class attributed by the neural network. 128 features from

and intermediate network layer are embedded with t-SNE into a 3 dimensional representation. There the measure on the axes is not as relevant as the clustering itself. The lower graph in Figure 3 illustrates the case where the user made a hypothesis which can verified to a high degree.

#### 4.2.2 Module 2: Keyword Detector

The keyword detector module is a customizable query engine, which needs to be trained on the user's data. It needs transcripts of text lines provided by the user to train a neural network. Once the model is trained the user can query the manuscripts for detecting language features which can be traced back to an individual style of writing in both meanings - author and scribe. In order to clarify how it works some examples are given.

● The use of lexis from colloquial language is concerned with linguistic register variation or dialectology, e.g. Czech pronoun <won>[1] which is marked by the initial prosthetic <w-> as clearly colloquial. That can be traced back to a specific author who used colloquial language in written texts.

But the keyword detector is not only detecting full word forms as might be expected. Word fragments or a single letter can be detected and analysed as well.

● Like <won> the adjective ending <-ej> is concerned with linguistic register variation or dialectology in Czech. The query for "-ej " as an ending of words would show which texts in a particular data sample are written in colloquial Czech or a dialect of Czech.

● Till the end of the 18th century it was possible to use different orthographic variations of the same consonant [š]. Mainly the use of <ſſ> or <ſy> instead of <š>. From the beginning of the 19th century it was standardised to use only one possible graphic realisation which is <š>. This means that manuscripts which use all three variations are probably from the turn of the 18th/19th century. In cases of an exclusive use of <ſſ> or <ſy> the manuscripts can't be from the 19th century and with exclusive <š> matches have been written definitely after 1820. Such partial analysis could not only help to date manuscripts, but in the case of handwritten copies show from which period the original might be.

● Another important means in linguistics is the statistics of upper and lower case appearances. This statistic is provided as well. It will show the distribution of absolutely and relatively upper, lower and other[2] characters for the whole data set. The interdependence of the number of upper and lower cases can be attributed to educational background and grammar competence of the scribe or as well the period of the origin of the manuscript, because at different times different rules or norms of Upper-Lower-Case-writing did exist.

Among single word forms, word fragments and single letters can be used for analysing a larger language unit as well.

● With the query for verb endings like feminine past affix <-ala> it can be detected, how many of these verbs are positioned in accordance to the Czech syntax and which are positioned in accordance to the German syntax in the end of the sentence. The second case can be interpreted as German structural influence on Czech grammar in a language contact situation.

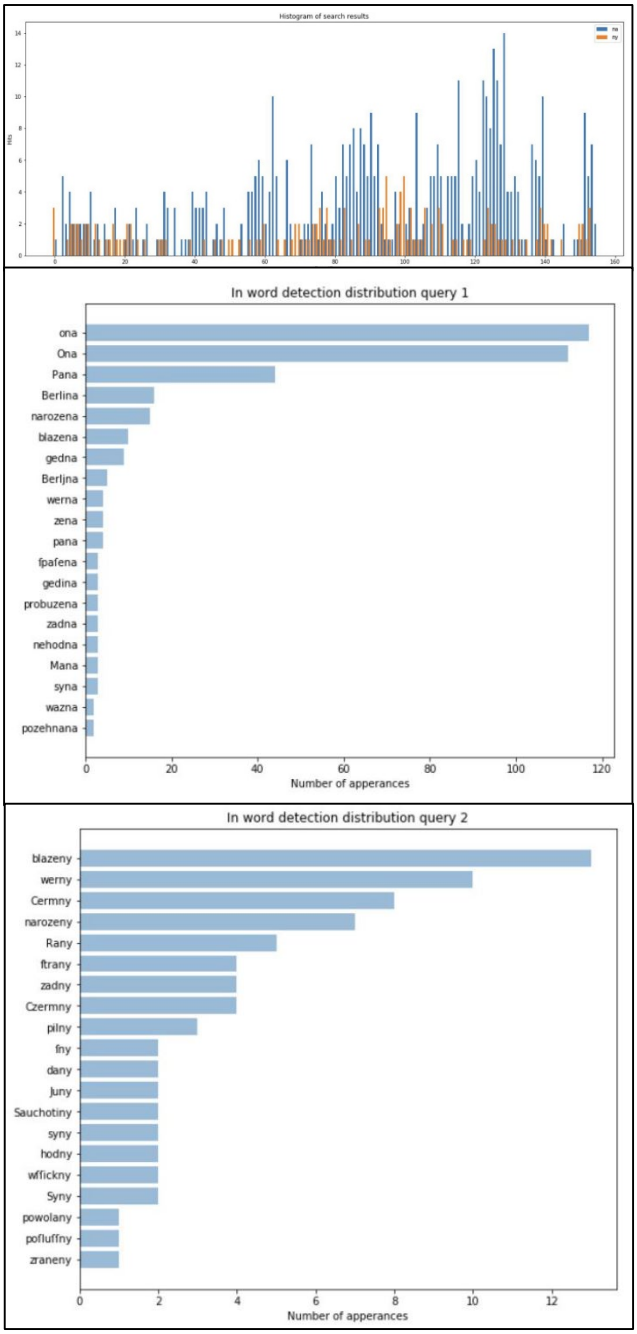

Figure 4. Output Module 2: Keyword Detector

The results for a competing query of Czech adjective feminine ending <-ná>[3] and masculine ending <-ný> can be seen in Figure 4. The top graph shows the distribution of detection over all documents (blue = feminine; orange = masculine). Figure 4 middle (query 1)/bottom (query 2) show the top 20 accrued words for each query. In this case

[1] standard Czech of 18th century & today: *on*; English: *he*

[2] Punctuation characters and numerals.

[3] There are more feminine / masculine endings in Czech. This is only one example.

there are as well some pronouns among the results (e.g. <ona> (Eng.: she)) and a few number of other word classes, which should be ignored in the analysis. This module also contains a manuscript viewer, which lets the user browse through the query results. Deleting individual results will dynamically adjust the statistical outputs.

#### 4.2.3 Module 3: Revision Detector

It is mainly a manuscript viewer, which can be used directly after pre-processing the data. It will highlight three different kinds of revisions in manuscripts. Crossed out areas, annotations made above a text line, and probable changes of single letters (e.g. if a scribe changes the letter <a> to <e>). The model underlying this module was generated by a convolutional neural network inspired by the U-Net[4]. It was tested on different historical manuscripts from different epochs and languages and is fairly robust, therefore it does not need to be trained on the user's data. This module is meant as a help for researchers working faster through their documents and giving them an overview of what kind of revisions can be found. It might

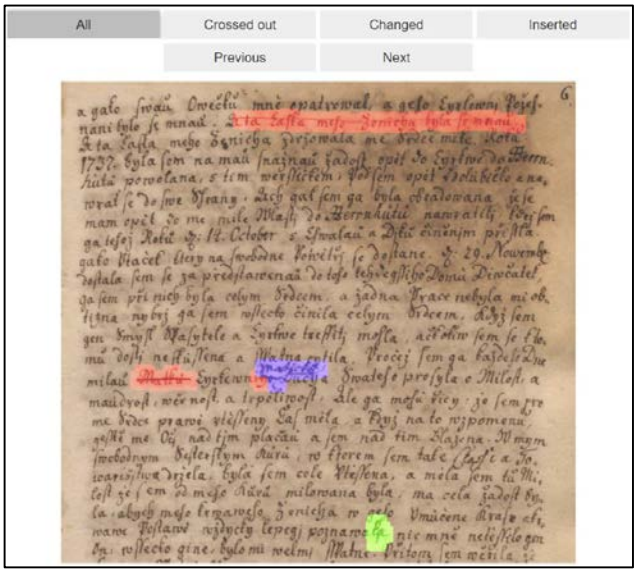


Figure 5. Output and Interface of Module 3: Revision Detector. Displaying all three kinds of revisions detected on a manuscript.

be used as a first indicator for the revisions made. The layout of the revision detector is displayed in Figure 5. The user can browse through the whole data set looking for all detected revision types or just individual ones. Each revision type is highlighted in a different colour, so they are easily distinguishable.

## 5. Use Case Evaluation

In our special research question about the scribes of the Czech immigrant manuscripts from the 18th century in Berlin LiViTo assisted us in making the following conclusions: (i) Significant revisions were made in the first half of the 19th century. LiViTo helped localizing the revisions without close reading of the manuscripts and categorized their kinds. Especially the revisions of <j> to <í> guided us to the revision moment not earlier than in the 1820s, because of the grammatical regulations of Czech standard language made by Josef Dobrovský in the first 20 years of the 19th century (Dobrovský, 1809, 1819). (ii) We could finally identify 10 scribes who wrote the analysed manuscripts with linguistic and visual features. LiViTo assisted the search for different linguistic features, as the archaic verb form <geſt> versus the modern <ge>. (iii) In the tradition of the Czech brethren in the 18th century these manuscripts should be an autograph of the people the CVs are dealing with. Altogether there are 183 people mentioned, but there are 10 scribes. All manuscripts are probably copies and not the originals. A deeper interpretation of the results will be finished by April 2020.

## 6. Conclusion

In this paper, we have described a non-commercial software tool, named LiViTo, that provides assistance system for the analysis of historical manuscripts. LiViTo is designed as an open source tool which will be released in the end of March 2020. It allows the users to modify it and assumes sharing of new possible modifications on GitLab. This tool relieves researchers of much technical work and allows them to focus on the analysis of their data and iterate faster through hypotheses. Moreover, the tool enables researchers with little knowledge of machine learning methods to apply them to their work. We have shown the main features of this software tool and we have distinguished three main parts: a module for scribe detection, a module for keyword detection and linguistic feature analysis, and a module for revision detection. Future expansion might include stylometric analysis of transcribed text with tools like *stylo* (Eder, 2016) and general NLP applications on the automatically transcribed texts generated by the keyword detector pre-processor. First tests on human vs. machine transcribed texts look very promising as stylometric analysis for both transcripts yield almost identical results. So far we have used the tool only on manuscripts written in languages using the Latin alphabet from the 18th century. Therefore, it would be interesting to see how well it performs on texts using other alphabets like Cyrillic, Arabic or Chinese.

## 7. Acknowledgements

This work was supported by a VolkswagenStiftung grant entitled "Tracing Patterns of Contact and Change" to Roland Meyer (Humboldt-Universität zu Berlin) and Bertram Nickolay (Fraunhofer IPK Berlin) under the Mixed Methods in the Humanities program (A118750, 04/2017-03/2020).

## 8. Bibliographical References

Afroz, Sadia, Aylin Caliskan Islam, Ariel Stolerman, Rachel Greenstadt, und Damon Mccoy. „Doppelgänger Finder: Taking Stylometry to the Underground". In *Security and Privacy (SP), 2014 IEEE Symposium On*,

[4] https://arxiv.org/abs/1505.04597 (Last access December 2, 2019)

212–226. IEEE, 2014. https://doi.org/10.1109/SP.2014.21.

Albe, E. E. Fournier d', und Oliver Joseph Lodge. „On a type-reading optophone". *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 90, Nr. 619 (1. Juli 1914): 373–75. https://doi.org/10.1098/rspa.1914.0061.

Baayen, Rolf Harald. *Analyzing linguistic data : a practical introduction to statistics using R*. Cambridge: Cambridge University Press, 2008.

Böß, Stephanie. *Gottesacker-Geschichten als Gedächtnis: eine Ethnographie zur Herrnhuter Erinnerungskultur am Beispiel von Neudietendorfer Lebensläufen / Stephanie Böß*. Studien zur Volkskunde in Thüringen BV035193383 Band 6. Münster ; New York: Waxmann, 2016.

Burrows, J. F. „'An Ocean Where Each Kind. . .': Statistical Analysis and Some Major Determinants of Literary Style". *Computers and the Humanities* 23, Nr. 4 (1. August 1989): 309–21. https://doi.org/10.1007/BF02176636.

———. „Word-Patterns and Story-Shapes: The Statistical Analysis of Narrative Style". *Literary and Linguistic Computing* 2, Nr. 2 (1. Januar 1987): 61–70. https://doi.org/10.1093/llc/2.2.61.

Dévière, Elisabeth. „Analyzing linguistic data: a practical introduction to statistics using R", Journal of Applied Statistics, Volume 36, Nr. Issue 4 (2009): 471–72. https://doi.org/10.1080/02664760802366783.

Dobrovský, J.: Ausführliches Lehrgebäude der böhmischen Sprache, zur gründlichen Erlernung derselben für Deutsche, zur vollkommenen Kenntniss für Böhmen. J. Herrl, 1809.

Dobrovský, J.: Lehrgebäude der Böhmischen Sprache: Zum Theile verkürzt, zum Theile umgearbeitet und vermehrt. Haase, 1819.

Eder, Maciej. „Does size matter? Authorship attribution, short samples, big problem", Digital Humanities 2010, 2010, 132–35.

———. „Does size matter? Authorship attribution, short samples, big problem". *computationalstylistics*, 2018. https://github.com/computationalstylistics/preprints/blob/master/Eder_Does_size_matter.pdf.

Eder, Maciej, Jan Rybicki, und Mike Kestemont. „Stylometry with R: a package for computational text analysis", R Journal, 8(1) (2016): 107–21.

Graves, Alex, und Jürgen Schmidhuber. „Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks". In *Advances in Neural Information Processing Systems 21*, herausgegeben von D. Koller, D. Schuurmans, Y. Bengio, und L. Bottou, 545–552. Curran Associates, Inc., 2009. http://papers.nips.cc/paper/3449-offline-handwriting-recognition-with-multidimensional-recurrent-neural-networks.pdf.

Hauger, J. Scott. „Reading Machines for the Blind: A Study of Federally Supported Technology Development and Innovation", 5. April 1995. https://vtechworks.lib.vt.edu/handle/10919/26321.

Hope, Jonathan. *The authorship of Shakespeare's plays : a socio-linguistic study*. Cambridge: Cambridge University Press, 1994.

Kahle, Philip, Sebastian Colutto, Günter Hackl, und Günter Mühlberger. „Transkribus - A Service Platform for Transcription, Recognition and Retrieval of Historical Documents". In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 04:19–24, 2017. https://doi.org/10.1109/ICDAR.2017.307.

Kay, Anthony. „Tesseract: an Open-Source Optical Character Recognition Engine | Linux Journal". Zugegriffen 29. November 2019. https://www.linuxjournal.com/article/9676.

Koppel, Moshe, Jonathan Schler, und Shlomo Argamon. „Computational Methods in Authorship Attribution". *Journal of the American Society for Information Science and Technology* 60, Nr. 1 (2009): 9–26. https://doi.org/10.1002/asi.20961.

Krajewski, Grzegorz, und Danielle Matthews. „R. H. Baayen, Analyzing linguistic data: A practical introduction to statistics using R.", Journal of Child Language, Issue 2, Nr. Volume 37 (2009): 465–70. https://doi.org/10.1017/S0305000909990080.

Lost, Christine. *Das Leben als Lehrtext: Lebensläufe aus der Herrenhuter Brüdergemeine / Christine Lost. Mit einem Geleitw. von Dietrich Meyer und einer Einf. von Ulrich Herrmann*. Baltmannsweiler: Schneider-Verl. Hohengehren, 2007.

Mettele, Gisela. „Weltbürgertum oder Gottesreich: die Herrnhuter Brüdergemeine als globale Gemeinschaft ; 1727 - 1857 / Gisela Mettele". Vandenhoeck & Ruprecht, 2009.

Meyer, Dietrich. *Geleitwort*, In : Christine Lost *Das Leben als Lehrtext: Lebensläufe aus der Herrenhuter Brüdergemeine / Christine Lost. Mit einem Geleitw. von Dietrich Meyer und einer Einf. von Ulrich Herrmann*. Baltmannsweiler: Schneider-Verl. Hohengehren, 2007.

Oakes, Michael P. *Literary Detective Work on the Computer*. Amsterdam: John Benjamins Publishing Company, 2014.

Zyl, Maryka van, und Yolande Botha. „Stylometry and characterisation in The Big Bang Theory", Literator - Journal of Literary Criticism, Comparative Linguistics and Literary Studies, 37(2) (2016). DOI: 10.4102/lit.v37i2.1282

## 9. Language Resource References

Archiv im Böhmischen Dorf. (2020). Czech CVs from the 18th-19th centuries. Signature marks A-IV-1 and A-IV-2, https://www.boehmischesdorf.de/archiv.html