Exploring the similarities and differences between VLM-driven and traditional OCR for Historical Swedish Data

Martin Johansson, Selma Waginder, Dana Dannélls
Department of Computer Science and Engineering,
Department of Swedish, multilingualism, language technology
Chalmers University of Technology, University of Gothenburg
stlmartin.j456@gmail.com, selmawaginder@gmail.com, dana.dannells@svenska.gu.se

Abstract

Recent Swedish OCR efforts rely primarily on traditional OCR methods, including deep CNN–LSTM hybrid neural networks and transformer-based models. Some approaches have also demonstrated the applicability of VLM-driven OCR to historical material. However, to date, no studies have examined in depth the performance of VLM-based OCR on historical Swedish sources. In this paper, we ask: How do transformers and VLMs differ in character- and word-level recognition performance across typefaces, and what qualitative differences can be observed in their error patterns? We show that fine-tuned versions of the Alibaba Cloud Qwen3-VL-8B-Instruct and Qwen3-VL-2B-Instruct, combined with a simple repetition-trimming step, outperform conventional OCR systems. Remaining errors are primarily attributable to challenges associated with the Blackletter typeface and formatting issues, such as missing or extra line breaks, characters, and spaces. Even when characters are correctly recognized, formatting inconsistencies can substantially increase transcription error rates.

Keywords: VLM, OCR, Historical newspapers

1. Introduction

Late Modern Swedish, the time period from late 18th century and 20th century, represents the final stage before the eventual standardisation of Swedish orthography. Similar to English (Baron, 2011), texts during this period are characterized by lack of fully established spelling norms, compounded by the absence of stable word lemmata, and morphological descriptions (lexicons) (Borin and Forsberg, 2011). In addition, older material is characterized by mixture of typeface (Blackletter, Antiqua) and low quality of the print. Finally, due to the changes in the language during this period, the texts are heterogeneous. For example, while some texts display a rich case system, others do not, and there is a large variation between text types, time periods, authors and scribes, as well as within texts, requiring manual close-reading of the material (Stymne et al., 2023).

Although digitized Swedish historical materials for late modern Swedish exist, the limited availability of gold-standard training data continues to constrain the development of accurate Optical Character Recognition (OCR) models, which extracts the pure text from the material., leaving this period comparatively low-resource.

Recent efforts of Swedish OCR rely on traditional OCR methods, including Deep CNN–LSTM hybrid neural networks (Brandt Skelbye and Dannélls, 2021) and transformers (Löfgren and Dannélls, 2024). Although the approach of Löfgren and Dannélls has shown remarkable improvements, their model could not locate all instances of a word in a text because of the limitations mentioned above, leaving much room for improvement. Vision Language Models (VLMs)-driven OCR, on the other hand, have been proven to outperform traditional text recognition models for numerous languages (Bao et al., 2025; Kolavi et al., 2025; Kim et al., 2025). However, to date, no studies have examined in depth the performance of VLM-driven OCR on historical Swedish material.

To our knowledge, this is the first study that examines the similarities and differences between traditional OCR models and VLM-driven OCR when tested on Swedish data. In this paper we ask: How do transformers and VLMs differ in their character and word recognition performance across typefaces, and what qualitative differences can be observed in their error patterns?

The novelty of our work is threefold: (1) we identify optimal configurations for adapting VLMs to Swedish as a use case; (2) we provide a systematic comparison of the CER and WER results achieved by state-of-the-art VLMs; and (3) we explore which fine-tuning approach is best suited to our specific use case.

2. Related Work

Adesam et al. (2019) found that the percentage of words found in Swedish dictionaries for modern newspaper material was approximately 80%, and Bouma and Adesam (2022) showed that there is a strong linear correlation between percentage of words found in dictionaries and a normalized word

Proceedings of The Fourth Workshop on the Role of Resources in the Age of Large Language Models (RESOURCEFUL 2026), pages 193–199 May 11, 2026. ©ELRA Language Resources Association (ELRA), 2026

Type	Number of Segments
Text	16,791
Images	191
Lists	234
Total	17,216

Table 1: Division between "Text", "Images" and "Lists" in the dataset. Lists include tables.

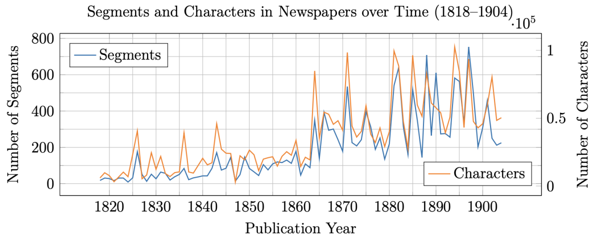


Figure 1: Combined visualization of number of segments and characters in newspaper pages from 1818 to 1904 according to our dataset.

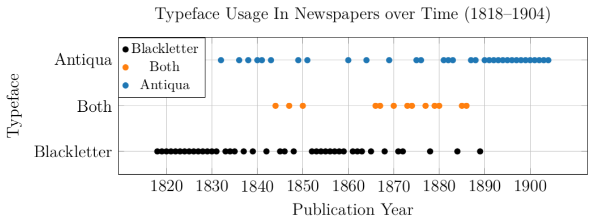


Figure 2: Typeface usage in newspapers over time ranging from 1818 to 1904.

error rate. This indicates that between 20-40% (depending on time period) of the OCR output words contain an error.

Dannélls et al. (2021) evaluated a two-OCR engine on a set of manually transcribed newspaper pages from 1818 to 2018. The two-OCR approach combined ABBYY FineReader (proprietary) (ABBYY, 2023) and Tesseract (open-source) (Smith, 2007) using a custom rule-based integration scheme and multiple period-specific word lists. However, neither the combined OCR system nor the incorporation of external word lists led to improvements in overall performance. In contrast, the baseline Tesseract model, used without any external word lists, outperformed all other system configurations, achieving a Character Error Rate (CER) of 11.86% and a Word Error Rate (WER) of 18.74%. Löfgren (2023) later demonstrated an improvement of the results for the same time-period by fine-tuning a post-OCR correction model based on the ByT5 byte-level Text-to-Text Transfer Transformer developed by Xue et al. (2022).

Likewise, recent approaches to VLM-driven OCR based on Vision–Language Models (VLMs) address character-level recognition tasks (Bai et al., 2025; Gemma Team et al., 2025). Owing to their generative nature, however, these models are prone to over-generating characters when encountering unknown glyphs or visually ambiguous input, a behavior commonly described as hallucination. Despite this limitation, VLMs have demonstrated strong overall performance across a range of benchmarks (Salla et al., 2025). Nevertheless, a systematic evaluation of their applicability to historical Swedish texts has not yet been conducted.

3. Data

Our dataset is the first half of the dataset prepared by Dannélls et al. (2021), consisting of 174 newspaper pages from a wide range of Swedish newspapers published between 1818 and 1904. Each page in the dataset was automatically divided into paragraph-level segments of varying length, which may be split due to changes in font or text size. A selection of pages and segments used as input is presented in the Appendix. The dataset comprises 17,216 segments, each with a corresponding ground-truth text file. Together, the ground truth files contain 3,149,860 characters. Most segments contain text, while a small fraction consists of images or lists (see Table 1).

Because segment size and length vary, the number of segments does not directly reflect the amount of text per newspaper. There is a close relationship though. This is illustrated in Figure 1, where segments (left axis) and characters (right axis) are compared. Segments typically correspond to sections or paragraphs, with more complex layouts producing more segments. Despite variation in segment length, the number of characters largely follows the same trend. In addition, we provide statistics about the division of typefaces throughout the years in Figure 2.

A manual inspection of the 10% of the most erroneous segments revealed approximately 150 segments with medium to large errors or ground-truth inconsistencies, ranging from cropped sentence beginnings to complete mismatches between image and text, or missing ground-truth text altogether. Most issues were fixable; however, five segments were deemed unfixable and were excluded from the project.

The dataset was randomly split at the segment level into 70% training, 15% validation, and 15% test sets. Training was used for fine-tuning, validation for early stopping, and testing for final evaluation. The test set includes 2,518 texts, 30 images, and 35 lists, reflecting the overall dataset distribution.

Model	Parameters	CER (%)	WER (%)
Qwen2.5-VL-7B-Instruct	7 B	3.73	13.99
Qwen3-VL-8B-Instruct	8 B	4.54	15.83
Qwen2.5-Omni-7B	7 B	5.11	15.15
Qwen2.5-VL-3B-Instruct	3 B	6.06	20.42
Qwen3-VL-4B-Instruct	4 B	8.14	23.12
Qwen2.5-Omni-3B	3 B	8.17	21.81
Qwen3-VL-2B-Instruct	2 B	13.42	26.76
Florence-2-large	0.8 B	17.27	44.00

Table 2: List of the stock models which were considered for fine-tuning. The models are sorted from lowest to highest CER.

4. Experiments and Results

All models selected for our experiments are open-weights VLMs meeting the following criteria:

1. compatibility with the vLLM inference engine (Kwon et al., 2023),
2. demonstrated community usage (measured by HuggingFace downloads),
3. have at most 32B parameters, and
4. compatibility with the developers' provided inference templates without major modification.

In total we evaluated 50 open-weights VLMs. Table 2 lists the CER and WER of the subset of models that were found to be strong enough relative to their size to be considered for fine-tuning. Priority was given to top-performing models within a given size class, while larger models were excluded when smaller counterparts achieved equal or better performance. Models that were either too large, too weak, or both were excluded from consideration. After some experimentation, this left us with the Qwen3-VL-8B-Instruct and Qwen3-VL-2B-Instruct, see Table 3.

Fine-tuning hyperparameters were selected based on prior findings (Hu et al., 2021) and preliminary experiments. Both the LoRA rank and scaling factor α were set to 16 for the Qwen3-VL-8B-Instruct and to 8 for the Qwen3-VL-2B-Instruct. A dropout rate of 0.1 was used for regularization, with no weight decay. Suitable learning rates and effective batch sizes were determined through exploratory runs on 5% and 20% subsets of the training data. Figure 3 and Figure 4 show the progress of the fine-tuning on the validation set after each epoch, which was done on a single Nvidia A100 80GB, using a batch size of 4.

VLMs occasionally produce repetitive outputs that loop until the maximum token limit is reached, a rare but significant source of errors. To mitigate this, a repetition-trimming function was developed to detect and remove repeated sequences at the end of generated text. It starts backwards from the very end of the generated text, and checks if there

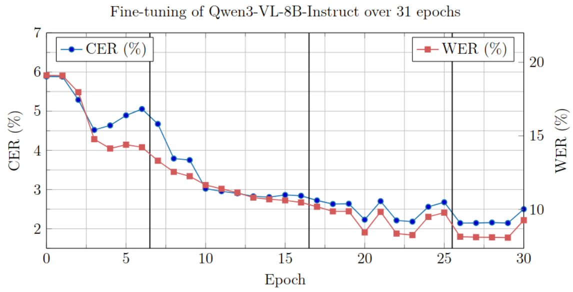


Figure 3: CER and WER achieved by each epoch on the validation set during the fine-tuning of the Qwen3-VL-8B-Instruct. Epoch 0 is the stock model. The vertical bars indicate where the learning rate was lowered.

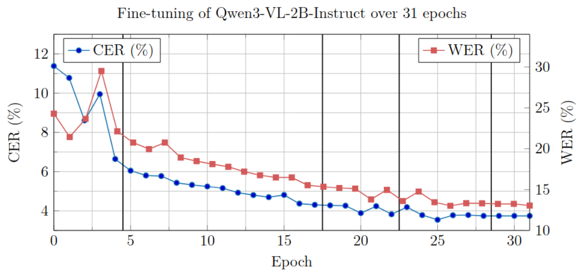


Figure 4: CER and WER achieved by each epoch on the validation set during the fine-tuning of the Qwen3-VL-2B-Instruct. Epoch 0 is the stock model. The vertical bars indicate where the learning rate was lowered.

is a repeating string of length $i \in 1, ..., I$, where I is a parameter called the maximum string length. The function also uses a parameter that sets a minimum number of characters that the whole repeating sequence has to contain in order to be detected. This ensures that no false positives are detected. Both parameters were set to 300 for all models, as these values performed well on validation runs during fine-tuning.

5. Analysis

5.1. Quantitative Analysis

Table 3 shows the performance of our VLM based OCR systems, and the corresponding performance of Abbyy FineReader ran by Dannélls et al. (2021) and Tesseract 5.5.1 as a reference point. The stock Qwen3-VL-8B-Instruct in combination with the repetition trimmer outperforms the 2021 FineReader by 47.8%. Table 4 shows how the CER of the VLMs compare with and without the use of the repetition trimmer.

We further examined the performance of the different models by typeface, Table 5, showing all the models perform better on texts written in Antiqua.

Model	CER	WER	Precision	Recall	F-score
Qwen3-VL-8B-Instruct (FT)	1.930	8.108	92.180	92.046	0.9211
Qwen3-VL-2B-Instruct (FT)	2.707	10.695	89.232	89.039	0.8914
Qwen3-VL-8B-Instruct (St)	3.198	12.453	88.388	87.349	0.8787
Qwen3-VL-2B-Instruct (St)	5.989	18.304	83.142	80.935	0.8202
FineReader (2021)	6.123	19.888	80.213	81.886	0.8104
Tesseract 5.5.1	7.089	21.882	78.195	80.161	0.7917

Table 3: The performance metrics achieved on the test set by our VLM based OCR systems (VLM + repetition trimmer) as well as the traditional OCR engines Abbyy FineReader and Tesseract 5.5.1. CER, WER, Precision and Recall are in percent. The fine-tuned models are labeled "(FT)" and the stock models are labeled "(St)".

	CER (%)		Reduction		
Model	wo.T.	w.T.	Abs.	Prop. (%)	N. Seg.
Qwen3-VL-8B-Instruct (FT)	3.565	1.930	1.635	45.863	4
Qwen3-VL-2B-Instruct (FT)	4.649	2.707	1.942	41.772	4
Qwen3-VL-8B-Instruct (St)	4.539	3.198	1.341	29.544	4
Qwen3-VL-2B-Instruct (St)	13.419	5.989	7.430	55.369	23

Table 4: Comparison between CER achieved by the models without trimmer and the whole VLM + Repetition Trimmer systems, as well as the absolute ("Abs.") and proportional ("Prop.") reduction in CER. The last column indicates how many of the 2581 segments in the testset that were affected by the repetition trimmer.

5.2. Qualitative Analysis

To assess the weaknesses of the fine-tuned Qwen3-VL-8B-Instruct model, the 30 segments with the highest CER are shown in Table 6. The results show a wide range of text types among high-error cases, with text quality classified as either "Good" or "Bad" based on visual degradation, and "Ordinary" referring to non-list, non-table text with minor layout variation. Both traditional OCR engines and VLMs perform worse on segments written in Blackletter than on those written in Antiqua. Several factors may explain this. First, Blackletter is an older typeface and is therefore likely underrepresented in training data. Second, Blackletter appears more frequently in earlier newspapers, which are often in poorer physical condition; however, poor performance is observed for both good- and bad-quality Blackletter segments, suggesting that degradation alone is not the primary cause. Third, older typefaces often reflect outdated spelling conventions, which may hinder recognition. Fourth, several Blackletter characters have visually similar forms, increasing transcription ambiguity.

Certain segments consistently produce high error rates across multiple models, contributing disproportionately to overall error. These cases show that even when characters are correctly recognized, text formatting can significantly hinder accurate transcription. Minor formatting differences—such as the use of ellipses instead of individual punctuation marks—are negligible for human readers but substantially affect CER and WER.

Model	Typeface	CER (%)	WER (%)
Qwen3-VL-8B-Instruct (FT)	Blackletter	3.332	14.726
	Antiqua	1.292	5.010
	Both	1.847	8.158
	Total	1.930	8.108
Qwen3-VL-2B-Instruct (FT)	Blackletter	4.646	19.373
	Antiqua	1.802	6.355
	Both	2.673	11.735
	Total	2.707	10.695
Qwen3-VL-8B-Instruct (St)	Blackletter	4.955	20.836
	Antiqua	2.381	8.375
	Both	3.157	13.052
	Total	3.198	12.453
Qwen3-VL-2B-Instruct (St)	Blackletter	8.562	28.820
	Antiqua	4.691	13.145
	Both	6.278	19.210
	Total	5.989	18.304
FineReader (2021)	Blackletter	10.254	35.274
	Antiqua	4.127	12.231
	Both	6.280	21.598
	Total	6.123	19.888
Tesseract 5.5.1	Blackletter	11.242	40.060
	Antiqua	5.033	12.475
	Both	7.417	25.175
	Total	7.089	21.882

Table 5: Performance of different models by typeface. Fine-tuned "(FT)", stock models "(St)". "Both" refers to both Blackletter and Antiqua segments, "Total" reports performance on the full test set.

Type	Typeface	Quality	N. Seg.
Ordinary	Blackletter	Bad	10
List	Antiqua	Good	6
Ordinary	Blackletter	Good	6
Ordinary	Antiqua	Good	3
Table	Antiqua	Good	2
Sideways table	Antiqua	Good	2
List	Blackletter	Good	1

Table 6: Manual classification of the 30 segments with the worst performance using fine-tuned Qwen3-VL-8B-Instruct. "Ordinary" refers to a normal text.

6. Conclusions

The VLM-based OCR models evaluated in this study outperform traditional OCR engines. Fine-tuned versions of the Qwen3-VL-8B-Instruct, and Qwen3-VL-2B-Instruct, combined with a simple repetition-trimming step, surpass conventional systems, with Qwen3-VL-8B-Instruct achieving 68.5% fewer errors than ABBYY FineReader, while remaining errors are mainly due to Blackletter typefaces and formatting inconsistencies.

Future work will focus on deploying VLM-based OCR systems in real-world digitization pipelines. Further gains may be achieved by combining the complementary strengths of different models.

7. Limitations

The scope of this paper is focused on 19th century newspapers, with material ranging from 1818 to 1904. This likely means that the performance of the fine-tuned models are poor on newspaper material much older than that. One of the reasons for the exclusion of 20th century newspapers is that copyright law still applies to most of them (the ones newer than 100 years). Since Swedish spelling and popular typefaces have remained mostly the same during the 20th century, performance might still end up being satisfactory on that material. Either way, this might not be very important since traditional OCR engines perform well on modern texts, at a much lower computational cost. The very most recent material will likely not be in need of digitization at all, since the material was already digital from the start.

8. Availability of Models and Code

The fine-tuned VLMs developed in this project are publicly and freely available on Hugging Face[1,2]. Some of the developed code and the dataset partitioning are available on GitHub[3].

9. Acknowledgments

This work has been funded by Språkbanken – jointly funded by its 10 partner institutions and the Swedish Research Council (2025–2028; project id 2023-00161).

10. Bibliographical References

ABBYY. 2023. Finereader PDF: Open, Read and Edit PDFs. Retrieved January 2, 2026.

Yvonne Adesam, Dana Dannélls, and Nina Tahmasebi. 2019. Exploring the Quality of the Digital Historical Newspaper Archive KubHist. In *4th Conference of The Association Digital Humanities in the Nordic Countries (DHN), Copenhagen, Denmark, March 5-8, 2019 / edited by Costanza Navarretta, Manex Agirrezabal, Bente Maegaard*, Språkbanken, University of Gothenburg, Sweden, Centre for Digital Humanities, University of Gothenburg, Sweden. CEUR Workshop Proceedings.

Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang, Shijie Wang, Jun Tang, Humen Zhong, Yuanzhi Zhu, Mingkun Yang, Zhaohai Li, Jianqiang Wan, Wei Ding, Zheren Fu, Yiheng Xu, Jiabo Ye, Xi Zhang, Tianbao Xie, Zesen Cheng, Hang Zhang, Zhibo Yang, Haiyang Xu, Junyang Lin, et al. 2025. Qwen2.5-vl technical report.

Xiaoyi Bao, Zhongqing Wang, Jinghang Gu, and Chu-Ren Huang. 2025. CalligraphicOCR for Chinese calligraphy recognition. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 4865–4877, Suzhou, China. Association for Computational Linguistics.

Alistair Baron. 2011. *Dealing with Spelling Variation in Early Modern English Texts*. Ph.d. thesis, Lancaster University.

Lars Borin and Markus Forsberg. 2011. A diachronic computational lexical resource for 800 years of swedish. In *Language technology for cultural heritage*, pages 41–61. Springer, Berlin.

Gerlof Bouma and Yvonne Adesam. 2022. Counting dirty words: The effect of OCR quality on token statistics in historical swedish corpora. In *Live and learn: Festschrift in honor of Lars Borin / Editors: Elena Volodina, Dana Dannélls, Aleksandrs Berdicevskis, Markus Forsberg, Shafqat Virk*, pages 17–24. University of Gothenburg, Gothenburg.

Molly Brandt Skelbye and Dana Dannélls. 2021. OCR processing of Swedish historical newspapers using deep hybrid CNN–LSTM networks. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, pages 190–198, Held Online. INCOMA Ltd.

Dana Dannélls, Lars Björk, Ove Dirdal, and Torsten Johansson. 2021. A Two-OCR Engine Method for Digitized Swedish Newspapers. Technical report, Linköping Electronic Conference Proceedings 180.

Gemma Team, Aishwarya Kamath, Johan Ferret, Shreya Pathak, Nino Vieillard, Ramona Merhej, Sarah Perrin, Tatiana Matejovicova, Alexandre Ramé, Morgane Rivière, Louis Rouillard, Thomas Mesnard, Geoffrey Cideron, Jean-Bastien Grill, Sabela Ramos, Edouard Yvinec, Michelle Casbon, Etienne Pot, Ivo Penchev, Gaël Liu, Francesco Visin, Kathleen Kenealy, Lucas Beyer, Xiaohai Zhai, Anton Tsitsulin, Robert Busa-Fekete, Alex Feng, Noveen Sachdeva, Benjamin Coleman, Yi Gao, Basil Mustafa, Iain Barr, Emilio Parisotto, David Tian, Matan

[1] `https://huggingface.co/J0hanski/Swe19centOCR-8B`

[2] `https://huggingface.co/J0hanski/Swe19centOCR-2B`

[3] `https://github.com/Martin31313/Swe19centOCR`

Eyal, Colin Cherry, Jan-Thorsten Peter, Danila Sinopalnikov, Surya Bhupatiraju, Rishabh Agarwal, Mehran Kazemi, Dan Malkin, Ravin Kumar, David Vilar, Idan Brusilovsky, Jiaming Luo, Andreas Steiner, et al. 2025. Gemma 3 technical report.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, et al. 2021. LoRA: Low-rank adaptation of large language models. *ICLR*, 1(2):3.

Seorin Kim, Julien Baudru, Wouter Ryckbosch, Hugues Bersini, and Vincent Ginis. 2025. Early evidence of how llms outperform traditional systems on ocr/htr tasks for historical records. *arXiv preprint arXiv:2501.11623*.

Adithya Kolavi, Samarth P, and Vyoman Jain. 2025. Nayana OCR: A scalable framework for document OCR in low-resource languages. In *Proceedings of the 1st Workshop on Language Models for Underserved Communities (LM4UC 2025)*, pages 86–103, Albuquerque, New Mexico. Association for Computational Linguistics.

Kungl. Biblioteket. 1835. *Wexjöbladet*, page 2. Libris-ID: 2831177.

Kungl. Biblioteket. 1861. *Umebladet*, page 4. Libris-ID: 2535033.

Kungl. Biblioteket. 1865. *Falköpings Tidning*, page 4. Libris-ID: 4112699.

Kungl. Biblioteket. 1888. *Göteborgs Handels- och Sjöfartstidning*, page 4. Libris-ID: 3678898.

Kungl. Biblioteket. 1889. *Hvad Nytt*, page 4. Libris-ID: 2732042.

Kungl. Biblioteket. 1891. *Hallandsposten*, page 2. Libris-ID: 4112716.

Woosuk Kwon, Zhuohan Li, Siyuan Zhuang, Ying Sheng, Lianmin Zheng, Cody Hao Yu, Joseph Gonzalez, Hao Zhang, and Ion Stoica. 2023. Efficient Memory Management for Large Language Model Serving with *PagedAttention*. SOSP '23, page 611–626, New York, NY, USA. Association for Computing Machinery.

Viktoria Löfgren and Dana Dannélls. 2024. Post-OCR correction of digitized Swedish newspapers with ByT5. In *Proceedings of the 8th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL 2024)*, pages 237–242, St. Julians, Malta. Association for Computational Linguistics.

Viktoria Löfgren. 2023. New tools for old news. Technical report, Department of Computer Science and Engineering, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden.

Rohit Kumar Salla, Manoj Saravanan, and Shrikar Reddy Kota. 2025. Beyond hallucinations: A composite score for measuring reliability in open-source large language models.

R. Smith. 2007. An overview of the tesseract OCR engine. In *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, volume 2, pages 629–633.

Sara Stymne, Carin Östman, and David Håkansson. 2023. Parser evaluation for analyzing Swedish 19th-20th century literature. In *Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)*, pages 335–346, Tórshavn, Faroe Islands. University of Tartu Library.

Linting Xue, Aditya Barua, Noah Constant, Rami Al-Rfou, Sharan Narang, Mihir Kale, Adam Roberts, and Colin Raffel. 2022. Byt5: Towards a Token-Free Future with Pre-trained Byte-to-Byte Models. *Transactions of the Association for Computational Linguistics*, 10:291–306.

11. Appendix

This section presents a selection of pages and segments used as input for the project. The pages were automatically segmented, and the segments in turn divide the pages into smaller images that serve as the actual input. Figures 5-7 show pages with added markings separating the text into segments and Figures 8-10 display possible segments.

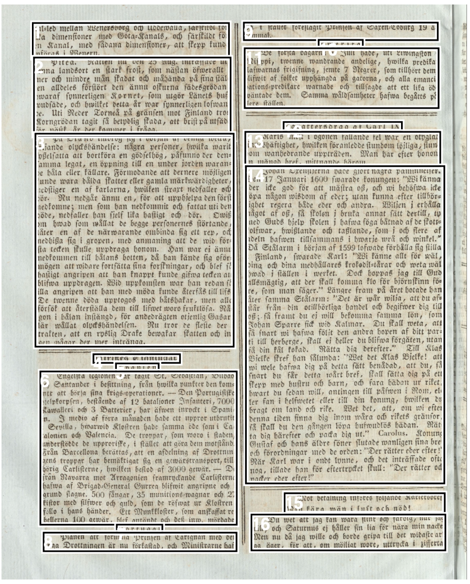


Figure 5: An excerpt from *Wexjöbladet*, 18th of September 1835 (Kungl. Biblioteket, 1835), with added markings to separate segments.

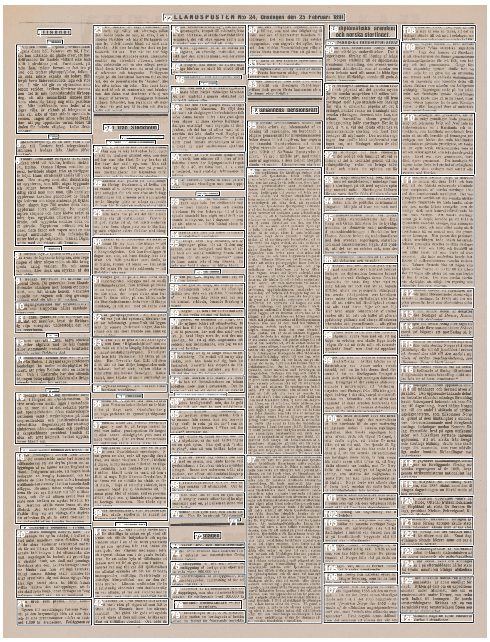


Figure 6: An excerpt from *Hallandsposten*, 25th of February 1891 (Kungl. Biblioteket, 1891), with added markings to separate segments.

Conrad Schildt

har härmed äran gifva tillkänna, att han nu, i likhet med föregående år, uti särskild lokal öppnat en

STÖRRE EXPOSITION

innehållande alla sorters Manufaktur- och Galanterivaror jemte Barnleksaker samt en mängd andra både nyttiga och till Julpresenter passande artiklar, om hvars i år mer än vanligt stora rikhaltighet och omvexlande beskaffenhet resp. kunder och vänner inbjudas att å utställningen taga närmare kännedom. De flesta varorna äro direkte inkomna från utlandet och allt säljes till så billiga priser som möjligt.

Umeå den 14 December 1861.

Conr. Schildt.

Figure 7: An excerpt from *Umebladet*, 14th of December 1861 (Kungl. Biblioteket, 1861), with added markings to separate segments.

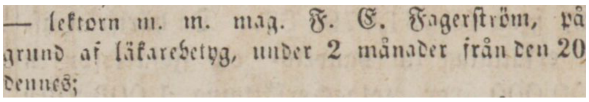
Andra våningen,

i f. d. Bronanderska huset vid Gustaf Adolfs torg, bestående af sju stora och välbelägna rum med kök, stor tambur och dithörande skåp, badrum, serveringsrum, pigkammare, handkammare, garderober, skafferi, veranda, rymlig och välförsedd källare, egen vind samt del i stor torkvind och brygghus m. m., uthyres att tillträda den 1 April 1889.

Våningen, som ej är lämplig för återuthyrning, innehafves nu af hr Otto Magnus och kan beses, efter tillsägelse hos värden, från kl. 12 till 2.

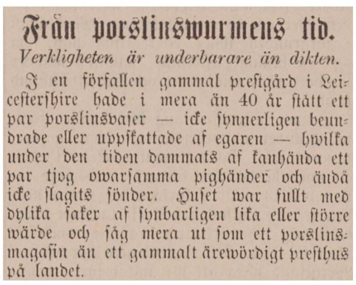
(9899.) 3

Figure 8: A segment from a page in *Göteborgs Handels- och Sjöfartstidning*, 12th of November 1888 (Kungl. Biblioteket, 1888).



— lektorn m. m. mag. F. E. Fagerström, på grund af läkarebetyg, under 2 månader från den 20 dennes;

Figure 9: A segment from a page in *Falköpings Tidning*, 23rd of June 1865 (Kungl. Biblioteket, 1865).



Från porslinswurmens tid.

Verkligheten är underbarare än dikten.

I en förfallen gammal prestgård i Leicestershire hade i mera än 40 år stått ett par porslinsvaser — icke synnerligen beundrade eller uppskattade af egaren — hwilka under den tiden dammats af kanhända ett par tjog owarsamma pighänder och ändå icke slagits sönder. Huset war fullt med dylika saker af synbarligen lika eller större wärde och såg mera ut som ett porslinsmagasin än ett gammalt ärewördigt presthus på landet.

Figure 10: A segment from a page in *Hvad Nytt*, 14th of June 1889 (Kungl. Biblioteket, 1889).