Al and Transkribus





Günter Mühlberger Digitisation and Digital Preservation group University of Innsbruck READ-COOP SCE

Agenda

- Archival documents
- AI and ML
- Examples
- Transkribus products
- Users
- READ-COOP SCE

Archival documents











Hundreds of shelf kilometres – billions of documents

Artificial intelligence? \rightarrow Machine Learning!

To my mind, a human intellectual competence must entail a set of skills of <u>problem solving</u> — enabling the individual to resolve genuine problems or difficulties that he or she encounters and, when appropriate, to create an effective product — and must also entail the potential for finding or creating problems — and thereby laying the groundwork for the acquisition of new knowledge. (Gardner, Howard: Frames of mind : the theory of multiple intelligences (1994) – Wikipedia "Intelligence" – June 2022) **Learning** is the process of acquiring new <u>understanding</u>, <u>knowledge</u>, <u>behaviors</u>, <u>skills</u>, <u>values</u>, attitudes, and <u>preferences</u>. The ability to learn is possessed by <u>humans</u>, <u>animals</u>, and some <u>machines</u>; there is also evidence for some kind of learning in certain <u>plants</u>. (*Wikipedia*, *Nov. 2021*) Types of machine learning

Supervised learning

- Labeled data are needed
 - As for Transkribus: a correct transcript is needed together with the line image
- Currently most wide-spread area of ML
- Well understood, though the neural network itself remains a black box
- Application areas
 - Text recognition
 - Speech recognition
 - Object recognition
 - •

Unsupervised learning

- No labelled data necessary
 - Input are "just" data
 - E.g. data, images, music, behaviour,...
- Often used to produce artificial training data for supervised learning
- Limited application areas
- Application areas
 - Clustering = grouping of input values of any kind
 - Producing "similar" output e.g. painting like Picasso, composing like Mozart, writing like Shakespear, etc...

Re-enforcement learning

- No labelled data necessary, but rules and targets
 - Computer is interacting with ist environment → direct feedback is used to improve/learn
- Typical application areas
 - Gaming, eSports,...
 - Robotics
- Most prominent example
 - AlphaZero is a generic <u>reinforcement learning</u> algorithm [...] that achieved superior results within a few hours, searching a thousand times fewer positions, given no domain knowledge except the rules. (*Wikipedia, June 2022*)

Self-supervised learning (transformers)

- No labelled data necessary, but learning is done in a similar way as with labelled data
 - Former President Donald Trump Says He Won't Come Back To Twitter: Pushing Truth Social (Forbes, June 2022)
 - XXX President Donald Trump Says...
 - Former XXX Donald Trump Says...
 - Former President XXX Trump Says...
 - Former President Donald XXX Says...
 - → Machine learns to make predictions on the missing words → in this way relationships among words (or entities in general) are discovered and learned
- Prominent examples: BERT, GPT3
 - Q&A, summarization, translation, coding,...
- Can be combined with supervised learning → e.g. for postprocessing or fine-tuning

ML for text recognition. Can you read this?

Dirc Jansz Schiouwer

> Op hHuijden den 18en. Decemb. @ 1638. Compareerde voor mij Hendrick Schaef Notaris pPub. etc. Trijn Barents huijsvr: van

16 Feb. Ao 1610 Ehrnveste, Achtbare vnd hochgelarte, Großgoönstige Herrn, Nehist meiner alzeit willig gefliessener Dienste erbiettung, kan E. E. g. Ich vnvormeldet nit lassen, Das vorschiener Jare Marx From image to text

Line detection

- Until 2017 "the" challenge
 - Complex non-standardized layout
 - Many different document types
 - Tables, notes, marginalia, etc...
 - Noise, borders, bleed through, strike through, drawings, separators, etc..
- Success story
 - New methods: machine learning
 - Representative and carefully selected data set (2000 pages from 12 archives and different epochs)
 - Service provider for creating "Ground Truth" (100.000+ lines manually drawn)
 - Scientific competition: ICDAR 2017
 - Research data are on ZENODO: cBAD
 - Best solution comes from University Rostock, Citlab Team (Tobias Grüning)
 - Grüning, Tobias, Neural text line extraction in historical documents : a two-stage clustering approach Rostock : Universität , 2018 https://doi.org/10.18453/rosdok_id00002427







mayon Wasterignung siter Sin Goilffe. 9) Sie ang In firighting Sinfag Hurtingan form and the states and for finger Jestalban money fanden States and for the farmer Sie aline 10) fir alla sug disfan there things Ing the disconting fir alla sug disfan thereage alma and gringende Hingshifun bafondaran Geright frand gaforan find. Jefterny Sin forflig gutternergen Chamfa som Big S protieratur in Stansbreeck wint flightig gotheredig. 11) Du Hartury wir in vinam Exemplane, maling in for yadartigh and a fill Grove Josef Alipsomaiter and a Mil berhel, and S Oliegund 1410 KR. Handsbaling Bricktin Junbuild Juff anfichanty Hagner my And baylow triging Register H. 235 for for its Aufletimatter & Surian Boundary Hoystille in Sief Hollow rig aufority interfferibet. Hon der Gerichets Rougher Son 18 Degitte Jenn 5. august 1900 Begenst my Official my Der mit _____ Urkunden und 7201//Ruguui/ Legelisirungestempel versehenen Urennes //Ruguui/ K. k. Bezirksgericht Kitzbühel m 17. Dagemale U. 191

mayon Harthingung other Sin Guilto 4) Sie ung das finifling diafag Hartungag formen lifen Den frifting desfallow ormen ffunden o nafun der Grietlianinge Hoffen, farmer die eter mig ver effetigenberneffing trigt Ing tek Anitighteritan enerlige migh tauf Chefelin Hispatisfan bafondarm Gerieftefhand gaforan find Jefting Sin parflief ginfambigen Ofamilia um diga mokineatur in Omnebruck und flingslief gathering. 11) Du Harburg wird in aiman Exemplane, malifism undertigh and arfield your Josef Anipomailer in Myberhel, and Durgund 1410 Somsbuck, an 2/ 10 total 1910 Jopf Cinfichenter Juker Theatsbally Million Lange 1/2 /2 /2 Rent beyloweligning Regiptor H. 235 for fan in Auflichmatter Sure Bourier Bring Highlich Siefe Halling signe forway meterffinitet Ton der yerichels tougle der ha generates Willowhele Jen 5 Mayust 1900 ind Byerst my lifty of my Der mit _____ Urkunden und /201 Legalisirungsetampal versehenen Urohme//IRugui) K. k. Bezirksgericht Kitzbühel m 17. Dagsmale UK 1910

Provence de l'equitre ville OBSERVATIONS MÉTÉORO	1 1 1 11	
	DLOGIQUES Termoire elle desquilhatoille	
Dinne le la Estingpa du mois de fancier 1935 (fa	uites a 6 heures.) Posse an Jaedin Bolanique el bala Observateur J. Denie.	
A BAROMETRE THERM S.A. PSYCHROMETRE ACTINOMETR	A GIROUETTE NURGES EAU	Ĩ
Date TEMP HAUTEUR HAUTEUR MINIMA EVAR. SEC HUMIDE HUMIDITE Q	DIRECTIMA FORCES NED. FORMES TOMBEE	
1 228 F370 7342 31 10 249 244 26 212 241	Calme D & ac - Bunilland Siger	
3 211 7161 737 200 11 11 207 16 201 300	R Bants de Blemillard	
1 240 7363 7342 340 03 240 244 25 245 246	te 1 8 die - De me	
5 230 7370 7343 24 61 230 246 46 244 245	Bala D # lin ha quelques gentes com	1000
7 18.2 7369 734,5 180 04 183 161 35 191 183	12, 2 10 14 -	
8 241 7369 F345 Dr 01 295 200 15 187 182	W I S Cis -	
10 218 2360 2342 205 0.9 21.9 21.4 36 18.9 201	8 1 6 Lis 19.0	
11 241 73×1 73×6 - 20 14 242 291 36 203 204	1 1 8 cic - Orage de 22 h. 20	1241
12 269 7361 1341 day at 14 441 441 441 441 441 441 441 445	No 1 2 Co - Annes de Requitan	1
19 11 110 THE LAN 11 251 Stat IS 221 Stat	Calme a 7 ca get Commence tointains	-
15 LIN FSKY ZULT SET 13 \$1,5 Sed IV LOT LOS	I 3 10 n 1.4 Phie de 4 h He ath	berne
17 126 7370 2343 245 13 129 243 45 221 213		
18 shit 73hg 2343 sell 1,2 shis sis shift 245	1. 2. 6 tes 260 Bance de Accuillance	4
19 440 7567 RAYN 18 45 411 249 13 206 208	The 2 10 th - This de 2h / w is the	Gene
21 230 2166 2116 211 21 21 21 21 21 21 21 21	No a 1 ac -	
22 141 7349 7141 34T 20 120 245 34 143 244	WNW 1 7 cie - Barnes de Barnikar	6
23 249 Files 2341 243 45 430 247 91 248 149 149	At a la la Ch Memilland leger	
25 223 F367 7340 Als al 243 249 76 245 246	dif 1 7 aic -	
26 24 7349 File - 202 0.9 223 241 92 243 241	Paris d 7 Cis - Barnes de Bromillaro	4
28 184 2120 FIRE 284 17 184 229 46 217 146	Walne O 7 as of panes de Browillans	6
29 135 Fite Fite flat as and and 15 the and	NNW 2 & au -	
The set also serve and and and the set	Sh 3 3 m m - Shit Former	
Assessment and Table 100 22.7 123 118 15 116 116	21 4 21 4	+
the ET MIN ABSOLUS DE T' II' CHI ON SHE SHE SHE SHE SHE	TOTAL LAU TOMBEL TIL ME	
48.17 48.7	NONDRE DE TOURS DE PLUIE 8	
FREQUENCE WARDES	ID DE DROUILLARD	
DES N., MNE. NE ENE E SE SE SE SEW SW WEW WWW NW NW WW COMME	O ZE GRÊLE	-

Text recognition

Training – Supervised learning

The process that allows a neural network to create a mathematical pathway from input to output is called training.

The process that allows a neural network to create a mathematical pathway from input to output is called training. We give the network training data consisting of input values and corresponding output values, and it applies a fixed mathematical procedure to these values.

The process that allows a neural network to create a mathematical pathway from input to output is called training. We give the network training data consisting of input values and corresponding output values, and it applies a fixed mathematical procedure to these values. The goal of this procedure is to gradually modify the network's weights such that the network will be able to calculate correct output values even with input data that it has never seen before. (How to Train a Basic Perceptron Neural Network November 24, 2019 by Robert Keim. https://www.allaboutcircuits.com/technicalarticles/how-to-train-a-basic-perceptron-neural-network/)



und Huge Veranftaltung/ Des Rapferl. General Feld. Marschall Lieutenants

innsve fign night nut ob his singsprifte. coul.





Fig. 3 Example of the sliding-window approach: Above, the text line to be analyzed is shown with the implicit sequencing direction assumed and some of the overlapping analysis windows superimposed. Below, the extracted image stripes are shown for part of the text line

und Huge Verankaltung/ Des Rapferl. General Feld. Marichall Lieutenants

und kluge Veranstaltung/des Käyserl.General Feld=Marschall Lieutnants

innove fin might nut ob his singsprifte. coul.

innere seyn mögte und ob die eingereichte. Druck.



יוחנן בן נורי וכי מה אכפת להם הע



Source: https://datasolut.com/neuronale-netzwerke-einfuehrung/

Character Error Rate

- Definition
 - Number of characters that need to be added, deleted or replaced to create a correct text from the incorrect text divided through the total number of characters multiplied by 100.
 - CER = Errors/Total Number of Characters*100
- Example
 - Das ist der richtige Satz mit 44 Buchstaben.
 - Das ist der falsche Satz mit 43 Buchstaben.
- In order to be able to form the word "richtige" from the word "falsche", we must
 - Replace f by r and a by i
 - Delete I and s
 - c and h can stay
 - Add t, i and g
 - e can stay
 - Replace 3 by 4
- 8 operations are needed to generate the correct sentence
 - Character Error Rate: 8/44*100= 18.18 Character Error Rate
 - Word Error Rate: 25% Word Error Rate (8 words in total, 2 need to be replaced).

Machine learning \rightarrow the engine finds a way how to connect the line image with the text of the line \rightarrow the model is the representation of this knowledge

Models

- Optical model
 - The core of the training process
 - Responsible for 90-95% of the results
 - Internal confidence levels for each character
 - Kind of knowledge included since neighbouring characters are learned as well
- Language model
 - Statistical distribution of characters is learned on basis of training data (and/or external resources)
 - Not a dictionary, but n-grams (up to 7 or 8) of characters
 - May improve results by 10%
- Base model
 - In order to speed up recognition a formerly trained model can be used as a base model
 - Training process is a kind of fine-tuning the base model to the actual training data
 - Especially important if there is no access to the training data or if training time shall be reduced

Choose a model

1-6/6 🙌 🖣 1 1 🕨	ы								
Name * <test_all_grenzbeschrei< th=""><th>Language german</th><th>Curator mirjam.el-at</th><th>Techno CITlab</th><th>Crea 29.1</th><th>nrOfW 213838</th><th>CER Tr 2.96%</th><th>CER Va 3.19%</th><th>Rating N/A</th><th>ID 29</th></test_all_grenzbeschrei<>	Language german	Curator mirjam.el-at	Techno CITlab	Crea 29.1	nrOfW 213838	CER Tr 2.96%	CER Va 3.19%	Rating N/A	ID 29
test3_grenzbeschreib	german	Mirjam.El-A	CITIab	13.1	181579	2.89%	5.19%	N/A	27
Kataster_Grenzbeschr	german	guenter.hac	CITIab	12.0	217847	2.23%	2.78%	N/A	26
Kataster German Kurr	german	guenter.hac	CITIab	09.0	584452	6.92%	5.23%	N/A	26
Kataster_Grenzbeschr	german	guenter.hac	CITIab	03.0	87435	2.71%	4.14%	N/A	25
₩Kataster_Grenzbeschr	german	guenter.hac	CITlab	03.0	87435	3.41%	4.19%	N/A	25
25 v grenzbe									8

lame:		Language:						
test3 arenzbeschreibung		german	derman					
escription:		Parameters:						
test model for description of l Fyrol, Austria	borders in 1850 -1900	Nr. of Epochs 200						
ogumant Tunai Handuuittan		v Chou advanced exemptor						
s of Words		Nr. of Lines:						
. or words.		24920						
		24520						
sibility: Private Model								
ating: -								
E Save	Show Train Set	Show Validation Set Show Charac	ters					
100% 90% - 40% - 50% - 50% - 40% - 30%	Lear	ning Curve						
20% 10% 0% 0 10 20 30 4	10 50 60 70 80 90	100 110 120 130 140 150 160 170 180 19 Epochs	0 200					
20% 10% 0% 10 20 30 4	10 50 60 70 80 90 	100 110 120 130 140 150 160 170 180 19 Epochs n — CER Validation	0 200					

 Examples for HTR results
Konzilsprotokolle

- University Archive Greifswald
 - Dirk Alvermann, German Science Funds Project, 2019-2020
 - Digitisation and recognition of court documents from 1570 to 1800
 - About 240,000 pages
- Documents
 - Challenging, hard to read
 - 800+ writers
 - Most parts are written in standard script, however, several parts in concept script
 - Kurrent and Latin script
 - Abbreviations
- Several models trained
 - Many writers, but every writer represented in the set
 - One example model: training data: 1432 pages = 226,717 words
 - CER on validation set: 3,31 (training set: 7,1), 1000 iterations

× **Version Comparator** Show line numbers Aan An den <mark>Sfertzo</mark> hertzog von Braudschwiegk Braunschwiegk wegen der Stadt Brauschwiegk Braunschwiegk Hochwurdiger Hochwürdiger vnd Durchleuchtiger Ha Dürchleüchtiger Hoch geburte geborner Furst Fürst, E f g sein vn sere ere vnderthenige vnd allezeit <mark>gehorsamb</mark>gehorsamp dienste zuuor Gnediger zuuuor, Gendiger Herr, Waß E. f.g. fg vnß-vns wegen der Speen vnd Mißhellig ketten, keiten, so zwischen E.e. f. g. vnd Vnterthanen hur Burgermeister, Rath, vnd Gemeine derselben Erb: eit Landttstadt sent Landthstadt Brunßwigk ytes m eingerißen, so woll will hiebeuor alß itzr itzer gnediglich zuegeschickt vnd mitgeteilet, haben wir vom Botten mit geburender gebürender Reuerentz vns ehren woll empfangen. Vnd thuen derowegen iegen E f. g. f g vnß vnderthenigk-vnderthenigk lich vnd vleißig bedanken, mit fernerm gernerns erbieten solches vmb dieselbigen in Vnterthenigkeit Zuuerdienen, Vndweil Vnd weil e f. g. daneben ahn vnß-vns gnediglich vorheren vehere solche abgedruckte Vrkunde zu <mark>zeng deng</mark>untergranting ende ense zuuerlesen, daß wir so viel darauß zu befinden, daß E-E. f. g. vnd dereselben derselben hoch lobliche Vorehtern lebliche Voreltern vnd vorfharen, Vorfharenn, ob gedachten E. f. g. Vnterthanen niemalen Summer daß geringste, welches E f g nicht auß hoher Landtsfurstlicher landtsfurstlicher Oberigkeit altem herkommen, herkommen vnd vndisputirlichenn vndispütirlichenn Siegell vnd briefen kundtlich befuegt befuegtt gewesen, vnd noch befugt sein, angemutet infait Quinwith gunn

CER = 5,04 WER: 21,63

Arabic test – Research Library Gotha

- One document
 - Nice Arabic handwriting, homogenous
 - Simple layout
- Ground truth
 - Training set: 72 pages = 17,327 words
 - Validation set: 2 pages = 506 words
- Training
 - HTR+ (University of Rostock), 200 epochs, Tensorflow implementation
 - Standard parameters, no tuning
- Results
 - CER on validation set = 3,51 (without dictionary)
 - Shows improvement very well: 13,7% with the first version of the HTR engine

	HTR 'MS A	.2654 M1+' – 🗖 🗙
Name:		Language:
MS A2654 M1+		Arabic
Description:		Parameters:
Retraining	^	Nr. of Epochs 200
	~	Show advanced parameters
Nr. of Words:		Nr. of Lines:
17327		1451
Visibility: Private Model		×
E Save	Show Train Set	Show Validation Set Show Characters
¢ 100% 1	Learnin	ng Curve
B 0% B 0% C	40 50 60 70 80 90	100 110 120 130 140 150 160 170 180 190 200 Epochs CER Validation
CER on Train Set: 0.27%	C	ER on Validation Set: 3.51%

امرالف لام ، فعال الملك لقد تم حادى عشريها ، فامر للك اجصا الفلام فاحصروه مخ التنتوا الوزرا وقالوا له ماردى الاصل يتج لك طبع والحياء بعدهذا اليم ترجوالنرج، فتاك لمم نغم، لاناس قادر باني بالنج فالشدة ، والحياه فريسط • سنا الحالة العامة بعن الما الما الم الم الم الم الم الم الم فتال لذاللك وكيف كان خباليسير فتال لذالغ لام ايها الملك ذكوا انتكاف سلكا وكان لذقصراء وكان ذلك المصر مشرقاع الحسن وكان فالحس جلعبو على ولم زالطل الزجزاس ويتول ، بايزوج، ويب فيج عن فغضب للك في ذاتد، فتالللك هذا احق بدالنج مزذبد، فسال للك وقال في هذا الحسر مجموس قالوالة ان محبور على دم فامرالملك باحصار لسمرالح بين درديه وقال لت العق باقليل لمتلكيف ترجي خلاهر نبسك وذ نبك عظيم فاننا معجاعة ليتتلوه خاج البلد وكان الوقت ليلافاخده والجند وخوجا بباخاج البلد فلماالدواان يتتلع واذابح لمسقد طلعواعليهم فسلوا سيوفم على تجند فتركوا اليسير وهديوا وههاليسين فزخون والتتلفغاص الرية فاحسر برج الأفى لدجله ، فطلع عليد سبع مخرف المنظر فخطف الحل واف ببالى وضع وحط وقلع شي وغط بعا الجل وانصرف لسبع

Version Comparator

×

Show line numbers

امر الغلام ، فقال الملك لقدتم خادى لقد تم حادى عسّر يومًا ، فامر الملك باحضار الغلام فاحضروه ، ثم التفتوا الوزرا وقالوا لهُ ، يا ردى الاصل بقى لك طمع في الحياه بعد هذا اليوم ترجو الفرح الفرج، فقال لهم نعم ، لإن الله قادر **باتي ياتي ب**الفرج في السّدة ، والحياه في وسط السَّدة كما إلى للرجل البير اليسير وكيف نجاه الله تعالى وفرج عنه ، فقال لمله الملك وكيف كان خبر اليسير ، فقال لمله الغلام ايها الملك ذكروا لنفانه كان ملكًا وكان لهُ مَصرًا ، وكان ذلك القصر مسرفًا على الحبس ، وكان في الحبس رجل محبوس على دم ولم يزال يطلب الفرح الفرج من الله ويقول ، يا من فرجه قريب فرح فرج على فغضب الملك في دائه ذائه ، فقال الملك هذا احمق يريد الفرح الفرج من ذنبه ، فسال الملك وقال من في هذا الحبس محبوس ، قالوا لهُ انه محبوس على <mark>نمر دم</mark> ، فامر الملك باحضار اليسير الى بين يديه وقال لهُ ، يا احمق يا قليل العقل كيف ترجو خلاص نفسك وذنبك عظيم ، فانفده مع جماعةٍ ليقتلوه خارج البلد وكان الوقت ليدايلًا فاخدوه الجند وخرجوا فخرجوا به خارج البلد ، فلما ارادوا ان يقتلوه واذا بحراميه قد طلعوا عليهم ، فسلوا <mark>سبوقهم</mark>سيوفهم على الجند فتركوا اليسير وهريوا و هرب اليسير ، فمن خوفه من القتل فغاص في البرية فما حس <mark>يزوج بروحه</mark> للإلاً في الدجله ، فطلع عليه سبع مخوف المنظر فخطف الرجل واتي يهِ الى موضع وحظه وحطه وقلع سَجر وعطى وغطى بها الرجل ، وانصرف السبع يطلب

Printed text - newspapers

- Wiener Diarium in cooperation with Austrian Academy of Science
- Newspaper from 18th century
- Bitonal scans with partly challenging quality
- Training set: 179.997 words or 345 pages
- CER on test set = 0,81
- WER on test set = 3,02

Mittags haben Allerhöchst befagt Se. Majestät der Rauser unter der gewöhnlichen Aufwartung des Päpstl. Hen. Muntii, und der Entbinding herannahete, fo wurden beimegen fogleich bey Sof aue nohtwendige Bierans Raltungen vorgekehret, auch Darumen sowol in benen Kalferl. Sof . tapellen, als in deuen albiefigen Pfare und mehrern Rirden bas Sochwürdigste Attars . Sacrament unter 314 lanf Deren alhiefigen getreuesten Inwohnern ju Erbittung einer bochit . beglutten Entbin. bung für Ihro Raifert. Ronigl. Maleflat aus. gejeget : auch hatte ein jahireicher hoher Ubel nach Sofe fich erhoben; morauf fodann um 9. Ubr Abends unfere Maermildelle Landes. Dut. ter mit einer iconen und gejunden Erg. Serlogin, BOit Lob, gludlich in allgemeiner Freus be deren unterthänigsten getreuen inwohnern alhtefiger Raiferl. Ronigl. Saupt ; und Refibeng . Stadt 2Bien, entbunden morden : auch feund viele Botten ju Pferd und ju Fuß von bier abgeschicket worden, Dieje gludliche Beitung auch auf dem Land tund ju machen. Fol-

Raifers, ber gesamten Raiserl. Königl. jungen Sperischaft, und des in größter Anzahl und prächtigster Gala, welche auch durch die a. solgende Lagen gedautet, erschienenen hoben Idels verrichtet.

Mittags haben Muerhöchst befagt Se. Majestät der Raiser unter der gewöhnlichen Aufwartung des Päpill. Sin. Muntii, und des sahlreichen hohen Adels offentlich gespeiset. Ubends haben Se. Ercen. (Lit.) der Sere Alo Graf von Bathiani denen Durchlasten iungen Serischaften samt einigen iungen Dames und Cavalieren in seiner Wohnung ber Hof einen Ball und prächtiges Abendmahl gegeben. Die Theatra nächst der Kaiserl. Burg und dem Kärntner Thor waren des Abeuds prächtig illuminiret, und seynd alle Stadt-Thore Sperisfrey gewelen.

Frentag , Den 6. Diefes

Rach dem gewöhnlichen GOttes . bienfi-in

Version Comparator	- • ×
Entbindung herannahete, so wurden deswe=	^
gen sogleich bev Hof alle nohtwendige Veran=	
staltungen vorgekehret, auch darumen sowol in	
denen Kaiserl. Hof=kapellen, als in denen	
alhiesigen Pfarr= und mehrern Kirchen das	
Hochwürdigste Attars=Sacrament-Altars=Sacrament unter Zu=	
lauf deren alhiesigen getreuesten Jnwohnern	
zu Erbittung einer höchst=beglükten Entbin=	
dung für Jhro Kaiserl. Königl. Majestät aus=	
gesetzet: auch hatte ein zahlreicher hoher Adel	
nach Hofe sich erhoben; worauf sodann um 9.	
Uhr Abends unsere Allermildeste Landes=Mut=	
ter mit einer schönen und gesunden Erz=Her=	
zogin, GOtt Lob, glücklich zu allgemeiner Freu=	
de deren unterthänigsten getreuen Jnwohnern	
alhiesiger Kaiserl. Königl. Haupt= und Resi=	
denz=Stadt Wien, entbunden worden: auch	
seynd viele Botten zu Pferd und zu Fuß von	
hier abgeschicket worden, diese glückliche Zei=	
tung auch auf dem Land kund zu machen. Fol=	
genden Tag darauf, nämlich	

Donerlag, Donnerstag, den 3.-5. dieses Jst der Tauf=Actus um 11. Uhr Vormittag

For this page: CER = 0,6 / WER = 3,0% (without dictionary)

V.

Inunor din Gammina Jazu wangflight wardan, fobila ig Jar, 189. um, Inf in Inni Dort on Andersfagnin, wo jødar Grubnigan. 1814, Ganne dub Full forthin, und brungten, allershi ga/indal aufz zu unfunni, din finsvillung galvotan warda, daß dain Mang ofun Snusilligung In6 Syntznu in Sub Dost aufgruomenn usnædnu durfn, Mnin næfnb Snynfænu unardn abga/llagar, and Inx gnunninn din Untrofallung Inx Lindny zurredant. Drynynn wurdn Inm G: Landraff v: Michlenfels, Inmbf: von Hochwaechter und mis aufgabragan, ninn Djulzan 2 ordnung fins dab dorf desaudbfagne zu untersighen.

CER = 2,2% / WER = 10,3% (with dictionary)



Version Comparator
sur ce par duvire produire, et les deposicions d'iceulx disposicions d'iceulx publient ou facent publier par tous les lieux qu'il qu'il appartendra et à perpetuele moire, memoire, à la consernacion conservacion de leurs dis biens, possessions et heritages. Car ainsi le voulons nous, pour consideracion des choses dessus dictes. Et que ce soit ferme et estable à tous jours, nous avons fait [mettre] nostre seel à ces lettres lettres. Sauf nostre droit en autres choses et l'autrui l'autrui en toutes. Donné au Bois-bois de Vinconnes, l'an Vincennes, l'an de grace mil ccc. lxxiilxxii, et de nostre regne le ixe
Par le roy, roy en
ses <mark>requestes requestes.</mark> Hennequin. Gt dederi G. de Seris.

CER = 6,02 / WER = 19,6 (without dictionary)

Observations

- One hand
 - Take pages from one writer
 - Satisfying results from 10.000 words (or even below) onwards
 - 2-5% Character Error Rate can be achieved with modest effort
- Several hands but all hands are known and go into one training set
 - Take e.g. 20 pages from each writer as a start
 - Satisfying results need e.g. 5000 words per hand
 - 3-7% Character Error Rate can be achieved with modest effort
- Many hands and many of them do not go into the training set
 - Take representative sample from the overall collection
 - Satisfying results from 250.000 words onwards, robust models have more than 1 mill. words
 - 5-8% Character Error Rate

What you can expect

- Language model may improve results by about 10% often more improvement in WER than CER
- Large models can be used as base model in order to speed up the training however training from scratch may in many cases provide the best results
- Doubling the amount of training data improves recognition by ca. 20%
- Below 10% Character Error Rate manual transcription goes faster, below 5% users benefit greatly
- Image quality has less influence on results than expected → images from microfilm and smartphones show surprising good results
- For the next years it is to expect that specialized models will remain necessary

Transkribus products

Services and products

- Transkribus AI public models
- Transkribus Lite
- Transkribus Expert Client
- read&search
- Transkribus Processing API
- ScanTent & DocScan
- citizen&science
- tables&forms

Transkribus AI - Public models

X R Members of REA 🗴 🚫 Transkribus im A 🗴 🚫 Transkribus im A 🗴 💪 mediāvistik - Go 🗴 🖓 philipperemy/na 🗴 📑 Surnames Mean 🗴 🥘 Public Models in 🗴 🌀 councal minute: 🗙 🕇 readcoop.eu/transkribus/public-models/ 6 \$ $\leftarrow \rightarrow C$ $\hat{\mathbf{n}}$ 🗯 🕜 🔢 Apps 🔟 Kalender 🔯 PartPortal 🔒 Analytics 🔡 Dashboard | Infinity 📙 Espacenet – Sucher... 🥥 Usersnap 📑 Transkribus Users »... 🛅 Leseliste READ Transkribus 🗸 ScanTent read&search The COOP 🖌 Resources 🗸 EN 🔻 Sign in co.op Public Al models in Meltie van der Does huigt V: **Transkribus** acren fannetje Explore more than 50 publicly available AI models our van for Transkribus. Have a look at this overview of the publicly available models in Transkribus we offer so far. For every model you will find a short description of the training material, which languages the model 87 can be useful for and who has created and trained it. We are working on making more and more Models currently available models available for Transkribus users, so they can benefit from the network effect and save work and time. **Featured Models** - Gebiedende een eeuwigh filenti HTR+ your going without uninglis onwaerachtige diffamatien van de Heerer Saf minne fington day in Staten van Utrecht, &c., den laetsten Mey mugu via findus fininga 43 ID: 29820 ID: 29418 TRANSKRIBUS GERMAN KURRENT M2 - HANDWRITTEN TRANSKRIBUS PRINT 0.3 - PRINT, TYPEWRITTEN Transkribus German Kurrent Transkribus Print Multi-Language By: Transkribus Team By: Transkribus Team





Transkribus Lite



Transkribus Lite

- All basic features
 - Upload Images and PDFs
 - User Management
 - Segmentation
 - Transcription and virtual keyboard
 - Annotation
 - Text recognition with public models
 - Training private models
 - Full-text search
 - Export

Transkribus Expert Client

🕼 Transkribus Expert Client v1.18.1 (02_03_2022_11:33), Loaded doc: TRAINING_VALIDATION_SET_Transkribus Polish Model 1, ID: 935290, Page 1, file: SKM_30819072309350_0079.jpg [Image Meta Info: (Resolution... — 🗇 🗘 🗙

			🥒 L	.ogou	t guei	nter			
	🖻 D	ocum	ent			鶕	Find		
	Docur	ment I	Manager			ኞ User	Manager		
	D	Versio	ns				Jobs		
Recent	Docum	nents		~		💰 Use	r activity		
ollectio	ons:		tS G	round	Truth (1, Owner)		Co	ol-I[
Docum	ents H	TR Mo	odel Data					-	
1-100	/ 236	14 4	1 3	► N	0 0	🕞 🖉 🐚 🖆	ി ↓ Doc	-ID	
ID	Title				Pa	Uploader	Uploade	ed	^
94	TRAIN	ING_V		DN	1	philip.ka	Wed Ma	ir	
94	TRAIN	ING_V	ALIDATIC	DN	1	philip.ka	Wed Ma	ır	
94	TRAIN	ING_V	ALIDATIC	DN	1	philip	Wed Ma	ir	
94	TRAIN	ING_V	ALIDATIC)N	1	philip	Wed Ma	ır	
93	TRAIN	ING_	VALIDAT	10	6	sebastian	Fri Mar		
93	doc				1	philip	Fri Feb 2	2	
91	TRAIN	ING_V	ALIDATIC)N	1	guenter	Sun Feb		
80	TRAIN	ING_V	ALIDATIC)N	46	guenter	Sat Oct	0	
79	TRAIN	ING_V	ALIDATIC)N	1	f.krull@r	Wed Oc	t	
73	Abbyy	2Page	e_omitNo	Со	2	guenter	Thu Jul 2	2	
66	t2i_exa	mple_	reichsger	richt	3	guenter	Fri Apr 3	3	
66	t2i_exa	mple_	reichsger	richt	3	guenter	Fri Apr 3	3	
66	t2i_exa	mple_	reichsger	richt	3	guenter	Fri Apr 3	3	
65	Sampl	e_235	_NZ_table	es	46	guenter	Fri Apr 1	۱	
65	Morav	ské_n	oviny 2		1	guenter	Tue Apr		
65	Morav	ské n	ovinv 1		1	auenter	Tue Apr		~
								>	

TR		
L BL W	Kapitan statku za pôžno dostnegi	
•	niebezpievzenstwo. Ognomna fala nineta na	
H V	politad.	
	W jeonej visiti mingrunde structure restry nadziei wickzar vozszalute morze. Na pokład	
	mineta ogromna fala	
	yonek chiałby zostać inżynierem. Zawód	-
1-1 89-01-27		^
2-1 Kapitan stat <mark>ku za</mark> 2-2 niebezpieczeństv	a późno dostrzegł wo. Ogromna fała runęła na	
2-3 pokład.		
2-4 chwili marynarze	stracili resztę	
2-6 nadziei widząc ro	ozszałałe morze. Na poklad	
		~

Transkribus Expert Client

- Many additional features compared to Lite
 - Text2Image matching
 - Tag search
 - Structural tags
 - Structure recognition (P2PaLa)
 - Training features for Osmanic (transliteration)
 - Expanded export
 - Convenience features
 - FTP Upload
 - DFG Viewer upload
 - IIIF Upload
 - Table2textregion, row detection
 - Copy of layout
 - ...

read&search



Great projects that have been using read&search







Datenschutz

Impressum

Societ Kalender - Woche X Industrieller 3D-Scanner | X Startseite - Grundbuchbla X Suche - Projekt historisch X Read&Search - Amsterda X + ~ D X C Industrieller 3D-Scanner | X Startseite - Grundbuchbla X Suche - Projekt historisch X Read&Search - Amsterda X + ~ D X C Industrieller 3D-Scanner | X Startseite - Grundbuchbla X Iso Suche - Projekt historisch X Read&Search - Amsterda X + ~ D X C Industrieller 3D-Scanner | X Iso Startseite - Grundbuchbla X Iso Suche - Projekt historisch X Iso Read&Search - Amsterda X + ~ Industrieller 3D-Scanner | X Kalender PartPortal Industrieller 3D-Scanner | X Iso Suche - Sucher... Iso Usersnap Transkribus Users >... Iso Suche - Projekt historisch X Iso Read&Search - Amsterda X + Industrieller 3D-Scanner | X Kalender PartPortal Industrieller 3D-Scanner | X Iso Suche - Sucher... Iso Usersnap Transkribus Users >... Iso Suche - Sucher - Sucher... Iso Usersnap Iso Suche - Sucher >... Iso Suche - Sucher >...

X Gemeente X Amsterdam Bladeren Zoeken X Stadsarchief

Taal 🔻 Over

Amsterdam City Archives

Doorzoek hier voor het eerst in detail honderdduizenden handgeschreven bladzijden uit de Amsterdamse notariële archieven van de 17e en 18e eeuw. Deze collectie bevat de resultaten van het project Crowd Leert Computer Lezen, waarin vrijwilligers transcripties maakten en corrigeerden om zo computermodellen te trainen in Handwritten Text Recognition (HTR): het zelfstandig ontcijferen van de historische handschriften. De set bevat zowel door de mens gecontroleerde transcripties als de automatisch gegenereerde HTR.

Crowd Leert Computer Lezen (2018-) is een project van Alle Amsterdamse Akten van het Stadsarchief Amsterdam.

Alle Amsterdamse Akten : https://alleamsterdamseakten.nl/

Amsterdam.





ito need

Tanilla

ran dexax

Afdruk Gegevensbescherming





Dokumente Suche

Über Sprache 🔻



Bei den Grundbuchblättern handelt es sich um Personalunterlagen aus dem Militärwesen, genauer gesagt der Armee der Habsburgermonarchie. Über jeden in der Grafschaft Tirol zum Wehrdienst eingezogenen Mann wurde ein Grundbuchblatt angelegt. Darin ist die militärische Laufbahn des Soldaten dokumentiert: seine aktive Wehrdienstzeit (Präsenzdienst), seine Waffenübungen als Reservist, seine Beförderungen und Auszeichnungen, seine Kriegsdienstleistung während des Ersten Weltkriegs etc..

Das Tiroler Landesarchiv besitzt rund 46.000 Grundbuchblätter von Wehrpflichtigen der k. u. k. Armee aus dem Bereich des heutigen Bundeslandes Tirol sowie rund 38.000 Grundbuchblätter von Wehrpflichtigen der k. u. k. Armee aus dem heutigen Südtirol und dem Trentino (Welschtirol). Letzterer Bestand weist allerdings Lücken auf.





Impressum Datenschutz Transkribus Processing API

C A I transkribus.eu/processing/swagger/ ender PartPortal Analytics Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infinity Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infinity E Espacenet – Sucher Duscription Dashboard Infinity Dashboard Infi	C A matylice of an transkribus.eu/processing/swagger/ C A matylice Deshboard Infinity Espacenet - Sucher Ousersnap Transkribus Users	pogle Kalender - Woche 🗙 🜍 (1) Neue Nachrichten! 🗙 🔤 Startseite - Grundbuchblä 🗙 🔤 Suche - Projekt historisch 🗙 ℝ Su	Swagger UI X	+ ~	_	Ľ
nder RAP PartPortal Analytics Reprocessing V1/openapi.json The API accepts single images as base64-encoded string or by http(s) URLs in a POST request and processes them using Transkribus HTR technology. Each processing request's response includes a processId that can be used to poll its status and retrieve the result. You can apply any of your HTR models trained in Transkribus as well as all public models by setting its htrId in the ConfigRepresentation section of the initial POST request. A list of public models including their ID can be found on the READ COOP website. The API accepts the image MIME-types image/jpeg , image/iff and image/png and file sizes up to 20 megabyte. Terms of service	<pre>der PartPortal Analytics PartPortal Analytics PartPortal Analytics PartPortal Analytics PartPortal Analytics PartPortal PartPortPortal PartPortal PartPortPartPortal PartPortal PartPortal PartPo</pre>	C ☆ a transkribus.eu/processing/swagger/		Q 🖻 ☆	*	
Transkribus Processing API (1.9-beta (A83) https://transkribus.eu/processing/v1/openapi.json The API accepts single images as base64-encoded string or by http(s) URLs in a POST request and processes them using Transkribus HTR technology. Each processing request's response includes a processId that can be used to poll its status and retrieve the result. You can apply any of your HTR models trained in Transkribus as well as all public models by setting its htrId in the ConfigRepresentation section of the initial POST request. A list of public models including their ID can be found on the READ COOP website. The API accepts the image MIME-types image/jpeg , image/tiff and image/png and file sizes up to 20 megabyte. Terms of service	Severs Authorize	ender 🔝 PartPortal 🔒 Analytics 👬 Dashboard Infinity 🗧 Espacenet – Sucher 🥥 Usersnap 📑 Transkribus Users ›				
	Servers /processing/v1 v	Transkribus Processing P	h processing request's response inc ne inital POST request. A list of publ	cludes a processId	that can b eir ID can	be
Retrieve processing result		Terms of service Contact the developer		Authoriz	ze 🔒	
Retrieve processing result GET /processed/{processId}/alto Get the processing result, identified by its processId, as ALTO XML	GET /processes/{processId}/alto Get the processing result, identified by its processId, as ALTO XML	Terms of service Contact the developer Servers /processing/v1 v Retrieve processing result GET /processes/{processId}/alto Get the processing result, identified by its processId, as ALTO XML		Authoriz	ze	
Retrieve processing result GET /processes/{processId}/alto Get /processes/{processId}/alto Get /processes/{processId}/page Get /processes/{processId}/page Get /processes/{processId}/page Get /processes/{processId}/page	GET /processes/{processId}/alto Get the processing result, identified by its processId, as ALTO XML GET /processes/{processId}/page Get the processing result, identified by its processId, as PAGE XML, version 2013-07-15.	Terms of service Contact the developer Servers /processing/v1 ~ Retrieve processing result GET /processes/{processId}/alto Get the processing result, identified by its processId, as ALTO XML GET /processes/{processId}/page Get the processing result, identified by its processId, as PAGE XML, version 2013-07-15.		Authoriz	ze 🄒	
Retrieve processing result GET /processes/{processId}/alto Get the processing result, identified by its processId, as ALTO XML GET /processes/{processId}/page Get the processing result, identified by its processId, as PAGE XML, version 2013-07-15. Retrieve processing status and result 	GET /processes/{processId}/alto Get the processing result, identified by its processId, as ALTO XML GET /processes/{processId}/page Get the processing result, identified by its processId, as PAGE XML, version 2013-07-15. Retrieve processing status and result ✓	Terms of service Contact the developer Servers [/processing/v1 ~ Retrieve processing result GET /processes/{processId}/alto Get the processing result, identified by its processId, as ALTO XML GET /processes/{processId}/page Get the processing result, identified by its processId, as PAGE XML, version 2013-07-15. Retrieve processing status and result		Authoriz		

ScanTent & DocScan



ScanTent + DocScan App http://scantent.eu/
Mincremborle.

Julaten: (dkg (Jutter, 15 dkg) Jucker Sewther 15 olky Mehr 12 Pet. N. Bachenwiver, 10th A Kanillin: Beiklar - CSchnee.

Crem Julaten: the vicker 1 Doller. 1 Klar-Schnee, 3 l. Wein 1/8 l Masser, '2 Ock N .- Brillingund: dingpulver" Subereihung: Masser mit Lucker 40 dkg Mehl, 1 Gek. N. Back OKochen. N. - a doingpulver pulver", 3 dkg Kakao, etnas mil dem mein u. dem Doster ver geriebene Libronenschule. rihren. Under standigen

Richren in die . Ercher de Flüssig heis geben is einige there and : roallen lassen. Kom Feuerneh: men u. den festen Schnee da runterziehen.

Marmongugel: Julaten: 10 dkg Ruther, 15 dkg Tucker, 2 Dotter, 00. 2/8 I Milch 40 dkg Mehl, 1 Gek. N. Back: Single



citizen&science



Jumbbofilion Wilten 2 01 Reiter Jofnf Wollfarbackard " . 02 Bauer fram & Pifron " 0 3 Bauer Jung & Tofna Grund- und Bauparzellen-" 04 Serviten Romand protokolle " . 05 Unterberger Juny Grislan " . 06 Tschurtschenthaler Jofaf finisal Innsbruck) Luin Witten " oy Tiener !: v. Adam: Milan

Mofn - mur Windffferflogabon 54 Par Alfeluft Syn bienda Bollowink Bolngabour 2/10 Blofm Morfluigh 253 Mofnogabirmin 53 . Balkmifte Min Speferflighboni 65







😰 Google Kalender - Woche 🗙 🕼 🏠 Industrieller 3D-Scanner 🗙 🛛 🔤 Startseite - Grundbuchblä 🗙 🕅 🔤 Suche - Projekt historisch 🗙 🎆 Kataster Tirol Digital Citiz 🗙	+	\checkmark	-	đ	\times
← → C △	Q		* 1	G	•
😰 Kalender 🔝 PartPortal 🔒 Analytics 🚼 Dashboard Infinity E Espacenet – Sucher 🥥 Usersnap 📲 Transkribus Users >					

.

倄 Start Gemein	iden M	litmach	ien B	ericht	Ük	ber				0
Zurück zur Projektseite	8		54	1			Abfallensbach Ringer -	Parz Zusatz Feedback	. 117	
	1 1		55	1		14	Troger Andon Jofan	Des Grundeigenthümers	303	
- +			56	1	-	.14	dello	Abfaltersbach Gemeinde	13	
		1	57	1			Abfaldensback Gamaine	Person (umschalten:Organisation)	402	
Aufgabe	21/		58	1			della -	Haus Nº	470	
Seite 6	1		0 59	/			della	Abfaltersbach Gemeinde	1 183	1
Einträge Bearbeiten			80	1			dello	Abfaltersbach 💿 🗱	1165	
Erstellen			61	. /		11	Himmen Dil Mul		-	
Bearbeiten Löschen		ied	ha	,		10	Telle I ch. Quin Q.			
Abgeschlossen		lon	R			10	a in	 Entfernen Hinzufügen 	. 34	
	-	a.	09	-	·	2	Callener Gauny	Cultur-Gattung: Oede	lh	
Navigation			04 P.	/		4/00	Walder /: Markler / Jofannis	Speichern	- 12	

tables&forms

🕻 Transkribus Expert Client v1.18.1 (02_03_2022_11:33), Loaded doc: BM_001711_001798_detectron, ID: 953295, Page 1, file: BM_001711_001798_069.jpg [Image Meta Info: (Resolution:300.0, w*h: 1770 * 2223)] 👘 🗇

 \times

= 🖉 🗉 🔻 🗔 👻 🕸 🗁 Search current document... 🔍 🕴 4 1 /1 🕨 🗷 😁 🖬 💌 In Progress 🗸 🗅 🖓 👁 😌 🔍 🤍 🔍 🖓 👾 💱 ... 🗐 🧶 😿

S	erver	Overview	Layout	Metadata	Tools			O TR			
			Ø L	.ogout guer	nter			© L			
		🖻 Docume	ent			🛱 Find		© BL			
	e	Document N	Manager		🐉 U:	ser Manager		©			
		🗅 Versio	ns			🗢 Jobs		۲			
	Recent	Documents		~	🦽 ເ	Jser activity		6			
С	ollecti	ons:	icrc tab	ole test (141	766. Adn	nin)	Col-ID				
6	Docum	ents HTR Mo	- odel Data			,		00 V			
	Jocuments HIK Model Data 00 1-16 / 16 1 1 1 0										
	ID	Title			Pa	Uploader	Up 🔨	*			
	95	test_2			68	holl@cvl	Fri				
	95	BM_001711	_001798_	detectron	1	holl@cv	Th	/			
	95	BM_001711_	_001798_	detectron	89	holl@cvl	We	Jan San San San San San San San San San S			
	95	VHBI00001_	tableGT_F	W3_detectr.	83	holl@cvl	We	2			
	95	VHBI00013_	tableGT_F	W2_training	101	holl@cvl	Tue	0			
	95	VHBI00001_	tableGT_F	W3_training	103	holl@cvl	Tue				
	95	VHBI00007_	tableGT_F	W1_training	95	holl@cvl	Tue				
	95	VHBI00001	tableGT_F	W3_duplica	103	guenter	Mc				
	95	VHBI00007	tableGT_F	W1_duplica	95	guenter	Mc				
	95	VHBI00013	tableGT F	W2 duplica.	101	guenter	Mc				
	95	BM 001799	001885	duplicated	88	holl@cvl	Mc				
	94	RF 004120 (004164 f	or training	. 1	holl@cvl	Thu				
	94	second dum	nmy –	- 5-	1	holl@cvl	Thu				
	94	dummy tabl	le test		1	holl@cvl	Thu				
	94	RF 003184 (_ 003236 d	luplicated	56	holl@cvl	Thu				
	94	RF 004120 (_ 004164 t	rainina dete.	45	holl@cvl	Fri 🗡				
	<						>				
	100	∽ Filter					2				

		1	BM 17	VQ (2
DUFRENCY Marcean	2333	Sergent	49 R.I.	5 Rus Marcelin Berthelot DRANCY
DHONT Pierre		Lieutenant	3298 R.A.	131 Av.Republique LAMADELETNE-LILLS
DEGROTTE Albert	810	Méd. Chef	2º Génie	SORBEY (Meuse)
FRETAndré		Latte Lieut	. 2898 R.A.L.	Rue Earbateux BRIGNOLES (Var)
FOREY Joseph	1846	Sarg. Chef	4º Génie	44 Rue de Paris AUXERRE
FERTE Georges	1401	Gendarme	Brig.Graveli	GRAVELINES ()
GALLAND Nicolas	5 - S	Lieutenant	B.T.D.p. 13	14 Rns Chanve ot PARIS (15?)
GUYENOT Robert	2222	21 Classe	809 R.I.	line Vve BERANGER GIGSY, par BEAUNE
GUE Roger	321	Caporal	2378 B.I.	JI Rus Esquirol PARIS (138)
GILLETTE Boger		Commendant	422 Corps	26 Rue Lemercier PARIS (171)
GERARD Auguste	1496	4448 B.P.	METLLE	Mine Vve MORICAN OIX-de-BRETAGNE (1
GUIDOUX François	1460	I? Classe	2061 R.A.L.	CHAMPTOSE S/LOIRE (M. & L.)
HUDE Joseph		Capitaine	542 D.I.	HONVILLE-1a-BRANC (E. & L.)
HEUNET Paul	2425	Adj. Chef	4063 E.A.	156 Rus Fout enay MONTRANGE
HOTEZ Louis	1617	Gendame	Brig. ARDRES-C CALESIS	- ARDRES-en-CALES
HACHET Robert	2186	28 C.S.M. 1091 R.A.L.	21 Classe	6 Hue Maniot KEHYADE (Morbihan
HERVE YVes	1193	2º Classe	4861 R.P.	ST-YOEAUN (C.D.N.
JOLY André		Lieutenant	4038 D.C.A.	II Rue Pétrograd PARIS (81)
JARDIN Fernand	2	Commandant	4169 R.P.	LAVAL (Mayenne)
JACUIN Albert	323	Adj. Chef	662 R.A.	ISI AV. Republique

🕻 Transkribus Expert Client v1.18.1 (02_03_2022_11:33), Loaded doc: VHBI00001_tableGT_PW3_detectron, ID: 952694, Page 1, file: VHBI00001000024.jpg [Image Meta Info: (Resolution:300.0, w*h: 6345 * 4427)] [c... — 🗇 🗸

😑 🍠 🗉 🔻 📮 💌 🐵 🖻 📚 🏶 😇 Search current document... 🔍 🕴 4 1 /83 🕨 א 🏖 🗄 🔻 In Progress 🗸 🗅 🖓 🗞 😒 🔍 🔍 🤤 👯 🖶 👯 🖶 🐩 🖷 🚳 🦻 🐲

	r Overview	Layout	wetadata	Tools			© TR			
		🥖 L	ogout guen	ter			© L ⊚ DI			
Document 👼 Find										
6	a Document I	Manager		🐉 U	ser Manager		⊙			
	D Versic	005								
_	- Versie	/13					10			
Recent	Documents	•		8	User activity		1			
ollect	ions:	icrc_tab	le_test (1417	66, Adr	nin)	Col-ID	₩Н			
Docum	ents HTR Mo	odel Data					٧ 🏍			
1-16	/ 16 🚺 🖣	1 1 🕨		🔊 🗅 🖻	' வி↓ Doc-l	D	ыĽ			
ID	Title			Pa	Uploader	Up 🔨	*			
95	test 2			68	holl@cvl	Fri				
95	BM 001711	001798 c	letectron	1	holl@cvl	Thu	/			
95	BM_001711	001798_c	letectron	89	holl@cvl	We	P			
95	VHBI00001	_tableGT_	PW3_dete	83	holl@cv	We	2			
95	VHBI00013_	tableGT_P	W2_training	101	holl@cvl	Tue	0			
95	VHBI00001_	tableGT_P	W3_training	103	holl@cvl	Tue				
95	VHBI00007_	tableGT_P	W1_training	95	holl@cvl	Tue				
95	VHBI00001_	tableGT_P	W3_duplica	. 103	guenter	Mc				
95	VHBI00007_	tableGT_P	W1_duplica.	. 95	guenter	Mc				
95	VHBI00013_	tableGT_P	W2_duplica.	. 101	guenter	Mc				
95	BM_001799	_001885_c	luplicated	88	holl@cvl	Mc				
04	RF_004120_	004164_fc	or_training	1	holl@cvl	Thu				
94	second dun	nmy		1	holl@cvl	Thu				
94 94	-	-		1	holl@cvl	Thι				
94 94 94	_ dummy_tab	le_test								
94 94 94	dummy_tab RF_003184_	le_test 003236_d	uplicated	56	holl@cvl	Thu				
94 94 94 94	dummy_tab RF_003184_ RF 004120	le_test 003236_d 004164 tr	uplicated ainina dete	56 . 45	holl@cvl holl@cvl	Thı Fri ❤				

NOULER wels inho STUK ontvange	E S YE W the secondaria	NOUNER dae cuite angele svenkow.	ALFORITHE OF PARTICULUE, df: DE OVERAUER CUREN Paid geochecen, of no wice men other 2.	ZAКRЪРКЕ INHOUD то JRЪКК STUK,	Vuirger behaudeld onder 38	AAR VILUINE VAS RIV HOOPD, VARMERIC DE RAR 18 GRIEGTA	ARN MINN 12 Z.A.A.K M.B.J.K.C.T. 15 VERLONDER.	BEHAMPELING NY VERDERE AFLIOP DER ZAAK.	Vardey bihandeta under Na
186	Jang 13	447	In Worth a ched	Tetus kanadang isa nenjadi m Petus kanadan shabalama ki Leng	1405. TATI	Munai	Ingr I	Joseph 9. 2: Jan 7 12	
1My	T 7	A41	Bedding	Mastering presse Merry 55		Perbeamingen	- AOZ - 6	Clock 24 feer 11	
118	n 9	1149	All Otto	Verbearing forces Manufatting	7	. Portmenayar	AOZ	Bard # 24 Jun 178	
14 9	* *	11101	A Rageman	Relayered and a all when sene Capperate from the College Device thash long	-	Magter	Mar . Troy Sny rule	Dearth Joh 718 Journal With H. Perkey M. Joh 199 Jacke Harman & Comer Jack 1. Shey 18	11314 1221 1979
190	* y	6	l it & Morpel Janmer Jon. 1. Marcon & Cotton Subu	Seogendens nægneddeneta ørner op y saf order ere Welskardenstar dar op Jusy progetsersk menem.		. Jacont Brillay 15 .			519 1.354
141	11 4	462	Jard Clincher	Walter of Example Int. Jacanes Har And fiftheants son Sty Dete		Leonandeler E.	Ing.	Gardie 7, Soil Carrow, Gynost, Carmon fo in Ia St. C. J. Joley Mil	1777
192		167	. Gomenderan _	Theleel warde yn obrinnel om N. M. Sil Sweederwer per he le elwet wer op et Mildersynderen twifeler en Servinet Sterrete fierde en werdette Goerten markt op twijden Mirt, waroopen,	240 1877	ellastinger of	Siè, black by stadd	to Alexand	2119.
193		169	dd c	Collect en arige dans an Oolffale welet en gewige danse, ween bee and gewige danse, we dan bee and en fer den Nagane Weeks Part en glee ween ook Part en gewige ween ook Malten ween dan St.	- <i>g</i> : ,	fillartin Mayor Association	tre- Augsthin	Norsternet fin sond Jani II. Seitund von Staff von De Janit II. on de Gestuden Goog etteckte bishedt	7/11/
								1	-

Medieval music notes recognition (proposition)



Newspapers







MANSTAG 12 PURIAD 1938 Leizte Meldungen Achtzehner man hak Auft 14 Meint

Der Brafibent ber Sanftionstonferent, be Bascancellos, hat, wie loeben befannt wird, ben Achtgebnerausichus biefer Konferens nummehr für den 2. Datz um 15.30 Uhr eine beruien.

ep Benf, 22. Februar.

Die Ingesordnung enthält jolgende beiden Buntte: 1. Anwendung des Borichlogs Rr. 4a (etwaige Ausdehnung des Ausjuhrveröcts auf Bongin, Roblen, Gifen, Guftelfen und Stab 0: 2. Bericht bes Sachweritanbigenausschuffen für die fibermachung ber Durchführung ber Sanftionen. Die Bedeutung diefer Tagung des Uchlashneraus chuffes ift noch nicht abzusehen, ba barüber in Bölferbundtrebien miberiprechente Unfichten laut werben. Go fteht noch nicht einmal feit, ob ber englifche Minifter des Buftern Cben felbit ju diefer Lagung noch Genf fonuner wird oder nicht. 3m erftern Fall würde vorausichtlich ber fraugofifche Minifter Des Auftern Alandin nach Genf tommen, im lebtern Sull bagegen biirfte-Frantreich mabl burch jeinen itandigen Bölferbundminifter Boul-Bon cour perireten merben.

head in CREENISCHE ZEITUNICICAN

Dhenveinwasfert erhliciond 1936

Di be Besprechung des grundlegenden Mimiterickerlaffes für die Haushaltführung der preußischen Gemeinden im Rechnungsjahr 1986 hatten mir bereits darauf hingemiefen (Rolnifche Beitung Dir. 67), daß ein Blafondgefet für 1936 im Ginne Des Gefetjes vom 26. Februar 1935 noch nicht ergangen jei. Dir hallen barouf hincemiejen, daß ber Ministerialerlaß für Breußen bei ber Festiepung ber Rechnungsantelle der Gemeinden an den drei in frage tommenden Uberweifungsfteuern den neuen Plafond jozujagen vorwegnähme.

In Nr. 4 des "Gemeindetags" wird feitgestellt, daß die besogien Rechnungsans teile ohne meiteres ertennen liefen, wie bie 26brangung ber gander und Gemeinben von dem Mehrauftommen on Einfommen-, Rörperichafts, und Umfakiteuer meiters gefuhrt werden. In diejem Bericht des Deutschen Gemeindetags wird ferner feitgestellt, ban ichon für die fetten Monate des noch laufenden Rechnungsjahres 1935 bie li ernenne gunger find, als fie

beträchtliche Sprtichritte gemocht. Die Leihung bes

nach dem Blafondgesetz vom 26. Februar 1935 eigentlich hätten ausfallen müffen, Das Blafand. gefet ift alto im Wege der Udministration perändert, b. h. für gänder und Gemeinden un: günftiger gestaltet morden.

Der Deutsche Gemeindetag falgert aus den Rechnungsanicilen des Ministerialerlasse, dah Die Aberweisungen bei ber Eintommen. fleuer etwas unter, bei ber Rörper= icaitsfteuer etwas über den Bit. einnahmen des Rechnungsjahres 1934 liegen werden. Bei ber Umfaultener fet eine folche Rechnung nicht möglich, da in Breußen für die Berteilung nach dem Umfahltenerschlüffel ein Mindeftbetrag fejtgelegt ift, der durch die Uberweifungen aus diefer Steuer nicht poll erreicht wird.

Inmiemeit das nuch zu erlaffende Blafond= gefen des Reiches für das Rechnungsjahr 1936 Anderungen gegenüber dem "Borgriff" des Ministerialerlaffes bringen wird, bleibt cbzuwarten. Die Gemeinden werden aber aut baran fun, feine Hoffnungen auf Befferung au hegen. Dofür find bie Bebürfniffe bes Reiches au ftart.

ABIE DBLAFT NR. 98 Weltwirtigantiszahlen

Statiftisches hondbuch der Weltt irticheft. Beard it e im S an til gen b el asem v erlin 1936. Ser a fir Cast toll ?, I bir Mail und Gias tiltif, Bmb.6, Worlin SW 65, Preis 18 207.

Diejes "Statiffiche Raubbuch der Befimirichait" mird von allen, die fich beruflich und miffenichaffe lich mit weltwittichaftiliden Bablen und Größen aut beichäftigen haben - und das fil beule ber Raufmonn, ber gabrilant, ber Journalift, ber politifch Inforoffictio, ber Gelehrte mie ber Clubierende --mit Freuden begrüßt marben, entheht es boch in vielen Dingen vom dem geitraubenden Nachlagen und Radsinden in ben oft ichmer erteichbaren ans. fändlichen Swellen Wer weiß, wie unfruchtbar und enthüufthend es für den einzelnen ift dann nach bos Gefundene in ben verichiedenen Berten Reits ichtiffen, Bellungen nim auf einen vergleichbaren Menner ju bringen, der mird bieje Mrbeitserleichtes rung um fo mehr ichaken

Die wirtichaftlich wichtigften Sochgebiete für 80 Bander und für bie Welt als Ganges merben in Jahreszahlen und Monatesahlen bargelegt. Ein Jahr ber Barfriegszeit (meift 1913) mich mit ben Jahren 1920 bis 1938 veralichen. In einer Beiloge find bie bisher varliegenben Jahreszahlen für 1984 enthalten. Es ift in Mugficht genommen. Die Bablenreihe in befonbern Rachtrugen fortzufegen. Die Bilang bes fiber 550 Quarffeiten umfaffenden Berts wird in einer Aufammenfalfung "Dell" gegeben. Gine abuliche ingppe und dach umfaffende Bergleichung der Gefamtaabien ber Welterzeugung an Robitoffen und Fertigwaren, ber Bablen des Bertebrs und bes Unfiene handels in rund 200 Underteilungen in ben Sahren 1913 und 1920 bis 1833 gibt es mobil bisher nach nicht. Die einzeinen Banberücherfterter find einneteilt in bie Sabresgablen ber Bepölferung. ber Erzeugung in Bandmirtfchaft, Forfimirtichaft. Berabau und Industrie, bes Auffenhandels, ber Preife, ber Urbeit, des Geibes und bes Rredits, Staatsfinangen. Rablungsbilang und in die entforechenden Monategablen. Das Borwort des Bandes fagt mit berechtigtem Siale: Cine Sulammenitellung mrildallise

1200 Letoutbeiter im saingerfier f De ware in bo nife ie lieft en iergbau dab Kallewig, 21. Rebruar, In Ratiowiy war ein Betriebsrätefongrift, an dem die polnischen Gewertschoftspertreter aus dem osioberschieftichen Dombrowger und Reatauer Revier teilnahmen. Es murden zwei

Entschließungen angenommen: die eine ftellt feft,

baß ber Schiedsipruch in dein Sohnftreit des oft-

oberichleftichen Berabaues micht den Worderungen

Frauffunck neuer Alughafen Vor egya ; des regehoäß gen Flug velle us Amar g dis Sormaers Sa Gund ber la Bieiten bei jeder beliedigen Bindrichtung feitzulegen. Durch eine Sectauibelwinde wird bann ber Sectwogen mit ME Frauffuri L. ML, 21. Webruar dem Buftichiff in die Richtung der Hallenachie gebracht und fo, voran der Untermaft, mit bem Bei bem ziemlich froffarmen Befler haben die Musbauarbeiten bes Slug- und Zeppelinhafens Buftichiff in bie Spalle eingefahren. Rhein-Main, der jeht im Gilben der Stadt Frant-2ln Stabltonftruftion - mobel es fich furt inmitten bes großen Stadtwaldes eniftebt.

wegen der außerordenitichen Windtröfte um

Who are the users?



Scholars







[Sammelband] (hg. zus. mit ACHIM RABUS) Vermitteln - Übersetzen - Begegnen. Transferphänomene im europäischen Mittelalter und in der Frühen Neuzeit. Interdisziplinäre Annäherungen, Göttingen 2011 (Nova Mediaevalia, Bd. 8)

by Balázs J . Nemes und Achim Rabus

"This anthology approaches premodern phenomena of transfer with various characters from an interd... more 💌

+ Save to Library + Download 245 Views



Search for papers, people, and interests • Search for papers, people, and interests • MARTING * MAR	💯 Kalender 📓 Home - EU Researc 🜻 READ Call 🔣 READ WEB 📉 READ WUI 🔇 KWS 🔇 Kalender 💥 PROJEKTE - Tran	kr 🔇 Sandbox 🖈 KWS SEARCH 💻 www.cardcomple	e	
According the properties of the second test of the	Search for papers, people, and interests		HOME ANALYTICS GRAN	TS READERS MENTIONS E UPLOAD
Becognizing handwritten text in Slavic manuscripts: A neural-network approach using Transkribus 1 Achim Rabus <l< td=""><td></td><td></td><td></td><td>ABOUT AUTHOR</td></l<>				ABOUT AUTHOR
Actimit Rabus ● Actimit Rabus ● Actimit Rabus ● Store of Actimit Rabus ● Store of Actimit Rabus ● Store of Actimit Rabus ● Actimit Rabus ● Actimit Rabus ● Store of Actimit Rabus ● Download PDF ● Download FUI PDF Package Image: Stare of Actimit State Actimit Rabus Image: Stare of Actimit Rabus Image: Stare of Actimit State Actimit Rabus Image: Stare of Actimit State Actimit Rabus Image: Stare of Actimit Rabus Image: Stare of Actimit State Actimit Rabus Image: Stare of Actimit State Actimit Rabus Image: Stare of A	Recognizing handwritten text in Slavic ma	nuscripts: A neural-		Achim Rabus 1.3 Albert-Ludwigs-Universität Freiburg Faculty Member
Addm Rabis So the state of the sta	network approach using Transkribus 1 Acl	nim Rabus		PAPERS VIEWS FOLLOWERS
 e Sto Yees _ d PaperBaix 1.5 _ 0.21 Paper 1 File - the schine Learning, old Charch Slavois, OC Pand Lest recognition The model are in the range of 3 to 4 percent, read-life performance shows that specifically trained to recognize different styles of Duruch Slavois, DOR and Lest recognition capabilities to PaperBaix 1.5 _ 0.21 Paper 1 File - Download PDF A Download Full PDF Package 1 The Download FDI PDF (42 minute read) The Summary The Summary Store I I PDF (42 minute read) <l< td=""><td>Rabus</td><td></td><td></td><td>53 854 154 + Follow</td></l<>	Rabus			53 854 154 + Follow
Machine Learning, 00 d/urch filsword, 000 and text recognized of neural network models specifically trained to recognized different styles of church Silsword handwriting within the software plotform Transitibus. Computed character error rates of the models are in the range of 3 to 4 percent, real-like performance shows that specifically trained modelsread more I Download PDF A Download FUII PDF Package I Download PDF A Download FUII PDF Package I Download PDF A Download FUII PDF Package I Download PDF (42 minute read) I Download FUII PDF (42 minute read) I Download FUII PDF (42 minute read) I Download PDF (42 minute	@ 855 Views _al PaperRank: 1.6 La 21 Pages 1 File ▼			52 MORE BY ACHIM RABUS
	Machine Learning, Old Church Slavonic, OCR and text recognition The paper discusses the automatic text recognition capabilities of neural network models specifically different styles of Church Slavonic handwriting within the software platform Transkribus. Computed cl of the models are in the range of 3 to 4 percent; real-life performance shows that specifically trained n	rained to recognize aracter error rates odels,read more		Biscriptality: A sociolinguistic typology Sandra Lippert, Daniel Bunčić, Achim Rabus, Carmen Brandt, Ekaterina Kislova, Henning Klöter, Alexandra von Lieven, Jürgen Spitzmüller und Anastasia Antipova • 878 Views ↓ Download •••• More Options
Poriginal PDF (42 minute read) Image: Standard	Download PDF A Download Full PDF Package			View All >
Original PDF (42 minute read) NEXT Summary Summary Solution RelateD PAPERS A summary (5 Minutes) Recognizing handwritten text in Slavic manuscripts: A neural-network 1 • Ownodal •• More Options Notes on Scribal Errors in the Earliest Slavic manuscripts: A neural-network 1 • Achim Rabus • Methodological issues in editing the Curzon Gospel*				2 RECOMMENDATION OF THIS PAPER
Original PDF (42 minute read) IETE Summary Show All - RELATED PAPERS Recognizing handwritten text in Slavic manuscripts: A neural-network A framework for improved access to museum databases in the semantic web Mariana Damova und Kiril Simov + 131 Views Mariana Damova und Kiril Simov + 131 Views Download More Options Recognizing handwritten text in Slavic manuscripts: A neural-network 1 Notes on Scribal Errors in the Earliest Slavic Manuscripts Notes on Scribal Errors in the Earliest Slavic Manuscripts Vaor Miltenov - 333 Views - PaperRank: 4.6 Download More Options Vaor Miltenov - 333 Views - PaperRank: 4.6 Vaor Miltenov - 335 Views - PaperRank: 4.6 Vaor Milte				Δανιήλ Δανιήλ (1.7) • April 20, 2019 University of Belgrade, Serbia, Doctoral student
ReLATED PAPERS Overview by Academia A summary (5 Minutes) Go to Summary Recognizing handwritten text in Slavic manuscripts: A neural-network 1 • approach using Transkribus 1 • Achim Rabus	Original PDF (42 minute read)			Show All +
Overview by Academia A framework for improved access to museum databases in the semantic web Mariana Damova und Kiril Simov + 131 Views Download Mariana Damova und Kiril Simov + 131 Views Download Manuscripts Notes on Scribal Errors in the Earliest Slavic Manuscripts Yavor Miltenov + 333 Views + PaperRank: 4.6 Download • Approach using Transkribus 1 • Achim Rabus				RELATED PAPERS
Overview by Academia A summary (5 Minutes) Recognizing handwritten text in Slavic manuscripts: A neural-network • approach using Transkribus • Achim Rabus				A framework for improved access to museum databases
A summary (5 Minutes) Go to Summary Recognizing handwritten text in Slavic manuscripts: A neural-network 1 • Approach using Transkribus 1 • Achim Rabus 1	Overview by Academia			Mariana Damova und Kiril Simov • 131 Views
A summary (5 Minutes) Dotes on Scribal Errors in the Earliest Slavic Manuscripts Yavor Miltenov · 333 Views · PaperRank: 4.6 Recognizing handwritten text in Slavic manuscripts: A neural-network 1 • approach using Transkribus 1 • Achim Rabus 1 • Achim Rabus 1	Over view by Academia			Download More Options
A summary (5 Minutes) Manuscripts Manuscripts Yavor Miltenov - 333 Views • PaperRank: 4.6 Vector Vector • approach using Transkribus 1 • Achim Rabus 1 • Achim Rabus 1		Go to Summary		Notes on Scribal Errors in the Earliest Slavic
Recognizing handwritten text in Slavic manuscripts: A neural-network 1 approach using Transkribus Achim Rabus 1 Methodological issues in editing the Curzon Gospel." Studii Slavistici 13: 257–71 (2016). Cynthia M Vakareliyska • 40 Views • PaperRank: 1.1	A Summary (5 Minutes)			Yavor Miltenov • 333 Views • PaperRank: 4.6
	Recognizing handwritten text in Slavic manuscripts: A neural-network			✤ Download
Achim Rabus Cvrthia M Vakarelivska • 40 Views • PaperRank: 1.1	■ approach using Transkribus			"Methodological issues in editing the Curzon Gospel."
	✓ Achim Rabus			Studii Slavistici 13: 257-71 (2016). Cynthia M Vakareliyska • 40 Views • PaperRank: 1.1

R ScanTent | Your mol: 🗙 | 🕂 Huber, Eugen, Mater 🗴 | G wibmer maurer gen: 🗙 | 🦝 https://transkribus.er 🗴 | G maria kallio archive - 🗴 | 🗞 DeepL Übersetzer - 🗙 | 💓 https://transkribus.er 🗙 | 💓 making a modern ar 🗴 | 🕒 Making a Modern Ar 🗙 🗛 (58) (PDF) Recogniz: 🗙

.

🕁 📕 🏄 👩 :

+

🗧 🔶 C 🔰 academia.edu/38835297/Recognizing_handwritten_text_in_Slavic_manuscripts_A_neural-network_approach_using_Transkribus_1_Achim_Rabus

🔢 Apps 🕺 Kalender 📓 Home - EU Researc... 🤤 READ Call 🔣 READ WEB 📓 READ WUI 📀 KWS 🕥 Kalender 💥 PROJEKTE - Transkr... 📀 Sandbox 🏚 KWS SEARCH 🚥 www.cardcomplete...

¥	Download PDF		A Download	Full PDF Package	I	Summary
	Name	Words	Lines	Epochs	Train CER	Test CER
	VMČ2	252179	53576	400	3.53%	3.91%
	Table 2: Features	of model VMČ2		•		

Interestingly, using more data leads to a slightly (albeit statistically not significantly) worse test CER. This may be due to a non-representative selection of test data. We can state that, with homogeneous data of the amount available in our experiments, doubling the training epochs does not lead to a significantly better model.

The real-world results of the model VMČ2 are as follows:

іманъ рече и слыша по сй гла	WAN BLEYEHEALUMANOTHPA
народа велика, на нбсъ	i i i i i i i i i i i i i i i i i i i
глюще аллил&и а. спніе и	Народавелнка . наньсть
слава и сила бга ншего. И	глющеаланавия спитен

Figure 6: Performance of VMČ2 on SIN 995, f. 999v

Interestingly, as opposed to VMČ1, the model VMČ2 renders the phrase rAÅ HAPPAA correctly. It also adds the correct titlo to HErek; however, it hypercorrectly adds a superscript ^{*}, which is quite understandable, since the plural form HErek is a valid Church Slavonic token that has apparently often been seen during training.

The second example from the VMČ manuscript is transcribed using VMČ2 as follows:

цбь и сбдїл достолщею. браку же црквноеку бгоу примъшенис, ембже сво– бодници стій апли быша

цевисодіядостоящею. EPAK THE EUPKE NOEK BE POV NONMEMENNE . EMOKECBO 7 **КОДНИЦИСТНИАЛАНБЫША**

Figure 7: Performance of VMČ2 on SIN 995, f. 1000r

As opposed to VMČ1 that transcribed this passage without any errors, here, the model fails in correctly rendering both the *title* and the blank space in UARBUARE. As one can see from these

53 8	354 1	I54	+ Follo	w 🦻
52 MORE BY	ACHIM RA	BUS		
Biscriptalit Sandra Lipp Ekaterina Ki Jürgen Spit: ↓ Download	y: A socio ert, Daniel íslova, Her zmüller un d More	blinguistic typology Bunčić, Achim Rabu nning Klöter, Alexand d Anastasia Antipov e Options	ıs, Carmen Brandt, ra von Lieven, a • 878 Views	Exceptedby Procedby Procedby Proc
View All ►				
2 RECOMME	NDATION 0	F THIS PAPER		
Δav Univ	r ιήλ Δανιή ersity of Be	jλ 1.7 • April 20, 2 Igrade, Serbia, Doctora	019 al student	
Show All -				
RELATED PA	PERS			
A framewo in the sema Mariana Dar Download	rk for imp antic web mova und d More	oroved access to m Kiril Simov • 131 Vie e Options	useum databases ws	USER Preside (1) Ficher Restruction of the second second of the second of the second second of the second of the second second of the second o
Notes on S Manuscrip Yavor Milter	ts nov • 333 \ More	ors in the Earliest S /iews • PaperRank: 4 e Options	lavic	
"Methodolo Studii Slavi Cynthia M V ↓ Download	ogical iss istici 13: 2 /akareliysk d More	ues in editing the C 257–71 (2016). :a • 40 Views • Paper e Options	u rzon Gospel." Rank: 1.1	
Towards Ki Archives. In Transcripti Marco Maio	nowledge n Codice I on of the prino • 30 V	Discovery from the Ratio – Episode 1: Manuscripts /iews e Options	e Vatican Secret Machine	
•				

 \rightarrow Public model for Old Church Slavonic



I HTR 'Combined_Full_VKS_2'									
Name: Language:									
Combined_Full_VKS_2		Church Slavonic							
Description:		Parameters:							
Training data consist of parts of the Russian Church century), Old Church Slavonic Codex Suprasliensis (manuscript of the Catecheses of Cyril of Jerusalem. transcribing a variety of Old Cyrillic script styles incl	Slavonic Great Reading Menology (16th 11th century), and the 11th century This is a generic model suitable for uding uncial and semi-uncial.	Nr. of Epochs 400							
	~	Show advance	d parameters						
Nr. of Words:		Nr. of Lines:							
393079		75422							
Visibility: Public model with private data sets			×						
E Save	Show Train Set	Show Validation Set	Show Characters						
	Learning Curve								
V 100% 80% 60% 40% 20% 0 20 40 60 80	100 120 140 160 180	200 220 240 260 280 300	320 340 360 380 400						
		Epochs							
		CER Validation							
CER on Train Set: 4.44%	c	ER on Validation Set: 3.94%							

ð X



Archives

National Archives Finland







← → C 🔒 tuomiokirjat.narc.fi		☆ 📕	🤌 🛛 🌀	1
📅 Apps 😰 Kalender 📓 Home - EU Researc 🜻 READ Call 🌃 READ WEB 🌇 READ WUI 😵 KWS 🕥 Kalender 💥 PROJEKTE - Transkr 🤣 Sandbox 🔹 KWS SEARCH 💶 www.cardcomplete				
Hae Selaa	Info 🗸	Ohjeet 🗸	FIN 🗸	Ì
int sit Under ting forought for acca tin date the date				
and and fig ike Onmart bidene ; in the ting in propieting of				
tout allons Confit of the parta made				
Tuomiokirjanaku				
Hae ja selaa renovoitujen tuomiokirjojen ilmoitusasioiden pöytäkirjoja vuosilta 1810-1870.				
A Hae allebrosta				
Can 1 a / no " the the			1	
a adjutt there				
and the second of the second o				
1 1 de la cara de la				
All the second and the second se				
The set				

R ScanTent | Your mol x 🕼 Huber, Eugen, Mate x G wibmer maurer gen: x 🚯 https://transkribus.e x G maria kallio archive x 🗞 DeepL Übersetzer x 🔻 https://transkribus.e x 🚱 Transkribus.e x 🙆 Making a Modern A x 🗛 (58) Achim Rabus | x +

_ 🗇 🗙



ree & Family History at	🔍 🗶 Transkribus - Keyword Spotting	 Iranskribus - Keyword Spottin 	g × Neuer Tab	× +			
🔒 tuomiokirja	at.narc.fi/en/search?term=helsingfors	nano 🔔 possione 👝 waa					☆ 🛃 🄌
Kalender 💟 Hon	me - EU Researc 🤤 READ Call 🛛 🕅 READ	O WEB 🚺 READ WUI 🔇 KWS 🤇	Kalender 🗶 PROJEKTE - Transki	🔇 Sandbox 🏚 KWS SEARCH	www.cardcomplete		
E NATIONAL ARCHIVES FINLAND	Search Browse						Info 🗸 Help 🗸
		A helsingfors					
				Findtorma	Duick Cuide		
				Find terms	Quick Guide		
	District	▼ Series	▼ Date	▼ Score			
	15,065 Results				10 🖌 🔳	Accuracy (descending)	v
	1826-1826 / Korsholman pohjoisen tuomiokunnan renovoidut tuomiokirjat / Ilmoitusasioiden pöytäkirjat / Page 284						
				D.1.1.11	··· 6/	D. Jigel	4
			-	Alfinflens en	light mid	Inginalit, beflyr	Kim
		ADDIMARIE	A.	beifingfors a l	uporliga de	nifka 12 den 3 Mar	m:/D.
				11	/ /.	Coffice	14.2
				0		Quall'1	
	10/1 10/0 ///						
	1861-1862 / Korsholman etelaisen tuomiokunnan renovoidut tuomiokirjat / Ilmoitusasioiden pöytäkirjat / Page 820						1
		1 Devigas invertion	I Under Ver	and Mattens Ord	forando har fa	(b emaklaga) en falgsande	
		Nin ill' Martalia	L' anfolzaw:		Securit 10		*
		scomeno gradente	- Ombudsm	annen à Sinlands	Rypaletas forenis	y. Heyingfors, den 28.	and the second se
🜀 Family Tree & Family History at 🗅 🗙 🌲 Transkribus - Keyword Spotting 🗙 🔹 Transkribus - Keyword Spotting 🗙 Neuer Tab × + 🗧 🔶 C 📔 tuomiokirjat.narc.fi/en/detail/5Ym7AcAQ?term=henriette%20knorring&result=0&search=eyJ0ZXJtljoiaGVucmIldHRIIGtub3JyaW5nIn0 🔢 Apps 👖 Kalender 📓 Home - EU Researc... 🥥 READ Call 📓 READ WEB 📓 READ WUI 🔗 KWS 🕥 Kalender 💥 PROJEKTE - Transkr... 🚱 Sandbox 🏚 KWS SEARCH 💻 www.cardcomplete. THE NATIONAL ARCHIVES Search Browse Info V Help V ENG V Hack Ylä-Satakunnan ylisen tuomiokunnan renovoidut tuomiokirjat / Ilmoitusasioiden pöytäkirjat 1869 < 682 815 Go > man's forenamence by (bom) man i samme by Home Borgmastaren Fredrick Tiocopi intermade) on so lydande om int uti Heit Bevittna: Johan Johans for Oblita, Que Rajsa Jakobed other Pura. Ð trita he i Thanks. striftlig intectnings ansokan: egare of Ollila skalle trem, a syne vaynar och sason formyn Johan Fredrik Esraimsson. Trengå " Til: taka by Ξ man i forturorde by, , dore for mine minderarige barn Om interpringsmeddelande whi Bonden Bernes Hen Wikalloclial Twitka med mig ega Purra hemmen form] rikofon Hickfula's egande Heikkile stalle hemman i Tear. i aflasagde by. Ibom/ säter. hantakes by af Birkala socken File Sakerhet for est emeller Eva Pristina Henripstaller Thats to a egne vagnaroch of upprallade harhos , original och bestyrkt afskrift bi (bom) sason formignedare for mine minderary chain Invilka laget hontrakt for idminkast jag antiales. Tammerfor med mig ega Claisto stalle hemman i Plantantake H. J. Sibelius. Inspikterå Saris i Saholaks due 15 Optober 1869. Fredrik Tolestam. My. (bom) socken. Bevillas: dere Herr Borgmastaren Prosopi aberhemtade ofvan Johan Fredrik Grainsfon. Frenza" With alloclice Med å andra sedan upptecknade contrakt intagne shippelses innehall samt ingaf det dere abe. sateria (bom) romade kontrakter som lyder valunde): förkla-Ho J. Libelius, Inspektor a Jaris ; Takolates souther) Farmed ofvertale si undertectmane, [se foregunde ft] Med a andre sidan upphecknase) contrapt fortheas rar jag mig till alla delar nöjd. Bonden Berndt Henriks for Heikkils fran Trankaan var jag mig tim alla delar nojo. Jaka by har i socker Villstadeskom genom fillsall on. Henrictui von. Tinpring. bud forre altulanderkanslisten Paul Fudrik Demander, Henriette um Pinorring Bevillne som medgaf den sokke inteckningen inen sade sig icke hun Bevittna: 6. Brakel. Ellrika Braked. no slyrke Bernalt Sleikkilas sganderale fice beroine 6. Brakel. Ulrika Brakel. Med a andra sidan upptecknade contratel fortha. hemmany dere vederborance fings afhada innan , Par ran jag mig tier alla delar nojd. Wik den 16 Oktober fen Med å andra sidan upptecknade contrakt Spain Anderson eque of Panthila trem. Belstots: förklaman : Fankaantaka by. (bom) Med vidare algard have lemmas beroence file anha Bevillyna rattegangs dagen afrast kommande ars i fings laget she rar såg mig till alla delar nöjd. Wik den 16 ende vinterting, de sokanden eger sin ansokan full C. Braker Ellika Brakel. Oktober Civilia Lindroos, egaine of Notice och afyran. folia afrunsom styrke Berndet Heithiles eganderall maa hemman. five Shipkila hemman, Apagot? 18686. Der Her Borgmasteren Recope' aterhemtade of van. Efraim Andersson egare af Pänstilä hemintagne intechningsansotan. Degenskap af ombad for Herr Bergsingenionen Ideetam Dovijad int: Bonden Josef Jabobym Sika fran Rankantika by i sten. na socket for anomformalder anstream, men sade sig icke oar intermense Herr Borgmastaren Fredrick Rosoni) on da. Wi Nahkale he man i Kankaantaka by. (bom) i Frankantaka by lydande skriftlig intectinings ansotran: re fised mid wheching on sin aganderall file the hommon Bevittna. for portratet. Om inteckningsmeddelande whi Bonden Mais Hen.

zeitpunkt – Zeitungsportal NRW



zeit.punktnrw

PROJECT

номе

ZEITUNGSPORTAL NRW

Discover historical newspapers from North Rhine Westphalia!

Funded by the federal state of North Rhine Westphalia, this project aims at digitizing local newspapers from North Rhine Westphalia from the period of 1801-1945. This portal presents the digitized newspapers for a broad public and will be expanded



м.≁

📀 ICDAR 2019 🛛 🗙 🔓 tobias grünin 🗙 👘 🐧 Transkribus in 🗙 🖨 alvermann di 🗙 👘 Transkribus 🗴 🖉 25. Archivwis 🗙 👘 Transkribus 🗴 🖉 neuronales n 🗙 👘 Funktionswe 🗙	🗶 zeit.punkt NF 🗙 🕂 🗸 🔶	
← → C 🔒 zeitpunkt.nrw/nav/index?facets=ftmode%3D"ocr"	☆ 🏚	6 E
🗰 Apps 👩 Kalender 🎛 Slack 👩 PartPortal 📊 Analytics 🚼 Dashboard Infinity 📶 Sales – Power BI 🍀 (+) NewsEye Slack 🔣 Startseite - Kataster 匡 Espacenet – Sucher 📧 Statisch	es Verzeich	Reading list
zeit.punktNRW HOME PROJECT ANNOUNCEMENTS	EN *	N
■ NEWSPAPERS ♥ MAP ♥ PLACES		
Sort by: Show: Order: More results: Newspapers 20 ascending I	21 TITLES	
	FEDERAL STATES	
AACHENER ANZEIGER 1885 - 1943	Nordrhein-Westfalen Niedersachsen	20 1
	DISTRICTS	
DORTMUNDER ZEITUNG 1828 - 1939	Düsseldorf Köln Arnsberg Detmold	9 9 2 1
DUISBURGER GENERAL-ANZEIGER 1893 - 1945	Kreisfreie Stadt Düsseldorf Rhein-Sieg-Kreis Städteregion Aachen Kreisfreie Stadt Bonn	7 3 3 2
DÜSSELDORFER ANZEIGER 1841 - 1855	Kreisfreie Stadt Köln Kreisfreie Stadt Duisburg Kreis Euskirchen Kreis Wesel Kreisfreie Stadt Mönchengladbach Kreisfreie Stadt Mülheim an der Ruhr ↓SHOW 9 MORE	2 2 1 1 1 1
DÜSSELDORFER BÜRGER-ZEITUNG 1892 - 1901	PLACES Düsseldorf	7



University Archive Greifswald



Universitätsarchiv Greifswald

- DFG Project
 - Ca. 240.000 pages of court records
 - 16. 18. Century
 - Ca. 2/3 clean script, rest concept writing
 - Zugänglichkeit wird über read&search möglich sein



Datei Bearbeiten Anzeige Fenster Hilfe

Start Werkzeuge alverman_Vortrag ... ×

🖹 🗘 🖶 🖂 🍳

⑦ ④ 1 /9 ► ⑦ ⊖ ⊕ 150% ▼ ☐ ♥ ♥ ℓ ℓ ℓ ℓ

Handwritten Text Recognition als Schlüsseltechnologie für integrierte Digitalisierungs- und Erschließungsprozesse

Dirk Alvermann

Dieses Kolloquium hat sich ja zum Ziel gesetzt, Digitalisierung auch von den Adressatenkreisen her zu verstehen. Schaut man auf die Zielgruppe klassischer Digitalisierungsangebote von Archiven, dann stellt man fest, dass es sich häufig an dieselben Adressaten richtet, die auch unsere Lesesäle aufsuchen, d.h. an Menschen die nicht nur über besondere Interessen, sondern auch über besondere Fähigkeiten verfügen, z. B. das Lesen alter Handschriften. Optimistisch geschätzt, reden wir also von ein bis zwei Prozent der Bevölkerung. Für alle übrigen sind die Digitalisate historischer Akten oder Urkunden eigentlich nur schön anzusehen.

Hält man sich das vor Augen, versteht man, warum Handschriftentexterkennung – das Thema über das ich hier sprechen möchte - in der Geschichte der digitalen Erschließung und Nutzung von Archivalien ein völlig neues Kapitel aufschlägt. Mit einem Satz könnte man sagen: HTR gestattet den Schritt von der einfachen Digitalisierung zur digitalen Transformation archivalischer Quellen. Dank der HTR wird nämlich nicht nur das digitale Abbild einer Handschrift sondern auch ihr Inhalt in einer für jedermann lesbaren und von Maschinen durchsuchbaren Form – und zwar über hunderttausende Seiten hinweg – verfügbar gemacht.

₽

B

Û

h





One model Two languages/two character systems same Accuracy Rate in both

m capita constante ma trimonio amonnon malita pona criam de inte fe matrimonio - nhecano mariti Jolins mortina nore fiant

1-6 halbbulechen vnndt Stiefvater in capita
1-7 teilen, vnndt das vbrige, so constante man
1-8 trimonio gewunnen, dem Stieffvatter
1-9 zulaßen: Zum Dritten, das also constan
1-10 te matrimonio quaesita bona etiam de iure
1-11 Lubecano mariti solius mortua uxore fiant,

GT

Model 2_8

1-6 halbbulechen vnndt Stiefrater in capita
1-7 teilen vnndt das vbrige so constante mar
1-8 dimonio genmmen, dem Stiffvatter
1-9 zulaßen. Zum Dritten, das also constan
1-10 te matrimonio quaesita bona etiam de inre
1-11 Enbecano mariti solius mortija nrore fian,

Model 3_1

1-6 halbbulechen vnndt Stiefvater in capita
1-7 teilen vnndt das vbrige so constante ma¬
1-8 trimonio gewunnen, dem Stiffvatter
1-9 zulaßen. Zum Dritten, das also constan¬
1-10 te matrimonio quaesita bona etiam de iure
1-11 Lubecano mariti solius mortua uxore fiant,



Family historians – volunteers in crowd-sourcing projects

Datei Bearbeiten Ansicht Navigation Nachricht Termine und Au	fgaben E <u>x</u> tras <u>H</u> ilfe		_ 🗆 🗙
📥 Posteingang - info@readcoop.e 🛛 🖸 Gibberish? - Posteingang -	ir 🗙	.	3
🐺 Abrufen 🔽 🖋 Verfassen 🔽 🖵 Chat 🛛 🖄 Adressbuch 🛛 🛇	Schlagwörter 🗸 🗑 Schnellfilter	Suchen <strg+k></strg+k>	< =
Von krijn: @telfort.nl> 숚	♦ Antworten	🖻 Archivieren ᠔ Junk	🛍 Löschen 🛛 Mehr 🗸
Betreff Gibberish?			24.06.2020, 12:54
An READ-COOP Help desk 🚖			
Hai thee,			
Want you to know that I really love Transkribus. Every time I use handwriting at double speed! I have one question that bugs me: when I use a trained model to there's one or two lines of absolute gibberish nonsense. After the Am I doing something wrong? Thanx for looking into this Best regards, Krijn	e the tool I am amazed about its o transcribe an old text, most of hat the transcription is normal ag	power. Great help in transcr it is transcribed very well. In ain.	ibing old
Das Konto "info@readcoop.eu" ist aktuell			(25) Tagesplan 🔨

Ŗ ScanTent Your mobile scanning 🗙 🐹 Huber, Eugen, Materialien für Kol 🗙 🐼 Project - Bohisto - Web Transkri 🗙 🐼 https://transkribus.eu/r/read/sani 🗙 🚺 Krijn Sijrier - Iroko	🗙 😝 DeepL Übersetzer 🛛 🗙 🛛 🗙 https://transkribus.eu/r/kati/citiz: 🗙 🛛 🕂 🗖 🗖
\leftrightarrow \rightarrow C (🔤 Q 🖈 💹 🤌 🌀 :
👯 Apps 😰 Kalender 📓 Home - EU Researc 🝳 READ Call 🔣 READ WEB 🔣 READ WUI 🔗 KWS 🕥 Kalender 🐰 PROJEKTE - Transkr 🔇 Sandbox 🏚 KWS SEARCH 🚥	www.cardcomplete
Wat zoekt	contact linked in
MANAGEMENT CONSULTANCY KENNISDELING PRODUCTEN ACTU	

PROFIELEN ///







Waar staat u voor bij Iroko?

Gecompliceerde vraagstukken simpel maken. Mijn kracht ligt in het combineren van bestuurlijke, beleidsmatige, juridische, en financiële invalshoeken tot een praktisch uitvoerbare oplossing. Ik ben resultaatgericht, maar heb oog voor de sociale component. Zonder daar al te idealistisch in te willen zijn geloof ik dat oplossingen alleen houdbaar zijn als ze worden gedragen door degenen die ze moeten uitvoeren.

Welke werkzaamheden zijn zoal op uw pad gekomen?

Doorgaans werk ik als adviseur en ondersteun wethouders en directieleden bij het opstellen en praktisch maken van een veranderende visie.

Ik voel me thuis in een omgeving waar ik creatief kan omgaan met de inhoud om praktische en werkbare oplossingen te genereren.

Een voorbeeld ter ondersteuning. Twee gemeenten die reeds ver gevorderd waren in een fusietraject, konden politiek niet komen tot een keuze voor een van beide gemeenschappelijke regelingen voor de uitvoering van de Wsw. Bovendien hing de financiële vergoeding voor uittreding als een zwaard van Damocles boven het fusietraject.

Ze legden uiteindelijk bij mij de vraag neer: kom met een advies voor de keuze van een van beide partners tegen een aanvaardbaar bedrag aan uittredingskosten. Dit is het type probleem waar ik me graag in vastbijt. Hoe kom je tot bestuurlijk en politiek draagvlak, maar ook hoe kom je tot een betaalbaar en vooral werkbaar resultaat?

. . . .

Be a set of the set of



Wat doet u in uw vrije tijd?What do you do in your spare time?Ik ben historicus, niet alleen van opleiding, maar
ook van belangstelling. Ik lees veel historische
literatuur, maar doe ook af en toe
archiefonderzoek en schrijf artikelen.I am a historian, not only by education, but also
by interest. I read a lot of historical literature,
but I also occasionally do archival research and
write articles.





😔 Online Chec 🗴 😔 Itinerary Rec 🗴 🚦 Drop-Inn 🛛 🗴 😵 DeepL Über 🗙 🚥 READ-COOF 🗙 😣 Endlich ist d 🗙 🧐 https://rech 🗴 🌃 Project - Bol 🗴 //trans × ■ transkribus.eu/r/read/sandbox/application/?mihyro=1&markers=1&colld=26748&docld=4382&pageld=3 C 0-Apps 🚆 Home - EU Researc... 🥥 READ Call 🛛 READ WEB 🌇 READ WUI 🚱 KWS 🕥 Kalender 💥 PROJEKTE - Transkr... 🚱 Sandbox 🏚 KWS SEARCH Kalender 14 \oplus Fin Bozen den zweisten Farmer 1.7. 9 θ \triangle Berrn Burgermeifter Foleps p: fonial. Dom Legieringsrats ind 7. S. M.

x² x₂ <u>U</u> ab ?! Unclear A

Notice	- + X Josep	h 💧	Text Region 1		#
Notice			ă.	1.	#
			2	<u>In Publiko = Politicis.</u>	#
Occupation			3	Aktum 🏦 Bozen den 🕓 zweyten Jänner 1794	#
Occupation	4	Vor Titl 🗄 etc: Herrn Bürgermeister 👗 Joseph 🗄 v: Remich	#		
Firstname			5	: kaiserl: : königl: : OÖ ^{en} = Regierungsrath und : T: L: M:	#
Firstname			6	In beyseyn der Herrn Herrn Magistratsräthe ,,	□ Ready for Review ▼
DateOfDeath			7	,, Oberlieutenant 🍐 Karl 🗄 v: Kuglern ,,	Save Changes
DateOfDeath	-		8	" 👗 : D ^r . Leopold Spiellmann "	

Transkribus community



How has working with Transkribus impacted your professional and or scholarly/activity?

- It has convinced me that it is workable and not vaporwar.
- Greatly. My most interesting job and my first funded project which I am now starting would not have been possible without it.
- Could do things that where not possible some years ago
- Expect it will affect my profession in the future dramatically.
- Gave our department the ability to start a whole new project
- Wouldn't have started my digital edition without Transkribus
- ground-breaking!

40 most used tools



Source: WHICH DH TOOLS ARE ACTUALLY USED IN RESEARCH? BY LAURE BARBOT, FRANK FISCHER, YOANN MORANVILLE AND IVAN POZDNIAKOV — 06 DEC 2019 https://weltliteratur.net/dh-tools-used-in-research/

Registered users in Transkribus 2015-02 - 2022-03



READ-COOP SCE



What to do after the project?

- READ Recognition and Enrichment of Archival Documents
 - 8,2 mill. Euro grant, Horizon 2020 EU project
 - 14 partners from all over Europe universities, archives, scholars
 - Excellent work
 - From 2016-2019 (3,5 years)
 - Before: tranScriptorium project 3 mill. Funding first prototype of Transkribus and Handwritten Text Recognition
- Sustainability?
 - Spin-off as company?
 - European Research Infrastructure Consortium (ERIC)?
 - Association or foundation?
 - Open Source?

Societas Cooperativa Europaea

- Collaborative
- Open
- Democratic
- Profit-oriented
- No shareholder value



Anmelden DE 🔻









2

Albert Development

Bettina Operations

Felix Development

Florian Development





Florian Marketing

Günter Development



Johanna

Operations

Joe



Mirjam Operations



Maria Development



Philip Development



Sebastian сто



Thank you for your attention!

