

Saubere, kontrollierbare und zuverlässige Daten- und Informationsextraktion für Archive

Beitrag zum 29. Archivwissenschaftlichen Kolloquium der Archivschule Marburg,

Michaela Prien, **Transkribus**, READ-COOP SCE









Die Herausforderung



1807

Acta

die auf Requisition des **Franzl. Procureu General** Tissot aus Mainz geschehene Aufsuchung eines Jean Becker aus Speyer in hiesiger Residenz **betrl.**

Polizeiakte

[Transkribus, Text Titan I, 16.05.2025]

D 51

Acta.

Betr. Auf Ersuchen des französischen Procureur Général Toffot aus Nancy getroffene Verfügung betreffend Jean Becker aus Bingen in hiesiger Haftanstalt betr.

Polizeiakten (Rechts oben handschriftlich: 1807)

[ChatGPT 4o, 16.05.2025]

1807

D 51 Acta

betr. Aufsichtssache des franz[ösischen] Procureurs General Tissot und Banquier Procureur Vogt. Aufsatz nicht Dean Becker aus Pfizen in hiesiger Polizeiakten betr[effend].

Polizeiakten

[Gemini Advanced, 2.0 Flash, 16.05.2025]





Tohnee Dypelpflug für tienlahnen. Solfo Guinnifa, Winingon, Umayalum, Brighailan, Antanhanfring no, - welfow the frances unoft from notyitainhumin Munfinsknifton walfu dirtyaban ind vaniculanta finnafran Vin Ofunny now ufring me in Ludwinder Ina fifure. Sufran unanifuefore, if with men dan Entwinds, boundow trafallow, forman muf som vaifour. Tehnee Dypelpflug für Eisenbahnen. Boulder Guranniga, Overingan, Umanyaluni, Brighting Untanhanfringen, - walfan thefund must plane mifgutanibanian Manffankniftane, walfa this gubow ind anamini nata finnafunn Sin Ofunny now ofring ne in and winder in fifure. mustikalum Ofinian, Jun in gafabana Ofunamaffa in your glingh Tilfbone frollat, and Linfalba buisha fails fullwind befinds, in find mit in maturelings Pefunoflish windunfullow in before, saidans, goone bridgeth nery unloverable medan neward Hinden wan whom 10 Gundan yayan dan Gunizout yaningta Stanner, walifu din gafabawa Orfinnmaffer un Sam Spirit winfred Mintenfulland in folorige varfinitioner, his fin wif Sin unfectuation different pritoriette gaftfoline ift. Vist grayer billet in winneringer Hagningenfiell,

Schnee-Doppelpflug für Eisenbahnen.

Welche Hindernisse, Störungen, Unregelmä ßigkeiten, Unterbrechungen; – welchen Aufwand an oft schwer aufzutreibenden Menschenkräften,welche Ausgaben und verminderte Einnahmen die Schneeverwehungen im Betriebe der Eisen-

[Transkribus, Text Titan I, 15.05.2025]

Schnee-Doppelpflug für Eisenbahnen.

Obgleich Spureinreißer, Räumungswagen, Rotationsräumer, Schneepflüge, Schneeanwerfmaschinen u. mehrere derartige um oft schwere mitführbare oder Menschenkraft erfordernde andere Hilfsmittel zur Freimachung oder Freihaltung der Bahnstrecken gegen den Schnee oder

[Claude 3.7, 15.05.2025]







Wie jedes gute Werkzeug nimmt KI Arbeit ab – Bedeutung und Kontext entstehen durch fachliches Wissen.





- Aktueller Stand: Wofür KI im Archiv bereits eingesetzt wird
- 2 Die n\u00e4chste Generation: End-to-End-Modelle
- Praxisbeispiele & Anwendungsszenarien
- 4 Ausblick: Integration großer Sprachmodelle (LLMs)



Aktueller Stand: Wofür KI im Archiv bereits eingesetzt wird





Texterkennung & Transkription



Wollten Sie mir, werthester Herr Hoffmann, mit wenigen Worten, einige Nachricht geben: ob die Zeichnung wieder zu Ihnen zurück gekommen ist, ob Sie das Bild angefangen haben und ob Sie noch glauben daß Sie es im August werden endigen und abschicken können? Bey dieser Gelegenheit wollte ich Sie ersuchen, mir ein Kästchen mit sechs Gläsern Eau de Cologne # zu überschicken, wofür ich den Betrag mit dem übrigen gern erstatten werde. Es ist dieses wohlriegende Wasser seit den Verwirrungen der Zeit schwer bey uns zu haben. Der ich recht wohl zu leben wünsche. Goethe

Weimar am 9. May 1802.

Mit dem Postwagen zu überschicken.





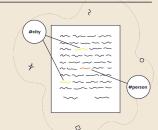
Strukturen erkennen (Formulare, Tabellen, Marginalien)

	-			_							
Keystone academy	Factor	yrill	le, Pa.	5000	aymnasium			16/11	Low		
	Philad				General			desta	Low		
	gamb				Bilding			2/18/11	Has done share, 2/20/11; D	22	
	Lincol		Tan. 10	000	Eudowment			2/20/11	Low		
Kentucky, State Univ. of (1) C	Louisi	Cle. 1	Ku.	600	Endowment Teacher's so	lan		3/4/11	Not sufficiently developed. 3/24	109	
Knox Institute + Ind. School(1) CT	athen	_					_			and the same	
Kolex Normal & Training Inst.	Koler	Į.	aupt:Gi (Affent		buchheft gang)	19.0.	1	Blatt	Mr. 625		
Kidder Institute	Kidd		Bor: un	d 3	uname	Name	ofir	mn E	Hipfair.	Hup	land
Kansas, Univ. of	Laur	_	1	Ort			1				
Kanaas Wesleyan University (1) F Kenka College (1)	Salin		Drt	yun	mbrick		ii.	Orts= gemeinde	Imsternich.	Geburts=	Jahrgang 18 83
Kingfisher College (1) 7	Kingfu				100/1100	,	1075356	0	OMMO JUNET	jahr	
Knox Institute o Ind. School (2) F	athe	irt8=	Bezirt	ym	worinek.	To be	Beimatsberechtigt	Bezirt	Innsbruck,	Religion	
V -00 C 00	Knox	Gebi	Comitat		7.		ıtsbe	Comitat	Y.	oterigion /	hathol.
							Ĕ		/	Runft,	
King College (1)	Brista	_	Land	(Tirol		Bei	Land	Tirol	Gewerbe, fonstiger Lebensberuf	Tiller
		Ini	1 190	1_					wad han Qasua	ine out buil	Okalima In ham
		D							nach der Losre		
		en ,	Jahre in	ier h	teserve und	zwei ?	šah	re in der	Landwehr, jum 3. 2.	after to livel.	ray sorget
	-									4	1





Entitäten festlegen (Namen, Orte, Daten)



It Wil Jer pain finnist, als brig dan in mag autionformy Jafo allfir in Tirbail an sgarfain Barray Jayiefon Dan ourbling die barraito fin in (ao: 1062. Zum gamainan, Ratfifla and Iga lates princta magan dan norfrinanda, Disto minden Olonfoffen, Samalo banfinatan gamingarary augast ifra holl bounder noroistanie nirgt anlanget, outerer dan unforme theil dans albon and ains hambana Zin ammading and 3go hallot worden, and darrainf mis and one fair halag circa medium Mais anni cur

Zu wisen se	ij hirmit, als beij der in negst¬
entwichene	n Iahr allhir in Lübeck ausgeschrie¬
benen hä	
bevor in a	Ortsname
auffgeset	GND http://d-nb.info/qnd/4036483-5
Städte wie	placeName
geringere	Lübeck
nicht erlai	type
selben au	Schließen
außgestel	
rer hänseta	g circa medium Maij anni cur¬
rentis ausge	eschrieben ist; das demnach
hierneast ve	ermeldete Erbh: Städte theils in





Durchsuchbarkeit & Recherchehilfe



ämtliches Fahrniß vermögen der Gemeinde soll gegen **Feuerschaden** bei dem Franz. Kon-versichert werden.

Q Search collection

Ausschuß.

Gemeinderath

Forch Unterkie-

Gassenbach



Die nächste Generation: End-to-End-Modelle





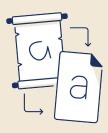
Die Herausforderung:

Wie kommen wir möglichst einfach zu sauberen und zuverlässigen Daten und Informationen?



Die Lösung:

Saubere, kontrollierbare und zuverlässige Daten- und Informationsextraktion mit End-to-End-Modellen







Was bedeutet End-to-End?

Bisheriger sequentieller Workflow:



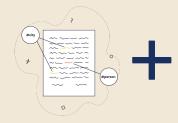


















Hochladen

Layouterkennung:

- Zeilen
- Felder
- Tabellen

Texterkennung

Entitäten markieren

Bearbeiten

Exportieren & Teilen





Was bedeutet End-to-End?

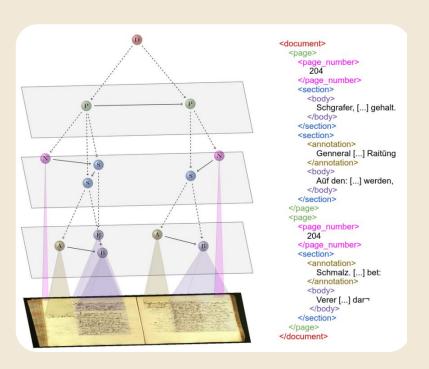
End-to-End Workflow:







Wie End-to-End Modelle arbeiten



- Betrachten das gesamte Bild in einem Schritt
- Erkennen Text, Layout und hierarchische Strukturen gleichzeitig
- Liefern strukturierte, kontrollierbare Informationen ohne Zwischenschritte

End-to-End heißt: Ein Modell, ein Durchgang, ein Ergebnis – direkt vom Bild zur Information.



Praxisbeispiele & Anwendungsszenarien





Beispiel: Baskische Schüleraufsätze

MAR	La sexa las gellas pestan ogolia festa sóstela suse elacia
	for eig low boson ellean lanc la asouther BESTKA
	(Bolo Horre a role gessa perka citea sur Docateccia min
	these evere con mosto ocropia pera de EESHA It operens
	morament a pela coopale reco gaures ganue 110
	May acous postra of stested a ula Mir par for EBEETA,
284	recta over de ofarpor person lo est ortula to some
a	ony true hustin new ousha o dalpha oo lesta (25th cesta)
	ste potea obje juvola sastra, ome pole parent tartità
	Muster you make devene pedang ma a garagean
	resing, partir also pasta, ala basisonare a centre de rescea
	loa nita land potina lama, EESRA EESTA poscrah
	accolaco to sesta cocrune outa esta a teriora
_	nusta lejásza desra noona cinceta. Ocus, per oexantiqua
186	when was out EESGOOD BESKA ICEZ OU METUR
	& acception EESICO Meme cate a ASTAIR -com
	pulla as pulsa pasta nata purata pe deta acesta ocules
	point ESTA actionaction allowante of resignate oris

Zeichenfehlerquote (CER): $7\% \rightarrow 3\%$

Layoutgenauigkeit: $45\% \rightarrow 90\%$

(Symbolbild)





Beispiel: Etiketten von Sammlungsobjekten



• Zeichenfehlerquote (CER): $27\% \rightarrow 8\%$ (bei gemischten Schriften)





Beispiel: Etiketten von Sammlungsobjekten





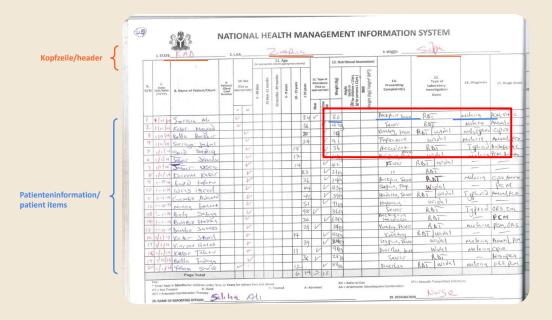
Specimen	Sex	Determination
Trypoxylon fronticorne Guss.	Q	det. A.V. Antropov 1987

- Zeichenfehlerquote (CER): $27\% \rightarrow 8\%$ (bei gemischten Schriften)
- Gute Datengrundlage für Entitätenerkennung





Beispiel: Tabellarische Daten



```
1 {
        "header": {
            "1_State": "kad",
            "2_LGA": "zaria",
            "3 Ward": "saku"
        "items": [
                "13_Weight": "36",
                "14_Complaints": "backpain, fever",
                "16 Diagnosis": "malaria"
14
                "13_Weight": "41kg",
                "14 Complaints": "fever",
                "16 Diagnosis": "malaria"
18
19
                "13_Weight": "19",
20
                "14_Complaints": "vomiting, fever",
                "16 Diagnosis": "malation-typhoid"
24
                "13 Weight": "91",
25
                "14_Complaints": "thypoid fever",
                "16 Diagnosis": "malaria"
26
28
29
                "13 Weight": "36",
30
                "14 Complaints": "accident",
31
                "16 Diagnosis": "typhoid"
32
```

- Zeichenfehlerquote (CER): 4%
- Wortfehlerquote (WER): 2%





Beispiel: Zeitungen

Oberste Ebene



Mittlere Ebene



Unterste Ebene

By Associated Press			RITA				from Edity Crystalian P. Soc. 15th 11,448
STATEMENT ED 1876		BET-BRITAIN, CONNEC	TICUT, TUESDAY, NOV.	ESTIGN SQ 2003 - ERST	IEN PAGEN		PROPERTY OF THE CENT
AILLAUX GETS FROM FRENCE	SENATORS	SHORTER HOURS FOR TAX CURIOUS	"STUTVESANT HE	LD FOR MURDER	UP IN BIG STORM	STORY OF AUG TOLD BY MIS	S SCHNEIDER
		ter ten listadies to	100	-	Add the Pessel in Sea-Tea	IS GIVEN TO	PUBLIC TODA
ote Was Surprise -	ORDINER HAS QUEER	Revenue Collectors	國人種	1	Known Brut, Mary Missing	HENE TWO YEARS, HE	
turned - May Be		DOSESSES SOUR STATES	6 6	The same	DOOR TOW NOTICE BOARS	BEFLEES TO PAY TAX	script Of Eviden Taken At Inquest In
Given Important Post — Paris Undisturbed.	Gut Into Furnace and	State And See the Familiania. The Care die Publication State		2 33	They Fall To Mike Lond - Line of Delayers, Are English for Pages	pears Refere Mayor	Rogers' Death A
		New Re Brought. Fundament Serv. 12 - Except of the Control of Control Serv. 18 - Except of the Control of the			School Vaniers.	and Alderssen	New Haven — Ha "One Hiebball."
THE ENEMY IN WAR	to You keelend From Crimeles, C., Spr., 11.—Gar., C., discover, for one pasts parts of Class Laboran shorter of Spr. a. select to the property of Spr. a. selects was applied by property of the selects of the property of the shorter of the select to the s		100			A Fig. Aller, head at a real value active species on the city on pages before an excess per head	MENNEN UN NOTHIN
The book of the property of the party of the				AN ASSESSMENT			OF WHAT HAPPENS
reports promounts of the wall- m core annexy make by the many fit to fid. The amounts his was going at the		of Bright Ad printing was being	The newsyl figure in the for of Orleanus Janes Hells	Name and Address of the Park	on mining in many the real	Common Tax classics Enterture Common to home for recovering or a no drawn for a tear class.	
The samples his was going we so a challenge that every . Early Solder, character satisface of the Color, countried of measured at the and translated from Place the state process to the process to the sample for the process to the p	for extreme of the period to be not to bring contacts to being Powers's fifth 1. Eve		The newest figure in the day of Polescous James, black Lourge Staymount, of Tary, N York olly, Staymount, if he is ably to Lought, to this olly it Illian alsows here in encounty of the New York point departies	on with the placed will pro- brough recording passessings.			
		Notice that below to a new to said the ball the below by we will be below by we will be below by the below by			The second second second	DESCRIPTION AND STREET OF THE PROPERTY AND STREE	
the court of the c		NA STREET, AND ACTION THE DAMPS OF AN ACTION AND ACTION AND ACTION AND ACTION AND ACTION AND ACTION AND ACTION ASSESSMENT	RART SDEPTICAL ON	TENTO OLYMAN	the real of a page which was been ing time from the mater to refug as a years guide name.	and the attention of the branch flow in the sphales, there is \$150. In page second try personnels in the con-	The Section of Contract Contra
of nationalise delice being by	reside to have barried to their the barried associated barried factors. The transport that for some ar- ele, but was strated for complex	the term of the angle to be responding to ration. It responds to resided	STUYVESAVT ARREST		A great man from the first to an A great man from the first to an A desired the ment when the are	Barriera automor and front to the court of the court of the court and the court of	
ordin ore product with the other me in his Egyptes to Ingenes. "William to King stough to be	which the was districted for named on the Control of Shirt schools as a second of the shirt schools are a second of the shi	to the half that which he specified to the heat that the state of the same after that year after that the same after that the same after the	fine Papers at Once,		and the same of the same of the same of		for each they led back many over week our cross to be t fact and amounted to live a
ploint," he mad, "[14] or long for our rise glower and political dis- month 1000, Earling for You are	tricks war lost word. Note the temporal part was en- received that his wire built loss on-		He Says	1 1	in any most card and for an each in the most of the card and the legal spiritual and the card an	or foreign provinced has no through private frames. If the Suprate Street was a second	
cotice and ill monatores year con- legion to more complete good at the after time for a financial of Feelin-	the six the give it to hope which it is not provided as another the bright of the soul-	There is no bottom to the state of the state	Popular for the expection of Strongs Mark man, of Days M. T., the new holysty bring had by Mi-		Parelinery has the Swidter for Swidter for the State of the Swidter for the Sw	Principle place and the Europe Control of the part of	
mine. The processors is price to a price of	And I p. to The minut's help the read of this b. it. when the photos network I'v man affiliar.		Petrologic Acada Signify in the Par-		from a coop with, will much one about jut on the law to the about the coop of the law to		
There is their man types, W. then in an mind, fallows first serve man reproduct (regard or serve and product with some first	the at heaving Area.			13-67	Wigned they will be risks being and processed to the process with the proc	Lind tale or eventual, sensor our eventual. State was been to be according to the courty do	Will gardened at the disk one died was "one Mesons Society." "Then the more of the soci
of the emission topic as all tree based in the Table joins of a far than at Appendix to se-	TENTENT SEGLER COMENC 1479 COUNTER SOOK	the gapter of the transport on givening. The agency is any other section. It is the transport of the transpo		A siding peral mean- liance between Geral Chry- mas, who was designed to the prize with transfering Perico- ture factors for tip, and Design	devices as they where the devices required. Here the the court of the	DETT TOOLSOM RESULTS	
er Proph others remajority, strik remo. The surery proping spring	And the Stephens Stephens	year such a district week! In any manufact last in the White the hear year was published at the streets would be invested.		prize with restlering Price-	The fat, become, very six of the correcting copy reported the loss of a life.	EBOW YOUR CONTRACK	
ay the for performed will be the fire, their wall interest entired to real the processed transfer to	Anni Maria Sana Sana at Maria Sana at Maria Sana at Maria Sana Anni Sana at Maria Sana at Sana	PRETA VATERILET		two pictures. Both face share	who designed to a st pro-	Arting Against Street St. Corn- unto of Standbury	
Mantha saltrantif to the atlante I design will fine feet from	the season of the tensor and the tensor and the tensor of the tensor and tensor	DOES \$75,000 DAMAGE		Surises, high Statuses and war places. At a cook places, one might to taken for the other.			the parting time. There was a record of the second
	de come a porset, from the a col- index granding exteriors in fire lock.	Chase Building Steps in those Building Is		CENTRALD IN SUIT -			
ngs gon do in the below of the common of the common for the beautiful in 18 to the last the common of the common o	tel sept from treating to ten	Heariest Loser		DIER WOODLAND DEAL			
the rise is again to the site of an act of the site of		Washing Nov. 21.—Place as- charged single for history of here married from A. area Cons-	Marie Committee of the	State R Select and	CONTRACTOR OF THE PARTY OF THE		
	ten. With the Kompan copy was being at union of the ten of the ten of the ten on the ten on the ten of ten of the ten of t		prints for the print for the order of the party of the prints and the prints of the prints of the party of th	Others See Saylooak Man for \$20,000			and the field off of the shift
on a change (reprint black) for the offerind below on to his fideric of about Chines arrowd for	TATOR STREET				processed on French Days of		
money, by ago name to the French. All the tax ago, as how whether the se closeless or the sector, in the 2.7	FILED IN THIS STATE		A design change the for engine are not to the common pages thing but was to see after the common pages that is not after		NYSTERY IN NURSER		and No. and Sim. All or key to also have been for the boats
Mind and Wise, afficiently on New London with parties, ago of the parties, ago of the parties of the contract of the parties of the contract o	New Hotelstein by Course				160 MILES IN FORESTS	Newman Milley Ther Deciment in Night Fire	
pe formation rough the to the organ region Pattern, to treated more of 100 with the trength of	Earlier, No., No., American Commission of American Commission of American Commission of American			Committee of which the fact of	versed. Was Married		From San San San Class C
The december was attended to hery fiche to produce you the lot options at bit was four fire.	e that of Charles St. Stones, Fig.	To be an formed to di-			Three Works Age		
per minimum of the behavior trees made handless and firms marring to play signs. To say the property	MARINE AT MOTO NATION OF THE PARTY OF T					Engage rate to the section of the se	PERCENT ANALLED AS
		caped the give to be beginned from the caped the give to be beginned as the caped the	Market M. Pop. 171	many and the second part		how, by the fixed dirty of the hite flux poors. Frame of the last man a going had not how white	AFTERNUITE OF MARK
"I Among great flat a formy position of the reposite row lo- tion again to region as produc-	r Wagan graf fire realization for the same realization of the same realization	rap to the first to the cold rap to the cold of the cold statement for the cold to the cold of the col	Dine N. B. H. S. Team	to the supplied to the first to the	the same of the part of the	to property control	TO DESCRIPTION OF THE PERSON O
make which contains degree for attendancy fee anaethering I may be than to my promonel formation using my promoterly no union of	are, 100, Fig. Tuber St. Arleans, autobares, 1991 Same & Printer, Spitzeris, 1991, John G. Petter,		Mark a Northern Street			process of a final state of the	
proteins." Nest in the frequencies, pro-		and a proof of a copy camber of sig- splet copy copy when a first long. The sanda restrict Press the families					
	ne the private property of the federal of Series, wave 2004.			British Swift Betrathal Is Announced Today		TOTAL PERSONS	
more the month of Peans with a leaded theory admitting means of the same tendence of the leader to t	PROPERTY AND ADDRESS.			regarded the employer. It has contain the transplant of his contain the first property for the factor of the first	The same and the same as the same	Racco, No., To-Capital con- also and that would not receive a see Harris Fa D	
titud laptur. Whose represe as a growing south him a figure of south completely in the details helps the real way, become the best of all	the flor that he built I live pro- mobile to be top of course disk! St. a strong and to the strong	and the first being being their states of the second to the second to the second test of		The State Section is a superior made for the State Section and the superior made for the State Section	Marie Co. Co. Co. Co. Co.		
		The short deader	Now hope two constants	AND SHOP AT PERSON AS A PARTY OF THE PARTY O	half fore control or fix age for gate, making the control of the gate, by passe by the control of the gate.	where he will be the sent of the force from a final of the point of th	timelind in Ferring Burned With Big L
Common part wit Corners The cities of the start from the lead of the start Michigan Cal- ary which had low Easter, Cal-			12 Material Street	THE RELEASE			parried With Edg L
			Stephen, You, (A-1) recent for New Settles and country Controlls: Sale could'd and		The strang of the same of the factor		





Beispiel: Zeitungen - Strukturierte Ausgabe

Input





Output

```
<heading-top>
               <heading-metadata>
               News of the World
               By Associated Press
               </heading-metadata>
               <heading-title>
               NEW BRITAIN HERALD
               </heading-title>
               <heading-ciculation>
               Average Daily Circulation For
               <tag_date>Nov. 15th...</tag_date>
               11,448
               </heading-ciculation>
           </heading-top>
           <heading-bottom>
               <heading-established>
               ESTABLISHED <tag date>1870</tag date>
               </heading-established>
               <heading-date>
               NEW BRITAIN, CONNECTICUT, TUESDAY, <tag date>NOVEMBER 18, 1924</tag date>.--EIGHTEEN PAGES
               </heading-date>
               <heading-price>
               PRICE THREE CENTS
               </heading-price
           </heading-bottom>
29
       </heading>
       <body>
               <article-heading>
               <tag_name>CAILLAUX</tag_name> GETS AMNESTY
               FROM FRENCH SENATORS
               AFTER VERY SHORT DEBATE
               </article-heading>
               <article-subheading>
               Vote was Surprise -
               Gets Civil Rights Re-
               turned -- May Be
               Given Important Post
               -- Paris Undisturbed
               </article-subheading>
               <article-subheading>
               </article-subheading>
               ACCUSED OF WORKING
               WITH ENEMY IN WAR
               <article-paragraph>
                  Paris, <tag_date>Nov. 18</tag_date>...<tag_name>Joseph Caillaux</tag_name>,
               former premier, who in October
               <tag date>1918</tag date> was found guilty of having
               "Impeded prosecution of the war.
               was voted amnesty today by the
               senate 176 to 104
               </article-paragraph>
           </article>
       chodys
59 </Page>
```



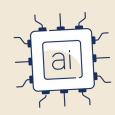


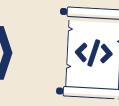
Vorteile von End-to-End-Modellen

- Vereinen Transkription, Layout-Analyse, Extraktion strukturierter Informationen, Erkennung von benannten Entitäten in einem Modell
 - Keine Fehlerfortpflanzungen mehr → höhere Genauigkeit
- Ermöglichen hierarchische Entitäten-/Layout-Extraktion
 - o Arbeit mit komplexen Leserichtungen, verschachtelten Strukturen etc. wird einfacher
- Zeitersparnis & Skalierbarkeit
 - Einfachere Wartung der Modelle möglich → weniger Overhead
 - o schnellere Verarbeitung großer Mengen, z.B. bei Massenakten oder Zeitungsjahrgängen
- Sind kontrollierter im Verhalten als große Sprachmodelle

















Anwendung von End-to-End-Modellen

Saubere, kontrollierbare und zuverlässige Daten und Informationen



Ausblick: Integration großer Sprachmodelle (LLMs)





End-to-End als Schlüssel für:

- die Weiterverarbeitung mit Hilfe großer Sprachmodelle (LLMs)
- das Trainieren großer Sprachmodelle
- die Unterstützung großer Sprachmodelle







LLM Dokumenten Analyse



Korrektur

Übersetzung

Zusammenfassung





Geschichte entschlüsseln mit Transkribus

m.prien@transkribus.org

https://www.transkribus.org/