

## Innovating Historical Scholarship: The Bullinger Digital Project

DH Colloquium @ SADiLaR January 31, 2024 Phillip B. Ströbel, Department of Computational Linguistics @phillipstroebel @phillipstroebel@techhub.social

## What will expect you in this talk:



## Shameless self-advertising

Ströbel, Phillip; Clematido, Sim	e tricica Schwitter, Rapha	ael; Ströbel, Phillip; Suter, Ber
/ I R models	Martin; Schwitter, Raphael; Hodel, Tobia Volk, Martin; Fischer, Lukas; Scheurer, Patricia; Schwitter, Rapha	e Corpus of 16th Century Let
n: LREC 2022, Marseille, 21 June 2022 - 23	Scheurer, Patricia; Raphael, Müller; Bernard, Schroffenegger; Ströbel, Phillip; Benjamin, Suter; Volk, Martin (2022).	)22, LREC.
	→ Ein Briefwechsel-Korpus des 16. Jahrhunderts in Frühneuhochdeutsch. In: Kupietz, Marc; Schmidt, Thomas, Neue Entwick the	(2022).
	In: Kupietz, Marc; Schmidt, Thomas, Neue Entrie II. Tübingen: Narr Francke Attempto Fischer, Lukas; Scheurer, Patricia; Schwitter, Raphael; Volk, Martin Fischer, Lukas; Scheurer, Patricia; Schwitter, Raphael; Volk, Martin Machine Translation of 16th Century Letters from Latin to Germ	an. d Ancient Languages (LT4HA
UZH	Fischer, Lukas; Scheuler, Future Prancke Attempto → Machine Translation of 16th Century Letters from Latin to Germ → Machine Translation of 16th Century Letters from Latin to Germ In: Second Workshop on Language Technologies for Historical and Marseille, 25 Juni 2022. LREC, 43-50.	3

## I will jump all over the place ...

# ... in the hope to provoke some thoughts and a fruitful discussion

### I want to start with our most recent contribution

Multilingual Workflows in *Bullinger Digital*: Data Curation for Latin and Early New High German

PHILLIP BENJAMIN STRÖBEL 
LUKAS FISCHER
RAPHAEL MÜLLER
PATRICIA SCHEURER
BERNARD SCHROFFENEGGER
BENJAMIN SUTER
MARTIN VOLK 
\*Author affiliations can be found in the back matter of this article

Journal of Open Humanities Data

> COLLECTON: DATA AND WORKFLOWS FOR MULTILINGUAL DIGITAL HUMANITIES

**RESEARCH PAPER** 

]u[ ubiquity press

UZH

### What is *Bullinger Digital* about?

5.0. How at an free himper parties an intelli-

Appendix of herd forwards analysis appropriate of a Copyrel constraint on the second second second second in the second second second second second second second is a second second second second second second second appropriate second The share the same wind find and going proved, site of the dynamic starting the same starting the spectra of the spectra of the same starting the spectra of the same starting the start of the same wind the going the spectra field of the same start of the spectra one starting for the starting the start of the spectra one starting for the spectra of the spectra of the starting for the spectra of the spectra of the spectra one starting for the spectra of bar Richard and Box garify misserer may Rim Aque San Birthan Co ward and warft ofter full and ward you you do not a start war war and warthan . I a fight and the sand th form in wing bigg and a property for the series Herein

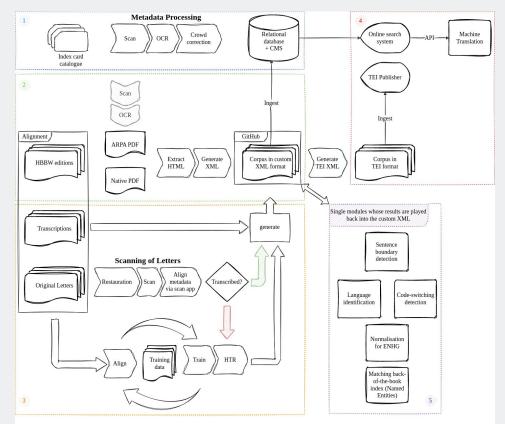




~ 2k

(3) Alter als sea parts designer, autor autority and an antipart designer autority and an antipart designer autority and an antipart designer autority and antipart and antipart designer autority and antipart and antipart designer autority and antipart and antipart and antipart and antipart designer autority and antipart designer autority and antipart and antipart and antipart and antipart designer autority and antipart and antipart and antipart designer autority and antipart and antipart designer autority antipart designer autorit The share between the first states they marked with of the first states of the states of the states of the of the first states of the states of the states of the the states of the states of the states of the first states of the states of the states of the the states of the states of the states of the the states of the states of the states of the the states of the states of the states of the states the states of the states of the states of the states the states of the states of the states of the states of the the states of the states of the states of the states of the the states of the states of the states of the states of the the states of the states of the states of the states of the the states of the states of the states of the states of the the states of the states of the states of the states of the the states of the the states of the stat san Ejlopen our gos garfy missered and this ford san to upon the new ain enough the figt and new in your sufficient and son work and referrer i die fifte And by another with a metal of the second state of the formation of the second state o Herein

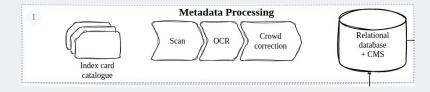
### Workflow in Bullinger Digital



### UZH

### **Metadata Processing**

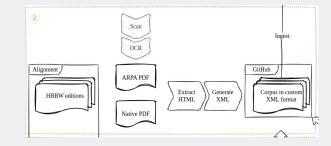
- Scanning of index card catalogue
- OCR
- Crowd sourcing
  - Correction of the OCRed index cards
  - 6 months during COVID
- Starting point for a relational DB



1 - 1 3		
Datum	Absender	Empfänger
1555 Oktober 22.	Beza Theodor	Bullinger Heinrich
	Genf	Zürich
¥ 12		
Autograph	Kopie	Photokopie –
Standort Zurich & A	Standort Ben Enger 8 - B III 61 (IMI , CA II 32) p. 934-928 - elundor p. 1096-1100	Bull. Corr.
Sign. EI 368, 63 ff.	Sign.	Abschrift ZB (Druck)
Umfang	J=nie 28 M: 5 85,215 Umfang	Bull. Corr. 20 B1.2, S.3
Sprache	Literatur	× .
Gedruckt - C.O., XV, \$36 ff. (W. 2331) - Com. A. Egg. 1:178-182 N. 67 D. J Com. A. A. M. 172 N. 67		
-TD : Torshard Brit. And Inin three W285, 289. (and En 28, (2 10 2)	Ex Calvini ## nost Bemerkungen inique et calumnic faciant	ri literis cognosces quam se, id est, quam suo more

# Add to this $DB \rightarrow Add$ already edited transcriptions

- ~20 already published editions since 1974
  - = 3'000 edited letters
  - The editors at the Institute for Swiss Reformation Studies (www.irg.uzh.ch) added around 80k (!!!) footnotes
- In 2016:
  - <u>www.teoirgsed.uzh.ch</u>
  - Digitised (already published) editions
    - Text available in the form of searchable PDFs
    - Limited search options



Heinrich Bullinger Werke

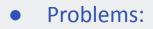
Zweite Abteilung: Briefwechsel Band 20: Briefe von April bis Dezember 1547

Anhang: Neue Briefe aus den Jahren 1523 bis 1546

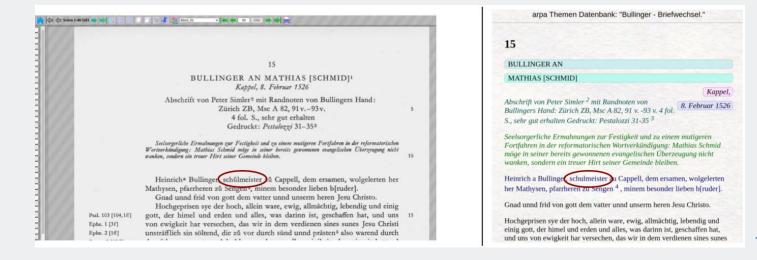
#### тνz

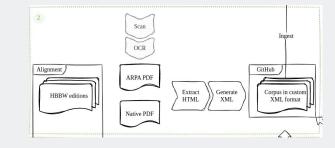
Theologischer Verlag Zürich

### teoirgsed



- Special characters (e-caudata *ę* used in Latin or the "u with a ring on top" *ů* for the Early New High German diphthong *uo* not represented :-(
- First 7 editions: PDF only, for the rest, TUSTEP (  $\rightarrow$  :-( ) and these digitised PDFs.





- + about 5'000 provisional transcriptions in ".txt" format
- Some leftovers of TUSTEP markup
- Extraction and structuring of text and metadata

- Together with the edited texts → transformation to custom XML
- Put everything in GitHub

```
Alignment

HBBW editions

Native PDF

Transcriptions

Generate

Mathematical Structure

Mathematical Structure

Mathematical Structure

Mathematical Structure

Mathematical Structure

Structure

Mathematical Structure

S
```

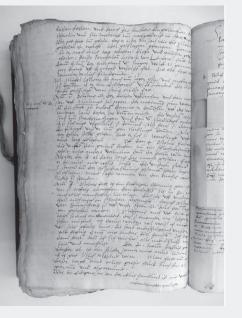
```
<regest fk bibliography="5" nr="167" page="25">
       <div ref="1">
               Bedankt sich für Bullingers Empfehlung an einen Gelehrten.
       </div>
</regest>
<letter lang="la" edition="HBBW" vol="3" fk bibliography="5" nr="167" page="25">
       <div ref_regest="1">
               <s n="1" lang="la" state="auto">Cum nuper a te discederem<fl>3</fl>, literis te m
                       <s n="2" lang="la" state="auto">Quod equidem aggredi vix essem ausus, ut scilicet
                       <s n="3" lang="la" state="auto">Unde spero, deo optimo maximo auxiliante, non mec
                       <s n="4" lang="la" state="auto">Tantum enim studium in docendo adhibet, ut verbi:
                       <s n="5" lang="la" state="auto">De quo memorare tibi multa possem, ni me inde nec
                       <s n="6" lang="la" state="auto">Meum tamen erit, quod referre nequeo, perpetuo m€
                       <s n="7" lang="la" state="auto">Verum utinam eam mihi facultatem aliguando conced
               <s n="8" lang="la" state="auto">Vale megue tibi semper guam commendatissimum habe
               <s n="9" lang="la" state="auto">Datum <placeName ref="28">Basilee</placeName>, ar
               <s n="10" lang="la" state="auto"><persName ref="8465">Melchior Virbius</persName
               <note type="address">[Adresse auf der Rückseite:]</note>
                       <s n="11" lang="la" state="auto">Et pietate et eruditione viro praeclaro d. Heinr
               </div>
</letter>
<footnotes fk_bibliography="5" nr="167">
       <alpha>
               <fn ref="a">scientia <i>am Rande nachgetragen.</i></fn>
               <fn ref="b"><i>in der Vorlage</i> praeditum; <i>auch <persName ref="2979">J. J. Simler
       </alpha>
       <num>
               <fn ref="1" meta="1" type="bio" fk_person="1365">Melchior Wirbens Herkunft ist unbekannt
```

12

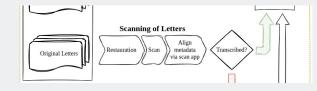
<fn ref="1" meta="1" type="bio" fk\_person="1365">Melchior Wirbens Herkunft ist unbekannt.



## Digitisation of letters @ Central Library & State Archive Zurich





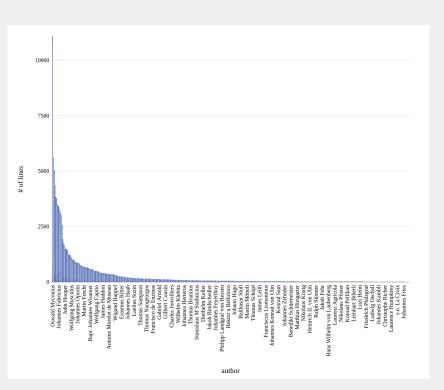


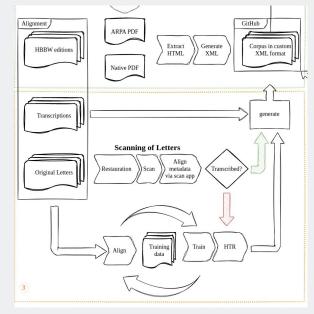
### Metadata alignment with Scan-App

Bullinger Digital Scan Matching   Zuweisung Dateien Statistiken C	SV-Import ZIP-Export
Start-ID: 7 Alle v 13159 Briefe mit die	esen Kriterien
Nächster Brief +1	
D HBBW- Empfänger / Absender / Datum Z Nummer Georg von Württemberg-Mömpelgard (Montbéliard) an Heinrich Bullinger 4. Januar 1548	Incipit / Transkription Georg, grafe zu Württemberg und zu Mumppelgart etc. Unsern grus zuvor, wolgelerter, lieber, besunder. Das wir euch uff ewer jungst schreiben, an uns gethon, wytters nichts zugeschriben, an solchem haben uns fürgefallene unsere geschefft verhindert. Schicken euch hiebei von unserm predicanten zu
Dokumente 🦪 Mit CMS abgleichen 🖪 Im CMS öffnen	
Autograph: E II 363,107 Korrekt Unklar Nachtrag in DB: Zugewiesene Scan-Dateien Korchau anzeigen Als korrekt markiert am 12/8/2021, 10:3	17:34 AM
E_II_363fol107_rtif /Scans_StA_transkribiert/StA_Scans_transkribiert_2021_07_23/E_II_363fol107_rtif	Quelle: CSV_IMPORT, erstellt am 12/7/2021, 10:00:31 AM Reihenfolge: 417 Quell-Signatur: E II 363 (fol. 107 r)
Brown Brown Bring	naire Ser meinsete emberg Sin Mannogeregasst 2000 5 100 ego cor tor - soborc hos inn den? wor innest of curichan an hore an

### Handwritten Text Recognition







### UZH

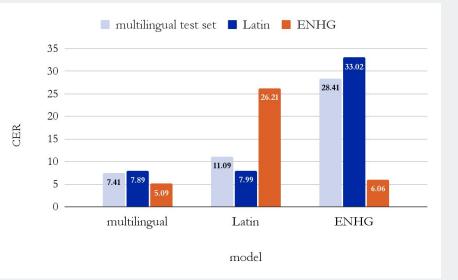
### **Different handwriting styles**

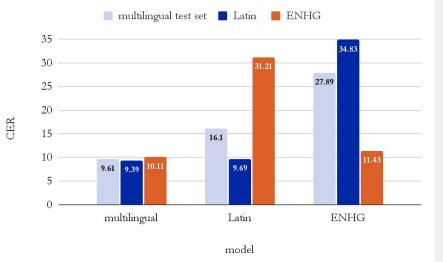
Grotin a par a des patre por christin Tra serie Remisso ad se liber aller, mit guide, fed som mon quisq farthe Leo Jud [Zürich ZB, Ms F 62,343] [Zürich ZB, Ms F 62,23 1 fol S.] his queen vides Carolfradius of Homo fant calis Latin's noting afflictionious. andit allour alia sim putiof. fruit of. **Huldrych Zwingli** Markus Bertschi [Zürich StA, E I 3.1, Nr. 63 1 S. 4°] [Zürich ZB. Ms A 51.10 2/3 fol. S.] Gerryi 2003 Good bis Buthampunt, going storia true > ques profime ad me fort fis mili non fort redu Velaty my Aporto **Johannes Rhellikan** Anne Hooper [Zürich ZB, Ms F 62,459 1 fol. S.] I find your bot ticker any transity winfind Grad and first was got goin watthe el-Canon Sigpe Wolfgang Rüpplin Ulrich Stoll [Zürich StA, A 230.1, Nr.159 a 1/3 fol. S.] ly New High German Bürgermeister und Rath von Basel

### 3 test sets and 3 different models

- A global test set  $\rightarrow$  10% of the overall data
- A test set consisting of lines from frequent writers → wrote >10 letters to Bullinger
   1,235 lines
- A test set consisting of lines from **infrequent writers**  $\rightarrow$  wrote <10 letters to Bullinger
  - **1,013 lines**
- Fine-tuned TrOCR on
  - $\circ$  all of the training data (  $\rightarrow$  multilingual model)
  - $\circ$  Latin lines only (  $\rightarrow$  Latin model)
  - $\circ$  ENHG lines only (  $\rightarrow$  ENHG model)

### Results





Frequent writers

Infrequent writers

UZH

### Result in the interface

#### Faksimile

#### Autograph: Zürich StA, E II 368,303 f. + 296e

Noy statuer am allas ad helideas dare, mei prins accesso altimit meis refinit si homen eas reformer dignas reputares. De accepte nona et interpertata strukent occassio, guan milis sont hieris ad xxi telemänis scripti persecut 3: Auleras. De nille ait expresse pro mondato. Come gallica reada l'atime reste ad reading ved his agrenere risom est. D. Bulingerns afrigais me russas ad se moding terbins vor anarrant accessand guerimemos et expertulationes com Ant. Galitar proper v illi apportenze, significarenze, mili an en ere con non reading, mis ved his ga duelong faces, significarenze, mis illi et une shire providere de alse galitar. Permit di iller reconciliandes, mi mani illi et une filie providere de alse galitar. Alur d'ere provide and experience allegentus est dicent, melestiam guan de print destre reconciliandes, mi mani illi et une filie providere de alse galitar. Alur d'ere provide me de facere allegentus est al constrant que destrant que de

#### Transkription

#### Automatisch erstellt

#### 3.

v 🖭

Non statueram illas ad te literas dare, nisi prius accepto ultimis meis responsum tamen eas responso dignas reputares. Sed accessit nova et inexpectata scribendi occasio, quam mihi suis literis ad 30. februarii scriptis porrexit d. Mulerus, idque ut ille aut expresso tuo mandato. Cuius Gallica verba Latine verbo ad verbum reddit, hic apponere visum est. D. Bulingerus (inovit) me rursus ad se, nudius tertius vocatus narrans nescio quas querimonias et expostulationes cum Antembitas propter illi appositum, iussitque, ut, quoniam ego illos in suam domini collocavi, id tibi, qui admodum facio, significarem, nihil in ea re cernens remedii, nisi ut tu ipse pemas ad illos reconciliandos, ni in manis illi et tuo filia proindere de alio hospitio. Ita filius d. Bullingeri perendie me de hac re, alloquutus est dicens, molestiam, quam pater desumpserat morbum illi concitasse. Quod ad me attinet, ego nihil in posterum

### Language identification, NER, normalisation

Faksimile Transkription 🕕 💵 Übersetzen 🗹 Sprachen markieren ~ • Kopie: Zürich StA, E II 351,14v.-15r. Farblegende: Deutsch (51.4 %) Latein (48.6 %) S. Mitto Osiandri responsionem ad carmen illud famosum, quo est lacessitus. Iam olla as formo of tarmone, quan ilis dasim inform: meres? habet operculum. Et iterum nobis iustae querimoniae oblata occasio est, ut nec quieta ab ulla in Literas vel Latinas vel vernacilas relata lex nec integra esse nobis ulla studia queant, laetante interim et sub pallio canente pontifice bestremok Abbatim historia ia. Interim Frechtus ad me scripsit tum Early New High German Vala. S: Galli . In. Calend: urbium mire in nostram salutem Vao h uß etwas widerwilles rö(misch) keiß. Nun spricht aber von selbigen apostolus, daß sölchs alles thugind umb rums willen, den menschen zu gefallen, unnd das sy dem crütz entlouffend. Dno Hermicho Bulling marime' soler ten und die hut doran ze binden; dann Mitto Diandre responsionen ano est lace sito fam olla ha nd, es wurde an anderen fryen stenden Normalised wiewol das bůbenvolk sampt ettlichen nobis infta maximoma re doch nur nicht ze gedenken, das die Nun spricht aber von selbigen Apostolus, dass solches alles Tücken um Rüms ren lassen werdind, das man ein mer integra else nobis vela for willen, den Menschen zu gefallen, und das sei dem Kreuz entlaufen. zer machen, dann meer Schwyzer et Part ballio candate Dont etat, tute ipse ex literis intelliges, ex tom alys of reog, tom una 10 Suchresultate von 61 Sortierung: Relevanz ~ 4 Georg von Württemberg-Mömpelgard (Monthéliard) an Heinrich Bullinger 🗎 10. April 1552 🚽 Schaffhausen Bref a) rootset Solchen und dergleichen missive haben wir euch den 3ten aprilis bey dem fürman, so den wein gefüert, toohot. Nhun vernemmen wir, das der wein zu Zürich und aber der brieve verloren worden sey (wie der würt zum tentor, Rappen alhie heut dato angezaigt, begern auch von euch zu wissen, ob der wein schadhafft worden sey), trotter derhalben üch dann bitzhär solcher wein nit überantwort worden. O Details

[Heinrich Bullinger] an [Hans Rudolf Bullinger]

water Wenn du mir wilt neißwas wyn gen, so

🚆 Zwischen 1573 und 1575 🚽 (Zürich) 🚺

b)

O Details

Generate Corpus in TEI XMI TEI format Single modules whose results are played back into the custom XML Sentence boundary detection Language Code-switching identification detection Normalisation for ENHG Matching backof-the-book index (Named Entities)

eidgnoschaft treffelich rüstind, uns by

cognoscere. Stumpffen Chronica die

[Johannes Bullinger] an Heinrich Bullinger

🛱 Ica. November, zwischen 1532 und 15371 🌱 (Ottenbach)

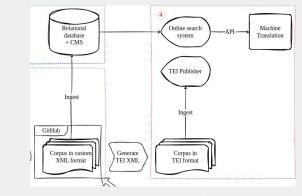
Boston immediate. Witter wirdt Michel Widmer den win diff

O Details

17H

20

### Machine translation





#### Impressum

Seite



Faksimile

Kopie: Zürich StA, E II 351,14



Latein - HBBW 2066

Dokumente

Latein-Deutsch-Übersetzer

#### S.

Mitto Osiandri responsionem ad carmen illud famosum, quo est lacessitus. Iam olla habet operculum. Et iterum nobis iustae querimoniae oblata occasio est, ut nec quieta nec integra esse nobis ulla studia queant, laetante interim et sub pallio canente pontifice Romano vel una dissensionis nostre confidentia. Interim Frechtus ad me scripsit tum aliis de rebus, tum una illa praecipua: de animis urbium mire in nostram salutem inclinatis. Ich hab im gschriben, ich besorg mich uß etwas widerwilles rö(misch) keiß. Mt. schwerer kriegen und das wir uns inn einer eidgnoschaft treffelich rüstind, uns by harkomnes unser fryheiten und libertet zu erhalten und die hut doran ze binden; dann wol zu gedencken, wann wir ußgemacht werend, es wurde an anderen fryen stenden ouch

#### Deutsch

Î

#### Ich grüsse Dich

Ich schicke Osianders Antwort auf das himmlische Gedicht, das er angegriffen hat. Die Töpfe haben eine Decke. Und wiederum bietet uns eine Gelegenheit zu gerechter Klage, dass weder Ruhe noch Unversehrtheit für uns sein können. während inzwischen mit Freude und mit einem Mantel, der den römischen Papst singt, noch mit einem einzigen Vertrauen auf unsere Meinungsverschiedenheit. In der Zwischenzeit schrieb mir Frecht sowohl über andere Angelegenheiten als auch über die grundlegende Angelegenheiten: Über die Gemüter der Städte neigt sich wunderbar zu unserem Gruss. Ich hab im gschriben, ich besorg mich uß etwas widerwilles rö(misch) keiß. Mt. schwerer kriegen und das wir uns inn einer eidgnoschaft treffelich rüstind, uns by

itus. Iam olla ut nec quieta nente pontifice ne scripsit tum salutem ö(misch) keiß. üstind, uns bv binden: dann fryen stenden ampt ettlichen enken, das die nan ein Schwyzer telliges, ex Chronica die

0 1 1

11/

UZH

### **TEI(-Publisher)**

Von Heinrich Bullinger, Wolfgang Rüpplin (Joner) an **Rudolf** Asper



30. November 1523 🗹 Kappel am Albis Brief

#### Metadata

Date	30. November 1523
Absender	Heinrich Bullinger, Kappel am Albis Wolfgang Rüpplin (Joner), Kappel am Albis
Empfänger	Rudolf Asper
Sprache	Latein
HBBW-Briefnummer	HBBW, Nr. 1b

#### Regest 6

Im Namen von Abt Joner fordert er dessen durch die Reformationsfreundlichkeit Joners entfremdeten Freund Asper auf, die alte Freundschaft wieder herzustellen und ihm (d.h. dem Abt) auf dem Wege zum Verständnis und zur Anerkennung des reformatorischen Schriftprinzips zu folgen. Er schildert diesen Weg als seine eigene Erfahrung, wie er von den Scholastikern, besonders den Dekretisten, durch die Kirchenväter endlich zur Heiligen Schrift vorgestossen sei, gibt eine theologische Begründung der «sola scriptura» und verteidigt diese gegen römisch-katholische Einwände.

#### Transkription () Show Language

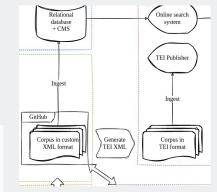
Epistola ad Rodolphum Asper de scripturae negotio. Egregio viro Rodolpho Aspero, amico syncerissimo Volcatius.

S. D. Haud certe iniuria de te conqueri possem, qui repente ex amico ardentissimo factus es admodum subcalidus. Quid nam est, quod in me desyderas? Quid tam venerandum vinculum rupit? Quis huius tam iniusti divortii author? Quis tam scelestus, tam nepharius tamque impius? Quae lingua tam virulenta ac letifera? Disperdat dominus universa labia dolosa! Quid non impia lingua potes? Quid, oro, neglectum ais? Queso, si quid est, quod alienum ab animo tuo geritur, indica et ne velis quorumvis insanis clamoribus et undique consarcinatis mendatiis fidem dare. Novisti in hunc usque diem malos demonas, invidos homines ac pugnaces amicitiae invidere. Si quid desyderas in me age, praesto sum, en veterem habes amicum et plus

#### Faksimile -)

+





### New avenues for digital editions ... Spoiler: it's all LLMs, or ...



### Translation

#### Latein-Deutsch-Übersetzer



Impressum

#### Latein - HBBW 2066

Dokumente

#### S.

Mitto Osiandri responsionem ad carmen illud famosum, quo est lacessitus. Iam olla habet operculum. Et iterum nobis iustae querimoniae oblata occasio est, ut nec quieta nec integra esse nobis ulla studia queant, laetante interim et sub pallio canente pontifice Romano vel una dissensionis nostrę confidentia. Interim Frechtus ad me scripsit tum aliis de rebus, tum una illa praecipua: de animis urbium mire in nostram salutem inclinatis. Ich hab im gschriben, ich besorg mich uß etwas widerwilles rö(misch) keiß. Mt. schwerer kriegen und das wir uns inn einer eidgnoschaft treffelich rüstind, uns by harkomnes unser fryheiten und libertet zu erhalten und die hut doran ze binden; dann wol zu gedencken, wann wir ußgemacht werend, es wurde an anderen fryen stenden ouch

#### Deutsch

前

#### Ich grüsse Dich

Ich schicke Osianders Antwort auf das himmlische Gedicht, das er angegriffen hat. Die Töpfe haben eine Decke. Und wiederum bietet uns eine Gelegenheit zu gerechter Klage, dass weder Ruhe noch Unversehrtheit für uns sein können, während inzwischen mit Freude und mit einem Mantel, der den römischen Papst singt, noch mit einem einzigen Vertrauen auf unsere Meinungsverschiedenheit. In der Zwischenzeit schrieb mir Frecht sowohl über andere Angelegenheiten als auch über die grundlegende Angelegenheiten: Über die Gemüter der Städte neigt sich wunderbar zu unserem Gruss. Ich hab im gschriben, ich besorg mich uß etwas widerwilles rö(misch) keiß. Mt. schwerer kriegen und das wir uns inn einer eidgnoschaft treffelich rüstind, uns by

#### ChatGPT

#### S.

Ich sende Osianders Antwort auf jenes berüchtigte Gedicht, durch das er herausgefordert wurde. Nun hat der Topf einen Deckel. Und wieder einmal wird uns ein gerechter Grund zur Beschwerde geboten, sodass keine unserer Bemühungen friedvoll oder vollständig sein können, während der römische Pontifex sich inzwischen freut und unter seinem Umhang singt, zuversichtlich in unserer Zwietracht. In der Zwischenzeit schrieb Frecht mir über verschiedene Angelegenheiten, darunter eine von besonderer Bedeutung: über die bemerkenswert günstige Gesinnung der Städte zu unserem Heil.

#### ChatGPT

#### S.

I am sending Osiander's response to that infamous poem by which he was provoked. Now the pot has a lid. And once again, we are presented with a just cause for complaint, so that none of our endeavors can be peaceful or complete, while in the meantime, the Roman Pontiff rejoices, singing under his cloak, confident in our discord. Meanwhile, Frecht wrote to me about various matters, including one of particular importance: about the remarkably favorable disposition of the cities towards our salvation.

### Automatic summaries

LETTER IN ENG GERMAN AND LATIN	HUMAN SUMMARY (GERMAN) FROM THE BULLINGER EDITION	GPT-4 SUMMARY (ENGLISH) [1] Greeting and New Year Wishes: Bullinger extends friendly greetings and wishes for a fortunate new year to Myconius, his wife, and household, noting the particularly perilous and critical state of current affairs. He calls for God's mercy in these challenging times.		
Früntlicher, fürgeliebter herr und bråder, ich wünschen üch, üwer lieben hußfrowen und dem gantzen hußgesind von gott durch unsern herren ein gåt glückhafft jar. Des bedörffend wir ouch wol, dann die sachen stand gantz spitzig und gfarlich. Gott der allmåchtig wö llesich unser erbarmen.	[1] Gute Wünsche zum neuen Jahr, auch an Myconius' Gattin [Margret] und an die Haus- angehörigen. Im Hinblick auf die gegenwärtigen Gefahren kann man Glückwünsche brauchen!			
Ulm wil nitt vil guts rhumbs haben der dappfferkeit. Man mumlet, sy vertrå gend sich gern mitt dem Sanherib; quod tibi dictum volo. Nå rlingen, Gien- gen, Boppffingen, Dinckelspühel, Rotenburg an der Tuber, Hall in Schwaben, alles und noch mee vomm keyser yngenommen! Ich bin gantz angsthafft umb den herrlichen gelerten Brentium.	[2] Die Stadt Ulm wird keine Lorbeeren für Tapferkeit ernten, denn angeblich hat sie sich recht leicht mit dem Sanherib" [Kaiser Karl V.] " öhnt! Dies als vertrauliche Mitteilung. Nördlingen, Giengen, Bopfingen, Dinkelsbühl, Rothenburg ob der Tauber und Schwäbisch Hall befinden sich in der Hand des Kaisers. Bullinger macht sich große Sorgen um den gelehrten [Johannes] Brenz.	[2] Military and Political Situation: He comments on the lack of bravery in Ulm and rumors of their alliance with Sanherib (possibly a metaphor for aligning with a powerful adversary). He reports the capture of several towns (Nördlingen, Giengen, Bopfingen, Dinkelsbühl, Rothenburg, and Hall) in Swabia by the Emperor. Bullinger expresses concern for the well-being of a distinguished scholar, Brentius,		

....

### Visual storytelling

[1] Bullinger wird sich wohl fragen, wie es der Kirche in Memmingen in dieser unruhigen Zeit geht. Alles steht noch auf der Kippe. Sollte Gott die Wünsche der Päpstlichen erfüllen, droht den Memminger [Pfarrern] die Todesstrafe! Momentan versucht man, der Stadt eine spanische Garnison aufzuzwingen. Einige Ratsherren wurden zum Reichstag nach Augsburg abgefertigt, um Kaiser Karl V. davon abzubringen. Sollten sie scheitern, ist es um die Religion, die Kinder und das Vermögen der Memminger geschehen! [2] Die in Biberach [stationierten Soldaten] führen sich völlig unmenschlich auf. In Ulm richten sowohl die Besatzer als auch die Pest Schaden an. In Augsburg regiert der Antichrist. Der Spanier tobt und der Landsknecht fletscht die Zähne, während die Kardinäle und Bischöfe sich um ihre Börse kümmern. [3] Wenn nur Christus mit Posaunenschall den Jüngsten Tag anbrechen lassen würde, denn Schuler sieht nicht, wie sonst den Gottesfürchtigen geholfen werden könnte! [4] Verschiedene Anzeichen deuten auf einen Feldzug des Kaisers gegen Konstanz hin. Allmählich werden Truppen gerüstet, die stets näher auf Memmingen, Kempten und Isny rücken. von wo dann ein Überraschungsangriff möglich wird. Da es unwahrscheinlich ist, dass jene Truppen beauftragt wurden, [den bereits gefallenen Städten] den Irrglauben der Vorfahren aufzuzwingen, ist mit ihnen wohl die Einstellung der eidgenössischen Freiheiten oder die Vernichtung von Konstanz beabsichtigt. Angeblich soll der Kaiser einen Vertrag mit dem französischen König Heinrich II. geschlossen haben; vielleicht gerade deshalb, um seinem Vorhaben bessere Chancen einzuräumen. [5] Kürzlich reiste durch Memmingen ein aus Augsburg kommender eidgenössischer Bote [...], der in der Öffentlichkeit mit großer Beredsamkeit behauptete, dass der Kaiser einen gewaltigen Krieg gegen Konstanz plane, und dass die aufseiten des Kaisers stehenden Fünf Orte sich dafür starkmachen werden, um dem Kaiser den Durchzug eines gegen Konstanz gerichteten Heeres durch ihr Gebiet zu gestatten. Stimmt das, dann stünde es auch um die Eidgenossenschaft schlecht! Die Zürcher sollen sich vor Verräterei hüten und aus dem Unglück der Deutschen lernen! Die mit dem Blut der Vorfahren erkaufte eidgenössische Freiheit wird beneidet; sie ist um jeden Preis zu verteidigen! Aus den Niederlanden sollen mit Schusswaffen beladene Wagen in Richtung Straßburg unterwegs sein. [6] Noch dürfen die Pfarrer in Memmingen ungehindert der Kirche Christi dienen: die Messpfaffen aber versuchen, dies mit allen Mitteln zu verhindern. Bullinger bete für die Kirche in Memmingen, und die Zürcher seien ihrerseits auf Attacken des Teufels und der Welt gefasst! Bullinger spende Schuler Trost, wann [immer] möglich, zumal viele Sorgen ihn bedrücken. Ansonsten geht es ihm und seiner Familie gut. [7] Grüße, auch an Anna [geb. Adlischwyler], an Rudolf Gwalther, Konrad Pellikan und Theodor Biblignder.





### For linguists ...

Part-of-Speech Tagging of 16th-Century Latin with GPPs proceedings of Elina Stüssi, Phillip Benjamin Sträppear in this Vear Social Department of Computational Linguistic Workshop on Cultural Heritage, Social University of Zurich 

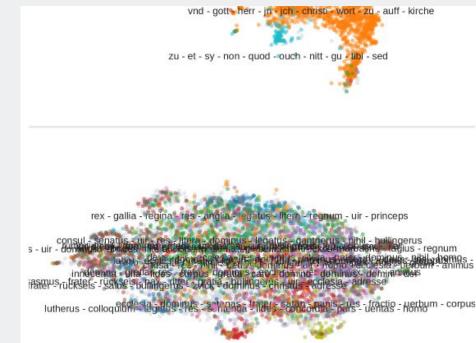
 Image: second procession
 Image: second procession

 Image: second procession

Tagger	Bullinger	ITTB	LLCT	UDante	Perseus	PROIEL	Avg TB
LatinCy	79.8	87.93	92.01	80.94	72.38	82.84	83.22
CLTK	77.1	88.45	79.32	77.36	72.44	80.63	79.64
UDPipe	72.8	71.59	71.45	70.51	69.2	84.49	73.45
RDRPOSTagger	64.5	82.67	67.41	71.72	64.79	75.22	72.36
TreeTagger	74.3	70.18	70.25	69.86	82.51	89.39	76.44
GPT-3.5-Turbo	62.2/80.2	74.82	78.82	74.33	68.81	79.22	75.2
GPT-4	74.3/83.5	79.73	84.89	77.62	73.9	84.38	80.1
train50	70.3/84.8	89.59	89.2	83.55	73.2	79.37	82.98
train100	69.4 <b>/85.5</b>	89.73	91.03	85.26	74.19	82.19	84.48
train200	68.2/80.0	91.57	90.93	85.8	73.46	83.5	85.05
train500	58.8/77.2	93.2	93.85	82.71	72.09	86.72	85.71
train1000	65.8/82.5	94.88	94.5	84.94	81.39	89.25	88.99
train2000	<b>78.2</b> /78.3	87.95	87.43	81.31	84.31	87.83	85.77
train5000	71.9/76.6	88.11	84.85	74.62	81.99	90.0	83.91
train10000	74.4/76.4	83.84	85.39	75.36	76.34	86.01	81.39

UZH

# Something that is not GPT, but still in development: Topic Modeling



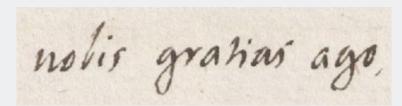
	vnd - gott - herr - jn - jch - christi - wort - zu - auff - kirche
	dies - litera - dominus - uir - res - annus - nihil - frater - gratia - ecclesia
	domino - gratia - res - frater - dominus - ecclesia - uerbum - pax - uir - anim
	ratio - res - litera - ecclesia - nihil - dominus - fides - opus - homo - gratia
	bullingerus - tigurus - uir - dominus - amicus - litera - iohannes - frater - hen
•	deus - uir - ecclesia - res - animus - litera - pater - dominus - nihil - homo
	causa - res - nihil - deus - dominus - uir - homo - ecclesia - uerbum - animu
٠	ecclesia - dominus - satanas - frater - satan - panis - res - fractio - uerbum -
	zu - et - sy - non - quod - ouch - nitt - gu - tibi - sed
۰	rex - gallia - regina - res - anglia - legatus - litera - regnum - uir - princeps
•	rumor - nihil - dominus - litera - urbs - res - uir - bullingerus - episcopus - fra
	italia - uir - res - frater - litera - nihil - bullingerus - pater - tigurus - curia
	lutherus - colloquium - legatus - res - sententia - fides - concordia - pars - u
0	labor - ecclesia - gratia - litera - uir - nihil - frater - dominus - animus - tempi
	doctor - ecclesia - doctrina - dominus - elector - deus - res - animus - litera
	erasmus - frater - rückseis - pax - ritter - gratia - bullingerus - uir - ecclesia -
	anglia - bucer - nr - bucers - ecclesia - gemeinschaft - eko - marbachs - fag
	innocentia - uita - fides - corpus - gratia - caro - domino - dominus - domini
۰	consul - senatus - uir - res - litera - dominus - legatus - gantnerus - nihil - bı
•	hainrichus - io - frater - rückseis - salus - bullingerus - zvick - dominus - chri

### What would we do differently?

- Work in TEI-XML from the beginning
  - Include the TEI-Publisher as well
- Do not underestimate the preprocessing steps!
  - Turning various input formats (PDF, TuStep, HTML) with historical symbols (e. g., ů) in multiple languages (in both Latin and Greek scripts) into a uniformly structured repository is time-consuming
- Unifying sources is not straightforward, even if call numbers are available (problems: typos, change over time).
  - Call numbers can be messy or outdated, but you need a system to keep track of your inventory
- Named Entity Recognition and Linking is challenging because of name variation (different spellings and names for the same person) and name duplicates (e.g., same name for father and son).
- Classifying the letters by topic is hard because of topic overlap and vagueness.

### If you want to use the Bullinger data:

https://github.com/bullinger-digital/bullinger-korpus-tei



### [vobis gratias ago]

### [Thank you]

