



University of
Zurich^{UZH}

Innovating Historical Scholarship: *The Bullinger Digital Project*

DH Colloquium @ SADiLaR
January 31, 2024

Phillip B. Ströbel, Department of Computational Linguistics



@phillipstroebel



@phillipstroebel@techhub.social

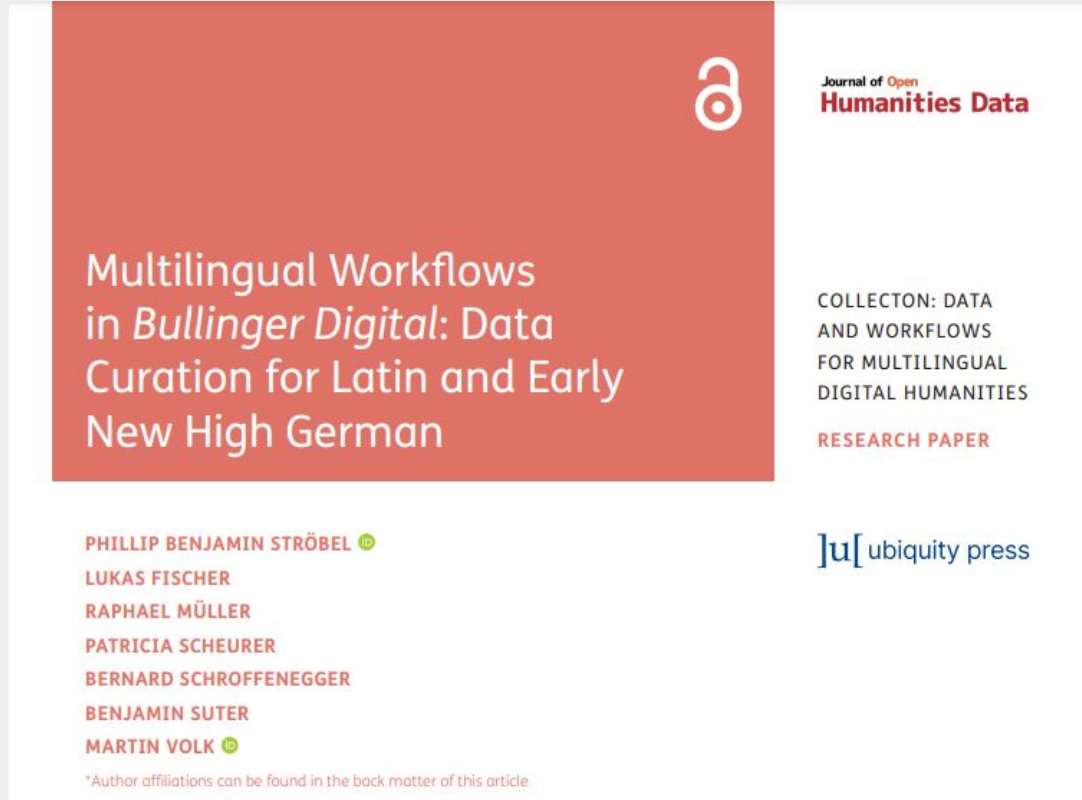
What will expect you in this talk:

Shameless self-advertising

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




The image shows the cover of a research paper. The top half has a red background with a white open lock icon in the top right corner. The title is written in white text on the red background. The bottom half has a white background with the journal title and authors listed in red and black text. The Ubiquity Press logo is in the bottom right.

**Journal of Open
Humanities Data**

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

**COLLECTON: DATA
AND WORKFLOWS
FOR MULTILINGUAL
DIGITAL HUMANITIES**

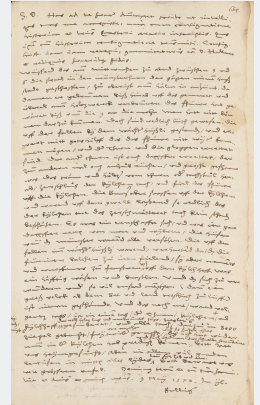
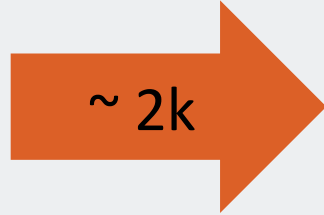
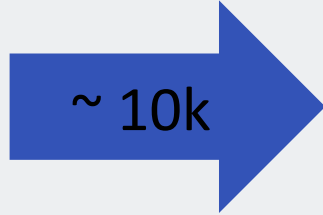
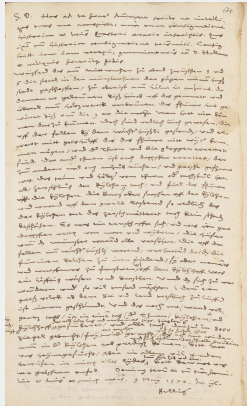
RESEARCH PAPER

PHILLIP BENJAMIN STRÖBEL 
LUKAS FISCHER
RAPHAEL MÜLLER
PATRICIA SCHEURER
BERNARD SCHROFFENEGGER
BENJAMIN SUTER
MARTIN VOLK 

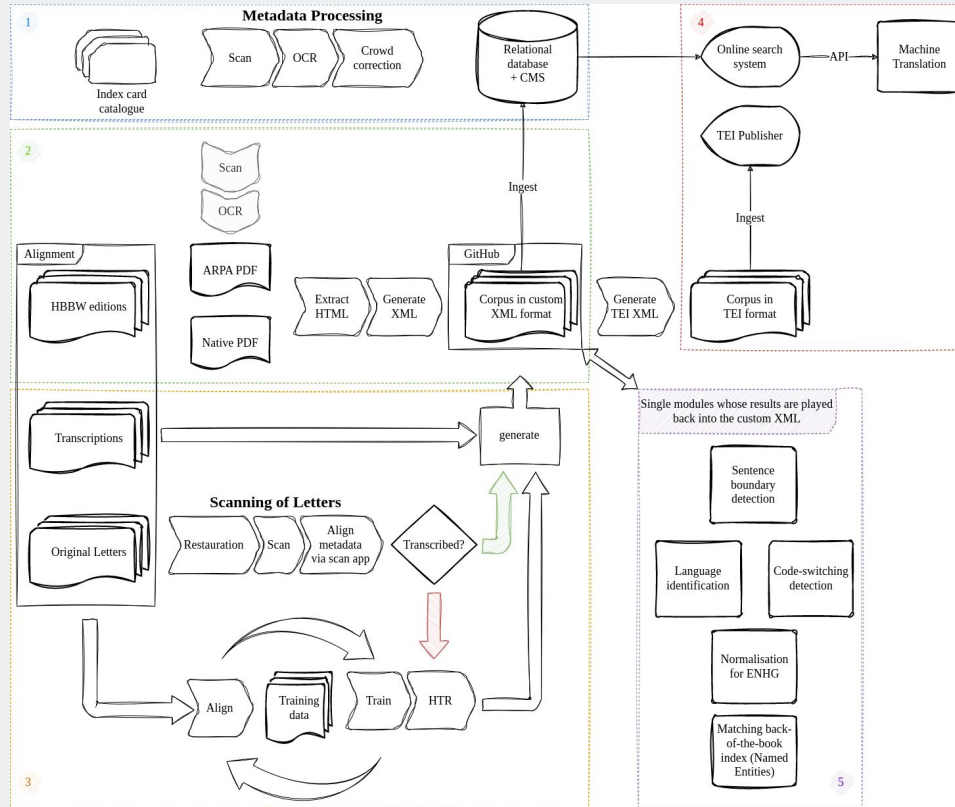
]u[ubiquity press

*Author affiliations can be found in the back matter of this article

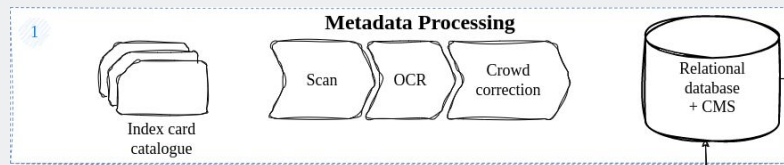
What is Bullinger Digital about?



Workflow in Bullinger Digital



Metadata Processing

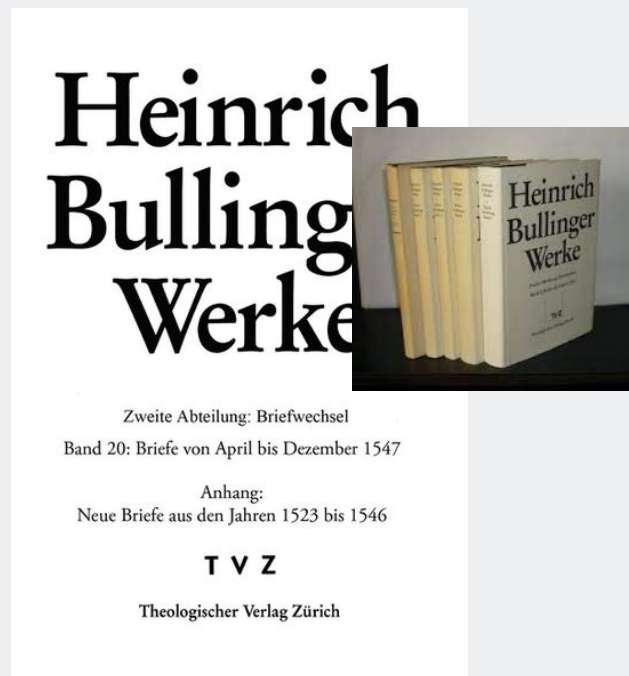
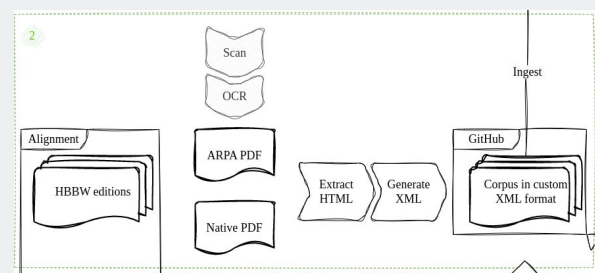


- Scanning of index card catalogue
- OCR
- Crowd sourcing
 - Correction of the OCRed index cards
 - 6 months during COVID
- Starting point for a relational DB

Datum 1555 Oktober 22.	Absender Beza Theodor Genf	Empfänger Bullinger Heinrich Zürich
Autograph Standort <i>Zürich H. A.</i>	Kopie Standort <i>Beza Burger B</i> <i>- B III 69 (H. CA VI 32) p. 934-938</i> <i>- skanda p. 1096-1100</i>	Photokopie - Bull. Corr.
Sign. <i>E II 368, 63 ff.</i> Umfang	Sign. Umfang <i>Janie 28. M: S 85, 215</i>	Abschrift ZB (Druck) Bull. Corr. 20 Bl. 2, S. 3
Sprache	Literatur	
Gedruckt - C. D. XV, 836 ff. (Nr. 2331) - Cons. de Reg. I 178-182 Nr. 67 - TD: <i>Treschard, Rev. A. Antiquaire II 285. 289.</i> <i>(Cons. de Reg. CA VI 32)</i>	Bemerkungen Ex Calvini A nostri literis cognosces quam inique et calumniose, id est, quam suo more faciant	

Add to this DB → Add already edited transcriptions

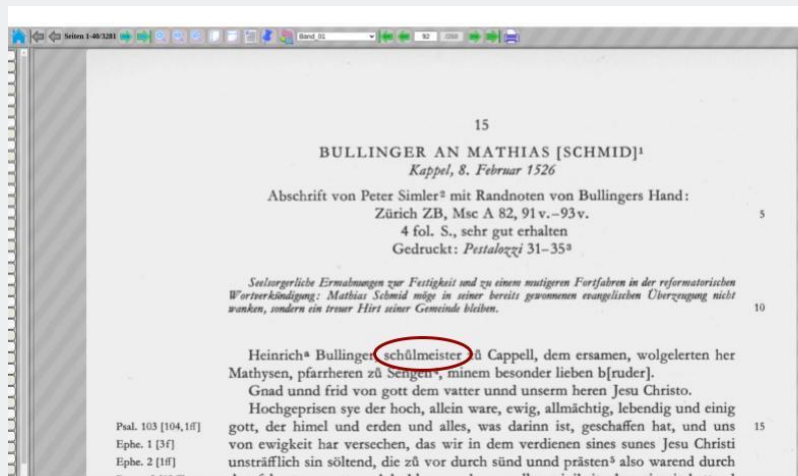
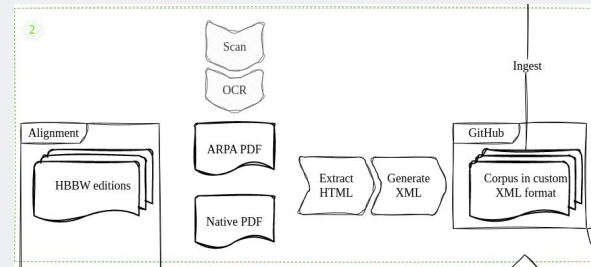
- ~20 already published editions since 1974
 - = 3'000 edited letters
 - The editors at the Institute for Swiss Reformation Studies (www.irg.uzh.ch) added around 80k (!!!) footnotes
- In 2016:
 - www.teoirgsed.uzh.ch
 - Digitised (already published) editions
 - Text available in the form of searchable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



teoirgsed

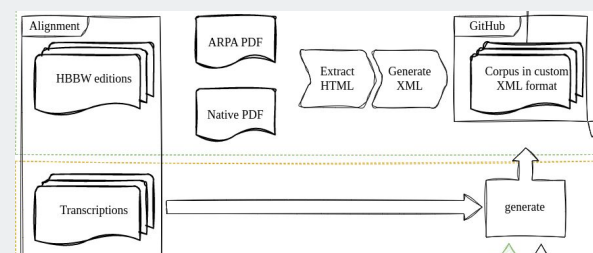
- Problems:

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



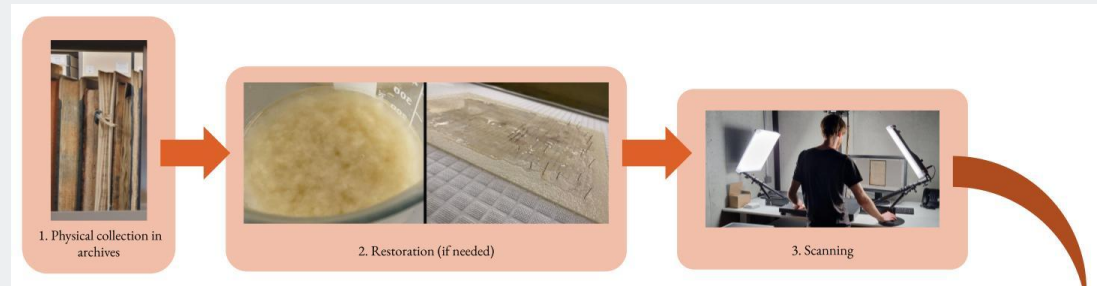
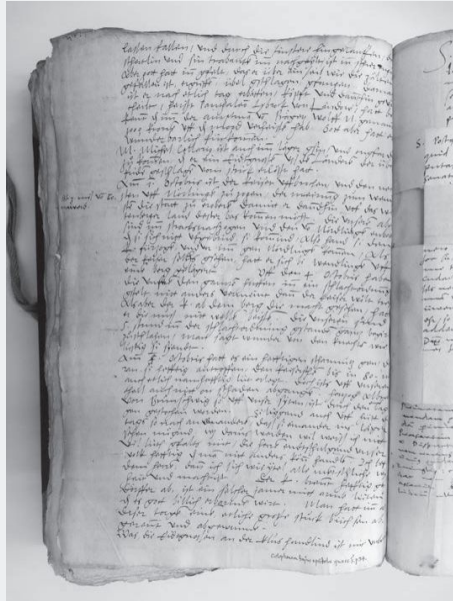
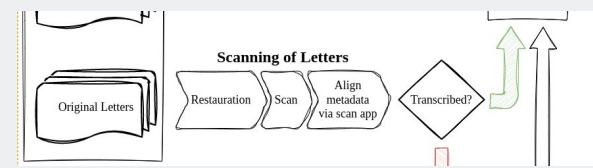
+ about 5'000 provisional transcriptions in “.txt” format

- 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

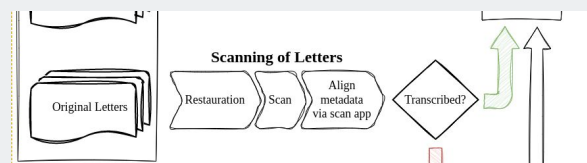


```
<regest fk_bibliography="5" nr="167" page="25">
  <div ref="1">
    Bedankt sich für Bullingers Empfehlung an einen Gelehrten.
  </div>
</regest>
<letter lang="la" edition="HBBW" vol="3" fk_bibliography="5" nr="167" page="25">
  <div ref_regest="1">
    <p>
      <s n="1" lang="la" state="auto">Cum nuper a te disciderem</s></p>
      <s n="2" lang="la" state="auto">Quod equidem aggredi vix essem ausus, ut scilicet</s>
      <s n="3" lang="la" state="auto">Unde spero, deo optimo maximo auxiliante, non me</s>
      <s n="4" lang="la" state="auto">Tantum enim studium in docendo adhibet, ut verbis</s>
      <s n="5" lang="la" state="auto">De quo memorare tibi multa possem, ni me inde ne</s>
      <s n="6" lang="la" state="auto">Meum tamen erit, quod referre nequeo, perpetuo me</s>
      <s n="7" lang="la" state="auto">Verum utinam eam mihi facultatem aliquando conce</s>
    </p>
    <p>
      <s n="8" lang="la" state="auto">Vale meque tibi semper quam commendatissimum habe</s>
    </p>
    <p>
      <s n="9" lang="la" state="auto">Datum <placeName ref="28">Basile</placeName>, ar</s>
    </p>
    <p>
      <s n="10" lang="la" state="auto"><persName ref="8465">Melchior Virbius</persName></p>
    </p>
    <note type="address">[Adresse auf der Rückseite:]</note>
    <s n="11" lang="la" state="auto">Et pietate et eruditione viro praeclaro d. Heinr</s>
  </div>
</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</i> <persName ref="2979">J. J. Simler</</alpha>
  <num>
    <fn ref="1" meta="1" type="bio" fk_person="1365">Melchior Wirbens Herkunft ist unbekannt.</fn>
    <fn ref="2" meta="1">Für eine genauere Datierung des Briefes liegen keine Anhaltspunkte v</fn>
    <fn ref="3">Über Wirbens Besuch bei Bullinger ist nichts bekannt.</fn>
  </num>
</footnotes>
```

Digitisation of letters @ Central Library & State Archive Zurich



Metadata alignment with Scan-App



Bullinger Digital Scan Matching | [Zuweisung](#) | [Dateien](#) | [Statistiken](#) | [CSV-Import](#) | [ZIP-Export](#)

Start-ID:

Alle ▼ 13159 Briefe mit diesen Kriterien

Nächster Brief +1

ID	HBBW-Nummer	Empfänger / Absender / Datum	Incipit / Transkription
<u>7</u>		Georg von Württemberg-Mömpelgard (Montbéliard) an Heinrich Bullinger 4. Januar 1548	Georg, grafe zu Württemberg und zu Mumpelgard 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 Vorschau anzeigen | Als korrekt markiert am 12/8/2021, 10:37:34 AM

E_II_363_fol_107_r.tif Quelle: CSV_IMPORT, erstellt am 12/7/2021, 10:00:31 AM
Reihenfolge: 417
Quell-Signatur: E II 363 (fol. 107 r)

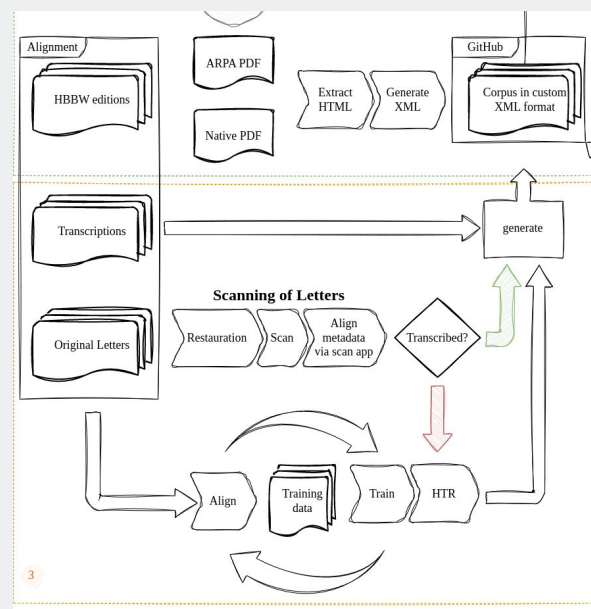
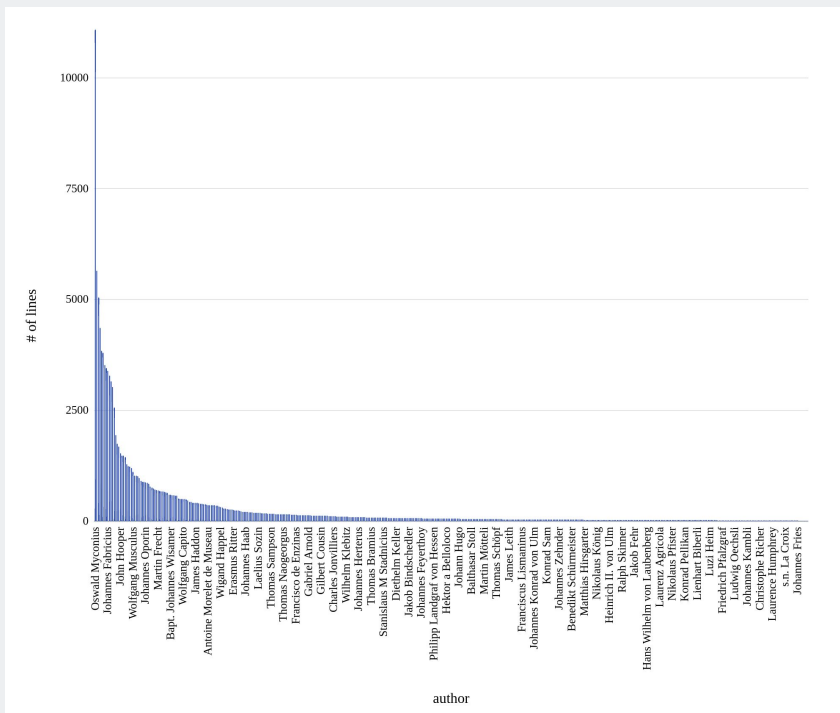
/Scans_StA_transkribiert/STa_Scans_transkribiert_2021_07_23/E_II_363_fol_107_r.tif

Georg Graf von Württemberg
und zu Mumpelgard 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

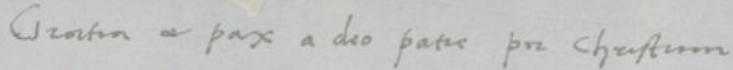
Handwritten Text Recognition



30k lines

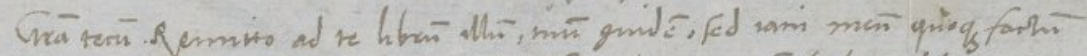


Different handwriting styles



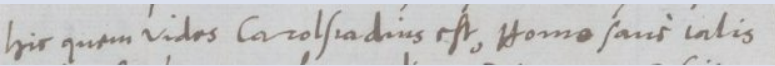
Gratia et pax a deo patre pro christum

Leo Jud
[Zürich ZB, Ms F 62,343]



Grā tuā Remitto ad te librum alium, tunc quidem, sed tunc non quidem factum

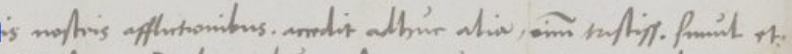
Ambrosius Blarer
[Zürich ZB, Ms F 62,23 1 fol. S.]



hic quem videt Carolus adius est, homo sanctus iustus

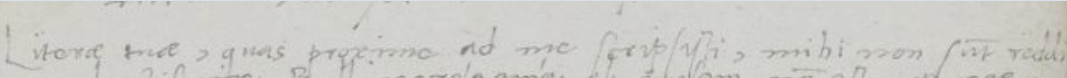
Huldrych Zwingli
[Zürich StA, E I 3.1, Nr. 63 1 S. 4*]

Latin



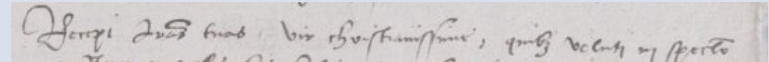
nostris afflictionibus. accidit aliter alio, cum tristis. fuit et.

Markus Bertschi
[Zürich ZB, Ms A 51,10 2/3 fol. S.]



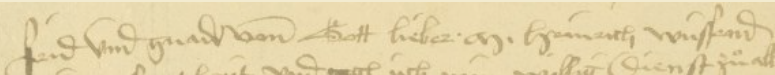
Litora tua, quas proximae ad me scripsisti, mihi non sunt reddita

Johannes Rhellikan
[Zürich ZB, Ms F 62,459 1 fol. S.]



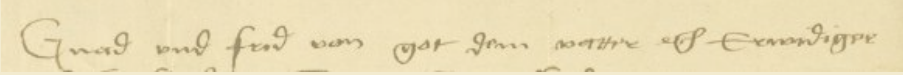
Georgi Jud Grad. Viri Christianissimi, quibus vobis in aperto

Anne Hooper
[Zürich StA, E II 369,203]



Seid und grad von Gott lobet. arg. bymuntz, wiffend

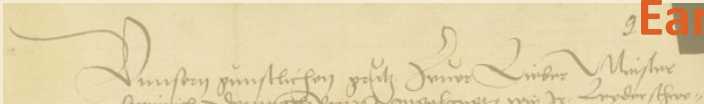
Ulrich Stoll
[Zürich StA, A 230.1, Nr.159 a 1/3 fol. S.]



Grad und fund von got dem watter eP. Ernandigere

Wolfgang Rüpplin
[Zürich StA, A 230.2, Nr.76 1 fol. S.]

Early New High German



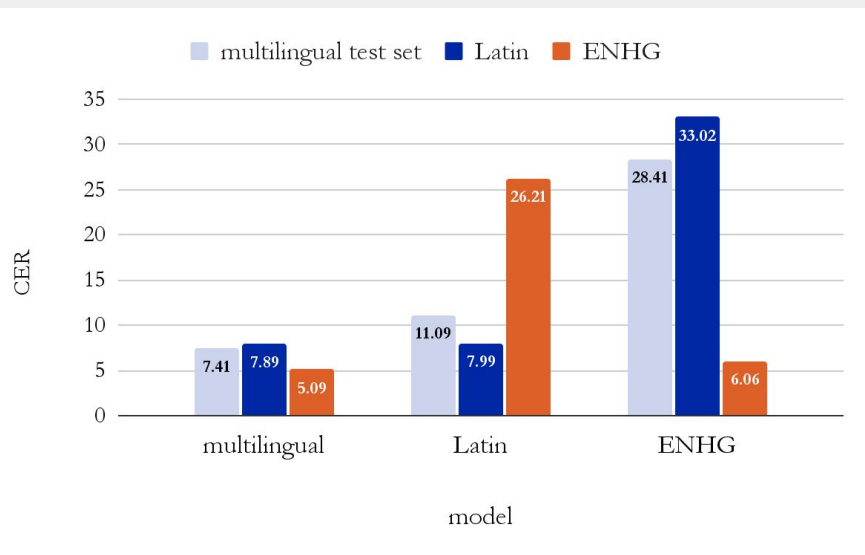
Immer zum besten geht immer. Immer. Immer. Immer. Immer.

Bürgermeister und Rath von Basel
[Zürich ZB, Ms A 51,9 1/2 fol. S.]

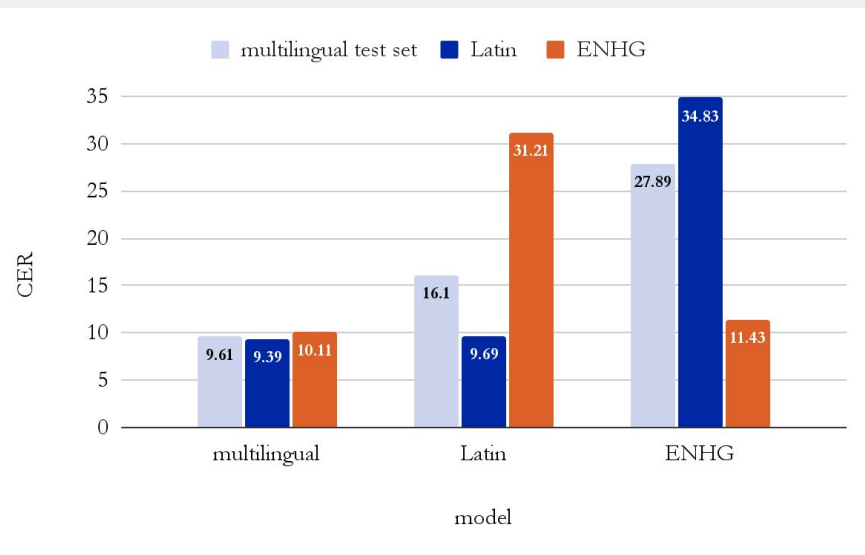
3 test sets and 3 different models

- A **global test** set → 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** → wrote <10 letters to Bullinger
 - 1,013 lines
- Fine-tuned TrOCR on
 - all of the training data (→ multilingual model)
 - Latin lines only (→ Latin model)
 - ENHG lines only (→ ENHG model)

Results



Frequent writers

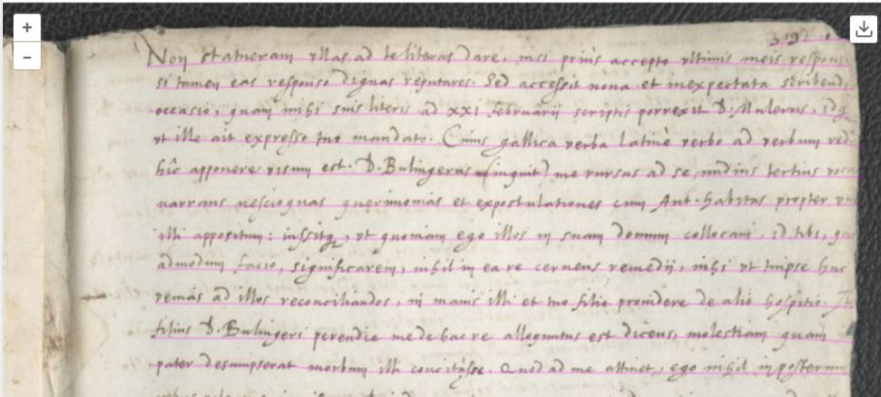


Infrequent writers

Result in the interface

Faksimile

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



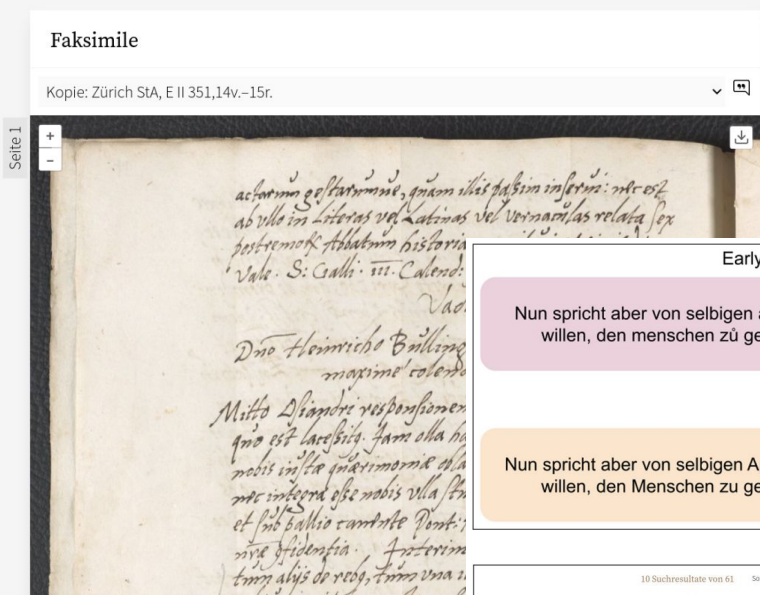
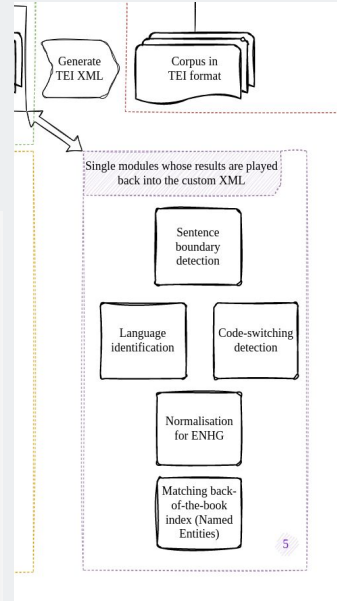
Transkription

Automatisch erstellt

3.

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 ait expresso tuo mandato. Cuius Gallica verba Latine verbo ad verbum reddit, hic apponere visum est. D. Bullingerus (inquit) 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 venias 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



Transkription i Übersetzen ☑ Sprachen markieren

Farblgende: Deutsch (51.4%) Latein (48.6%)

S. Mitto Osiandri responsonem ad carmen illud famosum, quo est laecessitus. 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. Interim Frechtus ad me scripsit tum urbium mire in nostram salutem h uß etwas widerwilles rö(misch) keiß. eidgnoschaft treffelich rüstind, uns by ten und die hut doran ze binden; dann ad, es wurde an anderen fryen stenden wiewol das bübenvolk sampt etlichen re doch nur nicht ze gedenken, das die gen lassen werdind, das man ein zer machen, dann meer Schwyzer etat, tute ipse ex literis intelliges, ex cognoscere. Stumpffen Chronica die

Early New High German

Nun spricht aber von selbigen apostolus, daß söchls alles thūgind umb rūms willen, den menschen zū gefallen, unnd das sy dem crütz entlouffend.

Normalised

Nun spricht aber von selbigen Apostolus, dass solches alles **Tücken** um **Rüms** willen, den Menschen zu gefallen, und das **_sei** dem Kreuz entlaufen.

10 Suchresultate von 61 Sortierung: Relevanz

a) Georg von Württemberg-Mömpelgard (Montbéliard) an Heinrich Bullinger

30. April 1552 📄 📧 📄

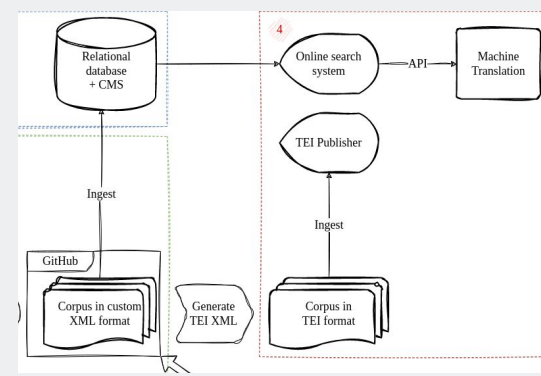
b) [Heinrich Bullinger] an [Hans Rudolf Bullinger]

Zwischen 1573 und 1575 📄 📧 📄

c) [Johannes Bullinger] an Heinrich Bullinger

14. November; zwischen 1520 und 1527 📄 📧 📄

Machine translation



Faksimile

Kopie: Zürich StA, E II 351,1

Latein-Deutsch-Übersetzer

Universität Zürich UZH

Impressum

Seite 1

Text Dokumente

Latein - HBBW 2066

S.

Mitto Osiandri responsonem ad carmen illud famosum, quo est lacessitus. Iam olla habet operculum. Et iterum nobis iustae querimoniae oblata occasio est, ut nec quietae nec integra esse nobis ulla studia queant, laetante interim et sub pallio canente pontifice Romano vel una dissensionis nostrae 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 kriegem 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 gedeenken, wann wir ußgemacht werend, es wurde an anderen fryen stenden ouch

Deutsch

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 kriegem und das wir uns inn einer eidgnoschaft treffelich rüstind, uns by

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

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

Transkription Show Language

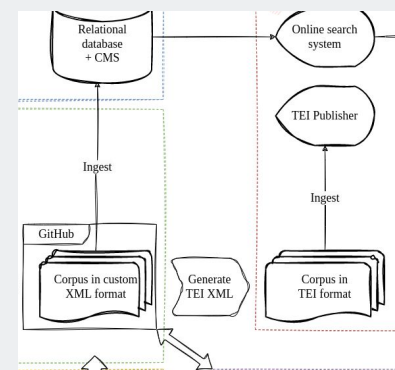
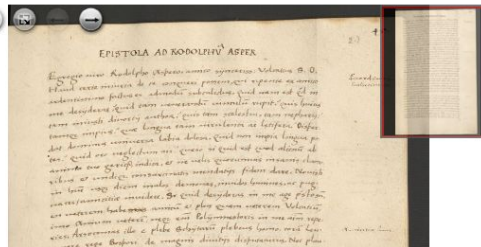
Epistola ad Rodolphum Asper de scripturae negotio. Egregio viro Rodolpho Aspero, amico syncerissimo Volcatius.

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 me velis quorumvis insanis clamoribus et undique consarcinatis mendatiis fidem dare. Novisti in hunc usque diem malos demonas, invidos homines ac pugnaes amicitiae invadere. Si quid desyderas in me, ago praesto sum, en veterem habes amicum et plus

Regest 1

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.

Faksimile 1



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



ChatGPT

Translation

Latein-Deutsch-Übersetzer



Impressum

Text Dokumente

Latein - HBBW 2066



Deutsch

S.
Mitto Osiandri responsonem ad carmen illud famosum, quo est lacesitus. Iam olla habet operculum. Et iterum nobis iustae querimoniae oblata occasio est, ut nec quieti nec integra esse nobis ulla studia queant, laetante interim et sub pallio canente pontifice Romano vel una dissensionis nostrae 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 kriegem 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 gedenccken, wann wir ußgemacht werend, es wurde an anderen fryen stenden ouch

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 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 kriegem und das wir uns inn einer eidgnoschaft treffelich rüstind, uns by

ChatGPT

S.

Ich sende Osianders Antwort auf jenes berühmte 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

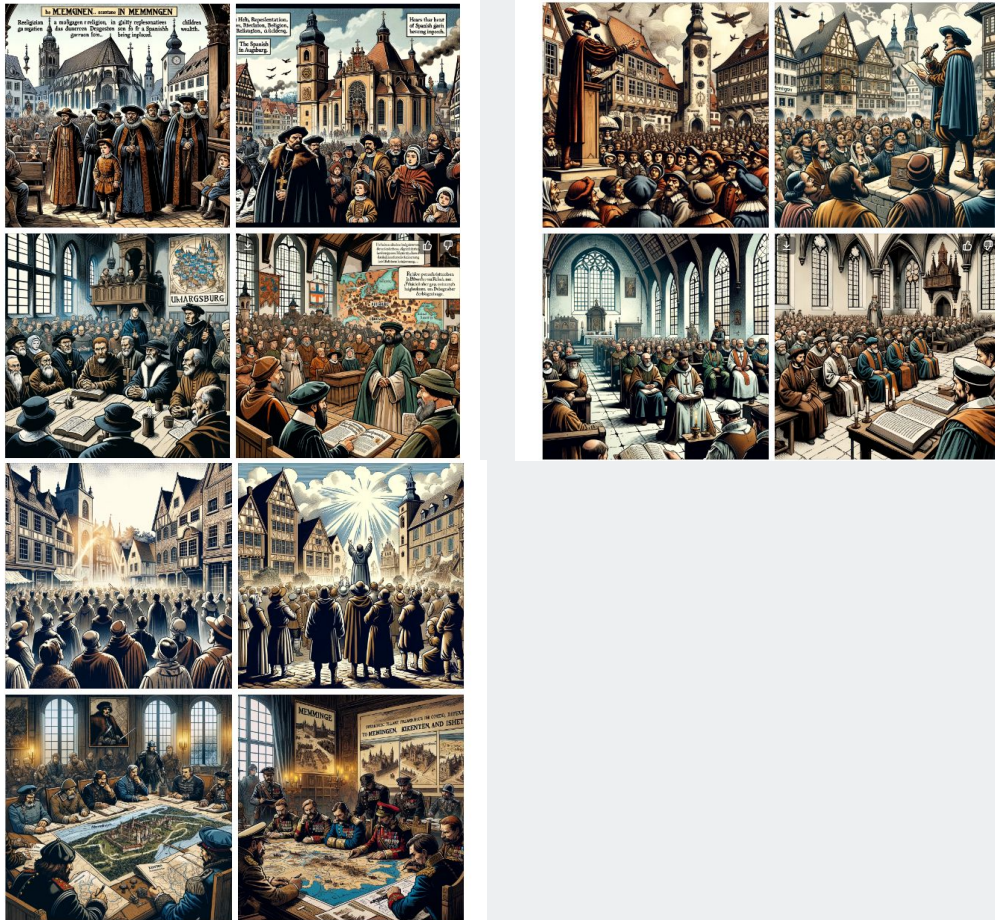
LETTER IN ENG GERMAN AND LATIN	HUMAN SUMMARY (GERMAN) FROM THE BULLINGER EDITION	GPT-4 SUMMARY (ENGLISH)
<p>Früntlicher, fürgeliebter herr und brüder, ich wünschen üch, üwer lieben hußfrowen und dem gantzen hußgesind von gott durch unsern herren ein güt glückhafft jar. Des bedörffend wir ouch wol, dann die sachen stand gantz spitzig und gfarlich. Gott der allmächtig wõllesich unser erbarmen.</p>	<p>[1] Gute Wünsche zum neuen Jahr, auch an Myconius' Gattin [Margret] und an die Hausangehörigen. Im Hinblick auf die gegenwärtigen Gefahren kann man Glückwünsche brauchen!</p>	<p>[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.</p>
<p>Ulm wil nitt vil güts rhũmbs haben der dappfferkeit. Man mumlet, sy vertrũgend sich gern mitt dem Sanherib; quod tibi dictum volo. Nõrlingen, Giengen, Boppffingen, Dinkelspühel, Rotenburg an der Tuber, Hall in Schwaben, alles und noch mee vomm keyser yngenommen! Ich bin gantz angsthafft umb den herrlichen geleerten Brentium.</p>	<p>[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, Bopffingen, 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.</p>	<p>[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, Bopffingen, Dinkelsbühl, Rothenburg, and Hall) in Swabia by the Emperor. Bullinger expresses concern for the well-being of a distinguished scholar, Brentius,</p> <p>...</p>

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 Posaemenschall 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 Verrätereihüten 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 Messpfeffen 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 Bibliander.



DALL-E



For linguists ...

Part-of-Speech Tagging of 16th-Century Latin with GPT

Elina Stüssi, Phillip Benjamin Ströbel

Department of Computational Linguistics

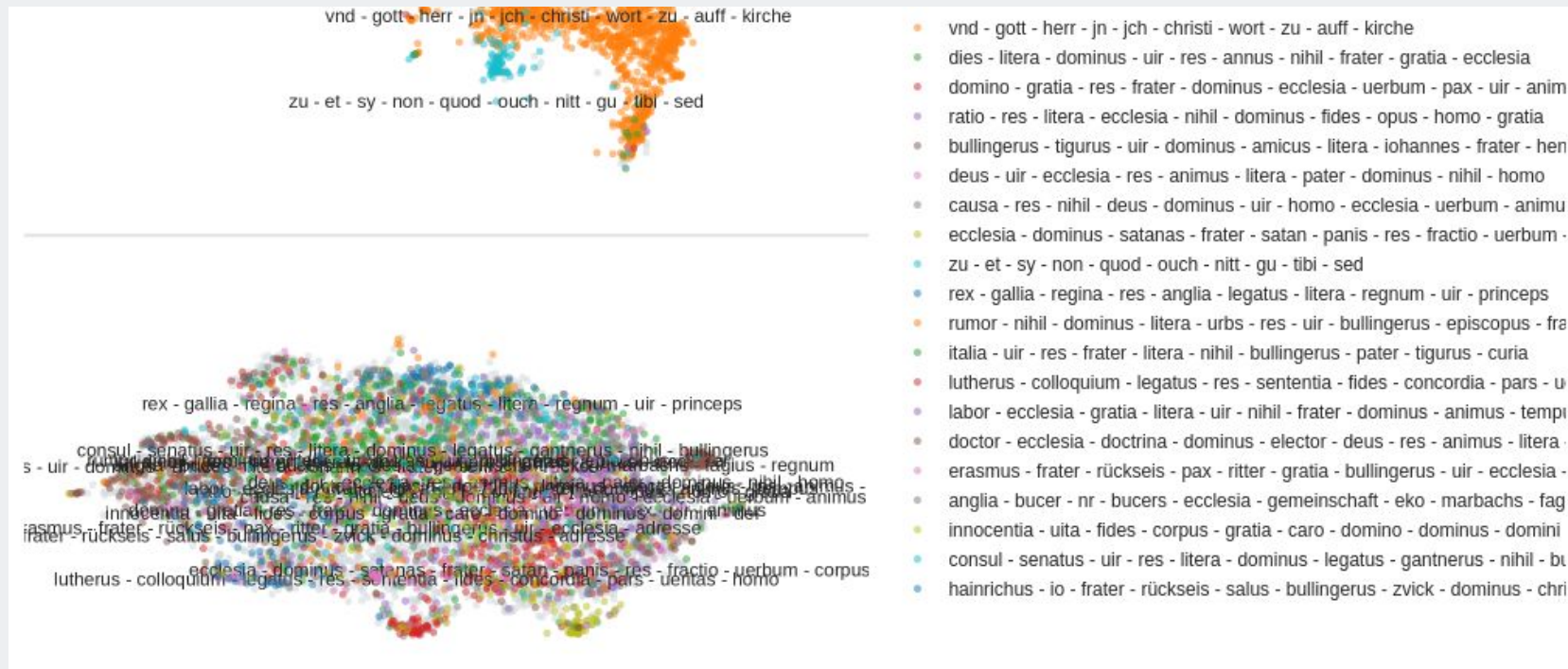
University of Zurich

{elina.stuessi, phillip.stroebel}@uzh.ch

To appear in this year's proceedings of
the Workshop on Computational
Linguistics for Cultural Heritage, Social
Sciences, Humanities and Literature

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/ 85.5	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	78.2 /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

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



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., ũ) 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,

[vobis gratias ago]

[Thank you]

