

Visualizing Language Models

Tobias Hodel, Sebastian Flicke, Reinhard Priber, Christa Schneider, Jonas Widmer
Digital Humanities, dh@wbk.kolleg.unibe.ch

Introduction

Language models are essential to succeed in NLP tasks such as Part-of-Speech-Tagging and Named Entity Recognition. Since the advent of *word2vec* and large transformer-based language models (such as *BERT* or *GPT-3*) a variety of specialized and fine-tuned language models is currently available. Nonetheless, our in-depth understanding of such models is limited at best.

In order to strengthen our knowledge of language models and to start the process of reflecting them, we are searching for creative ways of visualizing language models. We envision 3D-visualizations based on dimension reduction to identify the positioning of homonyms in vector spaces.

In our understanding, homonyms are words which share the same spelling but not the same meaning (homographs). Until today, this kind of homonyms is known to cause difficulties when processed in language models, as they are treated as one single element regardless to their semantic differences.

By using current frameworks [Akbik et al. 2019] which embed tokens (“words”) as parts of sentences, semantic and grammatical context plays an important role. Based on these contexts, homographs can for example be compared, when visualized even physically.

This leads us to new discussions and first very simple explanations of the inner functions of language models. As such models are currently implemented in tools (like search engines) and machines (like smartphones) that we use daily, we can expect to deal with other, even more complex questions.

Methods/Workflow

In order to make language models visually approachable we:

- Based on textual data from the Königsfelden corpus (for details see below)
- Trained a forward and backward character-based language model (vectorization of textual data) from scratch using *flair* NLP, not based on any pre-trained models [2096 dimensions]
- Vectorized 30 sentences containing homographs from the corpus
- Stored in a database for further use
- Reduced to three dimensions by applying PCA
- Visualized in a svelte frontend
- Deployed as a docker-image

Code

URL: https://github.com/DHBern/lm_viz

Dataset

URL: <https://koenigsfelden.uzh.ch>

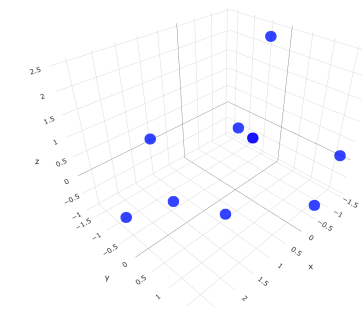
Charters and Records of Königsfelden Abbey (1308-1662). Prepared as a scholarly edition. Halter-Pernet et al. 2021: <https://koenigsfelden.uzh.ch>.

Results

URL: <https://nlp-hack-4.fdn-dev.iwi.unibe.ch>

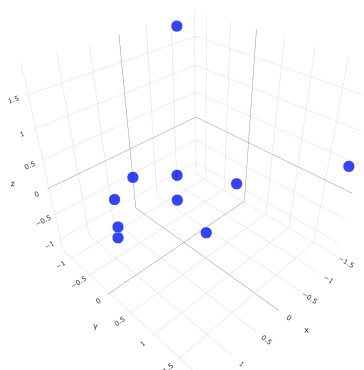
Label 1 “stat” (*city, place of execution, instead*)

Vor dem güt daz köft ward von den Mülner von Zurich diz ist als
stat0 von den selben güt untz an die sū « dn der nach stat » Hartmann
Müller « güt ill mallef hier en swin sol gelten » v « f xxx « eyler
Item Johannes Scherzer ill mallef
stat1 il fi Vo dez Mülliz güt von Zurich von dem güt daz hie nach stat «
ze Sarmentorf sol du eptisin Item ze Sarmentorf
« und daz dis war si und stet beilbe » so geben wir disen brief
stat2 unerr herrschaft versiget mit der stat insigel von Bruke « sid wir
nht eigen insigel haben
der vorgenanten lute bet « haben unser stat insigel gehenket an
stat3 disen brief « ze einem offren urkunde « der dingz die da vor
geschriben stant
ich des vorgenanden gütes des grunden conzen rechter were
stat4 vür eigen und vogtelge als da vorgeschriben stat an allen dien
stetten da si werschafft bedurfen von mir oder minen erben»
Chünrad den schultheiss « den rat und min buorger gemeinlich
stat5 von Arowe daz si ir stat insigel ze einer gezoßami die köfles
hant gehenket an disen brief Wir der Schultheiss
der rat und dū gemeinde vorgenande wan dir köf vor uns
stat6 beschach und mit urteile gewertigt wart nach nach unser stat
gewonheit und rechte mit allen dien gedigen als davor
geschriben
stat7 stat dur bette des vorgenen Ulrich unsers burgers han wir
enkein werschafft anders sunderlich an disen
stat dur bette des vorgenen Ulrich unsers burgers han wir
stat8 gehenket unser stat insigel ze einer gezoßami die köfles und
enkein werschafft anders sunderlich an disen
stat hat gewert und ez in minen nutz han bekeret daz ich
stat9 offrenlich vergihe an disen briefe « da von vergh doch ich « daz
ich



Label 2 “burg” (*castle, person & place name*)

Schenko « R » Vinsler « R » Sauerwiler « burger » ze Brugge «
burg0 Dietrich von Lenzburg burg » und andf erbt lute genüß « « «
Königsfelden
uf den gütern in dem Aygen ze Prukke « uf dem Pöz perg und
burg1 anderswo als die brief sagen die der vorgant unser cheim von
Nellem - burg « von
dazü verzigen si sich geistliches und welthes gerichtes
burg2 gewonheit der stete und des landes fries rechtes burg rechtes
und gemeinlich « aller der dinge da mite
Allen den die disen brief ansehet oder hörent lesen nu oder hie
burg3 nach « künden wir Johans Scherer schult heis ze der Nüwen
Regensperg » und die burger « die ze der selben burg
den gründen » in aller der welse « alz unser getruwe » brüder
burg4 Peter von Stoffeln « cömentur ze Tannenfels Hug von Güten
burg « Wernher Truchseß von Rinevelden und Heinrich von
Rinach « sundert und uscheidet » und under marksteinent
« und von Ulrichen und Albrechten iren sunen » ein holtz genant
burg5 « die Eichals » gelegen under unser burg Habsburg « dasselb
holtz und uns lehen ist « daz wir zü demselben kouffe « unsern
gunst und willen geben haben und
Königsfelden 333 wist umb j hus zü Lentzbürg « K H 8
burg6 Staatsarchiv AARGAU 366 « diz ist der grieff von dez huss
wegene Lenz burg
den burg graben ze Habsburg den Rudi Engelman buwet und
burg7 geschiedt in daz burglehen ze Habsburg und der egenannten
minir herrschaft von Österreich ledig « worden ist von Heintzlin
den lebenden « und geben auch in dem namen als davor
burg8 wissentlich mit disen briefe « für frye ledig eigen « den fränwalt
« und das holtz « gelegen under der burg ze Brunnegg « in Ergölw
« das man von aller nennet den Hag
und des ze urkund so henk ich der obgenanten lantvott min
eygen ingesigel an disen brief der geben ist ze Baden uff der
burg an sant Margareten tag « do man zalt von cristus gebürt
drühundert « eins und achtzig jar



Label 3 “schriben” (*written, writing*)

« daz dis allez daz hie vor ge- schriben ist « stette beilbe «
dar uber geben wir disen brief « versiget « mit unserm
schriben0 insigel « und mit dem insigel brüder Heinrich von Talhein der
ze den ziten
vorge- schriben güter « nütze » und « rehte ze beiden
schriben1 dorren « gar » und genzlich wider geben ze lösende « umb
« drissig » mark
geistlicher fröwen stat ze ire und ze ir nachkomen « händen
schriben2 dis nach « ge- schriben gelt « das lidig eigen ist und doch da
für verkört ist « vier mütt rogen « ein malter habern « ein
halb swin sol doch ein
wolt oder emhöht der sol und mach einen andern erben
schriben3 knecht in die vorgenante gelschaft schriben und legen der
an siner stat ze gleicher
und sinen brüder schriben mir jertzlich und inen ir jarzitt
schriben4 began jericlich und ewenlich disen obgeschribenen spruch
habend beid teil jetwedder für sich und die sinen gelobt und
verheissen by tröw an
ich darumb schriben das ich hoffen ir kemind an verziehen
schriben5 denn wir an ich als unseren oberen zu den sachen nitt
geantwurt künden noch zetzend wissen « meinten si nitt
getun künden sunder
schriben lassen das si und im teil der gült von juncker
schriben6 Albrecht von Rinach verweist stent nach inhalt disz briefs
etcf und die vi mütt kernengelts ist dahar « bisz uff
und hieruff zü merer sicherheit aller vor und nachgeschriben
schriben7 dinge « so von uns an disen brief ver- schriben stan « so
haben wir « den vlignanten unsern
ze Brugg die mich mitt Hansen Zölöf iren dickgenanten ir
schriben8 zu vogt geben hand sölich des Stapers angeben und schrift
vorgemeldeter in schriben lassen und anstat und für als
vor stat gelopt « sölich lichen und
« und her in schriben und setzen lassen » « Item des ersten
schriben9 ein gross acker ist der lüzpestreye widen hat vor Hansz
Geizler umb den landteil litt an dem kilchweg

