

ryan

April 12, 2022

```
[ ]: import semiolog as slg
```

```
[ ]: semiotic = slg.Cenematic("en_bnc_old_segments")
```

```
100%|          | 3/3 [00:00<00:00, 14.40it/s]
```

All model checkpoint layers were used when initializing TFBertForMaskedLM.

All the layers of TFBertForMaskedLM were initialized from the model checkpoint at models/en_bnc_old_segments/paradigms/tf_model.h5.

If your task is similar to the task the model of the checkpoint was trained on, you can already use TFBertForMaskedLM for predictions without further training.

```
[ ]: from random import randint
n = randint(0, semiotic.corpus.train.num_rows)
n = 577894
# n = 630149
sent = semiotic.corpus.train[n]["text"]
sent

sent_seq = semiotic(sent)
print(n)
sent_seq.tree.plot()
```

```
577894
```

```
interactive(children=(Output(),), _dom_classes=('widget-interact',))
```

```
[ ]: <function semiolog.syntagmatic.tree.Tree.plot.<locals>.inter()>
```

```
[ ]: sent_seq.parad_chain.df()
```

```
[ ]:
another_0 collective_1 title_2 the_3 angry_4 young_5 \
0 the serious victory of roleof young
1 inthe great problem for numbersof fisher
2 inits important question to two groupof
3 a obvious approach ofthe interestsof crafts
4 in new aspectof by supportof english
5 ina common factor forthe extenttowhich service
```

6	by	major	role	among	handsof	trades
7	forthe	second	senseof	tothe	proportionof	clergy
8	inhis	possible	majorityof	bythe	older	spokes
9	bythe	simple	attitude	ofthese	groupof	nessof
10	withthe	recent	attack	with	abilityof	black
11	his	significant	reason	against	deathof	rightsof
12	onthe	successful	advantage	in	younger	white
13	despite	famous	issue	the	numberof	french
14	with	final	point	between	majorityof	british
15	their	popular	example	amongthe	interestsofthe	police
16	underthe	striking	majority	from	performanceof	numberof
17	atthe	big	emphasison	which	absenceof	noble
18	despitethe	powerful	challenge	towards	treatmentof	work
19	ofthe	unusual	position	withthe	importanceof	women

	men_6	was_7	toprove_8	inthe_9	longterm_10	more_11	\
0	woman	are	reflected	tobe	handsof	more	
1	est	is	seen	theywere	least	as	
2	boy	was	now	themselves	process	less	
3	man	had	tobe	thatitwas	event	evenmore	
4	men	were	shown	the	times	most	
5	person	willhave	present	their	mind	tobe	
6	er	has	put	it	manner	in	
7	child	have	verymuch	theyare	best	muchmore	
8	girl	wouldhave	employed	thatheas	game	farmore	
9	mp	isnot	placed	them	press	themore	
10	victim	seem	based	thatthey	1970s	weremore	
11	candidate	tend	probably	as	hopethat	least	
12	student	mayhave	concentrated	thatitis	days	and	
13	king	appear	expressed	thattheyare	1980s	justas	
14	politician	havebeen	all	thatthe	extreme	is	
15	lady	havehad	growing	itwas	matter	amore	
16	leader	werenot	active	that	law	themost	
17	dog	wouldbe	perhaps	tohavebeen	processof	of	
18	reader	willbe	both	hewas	position	but	
19	s	arenot	forced	its	commons	was	

	accurate_12	as_13	predict_14	ion_15	than_16	description_17
0	difficult	self	perfect	ive	and	others
1	likely	mis	devot	ion	intheir	men
2	of	in	progress	ions	totheir	ever
3	powerful	physical	predict	iveness	inhis	anythingelse
4	effective	personal	confess	able	to	anyother
5	confident	dis	detect	ively	inthe	that
6	desirable	re	domin	ably	in	most
7	complex	social	exhaust	ability	or	women
8	seriously	sexual	access	ors	tothe	this

9	easily	senseof	insert	ing	their	theydid
10	radical	public	alb	ionof	fortheir	theother
11	accurately	political	complex	al	andhis	inthe
12	positive	their	cuss	or	ratherthan	usual
13	important	scientific	obsess	ional	by	one
14	clearly	formof	express	ory	than	them
15	widely	moral	select	ant	for	elsewhere
16	serious	meansof	gest	ivity	tohis	expected
17	strongly	intheir	accord	ers	forthe	otherwise
18	dangerous	levelof	conduct	ants	ist	thepolice
19	quickly	nessand	mans	ible	thanthe	those

```
[ ]: with open("/Users/Gianni/semiolog/models/en_bnc_old_segments/classes/kmeans_ids.
      ↪txt", "r") as k_ids:
        cluster_ids = [[int(id) for id in cluster.split()] for cluster in k_ids]

with open("/Users/Gianni/semiolog/models/en_bnc_old_segments/classes/
      ↪kmeans_corpus.txt", "r") as k_corpus:
        cluster_corpus = [[int(cl) for cl in cluster.split()] for cluster in
      ↪k_corpus]

with open("/Users/Gianni/semiolog/models/en_bnc_old_segments/classes/
      ↪kmeans_values.txt", "r") as k_values:
        cluster_values = [[float(val) for val in cluster.split()] for cluster in
      ↪k_values]
```

```
[ ]: cluster_labels = [[semiotic.vocab.decode[id] for id in cluster[:40]] for
      ↪cluster in cluster_ids]
slg.util.df(cluster_labels)
```

[]:	0	1	2	3	4	5	6	\
0	s	is	ing	inthe	and	the	john	
1	p	was	ed	the	of	their	mr	
2	the	are	edin	tothe	the	some	and	
3	t	were	edby	onthe	in	these	peter	
4	them	willbe	edto	fromthe	for	all	the	
5	f	hasbeen	edwith	forthe	to	its	david	
6	as	shouldbe	ingthe	ofthe	with	our	james	
7	d	canbe	edand	in	ofthe	other	robert	
8	l	wouldbe	ingin	withthe	inthe	his	michael	
9	g	maybe	edfor	atthe	from	new	paul	
10	b	hadbeen	edthe	intothe	on	allthe	he	
11	an	havebeen	edon	to	as	many	charles	
12	c	sare	s	and	by	those	richard	
13	at	mustbe	edas	andthe	tothe	two	martin	
14	am	couldbe	ingto	throughthe	onthe	both	george	
15	al	swere	edfrom	with	andthe	that	william	

16	h	mightbe	edat	bythe	forthe	any	stephen
17	m	being	ingand	for	fromthe	such	s
18	and	isnot	able	on	at	in	mary
19	thes	tobe	edinto	from	or	someofthe	harry
20	sh	wasnot	ers	of	s	more	simon
21	in	cannotbe	edout	overthe	withthe	a	as
22	theb	be	edbythe	by	a	this	tony
23	ac	and	edover	a	into	thenew	jack
24	w	arenot	and	at	that	your	mrs
25	ab	isnow	edthat	as	atthe	as	alan
26	thec	had	ingon	outofthe	sof	most	mark
27	thep	werenot	ingout	inhis	sand	my	mike
28	sp	isbeing	ingfor	underthe	but	and	bill
29	ap	isalso	es	that	through	for	in
30	ar	has	edup	ina	about	no	frank
31	thed	isstill	ings	duringthe	over	her	a
32	ad	as	ingup	aboutthe	like	with	philip
33	st	arealso	in	his	an	three	thomas
34	j	shouldnotbe	ingwith	into	bythe	anumberof	tom
35	ch	shavebeen	of	againstthe	their	various	miss
36	theh	became	ingof	this	after	to	this
37	k	remains	to	asthe	some	on	i
38	r	s	edthrough	an	its	good	she
39	n	areoften	edupon	their	ing	certain	that

	7	8	9	...	30	31	32	33 \
0	water	and	on	...	5	and	good	group
1	head	but	s	...	4	for	great	day
2	back	that	a	...	2	but	important	work
3	side	as	and	...	1	as	high	system
4	wall	when	er	...	3	in	simple	service
5	light	because	in	...	7	or	new	one
6	bed	if	an	...	6	with	serious	house
7	ground	so	ing	...	10	to	early	and
8	door	where	is	...	9	by	possible	team
9	house	although	le	...	8	from	different	company
10	glass	since	es	...	16	is	positive	man
11	table	before	as	...	15	of	special	time
12	hand	while	y	...	and	that	useful	course
13	surface	which	en	...	12	on	small	line
14	line	for	ley	...	11	not	poor	area
15	fire	though	ton	...	17	was	close	place
16	face	until	i	...	13	when	natural	club
17	way	sothat	e	...	18	so	true	book
18	place	after	o	...	20	the	complete	game
19	and	or	ers	...	p	only	practical	school
20	room	eventhough	to	...	25	even	difficult	er

21	body	sand	man	...	s	before	clear	s
22	air	once	the	...	30	at	effective	office
23	wind	then	el	...	14	tobe	obvious	party
24	window	andso	ed	...	in	about	common	business
25	s	unless	ie	...	to	now	large	record
26	mouth	yet	al	...	a	being	easy	side
27	land	inwhich	us	...	24	s	short	year
28	wood	evenif	ar	...	28	after	real	position
29	top	andthen	t	...	c	a	complex	show
30	floor	the	by	...	22	then	general	period
31	work	justas	am	...	21	just	low	programme
32	road	now	at	...	of	are	strong	field
33	car	in	ham	...	23	if	better	market
34	spot	to	son	...	26	all	appropriate	home
35	field	with	or	...	b	ratherthan	critical	stage
36	street	on	os	...	for	its	regular	issue
37	frame	all	ons	...	40	perhaps	essential	centre
38	roof	what	l	...	50	while	full	class
39	wheel	asif	land	...	19	without	specific	way

	34	35	36	37	38	39
0	to	seen	out	s	only	for
1	and	made	in	work	also	of
2	will	found	back	and	now	in
3	can	used	up	people	often	to
4	would	given	on	services	still	with
5	should	shown	again	problems	always	from
6	could	set	and	areas	quite	on
7	may	left	away	children	already	and
8	they	established	off	groups	all	by
9	or	done	down	activities	then	as
10	might	taken	there	members	both	at
11	that	known	it	companies	so	ofthe
12	donot	available	here	changes	usually	about
13	sto	carriedout	together	men	even	inthe
14	must	present	now	products	sometimes	against
15	then	lost	over	information	well	forthe
16	you	introduced	them	data	clearly	into
17	usedto	considered	for	workers	really	through
18	we	kept	to	conditions	actually	over
19	itwill	developed	well	events	more	or
20	didnot	held	that	times	most	tothe
21	dont	built	all	as	un	fromthe
22	but	placed	so	studies	probably	under
23	willnot	provided	about	staff	however	the
24	which	written	with	cases	very	between
25	youcan	there	only	issues	just	onthe

26	notto	arranged	from	others	obviously	after
27	also	brought	around	reports	too	including
28	now	drawn	through	that	rather	without
29	itwould	identified	as	records	generally	withthe
30	doesnot	published	right	units	therefore	suchas
31	the	treated	but	programmes	ofcourse	within
32	inorderto	prepared	more	students	almost	sof
33	i	put	before	sand	not	like
34	only	called	home	lines	never	towards
35	wecan	created	first	jobs	simply	that
36	haveto	chosen	one	schools	certainly	accordingto
37	theycan	based	too	research	tobe	but
38	cannot	important	by	numbers	nolonger	aswellas
39	itcould	reported	then	results	as	using

[40 rows x 40 columns]

```
[ ]: from collections import Counter

count_pairs = Counter()
count_triplets = Counter()
count_4tuples = Counter()

for sent in cluster_corpus:
    for pair in zip(sent, sent[1:]):
        count_pairs[pair] += 1

for sent in cluster_corpus:
    for pair in zip(sent, sent[1:], sent[2:]):
        count_triplets[pair] += 1

for sent in cluster_corpus:
    for tuple in zip(sent, sent[1:], sent[2:], sent[3:]):
        count_4tuples[tuple] += 1

top_pairs = [k for k, v in count_pairs.most_common(20)]
top_triplets = [k for k, v in count_triplets.most_common(20)]
top_4tuples = [k for k, v in count_4tuples.most_common(20)]
```

```
[ ]: slg.util.df([[cluster_labels[i][:5] for i in pair] for pair in top_pairs])
```

```
[ ]:
      0                                1  \
0  [to, and, will, can, would]  [s, p, the, t, them]
1  [see, get, find, take, have]  [in, ad, is, at, an]

      2                                3                                4  \
0  [mark, back, mix, present, view]  [5, 4, 2, 1, 3]  [s, in, p, and, f]
```

```

1      [ing, ed, edin, edby, edto]  [5, 4, 2, 1, 3]  [in, ad, is, at, an]

                                5  \
0  [new, local, public, the, international]
1      [s, work, and, people, services]

                                6              7  \
0      [the, a, this, that, their]  [and, a, in, all, one]
1  [work, action, position, time, order]  [ing, al, es, a, s]

                                8              9              10 \
0  [in, ad, is, at, an]  [s, in, p, and, f]  [s, in, p, and, f]
1  [on, s, a, and, er]  [ing, al, es, a, s]  [on, s, a, and, er]

                                11              12 \
0      [is, was, are, were, willbe]  [new, local, public, the, international]
1  [seen, made, found, used, given]  [group, day, work, system, service]

                                13 \
0  [work, action, position, time, order]
1      [for, of, in, to, with]

                                14 \
0  [new, local, public, the, international]
1      [work, action, position, time, order]

                                15              16 \
0  [good, great, important, high, simple]  [old, red, black, white, high]
1  [work, action, position, time, order]  [water, head, back, side, wall]

                                17              18 \
0  [the, a, this, that, their]  [the, a, this, that, their]
1      [and, a, in, all, one]  [group, day, work, system, service]

                                19
0  [new, local, public, the, international]
1  [new, local, public, the, international]

```

```
[ ]: slg.util.df([[cluster_labels[i][:5] for i in triplet] for triplet in
→top_triplets])
```

```

[ ]:
0      0              1              2  \
0  [5, 4, 2, 1, 3]  [s, p, the, t, them]  [to, and, will, can, would]
1  [5, 4, 2, 1, 3]  [in, ad, is, at, an]  [see, get, find, take, have]
2  [5, 4, 2, 1, 3]  [on, s, a, and, er]  [the, a, this, that, their]

                                3              4  \

```

0	[to, and, will, can, would]	[to, and, will, can, would]
1	[see, get, find, take, have]	[see, get, find, take, have]
2	[it, that, this, them, him]	[the, their, some, these, all]

	5	6 \
0	[is, was, are, were, willbe]	[to, and, will, can, would]
1	[mark, back, mix, present, view]	[only, also, now, often, still]
2	[ing, ed, edin, edby, edto]	[see, get, find, take, have]

	7	8 \
0	[s, in, p, and, f]	[the, a, this, that, their]
1	[in, ad, is, at, an]	[good, great, important, high, simple]
2	[on, s, a, and, er]	[work, action, position, time, order]

	9	10 \
0	[for, of, in, to, with]	[s, in, p, and, f]
1	[new, local, public, the, international]	[in, ad, is, at, an]
2	[s, work, and, people, services]	[ing, al, es, a, s]

	11	12 \
0	[s, work, and, people, services]	[s, p, the, t, them]
1	[to, and, will, can, would]	[in, ad, is, at, an]
2	[see, get, find, take, have]	[in, ad, is, at, an]

	13	14 \
0	[s, p, the, t, them]	[to, and, will, can, would]
1	[in, ad, is, at, an]	[see, get, find, take, have]
2	[ing, al, es, a, s]	[out, in, back, up, on]

	15	16 \
0	[new, local, public, the, international]	[s, work, and, people, services]
1	[new, local, public, the, international]	[is, was, are, were, willbe]
2	[s, work, and, people, services]	[seen, made, found, used, given]

	17	18 \
0	[and, that, which, who, but]	[the, a, this, that, their]
1	[mark, back, mix, present, view]	[work, action, position, time, order]
2	[ing, ed, edin, edby, edto]	[for, of, in, to, with]

	19
0	[only, also, now, often, still]
1	[mark, back, mix, present, view]
2	[ing, ed, edin, edby, edto]

```
[ ]: slg.util.df([[cluster_labels[i][:5] for i in tuple] for tuple in top_4tuples])
```



```

[ ]:          0                      1  \
0 [5, 4, 2, 1, 3]          [to, and, will, can, would]
1 [5, 4, 2, 1, 3]          [see, get, find, take, have]
2 [5, 4, 2, 1, 3]          [the, a, this, that, their]
3 [5, 4, 2, 1, 3] [work, action, position, time, order]

                                2                      3  \
0          [to, and, will, can, would] [s, work, and, people, services]
1          [see, get, find, take, have] [is, was, are, were, willbe]
2          [the, their, some, these, all] [mark, back, mix, present, view]
3 [work, action, position, time, order] [ing, ed, edin, edby, edto]

                                4                      5  \
0 [john, mr, and, peter, the]          [s, work, and, people, services]
1          [s, p, the, t, them]          [for, of, in, to, with]
2          [in, ad, is, at, an] [new, local, public, the, international]
3          [on, s, a, and, er]          [s, work, and, people, services]

                                6                      7  \
0 [s, p, the, t, them] [new, local, public, the, international]
1 [in, ad, is, at, an]          [s, work, and, people, services]
2 [on, s, a, and, er]          [to, and, will, can, would]
3 [on, s, a, and, er]          [see, get, find, take, have]

                                8                      9  \
0          [the, a, this, that, their] [s, p, the, t, them]
1 [good, great, important, high, simple] [in, ad, is, at, an]
2 [work, action, position, time, order] [in, ad, is, at, an]
3          [for, of, in, to, with] [in, ad, is, at, an]

                                10                     11  \
0 [new, local, public, the, international] [and, a, in, all, one]
1          [s, work, and, people, services] [5, 4, 2, 1, 3]
2          [is, was, are, were, willbe] [5, 4, 2, 1, 3]
3          [seen, made, found, used, given] [5, 4, 2, 1, 3]

                                12                     13  \
0 [work, action, position, time, order] [is, was, are, were, willbe]
1          [for, of, in, to, with] [only, also, now, often, still]
2 [new, local, public, the, international] [mark, back, mix, present, view]
3          [s, work, and, people, services] [ing, ed, edin, edby, edto]

                                14                     15  \
0 [s, work, and, people, services] [john, mr, and, peter, the]
1          [is, was, are, were, willbe] [s, in, p, and, f]
2 [seen, made, found, used, given] [in, ad, is, at, an]
3          [for, of, in, to, with] [on, s, a, and, er]

```

	16	17 \
0	[the, a, this, that, their]	[the, their, some, these, all]
1	[work, action, position, time, order]	[s, work, and, people, services]
2	[for, of, in, to, with]	[to, and, will, can, would]
3	[new, local, public, the, international]	[see, get, find, take, have]

	18	19
0	[to, and, will, can, would]	[on, s, a, and, er]
1	[see, get, find, take, have]	[5, 4, 2, 1, 3]
2	[the, their, some, these, all]	[5, 4, 2, 1, 3]
3	[s, work, and, people, services]	[5, 4, 2, 1, 3]

```
[ ]: slg.util.df(["|".join(slg.util.flatten([random.
    ↳sample(cluster_labels[l],1),random.sample(cluster_labels[c],1),random.
    ↳sample(cluster_labels[r],1)])) for _ in range(10)] for l,c,r in_
    ↳top_triplets[:20]))
```

	0	1	2	3 \
0	16 26 16	h at t	we meet thebest	itwould turn the
1	5 26 p	l il in	should show our	will turn life
2	c 50 c	l un al	might make our	and run some
3	and b 14	ab ip le	haveto play onthe	also start all
4	in 22 10	and ad the	would help some	they turn anything
5	8 4 for	ab i ons	then change tothe	haveto give here
6	25 8 20	and ay ed	will go agood	must enter him
7	19 4 50	ac ed an	only hold that	itwould meet me
8	c 9 of	n ed on	willnot leave asmall	you feel now
9	3 11 3	thed ath son	should play another	would meet us

	4	5	6 \
0	didnot see other	maybe record ingand	could just use
1	dont start his	canbe review s	then however provide
2	to be in	are return edat	doesnot therefore go
3	now avoid these	isbeing question edup	cannot now avoid
4	and meet the	isbeing form edup	notto un start
5	willnot change her	willbe help ingfor	sto therefore run
6	must show his	shouldbe mark edand	willnot generally avoid
7	willnot change some	willbe mix ingout	only always turn
8	itwould help most	is test ingof	then quite see
9	itwould bring three	arenot post edand	didnot obviously be

	7	8	9	10 \
0	at i l an	thenext real life	aswellas special s	an as or
1	l s us	thewhole low question	on technical issues	c er als
2	c i al	any different role	sof financial changes	p ath ed
3	or ang os	its regular change	but private as	mar s i

4	g all son	our positive authority	about existing that	ad il ation
5	in er or	in the simple condition	that their jobs	re ang ate
6	per as os	the last special order	that commercial that	in un al
7	and al land	the last simple purpose	about economic data	s un or
8	d es i	an different work	on and staff	t ar ates
9	ch at ie	the best early policy	without and data	sc un a

	11	12	13	14 \
0	s migh run	at ing at	ad is or	may show right
1	students only produce	the p us al	gl ip sand	it could hold through
2	activities now understand	g al ac	ad am s	we make one
3	problems now recognise	b un er	l it er	does not add on
4	reports it will hold	p et an	b un in	might turn out
5	members would turn	ap am on	sh an an	it would take them
6	as now turn	the i i il	ab ir ates	also hold more
7	as it will see	them i d ag	g al o	it could feel first
8	others you can stand	n all in	ar am s	only help through
9	services it will avoid	f i i i n	r af ating	dont see right

	15	16 \
0	industrial commercial people	jobs will be considered
1	home in staff	schools cannot be made
2	business public problems	research were not held
3	home all jobs	activities be seen
4	new the cases	people is set
5	local local others	products swere published
6	national free activities	workers is also identified
7	all financial as	information is still written
8	public european records	children and built
9	new local work	sand are taken

	17	18	19
0	then post to	their position in	only tim to
1	only search ed from	an authority on the	quite present ing to
2	having pack ing	any process that	well mark ing to
3	which fill ed over	each view over	never pick in
4	as point ed that	the first performance and	sometimes point ed as
5	was touch ed at	a small situation without	sometimes avoid to
6	but the help ing out	a situation through	always load ing on
7	when they lock ed on	another attention about	probably plac ed to
8	that question ed in	that question from	generally serv es
9	was deliver ing	as order using	even load able

```
[ ]: slg.util.df(["|".join(slg.util.flatten([random.
    ↳sample(cluster_labels[l],1),random.sample(cluster_labels[c],1),random.
    ↳sample(cluster_labels[d],1),random.sample(cluster_labels[r],1)))] for _ in_
    ↳range(10)] for l,c,d,r in top_4tuples[:20]])
```

[]:	0	1	2 \
0	10 to and 9	dont be thenext purpose	inorderto feel many situation
1	17 6 14 2	sto understand afull interest	i put such attention
2	in 16 19 25	and play each value	willnot feel in change
3	1 10 8 b	we see afull work	itwould stand his activity
4	25 2 of 2	they enjoy your record	but give our problem
5	40 14 17 15	notto take alarge policy	wecan run for role
6	16 16 16 a	will use thebest life	to go other strength
7	c 26 s c	cannot make forthe place	itwill hold in behaviour
8	for 11 8 24	that make that authority	doesnot meet any time
9	22 of 18 19	would avoid no purpose	dont start good life

	3	4 \
0	changes s connect ingin	tony an ay ar
1	information werenot pick edbythe	robert ap an is
2	programmes arealso mark able	martin ch ath ed
3	jobs tobe point edup	charles k as el
4	results s load ers	paul ap and os
5	programmes shouldnotbe question edthat	tom s un el
6	events werenot load s	tom m es an
7	times arenot head ingup	paul am af es
8	data being start ed	miss d al l
9	conditions remains sign edover	jack sh ing ons

	5	6 \
0	records without their as	ap s as e
1	that of personal work	ad on am ie
2	as onthe independent changes	n a el ley
3	data within political sand	theh ab by ie
4	activities at social s	thed ag l an
5	activities without educational sand	c el on at
6	units about high men	w is ing s
7	sand with technical problems	am al ham as
8	activities towards home problems	sp and ton le
9	records tothe international problems	them an es o

	7	8 \
0	regional research inorderto enter	an appropriate condition or
1	local activities sto follow	thewhole appropriate exercise ofthe
2	private services would lose	her important situation as
3	its changes willnot keep	thenext better concern after
4	economic schools or lose	their general statement inthe
5	its times didnot start	afull effective condition or
6	public issues they make	thenext better order onthe
7	independent events you turn	afull useful attention towards
8	public times might avoid	this useful power at
9	general events youcan enter	any small place towards

	9	10	11 \
0	thes el es or	all jobs havebeen established	all 13 a 11
1	d ing ac i	current data were left	may 22 12 13
2	g ath ang er	special records hadbeen there	as 8 6 23
3	thed on er is	private programmes mightbe provided	1990 c 40 in
4	at en es a	existing issues swere done	work 13 c 24
5	al am s es	in workers canbe important	later 8 24 11
6	ap al ot a	national research s written	work 28 7 30
7	n od ang an	social cases were found	later 3 7 21
8	l il ag ir	special events had made	in 3 3 p
9	at ot es od	individual problems wouldbe present	later 15 17 c

	12	13 \
0	development from high programmes	wouldbe still load es
1	position without other schools	isalso really sign edthat
2	value within all men	willbe well plac and
3	policy in business events	arenot however view ingup
4	performance from new children	became most load ing
5	attention aswellas british and	has sometimes deliver ingon
6	form on regional products	havebeen often not edby
7	attention that social results	isnot simply cross edas
8	purpose towards state groups	isbeing therefore review edto
9	strength about new records	couldbe clearly connect edwith

	14	15 \
0	services isnow brought but	james d ay al
1	units s brought accordingto	as al ip l
2	conditions arealso considered of	william an it er
3	reports wasnot published about	as le on us
4	services tobe prepared for	thomas ad ar land
5	people arenot considered with	david to am el
6	programmes couldbe found tothe	mary at ath es
7	men isalso there at	bill t al al
8	products wouldbe placed withthe	tony b ip land
9	changes isnow developed after	and n us ham

	16	17 \
0	asecond response to its	on lines we stand
1	this effect as legal	most groups i change
2	asingle time the legal	no men should do
3	asmall form after in	this records we put
4	thenew strength as european	three men might keep
5	alarge action within british	most men the turn
6	thenext effect within educational	that programmes you understand
7	asecond effect suchas technical	its numbers itwill put
8	asmall process tothe existing	with work youcan leave

```
9         our|activity|using|british    allthe|records|youcan|produce
```

```

                                18          19
0    now|learn|thenew|others    in|20|in|50
1        they|change|the|as    am|13|c|19
2        can|follow|as|lines    on|s|18|10
3    you|change|someofthe|that    on|15|10|and
4        also|add|with|sand    on|4|6|40
5    will|put|more|students    ing|40|3|14
6        can|lose|to|issues    and|for|for|13
7    or|hold|thenew|activities    us|in|50|2
8        might|hold|most|s    i|b|4|3
9        will|be|the|data    on|16|24|17

```

```
[ ]: import random
import plotly.io as pio
pio.renderers.default = "vscode" # "browser"#
```

```
[ ]: top_label = [c[0] for c in cluster_labels]
type_id2comptype = {id: comptype for id,comptype in enumerate(cluster_corpus)}
type_comptype2labels = sorted([(comptype,labels[:10]) for comptype,labels in
↪zip(cluster_corpus,cluster_labels)])
```

```
[ ]: for n in range(3):
    l,r=top_pairs[n]
    pair_cross = [l_i+"|"+r_i for l_i in cluster_labels[l] for r_i in
↪cluster_labels[r]]
    max_parad = 1000 # max_non_zeroes
    plot = slg.util.plot_scatter_line(
        list(range(max_parad)),
        cluster_values[l],
        top_label[l],
        add_trace=[(list(range(max_parad)), cluster_values[r],
↪top_label[r])],title=f"Rank of pair= {n}<br>Pair of classes:
↪{(l,r)}<br>Sample of pairs of terms: {random.sample(pair_cross,10)}")
    plot.show()
```

```
[ ]:
```