System_alignement3_example

October 25, 2018

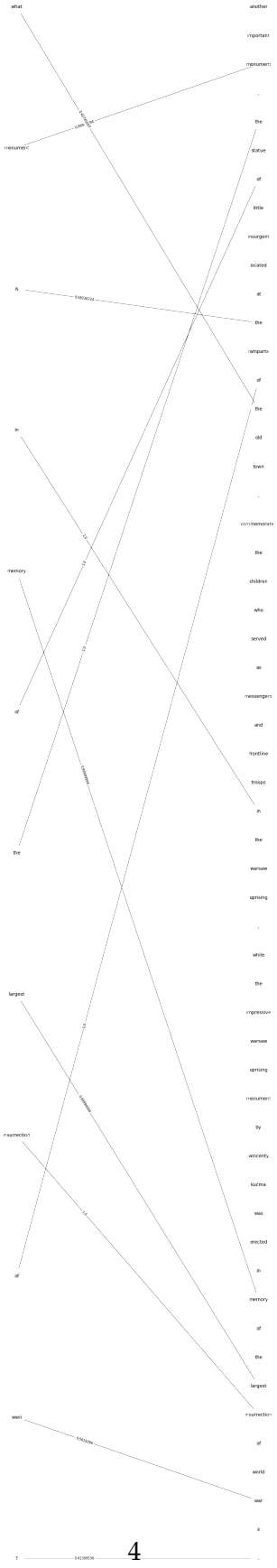
In [13]:
```
import networkx as nx
from networkx import *
import matplotlib.pyplot as plt
from networkx.algorithms import bipartite
from nltk import sent_tokenize,word_tokenize
import numpy as np
import pandas as pd

import sys
sys.path.append('../..')
sys.path.append('../../utils/')
from utils import *
import operator


import fastText
model = fastText.load_model('../../../Divers_Data_Maitrise/wiki.simple/wiki.simple.bin'

#ignorer toutes les depreciations de fonctions
import warnings
warnings.simplefilter('ignore')
```

In [14]:
```
def Alignement_graph_bipartite(question,sequence, list_alignements):
    '''
    Fonction qui crée un graphe bipartit (question, sentence) et qui affecte la similar
    cosine comme poids de chaque arete. L'arete en trait plein represente la plus forte
    les autres similarite seront affichées avec des traits discontinus.
    :list_alignements: les alignements reliser par l'algo
    '''

    G=nx.Graph()

    list_words_question = word_tokenize(question.lower())
    list_words_sequence = word_tokenize(sequence.lower())


    nb_words_question = len(list_words_question)
```

```
        nb_words_sequence = len(list_words_sequence)
        height_colum = max(nb_words_question,nb_words_sequence)
        y_linspace_question = np.linspace(0,height_colum,nb_words_question)
        y_linspace_sentence = np.linspace(0,height_colum,nb_words_sequence)

        fixed_positions={}
        labels={}


        list_edges = []
        edge_labels = {}
        for i in range(0,nb_words_sequence): # mots de la phrase
            num_node = nb_words_question + i
            G.add_node(num_node)
            labels[num_node]=list_words_sequence[i]

            fixed_positions[num_node] = (2,y_linspace_sentence[nb_words_sequence-i-1])

        for i in range(0,nb_words_question): # mots de la question
            G.add_node(i)
            labels[i]=list_words_question[i]
            fixed_positions[i] = (0,y_linspace_question[nb_words_question-i-1])

        for elem in list_alignements:
            G.add_edge(elem[0][0],elem[0][1] + nb_words_question,sim=elem[1])
            list_edges.append((elem[0][0],elem[0][1] + nb_words_question))
            edge_labels[(elem[0][0],elem[0][1] + nb_words_question)]=elem[1]

        pos = nx.spring_layout(G,pos=fixed_positions, fixed=fixed_positions.keys())
        plt.figure(3,figsize=(15,height_colum*1.75))
        nx.draw_networkx_nodes(G,pos,node_color='w',node_size=2500)
        nx.draw_networkx_labels(G,pos,labels,font_size=14)
        nx.draw_networkx_edge_labels(G,pos,node_color='b',edge_labels=edge_labels, label_po
        nx.draw_networkx_edges(G,pos,node_color='b', edgelist =list_edges,  width = 1.0)

        plt.axis('off')
    #     plt.savefig("Graph.png", format="PNG") # pour enregistrer l'image dans un fichier
        plt.show()

In [15]: def affiche_table_cosine(question,sequence):
        list_words_question = word_tokenize(question.lower())
        list_words_sequence = word_tokenize(sequence.lower())

        dict_table = {}

        for word_question in list_words_question:
            dict_table[word_question] = []
            vect_word_question = model.get_word_vector(word_question)
```

```
            for word_sentence in list_words_sequence:
                vect_word_sentence = model.get_word_vector(word_sentence)
                sim = cosine_similarity(vect_word_sentence, vect_word_question)
                dict_table[word_question].append(sim)
        df = pd.DataFrame(dict_table, index=list_words_sequence)
        print(df)
```

```
In [16]: Alignement_graph_bipartite( "What monument is in memory of the largest insurrection of
         affiche_table_cosine("What monument is in memory of the largest insurrection of WWII?",
         question=" What monument is in memory of the largest insurrection of WWII? "
         sequence=" Another important monument, the statue of Little Insurgent located at the ra
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: , while the impressive warsaw uprising monument by wincenty kuma wa
         print("reponses attendu: [{'answer_start': 606, 'text': 'Warsaw Uprising Monument'}, {'
```


what
monument
is
in
memory
of
the
largest
insurrection
of
world
war
ii
.
another
important
monument
.
the
statue
of
little
insurgent
located
at
the
ramparts
of
the
old
town
.
commemorate
the
children
who
served
as
messengers
and
frontline
troops
in
the
warsaw
uprising
.
while
the
impressive
warsaw
uprising
monument
by
wincenty
kucma
was
erected
in
memory
of
the
largest
insurrection
of
world
war
ii
.

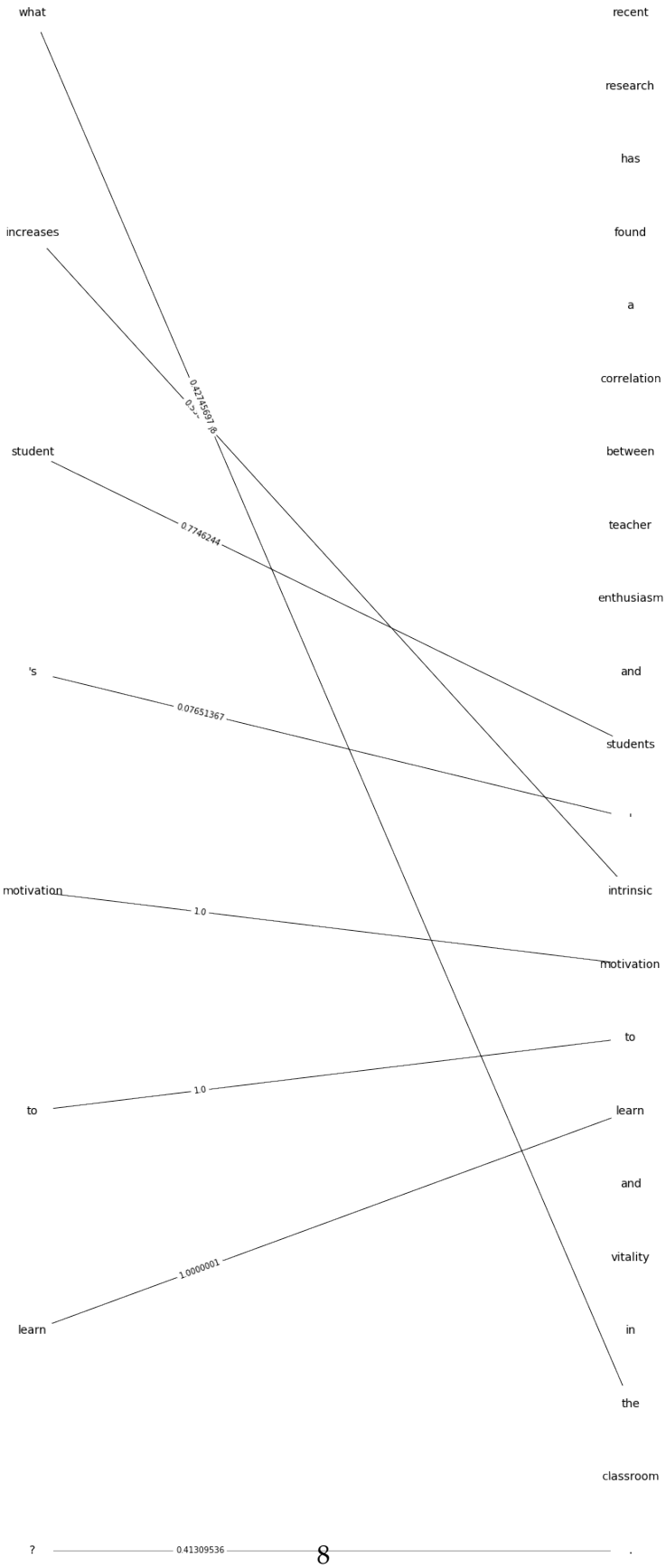
	what	monument	is	in	memory	of	\
another	0.419007	0.179076	0.412785	0.339769	0.190673	0.381471	
important	0.311649	0.268806	0.350485	0.379637	0.177604	0.433017	
monument	0.166299	1.000000	0.218277	0.245872	0.273939	0.235299	
,	0.308747	0.203744	0.469191	0.520147	0.149771	0.494608	
the	0.427457	0.233662	0.582367	0.666239	0.178509	0.733189	
statue	0.075944	0.505756	0.200333	0.177604	0.206246	0.228740	
of	0.332908	0.235299	0.485432	0.569639	0.179194	1.000000	
little	0.261418	0.139723	0.234369	0.225153	0.138461	0.201166	
insurgent	0.193829	0.144971	0.152074	0.210682	0.119952	0.213037	
located	0.203462	0.282696	0.357172	0.351498	0.130388	0.325209	
at	0.283861	0.190466	0.259063	0.437498	0.109882	0.425586	
the	0.427457	0.233662	0.582367	0.666239	0.178509	0.733189	
ramparts	0.236284	0.299419	0.233384	0.206664	0.134898	0.296171	
of	0.332908	0.235299	0.485432	0.569639	0.179194	1.000000	
the	0.427457	0.233662	0.582367	0.666239	0.178509	0.733189	
old	0.221178	0.202633	0.243326	0.286532	0.149489	0.305222	
town	0.197132	0.268992	0.300978	0.349755	0.123196	0.265857	
,	0.308747	0.203744	0.469191	0.520147	0.149771	0.494608	
commemorates	0.083365	0.433951	0.248240	0.341372	0.332882	0.305558	
the	0.427457	0.233662	0.582367	0.666239	0.178509	0.733189	
children	0.224695	0.198997	0.154957	0.206516	0.167174	0.197835	
who	0.409430	0.156056	0.242936	0.279211	0.177163	0.282786	
served	0.127113	0.168054	0.217768	0.242995	0.101051	0.296706	
as	0.317494	0.165687	0.432165	0.388784	0.128271	0.388661	
messengers	0.225767	0.060575	0.160498	0.191910	0.175126	0.209847	
and	0.313092	0.184708	0.385975	0.495406	0.202615	0.511697	
frontline	0.293165	0.224873	0.190570	0.239523	0.180779	0.266044	
troops	0.183117	0.233087	0.150640	0.258267	0.192057	0.236023	
in	0.308086	0.245872	0.539444	1.000000	0.203926	0.569639	
the	0.427457	0.233662	0.582367	0.666239	0.178509	0.733189	
warsaw	0.103738	0.222659	0.097412	0.226426	0.130419	0.192538	
uprising	0.190160	0.229108	0.161404	0.285772	0.062925	0.270334	
,	0.308747	0.203744	0.469191	0.520147	0.149771	0.494608	
while	0.355118	0.141717	0.271320	0.339928	0.163344	0.336693	
the	0.427457	0.233662	0.582367	0.666239	0.178509	0.733189	
impressive	0.290462	0.223749	0.163730	0.174277	0.134389	0.234688	
warsaw	0.103738	0.222659	0.097412	0.226426	0.130419	0.192538	
uprising	0.190160	0.229108	0.161404	0.285772	0.062925	0.270334	
monument	0.166299	1.000000	0.218277	0.245872	0.273939	0.235299	
by	0.245548	0.155371	0.389985	0.370934	0.179134	0.364681	
wincenty	0.259466	0.212276	0.266364	0.296929	0.153724	0.331438	
kuma	0.205720	0.110512	0.239649	0.192620	0.174648	0.173515	
was	0.221110	0.239102	0.581030	0.455141	0.164985	0.406907	
erected	0.123417	0.395306	0.114836	0.245508	0.266528	0.268452	

in	0.308086	0.245872	0.539444	1.000000	0.203926	0.569639
memory	0.206907	0.273939	0.218819	0.203926	1.000000	0.179194
of	0.332908	0.235299	0.485432	0.569639	0.179194	1.000000
the	0.427457	0.233662	0.582367	0.666239	0.178509	0.733189
largest	0.134632	0.221110	0.341342	0.301722	0.093422	0.366508
insurrection	0.235026	0.240116	0.166871	0.287283	0.231391	0.268343
of	0.332908	0.235299	0.485432	0.569639	0.179194	1.000000
world	0.216696	0.189339	0.244585	0.303820	0.175365	0.282893
war	0.225160	0.229449	0.185026	0.299041	0.191067	0.316802
ii	0.138667	0.222069	0.183010	0.214534	0.089356	0.284481
.	0.356790	0.215858	0.456966	0.559157	0.130566	0.535246

	the	largest	insurrection	wwii	?
another	0.501971	0.209383	0.168671	0.170949	0.295512
important	0.393481	0.318887	0.114808	0.189867	0.178125
monument	0.233662	0.221110	0.240116	0.211594	0.114681
,	0.493308	0.266263	0.186096	0.193261	0.310912
the	1.000000	0.375704	0.291021	0.263342	0.220681
statue	0.217462	0.214881	0.222554	0.130270	0.086727
of	0.733189	0.366508	0.268343	0.185755	0.182614
little	0.241261	0.099894	0.118569	0.173069	0.203365
insurgent	0.248968	0.177132	0.549381	0.402615	0.125134
located	0.407527	0.375091	0.118321	0.146873	0.089832
at	0.457213	0.170447	0.168162	0.126393	0.219786
the	1.000000	0.375704	0.291021	0.263342	0.220681
ramparts	0.342158	0.214311	0.254742	0.169403	0.105767
of	0.733189	0.366508	0.268343	0.185755	0.182614
the	1.000000	0.375704	0.291021	0.263342	0.220681
old	0.318498	0.100633	0.158698	0.210111	0.205771
town	0.297424	0.216125	0.186217	0.170582	0.099023
,	0.493308	0.266263	0.186096	0.193261	0.310912
commemorates	0.337768	0.253056	0.341096	0.253575	0.105349
the	1.000000	0.375704	0.291021	0.263342	0.220681
children	0.156541	0.134924	0.161479	0.172577	0.217882
who	0.373393	0.124223	0.213869	0.178442	0.262204
served	0.252093	0.202500	0.213691	0.162675	0.117141
as	0.483816	0.181992	0.163164	0.196150	0.270439
messengers	0.283224	0.113190	0.166541	0.074306	0.152077
and	0.517126	0.246027	0.176561	0.205932	0.205442
frontline	0.319491	0.135694	0.328129	0.413529	0.257962
troops	0.322996	0.155311	0.418041	0.346381	0.143298
in	0.666239	0.301722	0.287283	0.238832	0.204310
the	1.000000	0.375704	0.291021	0.263342	0.220681
warsaw	0.231372	0.151686	0.241528	0.326246	0.097699
uprising	0.285359	0.188552	0.486030	0.324267	0.136450
,	0.493308	0.266263	0.186096	0.193261	0.310912
while	0.398046	0.116322	0.193028	0.198578	0.210764
the	1.000000	0.375704	0.291021	0.263342	0.220681

```
impressive    0.233690  0.208846      0.112808  0.174877  0.204956
warsaw        0.231372  0.151686      0.241528  0.326246  0.097699
uprising      0.285359  0.188552      0.486030  0.324267  0.136450
monument      0.233662  0.221110      0.240116  0.211594  0.114681
by            0.440875  0.178394      0.180398  0.211429  0.179666
wincenty      0.309551  0.204803      0.246528  0.249482  0.273629
kuma         0.208374  0.168275      0.174726  0.076984  0.204851
was           0.487115  0.219871      0.302655  0.276011  0.184121
erected       0.236633  0.228712      0.342177  0.225445  0.080637
in            0.666239  0.301722      0.287283  0.238832  0.204310
memory        0.178509  0.093422      0.231391  0.121183  0.139768
of            0.733189  0.366508      0.268343  0.185755  0.182614
the           1.000000  0.375704      0.291021  0.263342  0.220681
largest       0.375704  1.000000      0.163594  0.188383  0.124012
insurrection  0.291021  0.163594      1.000000  0.384608  0.092945
of            0.733189  0.366508      0.268343  0.185755  0.182614
world         0.336043  0.344524      0.146594  0.321969  0.175796
war           0.366200  0.226804      0.439538  0.561510  0.155105
ii            0.263146  0.130036      0.246958  0.422187  0.119282
.             0.593451  0.188717      0.187174  0.230246  0.413095
question =   What monument is in memory of the largest insurrection of WWII?
sequence =   Another important monument, the statue of Little Insurgent located at the ramparts
span output: , while the impressive warsaw uprising monument by wincenty kuma was erected in
reponses attendu: [{'answer_start': 606, 'text': 'Warsaw Uprising Monument'}, {'answer_start': 6
```

In [17]:
```
Alignement_graph_bipartite( "What increases student's motivation to learn?","Recent res
affiche_table_cosine("What increases student's motivation to learn?","Recent research h
question=" What increases student's motivation to learn? "
sequence=" Recent research has found a correlation between teacher enthusiasm and stude
print("question = ", question)
print("sequence = ",sequence)
print("span output: research has found a correlation between teacher enthusiasm and ")
print("reponses attendu: [{'answer_start': 207, 'text': 'teacher enthusiasm'}, {'answer
```

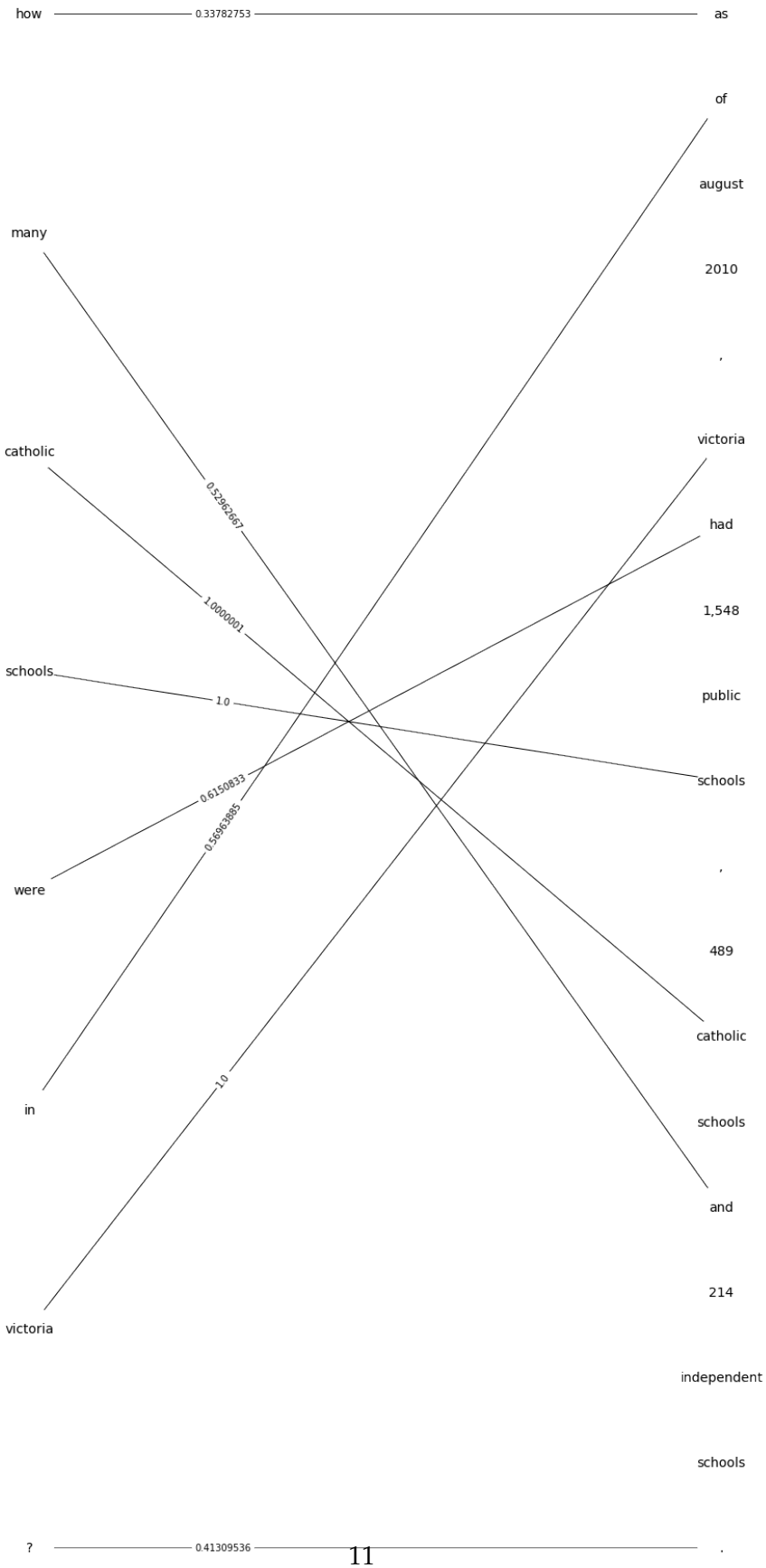

what
increases
student
's
motivation
to
learn
?
recent
research
has
found
a
correlation
between
teacher
enthusiasm
and
students
'
intrinsic
motivation
to
learn
and
vitality
in
the
classroom
.
0.42745697
0.7746244
0.07651367
1.0
1.0
1.0000001
0.41309536

	what	increases	student	's	motivation	to	\
recent	0.242562	0.264873	0.145005	-0.017463	0.197147	0.312057	
research	0.237140	0.179860	0.303215	-0.039236	0.241872	0.266731	
has	0.203381	0.326980	0.212762	-0.040684	0.128659	0.361999	
found	0.179633	0.189711	0.101555	0.005867	0.176985	0.252256	
a	0.332040	0.311974	0.320711	0.048670	0.234399	0.455036	
correlation	0.310826	0.439372	0.210162	-0.057671	0.424357	0.289198	
between	0.246216	0.314026	0.093434	0.006096	0.163029	0.363919	
teacher	0.204926	0.086776	0.575409	0.026858	0.319494	0.199134	
enthusiasm	0.246522	0.206068	0.158406	0.074785	0.372018	0.213332	
and	0.313092	0.280004	0.239157	0.068952	0.271786	0.489713	
students	0.238073	0.187418	0.774624	0.036191	0.246164	0.290582	
'	0.324581	0.117177	0.193568	0.076514	0.171786	0.345445	
intrinsic	0.386943	0.530421	0.178293	-0.107384	0.381959	0.279059	
motivation	0.363206	0.336243	0.283811	0.013235	1.000000	0.338755	
to	0.455848	0.319767	0.212202	0.037827	0.338755	1.000000	
learn	0.362374	0.223444	0.386204	0.034417	0.380363	0.348724	
and	0.313092	0.280004	0.239157	0.068952	0.271786	0.489713	
vitality	0.197406	0.282308	0.155492	0.027397	0.343004	0.229239	
in	0.308086	0.230205	0.254856	-0.037581	0.213828	0.450107	
the	0.427457	0.381410	0.212593	-0.015404	0.232487	0.535392	
classroom	0.212470	0.181987	0.435544	-0.043787	0.324054	0.264006	
.	0.356790	0.229873	0.209948	-0.003708	0.248310	0.464503	

	learn	?
recent	0.175797	0.254489
research	0.231157	0.134820
has	0.203600	0.159945
found	0.139591	0.167903
a	0.211142	0.239167
correlation	0.215278	0.234000
between	0.143937	0.128871
teacher	0.422318	0.179782
enthusiasm	0.260255	0.202661
and	0.266888	0.205442
students	0.449325	0.180173
'	0.182885	0.330942
intrinsic	0.181417	0.230021
motivation	0.380363	0.275321
to	0.348724	0.294633
learn	1.000000	0.311875
and	0.266888	0.205442
vitality	0.192169	0.190803
in	0.185811	0.204310
the	0.226916	0.220681

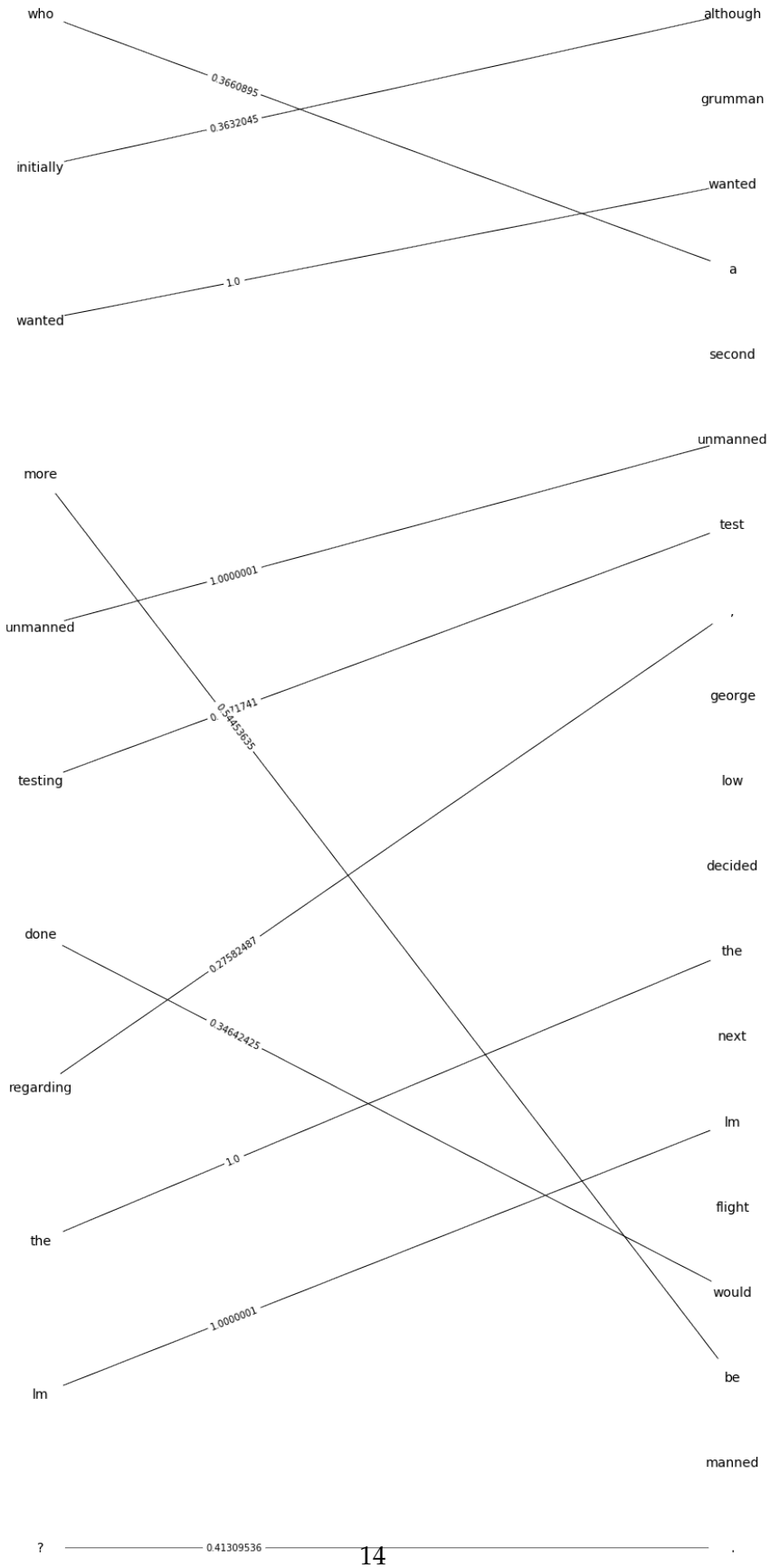
```
classroom    0.353849  0.231904
.            0.242660  0.413095
question =   What increases student's motivation to learn?
sequence =   Recent research has found a correlation between teacher enthusiasm and students' in
span output: research has found a correlation between teacher enthusiasm and
reponses attendu: [{'answer_start': 207, 'text': 'teacher enthusiasm'}, {'answer_start': 22, 'te
```

```
In [18]: Alignement_graph_bipartite( "How many Catholic schools were in Victoria?","As of August
         affiche_table_cosine("How many Catholic schools were in Victoria?","As of August 2010,
         question=" How many Catholic schools were in Victoria? "
         sequence=" As of August 2010, Victoria had 1,548 public schools, 489 Catholic schools a
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: schools ")
         print("reponses attendu: [{'answer_start': 54, 'text': '489'}, {'answer_start': 54, 'te
```


how
0.33782753
as
of
august
2010
,
victoria
had
1,548
public
schools
,
489
catholic
schools
and
214
independent
schools
.
many
catholic
schools
were
in
victoria
0.5296667
1.0000001
1.0
0.6150833
0.56963885
1.0
?
0.41309536

```
                 how      many  catholic   schools      were        in  \
as          0.337828  0.407014  0.218580  0.212330  0.302544  0.388784
of          0.248074  0.372540  0.345208  0.265506  0.340383  0.569639
august      0.133693  0.147423  0.189286  0.121073  0.204387  0.243079
2010        0.179273  0.338788  0.267908  0.269017  0.310222  0.269296
,           0.226594  0.406108  0.271398  0.276052  0.367555  0.520147
victoria    0.090041  0.192616  0.213801  0.114028  0.195463  0.208231
had         0.254914  0.401307  0.211014  0.219546  0.615083  0.385027
1,548      -0.111434 -0.023219  0.242494  0.093147 -0.020421  0.128553
public      0.156247  0.251150  0.273276  0.426081  0.169887  0.290747
schools     0.128744  0.310809  0.325232  1.000000  0.238761  0.263213
,           0.226594  0.406108  0.271398  0.276052  0.367555  0.520147
489         0.062446  0.087930 -0.025141  0.092507  0.012206  0.028045
catholic    0.114918  0.189360  1.000000  0.325232  0.229464  0.263139
schools     0.128744  0.310809  0.325232  1.000000  0.238761  0.263213
and         0.336480  0.529627  0.214783  0.293200  0.421755  0.495406
214         0.137365 -0.020941  0.029995 -0.069962  0.014858  0.055350
independent 0.174138  0.202611  0.235080  0.222316  0.195256  0.279461
schools     0.128744  0.310809  0.325232  1.000000  0.238761  0.263213
.           0.328722  0.397423  0.270731  0.275271  0.337824  0.559157

             victoria         ?
as          0.195159  0.270439
of          0.250544  0.182614
august      0.153771  0.300674
2010        0.252377  0.212328
,           0.227936  0.310912
victoria    1.000000  0.122469
had         0.199213  0.207284
1,548       0.143097 -0.101721
public      0.196297  0.198730
schools     0.114028  0.183493
,           0.227936  0.310912
489         0.047028  0.035369
catholic    0.213801  0.166099
schools     0.114028  0.183493
and         0.245073  0.205442
214        -0.070581  0.074473
independent 0.138029  0.119026
schools     0.114028  0.183493
.           0.230691  0.413095
question =   How many Catholic schools were in Victoria?
sequence =   As of August 2010, Victoria had 1,548 public schools, 489 Catholic schools and 214
span output: schools
reponses attendu: [{'answer_start': 54, 'text': '489'}, {'answer_start': 54, 'text': '489'}, {'a
```

```
In [19]: Alignement_graph_bipartite( "Who initially wanted more unmanned testing done regarding
         affiche_table_cosine("Who initially wanted more unmanned testing done regarding the LM?
         question=" Who initially wanted more unmanned testing done regarding the LM? "
         sequence=" Although Grumman wanted a second unmanned test, George Low decided the next
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: manned ")
         print("reponses attendu: [{'answer_start': 474, 'text': 'Grumman'}, {'answer_start': 52
```

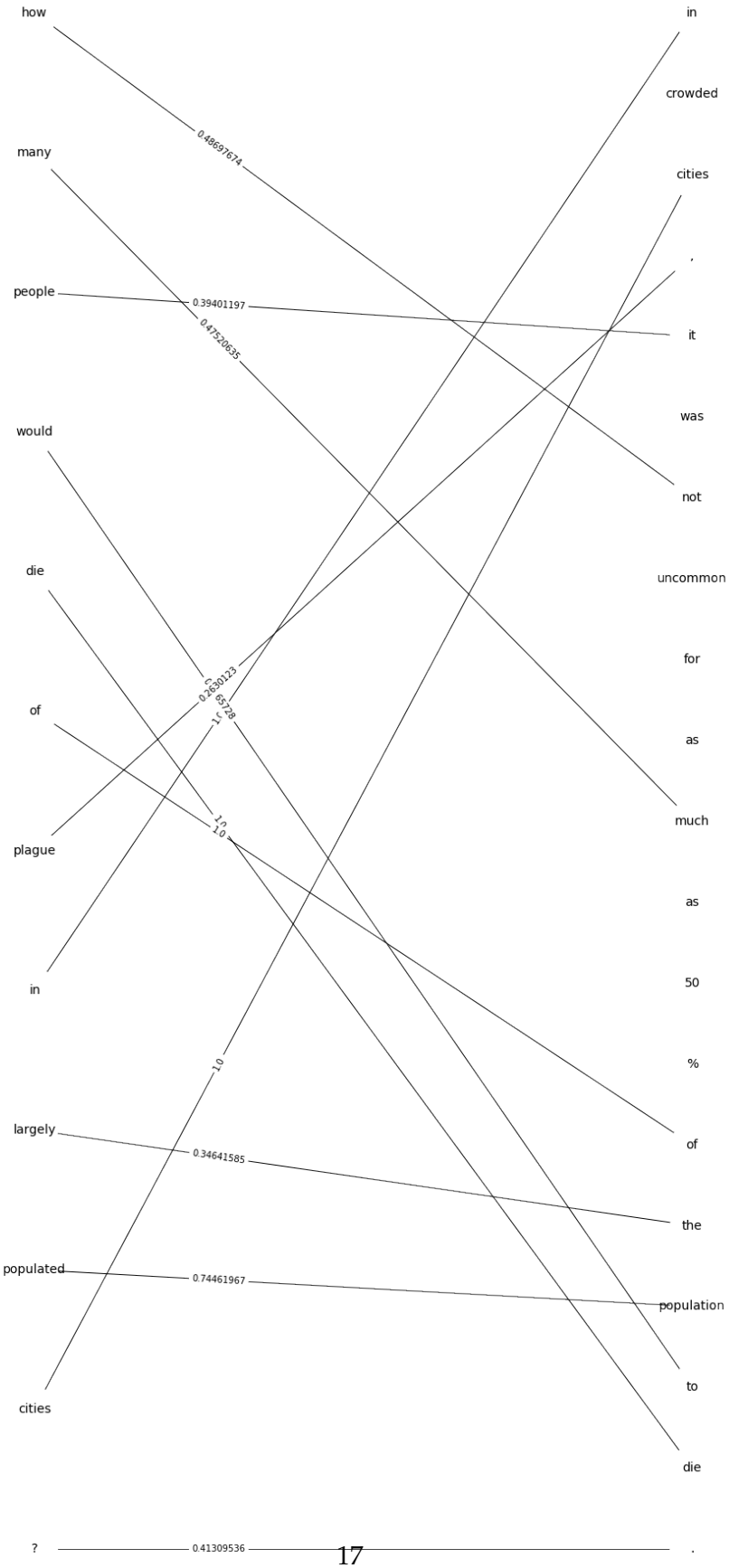

who
initially
wanted
although
grumman
wanted
a
second
0.3660895
0.3632045
1.0
more
unmanned
testing
test
,
george
low
decided
the
next
lm
flight
would
be
manned
1.0000001
0.771741
0.54453635
done
regarding
the
lm
0.27582487
0.34642425
1.0
1.0000001
?
.
0.41309536

	who	initially	wanted	more	unmanned	testing	\
although	0.300338	0.363205	0.218231	0.387962	0.163335	0.179154	
grumman	0.178640	0.277686	0.183927	0.151444	0.401299	0.236163	
wanted	0.386303	0.248153	1.000000	0.275864	0.150434	0.119765	
a	0.366089	0.255322	0.257022	0.345282	0.224304	0.229797	
second	0.219202	0.192042	0.169579	0.203630	0.176156	0.182498	
unmanned	0.184288	0.245118	0.150434	0.210095	1.000000	0.389544	
test	0.145408	0.211780	0.127473	0.165808	0.317706	0.717174	
,	0.336573	0.234033	0.275374	0.369652	0.175002	0.188090	
george	0.258050	0.210653	0.153590	0.153104	0.057924	0.075413	
low	0.161454	0.289939	0.097733	0.246337	0.195331	0.196295	
decided	0.300911	0.257369	0.682567	0.229268	0.197205	0.207890	
the	0.373393	0.299702	0.301105	0.397584	0.199820	0.250756	
next	0.191921	0.159012	0.245019	0.230932	0.154500	0.138481	
lm	0.142317	0.228577	0.055830	0.186795	0.426200	0.334509	
flight	0.077882	0.114526	0.105713	0.177104	0.441357	0.197172	
would	0.363298	0.298937	0.637761	0.446984	0.245938	0.259506	
be	0.337164	0.340661	0.254122	0.544536	0.248651	0.269383	
manned	0.207817	0.245327	0.177364	0.235672	0.827189	0.349764	
.	0.354134	0.239417	0.306808	0.407023	0.199463	0.230632	

	done	regarding	the	lm	?
although	0.286205	0.212723	0.326582	0.166178	0.155412
grumman	0.097063	0.118782	0.251509	0.293013	0.238288
wanted	0.282562	0.217677	0.301105	0.055830	0.243427
a	0.280001	0.224000	0.582920	0.181978	0.239167
second	0.115524	0.152855	0.462030	0.133511	0.172396
unmanned	0.223928	0.210110	0.199820	0.426200	0.208503
test	0.304133	0.159450	0.192353	0.279256	0.198327
,	0.320791	0.275825	0.493308	0.159341	0.310912
george	0.121988	0.148909	0.251076	0.108723	0.147902
low	0.188894	0.118124	0.254418	0.176372	0.106589
decided	0.300195	0.258125	0.375346	0.123464	0.258772
the	0.265759	0.292809	1.000000	0.255348	0.220681
next	0.173302	0.103406	0.358331	0.154549	0.167460
lm	0.226336	0.248708	0.255348	1.000000	0.227060
flight	0.131100	0.086342	0.249822	0.277470	0.205685
would	0.346424	0.323611	0.413375	0.203948	0.371790
be	0.396597	0.318203	0.424814	0.239490	0.332426
manned	0.296065	0.259270	0.239294	0.401778	0.175141
.	0.313232	0.273433	0.593451	0.229236	0.413095

question = Who initially wanted more unmanned testing done regarding the LM?
sequence = Although Grumman wanted a second unmanned test, George Low decided the next LM flig
span output: manned
reponses attendu: [{'answer_start': 474, 'text': 'Grumman'}, {'answer_start': 520, 'text': 'Low'

```
In [20]: Alignement_graph_bipartite( "How many people would die of plague in largely populated c
         affiche_table_cosine("How many people would die of plague in largely populated cities?"
         question=" How many people would die of plague in largely populated cities? "
         sequence=" In crowded cities, it was not uncommon for as much as 50% of the population
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: uncommon for as ")
         print("reponses attendu: [{'answer_start': 1202, 'text': 'as much as 50%'}, {'answer_st
```


how
many
people
would
die
of
plague
in
largely
populated
cities
?
in
crowded
cities
,
it
was
not
uncommon
for
as
much
as
50
%
of
the
population
to
die
.
0.49697674
0.47520635
0.39401197
0.7630123
0.65728
1.0
1.0
1.0
0.34641585
0.74461967
0.41309536

	how	many	people	would	die	of	\
in	0.260117	0.414742	0.396057	0.304612	0.186414	0.569639	
crowded	0.200326	0.278949	0.282145	0.300871	0.276262	0.257379	
cities	0.104348	0.286777	0.427417	0.177225	0.101918	0.295281	
,	0.226594	0.406108	0.366446	0.325868	0.239153	0.494608	
it	0.367150	0.396159	0.394012	0.363987	0.184069	0.387615	
was	0.166782	0.260356	0.251842	0.416631	0.155001	0.406907	
not	0.486977	0.436973	0.283957	0.585452	0.243372	0.330709	
uncommon	0.204828	0.344894	0.224780	0.201891	0.158757	0.307168	
for	0.330817	0.409728	0.249000	0.398775	0.120363	0.418599	
as	0.337828	0.407014	0.157047	0.382401	0.171330	0.388661	
much	0.465566	0.475206	0.294469	0.417203	0.129937	0.267110	
as	0.337828	0.407014	0.157047	0.382401	0.171330	0.388661	
50	0.014565	-0.037946	-0.061152	0.059570	-0.010429	0.006577	
%	0.123740	0.249484	0.261924	0.169932	0.156239	0.199534	
of	0.248074	0.372540	0.325104	0.335394	0.213242	1.000000	
the	0.338904	0.433521	0.368385	0.413375	0.224227	0.733189	
population	0.165106	0.220197	0.356241	0.137264	0.118754	0.306089	
to	0.441697	0.402437	0.372757	0.596573	0.147466	0.415485	
die	0.150835	0.212310	0.234705	0.229369	1.000000	0.213242	
.	0.328722	0.397423	0.331190	0.350672	0.265850	0.535246	

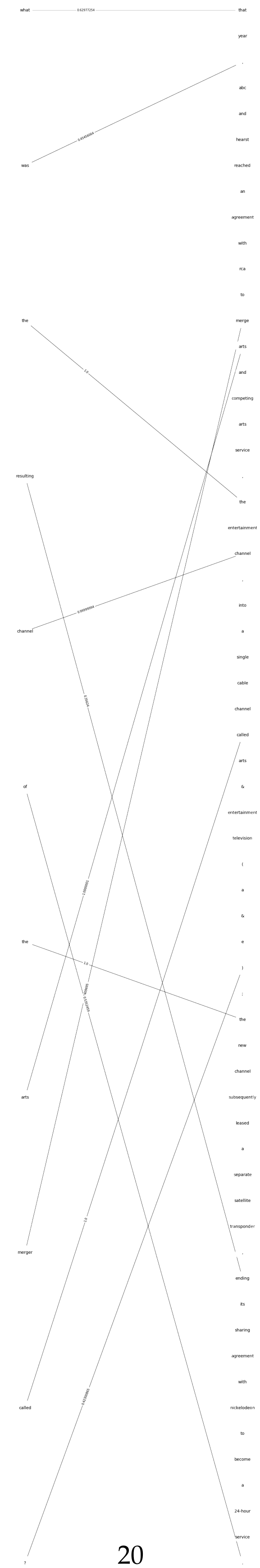
	plague	in	largely	populated	cities	?
in	0.248876	1.000000	0.237191	0.315952	0.322936	0.204310
crowded	0.227181	0.301665	0.141825	0.272621	0.284248	0.113262
cities	0.241249	0.322936	0.139489	0.419425	1.000000	0.133393
,	0.263012	0.520147	0.242128	0.305056	0.319276	0.310912
it	0.281223	0.513081	0.324360	0.355192	0.241297	0.257322
was	0.243120	0.455141	0.307300	0.211541	0.175984	0.184121
not	0.206087	0.340311	0.314571	0.241936	0.165524	0.419429
uncommon	0.241356	0.222836	0.278706	0.241371	0.114957	0.224110
for	0.147107	0.432223	0.220819	0.176304	0.212541	0.264628
as	0.150151	0.388784	0.241245	0.247184	0.191712	0.270439
much	0.188287	0.272522	0.348268	0.220197	0.204084	0.234422
as	0.150151	0.388784	0.241245	0.247184	0.191712	0.270439
50	0.043830	0.021892	-0.007866	-0.018021	-0.068443	0.030513
%	0.138379	0.194653	0.324030	0.299585	0.165964	0.214418
of	0.274556	0.569639	0.308716	0.285046	0.295281	0.182614
the	0.253776	0.666239	0.346416	0.330483	0.262117	0.220681
population	0.284637	0.326845	0.301879	0.744620	0.391951	0.144495
to	0.218457	0.450107	0.243482	0.219776	0.207353	0.294633
die	0.326130	0.186414	0.106014	0.069759	0.101918	0.168911
.	0.254765	0.559157	0.198394	0.239180	0.252522	0.413095

question = How many people would die of plague in largely populated cities?

sequence = In crowded cities, it was not uncommon for as much as 50% of the population to die.

```
span output: uncommon for as
reponses attendu: [{'answer_start': 1202, 'text': 'as much as 50%'}, {'answer_start': 1213, 'tex
```

```
In [21]: Alignement_graph_bipartite( "What was the resulting channel of the ARTS merger called?"
         affiche_table_cosine("What was the resulting channel of the ARTS merger called?","That
         question=" What was the resulting channel of the ARTS merger called? "
         sequence=" That year, ABC and Hearst reached an agreement with RCA to merge ARTS and co
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: new channel subsequently leased a separate satellite transponder ,
         print("reponses attendu: [{'answer_start': 715, 'text': 'Arts & Entertainment Televisio
```


what
was
the
resulting
channel
of
the
arts
merger
called
?
that
year
,
abc
and
hearst
reached
an
agreement
with
rca
to
merge
arts
and
competing
arts
service
,
the
entertainment
channel
,
into
a
single
cable
channel
called
arts
&
entertainment
television
(
a
&
e
)
;
the
new
channel
subsequently
leased
a
separate
satellite
transponder
,
ending
its
sharing
agreement
with
nickelodeon
to
become
a
24-hour
service
.

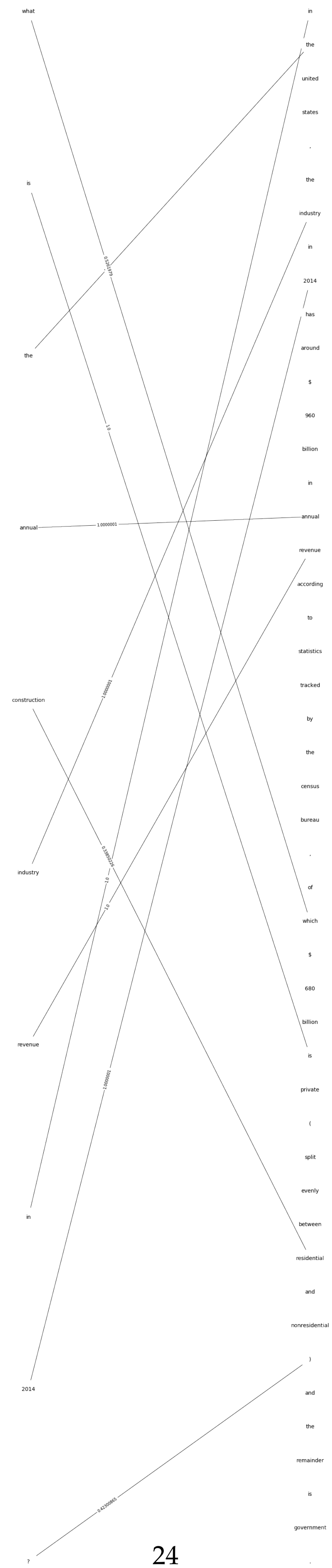
	what	was	the	resulting	channel	of	\
that	0.629773	0.336702	0.547399	0.303833	0.207358	0.414982	
year	0.151848	0.320984	0.360968	0.191836	0.147418	0.257831	
,	0.308747	0.454561	0.493308	0.248778	0.187303	0.494608	
abc	0.131734	0.215658	0.206876	0.148614	0.402018	0.146073	
and	0.313092	0.374096	0.517126	0.332499	0.248370	0.511697	
hearst	0.293649	0.335477	0.288464	0.131994	0.194476	0.277150	
reached	0.202025	0.311094	0.298610	0.260662	0.132198	0.186040	
an	0.282188	0.369437	0.420749	0.236423	0.159659	0.322809	
agreement	0.285109	0.239753	0.249070	0.243934	0.159829	0.212595	
with	0.246484	0.274119	0.422349	0.318028	0.175262	0.418273	
rca	0.142003	0.258932	0.126737	0.161458	0.152957	0.098286	
to	0.455848	0.363635	0.535392	0.268090	0.212991	0.415485	
merge	0.302105	0.157665	0.266165	0.151036	0.188092	0.192208	
arts	0.159752	0.254449	0.233746	0.068819	0.109084	0.292113	
and	0.313092	0.374096	0.517126	0.332499	0.248370	0.511697	
competing	0.246170	0.230552	0.258451	0.180763	0.165005	0.198615	
arts	0.159752	0.254449	0.233746	0.068819	0.109084	0.292113	
service	0.168572	0.226890	0.226610	0.087469	0.307428	0.184256	
,	0.308747	0.454561	0.493308	0.248778	0.187303	0.494608	
the	0.427457	0.487115	1.000000	0.390541	0.286004	0.733189	
entertainment	0.216488	0.216847	0.190779	0.107833	0.251549	0.139267	
channel	0.202959	0.156244	0.286004	0.194893	1.000000	0.167190	
,	0.308747	0.454561	0.493308	0.248778	0.187303	0.494608	
into	0.218062	0.232222	0.394932	0.288477	0.165948	0.322559	
a	0.332040	0.450847	0.582920	0.351084	0.209142	0.505250	
single	0.244583	0.227859	0.293386	0.262905	0.192047	0.224557	
cable	0.145828	0.186184	0.259975	0.176715	0.514901	0.186602	
channel	0.202959	0.156244	0.286004	0.194893	1.000000	0.167190	
called	0.323596	0.339277	0.489897	0.217800	0.216924	0.422290	
arts	0.159752	0.254449	0.233746	0.068819	0.109084	0.292113	
...	...	...	...	...	...	...	
entertainment	0.216488	0.216847	0.190779	0.107833	0.251549	0.139267	
television	0.198381	0.251668	0.245696	0.156620	0.463238	0.219958	
(	0.241184	0.298034	0.332891	0.137907	0.155724	0.350322	
a	0.332040	0.450847	0.582920	0.351084	0.209142	0.505250	
&	0.111419	0.121340	0.149922	0.097852	0.118255	0.199005	
e	0.132433	0.124777	0.218959	0.122180	0.129288	0.193855	
)	0.262599	0.326764	0.345673	0.139991	0.190970	0.368036	
;	0.109869	0.075437	0.091527	-0.008248	-0.021438	0.054534	
the	0.427457	0.487115	1.000000	0.390541	0.286004	0.733189	
new	0.213038	0.274831	0.384524	0.226721	0.177625	0.343458	
channel	0.202959	0.156244	0.286004	0.194893	1.000000	0.167190	
subsequently	0.108977	0.348680	0.288651	0.280953	0.150877	0.193300	
leased	0.160863	0.378232	0.260277	0.165269	0.170086	0.243337	

a	0.332040	0.450847	0.582920	0.351084	0.209142	0.505250
separate	0.332151	0.198023	0.361292	0.225846	0.163225	0.333886
satellite	0.177396	0.211373	0.225638	0.167694	0.344352	0.168383
transponder	0.311979	0.189819	0.293889	0.273699	0.210198	0.295923
,	0.308747	0.454561	0.493308	0.248778	0.187303	0.494608
ending	0.185261	0.216825	0.372684	0.356240	0.171693	0.266915
its	0.303937	0.351887	0.559237	0.312971	0.298660	0.486904
sharing	0.259406	0.128453	0.251057	0.232720	0.227877	0.300996
agreement	0.285109	0.239753	0.249070	0.243934	0.159829	0.212595
with	0.246484	0.274119	0.422349	0.318028	0.175262	0.418273
nickelodeon	0.224311	0.255264	0.240173	0.141520	0.417713	0.142558
to	0.455848	0.363635	0.535392	0.268090	0.212991	0.415485
become	0.271340	0.314744	0.317317	0.190452	0.105735	0.316780
a	0.332040	0.450847	0.582920	0.351084	0.209142	0.505250
24-hour	0.273288	0.236908	0.292632	0.258391	0.356644	0.229091
service	0.168572	0.226890	0.226610	0.087469	0.307428	0.184256
.	0.356790	0.401462	0.593451	0.294120	0.248915	0.535246

	arts	merger	called	?
that	0.211317	0.156310	0.416046	0.327163
year	0.142581	0.138497	0.154254	0.183150
,	0.252858	0.255696	0.265491	0.310912
abc	0.107372	0.076029	0.165625	0.190652
and	0.301350	0.168023	0.325176	0.205442
hearst	0.188492	0.224682	0.190230	0.150238
reached	0.007372	0.132234	0.183488	0.151802
an	0.141257	0.130944	0.285863	0.197939
agreement	0.104275	0.339847	0.174295	0.158302
with	0.199730	0.197417	0.313907	0.243331
rca	0.155596	0.224327	0.157214	0.073058
to	0.238828	0.110127	0.298714	0.294633
merge	0.114677	0.724097	0.148992	0.282160
arts	1.000000	0.184042	0.170352	0.094714
and	0.301350	0.168023	0.325176	0.205442
competing	0.230616	0.226714	0.176584	0.142817
arts	1.000000	0.184042	0.170352	0.094714
service	0.234320	0.204682	0.172571	0.167630
,	0.252858	0.255696	0.265491	0.310912
the	0.233746	0.255735	0.489897	0.220681
entertainment	0.280953	0.182840	0.192429	0.118552
channel	0.109084	0.146574	0.216924	0.201762
,	0.252858	0.255696	0.265491	0.310912
into	0.083773	0.282343	0.349519	0.156625
a	0.213889	0.207948	0.486881	0.239167
single	0.038870	0.080557	0.209311	0.215112
cable	0.143301	0.160291	0.167932	0.139939
channel	0.109084	0.146574	0.216924	0.201762
called	0.170352	0.196264	1.000000	0.189633

```
arts            1.000000  0.184042  0.170352  0.094714
...                  ...       ...       ...       ...
entertainment   0.280953  0.182840  0.192429  0.118552
television      0.266343  0.100486  0.175996  0.168258
(               0.178902  0.201024  0.254643  0.405615
a               0.213889  0.207948  0.486881  0.239167
&               0.176121  0.093532  0.122435  0.125964
e               0.117067  0.098537  0.170902  0.167578
)               0.201166  0.216075  0.273923  0.423009
;               0.054967  0.044640  0.022543  0.103498
the             0.233746  0.255735  0.489897  0.220681
new             0.195433  0.222092  0.258644  0.147828
channel         0.109084  0.146574  0.216924  0.201762
subsequently    0.118034  0.126347  0.116573  0.005594
leased          0.191185  0.174344  0.284329  0.165531
a               0.213889  0.207948  0.486881  0.239167
separate        0.165700  0.317513  0.365654  0.168357
satellite       0.119008  0.274605  0.170516  0.135684
transponder     0.166747  0.207539  0.306173  0.268521
,               0.252858  0.255696  0.265491  0.310912
ending          0.125010  0.153274  0.241662  0.141821
its             0.224218  0.186442  0.284070  0.151035
sharing         0.216759  0.211752  0.169044  0.162113
agreement       0.104275  0.339847  0.174295  0.158302
with            0.199730  0.197417  0.313907  0.243331
nickelodeon     0.177931  0.235892  0.132655  0.250988
to              0.238828  0.110127  0.298714  0.294633
become          0.143310  0.202623  0.297629  0.195113
a               0.213889  0.207948  0.486881  0.239167
24-hour         0.188453  0.095499  0.217806  0.265620
service         0.234320  0.204682  0.172571  0.167630
.               0.221740  0.232660  0.300345  0.413095

[61 rows x 10 columns]
question =   What was the resulting channel of the ARTS merger called?
sequence =   That year, ABC and Hearst reached an agreement with RCA to merge ARTS and competing
span output: new channel subsequently leased a separate satellite transponder ,
reponses attendu: [{'answer_start': 715, 'text': 'Arts & Entertainment Television (A&E)'}, {'ans
```

```
In [22]: Alignement_graph_bipartite( "What is the annual construction industry revenue in 2014?"
         affiche_table_cosine("What is the annual construction industry revenue in 2014?","In th
         question=" What is the annual construction industry revenue in 2014? "
         sequence=" In the United States, the industry in 2014 has around $960 billion in annual
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: and the remainder is government . ")
         print("reponses attendu: [{'answer_start': 54, 'text': '$960 billion'}, {'answer_start'
```


what
is
the
annual
construction
industry
revenue
in
2014
?
in
the
united
states
,
the
industry
in
2014
has
around
$
960
billion
in
annual
revenue
according
to
statistics
tracked
by
the
census
bureau
,
of
which
$
680
billion
is
private
(
split
evenly
between
residential
and
nonresidential
)
and
the
remainder
is
government
.

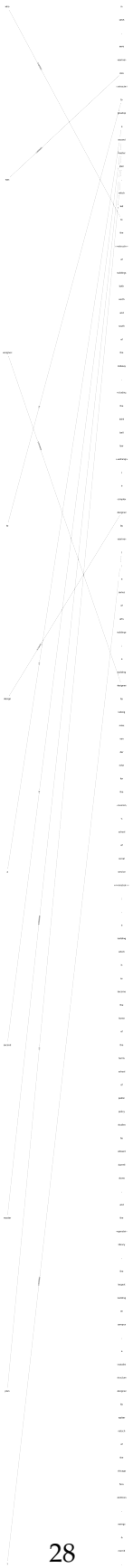
	what	is	the	annual	construction	\
in	0.308086	0.539444	0.666239	0.255778	0.303130	
the	0.427457	0.582367	1.000000	0.295898	0.267198	
united	0.200367	0.224085	0.343677	0.214941	0.202628	
states	0.215260	0.219390	0.346358	0.176051	0.169250	
,	0.308747	0.469191	0.493308	0.168394	0.237043	
the	0.427457	0.582367	1.000000	0.295898	0.267198	
industry	0.196630	0.257967	0.269429	0.246220	0.340754	
in	0.308086	0.539444	0.666239	0.255778	0.303130	
2014	0.041483	0.167459	0.059491	-0.028749	-0.078640	
has	0.203381	0.566875	0.417754	0.205563	0.158675	
around	0.274753	0.286477	0.400858	0.233271	0.208020	
$	0.248589	0.145779	0.233675	0.243553	0.217353	
960	0.059317	0.084467	0.191540	0.066510	0.110064	
billion	0.262669	0.160687	0.216233	0.231584	0.235215	
in	0.308086	0.539444	0.666239	0.255778	0.303130	
annual	0.147990	0.205384	0.295898	1.000000	0.174705	
revenue	0.130917	0.110669	0.225626	0.275431	0.299304	
according	0.277621	0.283899	0.320297	0.178177	0.151684	
to	0.455848	0.427451	0.535392	0.202341	0.248103	
statistics	0.282728	0.158820	0.239756	0.218206	0.122765	
tracked	0.225452	0.151178	0.271206	0.150188	0.196243	
by	0.245548	0.389985	0.440875	0.167768	0.192111	
the	0.427457	0.582367	1.000000	0.295898	0.267198	
census	0.206100	0.184032	0.252499	0.157463	0.109359	
bureau	0.183440	0.194182	0.263024	0.214122	0.219862	
,	0.308747	0.469191	0.493308	0.168394	0.237043	
of	0.332908	0.485432	0.733189	0.295350	0.281574	
which	0.526198	0.445740	0.652182	0.306855	0.316496	
$	0.248589	0.145779	0.233675	0.243553	0.217353	
680	0.234004	0.147405	0.154862	0.103945	0.081560	
billion	0.262669	0.160687	0.216233	0.231584	0.235215	
is	0.288376	1.000000	0.582367	0.205384	0.213566	
private	0.167781	0.142091	0.182914	0.162406	0.158791	
(	0.241184	0.353930	0.332891	0.179875	0.167597	
split	0.280210	0.147904	0.296763	0.083603	0.126897	
evenly	0.245928	0.274042	0.358467	0.278163	0.102893	
between	0.246216	0.268096	0.390670	0.231232	0.170004	
residential	0.098147	0.250221	0.255979	0.281198	0.338502	
and	0.313092	0.385975	0.517126	0.208473	0.227624	
nonresidential	0.151387	0.237045	0.221771	0.271416	0.278431	
)	0.262599	0.377322	0.345673	0.178905	0.183372	
and	0.313092	0.385975	0.517126	0.208473	0.227624	
the	0.427457	0.582367	1.000000	0.295898	0.267198	
remainder	0.164470	0.221311	0.367951	0.096457	0.211989	

is	0.288376	1.000000	0.582367	0.205384	0.213566
government	0.280745	0.203435	0.317981	0.187892	0.293066
.	0.356790	0.456966	0.593451	0.179685	0.236828

	industry	revenue	in	2014	?
in	0.267512	0.182089	1.000000	0.139962	0.204310
the	0.269429	0.225626	0.666239	0.059491	0.220681
united	0.275682	0.293408	0.393848	0.017225	0.172669
states	0.257594	0.306164	0.397744	0.009115	0.162727
,	0.245026	0.182250	0.520147	0.193502	0.310912
the	0.269429	0.225626	0.666239	0.059491	0.220681
industry	1.000000	0.334342	0.267512	0.028146	0.118133
in	0.267512	0.182089	1.000000	0.139962	0.204310
2014	0.028146	-0.060072	0.139962	1.000000	0.058128
has	0.269778	0.137467	0.417326	0.083330	0.159945
around	0.230359	0.152747	0.412941	0.036158	0.147297
$	0.166565	0.486181	0.216168	0.045880	0.278613
960	0.054786	0.036080	0.144221	0.033146	0.083525
billion	0.264320	0.462026	0.184758	0.028369	0.257918
in	0.267512	0.182089	1.000000	0.139962	0.204310
annual	0.246220	0.275431	0.255778	-0.028749	0.108349
revenue	0.334342	1.000000	0.182089	-0.060072	0.164132
according	0.153897	0.243770	0.292593	0.050928	0.211735
to	0.248767	0.186519	0.450107	0.016496	0.294633
statistics	0.232924	0.297671	0.243375	-0.006086	0.228869
tracked	0.148877	0.171894	0.196593	-0.021495	0.157663
by	0.195282	0.197761	0.370934	0.142259	0.179666
the	0.269429	0.225626	0.666239	0.059491	0.220681
census	0.145108	0.270799	0.257131	0.047312	0.125482
bureau	0.273511	0.330996	0.232861	0.013348	0.127161
,	0.245026	0.182250	0.520147	0.193502	0.310912
of	0.273856	0.224009	0.569639	0.049306	0.182614
which	0.286524	0.233123	0.484052	-0.052018	0.258146
$	0.166565	0.486181	0.216168	0.045880	0.278613
680	0.021278	0.107162	0.093723	-0.160675	0.293757
billion	0.264320	0.462026	0.184758	0.028369	0.257918
is	0.257967	0.110669	0.539444	0.167459	0.229614
private	0.175852	0.207345	0.211358	-0.012681	0.143890
(	0.108042	0.163020	0.308743	0.124349	0.405615
split	0.091899	0.112239	0.241718	0.070900	0.161339
evenly	0.147938	0.201063	0.256106	0.095166	0.106611
between	0.132994	0.157610	0.411004	0.009617	0.128871
residential	0.225145	0.256133	0.296868	0.004916	0.086033
and	0.261598	0.228098	0.495406	0.143337	0.205442
nonresidential	0.175597	0.261274	0.255144	0.030743	0.221057
)	0.140669	0.176032	0.332691	0.134229	0.423009
and	0.261598	0.228098	0.495406	0.143337	0.205442
the	0.269429	0.225626	0.666239	0.059491	0.220681

```
remainder       0.114162  0.246461  0.255095  0.040347  0.135795
is              0.257967  0.110669  0.539444  0.167459  0.229614
government      0.259290  0.386066  0.306679 -0.027861  0.188514
.               0.248578  0.257621  0.559157  0.145092  0.413095
question =   What is the annual construction industry revenue in 2014?
sequence =   In the United States, the industry in 2014 has around $960 billion in annual revenu
span output: and the remainder is government .
reponses attendu: [{'answer_start': 54, 'text': '$960 billion'}, {'answer_start': 54, 'text': '$
```

```
In [23]: Alignement_graph_bipartite( "Who was assigned to design a second master plan?","In 1955
         affiche_table_cosine("Who was assigned to design a second master plan?","In 1955, Eero
         question=" Who was assigned to design a second master plan? "
         sequence=" In 1955, Eero Saarinen was contracted to develop a second master plan, which
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: service administration ; , a building which is to become the home o
         print("reponses attendu: [{'answer_start': 89, 'text': 'Eero Saarinen'}, {'answer_start
```



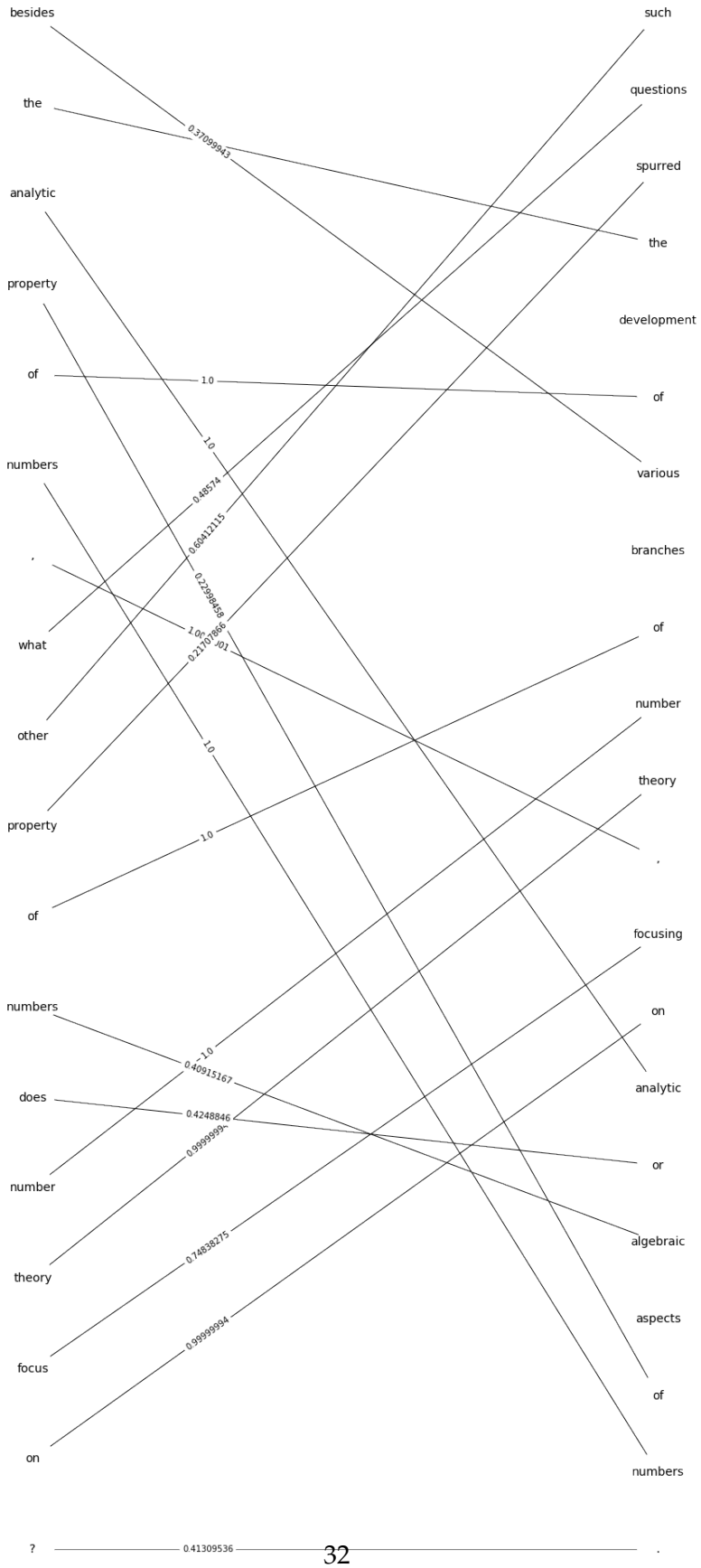
	who	was	assigned	to	design	a	\
in	0.279211	0.455141	0.227753	0.450107	0.218568	0.539381	
1955	-0.001535	0.047449	-0.037401	-0.031606	-0.079009	-0.059069	
,	0.336573	0.454561	0.211295	0.442459	0.193504	0.427140	
eero	0.238873	0.253060	0.141758	0.197497	0.383981	0.205425	
saarinen	0.200231	0.194668	0.173624	0.133130	0.363514	0.184962	
was	0.331882	1.000000	0.213246	0.363635	0.278080	0.450847	
contracted	0.263758	0.239288	0.294859	0.257016	0.161817	0.211350	
to	0.426566	0.363635	0.332006	1.000000	0.264150	0.455036	
develop	0.171545	0.153623	0.226177	0.337598	0.297038	0.271343	
a	0.366089	0.450847	0.286336	0.455036	0.252115	1.000000	
second	0.219202	0.385803	0.187896	0.309856	0.109765	0.365506	
master	0.301065	0.201551	0.218901	0.209494	0.211185	0.207899	
plan	0.295115	0.240772	0.225937	0.364219	0.335432	0.271570	
,	0.336573	0.454561	0.211295	0.442459	0.193504	0.427140	
which	0.368873	0.283096	0.266795	0.534506	0.302594	0.568945	
led	0.245080	0.344440	0.183214	0.218153	0.124603	0.190724	
to	0.426566	0.363635	0.332006	1.000000	0.264150	0.455036	
the	0.373393	0.487115	0.314982	0.535392	0.279135	0.582920	
construction	0.081259	0.224522	0.207198	0.248103	0.443448	0.250369	
of	0.282786	0.406907	0.314383	0.415485	0.252294	0.505250	
buildings	0.199000	0.235026	0.176342	0.252908	0.370938	0.227768	
both	0.316639	0.299077	0.285164	0.389771	0.225498	0.317179	
north	0.110267	0.290200	0.211321	0.274318	0.180395	0.250084	
and	0.412799	0.374096	0.201369	0.489713	0.218153	0.373925	
south	0.190689	0.279249	0.127807	0.256576	0.150871	0.236805	
of	0.282786	0.406907	0.314383	0.415485	0.252294	0.505250	
the	0.373393	0.487115	0.314982	0.535392	0.279135	0.582920	
midway	0.131617	0.188671	0.241539	0.229792	0.239216	0.171561	
,	0.336573	0.454561	0.211295	0.442459	0.193504	0.427140	
including	0.283584	0.218847	0.150954	0.240063	0.179294	0.210454	
...	...	...	...	...	...	...	
stone	0.230221	0.241109	0.105706	0.197679	0.197990	0.243663	
,	0.336573	0.454561	0.211295	0.442459	0.193504	0.427140	
and	0.412799	0.374096	0.201369	0.489713	0.218153	0.373925	
the	0.373393	0.487115	0.314982	0.535392	0.279135	0.582920	
regenstein	0.269324	0.221147	0.181300	0.183423	0.171755	0.219774	
library	0.158649	0.169163	0.185270	0.192527	0.184696	0.134801	
,	0.336573	0.454561	0.211295	0.442459	0.193504	0.427140	
the	0.373393	0.487115	0.314982	0.535392	0.279135	0.582920	
largest	0.124223	0.219871	0.129162	0.186872	0.117120	0.240423	
building	0.192643	0.237406	0.239501	0.265869	0.399189	0.257608	
on	0.234055	0.390922	0.263248	0.410314	0.214617	0.384837	
campus	0.181781	0.271132	0.168071	0.117957	0.244656	0.173323	
,	0.336573	0.454561	0.211295	0.442459	0.193504	0.427140	

a	0.366089	0.450847	0.286336	0.455036	0.252115	1.000000
brutalist	0.344139	0.301965	0.031133	0.098915	0.081207	0.193464
structure	0.151631	0.162635	0.136537	0.272003	0.349791	0.339685
designed	0.210970	0.283850	0.335366	0.307710	0.716781	0.284304
by	0.397278	0.475236	0.287081	0.335342	0.207551	0.320941
walter	0.281870	0.254563	0.083947	0.152030	0.181273	0.244793
netsch	0.223607	0.259766	0.228551	0.217388	0.204291	0.210987
of	0.282786	0.406907	0.314383	0.415485	0.252294	0.505250
the	0.373393	0.487115	0.314982	0.535392	0.279135	0.582920
chicago	0.204708	0.257484	0.188843	0.192781	0.218179	0.178999
firm	0.189499	0.226197	0.185588	0.155834	0.292021	0.219349
skidmore	0.271429	0.169269	0.193368	0.256226	0.352678	0.211608
,	0.336573	0.454561	0.211295	0.442459	0.193504	0.427140
owings	0.222050	0.271472	0.129751	0.192710	0.316541	0.140458
&	0.092446	0.121340	0.098872	0.125488	0.146416	0.072667
merrill	0.210183	0.273968	0.147280	0.134824	0.252434	0.168204
.	0.354134	0.401462	0.236554	0.464503	0.249958	0.504843

	second	master	plan	?
in	0.345103	0.178829	0.239247	0.204310
1955	0.000527	0.024275	-0.020363	-0.283727
,	0.299273	0.214511	0.212656	0.310912
eero	0.109546	0.258738	0.195883	0.197367
saarinen	0.093806	0.232694	0.164706	0.106816
was	0.385803	0.201551	0.240772	0.184121
contracted	0.254013	0.163203	0.165319	0.098552
to	0.309856	0.209494	0.364219	0.294633
develop	0.030722	0.177261	0.238526	0.157846
a	0.365506	0.207899	0.271570	0.239167
second	1.000000	0.170870	0.181786	0.172396
master	0.170870	1.000000	0.277016	0.119150
plan	0.181786	0.277016	1.000000	0.196105
,	0.299273	0.214511	0.212656	0.310912
which	0.336174	0.184312	0.250757	0.258146
led	0.244367	0.048479	0.199002	0.092739
to	0.309856	0.209494	0.364219	0.294633
the	0.462030	0.197691	0.285286	0.220681
construction	0.183107	0.163025	0.306712	0.134941
of	0.353416	0.170729	0.274213	0.182614
buildings	0.121761	0.128327	0.299899	0.111925
both	0.310541	0.161012	0.185744	0.194557
north	0.194267	0.111540	0.162352	0.118297
and	0.274391	0.198951	0.268445	0.205442
south	0.179156	0.067188	0.159480	0.123366
of	0.353416	0.170729	0.274213	0.182614
the	0.462030	0.197691	0.285286	0.220681
midway	0.133220	0.190087	0.287479	0.155566
,	0.299273	0.214511	0.212656	0.310912

```
including     0.116375  0.153802  0.158868  0.163330
...                ...       ...       ...       ...
stone         0.125417  0.148482  0.195053  0.119660
,             0.299273  0.214511  0.212656  0.310912
and           0.274391  0.198951  0.268445  0.205442
the           0.462030  0.197691  0.285286  0.220681
regenstein    0.133182  0.220661  0.174604  0.121171
library       0.136807  0.119623  0.187575  0.189387
,             0.299273  0.214511  0.212656  0.310912
the           0.462030  0.197691  0.285286  0.220681
largest       0.337076  0.132292  0.138042  0.124012
building      0.143386  0.147789  0.325715  0.135499
on            0.266591  0.153615  0.219097  0.271499
campus        0.127725  0.213612  0.258808  0.069527
,             0.299273  0.214511  0.212656  0.310912
a             0.365506  0.207899  0.271570  0.239167
brutalist     0.143236  0.156884  0.228116  0.103405
structure     0.185356  0.138992  0.295880  0.177077
designed      0.091171  0.201482  0.348450  0.112139
by            0.297768  0.210679  0.226636  0.179666
walter        0.103946  0.183103  0.192447  0.129721
netsch        0.154659  0.138779  0.186908  0.269641
of            0.353416  0.170729  0.274213  0.182614
the           0.462030  0.197691  0.285286  0.220681
chicago       0.179929  0.167889  0.167363  0.164470
firm          0.098634  0.120680  0.175281  0.087624
skidmore      0.056012  0.220971  0.324676  0.239913
,             0.299273  0.214511  0.212656  0.310912
owings        0.035990  0.154258  0.209128  0.163942
&             0.134268  0.140304  0.139263  0.125964
merrill       0.109724  0.189301  0.163223  0.208211
.             0.346034  0.168519  0.296624  0.413095

[118 rows x 10 columns]
question =   Who was assigned to design a second master plan?
sequence =   In 1955, Eero Saarinen was contracted to develop a second master plan, which led to
span output: service administration ; , a building which is to become the home of the harris sch
reponses attendu: [{'answer_start': 89, 'text': 'Eero Saarinen'}, {'answer_start': 89, 'text': '
```

```
In [24]: Alignement_graph_bipartite( "Besides the analytic property of numbers, what other prope
         affiche_table_cosine("Besides the analytic property of numbers, what other property of
         question=" Besides the analytic property of numbers, what other property of numbers doe
         sequence=" Such questions spurred the development of various branches of number theory,
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: aspects ")
         print("reponses attendu: [{'answer_start': 369, 'text': 'algebraic aspects'}, {'answer_
```

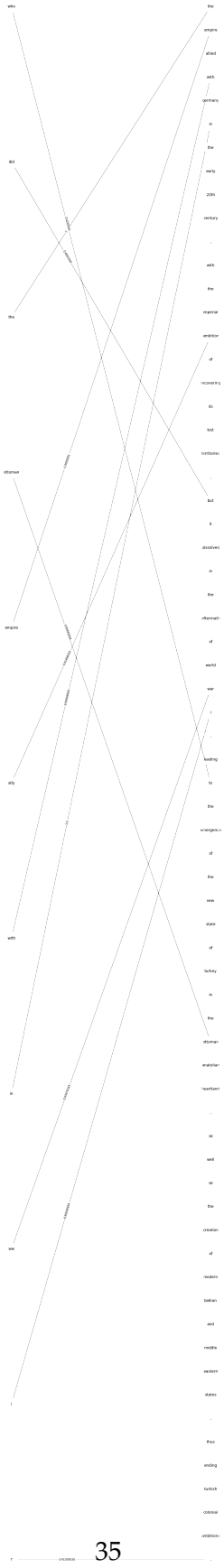

besides
the
analytic
property
of
numbers
,
what
other
property
of
numbers
does
number
theory
focus
on
?
such
questions
spurred
the
development
of
various
branches
of
number
theory
,
focusing
on
analytic
or
algebraic
aspects
of
numbers
.
0.37099943
1.0
0.48574
0.60412115
0.22998458
1.0
0.21707866
1.0
1.0
1.0
0.40915167
0.4248846
0.99999994
0.74838275
0.99999994
0.41309536

	besides	the	analytic	property	of	numbers	\
such	0.387094	0.357524	0.232333	0.127501	0.320953	0.273736	
questions	0.259984	0.235629	0.372428	0.181525	0.221028	0.247618	
spurred	0.157188	0.350657	0.256925	0.217079	0.322620	0.149850	
the	0.367077	1.000000	0.220675	0.206168	0.733189	0.293011	
development	0.216467	0.336204	0.307210	0.173706	0.287009	0.172028	
of	0.366572	0.733189	0.231133	0.229985	1.000000	0.228019	
various	0.370999	0.323504	0.291555	0.094127	0.319649	0.252948	
branches	0.253247	0.329501	0.237297	0.214789	0.301772	0.201326	
of	0.366572	0.733189	0.231133	0.229985	1.000000	0.228019	
number	0.249245	0.351457	0.254981	0.168261	0.277852	0.713212	
theory	0.210050	0.227716	0.489282	0.184990	0.235067	0.216120	
,	0.340059	0.493308	0.195011	0.195627	0.494608	0.264795	
focusing	0.326576	0.252008	0.295694	0.150969	0.292584	0.170037	
on	0.286339	0.510731	0.210605	0.110585	0.396693	0.204382	
analytic	0.213103	0.220675	1.000000	0.212546	0.231133	0.270060	
or	0.356546	0.417487	0.236622	0.295898	0.379260	0.287826	
algebraic	0.254096	0.204755	0.557807	0.283832	0.260537	0.409152	
aspects	0.277006	0.262029	0.375191	0.216174	0.261564	0.184980	
of	0.366572	0.733189	0.231133	0.229985	1.000000	0.228019	
numbers	0.250704	0.293011	0.270060	0.158231	0.228019	1.000000	
.	0.286362	0.593451	0.203980	0.229228	0.535246	0.259178	

	,	what	other	does	number	theory	\
such	0.396262	0.239896	0.604121	0.278719	0.300111	0.223165	
questions	0.280896	0.485740	0.292574	0.262910	0.195424	0.285696	
spurred	0.209165	0.208486	0.228582	0.210552	0.174617	0.284728	
the	0.493308	0.427457	0.441522	0.359280	0.351457	0.227716	
development	0.237769	0.190142	0.259964	0.193868	0.196862	0.360708	
of	0.494608	0.332908	0.385838	0.311982	0.277852	0.235067	
various	0.300932	0.260451	0.519729	0.211687	0.331177	0.212809	
branches	0.241965	0.172739	0.302467	0.194764	0.193942	0.239778	
of	0.494608	0.332908	0.385838	0.311982	0.277852	0.235067	
number	0.253570	0.230462	0.375476	0.183118	1.000000	0.241025	
theory	0.180958	0.297924	0.221777	0.180137	0.241025	1.000000	
,	1.000000	0.308747	0.444650	0.286174	0.253570	0.180958	
focusing	0.262909	0.342006	0.335764	0.217613	0.190325	0.194725	
on	0.490576	0.342261	0.341005	0.256880	0.297841	0.168739	
analytic	0.195011	0.329220	0.208989	0.245017	0.254981	0.489282	
or	0.376950	0.338671	0.532512	0.424885	0.283848	0.211219	
algebraic	0.207549	0.202852	0.242578	0.249905	0.306361	0.468668	
aspects	0.196502	0.278920	0.339196	0.319092	0.204470	0.354024	
of	0.494608	0.332908	0.385838	0.311982	0.277852	0.235067	
numbers	0.264795	0.245870	0.326913	0.214377	0.713212	0.216120	
.	0.681251	0.356790	0.415112	0.302299	0.325633	0.180478	

```
                focus        on         ?
such         0.288168  0.318116  0.178517
questions    0.248619  0.216658  0.429671
spurred      0.326286  0.274441  0.061637
the          0.253180  0.510731  0.220681
development  0.303452  0.238007  0.121136
of           0.242854  0.396693  0.182614
various      0.264633  0.267891  0.128235
branches     0.233068  0.153839  0.128583
of           0.242854  0.396693  0.182614
number       0.204462  0.297841  0.192568
theory       0.201940  0.168739  0.138729
,            0.238455  0.490576  0.310912
focusing     0.748383  0.351647  0.256239
on           0.315806  1.000000  0.271499
analytic     0.288592  0.210605  0.235798
or           0.247032  0.289723  0.323896
algebraic    0.150554  0.137919  0.231527
aspects      0.430713  0.232074  0.159247
of           0.242854  0.396693  0.182614
numbers      0.135685  0.204382  0.192980
.            0.211837  0.492570  0.413095
question =   Besides the analytic property of numbers, what other property of numbers does numbe
sequence =   Such questions spurred the development of various branches of number theory, focusi
span output: aspects
reponses attendu: [{'answer_start': 369, 'text': 'algebraic aspects'}, {'answer_start': 369, 'te
```

```
In [25]: Alignement_graph_bipartite( "Who did the Ottoman empire ally with in WW I?","The empire
         affiche_table_cosine("Who did the Ottoman empire ally with in WW I?","The empire allied
         question=" Who did the Ottoman empire ally with in WW I? "
         sequence=" The empire allied with Germany in the early 20th century, with the imperial
         print("question = ", question)
         print("sequence = ",sequence)
         print("span output: heartland , as well as the creation of modern balkan and middle eas
         print("reponses attendu: [{'answer_start': 358, 'text': 'Germany'}, {'answer_start': 35
```


who
did
the
ottoman
empire
ally
with
in
ww
i
.
the
empire
allied
with
germany
in
the
early
20th
century
.
with
the
imperial
ambition
of
recovering
its
lost
territories
.
but
it
dissolved
in
the
aftermath
of
world
war
i
.
leading
to
the
emergence
of
the
new
state
of
turkey
in
the
ottoman
anatolian
heartland
.
as
well
as
the
creation
of
modern
balkan
and
middle
eastern
states
.
thus
ending
turkish
colonial
ambitions
.

	who	did	the	ottoman	empire	ally	\
the	0.373393	0.347367	1.000000	0.262308	0.325322	0.236442	
empire	0.189779	0.188501	0.325322	0.586948	1.000000	0.310903	
allied	0.210107	0.222075	0.309215	0.420756	0.355121	0.365762	
with	0.328277	0.254721	0.422349	0.132191	0.161651	0.248658	
germany	0.202142	0.213283	0.318129	0.288397	0.284726	0.162048	
in	0.279211	0.259324	0.666239	0.170804	0.241106	0.169746	
the	0.373393	0.347367	1.000000	0.262308	0.325322	0.236442	
early	0.239857	0.320472	0.344422	0.186052	0.219489	0.198887	
20th	0.125972	0.129003	0.309751	0.155534	0.147225	0.162204	
century	0.299097	0.233374	0.361730	0.358072	0.363297	0.198886	
,	0.336573	0.301509	0.493308	0.227513	0.276469	0.273059	
with	0.328277	0.254721	0.422349	0.132191	0.161651	0.248658	
the	0.373393	0.347367	1.000000	0.262308	0.325322	0.236442	
imperial	0.110284	0.162544	0.307420	0.270636	0.388050	0.104127	
ambition	0.284204	0.302317	0.217980	0.194385	0.235968	0.414984	
of	0.282786	0.251390	0.733189	0.239910	0.305223	0.155785	
recovering	0.197352	0.234018	0.331515	0.194655	0.206284	0.220841	
its	0.132607	0.199227	0.559237	0.182377	0.229079	0.170820	
lost	0.220645	0.307793	0.337697	0.192927	0.227207	0.197088	
territories	0.223585	0.147332	0.301919	0.331681	0.428797	0.250493	
,	0.336573	0.301509	0.493308	0.227513	0.276469	0.273059	
but	0.448736	0.490337	0.485970	0.216872	0.225014	0.258336	
it	0.286840	0.299336	0.625118	0.199965	0.241148	0.257408	
dissolved	0.101195	0.204241	0.281504	0.202787	0.230849	0.233790	
in	0.279211	0.259324	0.666239	0.170804	0.241106	0.169746	
the	0.373393	0.347367	1.000000	0.262308	0.325322	0.236442	
aftermath	0.174232	0.274763	0.242276	0.211449	0.245998	0.212742	
of	0.282786	0.251390	0.733189	0.239910	0.305223	0.155785	
world	0.251583	0.194472	0.336043	0.210875	0.297638	0.246455	
war	0.242861	0.274220	0.366200	0.387969	0.394221	0.340457	
...	...	...	...	...	...	...	
the	0.373393	0.347367	1.000000	0.262308	0.325322	0.236442	
new	0.237805	0.227975	0.384524	0.117455	0.191503	0.136003	
state	0.156366	0.152080	0.309718	0.233186	0.307628	0.151819	
of	0.282786	0.251390	0.733189	0.239910	0.305223	0.155785	
turkey	0.191944	0.145019	0.245134	0.512392	0.283521	0.134544	
in	0.279211	0.259324	0.666239	0.170804	0.241106	0.169746	
the	0.373393	0.347367	1.000000	0.262308	0.325322	0.236442	
ottoman	0.202751	0.182489	0.262308	1.000000	0.586948	0.242170	
anatolian	0.257912	0.160853	0.279082	0.498025	0.351809	0.206598	
heartland	0.212350	0.115346	0.331635	0.257651	0.351693	0.241947	
,	0.336573	0.301509	0.493308	0.227513	0.276469	0.273059	
as	0.392305	0.359647	0.483816	0.218941	0.211083	0.287867	
well	0.348032	0.434174	0.426301	0.173305	0.205532	0.235482	

as	0.392305	0.359647	0.483816	0.218941	0.211083	0.287867
the	0.373393	0.347367	1.000000	0.262308	0.325322	0.236442
creation	0.242827	0.163651	0.346256	0.101034	0.189340	0.154499
of	0.282786	0.251390	0.733189	0.239910	0.305223	0.155785
modern	0.169373	0.213761	0.328397	0.232938	0.275668	0.140541
balkan	0.156019	0.130788	0.232827	0.508534	0.380166	0.246593
and	0.412799	0.344795	0.517126	0.245694	0.274796	0.272669
middle	0.160978	0.177322	0.340575	0.240094	0.259818	0.142478
eastern	0.180176	0.130090	0.392825	0.237449	0.314648	0.148139
states	0.213132	0.217546	0.346358	0.242419	0.285434	0.168643
,	0.336573	0.301509	0.493308	0.227513	0.276469	0.273059
thus	0.249033	0.236230	0.403688	0.200716	0.201662	0.259854
ending	0.114761	0.184661	0.372684	0.261850	0.255864	0.325064
turkish	0.223882	0.139103	0.195753	0.617157	0.297936	0.130244
colonial	0.186543	0.175757	0.271574	0.310520	0.362607	0.179354
ambitions	0.281479	0.322788	0.208270	0.201577	0.241962	0.386657
.	0.354134	0.352883	0.593451	0.252492	0.281172	0.225784

	with	in	ww	i	?
the	0.422349	0.666239	0.227490	0.266879	0.220681
empire	0.161651	0.241106	0.219102	0.247799	0.149095
allied	0.219941	0.222256	0.368629	0.235009	0.134110
with	1.000000	0.408093	0.149829	0.264071	0.243331
germany	0.195885	0.364284	0.284278	0.185101	0.138866
in	0.408093	1.000000	0.174242	0.229078	0.204310
the	0.422349	0.666239	0.227490	0.266879	0.220681
early	0.239199	0.346729	0.208508	0.127651	0.152729
20th	0.224451	0.207439	0.134938	0.208068	0.149856
century	0.146702	0.317403	0.200632	0.161677	0.102620
,	0.412074	0.520147	0.173453	0.346748	0.310912
with	1.000000	0.408093	0.149829	0.264071	0.243331
the	0.422349	0.666239	0.227490	0.266879	0.220681
imperial	0.150452	0.229536	0.251135	0.185937	0.109002
ambition	0.260509	0.182563	0.062169	0.247882	0.146467
of	0.418273	0.569639	0.212408	0.258494	0.182614
recovering	0.314673	0.306434	0.168112	0.165960	0.108044
its	0.308712	0.379378	0.184842	0.191599	0.151035
lost	0.189256	0.316766	0.197941	0.180126	0.152614
territories	0.211461	0.270011	0.204286	0.201113	0.112986
,	0.412074	0.520147	0.173453	0.346748	0.310912
but	0.449517	0.486745	0.126339	0.440766	0.279220
it	0.336905	0.513081	0.165436	0.344496	0.257322
dissolved	0.201174	0.265899	0.166744	0.145169	0.108018
in	0.408093	1.000000	0.174242	0.229078	0.204310
the	0.422349	0.666239	0.227490	0.266879	0.220681
aftermath	0.283537	0.245940	0.279415	0.248987	0.138482
of	0.418273	0.569639	0.212408	0.258494	0.182614
world	0.283282	0.303820	0.374645	0.172507	0.175796

```
war          0.205809  0.299041  0.434763  0.261459  0.155105
...               ...       ...       ...       ...       ...
the          0.422349  0.666239  0.227490  0.266879  0.220681
new          0.243256  0.375855  0.171110  0.213945  0.147828
state        0.206204  0.333161  0.134164  0.175085  0.116088
of           0.418273  0.569639  0.212408  0.258494  0.182614
turkey       0.169684  0.227271  0.114500  0.171737  0.198900
in           0.408093  1.000000  0.174242  0.229078  0.204310
the          0.422349  0.666239  0.227490  0.266879  0.220681
ottoman      0.132191  0.170804  0.209162  0.296352  0.144554
anatolian    0.161669  0.193623  0.130971  0.136366  0.105575
heartland    0.235736  0.287708  0.215088  0.155449  0.107331
,            0.412074  0.520147  0.173453  0.346748  0.310912
as           0.403874  0.388784  0.123579  0.285583  0.270439
well         0.372650  0.366127  0.133889  0.299626  0.248156
as           0.403874  0.388784  0.123579  0.285583  0.270439
the          0.422349  0.666239  0.227490  0.266879  0.220681
creation     0.236001  0.223948  0.213962  0.208014  0.240743
of           0.418273  0.569639  0.212408  0.258494  0.182614
modern       0.179692  0.298088  0.160195  0.127029  0.124754
balkan       0.185452  0.228608  0.181797  0.140445  0.083196
and          0.547773  0.495406  0.172706  0.319280  0.205442
middle       0.147290  0.267679  0.164729  0.116222  0.131396
eastern      0.227735  0.330636  0.205186  0.172646  0.200608
states       0.214241  0.397744  0.217745  0.131249  0.162727
,            0.412074  0.520147  0.173453  0.346748  0.310912
thus         0.329077  0.263582  0.092219  0.266315  0.187693
ending       0.284115  0.311942  0.163111  0.233657  0.141821
turkish      0.085761  0.157214  0.101700  0.155388  0.106385
colonial     0.076812  0.251962  0.150524  0.132829  0.101912
ambitions    0.262031  0.187138  0.085848  0.269434  0.157682
.            0.444615  0.559157  0.271846  0.336454  0.413095

[67 rows x 11 columns]
question =   Who did the Ottoman empire ally with in WW I?
sequence =   The empire allied with Germany in the early 20th century, with the imperial ambitio
span output: heartland , as well as the creation of modern balkan and middle eastern states , th
reponses attendu: [{'answer_start': 358, 'text': 'Germany'}, {'answer_start': 358, 'text': 'Germ
```