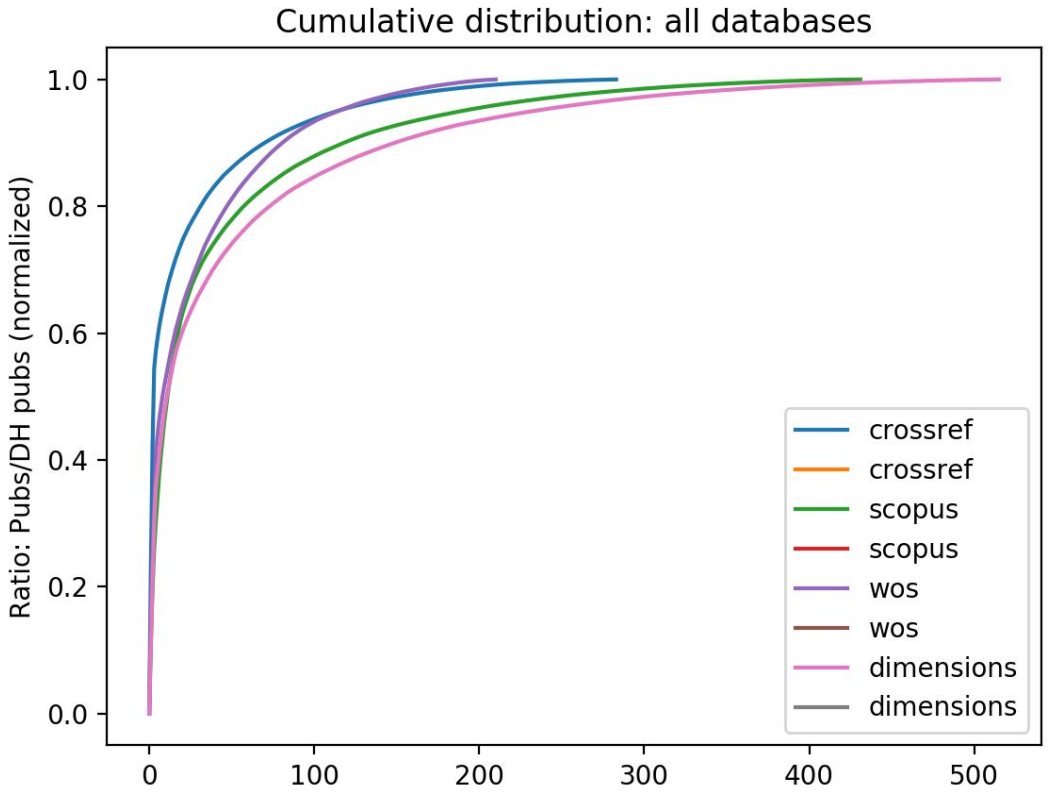
November 1st 2019

Cumulative distribution	DONE
Entropy-based measure => show distribution concentration respect to DH journals	
Group cluster with similars	
Compute euclidean and cosine distances	DONE
Normalize to the total of DH pubs (optional)	

Cumulative distribution with normalized ratio.
The ratio has been normalized over the number of DH



Computed with scipy Euclidean and cosine distances.

```
projects/thesis_chart/distances via [conda_env] charts_thesis
❯ python compute_distance.py
EUCLIDEAN DISTANCE

[[0.         0.00029359 0.0002     ... 0.04207613 0.07734178 0.03225917]
 [0.00029359 0.         0.00020641 ... 0.04197094 0.0772366  0.03225986]
 [0.0002     0.00020641 0.         ... 0.04205423 0.07732009 0.03225868]
 ...
 [0.04207613 0.04197094 0.04205423 ... 0.         0.03526586 0.0531596 ]
 [0.07734178 0.0772366  0.07732009 ... 0.03526586 0.         0.08396331]
 [0.03225917 0.03225986 0.03225868 ... 0.0531596  0.08396331 0.        ]]


COSINE DISTANCE

[[ 0.00000000e+00  4.45299804e-01  3.33333333e-01 ...  3.33333333e-01
   3.33333333e-01  1.00000000e+00]
 [ 4.45299804e-01  0.00000000e+00  1.67949706e-01 ...  1.67949706e-01
   1.67949706e-01  1.00000000e+00]
 [ 3.33333333e-01  1.67949706e-01  0.00000000e+00 ... -2.22044605e-16
   1.11022302e-16  1.00000000e+00]
 ...
 [ 3.33333333e-01  1.67949706e-01 -2.22044605e-16 ...  0.00000000e+00
   1.11022302e-16  1.00000000e+00]
 [ 3.33333333e-01  1.67949706e-01  1.11022302e-16 ...  1.11022302e-16
   0.00000000e+00  1.00000000e+00]
 [ 1.00000000e+00  1.00000000e+00  1.00000000e+00 ...  1.00000000e+00
   1.00000000e+00  0.00000000e+00]]
```