DESIDERATA FOR VISUALIZING UNCERTAINTY AND CONTEXTUALITY IN THE DIGITAL HUMANITIES

Experiences with the Digital Republic of Letters

Charles van den Heuvel & Scott Weingart
HuygensING-KNAW+Indiana University

Visualising Data Resources
The Potential of a Wikimedia Platform for the Digital Humanities
Forschungszentrum Gotha
27 – 28 April 2012

Introduction

Experiences with the Digital Republic of Letters (Circulation of Knowledge)

Data and Uncertainty

Data and Contextuality

Desiderata for visualizing uncertainty and contextuality

Circulation of Knowledge

Challenge

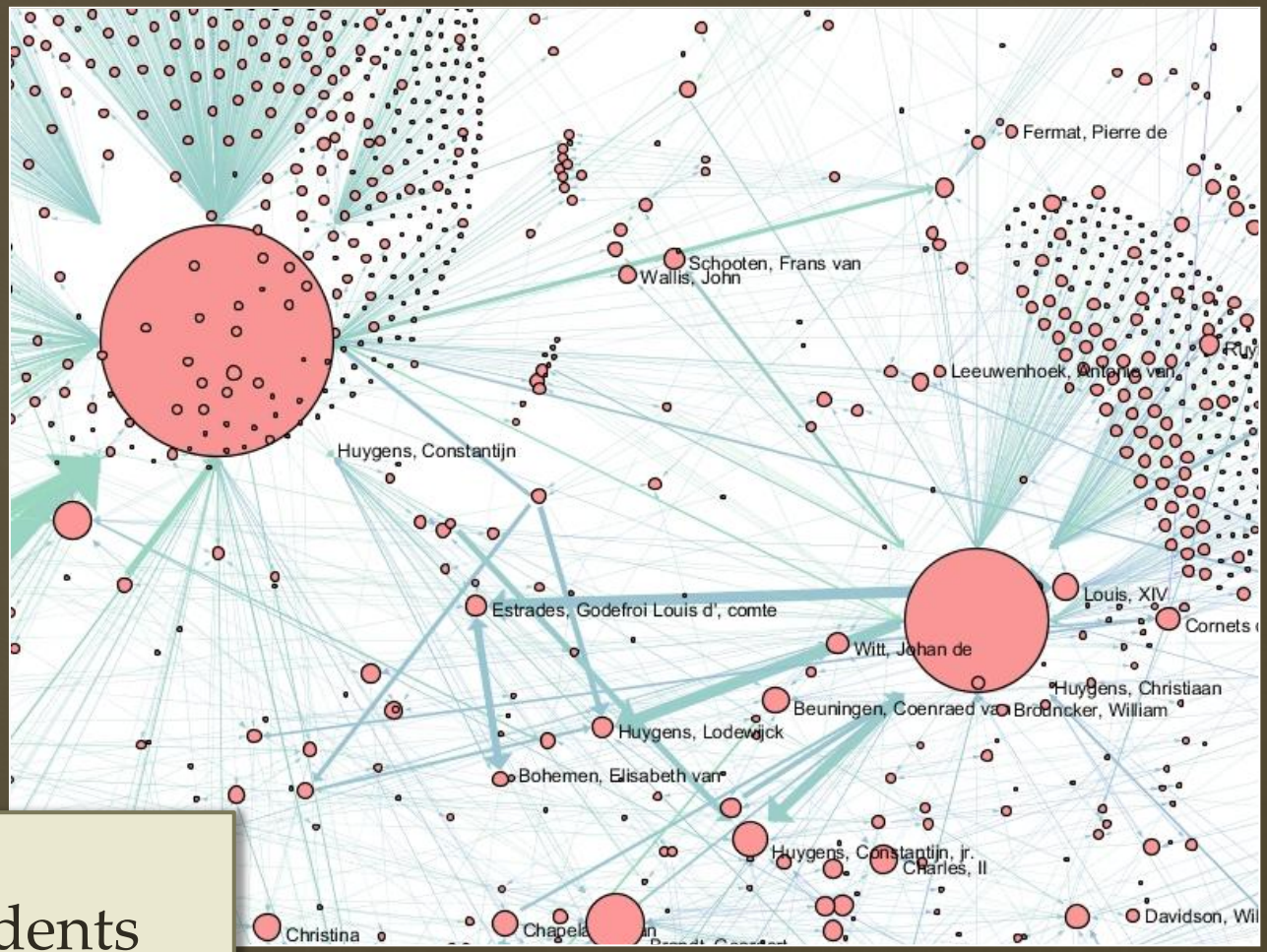
How can we combine and structure various sets of letters of 17th-century scholars and their correspondents in such a way that we can **analyze and visualize the circulation and appropriation of knowledge production** in a wider international context and recognize the **development of themes of interest and scholarly debate**s in space and time?

Letters and Data preparation

Corpus	Number of letters:	In posession?	Format TEI	Metadata	Normalized?
Grotius	7725	Yes		In Interp element	Yes, DBNL codes
Van Leeuwenhoek	337	Yes	TEI	In Interp element	Yes, DBNL codes
Descartes	375	Yes	XML (no TEI)	other markup	No, plain text
Barlaeus	1200	300 ready	Word	unknown	unknown
Swammerdam	80	Yes	Word	unknown	unknown
Constantijn Huygens	7295	Yes	xml	Probably Interp element	DBNL codes
Christiaan Huygens	2900?	No	probably TEI	Probably Interp element	DBNL codes

CEN Metadata Network

1550-1750



13 587 correspondents >700 in our corpus

Scott Weingart
Indiana University

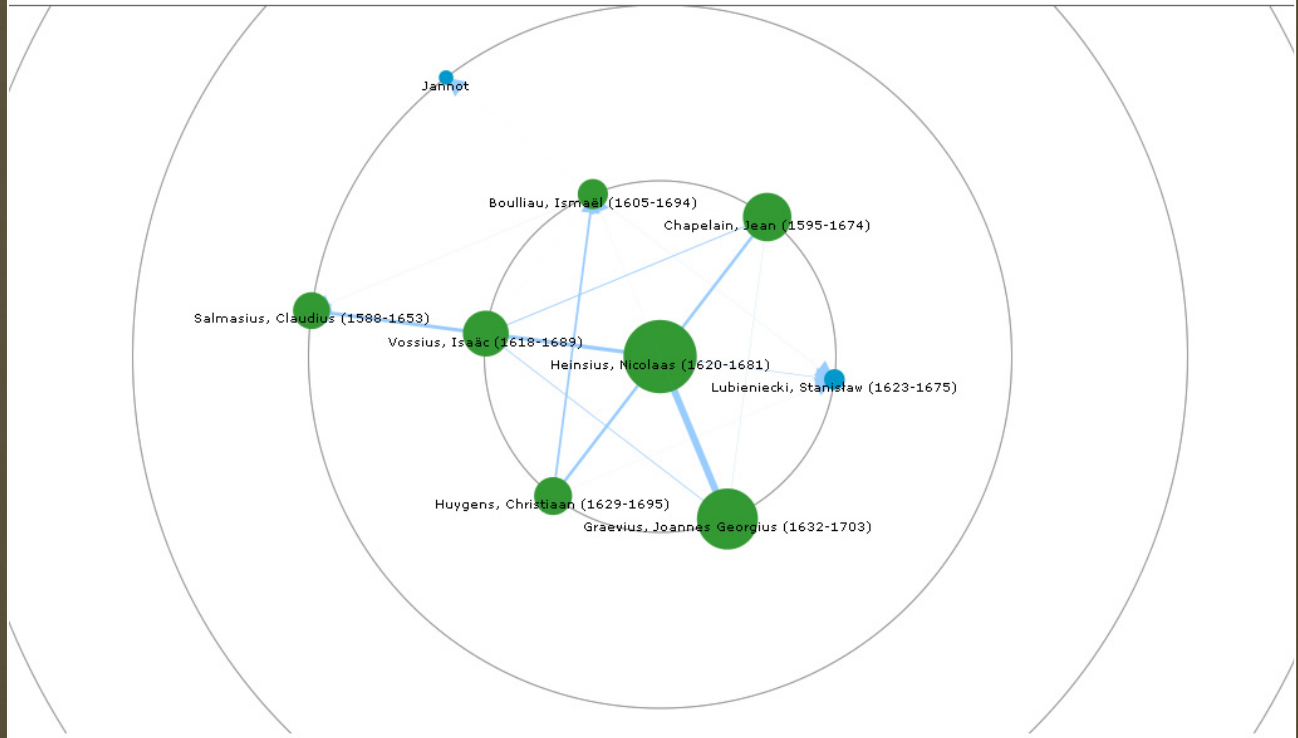
Letters | Network | Topics | Special visualisations

CKCC

Social network based on: Heisius, Nicolaas

See also in network: Salmasius, Claudius

Letter subset: C. Huygens to Leeuwenhoek



Topic Modeling

A topic is a probability distribution over words

Used to find

similar words

similar documents

documents similar to arbitrary text

Three methods in combination with language technologies

1. Latent Direchlet Allocation (LDA)
2. Latent Semantic Analysis (LSA)
3. Random Indexing

Peter Wittek/ Walter Ravenek. "Supporting the Exploration of a Corpus of 17th Century Scholarly Correspondences by Topic Modelling," Supporting Digital Humanities Conference, 17-18 November 2011 Copenhagen

Selector Reset

Text

luth luths Terms Search

No suggested terms...

Years

From: 1618 To: 1679 Unknown

Senders

All None

- Barre, Joseph de la, sr. (fl. 1648-1653) 1
- Bertet, Théodore (-1709) 1
- Bielke, Hinric (fl. 1644-1667) 2
- Burgh, Jacob van der (1599-1659) 3
- Chièze, Sebastien (1625-1679) 2
- Gaultier, Jacques(fl. 1617–1652) 4
- Huygens, Christiaan (1629-1695) 4
- Huygens, Constantijn (1596-1687) 19

Recipients

All None

- Barre, Anne de la (fl. 1648) 1
- Barre, Joseph de la, jr. (fl. 1654) 1
- Barre, Joseph de la, sr. (fl. 1648-1653) 1
- Beringhen, Henri de (1603-1692) 2
- Charas, Moise (1618-1698) 1
- Chièze, Sebastien (1625-1679) 3
- Cusance, Béatrice de (1614-1663) 3

http://localhost:8182/#

huygens/3928

Back to results Previous letter / Next letter

Metadata

1645

Gaultier, Jacques(fl. 1617–1652)
?

French

Huygens, Constantijn (1596-1687)
?

Text

..... Vous m'avez faict l'honneur de me commander de vous trouver un luth de Boulogne a neuf cottes. Je vous diray, Monsieur, que tous les luths de Bologne a neuf cottes sont de Laux Maller qui est mort il y a six vingt ans, et sont tous la pluspart de moiyenne taille et non propre pour chanter. Et croy qu'il n'y en a en tout le monde cinquante; pour icy je suis certain qu'il n'y en a pas six. Il y a eu deux autre maistres, l'un se nomme Anstrin et l'autre Sigismon Maller, qui on fait un peu apres plus grande cantité de luths et de grande taille, mais ilz ont honze cottes. Je me suis mis desja en devoir et fais recherche tant qu'il m'est possible. Il y a bien icy d'autre sorte de luths qui pourront servir et peut estre ausy bien rayusir pour le chanter, ausy bien qu'un luth de Bologne. Il en [est] de Sconuel (?) qui sont grands. S'il vous plaist me mander de quelle façon le manche soit, et sy vous voulez qu'il soit pour jouer des piesces et ausi pour chanter, je le feray acommoder .....

Analysis

Generated keywords: cottes luths chanter maller ausy bologne taille luth

Back to top

Topic-Author-Time

----------TopicKey

68 0.02294 synode gevoelen confessie gereformeerde predicanten lutherschen charenton dordrecht doctor poincten kerken schriften protestanten kercken breedt

79 0.02944 mogende heeren edele edelen zeelant saecke excellentie raden hollandt requeste staten hove waterbailliu tselve eynde verstaen principale lande raiden

82 0.02181 sweedsche generael heere deensche excellentie torstenson hamburg armee major keysersche mijnheere wegen sterck munster coningsmarck keyserlijcke

-	1605	1610	1615	1620	1625	1630	1635	1640	1645
68	1,02E+00	0.003279528	2,72E-01	0.002337202	0.002646312	0.003273037	0.006456783	0.004623124	0.005288161
79	0.116005472	0.124600115	0.067178413	0.056681704	0.06124088	0.059566158	0.052052532	0.044734561	0.043474241
82	9,65E-01	2,73E-01	2,59E-01	7,12E-02	3,74E-01	6,86E-02	1,50E-02	0.000968506	0.013411816

----------Topic 68 - religion

raphelengius_franciscus	0.055183695905
overall_john	0.053229471789
massis_g_of_eg	0.035321473298
hooft_pc	0.028161745133
pijnacker_cornelis_van	0.022090241378
bye_de_oude_nicolaas_de	0.017648170644
hooft_henrik	0.017526136684
reael_laurens	0.014726627250
muller_georg	0.013950523247

----------Topic 79 - nobility

boreel_johan	0.214461099051
bie_joris_de	0.182443436507
brasser_joost	0.151460997760
hessen-darmstadt_(landgraaf)_georg_ii	0.145934248583
bredero_ga	0.116617701630
blon_michel_le	0.099859869863
pijnacker_cornelis_van	0.099249709885
stadsbestuur_van_rotterdam	0.095804542183
colterman_david	0.093664067835

----------Topic 82 - scandinavian conflict

wael_van_moersbergen_adolf_de	0.031475737940
vondel_joost_van_den	0.020527103941
salm-kyrburg_johann_kasimir_von	0.006850411824
uyttenbogaert_johannes	0.003532663098
pommeren_en_croy_anna_van	0.003013425235
gecommitteerde_raden_ter_admiraliteit_te_amsterdam	0.002985989822
reael_laurens	0.002358086834
brasser_govert	0.002169135759
muys_van_holy_arent	0.001980522238

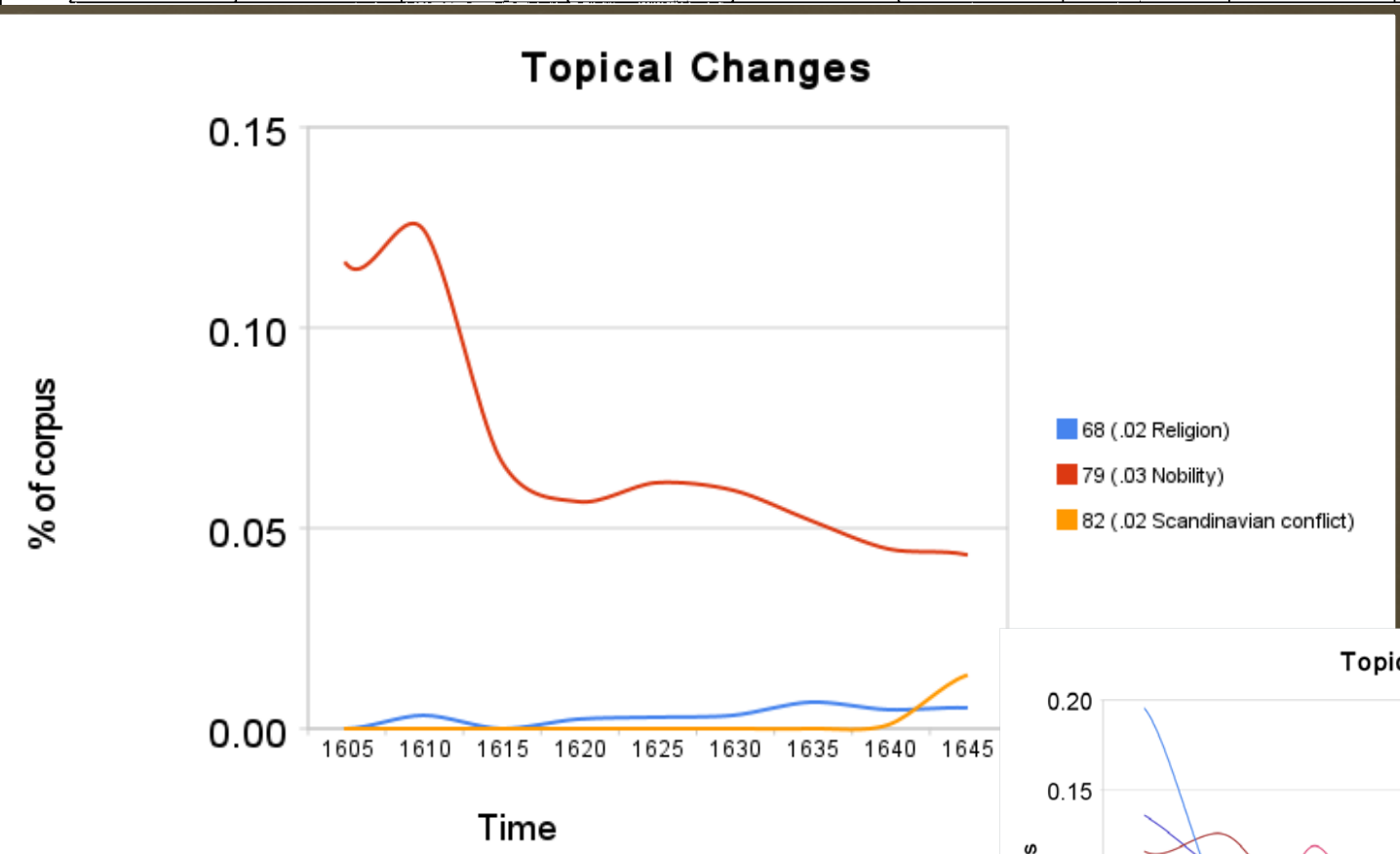
----------TopicKey

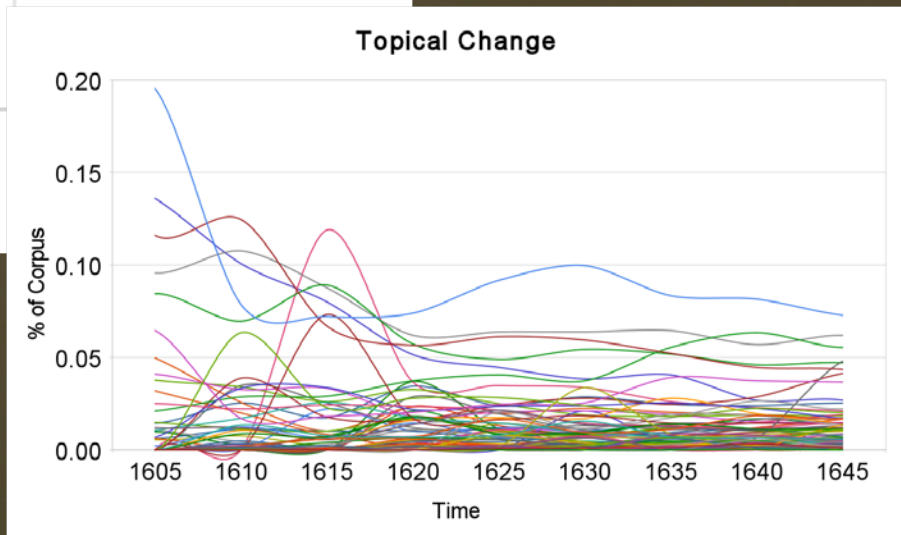
68 0.02294 synode gevoelen confessie gereformeerde predicanten lutherschen charenton dordrecht doctor poincten kerken schriften protestanten kercken breedt

79 0.02944 mogende heeren edele edelen zeelant saecke excellentie raden hollandt requeste staten hove waterbailliu tselve eynde verstaen principale lande raiden

82 0.02181 sweedsche generael heere deensche excellentie torstenson hamburg armee major keysersche mijnheere wegen sterck munster coningsmarck keyserlijcke

-	1605	1610	1615	1620	1625	1630	1635	1640	1645
68	1,02E+00	0.003279528	2,72E-01	0.002337202	0.002646312	0.003273037	0.006456783	0.004623124	0.005288161
79	0.116005472	0.124600115	0.067178413	0.056681704	0.06124088	0.059566158	0.052052532	0.044734561	0.043474241
82	9,65E-01	2,73E-01	2,59E-01	7,12E-02	3,74E-01	6,86E-02	1,50E-02	0.000968506	0.013411816

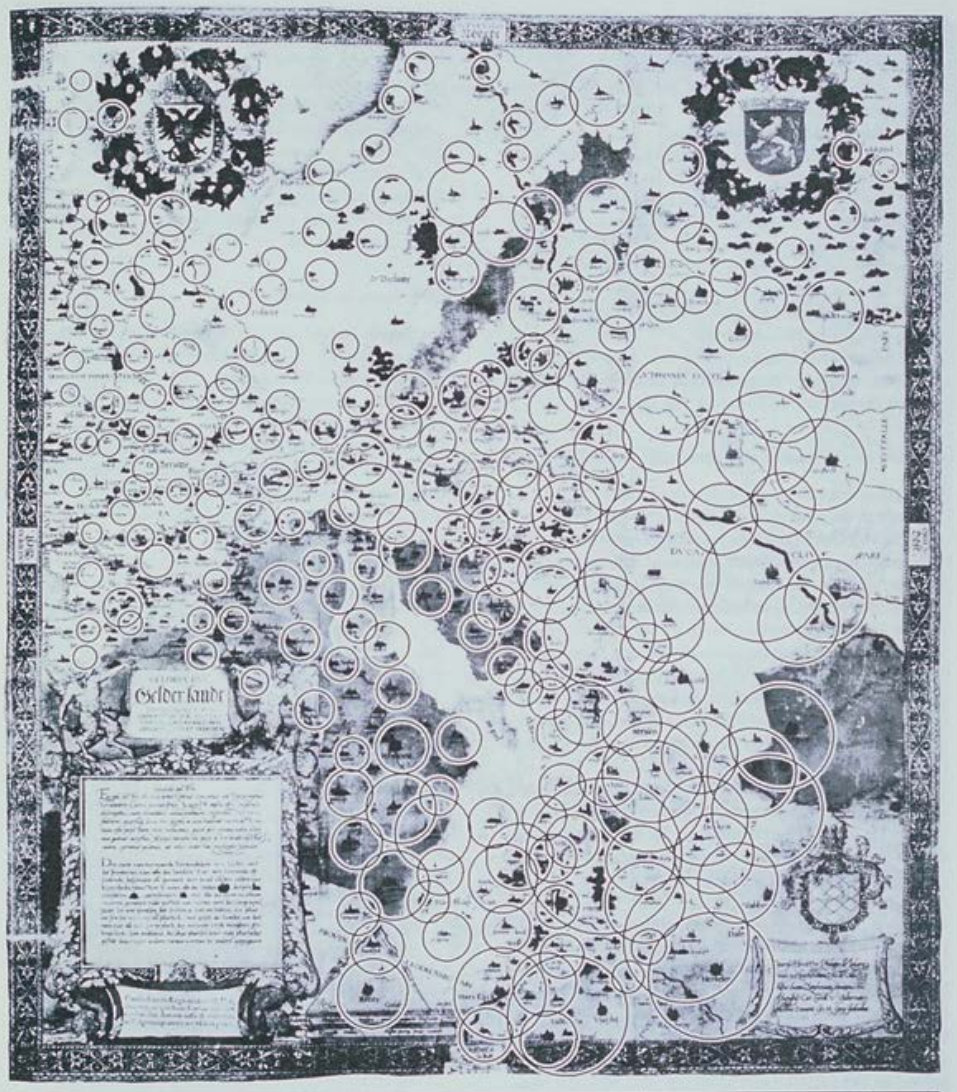




Source: Scott Weingart UIA

Data + Uncertainty

- Uncertainty in data
- Uncertainty in data and manipulation
- Uncertainty in interfaces in and between data sets



NAUWKEURIGHEIDSANALYSE

Standaardonnauwkeurigheidscirkel

$r(cm) = d_i \times f \times 5.10^4 \times 1cm$

r: cirkelstraal

d_i: puntonnauwkeurigheidswaarde

f: numerieke waarde voor de gebruikte kaartschaal

ACCURACY ANALYSIS

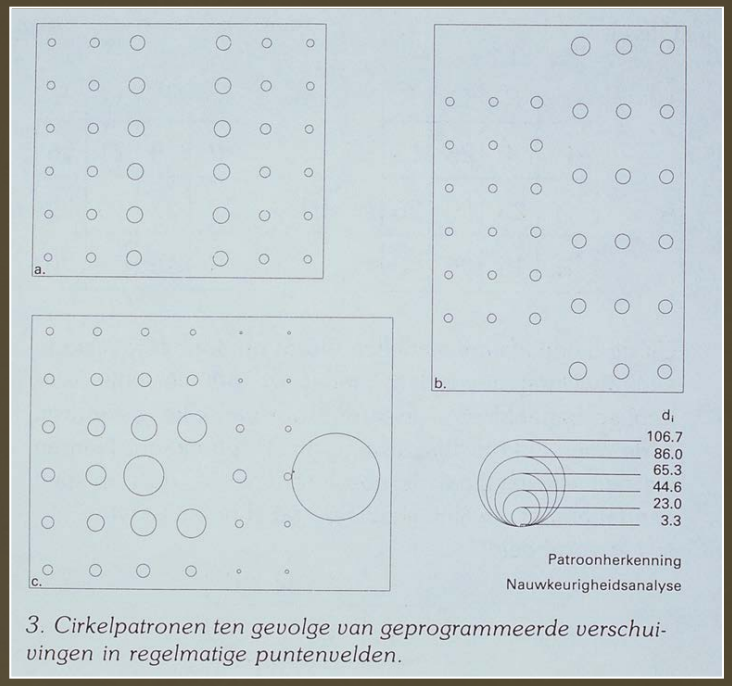
Standard inaccuracy circle

$r(cm) = d_i \times f \times 5.10^4 \times 1cm$

r: circle radius

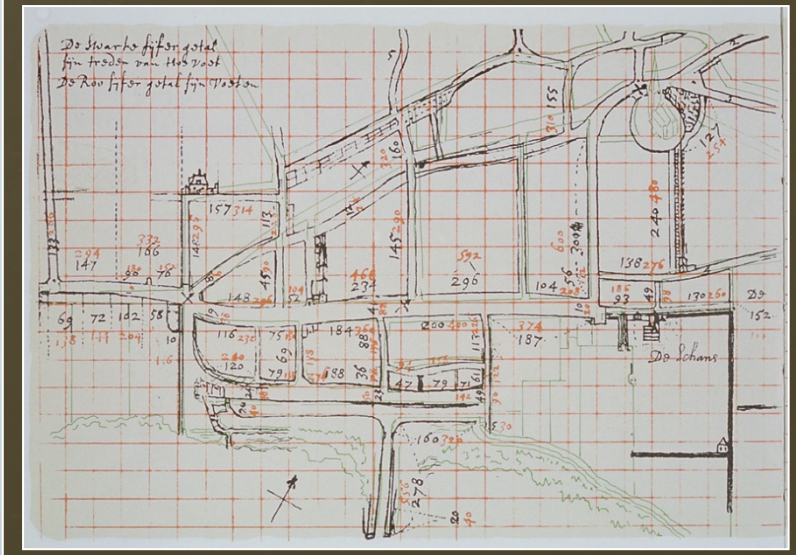
d_i: point inaccuracy value

f: numerical value for the used mapscale



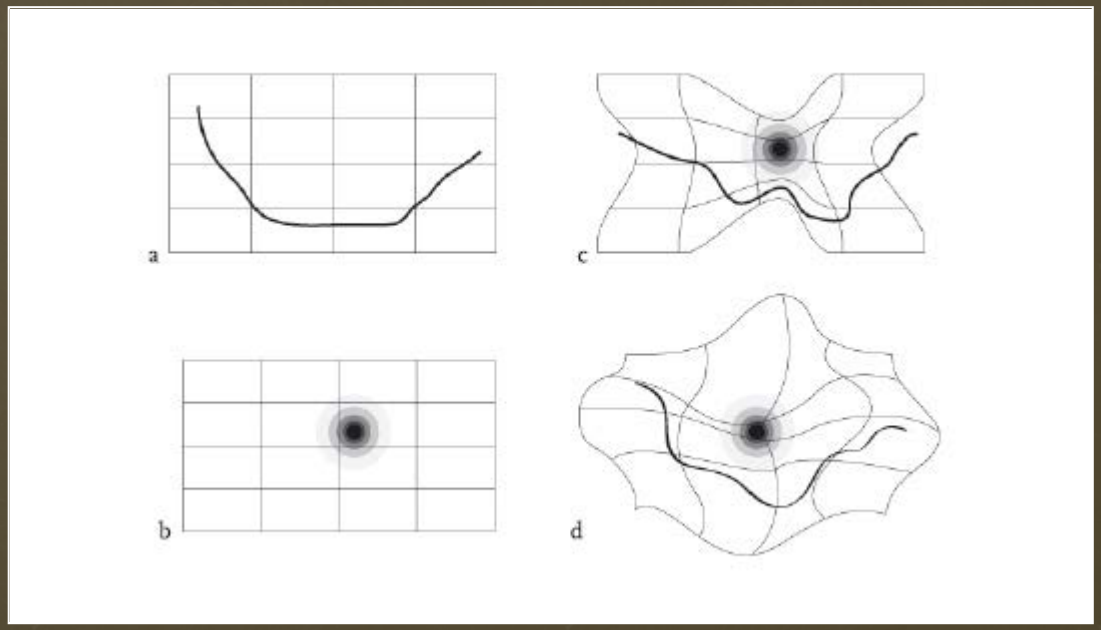
3. Cirkelpatronen ten gevolge van geprogrammeerde verschuivingen in regelmatige puntenvelden.

Source: P. Mekenkamp/O. Koop, "Nauwkeurigheidsanalyse van oude kaarten met behulp van de computer", *Caert-Thresoor*, 1986, nr. 3, pp. 45-52



Uncertainty in data and manipulation

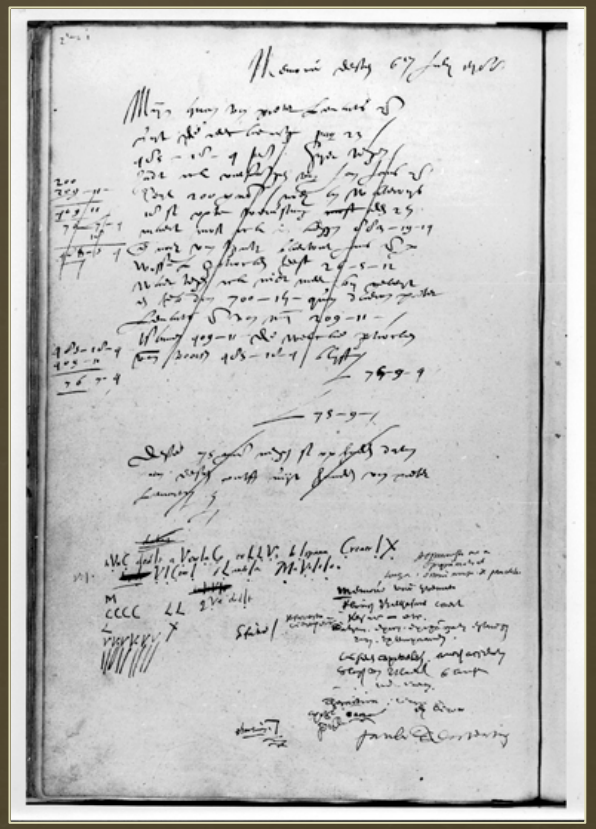
Data as Capta and Interpretative Visualizations to enhance experience and story telling



Johanna Drucker Digital Humanities Quartly 2011 Winter

Uncertainty in interfaces in and between data sets

Metadata	1618-07-14 Huygens, Constantijn (1596-1687) London (GB)	Dutch Huygens, Maurits (1595-1642) [Den Haag (NL)]
Text	Ik schrijf u maar niet dikwijls, want gij hoort toch alles uit de brieven aan onze ouders. Gij moet mij wat meer van dat duel vertellen. Ik kan nog niet goed eene laars aantrekken en ben dus nog niet naar Cambridge en Oxford geweest. Den 20sten van deze maand gaan wij naar Bagshot, waar de Koning ook zal zijn. 'Hier je menay de Gheyn veoir la galerie a peintures du feu prince par l'adresse de certains gentilshommes anglois qui m'ont grandement obligé. Celle du conte d'Arondel avec les antiquitez etc. de Gheyn m'avoit fait veoir par cy devant. Choses admirables en verité. Au reste nous voicy au païs le plus gentil du monde ou journellement nous voyons des palais et païsages *om op clavecingel-dexels te schilderen*'. Mijn knecht bevalt mij erg goed; wij spreken samen niets dan Engelsch. Ce 4^{e} de Juillet 1618. De Lambeth.	

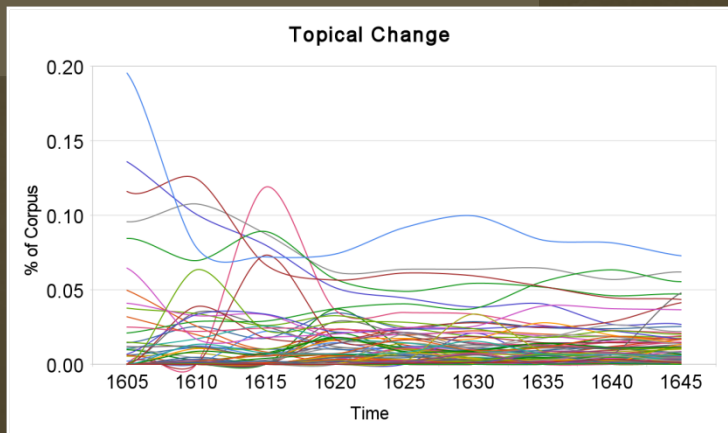


Data + Contextuality

Once you create a set of topics, plotting their frequencies is simple enough. But plotting the aggregate frequency of a group of words isn't the same thing as "discovering a trend," unless the individual words in the group actually correlate *with each other* over time. And it's not self-evident that they will.

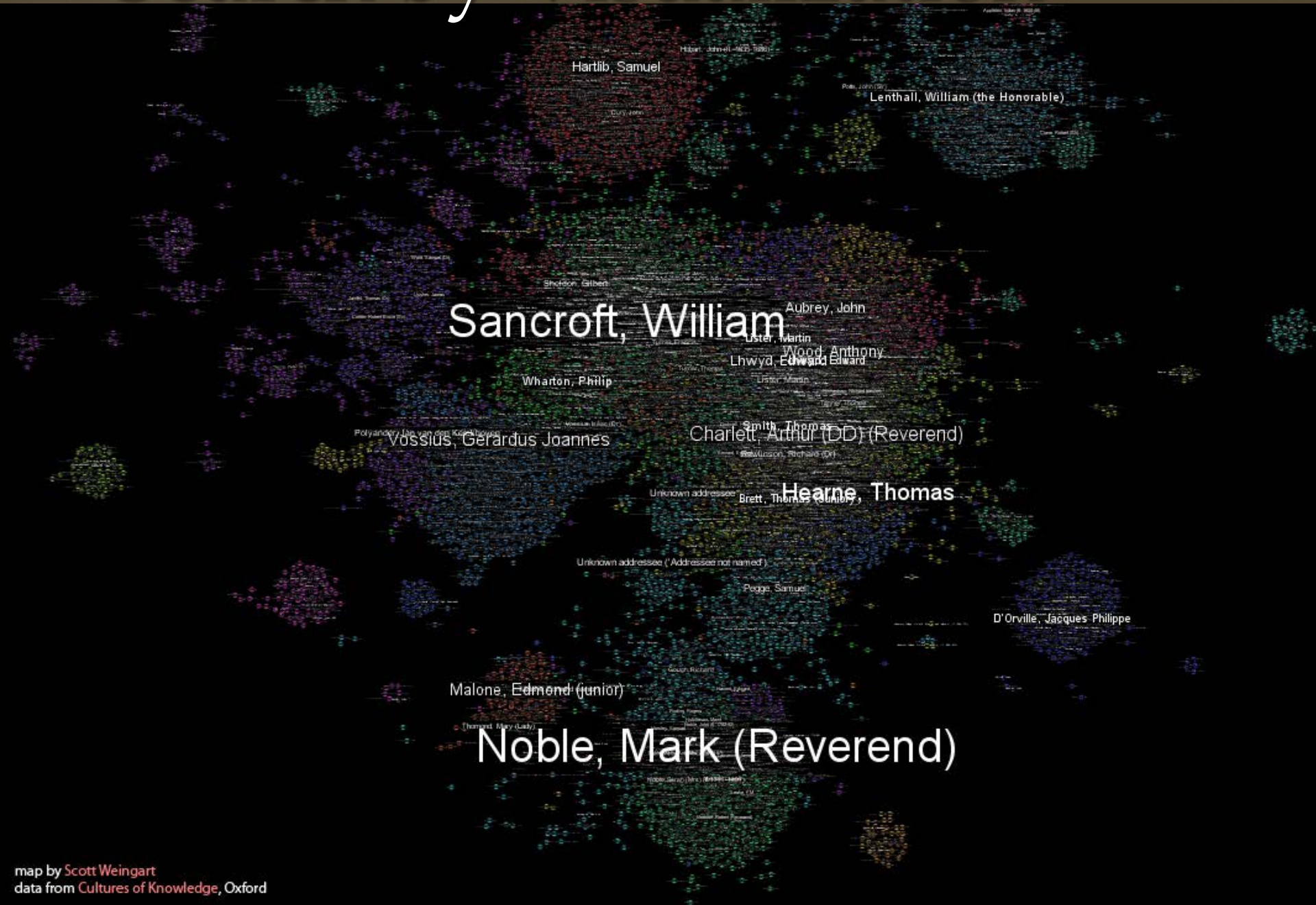
http://tedunderwood.wordpress.com/2011/09/16/topics-are-also-trends

Blog The Stone and The Shell
Ted Underwood: UIUC



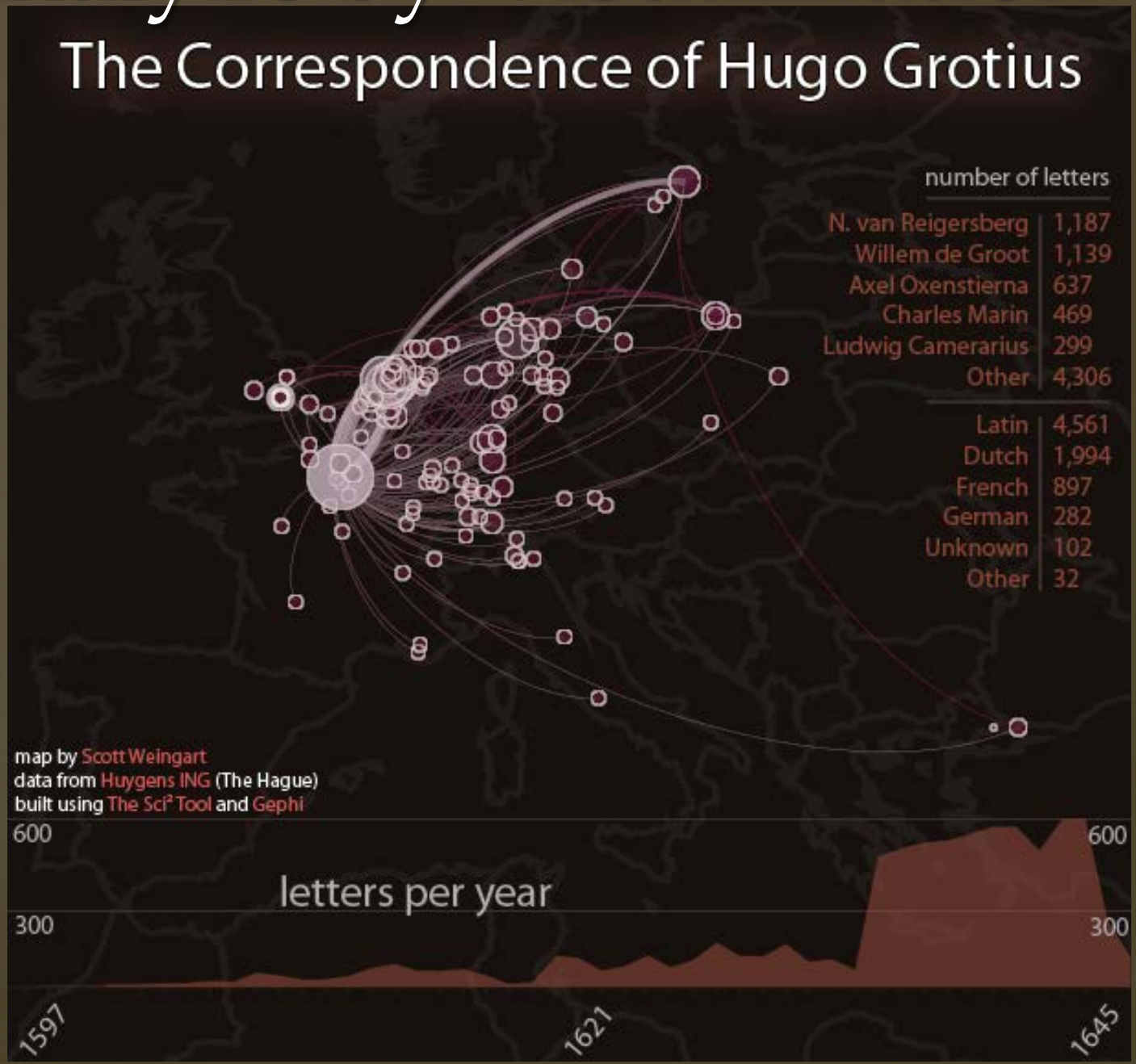
Desiderata

Search by Visualization

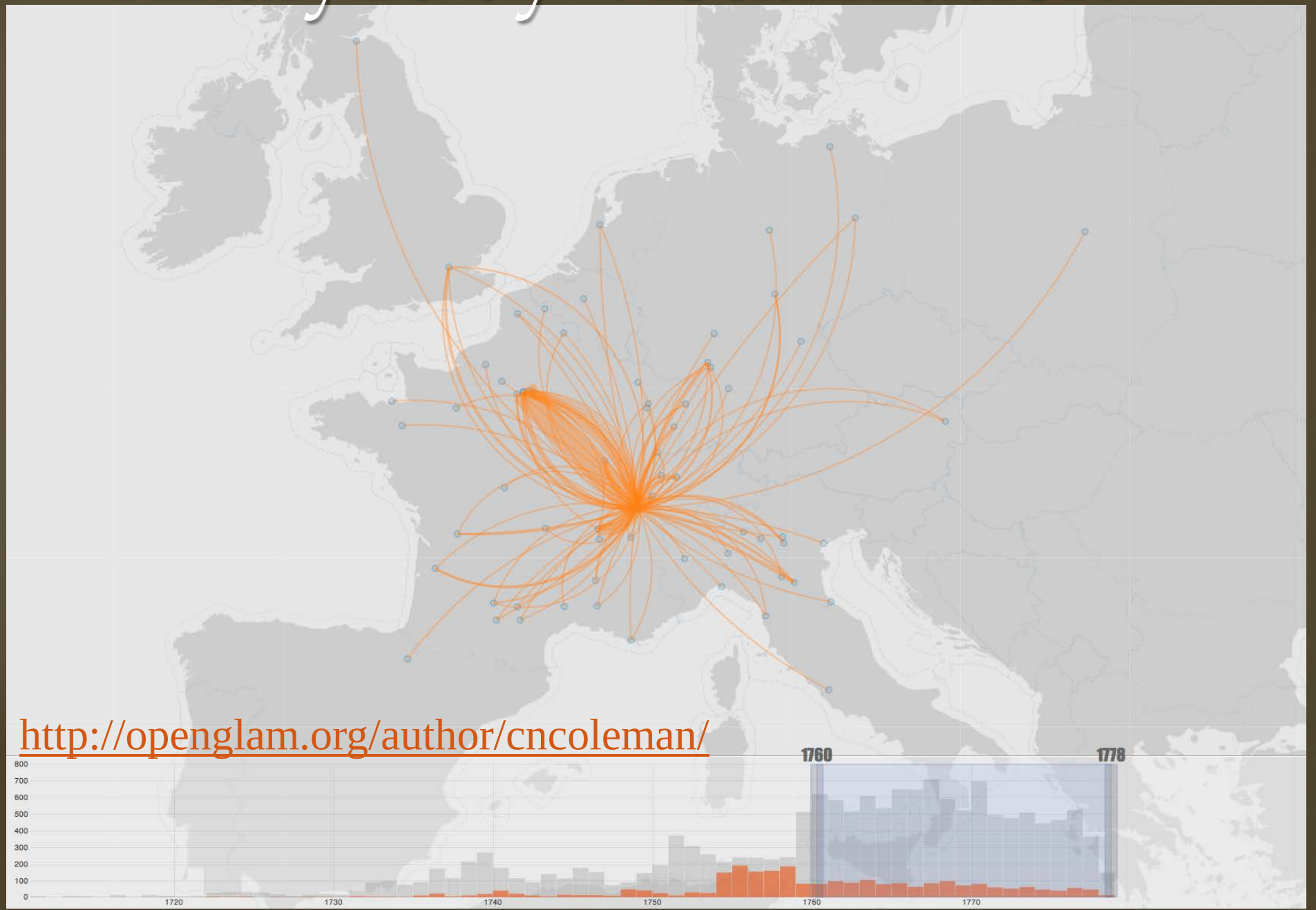


map by Scott Weingart
data from Cultures of Knowledge, Oxford

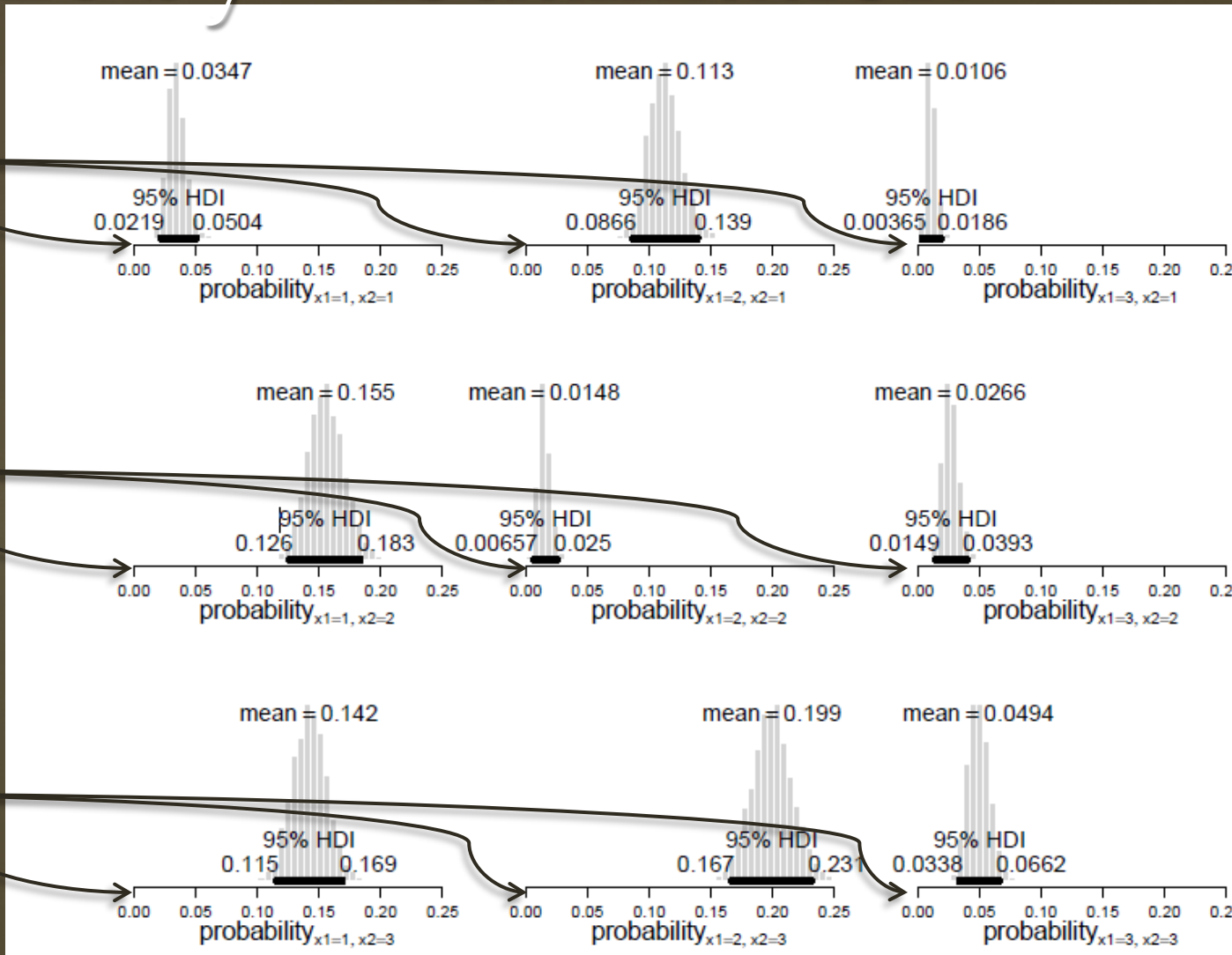
Analyze by Visualization



Analyze by Visualization

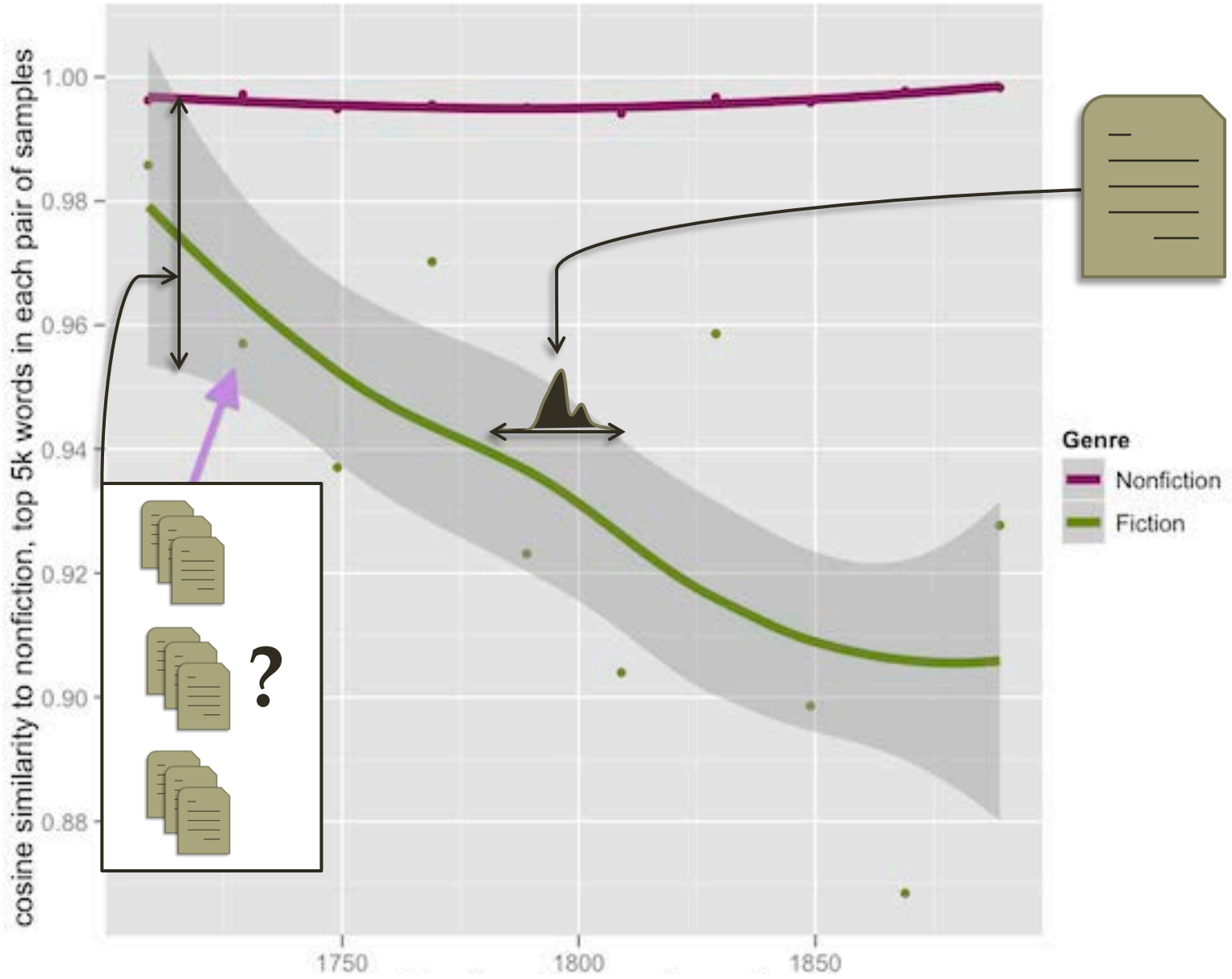


Analyze by Visualization



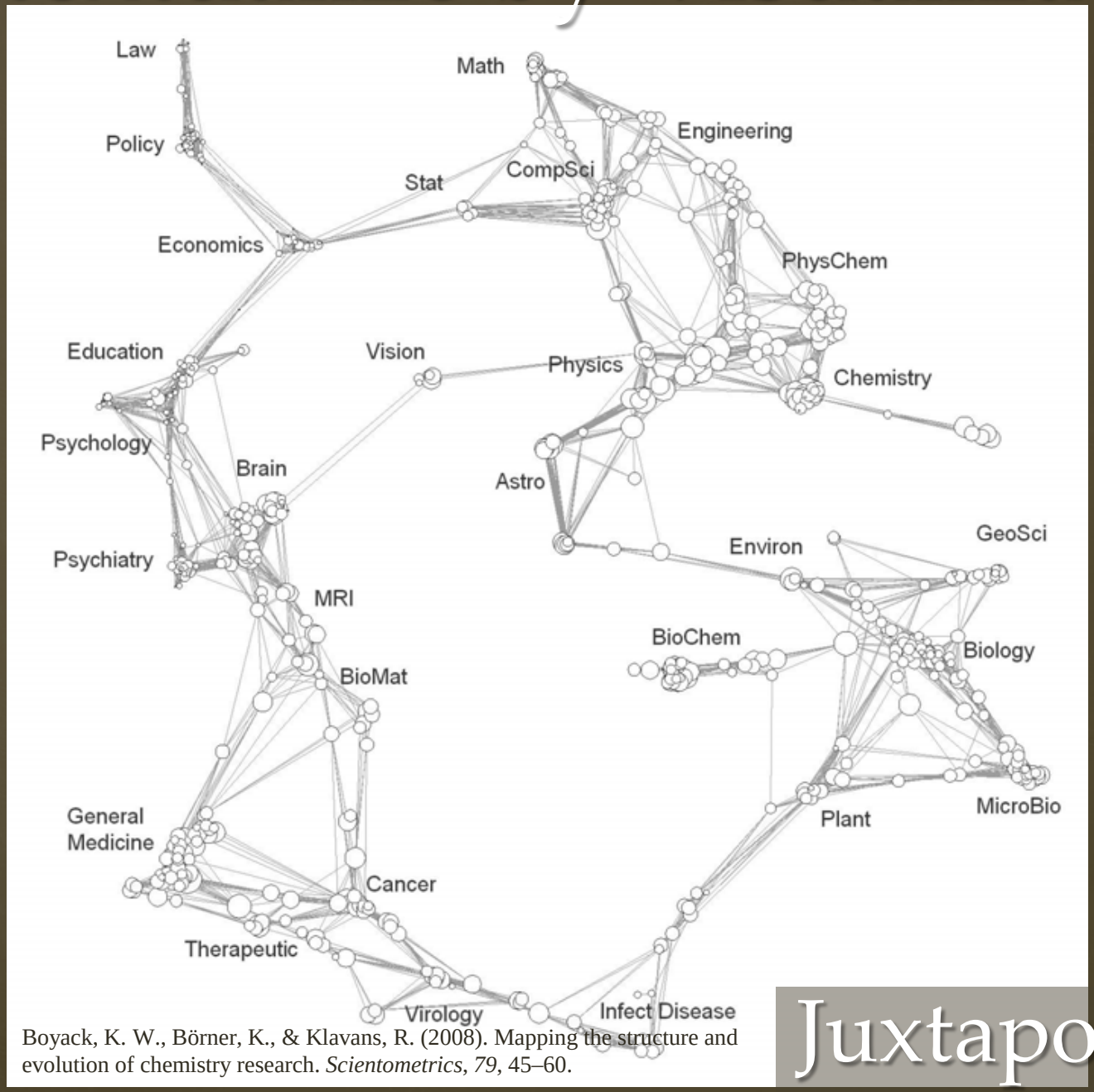
http://www.indiana.edu/~kruschke/DoingBayesianDat

Analyze by Visualization



http://tedunderwood.wordpress.com/2012/03/02/literary-and-nonliterary-diction-the-sequel/

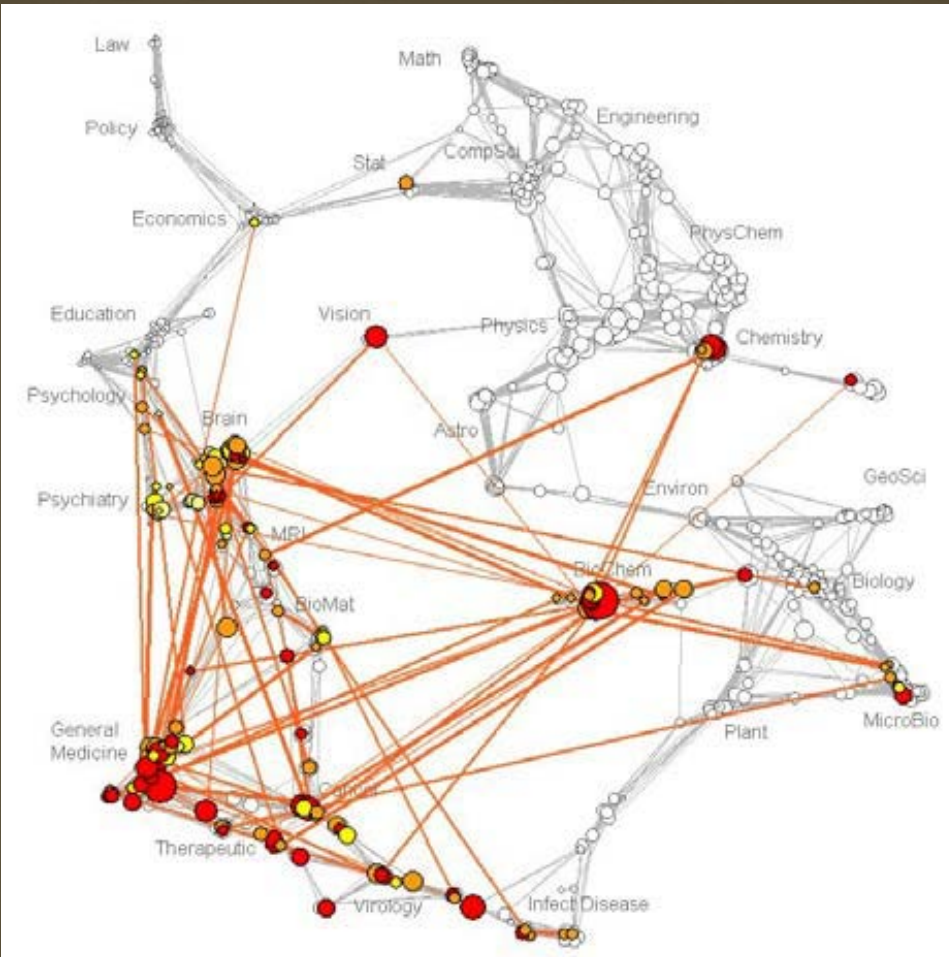
Contextualize by Visualization

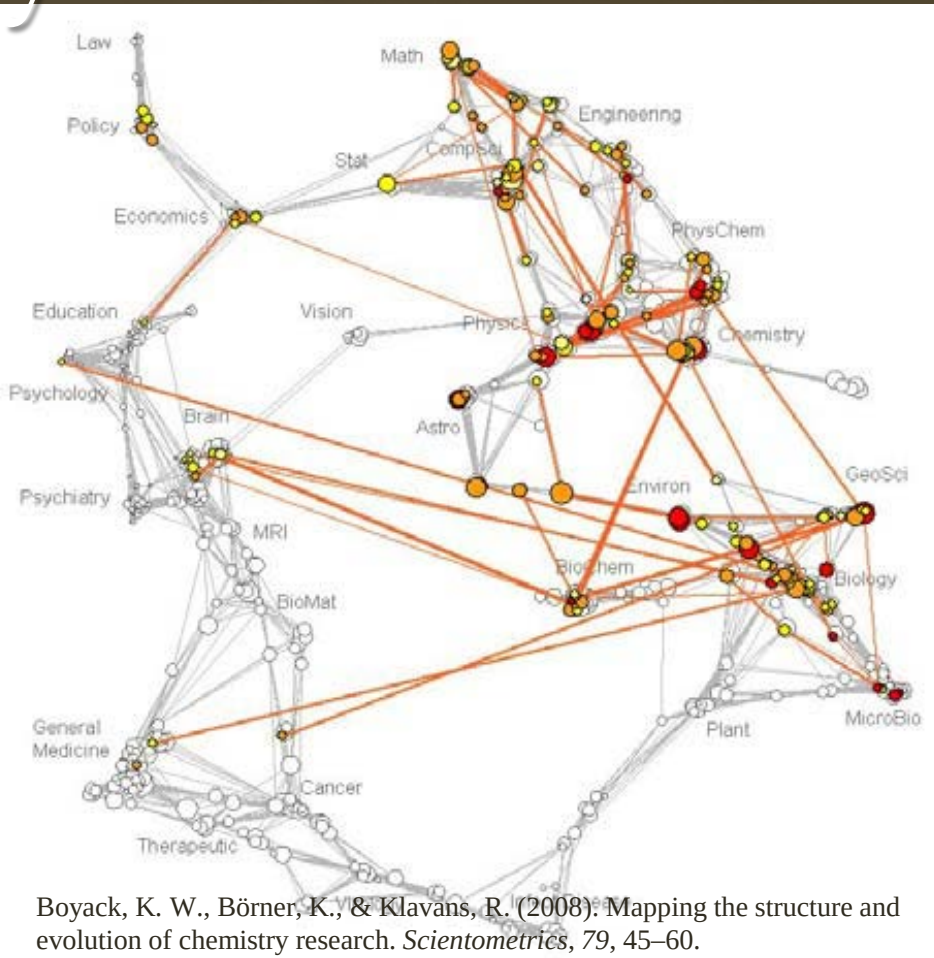


Boyack, K. W., Börner, K., & Klavans, R. (2008). Mapping the structure and evolution of chemistry research. *Scientometrics, 79,* 45–60.

Juxtaposition

Contextualize by Visualization

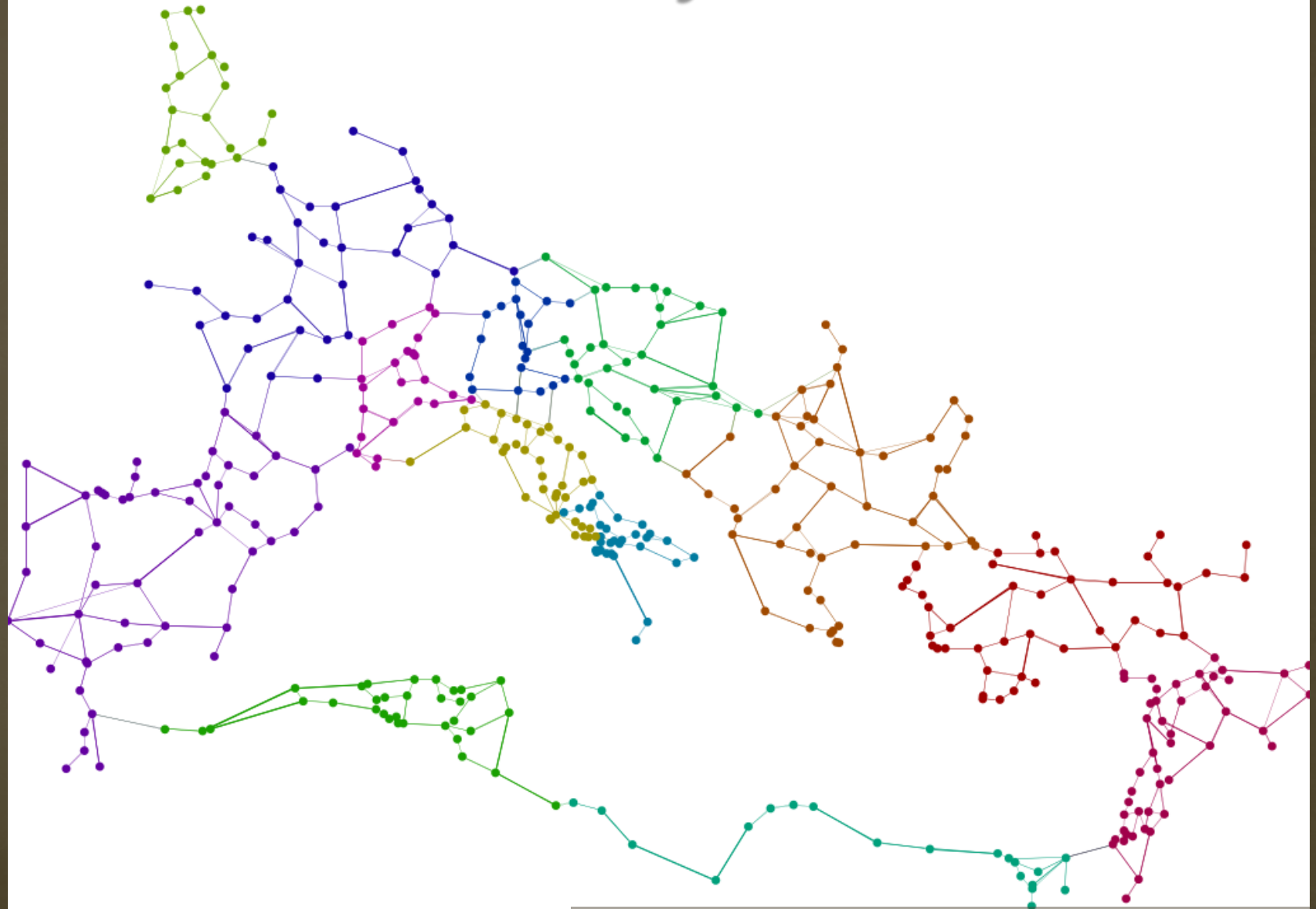




Boyack, K. W., Börner, K., & Klavans, R. (2008). Mapping the structure and evolution of chemistry research. *Scientometrics*, *79*, 45–60.

Juxtaposition

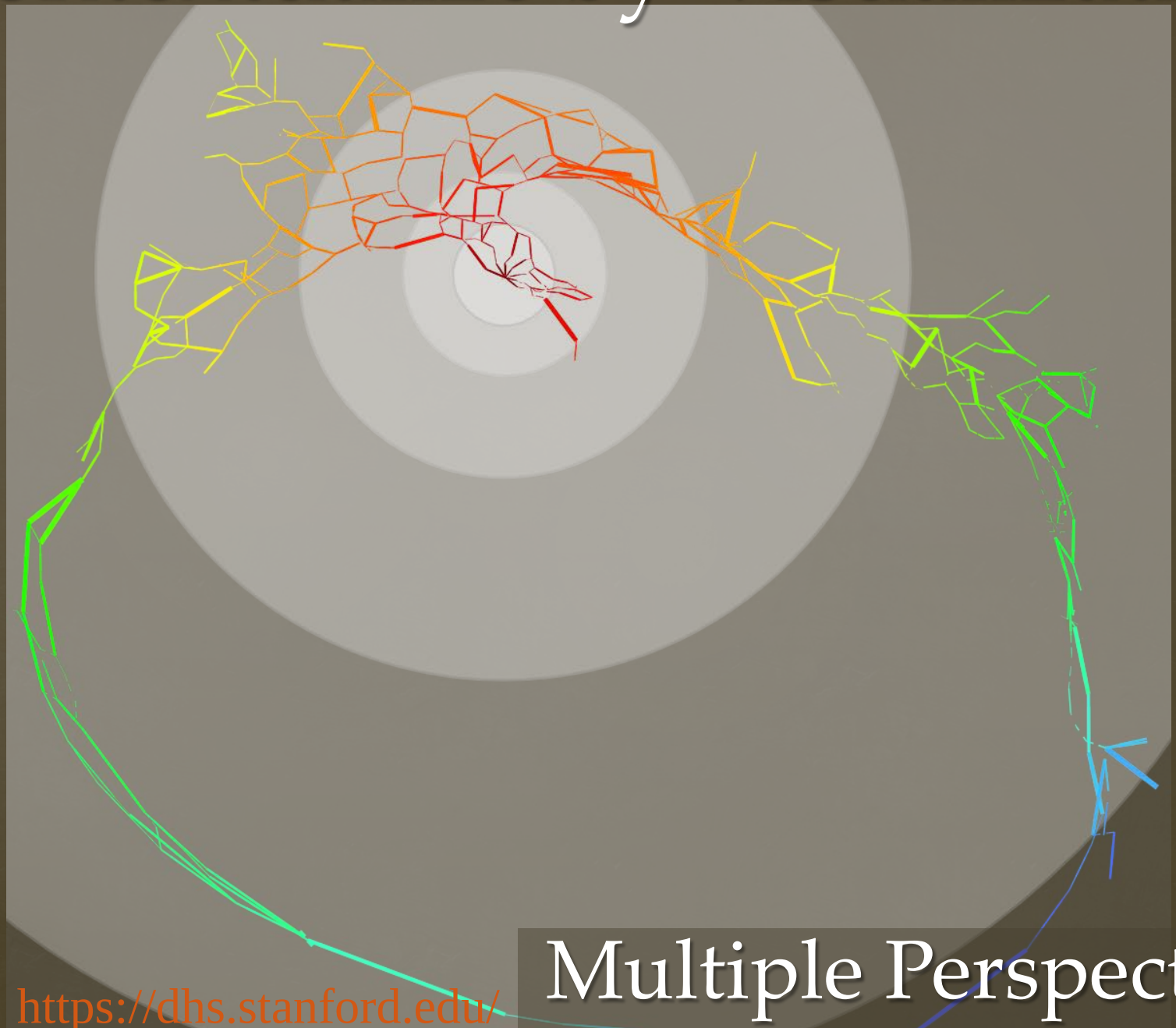
Contextualize by Visualization



https://dhs.stanford.edu/

Multiple Perspectives

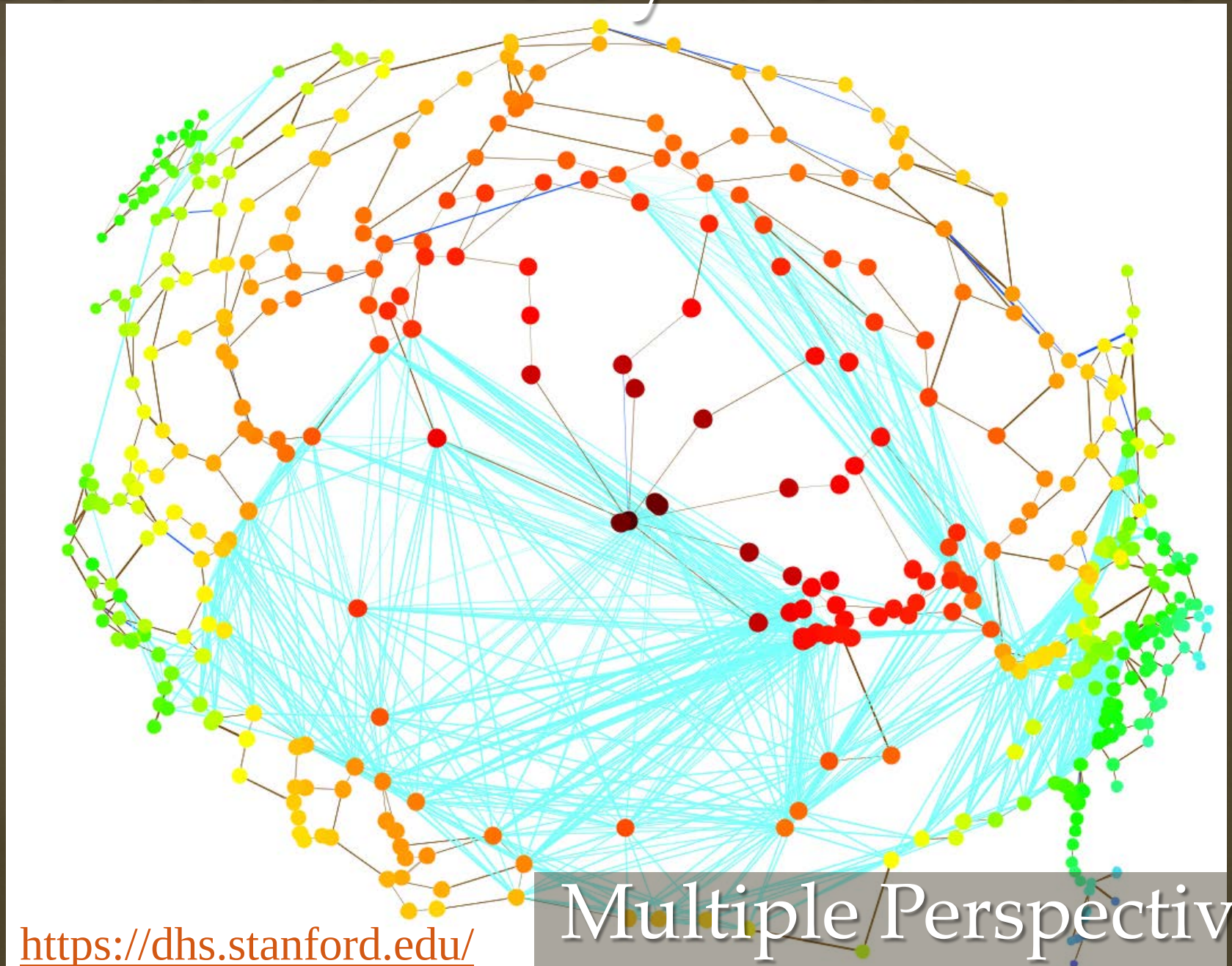
Contextualize by Visualization



https://dhs.stanford.edu/

Multiple Perspectives

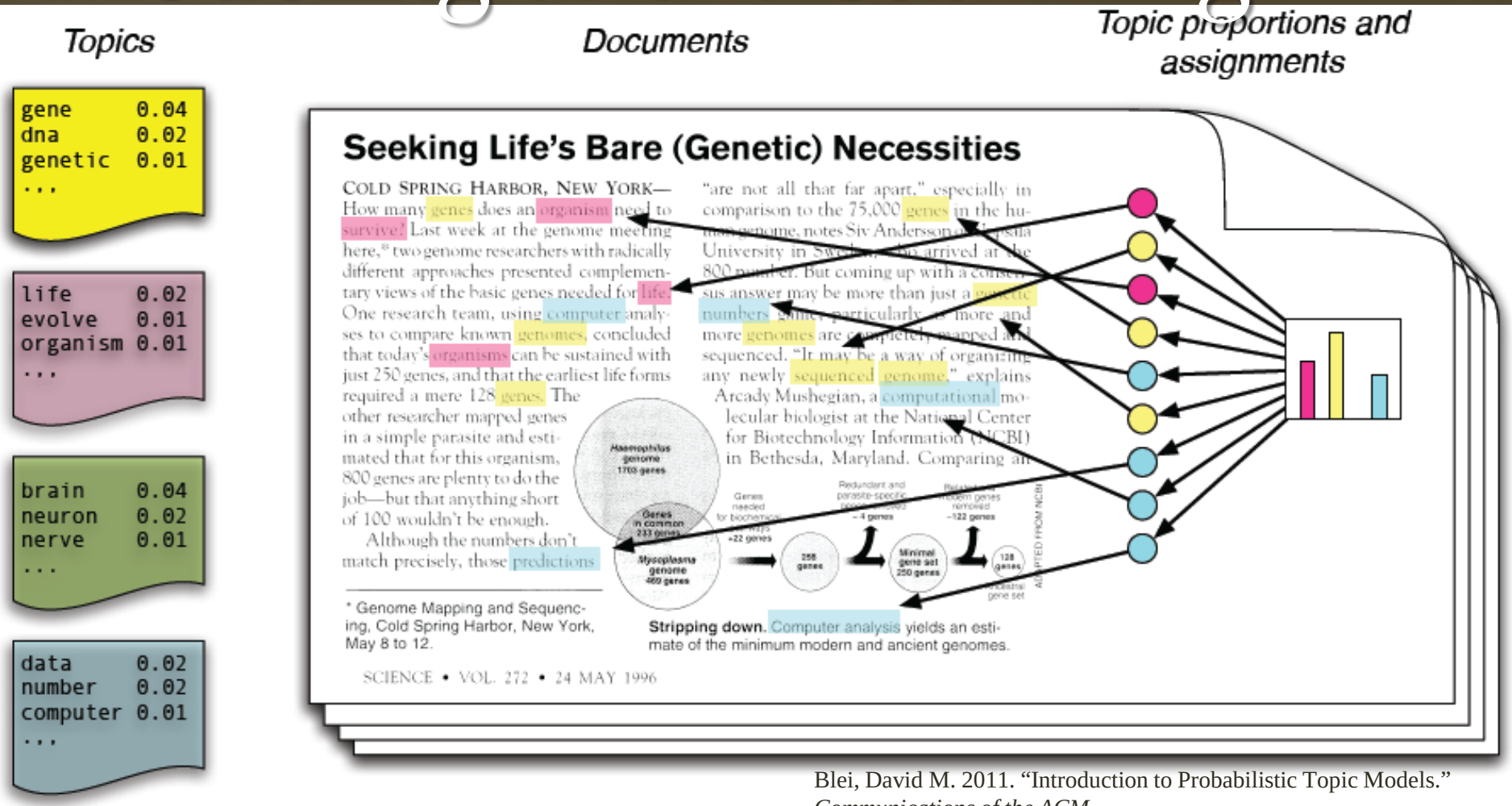
Contextualize by Visualization



https://dhs.stanford.edu/

Multiple Perspectives

Modeling and Visualizing



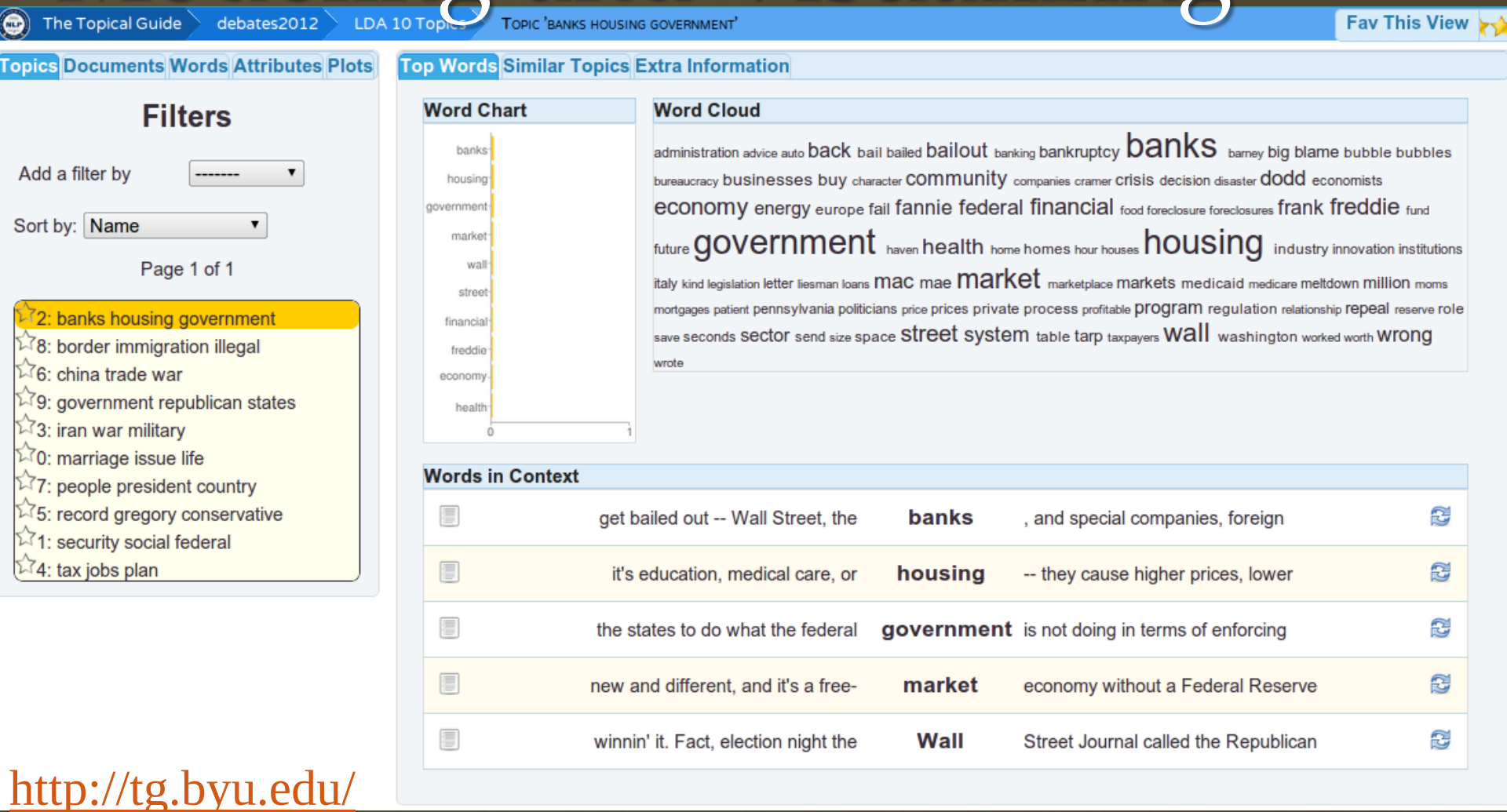
Blei, David M. 2011. "Introduction to Probabilistic Topic Models." *Communications of the ACM*.

Modeling and Visualizing

1880	1890	1900	1910	1920
fish	coast	islands	islands	museum
sea	sea	fish	fish	expedition
coast	fish	expedition	expedition	birds
island	island	island	birds	islands
expedition	islands	coast	animals	fish
energy	molecules	energy	energy	atom
molecules	energy	molecules	theory	atoms
atoms	atoms	atoms	atoms	energy
molecular	molecular	matter	atom	electrons
matter	matter	atomic	molecules	electron
action	action	solution	solution	solution
liquid	substance	action	solutions	concentration
substance	solution	solutions	substances	solutions
substances	substances	substance	water	ion
power	liquid	substances	action	water
government	government	university	university	public
scientific	national	institution	president	university
public	institution	government	professors	institutions
		onal	universities	government

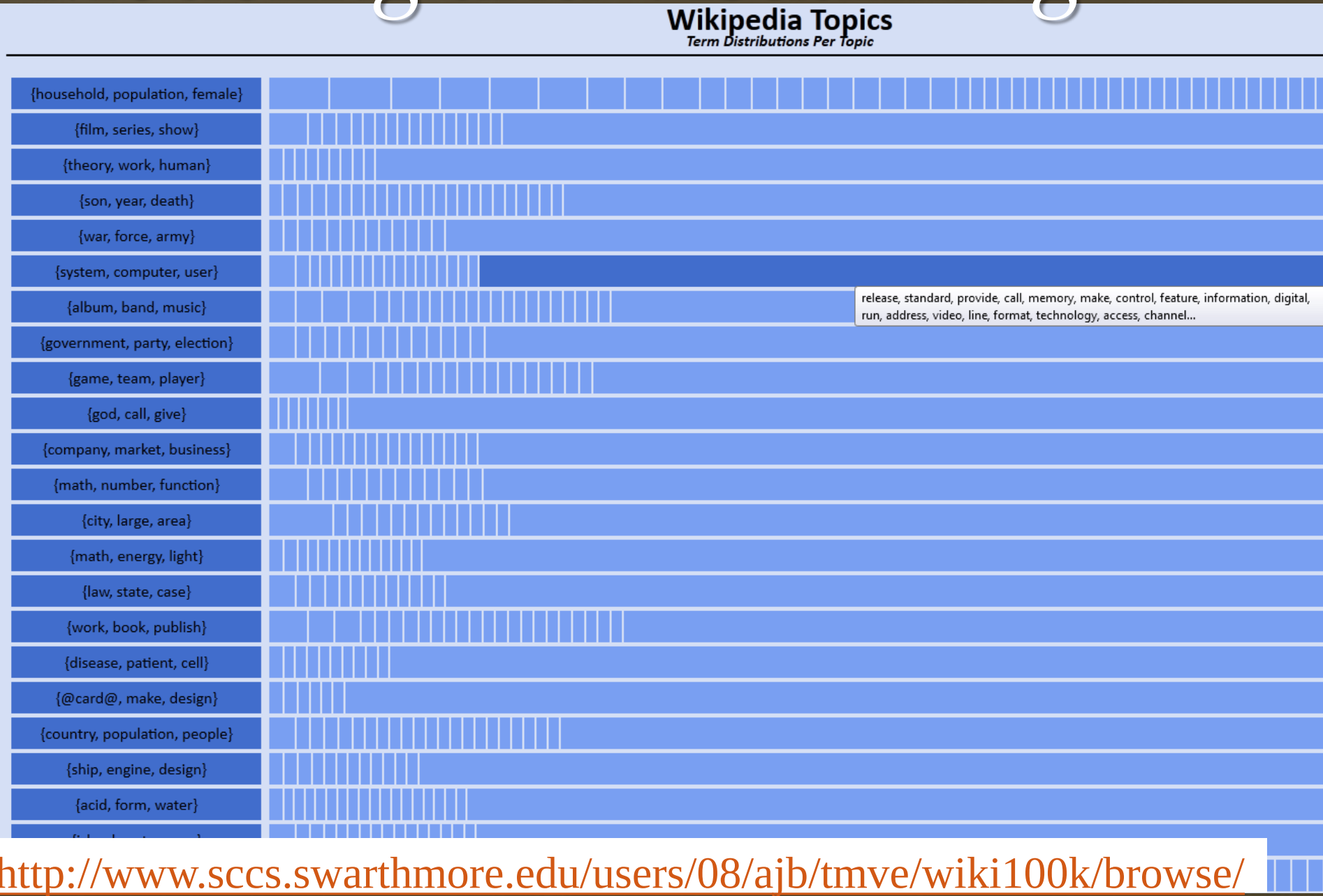
http://topics.cs.princeton.edu/Science/

Modeling and Visualizing



http://tg.byu.edu/

Modeling and Visualizing



http://www.sccs.swarthmore.edu/users/08/ajb/tmve/wiki100k/browse/

Desiderata

Search by visualization (navigating areas rather hopping from point to point)

Analyze by visualization (not seeing/seeing wrong)

Contextualize by visualization

juxtaposing

multiple points of view in visualization(annotations) etc. etc.

Modeling and Visualizing

Thank You

charles.van.den.heuvel@huygens.knaw.nl
scbweing@indiana.edu