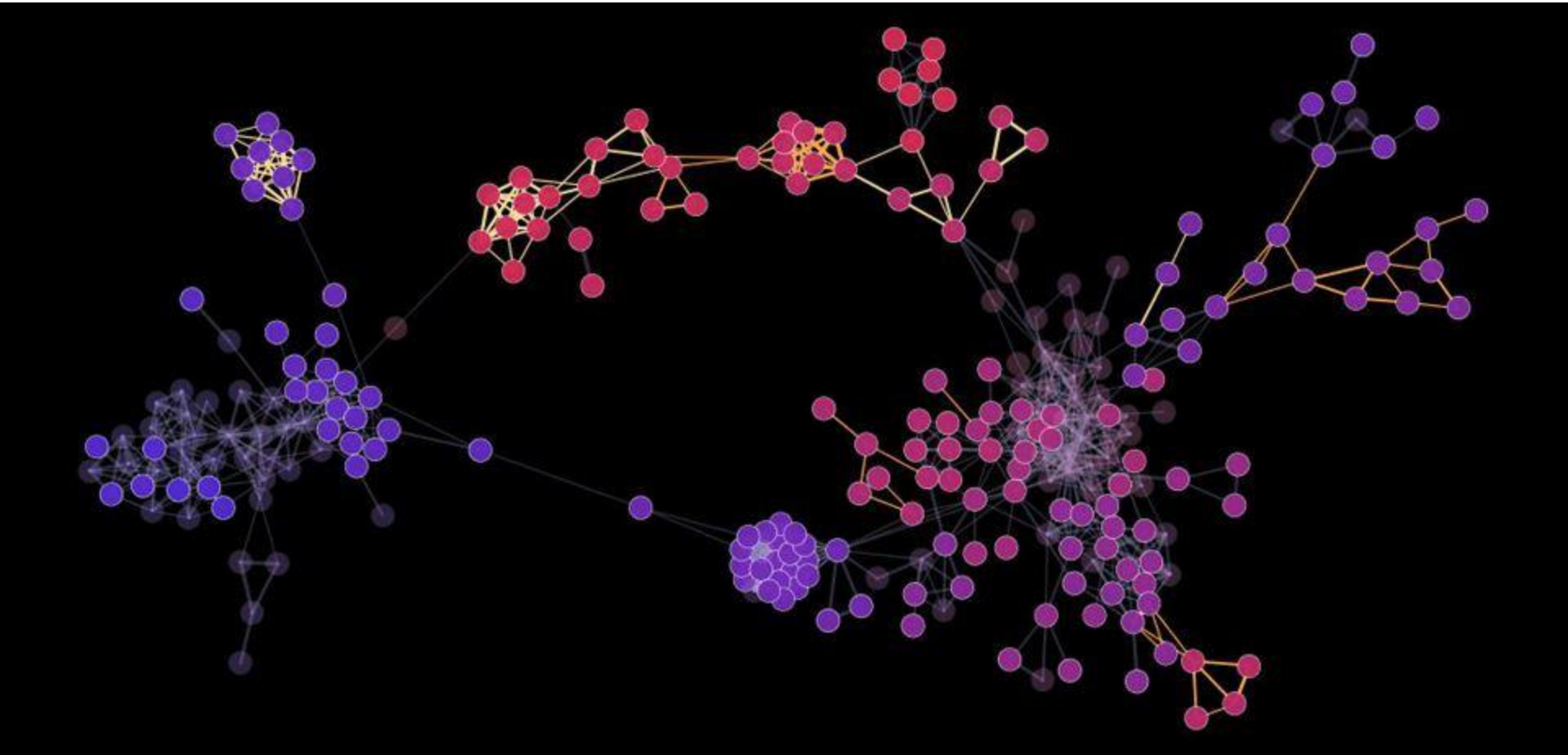
Tutorial

4

CiteSpace: Visualizing and Analyzing the Structure and Dynamics of Scientific Fields

Chaomei Chen
College of Information Science and Technology
Drexel University
Email: chaomei.chen@drexel.edu



How to Evaluate the Results

Example: A Superset of Citers to the Three Papers on CiteSpace

Clust...	Size	Silho...	mean...	Top Terms (tf*idf weighting)	Top Terms (log-likelihood ratio, p-lev...	Terms (mutual information)
3	218	0.942	1997	(16.64) citation; (14.92) metric; (14.5...	scientific dom...	...mic science
2	22	0.963	2001	(10.68) need; (10.68) cardiovascular ...	need (134.93, ...	...
0	21	0.857	2002	(9.3) systems literature analysis; (6.0...	systems liter...	rare disease
5	11	0.937	2005	(7.28) systems biology; (6.81) high q...	systems biol...	machine science
1	7	0.97	1999	(4.73) biological network; (2.79) novel;	biological net...	biological network
4	3	0.974	2003	...	self-correcting map (10.93, 0.001); m...	rare disease

No Pruning

Modularity Q=0.2684

Mean Silouette = 0.9405

Clust...	Size	Silho...	mean...	Top Terms (tf*idf weighting)	Top Terms (log-likelihood ratio, p-lev...	Terms (mutual information)
3	34	0.95	1993	(10.15) comment; (8.16) critique; (6.5...	evolution (18.85, 1.0E-4); citation net...	article influence score
9	26	0.812	1997	(6.59) clustering method; (6.59) soci...	clustering method (9.29, 0.005); soci...	profiling leading scientist
17	25	0.843	2001	(15.92) need; (15.92) cardiovascular ...	need (138.78, 1.0E-4); cardiovascula...	sustainability science
13	21	0.935	2005	(11.46) hirsch index; (9.25) psychiatr...	hirsch index (32.56, 1.0E-4); psychiat...	article influence score
5	20	0.787	1997	(7.47) network science; (6.59) tunabl...	modern science (25.13, 1.0E-4); pow...	multi-modal social network
1	19	0.964	1997	(9.28) future lis; (9.28) research field;...	visualizing change (20.71, 1.0E-4); in...	bibliographic record
0	18	1	2001	(13.87) systems literature analysis; (...	systems literature analysis (126.84, ...	visualizing network
14	18	0.939	1996	(6.59) delineation; (6.59) betweenne...	nanotechnology (20.23, 1.0E-4); evol...	hiv-1 host interaction
8	17	0.966	1998	(13.22) library database; (13.22) etho...	library database (85.14, 1.0E-4); mini...	bibliometric investigation
4	15	0.995	1989	(10.15) citespace ii; (10.15) transient...	citespace ii (37.41, 1.0E-4); transient...	bibliometric investigation
7	11	0.917	2005	(10.86) systems biology; (10.15) hig...	systems biology (43.03, 1.0E-4); valu...	hiv-1 host interaction
11	11	0.864	2003	(5.32) characterizing knowledge diffu...	nanotechnology (8.66, 0.005); korea...	analyzing temporal social n...
12	10	0.904	2002	(9.28) convergence; (9.28) usable cy...	convergence (32.64, 1.0E-4); usable ...	analyzing temporal social n...
16	8	0.992	2005	(8.19) tunable clustering; (8.19) hiera...	hierarchical lattice (29.13, 1.0E-4); ro...	visualizing network
10	7	0.981	2006	(4.26) graph-based data mining; (4.2...	discipline (14.66, 0.001); model (13....	profiling leading scientist
15	7	0.978	2000	(5.3) graph-based data mining; (5.3) ...	comparison (18.79, 1.0E-4); scientifi...	discovery
18	7	0.973	1999	(8.77) biological network; (7.47) novel;	biological network (55.52, 1.0E-4); n...	biological network
2	5	0.968	1970	(9.67) structuration; (7.49) meaning; ...	meaning (27.53, 1.0E-4); self-organi...	visibility
6	3	0.974	2003	...	self-correcting map (11.33, 0.001); m...	interdisciplinarity

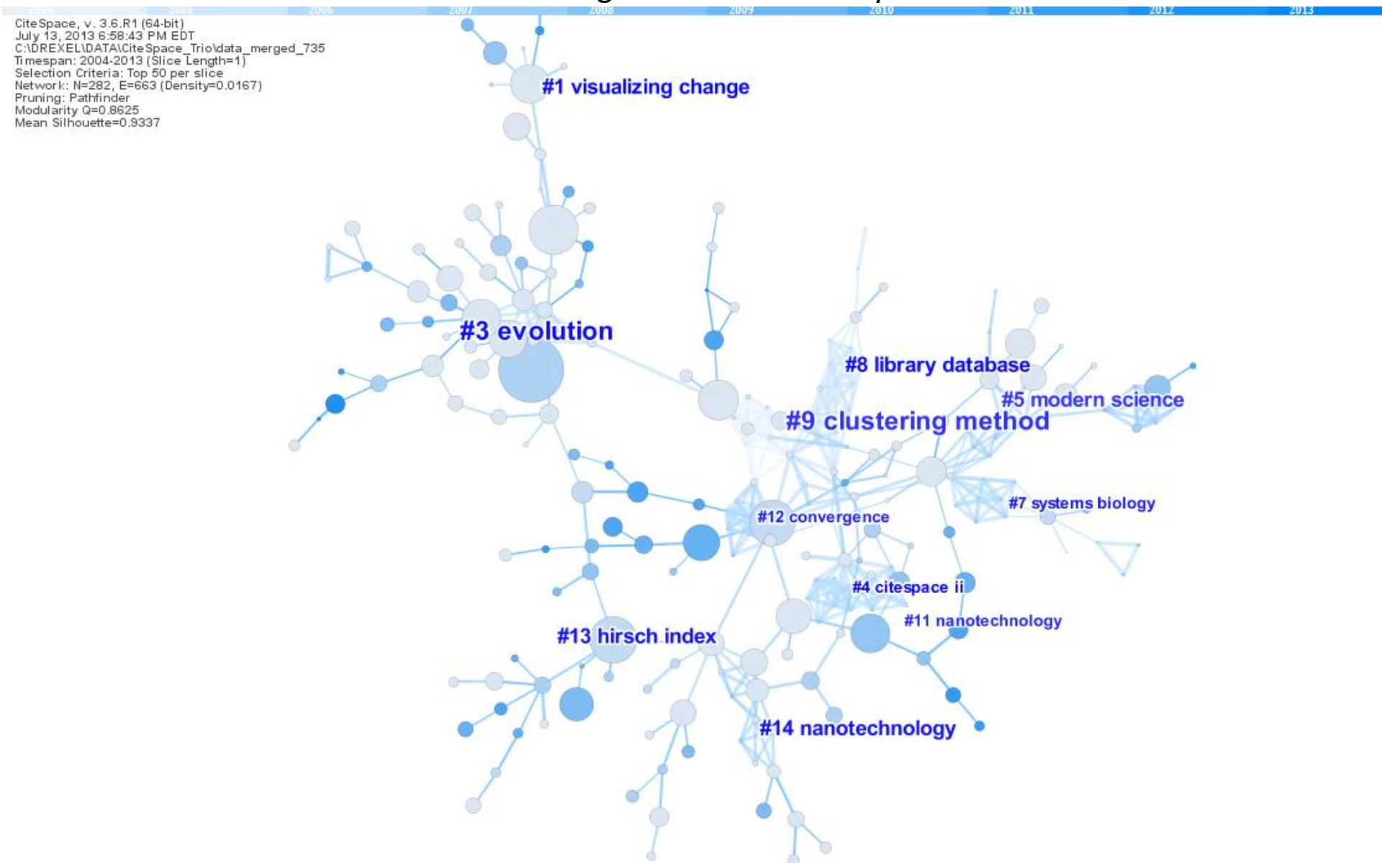
Pruned with Pathfinder

Modularity Q=0.8625

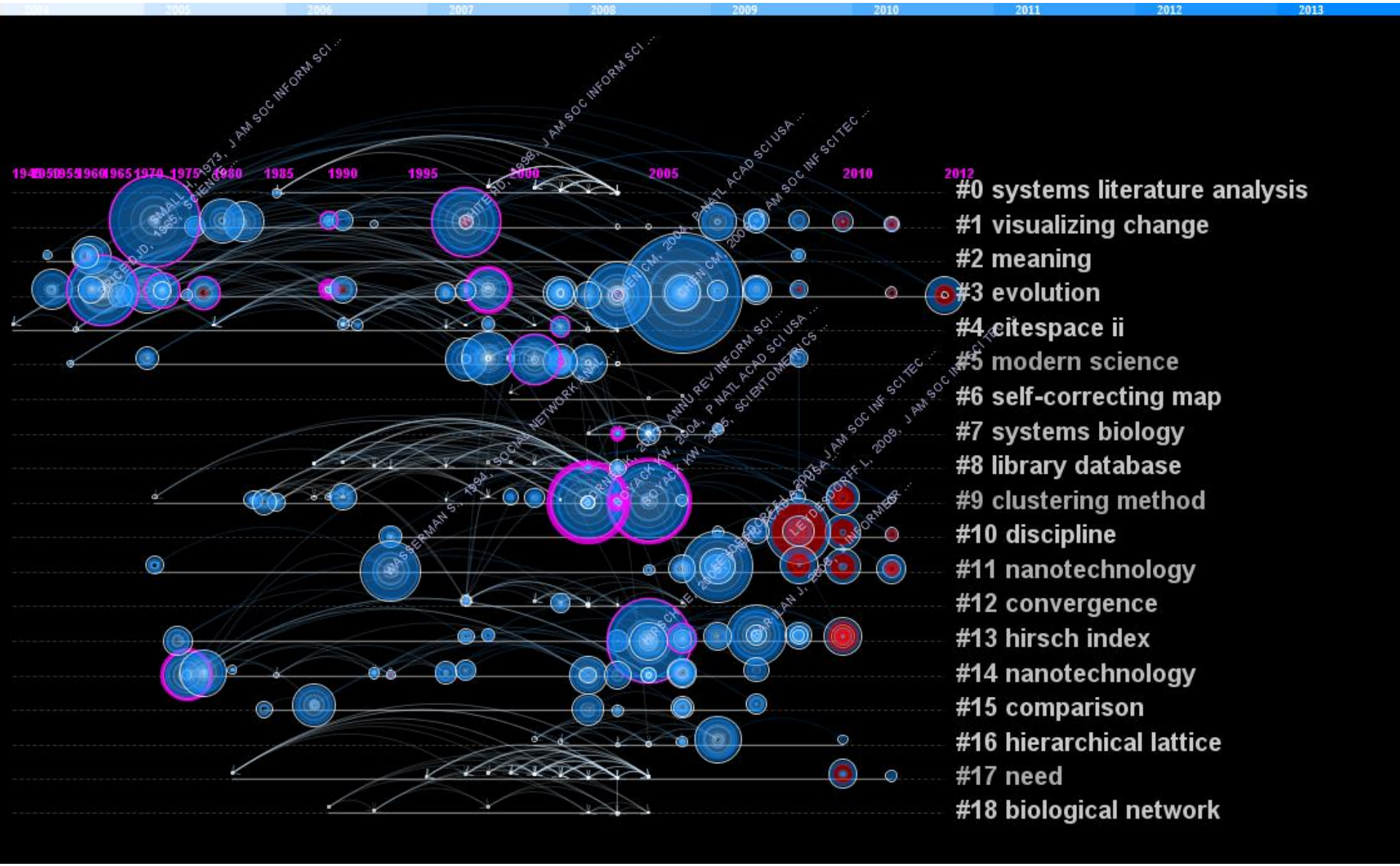
Mean Silouette = 0.9337

A Cluster View

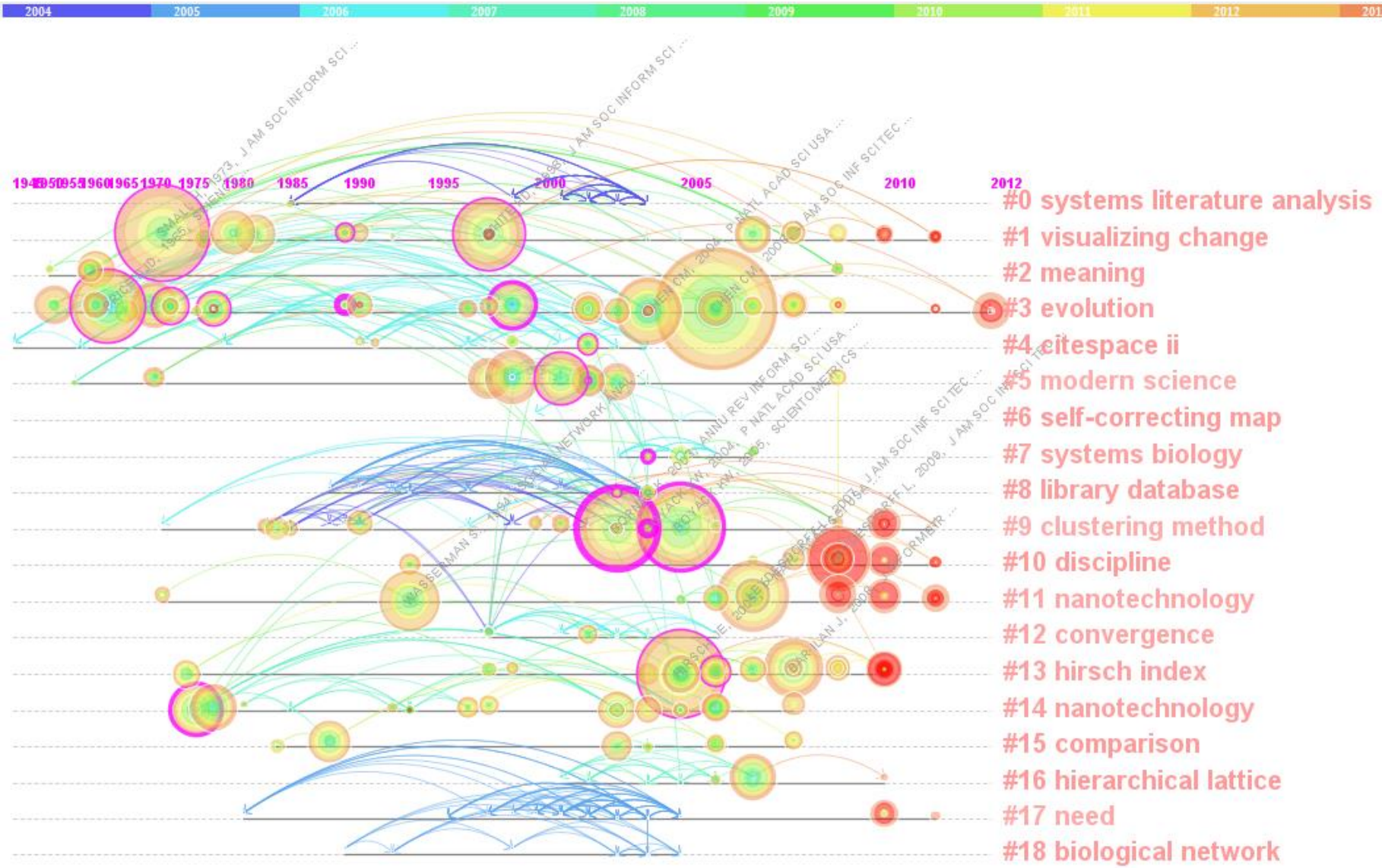
Showing Cluster Labels Only



Timeline View



Timeline View



Cluster #3: Evolution

Largest (*n*=34), Silhouette .95

Members of the Cluster

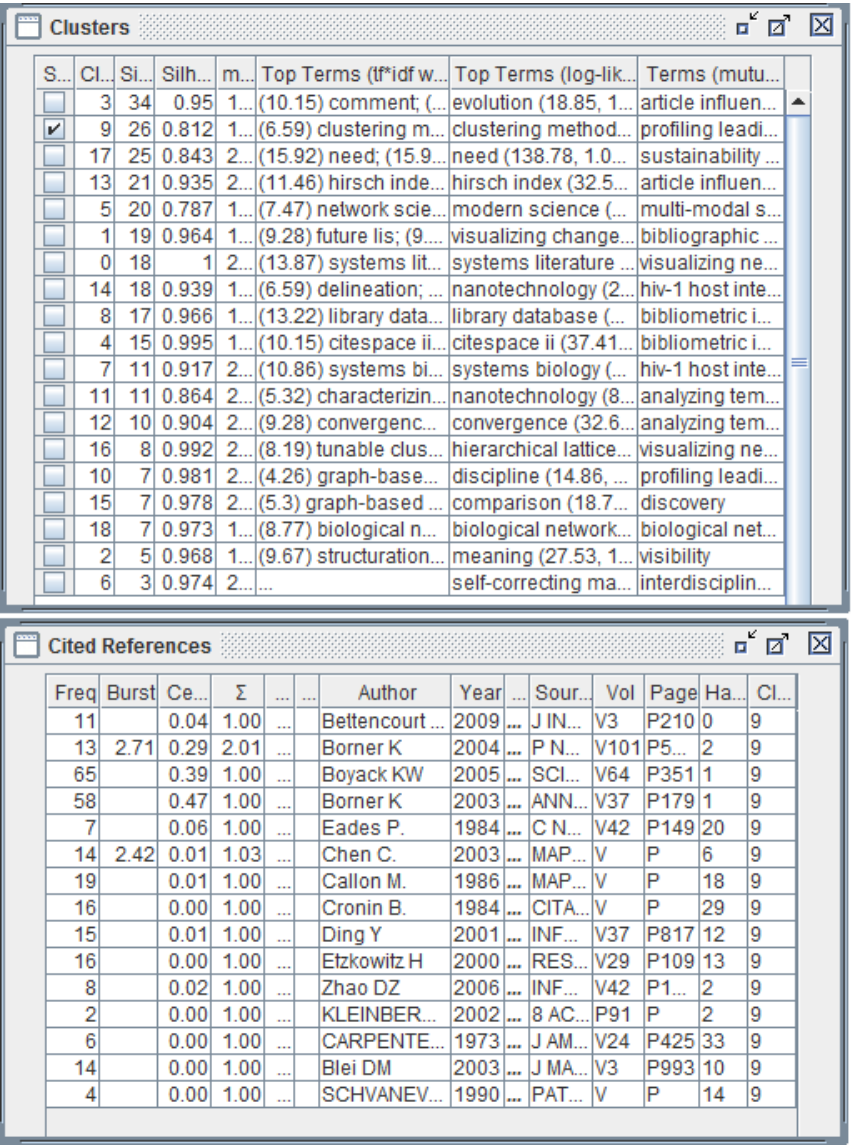
Freq	Burst	Ce...	Σ	...	...	Author	Year	...	Sour...	Vol	Page	Ha...	Cl...
94		0.00	1.00	...		Chen CM	2006	...	J AM...	V57	P359	1	3
57		0.19	1.00	...		PRICE DJD	1965	...	SCI...	V149	P510	41	3
55	3.43	0.05	1.17	...		Chen CM	2004	...	P N...	V101	P5...	0	3
37		0.01	1.00	...		GARFIELD E	1972	...	SCI...	V178	P471	35	3
35		0.29	1.00	...		Small H	1999	...	J AM...	V50	P799	7	3
32		0.06	1.00	...		GARFIELD E	1955	...	SCI...	V122	P108	53	3
31		0.12	1.00	...		SMALL H	1974	...	SCI ...	V4	P17	32	3
29		0.02	1.00	...		Small H	2006	...	SCI...	V68	P595	3	3
28	9.72	0.02	1.25	...		Vanclay JK	2012	...	SCI...	V92	P211	0	3
27		0.00	1.00	...		Girvan M	2002	...	P N...	V99	P7...	5	3
27	4.80	0.10	1.61	...		Garfield E.	1979	...	CITA...	V	P	29	3
24		0.00	1.00	...		Shibata N	2008	...	TEC...	V28	P758	2	3
24		0.05	1.00	...		Garfield E	2006	...	JAM...	V295	P90	4	3
23	2.74	0.01	1.03	...		BRAAM RR	1991	...	J AM...	V42	P233	18	3
23		0.06	1.00	...		KESSLER ...	1963	...	AM ...	V14	P10	43	3

Cluster #3: Evolution

Citing Articles

1. 0.26 VargasQuesada Benjamin (2010) showing the essential science structure of a scientific domain and its evolution
2. 0.21 Kurtz Michael J. (2010) usage bibliometrics
3. 0.21 Rafols Ismael (2010) science overlay maps: a new tool for research policy and library management
4. 0.21 Takeda Yoshiyuki (2010) tracking modularity in citation networks
5. 0.15 Chen Chaomei (2010) making sense of the evolution of a scientific domain: a visual analytic study of the sloan digital sky survey research
6. 0.15 Chen Chaomei (2010) the structure and dynamics of cocitation clusters: a multiple-perspective cocitation analysis
7. 0.12 Berendt B. (2010) intelligent scientific authoring tools: interactive data mining for constructive uses of citation networks
8. 0.12 Chen Tsung Teng (2010) citeseer: a system to visualize large cocitation networks
9. 0.12 Dolfsma Wilfred (2010) the citation field of evolutionary economics
10. 0.12 Herrera Mark (2010) mapping the evolution of scientific fields
11. 0.12 Quirin Arnaud (2010) graph-based data mining: a new tool for the analysis and comparison of scientific domains represented as scientograms
12. 0.12 Tonta Yasar (2010) diffusion of latent semantic analysis as a research tool: a social network analysis approach

Cluster #9 in Cluster Explorer



Clusters

S...	Cl...	Si...	Silh...	m...	Top Terms (tf*idf w...	Top Terms (log-lik...	Terms (mutu...
☐	3	34	0.95	1...	(10.15) comment; (...	evolution (18.85, 1...	article influen...
☑	9	26	0.812	1...	(6.59) clustering m...	clustering method...	profiling leadi...
☐	17	25	0.843	2...	(15.92) need; (15.9...	need (138.78, 1.0...	sustainability ...
☐	13	21	0.935	2...	(11.46) hirsch inde...	hirsch index (32.5...	article influen...
☐	5	20	0.787	1...	(7.47) network scie...	modern science (...	multi-modal s...
☐	1	19	0.964	1...	(9.28) future lis; (9....	visualizing change...	bibliographic ...
☐	0	18	1	2...	(13.87) systems lit...	systems literature ...	visualizing ne...
☐	14	18	0.939	1...	(6.59) delineation; ...	nanotechnology (2...	hiv-1 host inte...
☐	8	17	0.966	1...	(13.22) library data...	library database (...	bibliometric i...
☐	4	15	0.995	1...	(10.15) citespace ii...	citespace ii (37.41...	bibliometric i...
☐	7	11	0.917	2...	(10.86) systems bi...	systems biology (...	hiv-1 host inte...
☐	11	11	0.864	2...	(5.32) characterizin...	nanotechnology (8...	analyzing tem...
☐	12	10	0.904	2...	(9.28) convergenc...	convergence (32.6...	analyzing tem...
☐	16	8	0.992	2...	(8.19) tunable clus...	hierarchical lattice...	visualizing ne...
☐	10	7	0.981	2...	(4.26) graph-base...	discipline (14.86, ...	profiling leadi...
☐	15	7	0.978	2...	(5.3) graph-based ...	comparison (18.7...	discovery
☐	18	7	0.973	1...	(8.77) biological n...	biological network...	biological net...
☐	2	5	0.968	1...	(9.67) structuration...	meaning (27.53, 1...	visibility
☐	6	3	0.974	2...	...	self-correcting ma...	interdisciplin...

Cited References

Freq	Burst	Ce...	Σ	...	...	Author	Year	...	Sour...	Vol	Page	Ha...	Cl...
11		0.04	1.00	...		Bettencourt ...	2009	...	J IN...	V3	P210	0	9
13	2.71	0.29	2.01	...		Borner K	2004	...	P N...	V101	P5...	2	9
65		0.39	1.00	...		Boyack KW	2005	...	SCI...	V64	P351	1	9
58		0.47	1.00	...		Borner K	2003	...	ANN...	V37	P179	1	9
7		0.06	1.00	...		Eades P.	1984	...	C N...	V42	P149	20	9
14	2.42	0.01	1.03	...		Chen C.	2003	...	MAP...	V	P	6	9
19		0.01	1.00	...		Callon M.	1986	...	MAP...	V	P	18	9
16		0.00	1.00	...		Cronin B.	1984	...	CITA...	V	P	29	9
15		0.01	1.00	...		Ding Y	2001	...	INF...	V37	P817	12	9
16		0.00	1.00	...		Etzkowitz H	2000	...	RES...	V29	P109	13	9
8		0.02	1.00	...		Zhao DZ	2006	...	INF...	V42	P1...	2	9
2		0.00	1.00	...		KLEINBER...	2002	...	8 AC...	P91	P	2	9
6		0.00	1.00	...		CARPENTE...	1973	...	J AM...	V24	P425	33	9
14		0.00	1.00	...		Blei DM	2003	...	J MA...	V3	P993	10	9
4		0.00	1.00	...		SCHVANEV...	1990	...	PAT...	V	P	14	9

Citing Articles

1. 0.19 Rafols Ismael (2010) science overlay maps: a new tool for research policy and library management
2. 0.19 van Eck Nees Jan (2010) software survey: vosviewer, a computer program for bibliometric mapping
3. 0.15 VargasQuesada Benjamin (2010) showing the essential science structure of a scientific domain and its evolution
4. 0.15 van Eck Nees Jan (2010) a comparison of two techniques for bibliometric mapping: multidimensional scaling and vos
5. 0.12 Kurtz Michael J. (2010) usage bibliometrics
6. 0.12 Minguillo David (2010) toward a new way of mapping scientific fields: authors' competence for publishing in scholarly journals
7. 0.12 Takeda Yoshiyuki (2010) tracking modularity in citation networks
8. 0.12 Waltman Ludo (2010) a unified approach to mapping and clustering of bibliometric networks
9. 0.08 Chang Y. F. (2011) classification and visualization of the social science network by the minimum span clustering method
10. 0.08 Edmonds Bruce (2011) simulating the social processes of science
11. 0.08 Evans James A. (2010) industry induces academic science to know less about more
12. 0.08 JorgeBotana Guillermo (2010) visualizing polysemy using lsa and the predication algorithm
13. 0.08 Kiss Istvan Z. (2010) can epidemic models describe the diffusion of topics across disciplines?
14. 0.08 Mutalikdesai Mandar R. (2010) co-citations as citation endorsements and co-links as link endorsements
15. 0.08 Ohniwa Ryosuke L. (2010) trends in research foci in life science fields over the last 30 years monitored by emerging topics
16. 0.08 Persson Olle (2010) identifying research themes with weighted direct citation links
17. 0.08 Small Henry (2010) maps of science as interdisciplinary discourse: co-citation contexts and the role of analogy
18. 0.08 Vieira Pedro Cosme (2010) are finance, management, and marketing autonomous fields of scientific research? an analysis based on journal citations
19. 0.08 Zhu Bin (2010) visualization of network concepts: the impact of working memory capacity differences
20. 0.08 Zhu Bin (2010) visualizing social network concepts

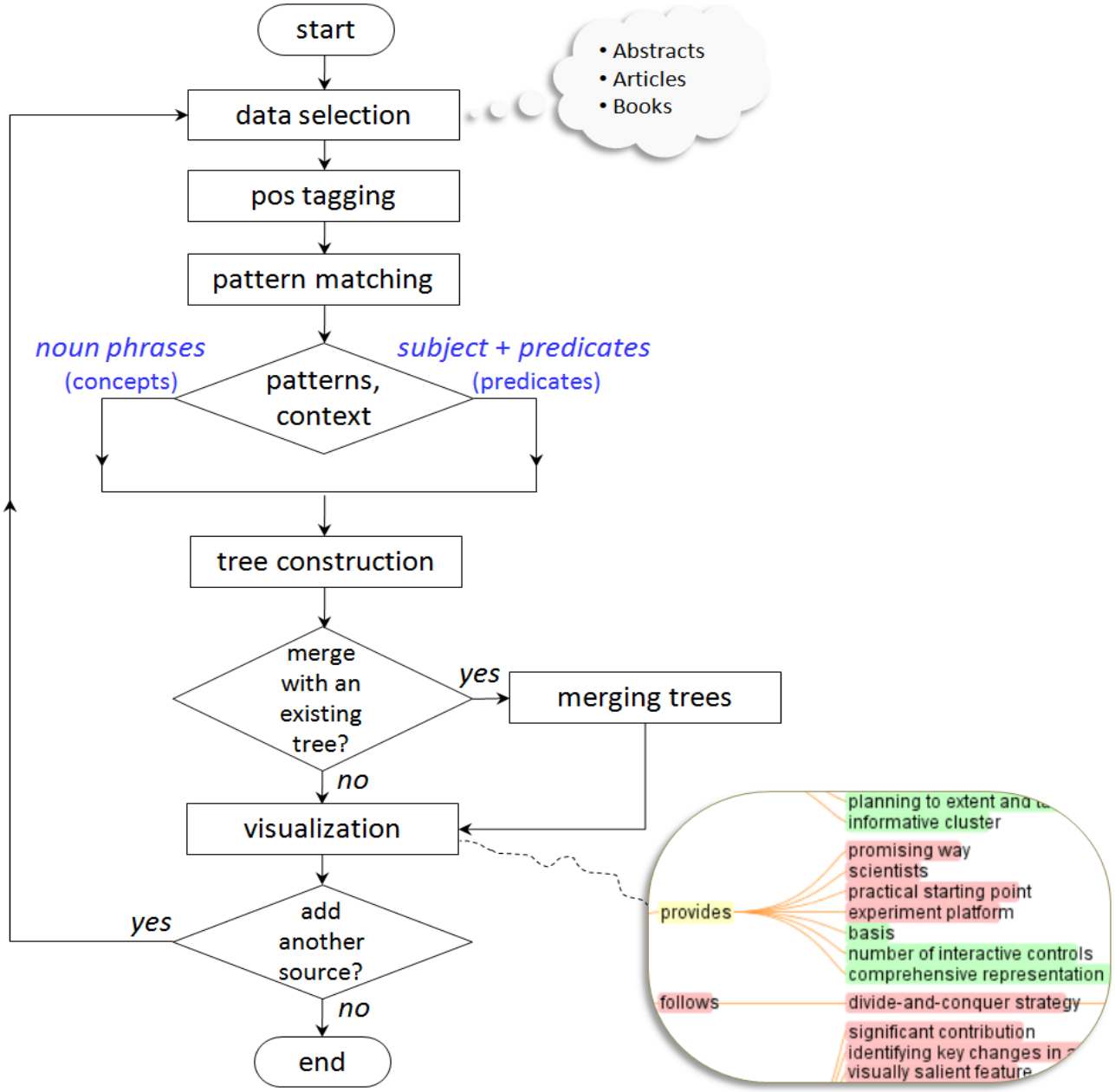
View Citation Burst History

Top 20 References with Strongest Citation Bursts

References	Year	Strength	Begin	End	2004 - 2013
CHEN CM, 2004, P NATL ACAD SCI USA, V101, P5303, DOI	2004	3.4314	**2004**	2007	
GARFIELD E., 1979, CITATION INDEXING IT, V, P	1979	4.7981	**2008**	2009	
WATTS DJ, 1998, NATURE, V393, P442	1998	3.8839	**2008**	2009	
SHIFFRIN RM, 2004, P NATL ACAD SCI USA, V101, P5183, DOI	2004	3.7456	**2008**	2009	
LEYDESDORFF L, 2006, J AM SOC INF SCI TEC, V57, P1616, DOI	2006	3.0914	**2008**	2008	
BOLLEN J, 2009, PLOS ONE, V4, P, DOI	2009	3.0354	**2009**	2011	
SMALL H, 1985, SCIENTOMETRICS, V8, P321, DOI	1985	3.4321	**2010**	2010	
LEYDESDORFF L, 2010, J AM SOC INF SCI TEC, V61, P1622, DOI	2010	3.3323	**2011**	2013	
SMALL H, 2009, SCIENTOMETRICS, V79, P365, DOI	2009	3.2571	**2011**	2011	
PORTER AL, 2009, SCIENTOMETRICS, V81, P719, DOI	2009	3.1699	**2011**	2013	
VANCLAY JK, 2012, SCIENTOMETRICS, V92, P211, DOI	2012	9.7218	**2012**	2013	
WAGNER CS, 2011, J INFORMETR, V5, P14, DOI	2011	6.2237	**2012**	2013	
RAFOLS I, 2010, J AM SOC INF SCI TEC, V61, P1871, DOI	2010	5.5182	**2012**	2013	
LEYDESDORFF L, 2009, J AM SOC INF SCI TEC, V60, P348, DOI	2009	4.8064	**2012**	2013	
LEYDESDORFF L, 2011, J INFORMETR, V5, P87, DOI	2011	4.1391	**2012**	2013	
RAFOLS I, 2010, SCIENTOMETRICS, V82, P263, DOI	2010	4.117	**2012**	2013	
COBO MJ, 2011, J AM SOC INF SCI TEC, V62, P1382, DOI	2011	3.7923	**2012**	2013	
LEYDESDORFF L, 2011, J AM SOC INF SCI TEC, V62, P846, DOI	2011	3.7331	**2012**	2013	
BOYACK KW, 2010, J AM SOC INF SCI TEC, V61, P2389, DOI	2010	3.4387	**2012**	2013	
JAHANGIRIAN M, 2010, EUR J OPER RES, V203, P1, DOI	2010	3.0762	**2012**	2013	

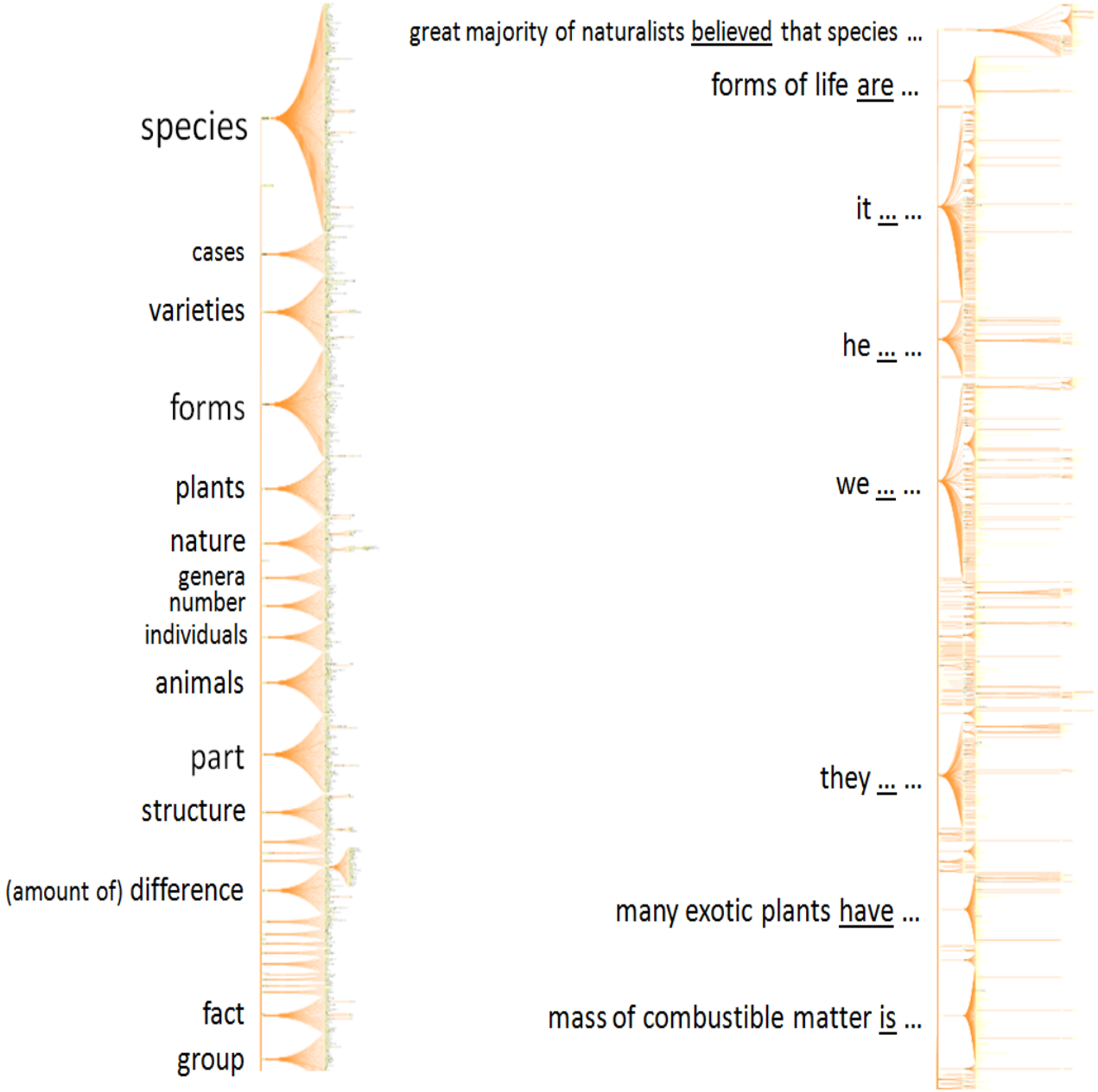
Sort by Strengths of Burst | Close

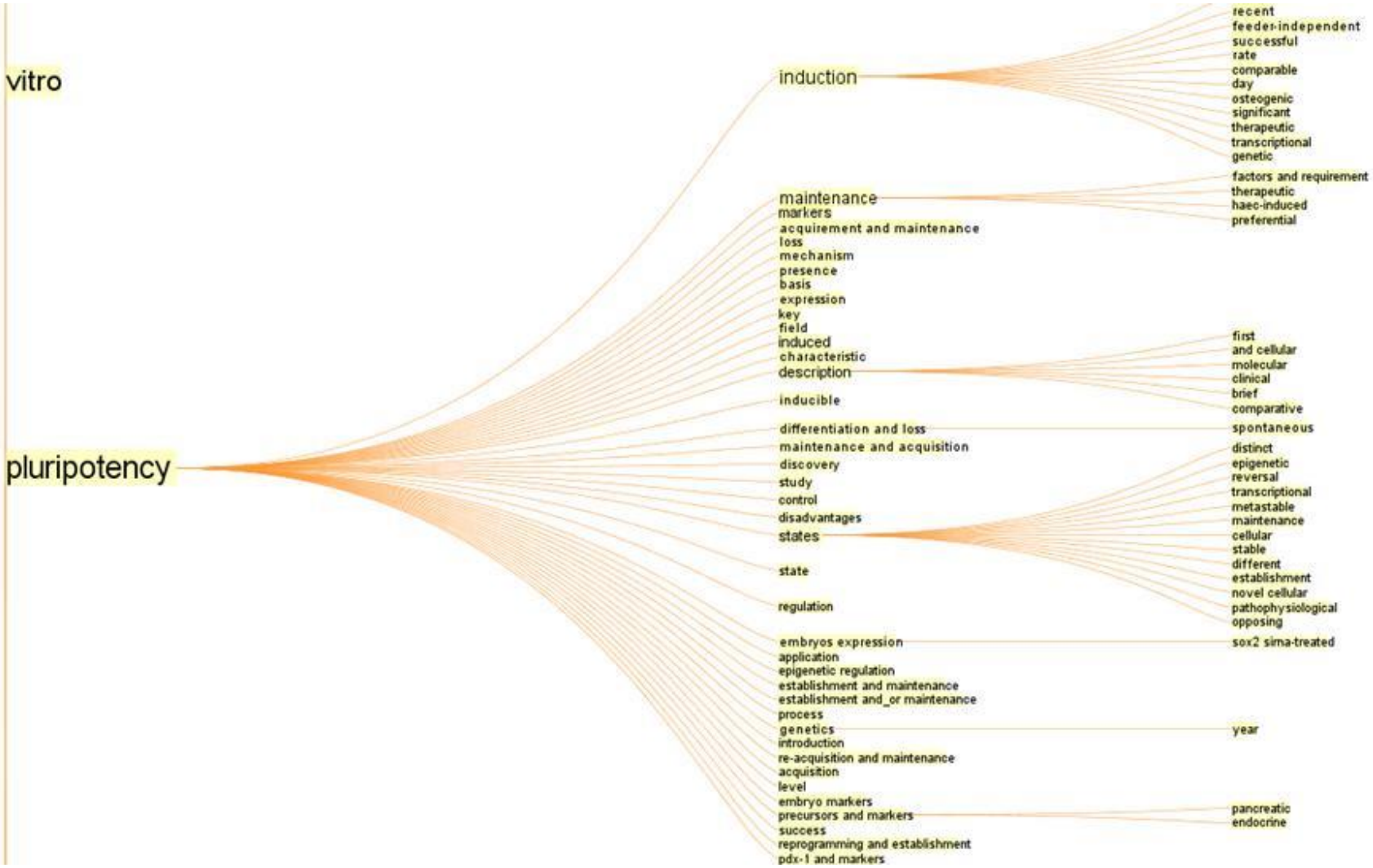
Concept Trees and Predicate Trees


start
• Abstracts
• Articles
• Books
data selection
pos tagging
pattern matching
noun phrases
(concepts)
patterns,
context
subject + predicates
(predicates)
tree construction
merge
with an
existing
tree?
yes
merging trees
no
visualization
add
another
source?
yes
no
end
planning to extent and
informative cluster
promising way
scientists
practical starting point
experiment platform
basis
number of interactive controls
comprehensive representation
provides
follows
divide-and-conquer strategy
significant contribution
identifying key changes in
visually salient feature

The Origin of Species by Charles Darwin (1872)

A 13,583-node concept tree (left) and a 5,180 node predicate tree (right) (both partially shown)



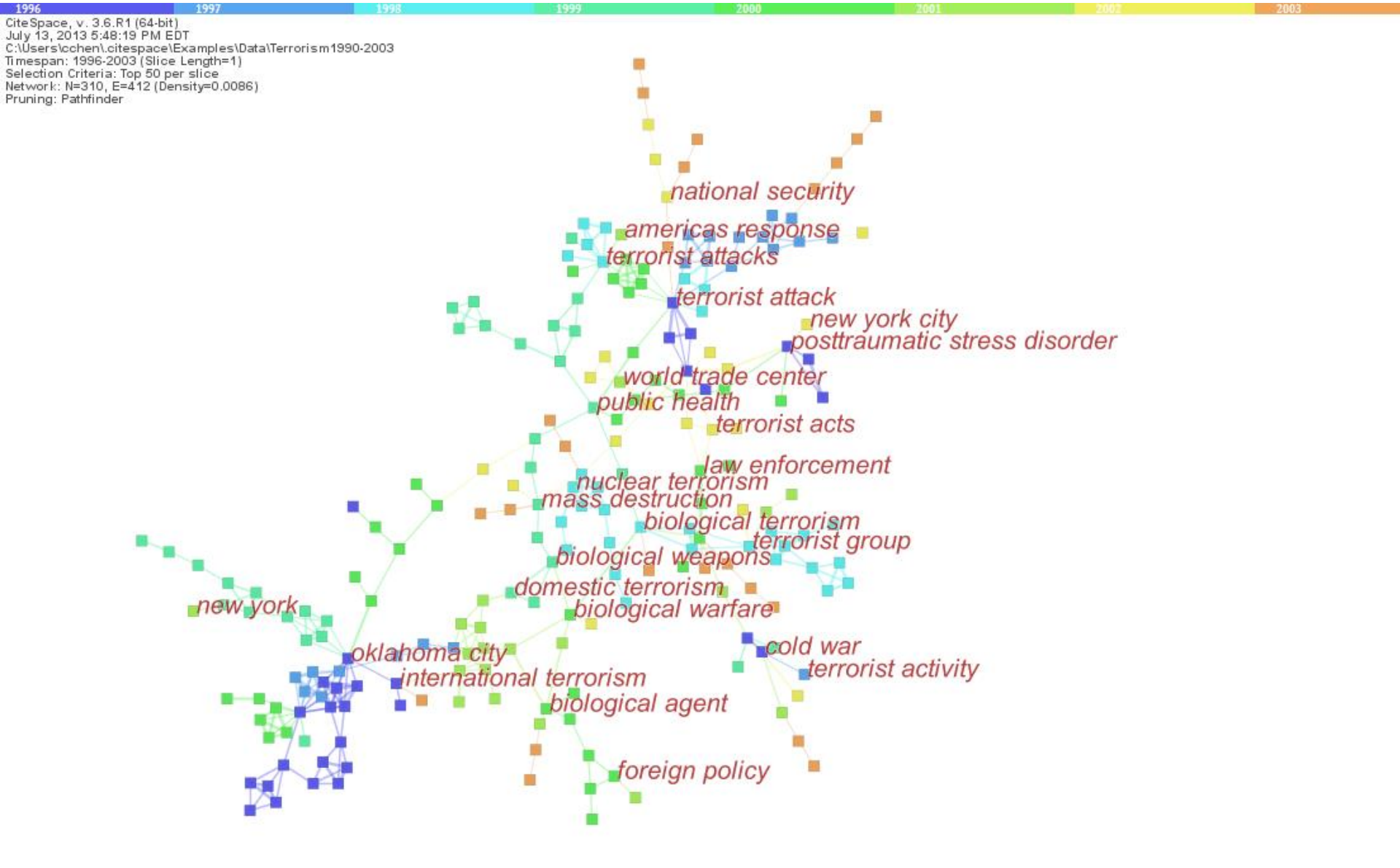

vitro
pluripotency
induction
recent
feeder-independent
successful
rate
comparable
day
osteogenic
significant
therapeutic
transcriptional
genetic
maintenance
factors and requirement
therapeutic
haec-induced
preferential
markers
acquirement and maintenance
loss
mechanism
presence
basis
expression
key
field
induced
characteristic
description
first
and cellular
molecular
clinical
brief
comparative
inducible
differentiation and loss
spontaneous
maintenance and acquisition
discovery
study
control
disadvantages
states
distinct
epigenetic
reversal
transcriptional
metastable
maintenance
cellular
stable
different
establishment
novel cellular
pathophysiological
opposing
state
regulation
embryos expression
sox2 sima-treated
application
epigenetic regulation
establishment and maintenance
establishment and_or maintenance
process
genetics
year
introduction
re-acquisition and maintenance
acquisition
level
embryo markers
precursors and markers
pancreatic
endocrine
success
reprogramming and establishment
pdx-1 and markers

[wos:000276286200028] here we report the efficient generation of induced pluripotent stem cells (ipscs) from mesenchymal cells of the umbilical cord matrix (up to 0.4% of the cells became reprogrammed) and the placental amniotic membrane (up to 0.1%) using exogenous factors and a chemical mixture.
[wos:000280921000057] although the induction of genome integration-free induced pluripotent stem cells (ipscs) has been reported, c-myc was still required for the efficient generation of these cells.
[wos:000284104100018] we used a completely defined (xeno-free) system that we previously developed for efficient generation of authentic dopaminergic neurons from human embryonic stem cells (hescs), and applied it to ipscs.
[wos:000284104100018] efficient generation of functional dopaminergic neurons under defined conditions will facilitate research and applications using pd patient-specific ipscs.
[wos:000284104100018] efficient generation of functional dopaminergic neurons from human induced pluripotent stem cells under defined conditions.
[wos:000273438400038] conclusion: this work is first to demonstrate the efficient generation of hepatic endodermal lineage from human ipscs that exhibits key attributes of hepatocytes, and the potential application of ipsc-derived he in studying human liver biology.
[wos:000276730400010] this study established an animal model for efficient generation of patient-specific es cell lines using cryopreserved oocytes.
[wos:000283048200085] here we report that a high density of human esc-derived fibroblast-like cells (hesdfs) supported the efficient generation of hepatocyte-like cells with functional and mature hepatic phenotypes from primate escs and human induced pluripotent stem cells.
[wos:000275086100004] however, besides the efficient generation of bona fide, clinically safe pscs (e.g., without the use of oncoproteins and gene transfer based on viruses inserting randomly into the genome), a major challenge in the field remains how to efficiently differentiate pscs to specific lineages and how to select cells that will function normally upon transplantation in adults.
[wos:000278953100013] these results demonstrate that innate differentiation propensity of hpscs could be overcome, at least in part, by modulation of intracellular signaling pathways, resulting in efficient generation of desirable cell types, such as neural cells.

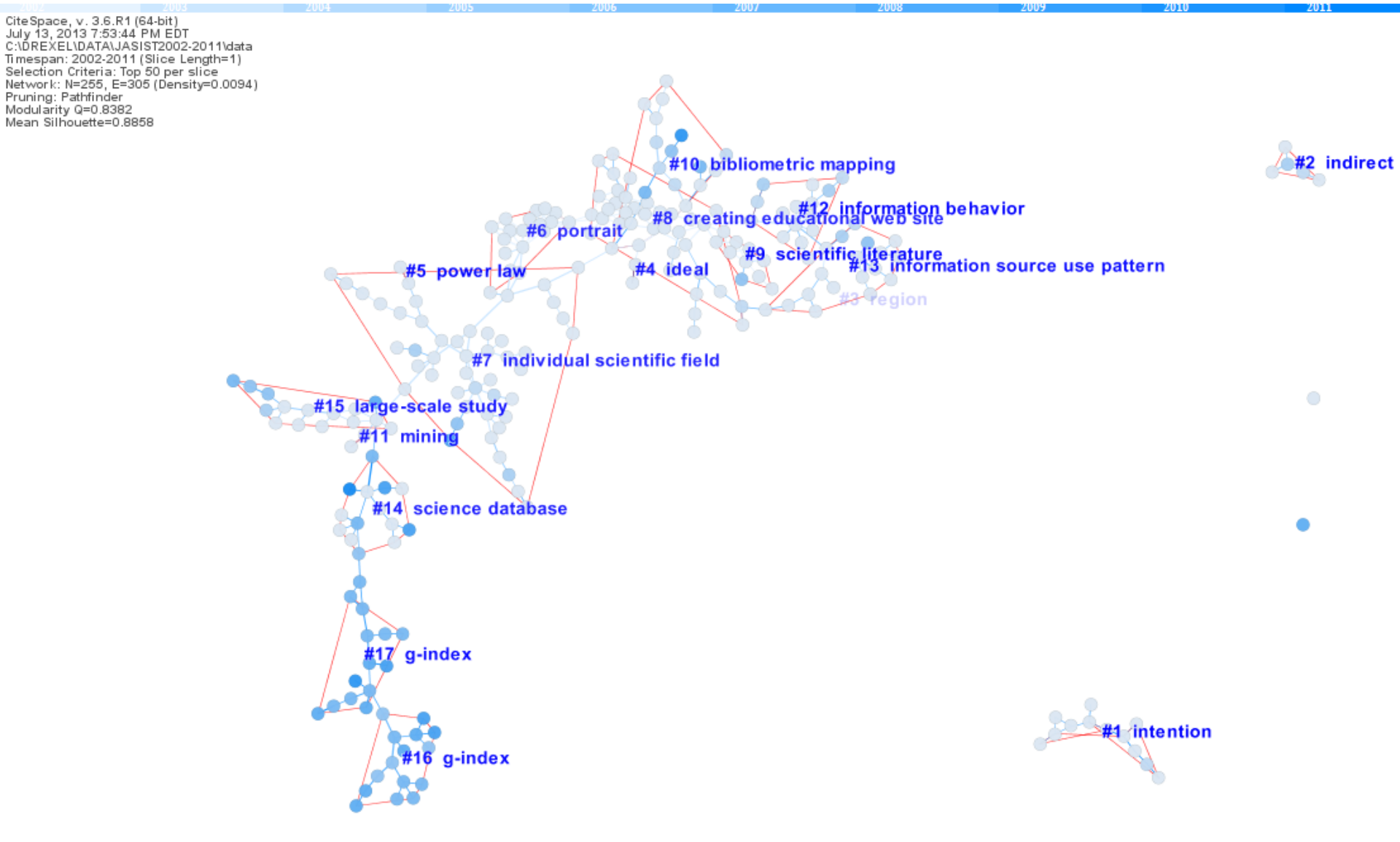
[wos:000273974600023] we conclude that the piggybac transposon system can be used to perform multiplexed stable gene transfer in cultured human cells, and this technology may be valuable for applications requiring concurrent expression of multiprotein complexes.
[wos:000274354400047] we conclude that microwell-engineered eb size regulates cardiogenesis and can be used for more efficient and reproducible formation of hesc-cms needed for research and therapeutic applications.
[wos:000276359200005] we conclude that cells of ectomesenchymal origin serve as an excellent alternative source for generating ips cells.
[wos:000276766600014] based on a comparison of in vitro and cell measurements, we conclude that cytoplasmic degradation by exonucleases can be a considerable barrier against efficient gene delivery.
[wos:000279103500006] therefore, we conclude that ce cells lack potential for photoreceptor differentiation and would require reprogramming to be useful as a source of new photoreceptors.
[wos:000281572700014] we conclude that epigenetic silencing of klf4 in b-cell lymphomas and particularly in chl may favor lymphoma survival by loosening cell-cycle control and protecting from apoptosis.
[wos:000283997800062] we conclude that phage integrase-mediated site-specific recombination can produce ips cells that have undisturbed endogenous gene function and could be safe for future human therapeutic application.
[wos:000284147700007] conclusion and significance: we conclude that the first essential function of sox2 in the preimplantation mouse embryo is to facilitate establishment of the trophectoderm lineage.
[wos:000291961200007] we conclude that nuclear transfer has still much to teach us about faithful nuclear reprogramming to pluripotency.

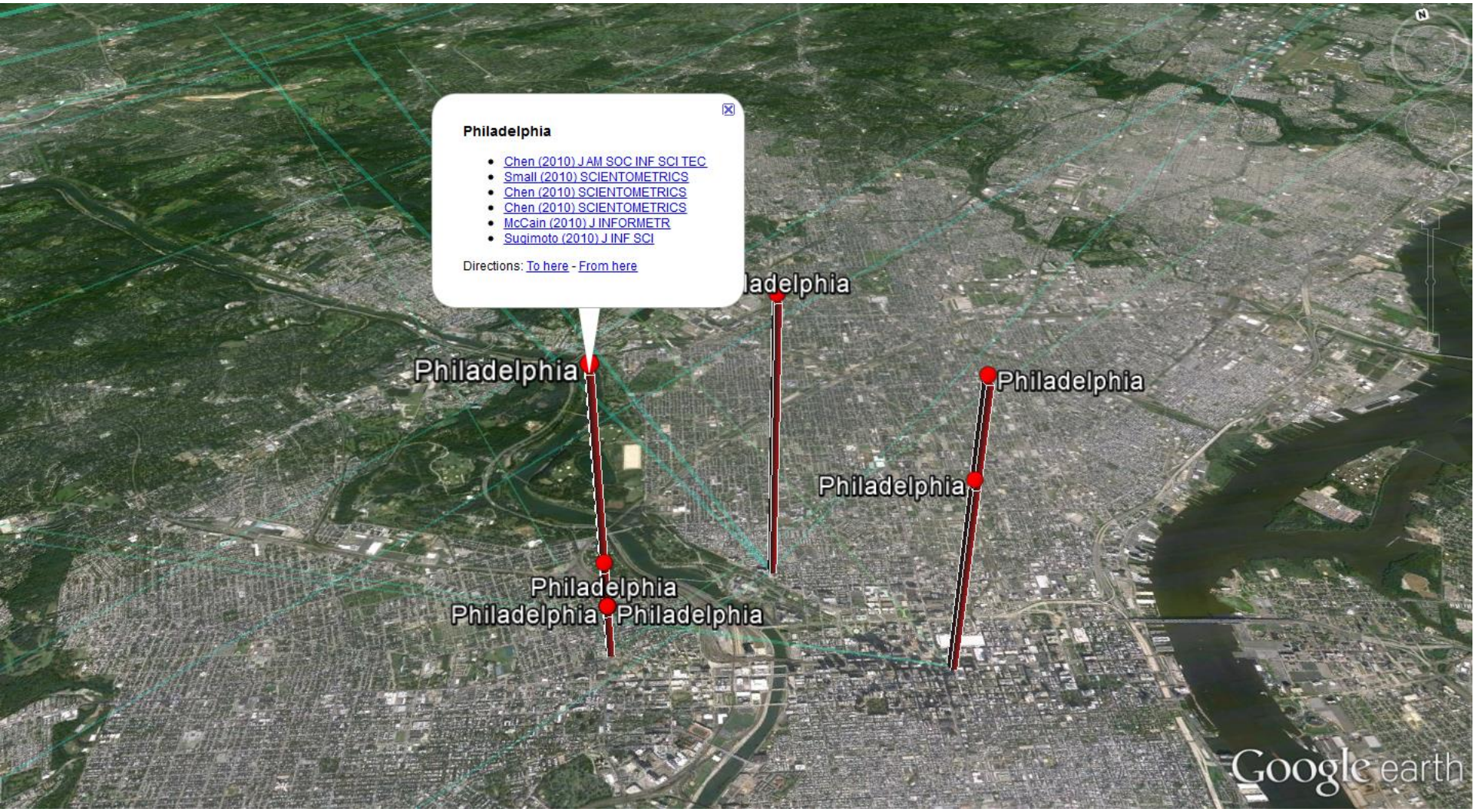
A Network of Extracted Noun Phrases

Source: Terrorism Research



Citers to JASIST (2002-2011)




Philadelphia
Chen (2010) J AM SOC INF SCI TEC
Small (2010) SCIENTOMETRICS
Chen (2010) SCIENTOMETRICS
Chen (2010) SCIENTOMETRICS
McCain (2010) J INFORMETR
Sugimoto (2010) J INF SCI
Directions: To here - From here
Philadelphia
Philadelphia
Philadelphia
Philadelphia
Philadelphia
Philadelphia
Google earth

Questions and Discussions