

文献阅读与综述方法 之 文献计量学

---以CiteSpace为例





汇报主要内容



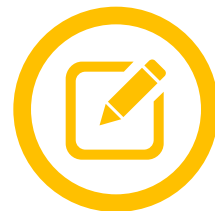
1 文献计量学



2 CiteSpace



3 案例分析



4 Q&A



汇报主要内容



1 文献计量学



2 CiteSpace



3 案例分析



4 Q&A



1 文献计量学

1.1 文献计量学工具

●Bicomb



●CiteSpace



●HistCite



HistCite™

Bibliometric Analysis and Visualization Software

●BibExcel





汇报主要内容



1 文献计量学



2 CiteSpace



3 案例分析



4 Q&A



2 CiteSpace

2.1 CiteSpace简介——是什么、用来干什么？



- 美国德雷塞尔大学 (Drexel University, Philadelphia, PA, USA) 信息科学与技术学院教授陈超美博士开发
- 科学图谱及知识可视化软件,采用谱聚类的方法对共被引网络进行聚类
 - 显示一个学科或知识域在一定时期发展的趋势与动向
 - 显示若干研究前沿领域的演进历程

CiteSpace能解决的问题:

- 在某个研究领域中，哪些文献是具有开创性和标志性的？
- 在某个研究领域的发展历程中，哪些文献起着关键作用？
- 哪些主题在整个研究领域中占据着主流地位？
- 不通的研究领域之间是如何相互关联的？
- 基于一定**知识基础**的**研究前沿**是如何发生演变的？

含有研究前沿术语词汇的文献的引文，实际上，它们反映的是研究前沿中的概念在科学文献中的吸收利用知识的情况。

某个科学领域中科学文献的暂时性成分。CiteSpace中，一个科学的研究前沿表现为涌现的施引文献群组。

2.3 CiteSpace文献来源

Web of
Science

CSSCI(Chinese
Social Science
Citation Index)

CNKI

NSF

Derwent

Scopus

arXiv-Print

Pubmed

SDSS([Sloan
Digital Sky
Survey](#))

1 共词分析

通过对能够表达某一学科领域研究主题或研究方向的专业术语共同出现在一篇文献中的现象的分析,判断学科领域中主题间的关系

展现该学科的研究结构,这种研究结构所代表的研究内容就是该学科的约定或重要约定,分析出这样的学科约定,也就找到了该学科的学科范式

例如,以Knowledge Management 在数据库中搜索出n篇文献;统计得出:A,B,C三个术语出现频次高,三者将作为领域内重要的术语。

术语A:innovations

术语B:Competitive advantage

Knowledge Management

术语C:organizational learning

2 共被引分析

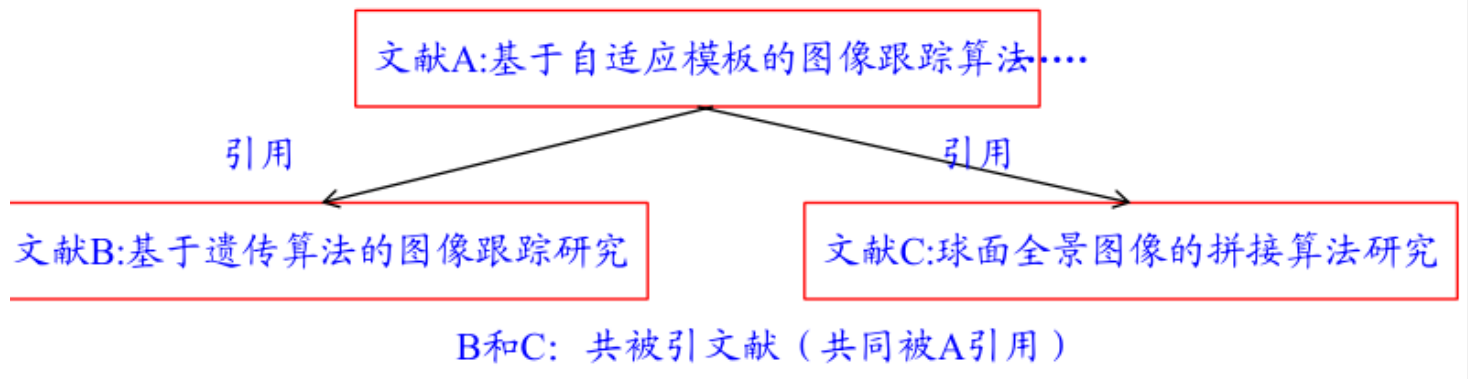
共被引分析包括作者共被引和文献共被引

虽然上述两者分析的对象不同,但原理一致:

※ 当两篇文献(作者)同时被第三篇文献(作者)引用时,这两篇文献(作者)之间就存在共被引关系

例如: 文献A, B, C, 文献A同时引用了文献B和C, 则B和C为共被引文献, 共被引频次为一次。

B和C存在共引关系

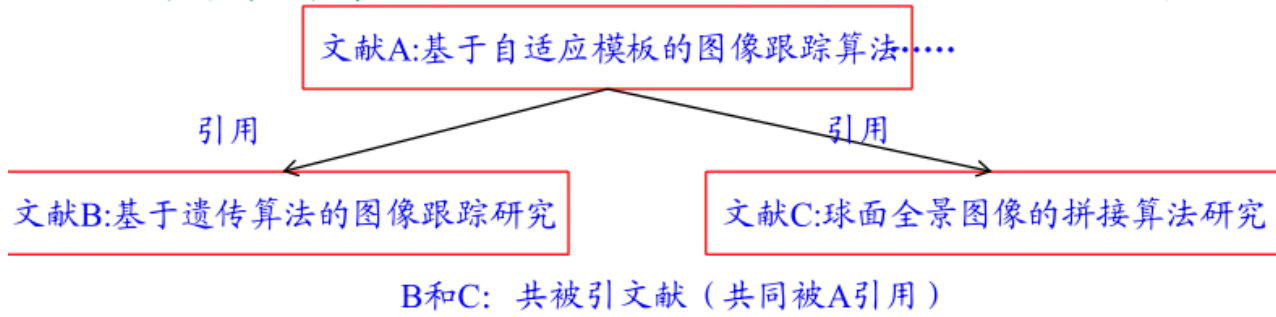


2 共被引分析

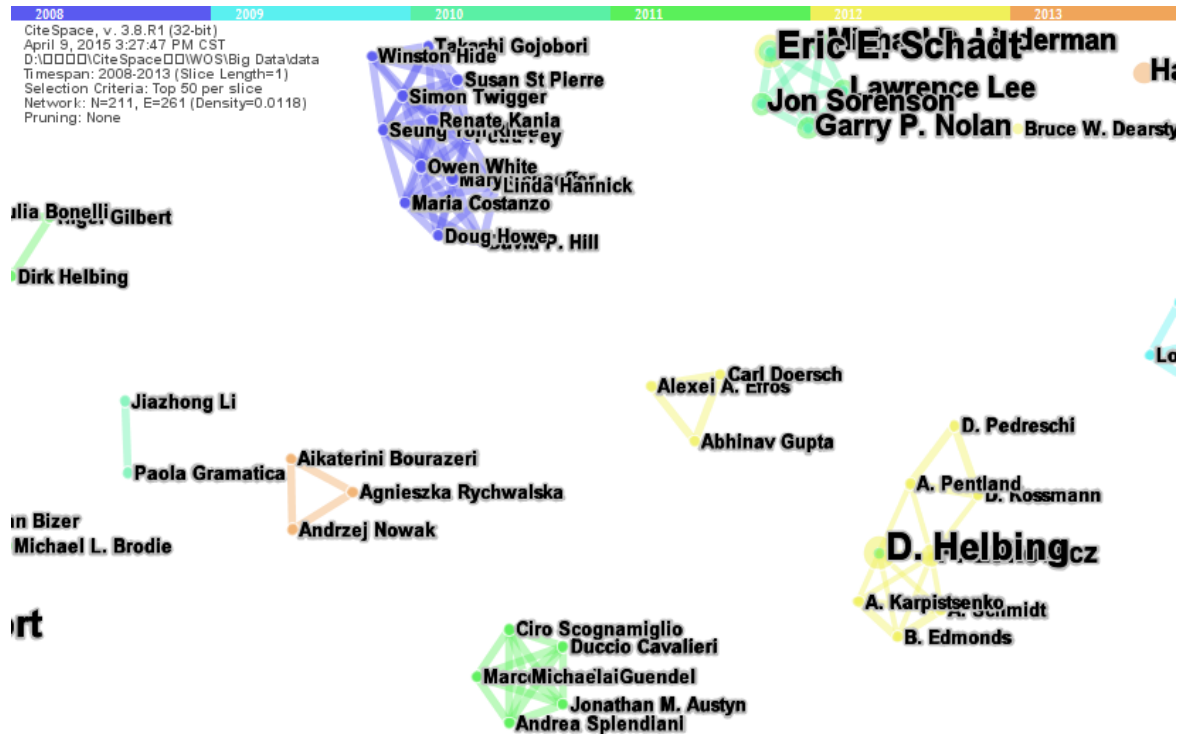
共被引文献(作者)：它(他)们在研究主题的概念、理论或方法上是相关的

如果文献(作者)的共被引次数越高,共被引强度越大,则证明二者之间的相关度越高(如图,除文献A,或许文献B和C共同被更多文献引用)

共被引分析的优势：数据的客观性、分类的科学性和有效性（与学者个人归纳、访谈调研等传统方法相比）

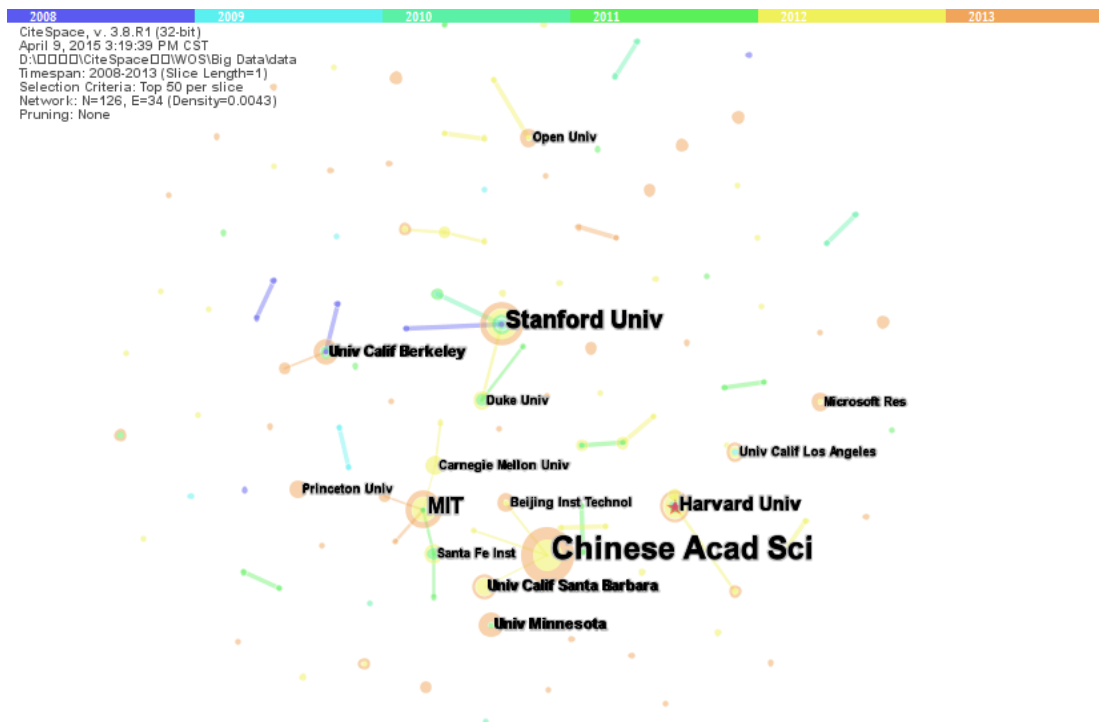


2.5 CiteSpace功能详述



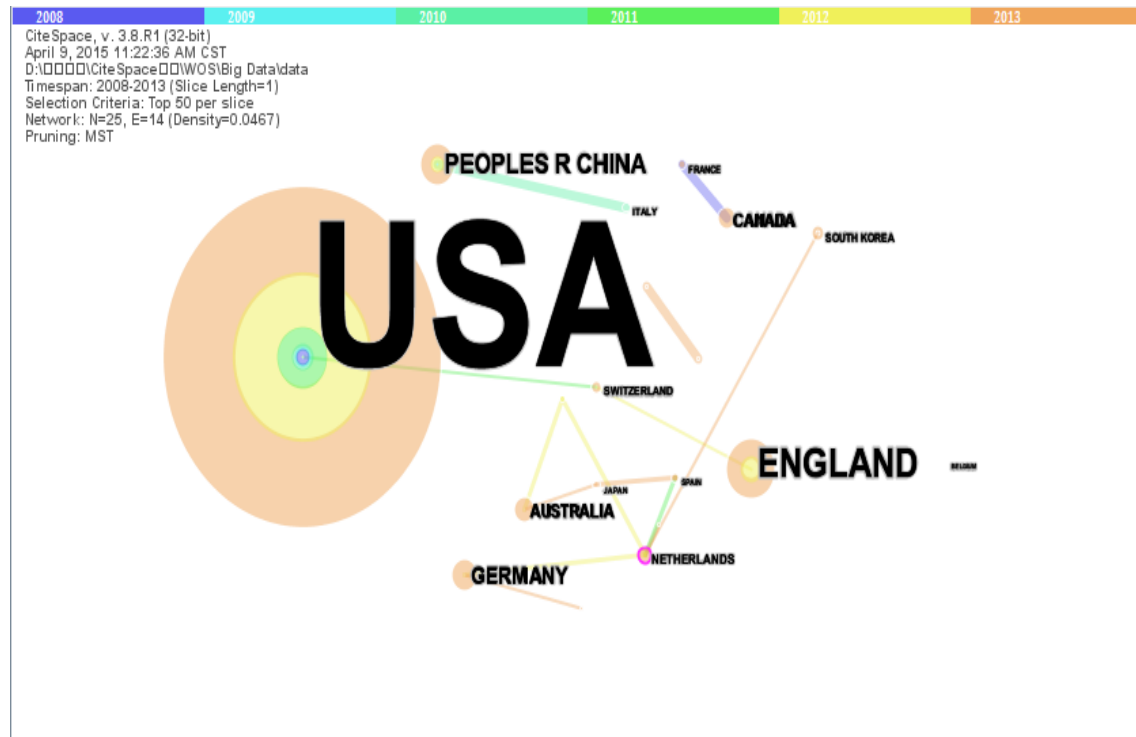
合作作者分析

2.5 CiteSpace功能详述



合作机构分析

2.5 CiteSpace功能详述



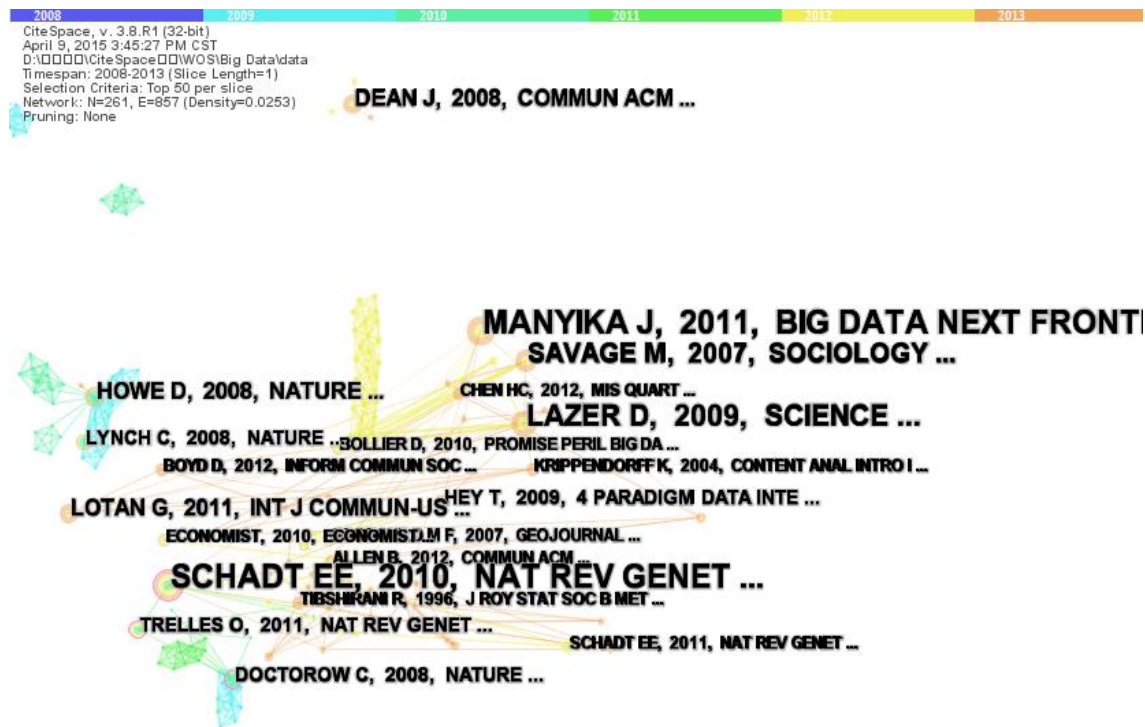
合作国家分析

2.5



学科领域知识图谱

2.5 CiteSpace功能详述



共被引文献知识图谱

2.5 CiteSpace功能详述

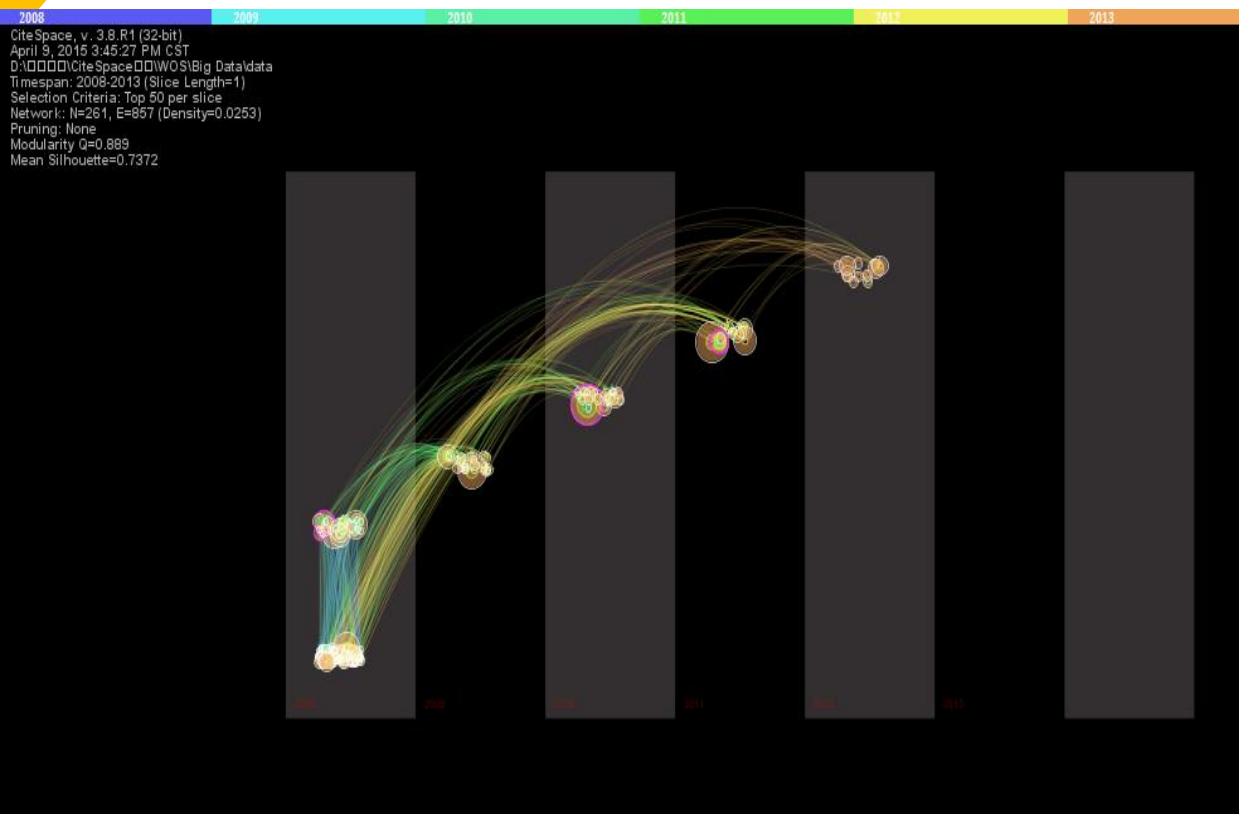


我们可以做如下分析：

1. 在哪些年份，该聚类开始出现，即有了该聚类的第一篇参考文献
2. 在哪些年份，该聚类的成果开始增多
3. 在哪些年份该聚类开始趋冷，关注度降低
4. 在整个聚类的发展过程中，在哪些年份出现了哪些标志性文献，这些文献怎样影响着整个聚类的走势

共被引文献——Timeline视图

2.5 CiteSpace功能详述

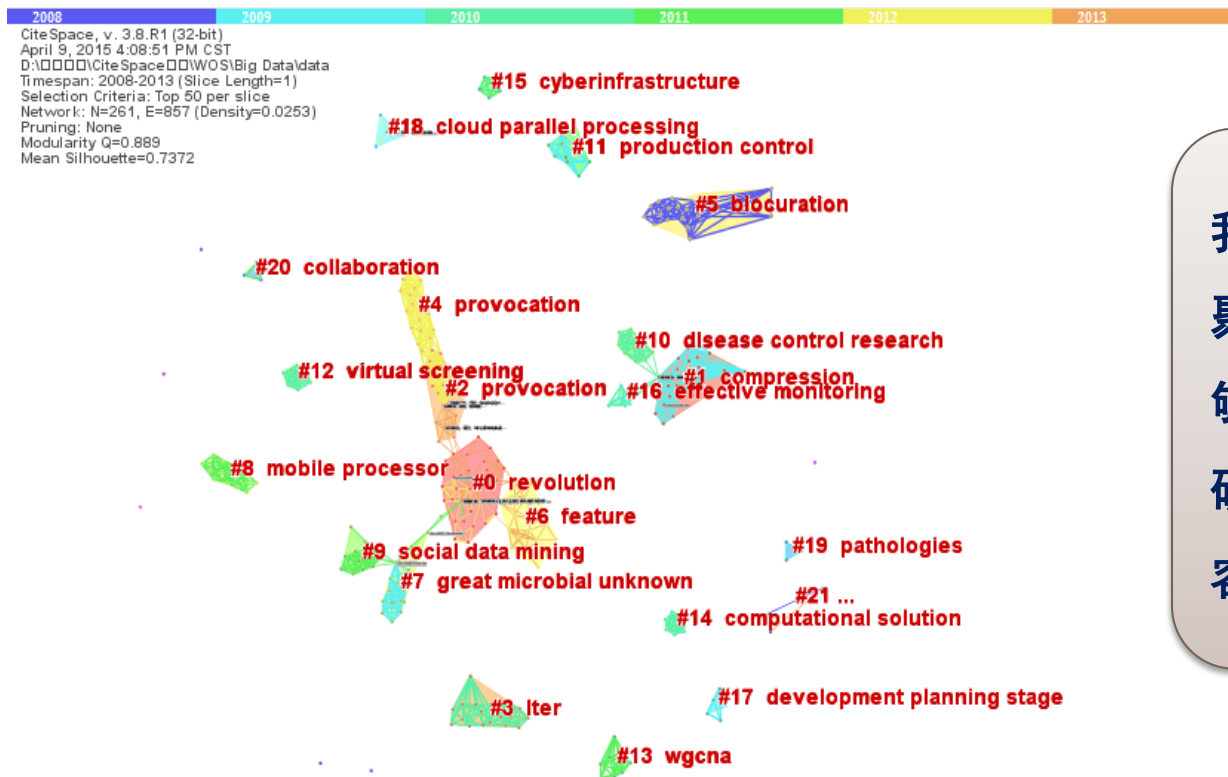


某一时区中的文献越多，说明这一时间段中发表的成果越多，该领域处于繁荣期；

某一时区中的文献越少，说明这一时间段中发表的成果越少，该领域处于低谷期。

共被引文献——Timezone视图

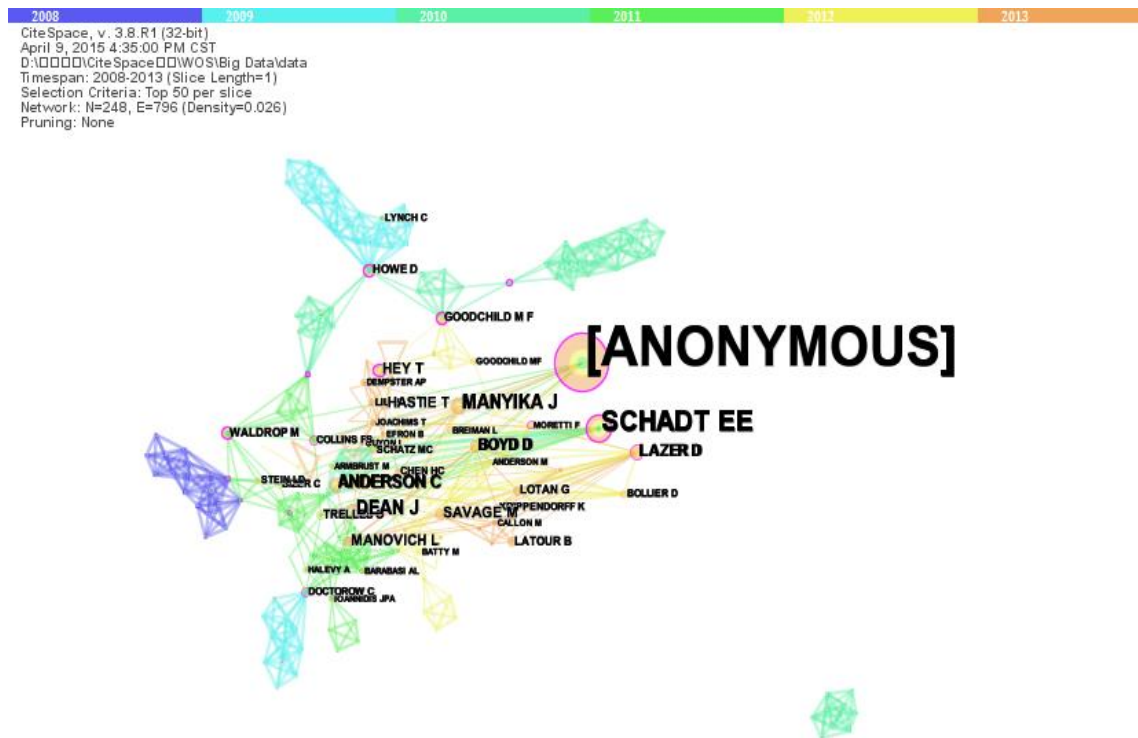
2.5 CiteSpace功能详述



我们可以认为
聚类的名称能
够间接反映出
研究前沿的内
容。

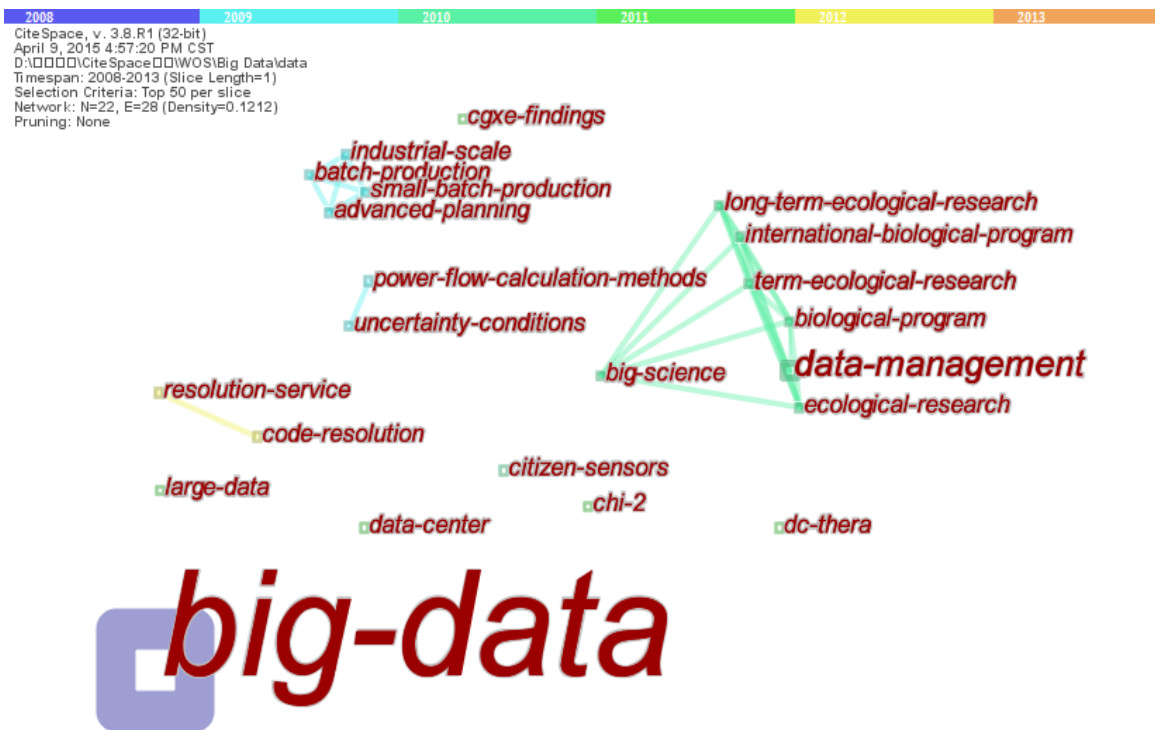
共被引文献聚类知识图谱

2.5 CiteSpace功能详述



共被引作者分析

2.5 CiteSpace功能详述



突现词分析

2.6 CiteSpace经典文献查询功能★

CiteSpace: Display Merged - (c) 2003-2015 Chaomei Chen - Project Home: D:\河北大学\CiteSpace资料\WOS\Big Data\project

File Metrics View Layout Display Network Overlays Filters Clusters Export Help

1. Clustering CiteNumPas-7
1. Clustering (Advanced) Cite-G

2. Label Clusters
3. Display Labels Selected by Different Algorithms
4. Summarization of Clusters
5. List Top Ranked Terms per Cluster by LSA
6a. View Similarity Networks of Citing Terms (VSM)
6b. View Citing Networks to Clusters (LSA)
Expectation Maximization (EM)
Enable/Disable Cluster Membership Export
Set the Minimum Number of Words of Cluster Label Terms
Set the Maximum Number of Words of Cluster Label Terms
Set the Maximum Number of Title Terms for Cluster Labeling
Set the Maximum Number of Index Terms for Cluster Labeling
Set the Maximum Number of t*idf Terms to display
Set the Maximum Number of Log-Likelihood Ratio (LLR) Terms to display
Summarize a Single Cluster
Select Cluster Summarizing Sentences
Cluster Explorer

Spotlight Citation Burs

2008
CiteSpace, v. 3.8.R1 (32-bit)
April 9, 2015 4:22:35 PM CST
D:\ODD\CiteSpace\DIWO4
Timespan: 2008-2013 (Slice
Selection Criteria: Top 50 pe
Network: N=251, E=857 (Der
Pruning: None
Modularity Q=0.889
Mean Silhouette=0.7312

VI, Fr., C., Y., Cited References

9 0. 2. SCHADT EE, 2010, ...
9 0. 2. MANYIKA J, 2011, BI...
8 0. 2. LAZER D, 2009, SCI...
7 0. 2. SAVAGE M, 2007, S...
6 0. 2. DEAN J, 2008, COM...
6 0. 2. HOWE D, 2008, NAT...
6 0. 2. LOTAN G, 2011, INT...
5 0. 2. DOCTOROW C, 200...
5 0. 2. WALDROP M, 2008...
5 0. 2. LYNCH C, 2008, NA...
5 0. 2. HEYT, 2009, 4 PAR...
5 0. 2. TRELLES O, 2011, ...
4 0. 2. BOLLIER D, 2010, P...
4 0. 2. GOODCHILD M F, 2...
4 0. 2. ECONOMIST, 2010, ...
4 0. 2. KRIPPENDORFF K, ...
4 0. 2. SCHADT EE, 2011, ...
4 0. 2. CHEN HC, 2012, MI...
4 0. 2. BOYD D, 2012, INFO...
4 0. 1. TIBSHIRANI R, 1996...
4 0. 2. ALLEN B, 2012, CO...
3 0. 2. BOYD S, 2004, CON...
3 0. 2. GONZALEZ MC, 200...
3 0. 2. ANDERSON C, 2008...
3 0. 2. GUYON I, 2003, JOU...
3 0. 2. DUHIGG C, 2012, N...
3 0. 2. MONS B, 2008, GEN...
3 0. 2. IOANNIDIS JPA, 200...
3 0. 2. LAVALLE S, 2011, MI...
3 0. 2. GOODCHILD M F, 2...
3 0. 2. KITAHARA R, 2000, ...
3 0. 2. GOLDER SA, 2011, ...
3 0. 1. DEMPSTER AP, 197...
3 0. 2. DAVENPORT T, 200...

collaboration
virtual
biocuration
mobile p
computational science
lgna

explore and summarize clusters and their member items

Input

Use a citation threshold for selecting sentences

5

确定 取消

CiteSpace Cluster Explorer

Clusters

S.	Cl.	SL	SL	m.	Top Terms (t*idf w.)	Top Terms (log-like)	Terms (mutua)
0	30	0	2	(11.3)	revolution, critical question, complex world		
1	22	0	2	(17.09)	compression, automated aut.		
2	21	0	2	(15.61)	provocation, critical question, complex world		
3	18	1	1	(17.09)	iter, iter (17.09), iter (15.48, 1.0E-4)		
4	18	0	1	(16.51)	provocation, critical question, provocation		
5	18	1	2	(15.44)	biocuration, biocuration (87.31), community int.		
6	16	0	1	(14.31)	feature, variance preservat, testing knowle		
7	16	0	2	(14.72)	great micro, great microbial unk, marine biotec.		
8	14	1	2	(15.77)	mobile proc, mobile processor, social data mi.		
9	12	0	1	(8.73)	social data, social data mining, year		
10	10	0	2	(10.43)	disease co, disease control res, disease contr...		
11	9	1	1	(13.85)	production, production control, ...		
12	8	1	1	(13.34)	virtual scree, classification (87.8		
13	8	1	2	(13.34)	wgicra, wgicra (67.96, 1.0E		
14	7	1	2	(12.76)	computatio, computational solut, analysis		
15	5	1	2	(9.4)	cyberinfrastru, cyberinfrastru, cyberinfrastru...		
16	5	0	2	(8.24)	effective mon, effective monitoring, effective monit.		
17	4	1	2	(10.34)	developme, development planni		
18	4	1	2	(7.34)	cloud paralle, cloud parallel proce, big data		
19	3	1	1	(9.09)	pathologies, pathologies (34.44, big data		
20	3	1	2	(7.56)	collaboration, collaboration (22.5, collaboration		
21	2	1	2		service operation (1, planetary nev.		
22	1	0	1		evaluation (13.67, 0...		
23	1	0	2		cyberinfrastru, cyberinfrastru...		
24	1	0	2		changing privacy la, big data		

Citing Articles

- 0.17 Singla, Deepak (2013) open source software and web services for designing therapeutic molecules
- 0.13 Iieda, Shun (2013) [systematization of the protein sequence diversity in enzymes related to secondary metabolic pathways in plants, in the context of big data biology inspired by the knapsack monocycle database](#)
- 0.13 Ties, James M. (2012) [the next industrial revolution: integrated services and goods](#)
- 0.1 Khoury, Moun J. (2013) [transforming epidemiology for 21st century medicine and public health](#)
- 0.1 Schadt, Eric E. (2012) [the changing privacy landscape in the era of big data](#)
- 0.1 Wichter, Doroja J. (2012) [theory and application in a post-systems world](#)

Cited References

Freq	Burs	Ce.	I	Pa.	Key.	Aut.	Year	Title	So.	Vol	Page/Hat.	Clu.
2	0.01	1.00	0.00		Bor.	2009		DI.	V3	P	4	0
2	0.00	1.00	0.00		Du.	2010		GE.	V2	P	2	0
2	0.00	1.00	0.00		Ho.	2008		PL.	V4	P	4	0
2	0.00	1.00	0.00		Ch.	2012		CE.	V148	P1.	0	0
1	0.00	1.00	0.00		Ar.	2007		PR.	V	P	4	0
4	0.00	1.00	0.00		Go.	2007		GE.	V69	P211	5	0
9	0.18	1.00	0.00		Sc.	2010		NA.	V11	P647	1	0
9	0.06	1.00	0.00		Ma.	2011		BIG.	V	P	1	0
5	0.04	1.00	0.00		He.	2009		4P.	V	P	3	0
4	0.00	1.00	0.00		Ec.	2010		EC.	V	P	2	0
2	0.02	1.00	0.00		Fra.	2004		BIO.	V20	P2.	9	0
2	0.00	1.00	0.00		Ab.	2012		ISO.	V337	P1.	1	0

2.6 CiteSpace经典文献查询功能★

Citing Articles

1. 0.17 Singla, Deepak (2013) open source software and web services for designing therapeutic molecules
2. 0.13 Ikeda, Shun (2013) [systematization of the protein sequence diversity in enzymes related to secondary metabolic pathways in plants, in the context of big data biology inspired by the knapsack motorcycle database](#)
3. 0.13 Tien, James M. (2012) [the next industrial revolution: integrated services and goods](#)
4. 0.1 Khoury, Muin J. (2013) [transforming epidemiology for 21st century medicine and public health](#)
5. 0.1 Schadt, Eric E. (2012) [the changing privacy landscape in the era of big data](#)
6. 0.1 Wright Dawn I (2012) [theory and application in a post-gisystems world](#)

该窗口显示的施引文献
(这些文献代表了**研究前沿**)

Cited References

Freq	Burst	Ce...	Σ	Pa...	Key...	Aut...	Year	Title	So...	Vol	Page	Half...	Clu...
2		0.01	1.00	0.00		Bor...	2009	...	DI...	V3	P	4	0
2		0.00	1.00	0.00		Du...	2010	...	GE...	V2	P	2	0
2		0.00	1.00	0.00		Ho...	2008	...	PL...	V4	P	4	0
2		0.00	1.00	0.00		Ch...	2012	...	CE...	V148	P1...	0	0
1		0.00	1.00	0.00		Ar...	2007	...	PR...	V	P	4	0
4		0.00	1.00	0.00		Go...	2007	...	GE...	V69	P211	5	0
9		0.18	1.00	0.00		Sc...	2010	...	NA...	V11	P647	1	0
9		0.06	1.00	0.00		Ma...	2011	...	BIG...	V	P	1	0
5		0.04	1.00	0.00		He...	2009	...	4 P...	V	P	3	0
4		0.00	1.00	0.00		Ec...	2010	...	EC...	V	P	2	0
2		0.02	1.00	0.00		Fra...	2004	...	BIO...	V20	P2...	9	0
2		0.00	1.00	0.00		Alb...	2012	...	SCI...	V337	P1...	1	0

该窗口显示的是被引文献 (这些文献反映的是**知识基础**) , 这些文献也是直接在图谱中显示的节点信息



汇报主要内容



1 文献计量学



2 CiteSpace



3 案例分析



4 Q&A



汇报主要内容



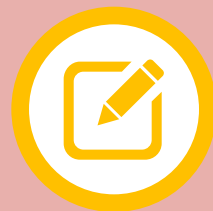
1 文献计量学



2 CiteSpace



3 案例分析



4 Q&A

The background features abstract yellow geometric shapes, including a large triangle on the left and several overlapping polygons at the bottom, creating a modern, minimalist design.

敬请老师和各位老师批评指正

T H A N K S F O R Y O U R A T T E N T I O N

文献阅读与综述方法 之 文献计量学

---以CiteSpace为例

