



Mapping Scientific Frontiers

The Quest for Knowledge Visualization

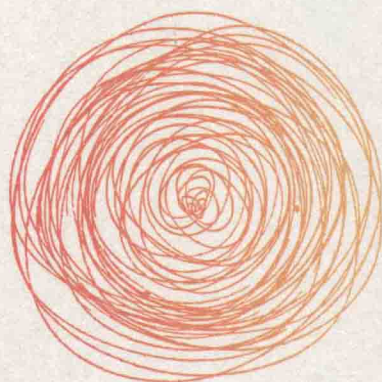
科学图谱将改变你看世界的方式

科学前沿图谱 知识可视化探索

(第二版)

[美] 陈超美◎著 陈悦 王贤文 胡志刚 侯海燕◎译

一图展春秋，一览无余；一图胜万言，一目了然。



科学出版社

大连理工大学长江学者讲座教授创新团队专题项目
教育部“985”三期哲学社会科学创新基地和
辽宁省高校人文社会科学重点研究基地项目

资助出版



Mapping Scientific Frontiers

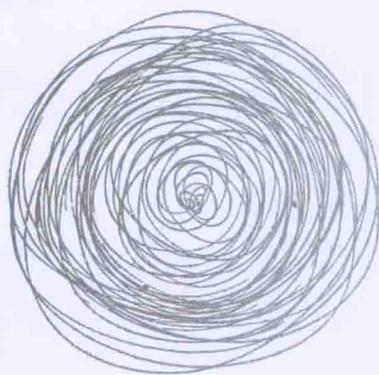
The Quest for Knowledge Visualization

科学前沿图谱 知识可视化探索

(第二版)

[美] 陈超美◎著

陈悦 王贤文 胡志刚 侯海燕◎译



科学出版社

北京

图字: 01-2013-8288号

Translation from English language edition:

Mapping Scientific Frontiers

by Chaomei Chen

Copyright © © Springer-Verlag London 2003, 2013

Springer London is a part of Springer Science-Business Media

All Rights Reserved

图书在版编目 (CIP) 数据

科学前沿图谱: 知识可视化探索 / (美) 陈超美著; 陈悦等译. —2版.
—北京: 科学出版社, 2014.7

书名原文: Mapping scientific frontiers: the quest for knowledge
visualization

ISBN 978-7-03-041007-8

I. ①科… II. ①陈… ②陈… III. ①科学知识-图集 IV. ①G302-64

中国版本图书馆 CIP 数据核字 (2014) 第 124430 号

责任编辑: 邹 聪 王茜艳 / 责任校对: 胡小洁

责任印制: 赵德静 / 封面设计: 无极书装

编辑部电话: 010-64035853

E-mail: houjunlin@mail.sciencep.com

科学出版社出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

中国科学院印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2014 年 7 月第 一 版 开本: 720 × 1000 1/16

2014 年 7 月第一次印刷 印张: 23 1/4

字数: 425 000

定价: 128.00 元

(如有印装质量问题, 我社负责调换)

译 序

高度抽象的科学前沿，通过知识可视化技术，以直观形象的图谱形式展现出来，正是《科学前沿图谱：知识可视化探索》（简称《科学前沿图谱》）一书的魅力所在。作者陈超美博士系信息可视化的最早开拓者之一，现任美国德雷塞尔大学计算与情报学院教授，大连理工大学长江学者讲座教授。

《科学前沿图谱》一书，作为世界上首部知识图谱专著，其突出特点在于作者引用大量实例对抽象知识给予形象表征进行了基础性分析，阐明了形象语言在知识创造、传播与交流中的必要性和普适性。从远古原始人的洞穴壁画到海上漂流的“瓶中信”的隐喻，从传说登上天堂的巴别塔绘画到飞向外太空航天器上供宇宙交流的形象语言，从莫基于魏格纳大陆漂移说的全球板块构造拟合到凯库勒梦中咬尾蛇隐喻的苯环结构，作者把这些古往今来的动人事例，同当代呈爆炸式增长的科学文献数据，以及日益远离直观观察的科学前沿领域联系起来，表明其实现可视化重要性，然而，这些抽象信息如何实现可视化呢？

绘制科学前沿图谱，关键在于信息可视化和可视化分析，实现“抽象—可视化—形象”的转化。作者是国际信息可视化的开拓者和领军者，1999 年率先发表了该领域的第一部专著《信息可视化》（2004 年再版），创办了国际期刊《信息可视化》（*Information Visualization*）并任主编，开发了 CiteSpace 信息可视化软件，使其成为科学知识图谱的主要工具之一。如果读者是知识图谱和知识可视化的爱好者，比较一下不同可视化软件制作的知识图谱，你一定会同意我对 CiteSpace 知识图谱形态的概括：

一图展春秋，一览无余；一图胜万言，一目了然。

CiteSpace 知识图谱之所以能达到如此神奇的效果，就在于作者把视觉思维、数学思维和哲学思维三种思维方式紧密结合起来，融合于 CiteSpace 操作的运行过程。在该书第一章题头处，作者引用了英国杰出的哲学家、数学家罗素的经典名言：

科学是那些我们已经知道的东西，哲学是那些我们还不知道的东西。

如何理解这句话的含义？当读完该书，你就会懂得：哲学就是从科学已经知道的东西中找到我们还不知道的东西。这里不妨再引用法国数学家波尔达·德莫林斯（Bordas-Demoulin）的一句名言来加以补充和深化^①：

没有哲学，固然难以得知数学的深度，然而没有数学，也同样无法探知哲学的深度，两者互相依存。还应特别指出，如果既无哲学也无数学，则就不能认识任何事物。

视觉思维乃是 CiteSpace 不言而喻的主要思维方式。视觉在人类感知外部信息中起绝对主导的作用，图像又是视觉信息的第一要素，但是不能把视觉误解为仅是一种感性认识。视觉思维是从感性视觉，到抽象思维，再到理性直观的螺旋式上升过程；它可以跨越感性视觉，直接把抽象信息与数据变换为可视化的空间结构与知识图谱。

CiteSpace 知识图谱是科学计量学的发展，离不开各种数学方法和计算方法。它洋溢着数学思维方式，实际上除了数量分析，还包括空间分析（space analysis）的几何方法。这意味着数学思维渗透着直观形象的视觉思维。CiteSpace 知识图谱是运用空间思维把非物理空间的抽象信息，映射为便于人们观察和理解的可视化形式。

哲学思维和知识图谱是什么关系？知识图谱似乎与哲学搭不上界。可 CiteSpace 的开发者陈超美告诉我：“CiteSpace 的背后需要有对库恩或类似宏观哲学思想体系的了解，才能明白 CiteSpace 到底在帮用户找什么。”这部著作吸收了库恩《科学革命的结构》中的核心概念范式（paradigm），在知识图谱中开创了竞争范式的案例分析。可以进一步说，库恩的历史主义科学哲学思维帮助作者找到了在第一代静态可视化中不知道的东西，开发了新一代多元、历时、动态的信息可视化技术，在以色带表征引文年代、以时间线显现聚类的动态知识图谱中，实现了“一图展春秋”的意境。

《科学前沿图谱》一书可谓集视觉思维、数学思维和哲学思维三种思维方式于一体，把科学学理论基础、科学计量学最新前沿、计算机科学、信息科学、科学史、科学哲学和科学社会学诸学科交叉渗透结合起来。从这部著作中，我第一次知道了陈超美把科学知识图谱起源，认可为科学学奠基人贝尔纳手工制

① 莫里兹. 1990. 数学家言行录. 朱剑英编译. 南京：江苏教育出版社：80.

作的科学技术史图 (historiographs) ①。

2008 年, 陈超美博士受聘为大连理工大学长江学者讲座教授到校履职时, 把《科学前沿图谱》英文版赠送给我。打开书本, 从色彩斑斓的巴别塔油画, 到 CiteSpace 展示的地质年代恐龙灭绝古生物史图、物理学超弦理论来龙去脉聚于一图, 一幅幅知识图谱, 令人目不暇接。我意识到《科学前沿图谱》理论厚实、方法独创、国际领先, 不仅对于我国深入研究和运用知识图谱理论与方法有重要的学术价值, 而且对于我国追踪和选择国际科技研究前沿、探索和制定科技发展战略及政策具有重大而深远的实际意义。

当时, 我向陈超美提议, 由我们 WISELab 团队翻译这部书, 同时倡议团队成员认真学习这本书, 掌握和运用 CiteSpace 技术与方法。在《科学前沿图谱》的引领下, 我们著述的国内第一部《科学知识图谱: 方法与应用》和《技术科学前沿图谱与强国战略》, 以及“知识计量与知识图谱丛书”10 本, 于 2008 年和 2011 年先后问世。可以说这 12 部知识图谱著作, 就是《科学前沿图谱》在中国播种所开的一簇知识图谱之花。如今《科学前沿图谱》(2003 年第一版) 历经 10 年之后, 经作者增补了自己最新成果的第二版, 由 WISELab 以陈悦博士率领的团队翻译并经作者校对同意, 交由科学出版社出版。

“科学图谱将改变你看世界的方式”, 这是陈超美教授为我们的《科学知识图谱: 方法与应用》一书所作序中的最后一句话。我把它作为箴言赠与《科学前沿图谱》的读者, 不妨在你阅览此书并自己动手绘制科学前沿图谱的实践中, 去领略知识图谱是如何改变观察世界的方式的。

大连理工大学 (中国) — 德雷塞尔大学 (美国)

知识可视化与科学发现联合研究所

刘则渊

2013 年 6 月

① 见贝尔纳《历史上的科学》中文版 (北京: 科学出版社, 1981) 书末第 727 页“表八”。这张大表实际上是一幅描绘从人类起源到 20 世纪中叶技术和科学中心在地域、各国与城市中扩展的历史图谱: 左侧分别为时间、时代、社会性质、技术与科学中心 4 个栏目, 由上到下做了文字说明。右侧通栏从上到下用连线描绘了不同时代及社会背景下技术与科学中心的变迁, 以及科学和技术的相互作用, 构成复杂的因果链, 引领时代的进步、社会的发展。有关此图表的内容见该书开头的序。

绘制科学图谱听着也许有些自相矛盾：科学这么抽象的东西能像地图一样画出来么？科学知识所代表的境界也许只有最有经验的科学家们才能捕捉到它的奥妙。然而，这本书所展现现在我们面前的是：科学知识图谱的研究不仅已经生根、发芽，而且正在开花、结果。不错，也许这只是一个漫长发展历程的开始，就像当年探险家们刚刚能画出第一张粗糙的世界地图一样。但是，从逻辑上来看，绘制科学知识图谱其实不过是把绘制普通地图时的对象变成了知识本身。

科学在当今世界中的重要地位不言而喻，科学为此所付出的努力也毋庸置疑。虽然我们的社会有时候小看了科学的作用，没有给它提供充足的资金支持，甚至试图阻碍科学教育计划的实施。但是，在当今人类智慧产物中，还有什么比科学更值得我们关注的呢？当然，科学本身既可用于行善，也能用于破坏。不管是谁掌握了科学原理，科学都会给其带来巨大的力量。有时候，个别科学家可能会滥用我们对他们的信任而争夺名利，超美在书中所研究的“撤稿”现象就反映了这种情况。尽管会存在这样或那样的例外，但科学仍然是我们了解宇宙的必由之路，也是人类社会和经济福祉的源泉和根基。

尽管我们用来描述科学的语言中，不乏类似于研究“领域”（field）、研究“范围”（area）这样带有空间色彩的隐喻。但当我们一旦开始绘制一幅科学图谱的时候，就会发现那些绘制地图所用的方法全都不再适用。我们所面对的，是“研究论点”“研究方向”“研究领域”或“研究学科”之间的抽象的关系和联系，而这些关联本身的存在与否也许还是个问题。库恩在一次采访中意味深长地说：“注意，现在的概念已经不再适用于过去的领域了。不单是其中的观点变了，学科的结构也变了。”^① 我们认识到，不论是在过去还是现在，科学都还有待探索。我们是否有足够理由把这些抽象的，甚至都不一定存在的关联表现为空间形式

^① Kuhn, T.S. (2000). The Road since Structure. Chicago: University of Chicago Press.

呢？我们是否在本能地把知识间的相互关联映射到实际空间？

科学知识图谱之所以难以绘制，也许是因为要想绘制出一幅有意义的图谱，需经历三个基本步骤，而其中两步都涉及一定程度的数学变换。第一步，我们需要选择一个合适的分析单元，它应该包括构成我们科学宇宙的基本粒子；第二步，需要定义如何度量这些单元之间的关联；第三步，需要找到一种方式把这些分析单元及它们之间的关联在低维空间（通常是二维）里显示出来。一旦能够实现这三个步骤的跨越，绘制科学知识图谱就变得水到渠成、自然而然了。

早先创立的科学史、科学社会学和科学哲学等学科，已经为我们了解科学思想的创生和演化，以及科学的社会建制和科学的哲学基础，提供了必要的方法和工具。但是，除了科学社会学的一些早期研究，这些学科所采用的方法一般都是定性的。科学史研究主要是利用史学的通用方法进行建构叙事，科学哲学则以逻辑基础和认识论作为研究特色。科学社会学，秉承创始人之一的默顿所用的研究范式，既重视理论基础，又鼓励使用定量证据。这一思路也被早期的社会网络研究者们所借鉴，进行无形学院的研究。然而到了20世纪70年代，科学社会学却背离了它早先的研究传统，转而信奉更为激进的科学社会学理论，如科学知识的社会建构，从而基本抛弃了早期科学社会学家所采用的定量研究方法。

此后，随着大型的科学文献数据库和引文数据库的出现，同时也为了反对建构主义的科学社会学，一个利用新的方法研究科学演化的学科诞生了。这个学科被称为科学计量学，或者信息计量学、文献计量学。这些称谓，一方面反映了其注重定量和计量的方法特征，另一方面也反映了其起源于图书馆学的历史渊源。不过，现在还不能说这个新兴的学科已经得到了学术界的认可，尤其是那些已经发展成熟的学科的学者的认可；对科学计量学而言，制度化的大学学科建设和学术地位也才刚刚开始建立。科学计量学的批评者们认为，单纯地将科学文献作为其主要数据来源，限制了对其他在科学研究中出现的数据的研究和分析。幸运的是，机器可自动识别的科学论文的全文格式的获取正变得越来越容易，为我们开辟了很多新的可供分析的数据来源，使我们有条件做一些用索引数据库不可能做到的研究。与科学计量学相关的软件包也开始出现，在这方面超美已经做出了榜样。种种证据表明，尽管在如何更好地与传统的方法，如科学史、科学社会学、科学哲学等研究方法，建立联系并进行融合方面，还有很长的路要走，但这一新兴研究领域已经站稳脚跟，并越来越多地在科技政

策层面发挥作用。

这本书的难得之处正在于，它构建在科学史、科学哲学的概念和发现的基础之上，而同时加入了新维度。在该书的第八章给出了一个关于诱导多能干细胞的案例，展现了科学知识图谱不仅可以帮助医药学的研究者了解自身研究领域的全貌，而且也为科学的史学研究提供了有益的参考资料。对科学知识图谱这一新兴研究领域的作用仍持怀疑态度的读者，可以读一下这一章。从中，我们可以看出新方法在探究和追踪当今科学重要进展的内部结构方面的巨大潜能。

该书还构建了一个宏大的学科框架，从计算机科学到信息科学，尤其是信息可视化技术。在第一版中，超美曾如此评价那些催生了科学知识图谱的学科：“最近几年，不同的科学知识图谱绘制方法，犹如百川入海，殊途同归……如何把知识的可视化变成一个统一的、跨学科的研究领域，还有很多的工作要做。”^①今天，这句话依然正确。科学计量学、计算机科学、网络分析方法仍然在围绕着相同的数据和问题各自为战。也许，作为学科壁垒的副作用之一，这一点无法避免。但是，我们希望该书能够有助于建立起这些不同的学科之间的联系。

绘制科学知识图谱与科学史息息相关。其例证之一是，超美以库恩的著作作为绘制科学知识图谱的一个重要依据，因为库恩所谓的科学革命可以视作一系列的时点截面图，科学在这些时点上发生着根本性的结构变革。截面分析也是科学史学家非常认可的一种方法，因为史学家深知要想理解一个历史时期的思想，就要进入那个时期的心理状态，“想他们之所想”^②，而不是用现在的理解去解释过去的科学。这个要求其实很难，因为一旦我们知道了已经出现的新发现和新结论是什么，就很难不被它们影响，我们常常迫不及待地要去找这个新发现的先驱和肇端。作为研究者，我们需要尽量避免这种“让现在影响过去”的态度。

绘制科学知识图谱是由不同的学科知识汇聚形成的，尤其是心理因素和社会因素之间的相互影响，前者如库恩的格式塔转换理论（一种看待概念转换的方式），后者如学院关系网和无形学院。社会关系是否决定了认知关系呢？或是认知关系决定了社会关系？在米尔格拉姆（Stanley Milgram）早期有关社会网络的研究中^③，受试者需要考虑如何通过熟人的熟人与陌生人建立联系。在斯旺森

① Chen, C. (2003). Mapping Scientific Frontiers. London: Springer.

② Kuhn, T.S. (1977). The Essential Tension. Chicago: University of Chicago Press.

③ Milgram, S. (1967). The small world problem. Psychology Today, 2, 60-67.

(Don Swanson) 关于隐藏的公开知识的研究中^①，科学发现依赖于寻找那些不相关概念之间的间接联系。可见，在社会结构和认知结构的研究中，使用的其实是同一种思路。如果我们以词语和引文作为我们绘制科学知识图谱的基本单元，那么心理学就会起作用，因为这其中涉及作者的记忆和回忆。但是记忆和回忆也受其他作者和同事的话语影响。如果我们绘制科学家的合作关系，那么社会因素显然会起作用，但是心理因素同样也会影响我们对合作者的选择。因此，社会和心理因素在科学的社会结构和知识结构中，不可避免地缠绕在一起。

在绘制科学知识图谱时，存在不同的分析单元可供选择，如词语、引文、作者、期刊等，每一个还有对应的共现分析，如共词、共被引、共作者，以及期刊引用等。而所有这些分析方法，归根到底是我们想要看到的结构类型和关系程度。比如，若要更好地理解科学发现在某个专题发展中的作用，我们可以应用共被引分析，因为很多发现都是与具体的文章和作者联系在一起的。而如果我们想关注更广泛的社会性，即非学术性因素时，便可以应用共词分析，因为这样我们可以更容易地获知公众或者政府对科学的态度。期刊，作为一种更宽泛的分析单元，或许更适合用来表示整个研究领域或学科。分析单元的选择还受限于其所处的历史阶段。文献共被引对于分析 1900 年之前的文献可能就无能为力，因为那时候还没有关于引文的规范。但是，对于这些早期文献，可以进行基于全文的共词分析。试想，如果我们可以对科学早期的文献，如 18 世纪的科学文献，进行共现分析或者其他图谱绘制，我们是否就可以识别出那个时期的主要思想派别和重要范式转移呢？这种设想是有意义的。

另一个重要的问题是怎样解读科学知识图谱。我们知道，科学知识图谱隐含的关系网络是高维度的，其在二维空间上的投射只是原始结构的一种近似，有可能会造成其实不那么相关的节点在图谱上的位置非常靠近。这意味着我们需要更加重视连线本身。连线是生成二维图谱的第一步，就像构成大脑世界的神经元一样。只有知道连线的含义，我们才能更好地理解科学知识图谱的内涵。这就要求我们充分了解连线发生的语境，用新的方法来表达并分类这些联系，比如通过功能或类型进行分类，分为逻辑相关、因果相关、社会相关、假设相关、隐喻相关等。在这方面，一个重要的进展是在该书最后一章描述的“可视化分析”系统，它的出现可以让我们对以支持决策为最终目的的图谱的绘制原

^① Swanson, D. R. (1987). Two medical literatures that are logically but not bibliographically connected. *Journal of the American Society for Information Science*, 38, 228-233.

则进行深入的探索。

绘制科学知识图谱中令人兴奋的一个原因是这个领域的不断变化：每年都有一些新论文发表，随着新兴研究领域的出现或现有研究领域的消亡，科学知识图谱的结构也会发生改变。这种状态也许会让有些人产生不安稳的感觉，而他们所愿意看到的科学知识应该是稳定的和可预测的。但是正如默顿所说的那样，机遇是科学所特有的^①。所以，对科学知识图谱也一样。我们目前还无从知道，科学发现究竟是否可以预测，是否具有可以识别的前兆，以及是否可以通过某种机制来促进新发现和新发明的快速出现。不过，由于在科学知识图谱上标识出已有的科学发现并非难事，所以我们还是有可能通过研究科学发现之前的图谱，从结构上寻找科学发现的端倪。

这是一本关于信息可视化的著作，涉及许多方面，尤其侧重于科学知识图谱绘制。绘制科学图谱作为一个新兴领域，还处在其婴儿阶段，还要面对许多理论和实践上的挑战。该书的最后一章列出了这些挑战。另外，新的修订版中添加了很多精辟的内容。该书不仅可以作为入门者的初级读本，也可以帮助资深专家学者全面了解这一领域的最新进展。

美国科学技术策略公司 (SciTech Strategies, Inc.)

亨利·斯莫尔 (Henry Small)

^① Merton, R. K. & Barber, E. (2004). *The Travels and Adventures of Serendipity*. Princeton: Princeton University Press.

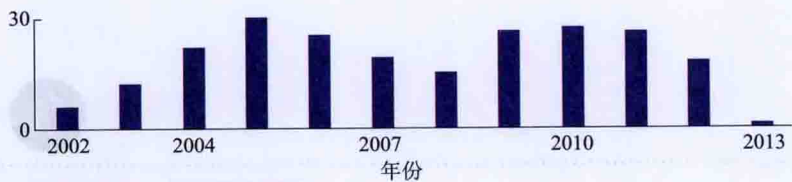
自序

从 2003 年《科学前沿图谱》第一版出版至今已有十多年了，这期间发生了许许多多的变化。社交媒体蓬勃发展的程度远远超出我们的想象。从 FaceBook (2004)、YouTube (2005) 到 Twitter (2006)，新闻、争论、恶作剧和学术博客等在新舞台上各显神通。iPhone (2007)、iPad (2010) 等移动设备使得这些社交媒体真正达到了无孔不入、无处不在的境地。

科学在过去的十年发展中有所突破。2002 年，格里戈里·佩雷尔曼 (Grigori Perelman) 证明了困扰数学界一个多世纪的“庞加莱猜想”；2006 年，山中伸弥 (Shinya Yamanaka) 和他的同事开启了诱导性多能干细胞 (iPSCs) 的研究并获得诺贝尔生理学/医学奖；2012 年，欧洲核子研究中心发现了“上帝粒子”，等等。

大科学正在变得更“大”。海量数据采集正致力于科学研究，如始于 2000 年的斯隆数字巡天 (Sloan Digital Sky Survey, SDSS, 2000 ~ 2014) 计划预计获取超过一百万个天体的多色测光资料 and 光谱数据而形成一种大数据源。随着已有科学领域的变革，新的研究领域不断兴起。2005 年，可视化研究作为新生事物进入我们的视野，极大地促进了将科学技术应用于解决现实问题，尤其当我们要处理的数据是复杂的、不确定的、不完整的并具有潜在冲突时，它的优势就越发显现出来。关于科学文献自身良好品质的维护是一个很好的例子。出版物种类与数量的增多掩盖了撤销论文在数量上的实际增长。要维护一个值得信赖的科学知识团体，我们需要如何作为？

《科学前沿图谱》这本书发挥着什么作用呢？通过谷歌学术搜索，该书在互联网上被引频次已达 235 次，而其施引文献又被更多的文献所引用。我们可以通过这些施引文献来窥见科学知识图谱研究的发展状况。有趣的是，从被引的情况看，科学知识图谱的发展经历了两个阶段。第一阶段从 2002 年至 2008 年，第二阶段从 2009 年至今 (图 1)。在第一阶段中，2007 年的被引频次达到峰值，而第二阶段头三年的被引情况相对平稳。在 Web of Science 中检索分析结果也是类似的。



科学前沿图谱：知识可视化探索
陈超美-2003
被引235次-相关文献-共11个版本

图 1 引用《科学前沿图谱》的文献数量年度分布
资料来源：Google Scholar

这种引用状况说明了什么？我们将《科学前沿图谱》的一系列施引文献作为一个整体来研究它们是如何反过来被随后发表的文章所引用的，尤其是当我们关注那些在 2002 ~ 2013 年被引率突增的文献时，情况就更加明朗了。图 2 列出了这种被引率突增的 25 篇文章。前一阶段的文献主题全部聚焦在“信息可视化”和“引文分析”上。在《科学前沿图谱》写作之初，确实就是出于这样的想法：将这两个不同学科的研究领域联系起来。

高被引突现的前 25 篇文献						
文献	年	突现度	突现		2002~2013	
			起始	终结		
					时间	
CHEN C, 2003, MAPPING SCI FRONTIER, V, P	2003	7.2508	2002	2006		
CARD SK, 1999, READINGS INFORMATION, V, P	1999	4.5742	2002	2006		
CHEN CM, 2002, J AM SOC INF SCI TEC, V53, P678, DOI	2002	4.1649	2002	2008		
CHEN CM, 2001, J AM SOC INF SCI TEC, V52, P315, DOI	2001	4.0433	2002	2004		
CHEN CM, 2001, IEEE T SYST MAN CY C, V31, P518, DOI	2001	4.0072	2002	2004		
CHEN CM, 2001, COMPUTER, V34, P65	2001	3.5648	2002	2005		
WHITE HD, 1997, ANNU REV INFORM SCI, V32, P99	1997	3.988	2003	2008		
GARFIELD E., 1964, USE CITATION DATA WR, V, P	1964	3.5679	2003	2008		
BATAGELJ V, 1998, CONNECTIONS, V21, P47	1998	3.6708	2004	2006		
CHEN CM, 2003, J AM SOC INF SCI TEC, V54, P435, DOI	2003	3.5478	2004	2006		
SUI DZ, 2004, PROF GEOGR, V56, P62	2004	3.7103	2005	2007		
KLAVANS R, 2006, J AM SOC INF SCI TEC, V57, P251, DOI	2006	5.0919	2009	2009		
KLAVANS R, 2006, SCIENTOMETRICS, V68, P475, DOI	2006	4.213	2009	2010		
SHIFFRIN RM, 2004, P NATL ACAD SCI USA, V101, P5183, DOI	2004	3.4738	2009	2009		
KLAVANS R, 2009, J AM SOC INF SCI TEC, V60, P455, DOI	2009	5.0572	2010	2013		
PORTER AL, 2009, SCIENTOMETRICS, V81, P719, DOI	2009	12.7909	2011	2013		
LEYDESDORFF L, 2009, J AM SOC INF SCI TEC, V60, P348, DOI	2009	7.3836	2011	2013		
VAN ECK NJ, 2010, SCIENTOMETRICS, V84, P523, DOI	2010	6.4516	2011	2013		
LEYDESDORFF L, 2010, J AM SOC INF SCI TEC, V61, P1622, DOI	2010	6.4334	2011	2013		
RAFOLS I, 2010, SCIENTOMETRICS, V82, P263, DOI	2010	4.6479	2011	2013		
PORTER AL, 2009, J NANOPART RES, V11, P1023, DOI	2009	4.1749	2011	2013		
CHEN CM, 2010, J AM SOC INF SCI TEC, V61, P1386, DOI	2010	3.9242	2011	2013		
VAN RAAN A. F. J., 2005, MEASUREMENT, V3, P1	2005	3.5074	2011	2011		
RAFOLS I, 2010, J AM SOC INF SCI TEC, V61, P1871, DOI	2010	8.1778	2012	2013		
WAGNER CS, 2011, J INFORMETR, V5, P14, DOI	2011	4.0553	2012	2013		

图 2 《科学前沿图谱》的施引文献的引文分析显示出两阶段，红条表示被引率突增

第二阶段的文献主题不同于前一阶段，主要是一系列在学科层面上绘制科学知识图谱的文献。从被引率突增的角度来看，这个阶段最具有影响力

的研究工作是 2009 年波特 (Alan L. Porter) 和拉法斯 (Ismael Rafols) 发表在 *SCIENTOMETRICS* 上关于“科学的跨学科”研究。其次, 是 2010 年拉法斯、波特和雷德斯道夫 (Loet Leydesdorff) 发表在 *The American Society for Information Science and Technology* 上的《科学图谱叠加》一文。目前, 我们仍然处在第二阶段。从分析的单元和规模来看, 跨学科互动研究是一种有助于更好地理解科学前沿动态发展的成效方式。

除了概念和理论方面的发展, 在分析、可视化和探索科学文献发展模式 and 趋势等计算工具的选择方面, 研究人员相对过去而言有更加广泛的选择空间。值得关注的工具包括 CiteSpace、HistCite、VOSViewer、Sci2、GeoTime、Jigsaw、Tableau、Gephi、Alluvial Maps、D3, 以及更通用的 WebGL 信息可视化软件等。今天, 随着可视化分析工具, 数据源以及研究范例越来越容易获得且兼容, 有一种力量正在积聚。科学前沿图谱已经迎来了新的起点, 存在前所未有的机遇, 它将跨越多种学科领域, 从而对科学研究活动产生更为广泛的影响。

《科学前沿图谱》第二版将会让您了解科学知识及其动态演化研究中一些最有意义的发现和进展。补充的内容如下:

第二章, 介绍了斯隆数字巡天计划, 以展示宇宙图谱的绘制。

第三章, 介绍了一系列主题演化图谱, 包括众所周知的 ThemeRiver, 使用简单的 TextFlow, 以及用途广泛的 Alluvial Maps。

第八章的内容是全新的。介绍了预测的分析框架, 并将其应用于再生医学研究领域, 其中诱导性多能干细胞 (iPSCs) 的研究因其突出的贡献而被授予 2012 年诺贝尔生理学 / 医学奖。在这一章中, 我们也讨论了论文撤销的现实意义。这一章的后半部分主要介绍科学全域图谱的设计、构建和分析, 包括我们新设计的“双图叠加”。

第九章的内容也是全新的。该章列出了一些最具代表性的可视化分析工具, 如 GeoTime 和 Jigsaw 等, 同时还描述了 CiteSpace 的主要分析功能。

在第一版结论中我们提出了 10 个挑战。我们有必要再重新审视 10 年前提出的这些问题, 看看有哪些变化, 还有什么新情况的出现。

在第二版的最后, 将提出科学前沿图谱面临的一些新的挑战 and 重要事件。

陈超美
于美国宾夕法尼亚州维拉诺瓦
2013 年 4 月 15 日

CONTENTS 目 录

译序	刘则渊
序	亨利·斯莫尔
自序	陈超美
第一章 科学知识的动态特征	1
1.1 科学前沿	2
1.1.1 竞争范式	6
1.1.2 无形学院	10
1.1.3 概念革命	11
1.1.4 追溯	17
1.2 视觉思维	20
1.2.1 格式塔理论	20
1.2.2 著名的地图	22
1.2.3 巴别塔	24
1.2.4 外层空间的信息	26
1.2.5 “这不是一个烟斗”	29
1.2.6 格式塔心理学	32
1.2.7 信息可视化与可视化分析	33
1.3 绘制科学前沿	38
1.3.1 科学图谱	39
1.3.2 竞争范式的案例	40
1.4 本书的结构	41

第二章 宇宙图谱	45
2.1 绘图学	47
2.1.1 主题图	50
2.1.2 地形图和照相绘图法	51
2.2 陆地地图	52
2.3 天体图	55
2.3.1 天体模型	55
2.3.2 星座图	60
2.3.3 宇宙图谱	63
2.4 生物图谱	74
2.4.1 DNA双螺旋结构	74
2.4.2 针灸图谱	76
2.4.3 基因组图谱	78
2.4.4 流感病毒蛋白质序列图	79
第三章 关联映射	81
3.1 关联的作用	82
3.1.1 我们可以试想	83
3.1.2 认知图的起源	84
3.1.3 信息可视化	87
3.2 结构识别	87
3.2.1 主题模型	88
3.2.2 寻径网络尺度	89
3.2.3 图像间相似性的测度	92
3.2.4 抽象结构的可视化	96
3.2.5 知识演进模式和趋势的可视化	102
3.3 降维	106
3.3.1 几何相似	106
3.3.2 多维尺度分析	107
3.3.3 个体差异分析	113
3.3.4 线性近似——等距映射	115
3.3.5 局部线性嵌入	118

3.4 概念图	119
3.4.1 卡片分类法	119
3.4.2 聚类	120
3.5 网络模型	122
3.5.1 六度分离理论	123
3.5.2 随机网络模型理论	124
3.5.3 厄多斯数	125
3.5.4 语义网络	126
3.5.5 网络可视化	127
3.6 小结	130
第四章 探寻轨迹	131
4.1 信息空间的足迹	132
4.1.1 旅行推销员问题	132
4.1.2 虚拟世界中的搜索	134
4.1.3 信息觅食	135
4.1.4 觅食过程建模	136
4.1.5 用户的轨迹	142
4.2 小结	147
第五章 科学知识的结构和动态	149
5.1 马太效应	151
5.2 词的图谱	154
5.2.1 共词图谱	154
5.2.2 包容指数和包容图谱	156
5.2.3 RISC个体发生学	157
5.3 共被引分析	159
5.3.1 文献共被引分析	160
5.3.2 作者共被引分析	167
5.4 HistCite	176
5.5 专利共被引	177
5.6 小结	179