



附赠
完整版电子光盘

汉语科技词系统

(新一代工业生物技术卷)



CHINESE SCIENTIFIC & TECHNICAL
VOCABULARY SYSTEM

NEXT GENERATION
INDUSTRIAL BIOTECHNOLOGY

■ 中国科学技术信息研究所 编著



科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

“十一五”国家科技支撑计划项目“知识组织系统的集成及服务体系研究与实现”(2006BAH03B03)

中国科学技术信息研究所重点工作项目“汉语科技词系统建设与应用工程”(2008KP01-3-1, 2009KP01-3-2, ZD2010-3-2, ZD2011-3-2, ZD2012-3-2)

汉语科技词系统

(新一代工业生物技术卷)

中国科学技术信息研究所 编著

 科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

· 北京 ·

图书在版编目 (CIP) 数据

汉语科技词系统. 新一代工业生物技术卷 / 中国科学技术信息研究所编著. — 北京: 科学技术文献出版社, 2014. 5

ISBN 978-7-5023-8577-4

I. ①汉… II. ①中… III. ①科学名词—检索系统 ②生物工程—名词术语—检索系统
IV. ①G354

中国版本图书馆 CIP 数据核字 (2014) 第 008449 号

汉语科技词系统 (新一代工业生物技术卷)

策划编辑: 周国臻 责任编辑: 周国臻 张 丹 责任校对: 张叫咪 责任出版: 张志平

出 版 者 科学技术文献出版社
地 址 北京市复兴路15号 邮编 100038
编 务 部 (010) 58882938, 58882087 (传真)
发 行 部 (010) 58882868, 58882874 (传真)
邮 购 部 (010) 58882873
官 方 网 址 www.stdp.com.cn
发 行 者 科学技术文献出版社发行 全国各地新华书店经销
印 刷 者 大恒数码印刷(北京)有限公司
版 次 2014年5月第1版 2014年5月第1次印刷
开 本 787×1092 1/16
字 数 454千
印 张 20
书 号 ISBN 978-7-5023-8577-4
定 价 158.00元 (附1张CD)



版权所有 违法必究

购买本社图书, 凡字迹不清、缺页、倒页、脱页者, 本社发行部负责调换

《汉语科技词系统》指导委员会

主任委员 郑国安

委 员 张晓原 袁海波 吴学梯 陈家昌 武夷山 沈仲祺 吴波尔
孙增奇 马晋并

《汉语科技词系统》编写委员会

主任委员 贺德方

委 员 陈家昌 乔晓东 蒋勇青 钱起霖 赵建华 马张华 侯汉清
张 旭 曾建勋 郑彦宁 姚长青 朱礼军 张运良

《汉语科技词系统（新一代工业生物技术卷）》编写组

主 编 贺德方

副 主 编 乔晓东 朱礼军 张运良 周 群 徐国兴 王丽华

参编人员 李 颖 许德山 张兆锋 钞亚鹏 王国业 黄健毅 桂 婕
李 鹏 张 东 于海燕 张 凌 李 楠 闫莹莹 牛 芳
刘亚洁 蔡配平 王 晨 徐 硕 刘 耀 张均胜 席 敏
张寅生 韩红旗 李琳娜 曾 文 潘继财 何彦青 高影繁
任智军 史庆伟 农国武 石崇德 徐红姣 屈 鹏 于 薇
赵 婧 剧伟伟 杨纷纷

前 言

在科技创新活动过程中，科研人员选题策划、立项评审、研发实施、结题验收、成果传播和成果评价都需要科技信息的支撑。面对大规模增长的科技信息资源，科研人员对及时、准确地获取信息的需求日益强烈，也对科技信息工作提出了新的、更高的要求，信息资源加工、组织和服务的理论方法及技术亟须创新。

科技信息工作的开展离不开信息技术尤其是知识技术的进步。知识组织系统是知识技术的核心，是提高各类信息资源开发利用效率的重要工具，也是推进知识服务的关键基础。知识组织系统包括图书分类法、主题词表、本体、词系统等多种类型。中国科学技术信息研究所（以下简称“中信所”）作为科技部直属的国家级科技信息机构，在知识组织系统研究和建设的不同历史阶段做了大量工作，曾先后牵头组织编制了《汉语主题词表》和《电子政务主题词表》，并主持制定了《汉语叙词表编制规则》（GB/T 13190—1991）、《文献多语种叙词表编制规则》（GB/T 15417—1994）、《电子政务主题词表编制规则》（GB/T 19486—2004）等国家标准。

当前，海量的网络化数字化信息资源使得知识组织系统产生了强烈变革需求，为了适应更加深入的语义应用和大规模信息即时处理的需求，知识组织系统不但更注重实用化和针对性，而且在某些局部变得日益丰富细化。中信所正是在这样的背景下，进一步加强了知识组织系统相关研究工作，针对科研活动跨学科特点，提出以领域为视角构建全面反映各学科知识相互渗透、交叉融合现状的汉语科技词系统。这是一种新型的知识组织系统，其本质是通过概念、概念内涵和概念之间的关系来构建领域知识体系，并通过相应算法对海量科技信息进行语义分析，实现科技信息的深度加工与服务。

汉语科技词系统从萌芽到发展，从理论到实践，前后历经五载。在研究开

发过程中,中信所及其合作单位做了大量的研究论证和试验,努力解决了汉语科技词系统的理论研究、工具开发等环节中的一些关键性问题,包括领域概念知识的整理加工、概念知识的构建、管理和服务平台系统搭建等。2010年以来,中信所先后开展了新能源汽车、新一代工业生物技术、智能材料与结构技术、重大自然灾害监测与防御、新能源等领域汉语科技词系统的建设。同时,在汉语科技词系统的基础上还开展了科技监测与评价、专利内容深度分析、企业知识管理和移动知识服务等方面的应用实践。在研究实践过程中,中信所也积累了一定的经验和教训,希望它们能够为今后领域汉语科技词系统和其他类型的知识组织系统建设应用提供借鉴。

《汉语科技词系统》有纸质版和电子版。本书即纸质版,由于内容体量较大,本书只收录了一部分内容,完整的电子版内容可以从配套光盘或者官方网站(www.vocgrid.org)上获取。为保证对本书概念理解的一致性,对必要的核心词我们附加了定义解释。在此过程中,编写组参考了领域相关的工具书、标准、教科书等,以保证定义的准确性和权威性,在此一并向这些文献的作者、编者和出版单位表示诚挚的感谢。

《汉语科技词系统》的编写是一项非常艰巨的工作,在本书出版之际,感谢在汉语科技词系统研发和内容建设中付出大量心血的科研人员,同时也感谢一直关心和支持这项工作的同行专家。由于时间有限,工作量大,虽然几经审校,但不妥和错误之处在所难免,欢迎同行专家和广大读者批评指正,以便我们在后续的更新版本中修正、不断完善。

衷心希望本书的出版能够推动其他领域词系统及知识组织系统相关研究和实践的发展。汉语科技词系统作为一棵小小的幼苗已经开始抽枝展叶,希望在国内外同行专家的帮助和指点下,在不远的将来能够成长为一棵参天大树。

中国科学技术信息研究所所长 贺佳方

分卷前言

本书围绕新一代工业生物技术，比较完整地覆盖了工业生物技术乃至生物技术的各大类基础概念和应用领域，而且涵盖了当前生物学在应用研究中的热点词汇。将该领域的科技词汇分为工业生物技术、基因工程、细胞工程、蛋白质工程、发酵工程、酶工程、生化分离技术、医药、农业、化工、能源、生物学基础、生物工程设备共十三大类。新一代工业生物技术作为新兴和前沿的科技领域，正推动着一个以生物催化和生物转化为特征，以生物能源、生物材料、生物化工和生物冶金等为代表的庞大的现代工业王国的形成，以及一场以化石为原料的工业经济迈向生物质经济的现代工业技术革命的到来。因此，汉语科技词系统中新一代工业生物技术部分选词范围广泛，词汇类型覆盖全面，能够结合生物技术这一高速发展的行业特点，既能满足一般人员参考阅读，又能为科研工作者提供准确全面的解释。

21 世纪，人类面临化石资源不断枯竭、环境污染日益加剧的严重局面，转向以可再生生物资源为原料，可再生生物能源为动力，高科技含量、环境友好、过程高效的新一代物质加工模式是必然趋势，这种加工模式的核心技术就是工业生物技术（Industrial biotechnology）。我国应针对自身社会经济发展的战略需求，充分挖掘能源植物资源和极端环境微生物资源的特色，发挥传统工业生物技术产业的优势，把握新一代工业生物技术创新发展的机遇，形成可持续的自主科技创新能力，形成我国工业生物技术综合性研发基地和国际化企业集团，以此带动形成规模化经济示范区。

世界经合组织（OECD）指出：“工业生物技术是工业可持续发展最有希望的技术”。我国高度重视工业生物技术的发展，2005 年 9 月，由国家科技部中国生物技术发展中心组织的“首届国际生物经济高层论坛”在北京召开。在 2006 年颁布的《国家中长期科学和技术发展规划纲要（2010—2020 年）》中把“新

一代工业生物技术”作为“前沿技术”列入规划。当前生物催化和转化技术中具有代表性的研究热点包括以酶法生产甜味剂阿斯巴甜、L-苯丙酸、抗生素中间体 6-APA、D-对羟苯甘氨酸和 1,3-丙二醇等。作为生物材料的代表产品聚乳酸,目前已有多家公司大规模生产。2000 年全球以生物法生产的可再生生物能源——液体燃料,已近 1000 万吨。

虽然生物催化技术的前景非常广阔,但其发展还受到现有生物技术和研究水平的制约。中国工业生物技术是国际上发展较早、产业规模较大的领域之一。近年来,在微生物资源、基因组学、蛋白质组学、代谢工程、酶分子进化等研究领域呈现出快速成长的良好势头。产业技术水平与产品产量也呈快速增长趋势,如丙烯酰胺、谷氨酸、柠檬酸、维生素 C、青霉素等产品的产量已进入世界的前列;又如丙烯酰胺、维生素 C、黄原胶、1,6-二磷酸果糖、L-苹果酸等产品的技术水平已达国际先进或领先;在生物能源方面也取得了较大的进展,如目前我国酒精产量为 300 多万吨,仅次于巴西、美国而位列世界第三。但是,我国生物技术的研究和产业技术总体水平与世界先进水平仍有较大的差距,国外代谢工程菌的工业化成功例子虽然不多,但在我国尚无一例代谢工程菌株进入中试试验。由于国外已积累了较丰富的构建及使用代谢工程菌株的经验,未来 5 年内预期将会有一批产品至少在中试规模用工程菌生产。我国相关大学和研究机构虽已开展了一些微生物功能基因组、代谢工程等的研究,但大多数研究仍停留在前期的单基因操作阶段,而生物能源、生物材料和精细化学品的生物制造的研究和技术水平落后于发达国家。无论是原始创新的基础研究,还是技术创新性改造,整体水平都落后于发达国家,因此急需加强与工业生物技术相关的应用基础和技术开发的研究。

新一代工业生物技术的属性类型包括功用、效应、方法手段、时间属性、物理特性、特征、状况、目的、空间属性、约束等 10 种一级属性类型,一级属性又进一步细分为二级子属性类型。子属性类型划分较为细致全面,能够比较完整地覆盖词汇的概念外延,侧面说明了词汇的用途、功能和特点。编写本书时,核心词的选词领域广泛,包括了有机物、酶类、昆虫、生物化学和分子生物学、工业技术、激素类、生物技术等几方面词汇,是上述一级类型的代表词汇;核心词的选择层面广泛,包括了基础理论、技术名词、实践应用等方面。

核心词词汇的定义涉及英语汉译的问题，译文多来源于《英汉生物学大词典》，来源权威，为广大专业人员普遍认可；其次来源于生物学专著和百度百科，如百度科技词定义等。希望本书能够为完善科技词汇系统，促进我国新一代工业生物技术进一步发展做出贡献，同时为国内从事相关领域的研究者提供借鉴和参考。

中国科学院院士

中国生物工程学会副理事长

高福

中国科学院微生物研究所原所长（2004—2008）、研究员

目 录

第一部分 词系统背景知识	1
1 汉语科技词系统的发展	3
2 汉语科技词系统的知识内容配置	4
3 汉语科技词系统相关平台工具进展	5
3.1 汉语科技词系统协同构建平台改进	5
3.2 汉语科技词系统语料库平台建设	5
4 汉语科技词系统的建设流程变化	7
4.1 汉语科技词系统的人员配置建议	7
4.2 汉语科技词系统的建设流程建议	8
4.3 汉语科技词系统语料库平台使用	9
5 汉语科技词系统的领域扩展尝试	10
6 汉语科技词系统发展中的争议问题讨论	11
6.1 知识的权威性	11
6.2 知识的准确性和客观性	11
6.3 同义词环引入词系统的分析	12
7 汉语科技词系统的 SKOS 格式转换	13
7.1 SKOS 格式转换	13
7.2 导出格式中简化处理的说明	16
第二部分 新一代工业生物技术领域汉语科技词系统的建设	17
1 领域选择	19
1.1 农业方面	19
1.2 食品工业方面	19

1.3 医药医疗方面	20
1.4 能源领域	20
1.5 生物材料开发方面	21
1.6 环保领域	21
2 生物技术的发展阶段	22
2.1 传统生物技术阶段	22
2.2 工业化生物技术阶段	22
2.3 现代生物技术阶段	22
3 新一代工业生物技术	24
3.1 新一代工业生物技术的内涵	24
3.2 新一代工业生物技术核心内容	24
3.3 新一代工业生物技术在我国的应用现状	25
4 新一代工业生物技术词系统的建设	27
4.1 领域界定	27
4.2 核心词选择	27
4.3 关系和属性确定	27
4.4 关系和属性统计	29
第三部分 新一代工业生物技术领域汉语科技词系统实例	31
1 格式说明	33
2 实例正文	35
第四部分 附 录	245
A 实例词条音序索引	247
B 实例词条笔画索引	252
C 工业生物技术分类法	257
D 全部核心词列表(4994 条)	259
E 示例中定义来源参考文献	303
《汉语科技词系统(新一代工业生物技术卷)》完整电子版光盘	308

第一部分

词系统背景知识

1 汉语科技词系统的发展

继新能源汽车领域汉语科技词系统之后，新一代工业生物技术、智能材料与结构技术、重大自然灾害监测与防御、新能源等四个领域汉语科技词系统的建设也陆续完成。经过这五个领域汉语科技词系统的建设，建设思路和方法已经比较成熟，完成了词系统建设之初所设定的目标。为此，中国科学技术信息研究所(以下简称“中信所”)调整了词系统的发展战略，不再以自身为主单独启动词系统建设项目，而是结合其他项目综合需要或者根据合作单位的需求，开展建设任务或者建设咨询工作。近期，中信所开展了电动汽车领域科技词系统建设，也根据合作单位的需求，先后启动铝业、金融银行、航空等领域词系统建设的论证工作。这些领域词系统的建设有可能陆续开展。在相关工作开展过程中，中信所的研究人员先后多次与多家企业和科研机构的有关人员进行了深入细致的交流，了解业务需求，拓宽发展思路。

回顾汉语科技词系统近两年的发展，我们主要有以下三点的认识。第一，汉语科技词系统等知识组织系统的建设有巨大的现实需求，主要集中在企业知识管理和竞争情报方面。当然，不同行业不同领域的应用需求有一定差异，如有的需求强调同义词，有的需求强调主题分类一体化。不同领域涉及和关注的实体也会有所差别，如有的需要增加组织机构，有的需要体现政策方针。第二，汉语科技词系统作为一种典型的知识组织系统，具有一定的领域适应性。而且，经过过去的建设实践，相关的流程方法进一步规范和成熟，工具平台也进一步完善。可以预见，今后构建词系统的时间将进一步缩短，成本将进一步下降。第三，汉语词系统也不一定局限在科技领域。在其他领域，只要是涉及知识的组织、管理和服务，就会有词系统的一席之地。由于汉语科技词系统的知识类型是灵活的，只要配置得当，词系统就能发挥作用。

2 汉语科技词系统的知识内容配置

汉语科技词系统包含的各项知识内容,由于需求本身的差异,以及建设难易程度和对建设者的要求不同,所以不能也不应该在所有的领域都求全责备,这里涉及一个词系统知识内容灵活配置的问题。随着对不同领域词系统的建设实践及其研究,在建设过程中我们对相应的知识内容配置也可以做一些调整。在五个已经完成的词系统中,必备的部分包括核心词的英文翻译、核心词的定义、以核心词为中心的关系和属性、核心词和领域粗分类之间的映射关系。当然,这并不意味着这些内容是不可或缺的。

此外,尽管不同领域的知识内容配置的组成类似,但是在配置的比例上可能存在较为显著的差异。例如,比较已经完成的两个领域汉语科技词系统,就会发现:新一代工业生物技术领域关系实例数量较少而属性实例数量较多,而重大自然灾害监测与防御领域关系实例数量较多而属性实例较少。

3 汉语科技词系统相关平台工具进展

3.1 汉语科技词系统协同构建平台改进

随着汉语科技词系统各领域的建设,协同构建平台系统出现的一些问题已修改完善,其中典型的就检索结果的分页浏览,以及对检索结果的排序处理。另外,在功能上也有少量的增加,主要有以下两点:一是在属性和定义添加中,允许增加图片等辅助说明材料,以富文本(Rich Text)方式进行编辑(如图1-1所示);二是从编辑和审核细节角度进行了一些改进,如可以补充是否对删除信息进行检索查询,以满足对保留和删除状态的知识进行检索的个性需求。



图 1-1 添加属性界面

3.2 汉语科技词系统语料库平台建设

本语料库平台是为了配合汉语科技词系统建设流程而专门开发的。开发的核心思想是利用相关领域的期刊论文、会议论文、学位论文、专利等科技文献进行计算分析,提供对选词和关联关系构建的统计支持。当然,本语料库平台也可以基于企事业单位的自有资源进行有针对性的分析,只需对相关资源的元数据做一定的转换处理。

本平台具有三项前台功能和四项后台功能。前台功能包括关键词词频统计、关键词共现和语料全文检索。关键词词频统计,主要从文献中提取关键词,并对词频做分析;分析可以从所属领域、语料类型、出现位置、时间范围、出现频率等五个角度进行,从而实现初步的选词功能。关键词共现则是在选定的词条基础上(包括机器筛选的和人工修订完善的)进行两两共现分析,从而支持关系构建。共现类型分为关键词共现、句内共现和文献内共现三种,可以根据需要为不同的共现类型赋予不同的权重。全文检索基于 Lucene 构建,主要用于构建过程中临时的共现分析。这是因为知识工程师在知识构建过程中,总会引入一些已经筛选好的词表中没有的词,这些词与已有的词的共现信息并没有预先计算,需要以全文检索来补充。后台功能包括领域范围管理、语料管理、专业词典维护和数据计算。领域范围管理用于管理语料库平台中多个领域的增加、删除和修改;语料管理主要用于浏览、增加和删除各领域的语料;专业词典维护可以将人工修订的筛选词导入系统,使得系统能够在这基础上进行共现计算;数据计算主要分为三个子功能,分别是索引建立、词频统计和共现计算,还有一些辅助计算功能。