

张琪玉情报语言学文集

YUYANXUE WENJI

ZHANG QI YU QINGBAO



北京图书馆出版社

张琪玉情报语言学文集

北京图书馆出版社

图书在版编目(CIP)数据

张琪玉情报语言学文集/张琪玉著. - 北京:北京图书馆出版社, 1999.5

ISBN 7-5013-1591-4

I. 张… II. 张… III. 检索语言-文集 IV. G254.0-53

中国版本图书馆 CIP 数据核字(1999)第 09410 号

书名 张琪玉情报语言学文集

著者 张琪玉 著

出版发行 北京图书馆出版社(原书目文献出版社)
(100034 北京西城区文津街7号)

经销 新华书店

印刷 北京市双桥咸宁侯印刷厂

开本 850×1168 毫米 1/32

印张 15.3

字数 380(千字)

版次 1999 年 5 月第 1 版 1999 年 5 月第 1 次印刷

印数 1—3000

书号 ISBN 7-5013-1591-4/G·432

定价 30 元

目 录

(一)按著作内容分类顺序

张琪玉与情报语言学学科建设(代自序)·····	(1)
-------------------------	-----

■情报语言学总论

论情报检索语言的研究、创制与普及·····	(10)
情报检索语言·····	(21)
情报语言学基础·····	(50)
情报检索语言·分类检索语言·主题检索语言·····	(76)
情报语言学文献分类表·····	(80)
情报检索中的语言保证问题·····	(103)
检索效率及其影响因素·····	(112)
情报检索语言的发展趋势(与吴建中的对话)·····	(118)
世纪之交中国情报语言学发展之路·····	(124)
情报检索语言方法综述·····	(131)
情报检索语言原理的一致和方法的差异·····	(145)
情报检索语言语法体系初探·····	(152)
情报检索语言中聚类的原理和方法·····	(160)
组配及其演变·····	(174)
情报检索语言的国家特点、时代特点和自然语言特点 ·····	(184)
学科-事物概念组配型检索语言——关于情报检索语言 的遐想与求索·····	(187)

探索 21 世纪的情报检索语言	(194)
文献主题的构成因素及层次	(200)
试论隐含主题	(205)
情报检索与用户教育	(212)

■分类检索语言

分类法与分类表	(220)
体系分类法的准则和惯例	(229)
体系分类法中“集中与分散”的矛盾	(255)
关于《中图法》增加组配成分的可能性和方法的探索	(263)
体系分类法中的交替法	(276)
分类标记原理与方法概述	(285)
分类目录产生漏检和误检的原因	(295)

■主题检索语言

主题法与词表	(300)
情报检索语言中语词标识的功能与局限—关于主题法性 能的几点分析	(308)
充分利用入口词原理	(325)
汉语检索词词素轮排索引编制法探索	(331)
主题标引的原理和方法	(340)

■自然语言在情报检索中的应用

自然语言在情报检索中的应用	(358)
---------------------	-------

自然语言检索中各种因素对检索效率的影响·····	(364)
自然语言与人工语言对应转换——情报检索语言走向 自动化之路·····	(372)
分类法主题法一体化自动标引系统的基本原理和方法 ·····	(379)
人—机结合的题内关键词索引可回避汉语分词难题 ·····	(386)
论自由标引·····	(389)
论后控制词表·····	(395)

■索引学

关于索引学研究和索引工作开展的设想与建议·····	(403)
论索引项·····	(414)
工具书功能索引——关于编制“工具书之工具书”的设想 ·····	(420)
目录的特征——谈谈充分发挥书目检索工具的作用 ·····	(427)

■其他

钟情事业就是走向成功之路·····	(431)
张琪玉情报语言学著作目录·····	(445)

目 录

(二)按著作完成时间顺序

目录的特征——谈谈充分发挥书目检索工具的作用	1964□427
情报检索语言	1980□ 21
分类目录产生漏检和误检的原因	1980□295
体系分类法中“集中与分散”的矛盾	1981□255
体系分类法中的交替法	1982□276
论情报检索语言的研究、创制与普及	1983□ 10
情报语言学文献分类表	1983□ 80
情报检索语言方法综述	1984□131
文献主题的构成因素及层次	1984□200
情报检索语言中语词标识的功能与局限——关于主题法性能的几点分析	1984□308
关于《中图法》增加组配成分的可能性和方法的探索	1985□263
情报检索语言语法体系初探	1986□152
情报检索语言·分类检索语言·主题检索语言	1987□ 76
情报语言学基础	1987□ 50
情报检索语言的国家特点、时代特点和自然语言特点	1989□184
工具书功能索引——关于编制“工具书之工具书”的设想	1991□420

体系分类法的准则和惯例·····	1992□229
分类标记原理与方法概述·····	1992□285
汉语检索词词素轮排索引编制法探索·····	1992□331
充分利用入口词原理·····	1992□325
试论隐含主题·····	1992□205
关于索引学研究和索引工作开展的设想与建议 ·····	1992□403
人—机结合的题内关键词索引可回避汉语分词难题 ·····	1993□386
论后控制词表·····	1993□395
论索引项·····	1994□414
情报检索语言原理的一致和方法的差异·····	1994□145
情报检索中的语言保证问题·····	1994□103
检索效率及其影响因素·····	1994□112
分类法与分类表·····	1995□220
主题法与词表·····	1995□300
论自由标引·····	1995□389
分类法主题法一体化自动标引系统的基本原理和 方法·····	1995□379
钟情事业就是走向成功之路·····	1995□431
主题标引的原理和方法·····	1995□340
自然语言与人工语言对应转换——情报检索语言 走向自动化之路·····	1995□372
组配及其演变·····	1995□174
自然语言在情报检索中的应用·····	1995□358
情报检索语言的发展趋势(与吴建中的对话)·····	1996□118
情报检索与用户教育·····	1996□212
情报检索语言中聚类的原理和方法·····	1996□160

张琪玉情报语言学著作目录·····	1996□445
自然语言检索中各种因素对检索效率的影响·····	1996□364
学科-事物概念组配型检索语言——关于情报检索 语言的遐想与求索·····	1997□187
探索 21 世纪的情报检索语言 ·····	1997□194
世纪之交中国情报语言学发展之路·····	1997□124
张琪玉与情报语言学学科建设·····	1998□ 1

张琪玉与情报语言学学科建设

(代自序)

注:本文原为黑龙江教育出版社《中国当代社科精华》丛书文化教育卷的约稿,体例亦按编辑部要求。

人物简介

张琪玉,68岁(1930年6月7日生),上海南汇人。北京大学图书馆学系1954年7月毕业。职称教授。现在空军政治学院信息管理系任职。曾任武汉大学图书情报学院教授、图书馆学情报学研究所所长、空军政治学院图书档案系主任等。在学术活动方面,现任中国图书馆学会理事、中国索引学会副理事长暨学术研究委员会主任、上海市图书馆学会副理事长、世界知识组织学会会员。

主要研究领域是情报语言学,是该学科领域的开拓者。曾因在编制《中国图书馆图书分类法》中作出较大贡献而获1985年国家科学技术进步奖一等奖,因在编制《中国分类主题词表》中作出的贡献而获1996年国家优秀科技信息成果奖二等奖,1991年7月获国务院颁发的政府特殊津贴。简传曾载入《中国大百科全书》(图书馆学情报学档案学卷)、《世界名人录》(中国卷)等。

学术轨迹、理论成果和学术观点(自述)

我的主要研究领域是情报语言学。这是图书馆学和情报学的

共同分支学科。其主要研究对象是情报检索语言,同时也研究自然语言在情报检索中的应用问题。

我的学术观点和治学道路,写于1983年8月30日的《论情报检索语言的研究、创制与普及》一文最能作说明。我在该文中说:“情报检索语言的职能,是在各种情报检索系统中起‘语言保证’作用,使标引人员对文献情报内容的表达(标引用语)和检索人员对相同内容的情报需要的表达(检索用语)取得一致,以减少情报在存贮—检索过程中的损失,达到较佳的检索效果。保证较高的检索效率,这是对情报检索语言的基本要求;而提高检索效率,则是情报检索语言理论研究工作的根本目的。可是,从过去很长一段时期内的情况看,我们对情报检索语言理论研究工作的这个根本目的并不是很清晰的……我们的注意力大部分都用在体系分类法的思想性方面了,忽视了它更基本的是文献检索的语言工具这一职能”,因此提出“我们的情报检索语言理论研究工作必须开辟新路”:

首先,要改变研究方向,把研究的重点转移到如何提高情报检索语言的检索效率方面来;

其次,要扩大研究范围。研究范围的偏窄,对解决我国的情报检索问题极为不利。目前,要特别重视研究情报检索语言中的组配原理,研究组配型的分类法和主题法。

关于扩大研究范围,似乎也应把对自然语言检索法的研究包括进去。电子计算机检索技术的发展,为自然语言检索法开拓了光明的前景。自然语言检索法最终是否能取代以及何时能取代人工语言即情报检索语言,目前下结论还为时过早。但是,人工语言和自然语言相结合,既是严密的人工语言,又尽量利用自然语言作为辅助,无疑有助于检索效率的提高,因而是情报检索语言发展的正确方向;

第三,要改进研究方法。在该文中所列举的研究方法有结构

功能分析法、历史演进研究法、比较研究法,理想语言设计法、现用语言改进法、原理或方法的移植法、计算机方法、数学方法和统计方法等;

第四,要吸取国外的研究成果,但不要以此代替对中国情报检索语言特殊问题的研究。

我的学术研究轨迹,一直没有偏离这个方向。于1980年3月写成的《情报检索语言》一书,是循着上述学术思想的第一次探索,并构成了情报语言学的初步框架。该书对情报检索语言(这种人工语言的历史在我国几乎可追溯到公元前7~5年的《七略》分类体系)的论述给人有璨然一新的感觉,被誉为“独辟蹊径”,“创一家之说”。通过在高等学校的教学和该书在社会上的传播,我的学术观点逐渐得到了专业学术界的普遍认同。其实所谓“独辟蹊径”,不过是我第一个跳出当时毫无结果的“三性”(思想性、科学性、实用性及其相互关系)争论的怪圈,第一个按情报检索语言的本质属性来研究和阐述情报检索语言而已,这条“蹊径”是迟早会被人发现的。但当我朝着这个方向跨了出去之后,我看到了,前面是一片广阔的天地,那里地面上到处散布着矿苗,大有开发的前途,只要辛勤勘探和采掘,肯定会有更多成果的。所以,我二十年来一直在这片土地上勘探和采掘,获得了较多的成果。

我所做的工作,犹如在植物学和动物学的基础上建立生物学。植物学和动物学虽是生物学的分支学科,但他们并不是从生物学中分化出来的。它们在生物学形成之前,本是两门独立的学科,后来发现它们的研究对象有统一的规律,而到19世纪末才产生统一这两门学科的生物学。我也发现,分类法也好,主题法也好,其他文献内容的检索方法也好,它们都是情报检索系统的组成部分,都是在寻求更佳的检索效果中创制出来的,都是表达一系列概括文献情报内容的概念及其相互关系的概念标识系统,其职能是作为情报检索系统的语言保证,其核心问题是检索效率。所以,它们的

基本原理是一致的,只是它们在表达各种概念及其相互关系时,和在解决对它们提出的那些共同要求时所采取的方法不同,才形成了不同的类型和语种。在情报检索语言的概念下,对它们进行综合研究,可以找出它们之间最本质的东西,以及结构和功能上的相同或相异之处,概括出它们影响检索效率的共同规律,以及有效地改进和创新的途径。我正是在这种认识指导下,对情报语言学的各个方面都进行了研究和探索,所取得的成果充实了情报语言学的学科理论体系。

情报语言学理论从更高更宽的角度和更深的层次来考察情报检索语言,从而对各种情报检索语言和自然语言检索法的研究更为透彻,提供了对它们进行改进和创新的明确道路。

我在《情报语言学基础》第一版最后一章中,根据自己的考察,综述了计算机技术的发展对情报语言学的重大影响,即:50年代开始的情报检索计算机化,促进了情报检索语言的创新和改造,使词表、分类表向机编化和机读化方向发展,使文献标引和索引编制走向自动化,使自然语言检索得以实现,使多种语言的结合使用成为可能,使检索方法有了很大的进步,并正在使情报检索语言的应用范围扩大(例如开始应用于情报研究和文献计量)。特别是自然语言在情报检索中的应用,使情报检索系统的语言不再局限于情报检索语言。同时,情报检索计算机化对情报检索语言研究提出了许多新课题,并提供了许多新方法和新条件。总之,情报检索计算机化对情报语言学的发展产生了极为深刻的影响。甚至可以说,情报检索计算机化是加速情报语言学形成过程的一个重要因素。可以预见,情报检索计算机化今后将会更快、更广阔、更深入地得到发展,情报语言学也将会有更快的进步。因此,我认为,在这世纪之交,在计算机检索正在逐步发展到互联网络阶段的新形势下,情报语言学研究的基本课题应是如何使情报检索语言适应新的检索环境。

对于自然语言在情报检索中的应用即自然语言检索法,我一直认为,自然语言检索系统与情报检索语言检索系统并不是绝然对立的,它们各有长处和短处,可以并行发展,可以互相结合,互相补充。对于自然语言检索效率全面或总体高于情报检索语言,自然语言将替代情报检索语言或最终将替代情报检索语言,情报检索语言或情报检索语言研究已经过时的论点,我都并不以为然。因为那些论点如果是正确的话,全世界的情报检索早已全面自然语言化了。而过去了十多年,这种情况并未发生,说明那些论点虽然振奋人心但是言过其实。但我并没有忽视自然语言检索法,我认为自然语言目前还处于其发展的初级阶段,它并非十全十美,尚有许多地方有待改进,在计算机检索日益发展的条件下,自然语言具有不可阻挡的发展前途。特别是在互联网的检索环境中,它将成为一种必然的优先选择。另一方面,对于高要求的情报检索来说,控制是绝对必要的。而“对检索过程进行控制”,正是情报语言学的精髓。我们应当把情报语言学的理论成果运用到自然语言方面去。所以我认为,当前亟待从情报语言学角度深入研究自然语言,情报检索语言研究者应当积极参与自然语言检索法的研究。正是基于这种认识,我近年来的研究重点转向了自然语言。

至于自然语言或情报检索语言的未来,我认为将是自然语言的情报检索语言化或情报检索语言的自然语言化。未来的检索语言将既是情报检索语言,但不是现今的情报检索语言模式;它既是自然语言,但也不是现今的自然语言检索模式;它既具有自然语言的优点而优于现今的情报检索语言,又具有情报检索语言的优点而优于现今的自然语言。

我所以能在情报语言学研究取得较多成果,在很大程度上是因为找到了合适的研究方法。我所使用的主要研究方法是结构功能分析法。在本文开始时所列举的多种研究方法,可以说都是从结构功能分析法衍生出来的。脱离了结构功能分析法,就等于

抽去了那些方法的精髓。我正是利用了那些研究方法,找出了各种各样情报检索语言的许多共同规律,使其融会贯通,构成情报语言学理论体系。

十多年来,我一直遐想着运用理想语言设计法,创制一种完全新颖的情报检索语言结构模式,这种情报检索语言模式包含了现今人们对情报检索语言的绝大部分要求。这种理想的情报检索语言应是:学科聚类系统与事物聚类系统的结合(事物聚类也应当有系统性,字顺序列可作为进入事物聚类系统的手段),先组式语言与后组式语言的结合,体系分类法与组配分类法的结合,人工语言与自然语言的结合,号码标识与语词标识的结合,系统序列与字顺序列的结合,不变概念代码与可变概念体系的结合。

这种情报检索语言应是:分类法与主题法彻底一体化的,充分发挥情报检索语言对知识进行系统组织和对自然语言进行规范控制的功能的,用户可十分方便地进行标引和检索的,概念可不断增补及概念的代表词可进行更换的,用户区别不出是自然语言还是人工语言而其实是由严密的人工语言控制的,修订不受已标引文献所牵制,故分类体系可逐步完善的,并可以挂接英文索引、分子式索引等以及可用于机助标引的。

应当说,构成这种理想情报检索语言性能的原理和方法都已存在,难题仅在于找到它们的结合方案。我一直未找到达到上述要求的方案。但在去年年初终于找到了,那主要就是“方面分析+概念代码+概念对应转换+数据库技术”。我把这种模式命名为“学科-事物概念组配型检索语言”,其具体原理和方法见本文所附的主要论著(5)和《北京大学学报:信息管理系建系五十周年专刊》(1997年)所载《探索21世纪的情报检索语言》一文,在此不再详细叙述。

索引是情报语言学研究对象的应用场合,所以我对索引学也有一定研究。所写文章较满意的,有本文所附主要论著(9)和《工具

书功能索引——关于编制“工具书之工具书”的设想》(载《图书馆杂志》1991年第2期),《关于索引学研究和索引工作开展的设想和建议》(载《江苏图书馆学报》1993年第1期),《推广文献索引计算机编制法是促进我国索引事业发展的一项重要措施》(载《图书与情报》1996年第4期)等文。

我所构建的学科命名为“情报语言学”,被认为是“中国人自己创建的学科”。在《中国大百科全书》图书馆学情报学档案学卷的概观性文章《图书馆学》(周文骏、金恩晖著)中指出:“70年代末,中国学者从情报检索角度对分类法和主题法以及代码检索语言、自然检索语言等进行综合比较研究,探索它们影响情报检索效率的规律和提高检索效率的途径,并以此建立情报语言学这一新的分支学科”。

范并思在《论图书馆学学科前沿的转移》一文中写道:“与图书馆现代化相比,检索语言研究有更多的创新意义。图书馆现代化研究是学习西方建立起来的领域,而检索语言研究却基本属于中国人自己创立的领域。在张琪玉以前,国外还没有人能“在‘情报检索语言’的书名下将检索语言理论的内容讲述得如此系统而精密。检索语言领域吸引了一批非常勤奋且治学严谨的理论家,除张琪玉外,中年人有刘湘生、丘峰、侯汉清,青年人有曾蕾、戴维民等,他们使检索语言研究成为80年代图书馆学内最活跃、最具有创造性的理论前沿之一。”(见《图书馆》1993年4期4页)。

下面摘录几段评论,说明创立情报语言学的学术影响:

吴建中:“80年代初,张教授针对当时人们把注意力集中在体系分类法的思想性方面,使图书分类研究陷入所谓‘三性(思想性、科学性、实用性)’怪圈问题时指出:‘要改变研究方向,把研究的重点转移到如何提高情报检索语言的检索效率方面来’,并独辟新路,开创‘情报语言学’,提出从更高的层次对分类法、主题法等各种检索方法进行统一研究。张教授这一开拓性的研究,对我国情

报检索语言领域的理论与实践起到了积极的导向和推动作用。”(见《21世纪图书馆展望——访谈录》,上海科学技术文献出版社,1996年7月,第91页)。

倪波:“张琪玉教授是我国图书情报界知名学者,他在情报检索语言研究领域内有着不可磨灭的贡献。”(同上书98页)。

曾蕾:“近年来,我国在这方面出现了一次巨大的飞跃,其特点反映在以下方面……第二,对检索语言的系统研究成果突出。继刘国钧先生的《现代西方主要图书分类法评述》一书之后,以张琪玉教授编著的《情报检索语言》教材为代表,由他本人及众多研究者所取得的成果,使我国在检索语言的研究水平方面迅速接近世界水平。”(《关于近年来国内外情报检索语言研究方向与方法的比较研究》,载《图书情报知识》1988年3期17页)。

1980~1997年期间,《情报检索语言》及其增补修订本《情报语言学基础》以5种版本累计印刷11次,共99900册,被288位作者的专著和论文共引用了约450次。作为教材,在图书馆学、情报学的本科、函授和电大等教学中被广泛采用,某些档案学专业也开设情报检索语言课程,估计学习过这门课程的不少于25000人。我的其他情报语言学专著和论文也被引400多次。我编著、主编、参编的专著有20多种,发表论文译文约160篇,论文被各种文集等转载约60篇次,得到了比较广泛的传播,对我国情报语言学研究及情报检索语言创制和文献标引检索实践产生了较大的影响。

我在情报语言学领域艰苦治学20年,作为一名中国学者,能够对世界关于“知识组织”的知识领域添加一份贡献,感到高兴。

主要论著

- (1)《情报检索语言》(专著,武汉大学出版社1983年6月出版;增补修订本改名《情报语言学基础》,武汉大学出版社1987年11月出版,后者增订二版于1997年9月出版);