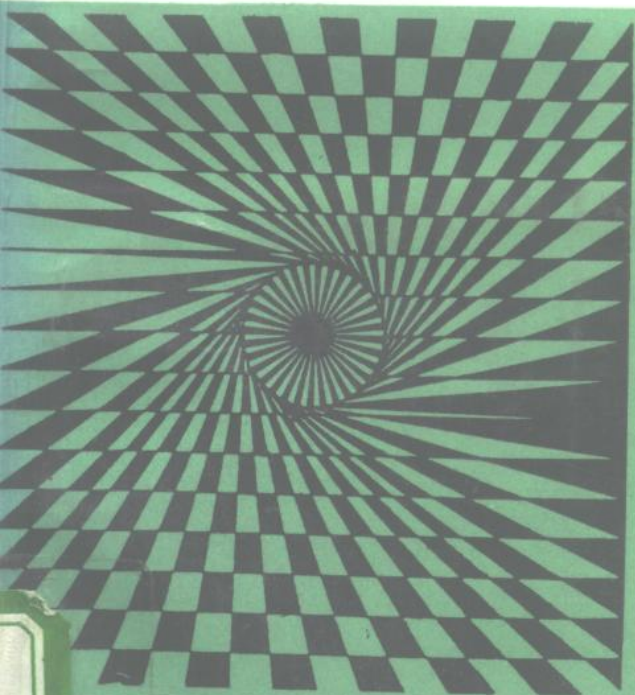


情报组织 与利用

书目文献出版社



●〔美〕达格伯特·索尔格尔 著
● 谢新洲 陈 睿 常 彤 译

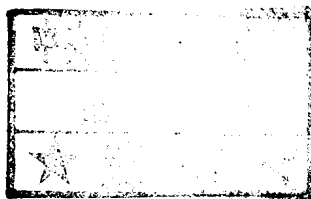
03421

情报组织与利用

(美)达格伯特·索尔格尔 著
谢新洲 陈睿 常彤 译



200079020



书目文献出版社

(京)新登字189号

情报组织与利用

(美)达格伯特·索尔格尔 著

谢新洲 陈睿 常彤 译

书目文献出版社

(北京文津街七号)

北京顺义牛富屯福利印刷厂排版

涿州市新华印刷厂印刷

书目文献出版社发行 新华书店经销

787×1092毫米 32开本15.25印张 324千字

1992年11月北京第1版 1992年11月北京第1次印刷

印数: 1—4000册

ISBN 7-5013-0939-6

G·249 定价: 11.50元

前 言

本书绘出了情报系统的分析、设计及实现的理论基础和前景，尤其是情报系统的情报存贮与检索(ISAR)这一组成部分，并对手工系统和机械系统都进行了论述。情报系统要处理许多类型的问题：事件、人、文献、商业事务、博物馆展品、研究计划等。情报系统提供服务的目的是将信息提供给大众，帮助管理人员、研究人员和工程师，并为人工智能项目提供知识基础。本书中讨论的原理都适用于上述环境。通过描述两个交叉领域的观点，即数据库管理和图书馆中知识的组织与利用，并将这些观点加以综合，成为一个统一的机体，本书得出了这些普遍性原理。这些原理也适用于新系统的设计，尤其适用于现存系统的分析，以便进一步开发其潜力，克服其缺点和进行改进。

本书是一本关于组织和检索情报的导论性教程（例如，可称之为“情报组织导论”，“情报存贮与检索导论”或“情报学导论”），可用于图书情报学院、商学院，或用于更广泛的基层情报研究和情报管理机构。除此以外，它还有助于图书馆学和情报学课程的现代化和集成化。本书可用作讲授情报存贮和检索基本原理及开展作为专门应用领域的编目和参考咨询工作的基础教程。这样的教程不但能克服编目和参考服务的人为分离，而且也为学生选择第一职业提供了广泛的灵活性，并为他们专业的进一步发展在许多方面打下了良好基础。如果不改变任何课程量，本书也可以节选出编目和参考服务课程的指定章节。这样的教程可进一步应用于商业情

报系统、新闻专业以及类似的领域；它们都具有共同的理论基础，但是，其应用各有所别。本书还适合于那些希望为他们日常工作打下更好的理论基础的实际工作者自学。我们还正在编辑一本带有练习和补充示例说明的工作手册；从作者处可取得原稿。

情报学研究是一个新领域，这表明其术语是十分混乱的，部分原因是由于缺乏通用的概念机制：同一术语用于不同的概念；不同的术语用于同一概念。本书采用大多数著者在本领域中使用的术语，有时也引入了新概念的新术语，或用新术语取代反映错误概念分析的现存术语。

全书的思想是按如下方式展开的，情报专家是专职工作人员，他们在测定情报需求时与用户合作，并且利用他们的知识设计系统或进行检索来满足这些需求，而不仅仅是了解用户需求。

本书的第一部分讨论了情报的性质和结构，并对情报系统的所有性质进行了展开论述。第二部分根据可能实现的目标为情报存贮与检索系统的设计与利用提供了基础，并对第三至五部分中的设计方法的优缺点进行了讨论。第三部分论及了数据的逻辑表达式及数据存取结构，还论述了情报存贮与检索系统中必要的规则与约定。第四部分集中讨论了主题检索（许多原理具有更广泛的应用范围），讨论了索引语言的性质——词汇控制，基本功能及概念结构。第五部分讨论了标引和检索方法及系统分析与设计。

本书是根据在马里兰大学上课的讲稿写成的。学生的许多提问引起了我的深思，他们对几易手稿提出的意见十分有益。诺曼·罗伯特、哈罗德·博考、拉雅·艾德尔都提出了很

好的建议，其中有的建议对把本书压缩为适当篇幅很有帮助。简·伯格林和曼·萨默多次打印了初稿，并纠正了一些段落中的模糊文字。我的妻子莉萨自始至终检查讨论了本书的思想，并对本书的内容和形式提出了建议；她还花费大量时间进行编辑和校对。我还要感谢本领域中的许多同行，尤其是加尔文·摩尔斯的开拓精神。

译 序

迄今为止,国内出版的情报学理论、情报检索理论和情报检索语言方面的专著和译著已经很多,但我们仍然认为,这本书是很有特色,值得一读的。它最突出的特色是,没有太多的学科局限,作者没有过分地沉湎于技术细节,而是凭借自己的多学科知识和驾驭文献的能力,抓住了情报的本质,从广泛联系的立场出发阐明了情报存贮与检索系统的基本理论问题和设计问题。我们认为本书对于情报学研究人员融会贯通地理解情报工作各环节的理论基础、对于情报工作人员补充理论知识、对于情报科学各分支领域的融合都是大有裨益的。此外,本书中还包含了许多新颖的观点,有些是值得深入研究的,比如以用户立场出发的面向提问的标引方法。

在本书的翻译过程中,得到了陈光祚老师的悉心指导;张奇涵、李东风同志及姚富强、白学锋、徐正伟博士也给了我们许多有益的启发与帮助。在此,我们向他们表示衷心的感谢。

参加本书翻译的有谢新洲、陈睿、常彤,具体分工如下:常彤:第一部分(1、2、3、4章),谢新洲:第一部分(第5章)、第二部分(6、7、8章)、第五部分(16、17、18章),陈睿:第三部分(9、10、11章)、第四部分(12、13、14、15章),全书由陈睿、谢新洲统稿,最后由陈光祚进行了审校。

由于译者水平有限,书中错误在所难免,恳请读者批评指正。

译 者

1991年2月

目 录

前言.....	(I)
第 I 部分 情报传递的系统方法.....	(1)
第一章 解决问题的情报系统.....	(1)
第二章 情报的性质.....	(8)
目的.....	(8)
2.1 意象的作用.....	(8)
2.2 孤立个人意象的形成.....	(11)
2.3 通过交流的意象形成.....	(11)
2.4 情报传递和利用模型.....	(14)
2.5 数据、情报和知识.....	(16)
2.6 按所提供服务分类的情报系统类型.....	(18)
2.7 小结与推出的原则.....	(21)
第三章 情报的结构.....	(22)
目的.....	(22)
导言.....	(22)
3.1 大学数据库：一个实例.....	(22)
3.2 情报结构的元素.....	(25)
3.3 分析参考工具：一个实例.....	(28)
第四章 情报传递网络.....	(30)
目的.....	(30)
导言.....	(30)
4.1 情报传递网络中的环节.....	(31)
4.2 环节模型.....	(32)
4.3 环节的特性.....	(35)

第五章 情报系统结构.....	(41)
目的.....	(41)
5.1 情报系统的总体结构.....	(41)
5.1.1 识别特定用户需求	(44)
5.1.2 获取实体或有关情报	(45)
5.1.3 情报存贮与检索系统	(45)
5.1.4 将情报或实体提供用户利用	(45)
5.1.5 情报的进一步处理	(46)
5.1.6 识别一般需求; 需求指南	(48)
5.1.7 公共关系	(50)
5.1.8 系统的功能划分与组织划分	(51)
5.2 检索问题: 导引	(51)
5.2.1 索引语言和文档的结构: 导引	(59)
5.3 ISAR系统的结构	(61)
5.3.1 作为整体的 ISAR 系统.....	(61)
5.3.2 系统规则与约定: 概念模式.....	(63)
5.3.3 对 ISAR 系统其它组份的说明.....	(64)
5.4 定义.....	(65)
5.4.1 叙词、入口词、索引语言、入口词表、叙词表.....	(65)
5.4.2 检索提问、提问、提问表述、提问表达式.....	(66)
5.4.3 标引、编目、编码.....	(67)
第 I 部分 ISAR系统的目标.....	(71)
第六章 系统分析.....	(71)
目的.....	(71)
导言.....	(71)
6.1 制定决策的方法.....	(72)
6.2 系统分析过程的功能.....	(74)
6.3 系统分析的阶段: 系统生命周期.....	(84)

6.4	系统分析中情报与数据的搜集.....	(90)
6.5	选择决策.....	(93)
6.6	面向资源与面向问题的系统分析.....	(96)
6.7	性能与情报服务的影响.....	(97)
第七章	用户问题与需求评估.....	(99)
	目的.....	(99)
	导言.....	(99)
7.1	作为系统设计基础的用户研究.....	(100)
7.2	需求研究的原则.....	(104)
7.2.1	指定优先权.....	(104)
7.2.2	用户与情报人员的责任分担.....	(105)
7.2.3	需求、期望、要求或被认知的需求、利用与效果.....	(105)
7.2.4	无限制的需求评估.....	(107)
7.3	需求研究的方法.....	(107)
7.3.1	基于记录的问题分析.....	(110)
7.3.2	提问与检索行为分析.....	(111)
7.3.3	查询(潜在)用户的需求.....	(114)
第八章	情报存贮与检索系统的目标.....	(117)
	目的.....	(117)
	导言.....	(117)
8.1	结果质量的评价(局部性能).....	(125)
8.1.1	结果质量各方面的评价.....	(125)
8.1.2	结果质量的一个单一综合量度.....	(134)
8.2	ISAR系统总性能的评价量度.....	(136)
8.3	测试与评价.....	(138)
8.4	相关性和相关性判定.....	(140)
8.5	ISAR系统设计和运行的启示.....	(142)
8.5.1	系统设计的启示:有创新意义的特性.....	(142)

8.5.2	系统运行的启示	(145)
8.5.3	相关性判定与职业化	(146)
第Ⅲ部分 数据模式和数据结构		(149)
第九章 数据模式和格式		(149)
目的		(149)
引言		(149)
9.1	概念模式的设计	(151)
9.1.1	实体标识符形式的规则和约定	(155)
9.1.2	建立关系的一般规则：用一些简单规则来表明这一概念	(158)
9.2	记录格式：一般问题	(158)
9.2.1	记录在数据库操作中的功能	(158)
9.2.2	结构：一个关键概念	(162)
9.2.3	固定字段记录和可变字段记录	(163)
9.3	数据模式设计和评价标准	(167)
9.4	输入格式	(169)
9.4.1	输入格式的设计问题	(174)
9.5	输出格式	(176)
附录9.1	记录格式举例	(178)
第十章 基本提问表达式		(185)
目的		(185)
10.1	逻辑“与”(AND)	(185)
10.2	逻辑“或”(OR)	(186)
10.3	逻辑“与”和逻辑“或”混和使用	(187)
10.4	自然语言中“与”的模糊性	(189)
10.5	括号的层次	(190)
10.6	逻辑“非”(NOT)及其隐含的危害	(191)
第十一章 数据结构和存取		(193)

目的.....	(193)
11.1 数据结构的研究.....	(193)
11.2 数据结构的功能与特性.....	(224)
11.3 作为数据结构的主文档和索引文档.....	(226)
11.3.1 术语注释.....	(226)
11.3.2 一些用法及设计思想.....	(227)
11.3.3 作为数据结构一部分的实体主文档.....	(228)
11.4 序的概念.....	(229)
11.5 文档类型的二维连续统.....	(234)
11.6 数据库成本与检索成本之间的权衡取舍.....	(236)
11.6.1 情报量.....	(236)
11.6.2 有序程度.....	(237)
11.6.3 关于设计的一些考虑.....	(237)
11.7 定义.....	(239)
第 IV 部分 索引语言功能与结构.....	(245)
第十二章 词义控制	(245)
目的.....	(245)
引言: 词义控制的问题.....	(245)
12.1 概念和语词: 同义词—多义词结构.....	(251)
12.2 密切相关的概念的聚类: 等价结构.....	(252)
12.3 分类结构.....	(254)
12.4 索引语言.....	(255)
12.5 叙词表.....	(257)
第十三章 索引语言功能.....	(261)
目的.....	(261)
13.1 回顾: 情报检索问题	(261)
13.2 索引语言在标引中的作用.....	(263)
13.2.1 面向实体标引的缺陷.....	(264)

13.2.2	面向提问的标引: 一般方法	(267)
13.2.3	面向提问的标引: 实现	(271)
13.2.4	增补的面向实体的标引	(276)
13.3	索引语言在检索中的作用	(277)
13.3.1	适用于提问表达式的检查表技术	(277)
13.3.2	对面向提问标引不足的弥补	(280)
13.4	索引语言在数据库组织中的作用	(281)
13.5	选择最佳的标引方法	(284)
13.5.1	标引费用	(284)
13.5.2	标引质量	(285)
13.5.3	检索费用与质量	(286)
13.6	等级体系的作用: 小结	(287)
13.7	标引和分类的方法论	(289)
第十四章 索引语言结构 1: 概念结构		(292)
目的		(292)
导言		(292)
14.1	等级体系	(293)
14.2	概念组配与语义分解、组面分析	(299)
14.3	概念组配与等级体系间的相互作用	(305)
14.4	应用和例证: 检索	(316)
14.5	概念分析、组面分析: 评论	(323)
14.5.1	组面表的设计	(323)
14.5.2	对一般概念的认识	(324)
14.5.3	亚面	(326)
14.5.4	组面分析与先组叙词的词间关系	(327)
14.5.5	语义分解和组面分析的优越性	(327)
14.6	等级结构: 评论	(328)
14.6.1	等级关系与相关关系	(328)

14.6.2 等级关系的类型.....	(329)
14.6.3 收录新的上位概念.....	(331)
14.7 叙词表编制中概念的形成.....	(333)
第十五章 索引语言结构 2: 数据库组织	(337)
目的.....	(337)
导言.....	(337)
15.1 问题.....	(338)
15.2 实体成组, 成组文档中的检索.....	(340)
15.2.1 成组的思想与先组式叙词.....	(340)
15.2.2 从思想到现实: 限定先组.....	(352)
15.2.3 成组文档的存取优越性.....	(356)
15.3 成组聚类与实体描述.....	(357)
15.4 后组与先组.....	(359)
15.4.1 后组与先组的程度.....	(359)
15.4.2 确定总体先组程度.....	(362)
15.4.3 确定各个先组叙词.....	(363)
15.4.4 标引和检索中的先组叙词.....	(366)
15.5 出于存取目的的索引语言的组织.....	(368)
15.5.1 查词索引.....	(369)
15.5.2 叙词的设计与排列.....	(375)
15.6 适用于不同检索装置的集成索引语言.....	(382)
第 V 部分 ISAR 系统运行与设计	(385)
第十六章 标引的专指度与穷举度	(385)
目的.....	(385)
16.1 标引对系统性能的重要性.....	(385)
16.2 标引的穷举度与专指度定义.....	(386)
16.2.1 穷举度的定义.....	(386)
16.2.2 专指度的定义.....	(389)

16.3 标引的穷举度和专指度对检索性能的影响.....	(390)
16.3.1 穷举度的影响	(390)
16.3.2 专指度的影响	(393)
16.3.3 关于穷举度和专指度影响的错误观念.....	(395)
16.3.4 小结.....	(396)
16.4 设计标引规则与步骤.....	(396)
16.4.1 影响穷举度的因素.....	(396)
16.4.2 影响专指度的因素.....	(399)
16.4.3 费用问题.....	(400)
第十七章 检索.....	(402)
目的.....	(402)
导言.....	(402)
17.1 认识和表述需求、表述检索需求.....	(405)
17.1.1 认识一个需求的存在.....	(406)
17.1.2 给出提问表述.....	(407)
17.1.3 决定明确的检索要求	(410)
17.2 制定检索策略	(411)
17.2.1 构造概念提问表达式	(412)
17.2.2 选择情报源并按检索顺序排序.....	(421)
17.2.3 将概念提问表达式转换成情报源的语言.....	(424)
17.2.4 自由文本检索	(429)
17.2.5 概念提问表达式与面向特定情报源的提问表达式间的交互作用.....	(432)
17.3 检索策略的执行.....	(432)
17.4 回顾检索结果并改进检索.....	(433)
17.5 编辑检索结果并传递给用户.....	(435)
17.6 检查结果是否对用户有帮助.....	(435)
17.7 交互.....	(436)

17.8 监控检索过程和评价检索结果	(442)
第十八章 情报系统设计与评价	(444)
目的	(444)
导言	(444)
18.1 确定用户要求与能力	(448)
18.1.1 确定情报系统的范围	(448)
18.1.2 获取检索提问	(449)
18.2 建立馆藏并进行相关性判断	(450)
18.2.1 建立实体馆藏	(450)
18.2.2 进行检索前相关性判断	(451)
18.3 设计和建立ISAR系统	(452)
18.4 运行ISAR系统	(454)
18.4.1 标引实体	(454)
18.4.2 构造提问表达式	(455)
18.4.3 检索实体并判断其相关性	(457)
18.5 评价ISAR系统性能	(457)
18.5.1 宏观分析: 性能度量	(457)
18.5.2 微观分析: 检索成功与失败	(458)
18.6 检索测试及系统设计与运行	(458)
18.7 作为设计原则的成本-效益分析	(461)
18.8 作为设计原则的面向问题的立场	(464)

第I部分 情报传递的系统方法

第一章 解决问题的情报系统

情报系统存在于我们周围：管理信息系统、消息传播系统，情报和咨询中心，图书馆，文摘/索引机构，法律数据库，职业介绍所，商店和库存目录(商业情报系统)等。为了解决问题，情报系统在不断增长，情报的目的就是为了解决问题。我们生活的世界越复杂，我们就需要越多的情报来解决问题。图1.1表示了某些问题和由这些问题而产生的情报需求。

一个情报系统可以识别问题，提供有用的情报并将这些情报传递给需要的用户。同时，情报系统也可以处理和传递有用的实体，如商品和技术部件，甚至人才(例如职业介绍所)。请看下图：

