

日本科学技术厅编

科技情报工作
现状和展望丛书

科技情报的 机械化和自动化

科学技术文献出版社

科技情报工作
现状和展望丛书

科技情报的机械化 和 自 动 化

日本科学技术厅 编

高崇谦 赵仲儒 译

曾民族 校

科学技术文献出版社

1986

内 容 简 介

本书是《科技情报工作现状和展望丛书》之一，是日本科学技术厅为向广大科技人员普及情报基本知识委托日本文献工作协会组织科技情报界知名人士编写的。本书以通俗易懂的语言，全面系统地介绍科技情报工作中技术革新的现状和动向。从情报检索入手介绍检索语言，联机系统，数据库，汉字系统，制作情报资料的自动化，新型资料发行形式，新型流通方式，并对未来技术给予展望。书中提供许多具体数据，文后附有80个专业名词解释，15个日本重要情报机构概况以及有关标准资料。

本书可作为科技人员情报知识的普及读物，也可作为情报人员从事情报技术开发和应用的参考资料。

2010/14
科学技术厅编

科学技术情报の机械化

大藏省印刷局，昭和56年5月

科技情报的机械化和自动化

日本科学技术厅编

高嶽謙等 译

科学技术文献出版社出版

中国科学技术情报研究所印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

开本：787×1092¹/₃₂ 印张：8.25 字数：177千字

1986年3月北京第一版第一次印刷

印数：1—9420册

科技新书目：115—47

统一书号：17176·455

定价：1.75元

出 版 说 明

《科技情报工作 现状和展望丛书》是日本科学技术厅主编的，内容包括：化学情报；医学情报；科技情报的机械化和自动化；农林水产科技情报；冶金情报。我社翻译出版的这套丛书，先后次序与原书有所变更，这次出版《科技情报的机械化和自动化》，其他几本将陆续出版。

1985年4月

前 言

日本正在迎接经济社会稳定增长期的到来，对科学技术的期望日益增大；同时也在按照国内外环境条件的变化，把积极、主动地开发技术作为当前的课题，在各方面制定相应的政策。

从科学技术健全发展的角度看，科学技术情报作为科研活动的成果，凝聚了高度的知识和大量的投资，对最大限度地使用有限的资金和人力，作用十分重要。

但是，随着近年来科学技术的迅速发展，科学技术情报的种类和数量急剧增加，因此要全面理解并掌握科学技术各领域的发展动向也越来越困难。

日本科学技术厅为了对综合性科学技术领域的发展过程、现状、存在的问题以及未来展望进行全面评估，编出可供政府、企业和研究管理等部门参考的资料，以广泛交流科学技术情报。从1972年开始，开展了综合述评工作。

这项工作从海洋科学技术、材料科学技术、电子通信、生命科学以及系统科学的各个领域，有计划地选取若干主要课题，以丛书的形式出版。近几年来，由于对科学技术情报工作重要性的认识不断提高，1977年以后“科学技术情报工作”也列入综合述评的对象，1977年出版了《化学情报》、1978年出版了《医学情报》。

1979年，作为科学技术情报工作丛书的第三部分，选取“科学技术情报的机械化”为题，委托社团法人日本文献协会，进行编写。

为了帮助人们全面了解促进科学技术情报流通的技术革新的动向,本书就情报的存贮技术、依赖人脑智能的标引作业的自动化技术、联机情报检索技术、分散在各种情报中心的数据库综合利用技术以及日本独自开发的汉字处理技术等现状以及动向、存在问题、未来展望等进行述评。

希望读者结合其他学科的综合述评阅读本书,同时对本书提出宝贵意见。

最后,谨向对本述评工作提供帮助的有关人士表示深切谢意。

1981年3月

科学技术厅振兴局长 宫本二郎

编辑和执笔委员会名单

- | | |
|-----------|------------------------|
| 委员长: 中村幸雄 | (日本文献协会副会长) |
| 委员: 草间 基 | (东北金属株式会社事业计划室) |
| 笹森勝之助 | (富士照相胶片株式会社缩微系统部) |
| 泽本 孝久 | (前庆应义塾大学文学部教授) |
| 中井 浩 | (日本科学技术情报中心技术管理室主任) |
| 南部和夫 | (花王石鹼株式会社办公室自动化推进本部部长) |
| 山本毅雄 | (东京大学大型计算机中心副教授) |

目 录

第1章 科学技术情报的检索	(1)
1.1 情报检索机械化的前提条件	(1)
1.1.1 什么叫机械化.....	(1)
1.1.2 标引	(2)
1.1.3 标题	(3)
1.1.4 分类法	(5)
1.1.5 查找	(8)
1.2 初期的机械化	(10)
1.2.1 叙词表的发展	(10)
1.2.2 叙词表的层次关系.....	(16)
1.2.3 叙词的组配	(22)
1.3 穿孔卡片	(24)
1.4 批量处理和通用计算机	(28)
1.4.1 计算机的利用	(28)
1.4.2 字符	(29)
1.4.3 匹配	(31)
1.4.4 存贮器	(33)
1.4.5 磁介质文档	(34)
1.4.6 磁带检索	(36)
第2章 检索语言	(39)
2.1 新型分类法	(39)

2.1.1	分类法中的观点问题	(39)
2.1.2	组面分类法	(44)
2.1.3	组面分类法的实用	(49)
2.2	分类法的词语问题	(51)
2.2.1	概念与词语的关系	(51)
2.2.2	组面叙词表	(51)
2.2.3	分类表主题索引的改进	(54)
2.3	主题结构	(59)
2.3.1	主题的定义	(59)
2.3.2	EJC叙词表的11种区别	(60)
2.3.3	主题图解法(1)	(63)
2.3.4	主题图解法(2)	(65)
2.4	检索语言的综合比较	(66)
2.4.1	主题词法	(67)
2.4.2	分类法	(68)
2.4.3	两种方式的对应关系	(69)
2.4.4	两者的搭配	(70)
2.5	情报检索和用词	(71)
第3章	联机情报检索	(75)
3.1	情报检索联机化	(75)
3.1.1	什么叫联机	(75)
3.1.2	文档结构的变化	(76)
3.1.3	检索手段	(77)
3.1.4	联机方式下的批处理	(80)
3.1.5	联机检索的经济性	(81)
3.2	联机的制约条件	(87)

3.2.1	通信线路	(87)
3.2.2	情报中心的网络	(88)
3.2.3	用户对联机服务的反应	(92)
3.3	多种文献库和多种中心的利用	(94)
3.3.1	数据库的产生	(94)
3.3.2	多种情报中心的联合	(95)
3.3.3	计算机通信网的利用	(98)
第4章	数据库	(101)
4.1	数据库的范围	(101)
4.1.1	数据库的定义与特性	(101)
4.1.2	数据库的内容	(102)
4.1.3	数据库的制作单位	(103)
4.1.4	自然形成的数据库利用方式	(104)
4.2	数据库利用的有关问题	(106)
4.2.1	多个数据库的利用	(106)
4.2.2	情报中间人存在的意义	(108)
4.2.3	数据库制作的发展方向	(109)
4.3	事实数据库	(110)
4.4	国际关系	(112)
4.4.1	数据库的建立	(112)
4.4.2	一次资料的提供	(119)
第5章	汉字系统	(121)
5.1	汉字系统的特点	(121)
5.1.1	汉字的特性及其对策	(121)
5.1.2	汉字处理系统的难点	(126)
5.1.3	汉字联机检索系统	(127)

5.2 汉字输入输出设备	(130)
5.2.1 只用于情报检索场合	(130)
5.2.2 汉字文字处理场合	(131)
5.2.3 打印机的速度与高速化	(132)
第6章 制作情报资料的机械化	(134)
6.1 印刷、编辑的机械化	(134)
6.1.1 铅字向照相排版的过渡	(134)
6.1.2 文字处理机的应用	(136)
6.1.3 小批量印刷出版物	(140)
6.2 资料形态的变化	(141)
6.2.1 各种专刊的增加	(141)
6.2.2 索引刊物的大量出现	(142)
6.2.3 大型著作由于节省作者的劳力而得以加快 出版发行	(143)
6.3 自动标引与自动编文摘	(144)
第7章 资料发行形态的变革	(147)
7.1 学术杂志、专业杂志的发展趋势	(147)
7.1.1 文献发生量的增加	(147)
7.1.2 对文献的需求与形式	(148)
7.1.3 对学会杂志、专业杂志的影响	(149)
7.1.4 发行形态改革的方向	(151)
7.2 缩微品的应用	(152)
7.2.1 发展阶段	(152)
7.2.2 在文献工作方面应用的特点	(154)
7.3 选择性流通方式	(155)
7.3.1 选择性的必要理由	(155)

7.3.2 选择范围和提供方法	(157)
7.4 杂志多级出版方式	(162)
第8章 专业情报流通过程的问题	(164)
8.1 不同情报形态的流通过程	(164)
8.1.1 图像情报的流通	(164)
8.1.2 口头情报	(171)
8.2 大范围交流与国际交流	(173)
8.2.1 有关搜集情报的问题	(173)
8.2.2 向国外输出情报	(175)
第9章 未来展望	(178)
9.1 电子计算机应用的目的	(178)
9.2 资料概念的变化	(181)
9.2.1 情报载体的改变	(181)
9.2.2 情报流通方面的变化	(182)
9.3 辅助情报生产的各种自动化设备	(185)
9.3.1 促进脑力活动的设备	(185)
9.3.2 节约情报生产所需人力的设备	(187)
用词解释	(188)
附录1 日本主要情报机构一览表	(227)
附录2 信息方面的日本国家标准 (JIS)一览表	(240)
附录3 情报专业的ISO标准一览表(1978)	(244)
附录4 ISO646字符集规格	(248)
注释	(252)

第1章 科学技术情报的检索

1.1 情报检索机械化的前提条件

1.1.1 什么叫机械化

情报检索的机械化, 简单地说, 就是使用通用电子计算机。但是, 在发展机械化的过程中, 出现过各种专用机械, 即使今天, 在某种特定条件下, 放弃某些专用机械也是行不通的。目前计算机主要用途为顺序处理, 但也未必都是方便的。

因此, 不考虑情报检索的基本问题, 一上来就考虑使用计算机, 往往搞错问题的本质, 导致在估价将来发展方面发生差错。

在处理具有高度专业内容并且需要专家从事处理的专业情报*时, 适于使用机械尤其使用计算机, 不仅是情报检索, 而且还应该考虑其他方面。因此, 本书不局限于狭义的情报检索, 在后半部分, 对其他方面, 特别是情报的生产(资料的制作)也都给予注意。这些问题与情报检索问题紧密相关, 今后的地位将越来越重要。

* “科学技术情报”是专业情报的一种。专门情报除了科学技术以外, 还包括其他学科。对科学技术以外也适用的情报, 本书统一采用“专业情报”这一概念。

讨论机械化，必须从机械化开始之前讲起。机械化就是使用机械，也就是当时手工完成的事情部分用机械来代替，因此，必须理解机械化业务是指当时人们不用机械所做的日常业务。当然，这些业务的实施方法世界各国各不相同。但是机械化首先是从美国开始的，所以全部按照美国的实践进行设计。

欧洲的机械化起步较晚，如果欧洲首先采用机械化，人们想像当会出现与美国做法不同的欧洲方式的机械化，然而实际并非如此。所以，美国方式之外的做法，是指与初期机械化的做法完全不同的做法。以上事实今日仍然适用。

与情报检索工作首先关系最为密切的是图书馆。美国和欧洲图书馆的实践最大不同点在于：在编制图书目录时，欧洲主要采用分类法，而美国则采用标题法。差别的原因很多，欧洲直接处理概念，而实用主义盛行的美国则把表示概念的自然语言作为处理对象。美国方式的特点在于入门容易、达到目的难。与此相反，欧洲方式的特点在于入门难而容易实现目的。

1.1.2 标引

图书馆和资料室，通常收藏大量的图书与资料，必须根据需要而能提取少量所需的图书与资料。一般来讲，图书由其内容（称作主题，参照2.3）和作者来表示。（还可包括其他项目，但不太重要，故略。）检索时把这些作为线索是很方便的。这种线索称为标识（tag）。图书以外的情报，也多以内容（即主题）作为主要标识；至于作者是否作为主要对象来表示，要取决于情报的种类。

总之，选什么项目作为标识，根据情报的种类不同而异，然而不选内容（即主题）作为标识者却是很少见的。

这种简单地表示情报内容的方法统称为索引（index）。索引有各种各样的索引，如分类索引，关键词索引，以及由某种体系所规定的代码索引等等。像图书的书名，论文的题目以及新闻消息的标题等绝大多数是用来表示特定的情报，不能直接称为索引。索引通常不是表示特定的情报的，而是表示一组同类的情报（叫做一类）。一类包含多件情报。如果个数太多（例成百成千时），一类则失去有效的作用，所以通常需要把那个类再细分为多个小类。也就是说，一个类如果只有一件情报，这样处理效率低。究竟几件合适，没有具体的统计数据也难以确定。手工操作时，根据经验，一般以20—30件左右为宜。

例如，常见的分类卡片目录，1张卡片代表1件情报，一类则是具有同一类号的几张卡片的集合。

指派与各件情报内容相一致的类，称为标引*（Indexing）。

1.1.3 标题（Subject heading）

标志的一个类，叫做一个标题。一个标题可用一个单词或若干单词的组合表示，要求概念明确易懂。例如“飞机”，“客机”，“特别快车”，“深海潜艇”等。把成百上千个这样的标题汇集在一起，就成为“标题表”。人们把

*日本也有人管Indexing叫做“编制索引”，为了避免产生误解，日语对标引采用“インデクシング”一词表示，对编制索引采用“索引付け”一词表示。

这个表中的一个或几个标题，分配给整理对象的情报（或资料）加以存贮，查找情报资料时，把认为合适的标题做为标识（线索）来使用。因此，标题之间如果没有唯一性，就会给使用造成困难。这样，除特殊情况外，如果给标题的人和为了查找而使用标题的人不是同一个人*，就会遇到困难。

标题出现“飞行器”和“航空器”，可能发生混乱，因为一般人都认为这两个词是同义词。而“客机”和“飞机”之间，意义的范围不同，“客机”是“飞机”的一种，“飞机”不一定是“客机”。因此，有关“客机”的情报，放在“客机”一类。有关“军用飞机”的情报，放在“军用飞机”一类。没有客用、军用之分的问题，放入“飞机”一类。这样就不会引起混乱。

标题的选择方法，由于标题表已注意采取避免混乱的措施，可以是任意的。标题的总数太多或太少都不好。标题的数量应取决于需要整理的情报的总数。有几百万件资料的单位和不满1万件资料的小规模文档，即使使用相同的标题，也会出现差异。

标题必须简洁。但是过于简单，意义可能不明确。例如，“采购法”这一标题，其意义好象谁都懂。但是购买这一行为的主体范围却很广。企业的采购员使用的“方法”与主妇在商店所考虑的“购物法”则有很大的不同。因此，“采购法”这样的标题如果不指出范围、不加注释、不给具体示例，不表示主体或环境，则含义是不清楚的。

* 即使同一个人做这两方面的工作，如果时间不一样，也不能保证所用的标题相同。

因此,对于重视以少量单词简洁表达概念的标题法,必须妥善处理使用上的方便和意义含糊带来的不方便这一对矛盾。

标题表收录的是以自然语言的单词构成的“用词”(term)。各个标题表的“用词”总数有一定限度。标题法限定只能使用表中的词进行检索。

例如,某标题表有关飞机的用词只有表1.1所示的几个。飞机的种类只有这么几种,而且除此以外的用词又不允许使

表1.1 标题表片断

普通飞机
客 机
军 用 机
小 型 机
直 升 机

用,因此要想按照资料的篇名进行标引,就必须动动脑筋。

“涡轮喷气机的……”入“客机”,“P3C 许可制造问题的……”入“军用机”,这些可以立刻处理。而“军用直升飞机……”的问题,究竟入“军用机”,还是入“直升机”就不那么好办。当然,人们可以有即席的解释,然而“因人而异”怎么办。为了克服这个困难,只有加上定义、注释,越搞越复杂。这样一来,与标题法“简洁易懂”的特性则背道而驰。

1.1.4 分类法

分类法同标题法一样,也是处理“类”,但是在确定类

时，分类法不考虑类的名字，而考虑类代表什么概念。这个概念有的可用简洁名字表示，有的则不然。

例如，当产生过电流时，为了不损坏整个机器设备，用于切断部分电流回路以达到安全目的的电气设备，一般叫做“断路器”或“保护装置”。检查过电流有种种方法。其中一种是“保险丝”。这是把低熔点的金属丝放到电路中，由于过电流的焦耳热，金属丝熔化而切断线路。但是，还有其他类似设备，可以例举如下。

对于电路内引起的过电流进行保护：

- (1) 保险丝；
- (2) 保护继电器（利用电流的电磁力的保护装置）；
- (3) 利用绝缘材料等温度升高的保护装置。

以上（1）、（2）、（3）在概念上属于同一类（同位）。从自然语言角度看，（1）的用词直接了当，（2）也是一个词汇，但加注释，即它是继电器的一种，是进行过电流保护的“保护继电器”，（3）由于没有合适的名称，只能用长句表示。如果采用标题的话，可取代表整体的“保护装置”，“保险丝”，“保护继电器”作为标题，而（3）则无法取标题。

但是，从分类法的角度看，（1）、（2）、（3）都是可取的。这是因为在概念上不受“表达长短”的限制（只要内容上有可取价值就行），即凡是可以表达的概念，都可选为类目。

在多数情况下，简洁的名词，使用频率也高，第（3）种情况似乎不太实用。但从学科体系上考虑事物或考虑出现新的事物时，如果只图表达方便的话，恐怕是不合适的。