

第一章 情报学经验规律

一、布拉德福定律、洛特卡定律、齐夫定律——P1-

1. 布拉德福定律 p1-p4

定义：

是描述专业论文在期刊中分布情况的经验规律的规律。

应用范围：

确定核心期刊，以指导期刊订购和期刊利用，并由此扩展到核心馆藏维护、核心检索工具选择等。是“二八律”的体现。

与此同时还应用在考察专著分布；引用期刊的研究，发现影响因子。

2. 洛特卡定律 p4-p7

定义：

是描述作者与论文数量之间关系的经验规律。

应用范围：

主要用于研究“科学生产率”，可用于预测发表不同篇数论文的作者数量和特定学科的论文总量，或根据特定领域作者数量估计该领域相关科学论文数量等。

3. 齐夫定律 p7-p9

定义：

是描述文献中用词与其出现频次之间关系的经验规律。

应用范围：

主要用于指导文献标引、词表编制和情报检索等。

4. 文献增长率 p9-p11

定义：

是描述文献信息量随时间而不断增加的一组经验规律。

其中包括：指数曲线、逻辑曲线、勒歇尔模型

5. 文献老化律 p11-p12

定义：

是描述文献信息价值随时间推移而不断下降的一组经验规律

其中包括：负指数老化模型、博顿-凯卜勒方程

第二章 情报学基础理论

1. 概念属性 p32

数据、信息、情报、知识、文献属性之间的区别

第三章 情报学方法

1. 学术研究程式 p43（注意与后面章节结合）

包括：选题 检索 研究 结果

对于选题：

选题为何重要？

选题具有实际的意义和实用价值。从大方面而言，对于社会、经济发展有贡献；抑或是从小方面而言，对某个领域有支持贡献。或者选题可以解决某个领域内尚未解决的问题。只有具有上述作用的选题才算得上是有意义、有价值的选题，因此确定一个合适的选题是非常重要的。

对于检索：

检索策略的制定、以及有什么样的检索流程？——第六章

6.1 情报检索基本原理

The diagram illustrates the basic principles of information retrieval, showing the flow of information from storage to retrieval. It is divided into two main horizontal sections: "信息存储" (Information Storage) at the top and "信息检索" (Information Retrieval) at the bottom, both indicated by large double-headed arrows.

Information Storage Process:

- 信息源 1、2、3...** (Information Sources 1, 2, 3...): The starting point of the storage process.
- 信息选择、采集** (Information Selection and Collection): The process of selecting and gathering information from the sources.
- 信息描述 (特征提取、标引、规范)** (Information Description (Feature Extraction, Indexing, Standardization)): The process of describing the information, including feature extraction, indexing, and standardization.

Information Retrieval Process:

- 用户** (User): The starting point of the retrieval process.
- 需求分析** (Requirement Analysis): The process of analyzing the user's needs.
- 需求表达** (Requirement Expression): The process of expressing the user's requirements.

Central Components and Flow:

- 索引交互系统** (Index Interaction System): A central component that receives input from the "信息描述" and "需求表达" boxes and interacts with the "数据库" (Database).
- 数据库** (Database): A central component that stores information and provides it to the "检索结果" (Retrieval Results) box.
- 检索结果** (Retrieval Results): The output of the retrieval process, which is then sent to the **用户** (User).

The flow of information is as follows: Information Sources (1, 2, 3...) lead to Information Selection and Collection, which leads to Information Description (Feature Extraction, Indexing, Standardization). This leads to the Index Interaction System, which interacts with the Database. The Database provides Retrieval Results, which are sent to the User. The User's Requirement Analysis leads to Requirement Expression, which interacts with the Index Interaction System. The Index Interaction System also interacts with the Database.

```
graph TD; subgraph "信息存储"; IS[信息源 1、2、3...] --> ISA[信息选择、采集]; ISA --> ID[信息描述<br/>(特征提取、标引、规范)]; end; subgraph "信息检索"; U[用户] --> DA[需求分析]; DA --> DE[需求表达]; end; ID --> IIS{{索引交互系统}}; DE --> IIS; IIS <--> DB[(数据库)]; DB --> IR[检索结果]; IR --> U;
```

广义信息检索基本原理

数据分析的研究流程是什么？

对于结果:

分析数据结果的正反方面意义、启示、价值等。

包括实体规范和形式规范

要求学术研究尊重学术传统、维护学术道德，从源头上防止并杜绝抄袭、剽窃等学术不端行为，以创造坚持学术真理、保护学术自由的学术文化氛围。

主要作用于学术成果的表现形式和格式上,

表现形式主要有：学术论文、学术专著、研究报告、专利说明书等

格式规范有：论著格式规范、引文格式规范（哈佛制、温哥华制、牛津制、芝加哥制）

实体规范侧重约束内容和方法，形式规范侧重约束形式和格式。

统计分析法:

包括假设检验、时序分析、回归分析、多元分析

知道上述四个分析的概念。

4.1 引文分析法 p62

基于施引和被引关系发展起来的具有情报学特色的方法就是引文分析

共被引：当两个作者的文献同时被第三个作者的文献引用，则称这两个作者存在共引关系；如果这两位作者共被引频次越高，则说明他们的学术关系越密切，“距离”越近。

耦合：若两篇文章同时引用了同一篇文章的话，那这两篇文章之间称作具有耦合关系。

共现：对于作者而言，如果有多个作者共同出现在一个文献中，那么就称这些作者之间存在共现合作关系；对于关键词而言，如果有多个关键词共同出现在同一个文献中，那么就称这些关键词之间存在关键词共现关系。

指标

h 指数：是一个混合量化指标，可用于评估研究人员的学术产出数量与学术产出水平。一名科学家的 h 指数是指其发表的 N_p 篇论文中有 h 篇每篇至少被引 h 次、而其余 N_p-h 篇论文每篇被引均小于或等于 h 次。

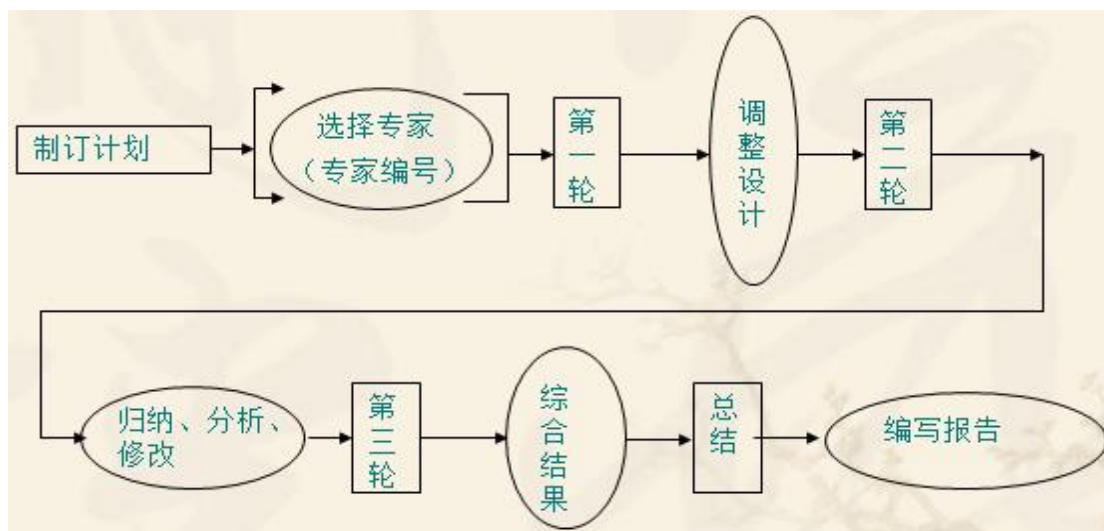
期刊影响因子：期刊影响因子，是代表期刊影响大小的一项定量指标。也就是某刊平均每篇论文的被引用数，它实际上是某刊在某年被全部源刊物引证该刊前两年发表论文的次数，与该刊前两年所发表的全部源论文数之比。

被引频次：以一定数量的统计源（来源期刊）为基础而统计的特定对象被来源期刊所引用的总次数。

4.2 德尔菲法 p65

概念：德尔菲法实际上是规定程序的专家调查法，具有匿名性、反馈性、统计性三大特点。它是由调查组织者拟定调查表，按照规定程序，通过函件分别向专家组成员征询调查，专家组成员之间通过组织者的反馈材料匿名地交流意见，经过三轮征询和反馈，专家的意见逐渐集中，最后获得有统计意义的集体判断结果。

流程：



4.3 头脑风暴法（BS 法）-p67

是用会议形式通过专家间的相互交流，激励智力碰撞，产生思想火花，使专家的意见得到集中和精化。

头脑风暴法的参会人数以不超过 15 人为宜，时间约 60-90min。

第四章 情报信息源

1. 情报信息源的分类

按照情报信息源是否为文献形式，可分为文献情报信息源和非文献情报信息源

文献情报信息源中包括：图书、期刊、专利文献、标准文献、科技报告、学位论文、会议文献、政府出版物、产品样本、档案文献。

非文献情报信息源中包括：电子传媒情报信息源、口头情报信息源、实物情报信息源、技术搜集。

按照载体形态，可分为印刷型和非印刷型

按照应用领域，可分为学术情报信息源和非学术情报信息源

按传播方式，可分为直接情报信息源和间接情报信息源

2. 情报源去哪可以检索

获取中文文献的数据库：万方、维普、cnki

获取英文文献的数据库：web of science、IEEE/IET Electronic Library

3. 非文献情报信息源 p98

替代计量学

替代计量指标

4. 专利与文献的区别 p88-p89

专利：专利是对技术发明给予法律保护的一种制度，专利权与商标权、工业品外观设计等一道构成工业产权，工业产权又与版权一道构成知识产权。

专利文献无论是形式上还是内容上都具有区别于其他文献类型的特殊之处。由于构成专利起码要符合新颖性、先进性和实用性三个条件，因此，专利反映的发明都是首先取得、在此之前不曾发表过的有关文献，在技术上有独到之处并对实际应用有价值。此外，专利文献还具有以下特点：详尽、实用性强；涉及技术领域范围广；专利文献集技术、法律和经济情报为一体；可靠性强，质量高；内容新颖，出版迅速；内容详尽，格式规范化；某些国家的专利申请书含有对期刊文献的引证，因此，可据此分析科学与技术的关联。

5. 用专利分析与文献分析分别侧重哪些方面

专利分析法对专利说明书、专利公报中大量零碎的专利信息进行分析、加工、组合，并利用**统计学方法和技巧**使这些信息转化为具有总揽全局及预测功能的竞争情报，从而为企业的技术、产品及服务开发中的决策提供参考。

文献分析法是指通过对收集到的某方面的**文献资料**进行研究，以探明研究对象的性质和状况，并从中引出自己观点的分析方法。

6. 学位论文

学位论文的篇幅不定，内容一般从历史、背景、述评开始，详细介绍研究经过、实验记录 and 具体数据等。

7. 非文献情报信息源可以用在哪些方面

非文献情报信息源虽然不常用于学术研究，但却往往是市场情报、竞争情报等的基础。

8. 开放存取信息资源 p101

OA（开放存取 open access）：文献可通过公共互联网免费提供，允许任何用户阅读、下载、复制、传播、打印，搜索或者对论文的全文进行链接，建立索引，作为素材编入软件，或者对论文进行任何其他处于合法目的的使用，不受除了与互联网本身无法接触的那些障碍之外的法律或技术障碍。唯一的限制是在再生及散布时，必须提到作者的名字，并将著作权归属于作者。

OA 期刊（OAJ）：采取不向读者收取费用的资金模式，用户可以不受限制地访问期刊论文全文。

收录 OAJ 的数据库有 Scopus, Web of Science, ESI 等。

OAJ 的分类

①根据期刊的创办方式划分为原生型 OAJ 和转化型 OAJ，前者是指采用 OAJ 出版模式新创办的一种期刊，后者是指由传统的学术期刊转化而成的 OAJ

②根据期刊的访问方式和访问权限划分为完全 OAJ 与部分 OAJ，前者是指**出版之后即可免费获取**的期刊，如（美国公共科学图书馆 PLOS）和生物医学期刊出版中心（BMC）出版的期刊。部分 OAJ 又称复合 OAJ，指对期刊的部分内容 OA 由作者自由选择论文是否开放，**但需要作者支付一定的费用**

③延迟 OAJ，是指超过预设的时间之后采用 OA 模式的期刊，如《中国图书馆学报》

OA 仓储（OAR）：开放仓储库，是 OA——开放存取的新兴实践方式，是为了解决传统出版模式滞后，对许多在学术活动中产生的非正式信息不能方便地加以查询利用的问题而诞生的。根据内部资源的收录范围和对外开放程度，OA 仓储可以大致分为：学科范围开放的学科 OA 仓储和机构专属的机构 OA 仓储两类。

①学科 OA 仓储

学科 OA 仓储是按照学科领域进行组织的 OA 仓储，开放仓储库中的文献可以是电子预印本，也可以是后印本。

②机构 OA 仓储

机构专属的开放仓储，即机构库，或者机构知识库。机构仓储汇聚了某个大学、研究机构研究人员在进行学术活动中产生的各种学术信息，提供信息提交、发布、保存、浏览、检索和全文下载等服务，往往由大学、研究机构或者大学图书馆、政府部门等创建和维护。

9. 对 OA 的评价

优势：开放存取期刊是由开放存取衍生来的，由于传统的订阅期刊出版模式阻碍了科学研究的交流，科研人员很难了解自己的研究内容在同行领域的研究情况，传统期刊出版方式会导致科研工作的重复劳动，导致学术资源恶性循环，不利于科技进步。另一方面，开放存取期刊能保障读者免费获得期刊全文，这能保障读者获得准确全面的信息。在开放存取的出版模式下，开放存取期刊各种优势特征使科研成果得到更广泛的传播，学术影响力的范围越来越广，开放存取期刊必将对科研领域产生深远的影响。

劣势：面对开放存取期刊数量的不断扩大，具有价值的开放存取期刊可能会被大量的水平较低的期刊所冲击，影响开放存取期刊资源整体质量。

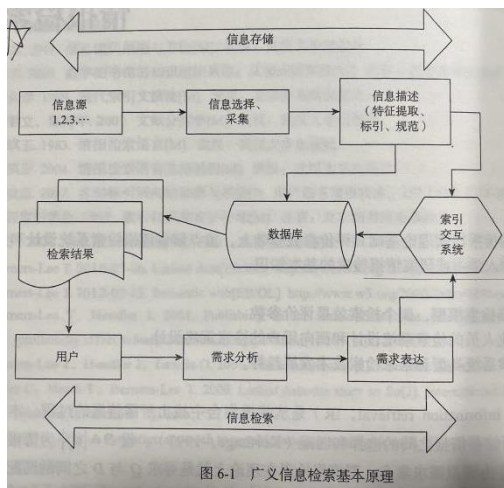
第五章 情报组织

第六章 情报检索

1、掌握情报检索的基本原理，理解流程 P148

狭义的情报检索：“查询”或“搜索”的单一过程

广义的情报检索：包含情报的存储与检索两个互逆过程



2、理解基本逻辑算法含义 (and or not) P153

And, *: 逻辑与。A 与 B 都为真时，结果才为真

Or, +: 逻辑或。A 或 B 中只要有一个为真时，结果就为真

Not, -: 逻辑非。A 为真、B 为假时，结果才为真

3、掌握使用检索技术的注意事项 P154

不同检索系统对检索技术的应用规则、表示方式、采用符号等都不相同，使用时必须首先了解检索系统相关的规则。

(常用检索系统的检索技术 见表 6-1)

4、检索系统选择 P156→决定数据质量

(1) **最简途径**：在对所需信息并不很明确，也无法判断用什么数据库最合适的情况下，最简单地可以使用搜索引擎进行试查。

(2) **原文直接检索途径**：如果对所需信息非常明确，如需查阅某个出版社，某个学协会出版的某种期刊刊载的论文都可以利用出版社、学协会或包含全文的数据库直接检索。

(3) **分领域检索途径**：很多专业领域都有专业的检索系统，如专利领域的 DII，工程技术领域的 Engineering。这些检索系统是该学科领域和专门文献类型的首选检索工具。

(4) **综合检索途径**：有些检索系统，涵盖了各领域的文献内容，有时也包含了多种文献类型，提供较为强大的检索功能（如引文检索），可方便用户进行“一站式”检索，进行广泛的综合检索。如中国知网，维普中文科技期刊数据库，万方数据。

(5) **跨库检索途径**：除了选择单一的数据库检索，还可以选择一些多库检索平台，一次性对多种文献类型，多个科学领域的文献进行检索。如中国知网文献总库

5、逻辑检索式编制→决定数据是否可靠 P157

逻辑检索式编制检索式：用户表达检索提问的逻辑表达式

6、情报检索评价，理解公式 P158

R (查全率) = $b/a \times 100\%$

P (查准率) = $b/m \times 100\%$

M (漏检率) = $(1 - b/a) \times 100\% = 100\% - R$

N (误检率) = $(1 - b/m) \times 100\% = 100\% - P$

n 为检索系统中文献总量， m 为检索输出的文献量， a 为 n 中与检索课题有关的文献量， b 为 m 中与检

索课题有关的文献量

7、综合分析功能 P168

第七章 情报分析

1、情报分析含义（掌握一种即可） P177

(1) 尤正平认为，情报分析是一种“环境分析和扫描工作”，即对那些能为某组织提供发展机会或产生威胁的外部因素进行分析。

(2) 包昌火认为，情报研究是根据社会用户的特定需求，以现代的信息技术和软科学研究方法论为主要手段，以社会信息的采集、选择、评价、分析和综合等系列化加工为基本过程，以形成新的、增值的、情报产品，为不同层次的科学决策服务为主要目的的一类社会化的智能活动。

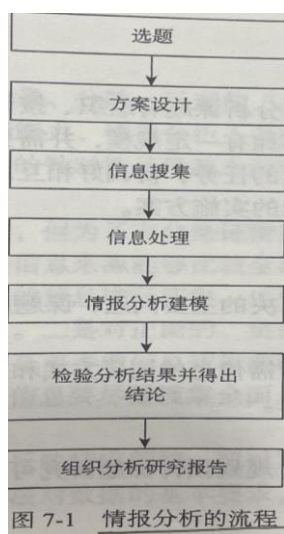
(3) 卢泰宏认为，情报分析可以归纳为从混沌的信息中萃取有用的信息，从表层的信息中发现相关的隐蔽信息，从过去和现在的信息中推演出未来的信息，从部分信息中推知总体的信息，揭示相关信息的结构和变化规律。

(4) 范并思认为，情报分析是利用科学的研究方法对现有信息进行鉴定、分解、综合和推断，得到新的信息产品的工作。

(5) 朱庆华和陈铭认为，情报分析是指以社会用户的特定需求为依托，以定性和定量的研究方法为手段，通过对文献信息的收集、整理、鉴别、评价、分析、综合等系列化加工过程，形成新的、增值的信息产品，最终为不同层次的科学决策服务的一项具有科研性质的智能活动。

2、情报分析的流程和步骤 P178

一次典型的情报分析活动，一般由选题、方案设计、信息搜集、信息处理、情报分析建模、检验分析结果并得出结论，组织分析研究报告等几个环节组成。



信息搜集:

原则 P180:

针对性、全面性、系统性、新颖性、可靠性、科学性、计划性

方法 P181:

1、文献调查：采取科学、系统的检索方法

2、社会调查：是一切以信息搜集为目的的社会实践活动的总称。社会调查又分为现场调查、访问调查、问卷调查、样品调查

3、科学技术情报分析

(1) 科学情报分析 P185：以科学研究论文为主要发展对象；以提高科研和设计效率为目的；离散分布性、相关性和有序性是科学情报的基本属性；科学计量学方法为开展科学情报分析的主要方法

(2) 技术情报分析 P186：利用专利信息开展技术情报分析。通过对专利信息分析，可以获得技术能力指标，指导并聚焦新技术，了解技术开发及其资源分布，评价专利技术隐含的经济价值，用作战略规划和技术开发的精确的数据，以及作为改善和创造其他新技术专利的重要基础。

4、竞争情报分析 P189

竞争情报：关于竞争环境、竞争对手和竞争策略的信息和研究

竞争情报分析方法：如用于行业结构分析的波特竞争力模型、用于市场地位分析的 SWOT 分析

5、情报分析的工作类型

(1) 跟踪型情报分析 P192

跟踪型情报分析方法:

i. 人工跟踪型情报分析

由情报搜集人员自己从各种情报源中搜集、加工和报道情报的过程。

ii. 自动化跟踪型情报分析

利用 RSS 聚合技术、信息抽取和网页提取与监测技术等自动完成情报的搜集跟踪工作。

iii. 人工与自动化相结合的跟踪型情报分析

充分利用自动化情报跟踪技术的检索效率和检全率，在系统完成情报检索之后，再利用人工情报分析再检准率上的优势，由情报分析人员对自动化情报系统跟踪的情报信息进行进一步的筛选。

跟踪型情报分析的成果与报道：

i. 跟踪型快报

最常用的跟踪型情报分析报道形式。即根据需求，即时跟踪、监测和传递有关研究对象的情报。

ii. 跟踪分析报告

通过即时的反馈，了解决策和管理者的深入和现实需要，将跟踪的信息围绕某主题或研究对象系统地组织起来提出的更深入的专题分析报告。

iii. 跟踪数据库

在跟踪型情报分析的基础上，将跟踪获取的情报、数据和信息给“固化”下来所形成的面向数据的数据库。

(2) 评价型情报分析 P195

是针对特定主题，选择特定的研究方法，充分利用特定的分析工具，对来源于各种信息渠道的情报进行分析说明，归纳提炼，比较测评等。

(3) 预测型情报分析 P199

预测性情报分析方法：P201 表 7-3

如：统计分析（eg. 风险分析 相关分析）

逻辑/因果分析（eg. 需求分析、机构分析、利益相关者分析）

综合（eg. 情景模拟、趋势影响分析）

常用分析方法大类可分为：基于时间序列外推的方法、模型类方法、专家观点和情景分析类方法

预测性情报分析实例：P203

第八章 情报技术

1、数据挖掘

数据挖掘是指从大量的数据中提取有用的信息和知识的过程

数据挖掘对象 P220：可以针对任何类型的数据源进行，包括关系数据库（主要对象），也包括数据仓库、文本、多媒体数据、空间数据、时序数据、Web 数据等。

数据挖掘的功能 P220：主要包括概念描述、关联分析、分类、聚类、偏差检测、时序演变分析等

2、文本挖掘技术 P225

文本挖掘是一个从大量文本数据中提取以前未知的、有用的、可理解的知识的過程。可以视为数据挖掘的一种，或者是数据挖掘在文本数据中的应用。

文本挖掘对象：半结构化或非结构化的文本

第九章 用户行为与情报服务

服务目的、对象

用户需求

科技情报服务：做什么？用什么数据？用户需求，任务？（评价、预测、多元异构、用户画像）

竞争情报服务：过程

基于数据挖掘的情报服务：

基于大数据的情报服务：可以做什么、用途、实现方式？

第十二章 当代情报学发展

补充计量学

科学知识图谱：绘制方法…

更多额外

ESI DB

德温特专利数据库

检索流程——掌握

获取数据

知识图谱：流程

计量学定律：应用、局限