**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定律。

**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。20%的核心期刊上刊载了80%的相关论文。与此同时还应用在考察专著的分布；引用期刊的研究，发现影响因子。

**二八律：80%的任务由20%的人来完成，剩下的20%由另外80%的人完成。**

例如：一位领导通常花全部工作时间的20%完成80%的工作。如领导主要工作是决策，但决策所需要时间不会超过20%，其余80%的时间他只会做20%的工作。 一位工程师、设计师或发明家，他的一生中80%的成果、发明，是用20%的黄金岁月完成的，另外20%的成果、发明是其它80%的时间完成的。 一个单位的年支出费用的20%完成80%的工作任务，而另外80%的费用只干了20%的工作。

**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。

**描述：**

若发表1篇论文的作者占作者总数的频率为ƒ(1)，发表2篇论文的作者是ƒ(1)/22,

发表n篇论文的作者数量的ƒ(1)/n2

从公式 以及公式

(其中nmax为杰出科学家中最高产作者发表论文数；m为杰出科学家中最低产作者发表论文数；R为杰出科学家人数与全体科学家总数之比)

**得出结论：**

杰出科学家中最低产作者发表论文数是最高产作者发表论文数平方根的0.749倍；杰出科学家人数与全体科学家总数之比是杰出科学家中最高产作者发表论文数平方根倒数的0.812倍。

**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。

描述：

设有一篇含有N个词的文献（N≥5000），用自然数1，2，3…给文献中的词编级，出现频次最高的为1，其次为2，直到r（r<N），则高频词的频率ƒ与词级r的乘积是一个常数。 **ƒr=c**

**应用范围：**

主要用于指导文献标引、词表编制和情报检索等。尤其是在自动标引中具有应用价值，因此使用频率太高的词和使用频率太低的词都没有多少检索意义，故应当使用词频适中、词义强健的词作为标引词和检索词。

**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

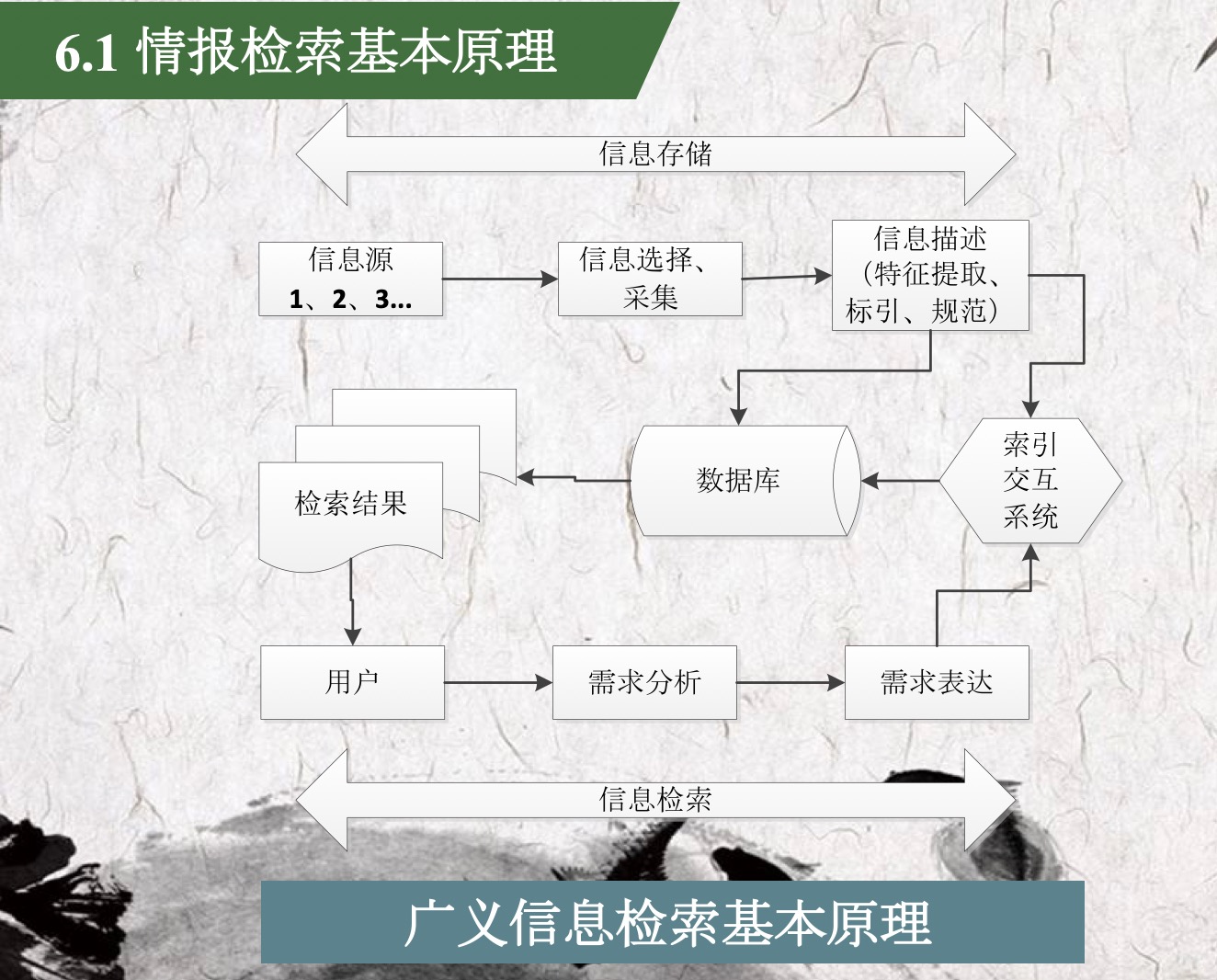
**对于选题：**

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



对于研究：

数据分析的研究流程是什么？

首先确定分析对象，第二找数据、确定数据源，第三对数据进行分析（见8.2、8.3内容p219-p233）第四解读数据结果。

对于结果：

分析数据结果的正反方面意义、启示、价值等。

1. **学术规范概要p47**

包括实体规范和形式规范

**实体规范：概念p47**

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。

**形式规范：概念p48**

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）

**二者区别：**

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。

1. **数学方法p62**

**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。

1. **特色方法——适用于情报学的方法p62**
   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Np篇论文中有h篇每篇至少被引h次、而其余Np-h篇论文每篇被引均小于或等于h次。

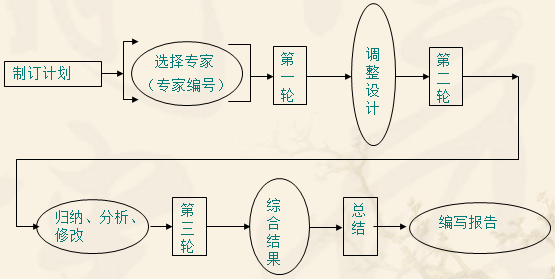
**期刊影响因子：**期刊影响因子，是代表期刊影响大小的一项定量指标。也就是某刊平均每篇论文的被引用数，它实际上是某刊在某年被全部源刊物引证该刊前两年发表论文的次数，与该刊前两年所发表的全部源论文数之比。

**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。

* 1. **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：



* 1. **头脑风暴法（BS法）-p67**

是用会议形式通过专家间的相互交流，激励智利碰撞，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

1. **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

1. **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。

1. **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。

1. **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、实验记录和具体数据等。

1. **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。

1. **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**

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

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**

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。

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。

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。

1. **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

1. **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

1. **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.

1. **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

1. **情报检索评价，理解公式 P158**

R（查全率）=b/a\*%

P（查准率）=b/m\*100%

M(漏检率)=（1-b/a）\*100%=100%-R

N(误检率)=（1-b/m）\*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

**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。

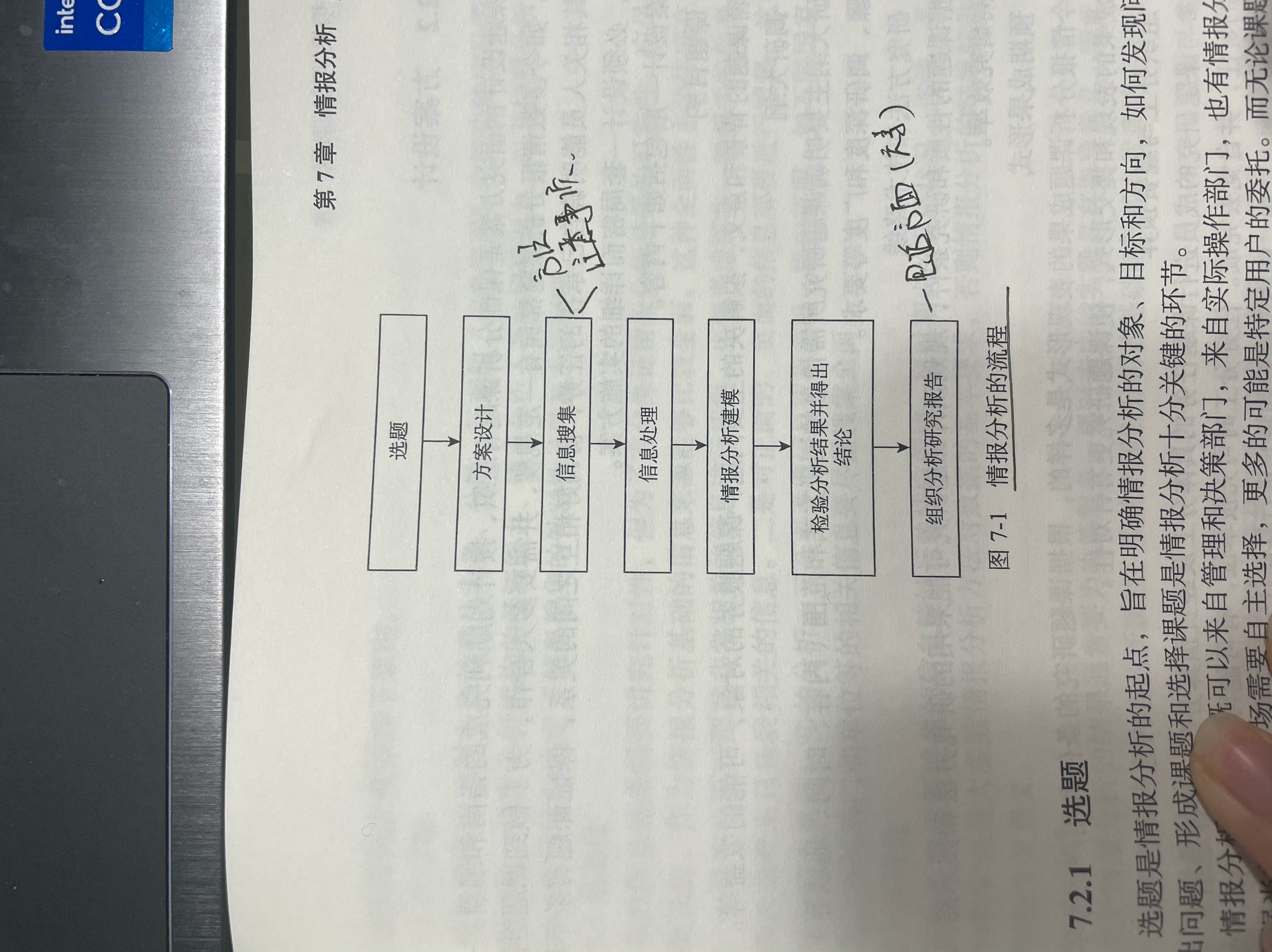
（3）卢泰宏认为，情报分析可以归纳为从混沌的信息中萃取有用的信息，从表层的信息中发现相关的隐蔽信息，从过去和现在的信息中推演出未来的信息，从部分信息中推知总体的信息，揭示相关信息的结构和变化规律。

（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。通过对专利信息的分析，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、可理解的知识的过程。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 用户行为和情报服务P237**

**1、情报服务**

**（1）目的、对象**

情报服务的根本目的是帮助用户克服搜集、选择过程中的困难，及时获取情报信息。

**2、用户需求**

**（1）用户需求**

①本能需求：未用语言表达出来的实际信息需求。

②意识需求：大脑已经意识到的信息需求，但是还不能正确的表达出来。

③正式需求：成形的信息需求，可以用有条理的、有逻辑的语言表达出来。

④折中需求：信息来源所能提供的信息与初始需求之间的折中。

**（2）用户行为研究方法**

①问卷法

问卷是由参与者自行回答预设问题的调查方法。基于问卷方法的研究一般都体现出调查研究的典型过程，包括确定研究目标和形式、抽样、创建问卷表格、招募参与者、分发并回收问卷、分析数据、报告结果等阶段。印刷问卷是问卷的传统形式，通过邮寄的方式发送给符合条件的人，其不足之处在于回收周期长、回收率低。近年来，问卷在信息行为研究中的应用大多采取在线分发的形式，效率高、成本低、样本分布广泛。

②访谈法

与问卷一样，访谈也是以问答为核心的研究方法，但两者的优势分别在于广度和深度。同样的问题在问卷中可能只会得到很简短的回答，但在访谈中可以探讨得很透彻，因为口头对话让访谈对象享有较大的自由度，允许他们充分表述自己的理解和思考，很少受到篇幅限制，而研究人员在聆听的过程中可以随时发现有价值的言论并据此提出后续问题，同时留意访谈对象的非语言反应，如表情、语调、动作等。当然，访谈的灵活性也对研究人员的提问技巧和控制能力提出了更高的要求，而访谈后期的数据处理和分析工作也更加耗时、复杂。

③焦点小组法

焦点小组是经过正式组织、严密规划的讨论小组，小组成员聚集到一起，针对特定主题展开讨论，分享想法和观点。

④日记法

日记是一种基于时间抽样的自陈方法，一般也与其他方法一起使用。参与者在目标事件发生时对其进行记录形成日记，记录的内容可以是对事件的简单描述、解释甚至是他们的个人看法。参与者在日记研究中通常都需要根据预设的问题来提供日记条目，而不是自由记录，同时他们可能需要在研究指定的或是收到提醒的时间点上进行记录。他们在记录时所用到的工具包括纸笔、移动电子设备、在线日记研究工具等。

⑤实验法

在实验研究中，研究人员通常会抽取一定数量的样本，并人为控制所有其他因素对结果的影响，旨在探索客观事实发生的原因，揭示具有解释意义的因果关系。

⑥内容分析法

内容分析法是基于结构化的研究手段和系统的研究设计，目的在于从文本中做出可复制的、有效的推理。内容分析中的文本不仅仅指书面或口头文字，而是泛指所有可以用来传递信息的客体，包括文字、声音、图像等一切可观察的交流产物，有的是实物，有的是数字化的。

⑦元分析法

元分析是对以往多项研究的结果进行综合与解释的统计工具。在本质上，元分析是文献回顾的一种形式，但属于定量综合，研究人员采用系统的方法和程序对单个研究所得到的效应量（Effect Sizes）进行统计分析。效应量是元分析的核心概念，是受观察效应大小的标准化测度，反映了现象的强度。最常见的效应量指标主要有三种，包括标准均数差d、相关系数r和对数差异比OR。一般来说，元分析的过程可以分为6个阶段：查找文献、研究特征编码、效应量计算、基本元分析、调节分析、报告结果。

⑧观察法

观察法强调的是行为的过程而不是结果，在实验室外自然发生，对观察对象的反应几乎没有什么限制。观察作为一种独立的方法具有独特的自然性。这种自然性首先表现在行为方面，参与者的行为是无意识或不自觉做出的，而不是对研究的回应；其次表现在条件方面，研究的条件可以是人为设定的，但是参与者并不知道；最后表现在环境方面，参与者感知到的环境并不是专门为了开展研究而建立的。

**（3）科技情报服务**

①科技情报主要是指反映科学技术相关领域和学科最新的研究成果以及研究动态的信息，具有及时、新颖、方便、简洁、信息密度大等特点。

②科技情报服务是指情报服务机构或人员为用户搜集、处理、研究并提供科技情报的工作，以帮助用户规避侵权、降低成本、缩短研发周期等。

**（4）竞争情报服务过程**

搜集数据、分析数据、面向企业决策使用

**（5）基于数据挖掘的情报服务能做什么**

将数据挖掘技术应用于情报服务中，可以扩展服务范围、提高服务效率和准确性，具体体现在以下两个方面：

①情报获取方式从人工搜索扩展到机器自动抓取

在大数据时代，情报人员通过手工采集、获取情报已变得十分困难，元搜索技术、爬虫技术、数据清理和噪声压制等数据挖掘技术的应用带来了全新的情报思维模式和服务模式，例如Maltego等互联网情报聚合工具可以协助完成开源情报收集和取证等工作，提高互联网情报收集效率。

②情报分析上升为多维分析

利用数据挖掘技术，可以对多种资源进行自动采集、分类和去重等处理，从而从海量数据中及时准确地筛选出关键的情报信息。

**（6）基于大数据的情报服务能做什么（看文献，究竟在做什么）**

例：能做什么？有什么用？怎么实现？

目前，根据大数据环境特征构建的情报服务模式包括蛙跳模式、众包模式、动态运行模式、一站式服务模式等。这些服务模式表现出以下特征：

①动态性、实时性

在大数据时代，借助RFID射频、传感器、社交网络及移动互联网等方式，情报服务机构可以采集并实时传送海量数据，这使得情报服务更强调服务能够即时满足需求。

②智能集成性

由于已有的情报服务信息源充斥着各种结构化、半结构化和非结构化的信息，因而有必要通过智能化的手段将不同形式的信息集成在统一的平台上。

③定量分析与定性分析结合

传统的情报服务以定性分析为主，而大数据环境中定量分析也开始扮演同样重要的角色，通过对大规模数据的分析统计实现对情报的处理与解读。

**第十章 情报教育（了解 读信息）**

**（1）美国日本法国都做了什么？ 课程体系、设置原因P280**

①美国

一些与图书情报有关的学院本科生教育所涉及的领域有：信息技术、情报学、信息系统、信息技术与信息学。

总体来说，美国许多大学的图书情报学教育在培养目标上存在不同的导向。

美国伊利诺伊大学厄巴纳－香槟分校图书情报学院的斯蒂芬•唐尼（J．Stephen Downie）认为，图书情报学的研究、实践与教育总体来说可以分成两派：

一派以用户为中心，侧重点为图书馆学；

另一派以系统为中心，侧重点为情报学。

图书馆学与情报学学科内部也明显分为两派。情报学学科教育中强调用户为中心的一派倾向于人机界面问题，而强调系统为中心的一派则关注考察各种检索算式或标引方案的效益。

图书馆学学科教育中又分为图书馆员类和技术类。

②日本

日本大学的图书情报学教育，主要由图书馆学情报学专门教育、司书教谕（教师）专门课程及以普及图书馆情报学知识为主要目标的一般教育构成。研究生教育主要分为硕士和博士两个层次。硕士教育一般学制2年，招生来源比较广泛，不考虑学生的专业、职业、年龄或国籍，只要有相应的本科学位并通过硕士生资格考试便可录取。

博士教育学制为3年，要求必须是图书馆与情报学专业的硕士生才可申请。

③法国

法国没有图书馆的博士课程，学生要想申请博士学位需要参与和里昂综合大学合作的课程，以获得博士学位。

**（2）情报素养是什么？P294**

1989年，美国图书馆协会的报告中给出了一个使用最广泛、最权威的信息素养的定义：一个具有信息素养的人能够确定何时需要信息，能够有效地获取、评价和利用所需要的信息。

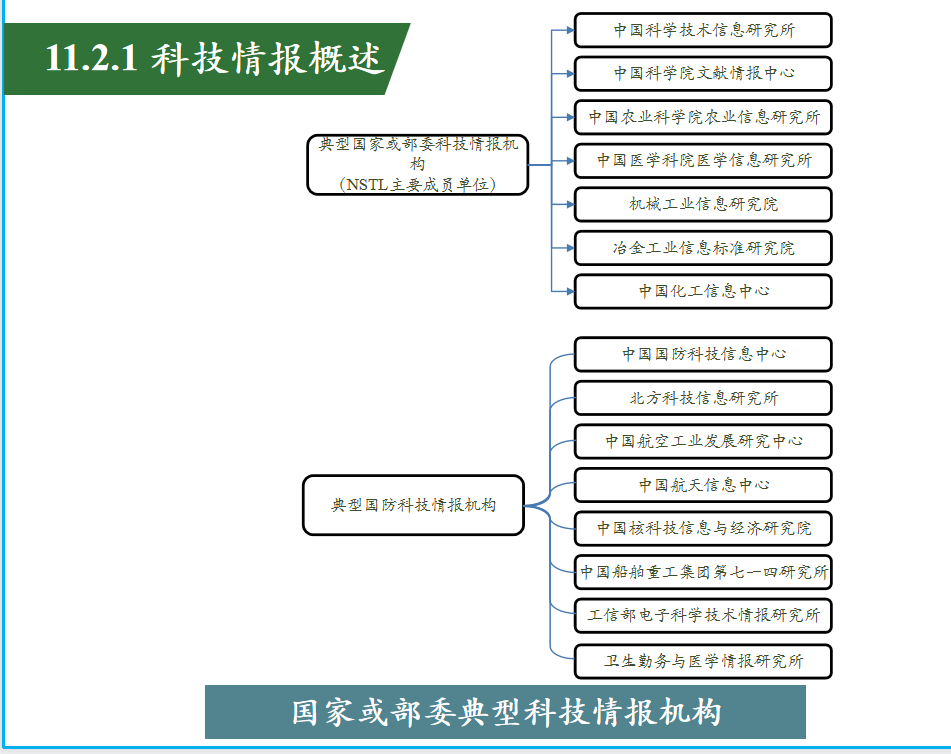
情报素养应是对活化知识掌握和运用的意识与能力，即其内涵包含了人文和技术两个层面的含义：从人文角度来讲，情报素养反映了人们面对情报的心理状态，也就是面对情报时所具备的修养；而从技术层面来讲，反映的是人们把握情报的意识和能力。

**第十一章 行业情报**

**1、科技情报P300**

（1）科技情报是什么？可以做什么事？

情报是指传递着的有特定效用知识。



（2）科技情报使用什么数据——科技信息大数据

（3）用户的需求和任务

**（一）需求**

情报需求分析离不开情报用户的研究，情报用户的研究有任务模型、用户分类、用户信息行为等多个视角。

情报需求是依赖于情报用户而存在的。传统的情报需求往往是由情报用户提出来，或者情报工作人员通过调查问卷或访谈的方式获取。而在大数据环境下，有可能基于多源异构的数据，既包括传统的论文、专利、报告等文献数据，也包括网络、微博、新闻等网络或新媒体数据，通过政策、科学、技术、产业、市场等多方面数据的融合分析，运用主题模型、关联规则等深度学习及文本挖掘等技术可能会捕获出一些情报需求，甚至是一些潜在的情报用户需求。情报用户的情报需求受很多因素影响，主要包括情报用户信息与情报需求情境信息。情报用户信息既包括专业背景、外语水平等静态信息，也包括最近活动、工作关系网络等动态信息。情报需求情境信息包括政策环境、前沿热点、社会焦点问题、行业形势、竞争格局等。

**（二）任务**

①情报识别

为了发挥情报的耳目功能，就需要增强对情报信息的敏感性，从各类信息源中快速识别有用的信息。情报识别的范围比较广，类型颇多，主要分为判别性识别、目标类识别、科技内容识别。判别性识别包括虚假情报的识别，敏感或关键信息的识别；实体目标类识别包括竞争对手识别，合作伙伴识别，专家人才识别，情报用户识别、新产品与新型号识别；科技内容识别，包括突破性创新的识别、研究前沿识别、关键技术识别、共性技术识别、市场机会识别等。这些识别任务需要广泛搜集论文、专利、网页等多源信息，对信息进行比对与甄别，从广泛的信息源中识别出有用的真信息。

②情报跟踪

为了发挥情报的尖兵功能，就需要及时和准确地捕捉最新、最快的情报信息，持续地对关键技术问题、重要国家与机构进行定期或长期跟踪，把握最新进展与动向，梳理技术路线与演化规律。情报跟踪包括科学技术跟踪、战略计划跟踪、重大事件与问题跟踪、国家机构跟踪。科学技术跟踪包括跟踪世界先进科学技术，注重前沿热点监测、关键技术跟踪。战略计划跟踪主要重大计划跟踪、战略实施进展跟踪、政策走向跟踪等。重大事件与问题跟踪如阿波罗计划跟踪、日本福岛核泄漏信息监测等。国家机构跟踪主要指以重要国家或重要机构为跟踪对象，全面跟踪相关领域科技发展、战略规划等。情报跟踪需要及时采集各种渠道的信息，全面跟踪与扫描最新的信息，快而全是大数据时代科技情报跟踪的重要特点。

③情报比较

情报本身具有非常强烈的敌我对比的色彩，所以情报比较在情报工作中占据非常重要的地位，与情报识别和情报跟踪相比，情报比较更加注意竞争关系与态势的分析。情报比较包括竞争对手比较，国家或区域的技术水平对标、核心竞争力比较、创新能力比较。

④情报评价

有时情报评价是情报比较的前提或基础，但有时情报评价并不是为了比较，如对合作伙伴的评价，或对自身的评价。科技情报的评价主要包括科技现状的述评、科技实力的评估、科技机构或人才的能力评价与潜力评估、以及科技政策或举措实施以后的效果评价，如对省部会商制度的评价。由于科学计量学的快速发展，科技评价在情报中的作用很大，目前在情报学研究以及情报实践中得到了广泛的应用，许多科技情报机构都专门建有科技评价或评估中心。

⑤情报预测

预测包括时间节点预测、发展趋势预测、重大事件预测，机会预测等。时间节点预测如基金申请量预测、流量预测、峰值预测等。趋势预测包括技术发展预测，舆情演化预测、政策走向趋势分析等等。重大事件预测包括奖项预测、结点事件预测。机会预测包括技术预见、未来技术机会分析、市场机会分析等。

**（4）科技情报分析和方法**

①多元异构数据融合技术

在数据时代，快速准确地获取多源数据，成为许多工作的基础与前提。多源数据从不同视角反映国家或行业的相关信息，把这些数据融合汇聚在一起进行相关分析，可以相互交叉印证与补充完善，从而更全面地揭示事物联系，为实力对比与评估、竞争环境扫描与态势分析、战略机会的选择与拓展提供有力的数据支撑与决策参考。不同数据源之间如何关联映射，异构数据之间如何统一揭示与加权分析，数据融合中所产生的数据冲突、数据遗失、数据重复等问题都需要解决，如何围绕主题或情景构建一个多源异构的数据资源，涉及到一系列的技术问题。

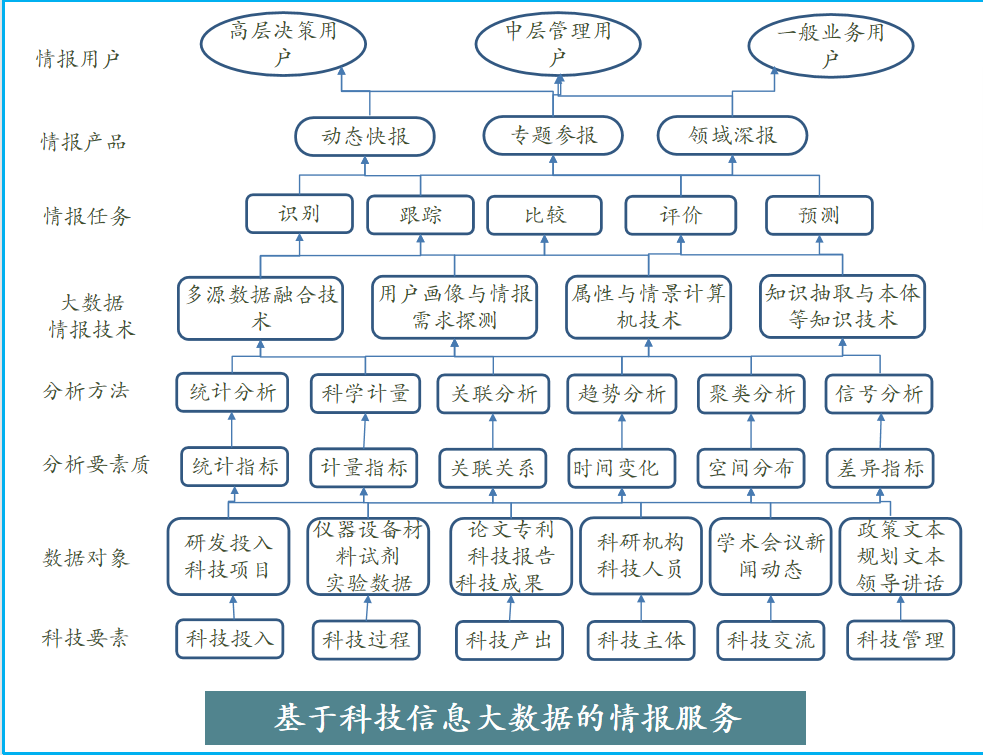
②用户画像及需求探测技术

为了出色地完成情报任务，需要对情报用户进行分析建模，对情报需求进行画像描述。情报用户的需求画像研究包括情报用户的类型确定、情报用户静态属性获取、情报用户动态信息监测，从不同维度为用户赋予合适的标签，既能把用户准确地分到相应的类目下，提取用户的共性群体特征，又能对情报用户的个体特征与偏好进行充分地揭示，准确地刻画出“千人千面”，以便准确地提供个性化服务与精准推荐。

**2、商业 了解**

商业情报往往是为了提高企业的竞争能力、改善竞争环境，所以在某些场景下又称为竞争情报。

**3、基于科技信息大数据情报服务**



**第十二章 情报学的发展**

**1、科学计量学与学术评价P338**

（1）补充计量学是什么？

补充计量学用于评估学术研究的社会影响力，是新的评价方法，用来描述基于网络的学术材料，将社会媒体视为数据来源，区别于传统计量学大多仅是对论文被引用情况进行研究，补充计量学主要用于统计科学研究成果在网络中被浏览、下载、点击、注释、保存、推送、分享、收藏、分类、评论、讨论、引用等的情况。

狭义的补充计量学特指其对社会网络中学术交互行为的测度，而广义上则涵盖了引用数据以外的所有信息源，即使用、获取、提及、社交媒体等。

（2） 为什么提出补充计量学？科技计量学不够用吗？

**（一）补充计量学内涵及特征**

尽管补充计量学的基础是为评价论文影响力的测度指标，其在理论上应用于评价研究绩效的多种层面，与传统计量学引用指标相比，补充计量学具有以下特征及优势。

①信息源多样化

补充计量学不仅能够用于对已出版论文进行计量评价，而且可以对开放式获取论文、研究数据集、电子类书籍、开源软件、演示文稿、课件、网络开放性文章等网络学术资源进行计量评价，从不同层面对所有研究成果的被浏览、点击、下载、评价、引用等情况进行分析，以评价其社会影响力。

②计量时效性强

补充计量学的计量数据来源于网络，充分利用互联网的便利，实时获取和跟踪用户使用网络产生的操作数据，并计算结果，发现规律，得出结论。充分弥补了传统计量学评价结果时滞性的缺陷。

③计量影响力全面

随着学术社交网络的开放和普及，获取学术成果的途径不断增加，更多的研究成果对学术研究者和非学术人员全面开放，专业学术研究者通过浏览、下载、评论、引用等行为来利用网络中的学术成果而非专业学术研究人员大多对学术成果进行阅读、下载、讨论等互动行为，大多情况下，对于学术成果的使用并非只有引用才会给学习者带来影响，阅读、推荐、传播、讨论等也会扩展人们的认知，影响社会的发展，同属于研究成果的社会影响力表现。补充计量学正是发挥其计量独特性测量这类影响力，扩大影响力范围，是评价结果更加准确可靠。

④应用价值更大

除了同传统计量学一样可应用于学术评价，挑选优秀的学术研究成果，为决策科学基金资助提供意见，为科研人员的评估考核提供依据等，补充计量学具有更多的应用价值：展现科研人员的社会影响力；发现潜在合作机会，参考补充计量学的研究结果来寻找对研究成果感兴趣或对其影响力的产生提供贡献值的人，建立合作关系或带动更多的人参与学术研究；促进科学研究的开放、公平，以补充计量学的结果来反映相关研究成果的开放性水平，进而通过对比提高学术研究的整体开放性水平。以补充计量学多元化的计量数据源、计量方法及广泛的影响力表现，使得学术研究更加的公平、公正；这种方法不仅适用于学术研究人员，而且适用于公司企业机构等使用各种在线资源实时地关注其相关信息发布后的影响等。

**（二）补充计量学指标研究及意义**

①为读者筛选优质文章

补充计量学相当于为读者提供了一个实时、开放、个性化的过滤器，帮助读者选择各个领域的优质研究成果，读者不仅可以发表自己的见解还能够与其他读者进行讨论，促进读者对文章的理解、对知识的掌握。

②为作者展示作品影响力水平

补充计量学更为广泛和开放的影响力让作者了解其读者群体和受关注程度，通过捕捉读者浏览、使用等数据来获知影响力产生过程。作为对期刊影响因子的补充，替代计量指标定位自己的科研成果影响力水平，便于其总结自己和制定个人简历，同时作者和读者有关研究成果的交流互动有利于相关研究的宣传和有效利用。

③为资源管理者个性化服务提供参考

补充计量学为资源管理系统用户提供针对性的个性化服务，有利于巩固老用户开发新用户，资源库中成果不断增加的使用率和被引率机理科研人员更多的高品质研究及开放获取，补充计量学相关指标数据为馆藏建设、资源分配、市场推广等提供重要的参考依据。

④为科研管理层人员开展评价工作提供便利

补充计量学为衡量学术影响力提供可靠说明，对于有效评价科研成果、科研人员、科研组织机构、科研院校等的学术水平具有重要意义。

**2、科学知识图谱P345**

**（1）什么是科学知识图谱**

科学知识图谱（Mapping knowledge domain）是一种显示科学知识大发展进程、现状和结构关系的图形，属于数据科学计量学分范畴，是对科学知识的信息可视化表示。最初的知识图谱是在以数学方程表示科学发展的某些规律，进而以曲线图的形式绘制二维图形来表示的。

**（2）绘制方法及常用软件揭示的是什么？（课本）**

①科学知识图谱绘制方法

类似于信息可视化研究，科学知识图谱的使用方法具体包括有引文分析、共引分析、共词分析、词频分析、社会网络分析、聚类分析、多维尺度分析等。其中，聚类分析是以一定的方法将为分类信息按相似程度进行归类的过程，是数据挖掘和知识发现的重要工具之一。多维尺度分析主要凭借观测量之间距离的测定来发现各观测量之间的结构，通过观测量来定位概念空间中的特定位置，使得空间中各观测量之间的距离相似性越来越好。科学知识图谱分析方法繁多，各类方法之间并不完全独立而是相互配合使用。

②科学知识图谱绘制常用软件

目前图书情报领域科学知识图谱绘制工具大致分为两大类：通用软件、专用软件。通用软件如统计软件SPSS，知识图谱绘制时常用到聚类分析和多维尺度分析等、社会网络分析软件如Ucinet和Pajek，常用于展示科学知识结构、GIS等可视化软件；专用软件有文献计量软件如Bibexcel和Histcite；知识图谱绘制分析工具如Citespace和VOSviewer；文本挖掘软件如TDA和Loet Leydesdorff系列软件等。

③揭示？

陈超美认为知识图谱是揭示知识演化进程及机制的信息可视化方法。

额外PPT整理：

**1、ESI数据库是什么？**

基本科学指标数据库(Essential Science Indicators，简称ESI)是由世界著名的学术信息出版机构美国科技信息研究所(ISI)于2001年推出的衡量科学研究绩效、跟踪科学发展趋势的基本分析评价工具。

**2、德温特专利数据库由什么构成**

数据库主要由以下两种文档组成:WPI 文档收录1963~1980年世界20多个国家、地区和国际组织专利局公布的专利文献记录，每月更新同族专利数据;WPIL文档收录1981年以后世界30余国专利局公布的数据，每周更新一次。上述两种文档收录专利文献记录1000万条。

**3、专利可以做什么**

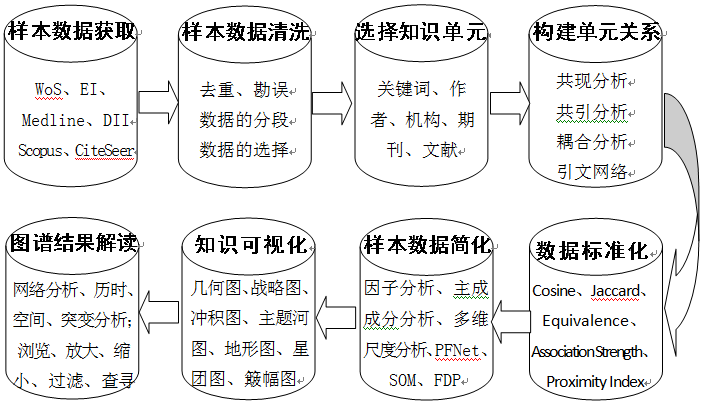
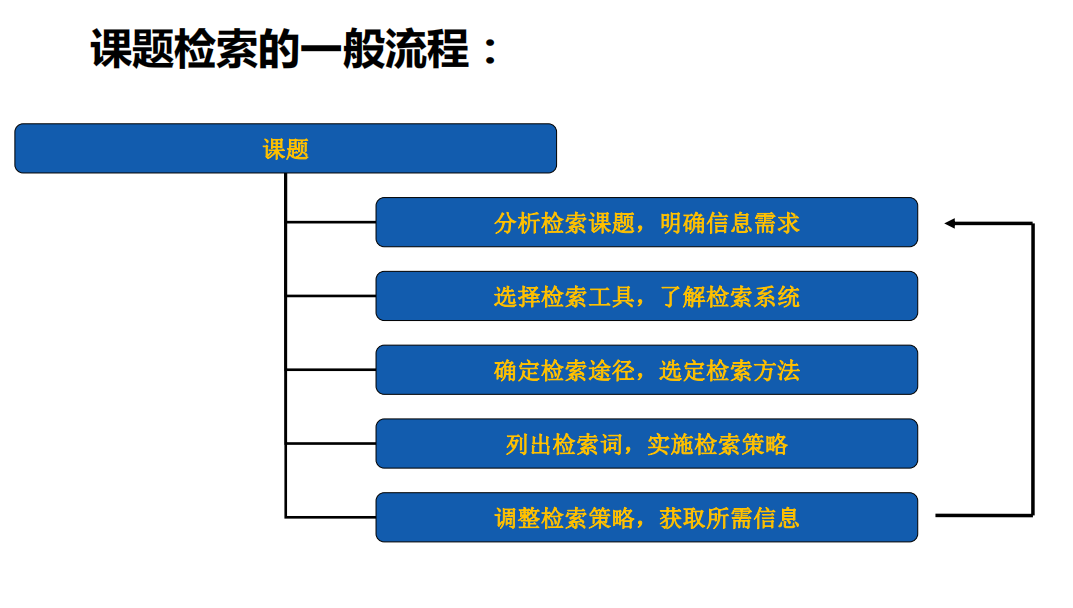
及时了解最新技术研究进展，启发思路，提高科研起点

洞察技术发展趋势，预测技术发展动向

洞悉竞争对手、同行的研究进展

避免重复研究和开发

**4、检索流程**



**5、知识图谱方法 （右图）**

共被引：文献共被引分析是计量文献之间关系的一种新方法。即2篇文献共同被1篇文献引用，这2篇文献就构成共引关系。共被引频率定义为这2篇文献一起被引用频次——马沙科娃、斯莫尔（苏联、美国）