



信息资源管理技术动向

苏新宁





什么叫信息资源管理技术



- 信息资源管理技术就是“获取、处理、存储、检索信息及传递文字、数字、图像、声音信息而采用的方法、设备的总称。
- 简单的说，信息管理工作中涉及到的技术、方法和设备。
- 具体的说，信息采集技术、情报处理技术、信息组织技术、信息检索技术、信息分析技术、信息服务技术。



信息采集技术动向



- 搜索引擎针对整个Web的采集。
- 信息采集特征：针对性、积累性、时间性、预见性、计划性。
- 技术手段：面向主题采集，面向问题采集。
- 未来动向：面向个性化服务的采集、专项服务的采集
用于企业竞争服务、职能部门宏观决策服务。
为社会稳定服务：舆论信息采集，**突发事件信息采集**。
为国家安全服务：危害国家言论、重大事件相关言论的采集。
为科研规划、政府战略决策服务。



信息处理技术动向



- 信息处理就是信息的加工过程，涉及到信息的主题标引、分类聚类、自动摘要等技术。
- 自动标引动向：主题词标引将取代关键词标引
- 分类聚类动向：聚类处理将占据主要份额
- 自动摘要动向：多文档摘要将成为信息分析主流
- 信息的关联处理、语义网络技术将继续得到信息工作的重视



信息检索技术动向



- 传统的信息检索技术将被智能检索技术取代
- 明确检索词检索将发展到语义推理检索
- 云环境下实现分布式检索
- 排除检索技术将在中文信息检索中广泛应用
- 可视化检索技术实用化大众化



信息分析技术动向



- 联机分析OLAP技术将在信息分析中广泛应用
- 多因素分析将成为信息分析中主要方法
- 分析不再限于表面，将寻求数据蕴含的内在规律
- 将更加重视面向战略决策的信息分析



深度分析和规律发现案例



- 引用文献数量

学科学术规范、学科严谨度、学科研究深度、学科研究性质。

- 引用文献类型

学科成熟度、学科成长性、学科资源特点。

- 引用文献文种

学科研究国际化、学科学者的语言能力、学者国际资源获取能力。

- 如何判断学科创新力

新理论、技术采用、交叉学科研究成果分析、跨学科合作、引用文献的回溯与跨学科引用分析、新的关键词出现频率、项目论文的比例、机构在各学科的创新力、影响力、学科研究的资源和环境。



信息服务技术动向



知识服务、个性化服务、战略信息服务将带来更多技术进入信息服务领域

■ 基于语义推理的知识服务

通过推理检索实现的服务。本体技术、知识地图、语义网络技术是实现语义推理的理想方式。

■ 知识点知识推送服务

在线动态提供阅读中专有名词的知识的服务。

■ 基于信息关联的知识服务

将信息经过有机的关联而呈现出的规律和产生出知识提供给用户的服务。

■ 用户行为知识推荐服务

将有经验的用户行为记录，推荐给相关用户的服务。



信息组织技术动向



- 信息检索中语义检索带来数据组织技术的变化
- 信息分析的深入对信息组织的新要求
- 信息服务的发展的信息组织的要求
- 在线阅读对数据组织的要求
- 战略信息分析对信息组织的要求



数据间语义关系的组织



- 数据的概念划分，也就是将数据分门别类归集，制定类与类之间的规则，并确定类与类之间相互联系的语义关系。
- 建立数据单元之间的语义规则、语义关系和语义迭代关系。
- 数据间语义关系的建立应以可扩展、易组合、易操作的数字编码标注，为了简化语义关系标注的工作量以及语义关系的简单化，对于复杂性的语义关系应当由简单的语义关系复合而成。
- 信息的语义关系能够以网络形式展示。



实现语义关系的知识组织



- 数据库中组织



- 关系代码的设置（要求可重用和可叠加）

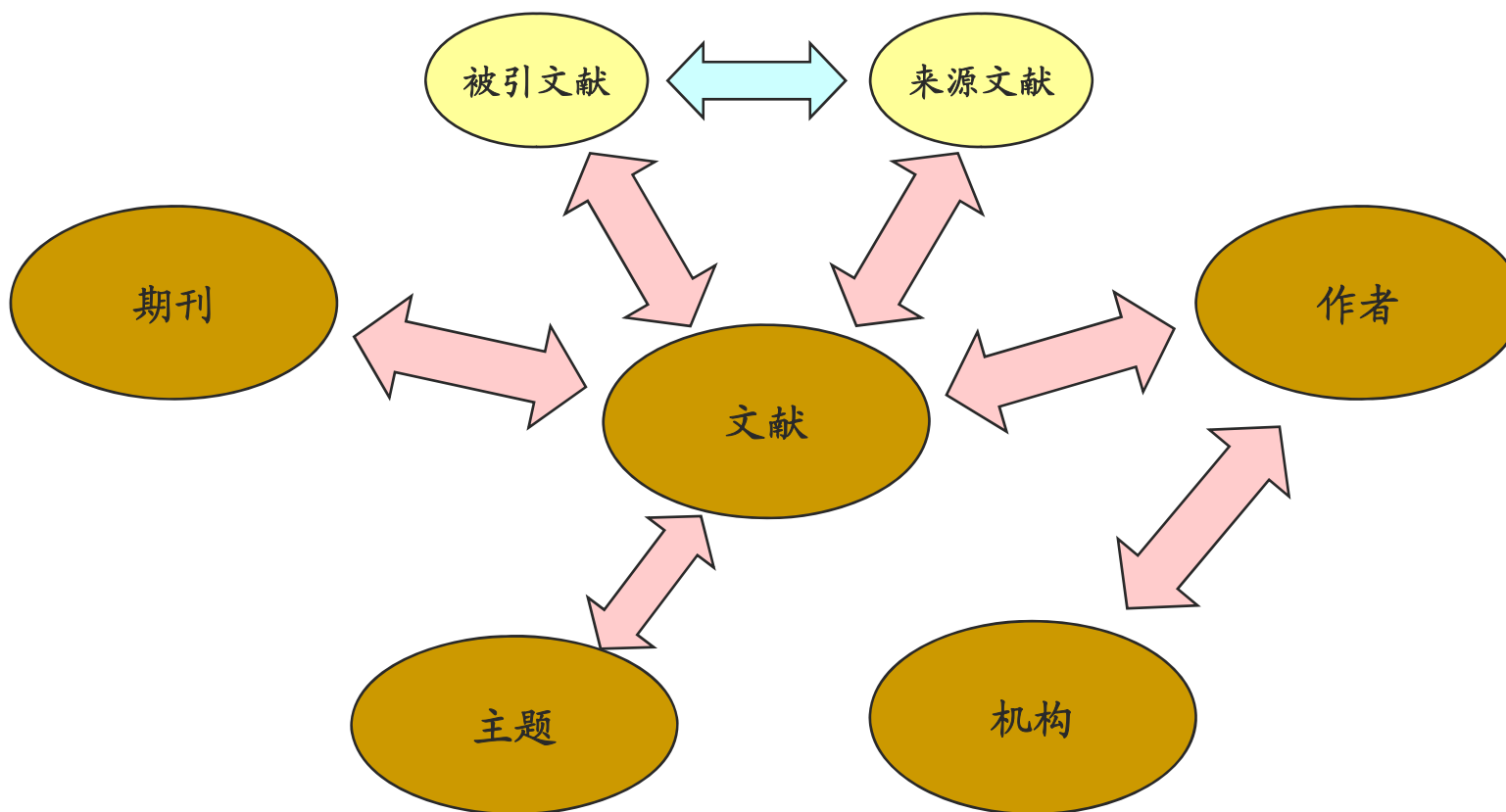
如：1, 2, 4, 8, 16, ……

- 关系的设置

要求简单，能够叠加生成的关系不跨越



引文索引语义本体框架模型





阅读中知识点的知识组织



由来:

人们在阅读文献时常常会遇到一些不太熟知的知识，如果对这些知识的不甚理解，必然影响阅读效率，也不能对文章内容深刻理解，甚至会影响阅读情绪。

知识点的知识组织，就是动态的为这些知识点建立链接，使读者在阅读的过程中可以及时的获取这些知识。



文本中知识点生成示例



■ 阅读文本中的知识点

金砖国家领导人第三次会晤今日将在海南三亚举行，会晤由中国国家主席胡锦涛主持。巴西、俄罗斯、印度、南非领导人应邀与会。南非领导人作为新成员首次参加会晤

传统“金砖四国”（BRIC）引用了巴西、俄罗斯、印度和中国的英文首字母。由于该词与英语单词的砖（Brick）类似，因此被称为“金砖四国”。南非加入后，其英文单词将变为“BRICS”，并改称为“金砖国家”。



知识点知识提供三要点



- 建立专业术语的知识库

知识库中的词条应能够动态增减的，可以通过词条的被阅读的频率决定舍取，也可以及时增加新概念词条。

- 建立知识点生成链接数据库

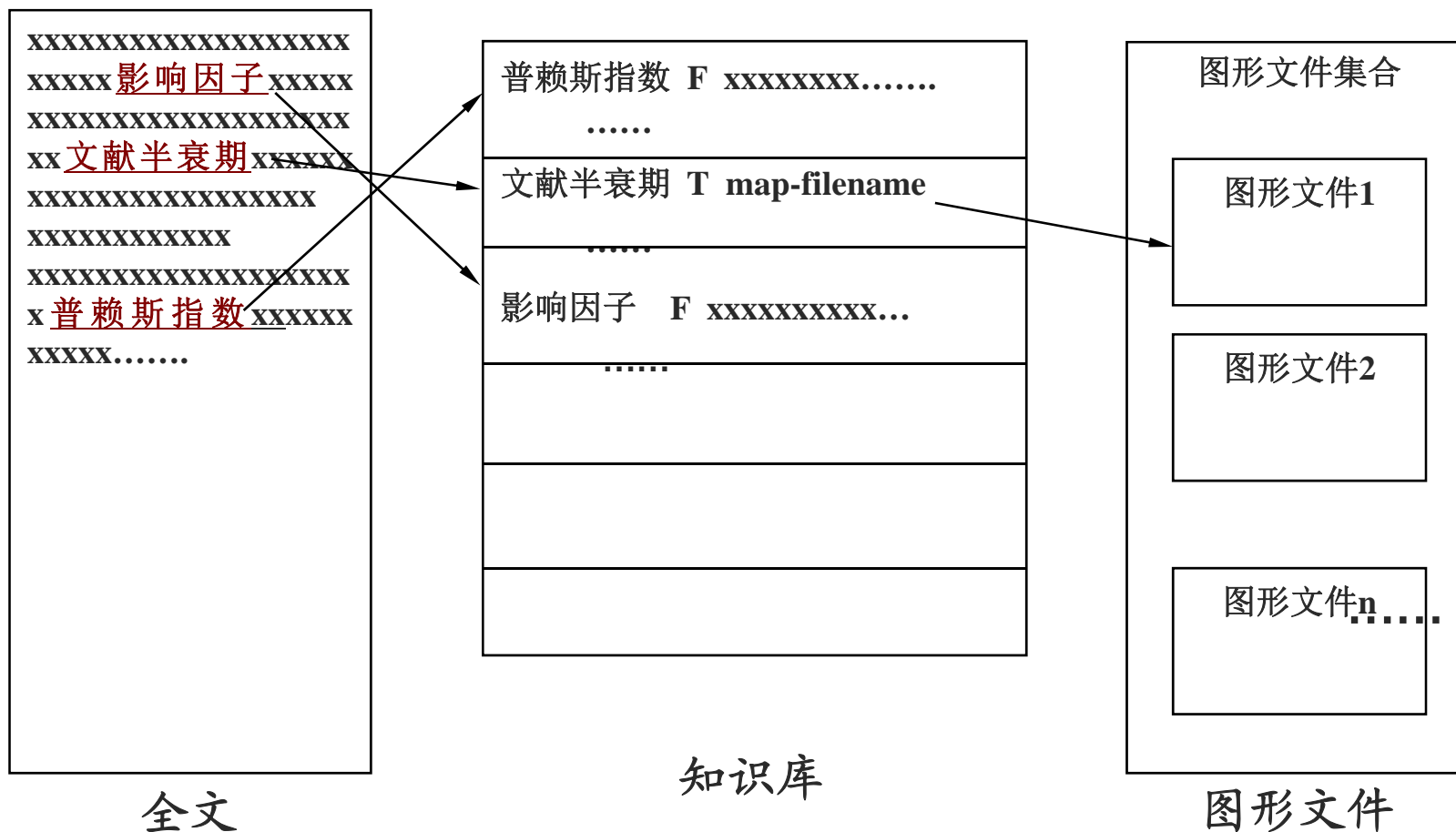
建立与知识库词条对等的动态知识点生成的知识链接数据库，保证动态知识点能够快速生成。

- 知识点动态生成程序，

保证文本输出时知识点的链接自动快速输出。



全文与知识库的关系结构图





知识点服务实现过程



1. 建立知识点动态生成超链之数据库

结构:

词汇	链接 (指向知识描述内容)
----	---------------

2. 欲浏览的全文与知识点数据库动态匹配, 生成具有超链的浏览文本

3. 阅读时可以点击所生成的超链阅读知识点知识。



复杂关系的知识关联组织



- 数据的复杂关系经过有机的联系，可能形成对解决问题非常有用的知识，这些数据和联系关系存储于数据库中，就形成了促进知识服务的高效知识库。



复杂关系的知识组织



■ 数据仓库的逻辑组织

实现数据的多维数据关联，将数据分割成为多维度、多模块的知识元体系。实现通过按照知识元的检索，能够将不同类型却有着密切联系的知识内容的资源跨维度建立联系。

■ 知识地图的数据组织

研究不同类型的知识在数据库中的表现形式，不同知识间的关联如何在数据库中建立起来，知识地图的图状结构如何在数据库中刻画出来，等等。确切的说，就是研究知识地图的存储结构和组织结构。



知识地图数据组织流程



- 从文本或处理信息中抽取对象或主题，并建立和分析之间的关系。
- 根据对象之间的关系，构造向量矩阵。
- 分析关系的相似度，确定阈值，并删除关系不密切的联系关系。
- 根据处理后的向量矩阵建立知识地图关系数据库。
- 根据用户要求，动态生成所需知识地图。



用户行为知识的数据组织



- 用户行为知识的组织是指将用户使用网络信息系统过程中的行为轨迹加以清洗、归纳和组织的过程。
- 目的：为进行经验知识的推荐提供数据基础



用户行为知识记录注意事项



- 信息系统的服务日志数据结构的科学组织

确保用户使用中的信息都能记录下来，尤其是能够反映经验和背景的数据都能记录下来，以方便未来获取、处理和知识升华；

- 构建合理的行为知识存储结构

该结构应能够方便推荐和智能分析，并能够充分体现行为知识的规律性信息。

- 处理并存储有序的行为知识

对用户的行为信息进行知识挖掘，对其进行使用行为归纳聚类，寻找规律，使之具有知识推荐价值，并将处理后的用户行为知识存储到用户行为知识库中。



知识网络的构建



为了提升知识服务系统的决策能力，需要知识服务系统内知识能够形成知识网络：

- 主题知识关联网络的构建
- 知识分类关联网络的构建
- 专家知识网络的构建
- 期刊引用网络的构建
- 文献关联网络的构建
- 学者研究网络的构建

等等，这些知识网络的构建研究都是知识组织研究的重要内容。



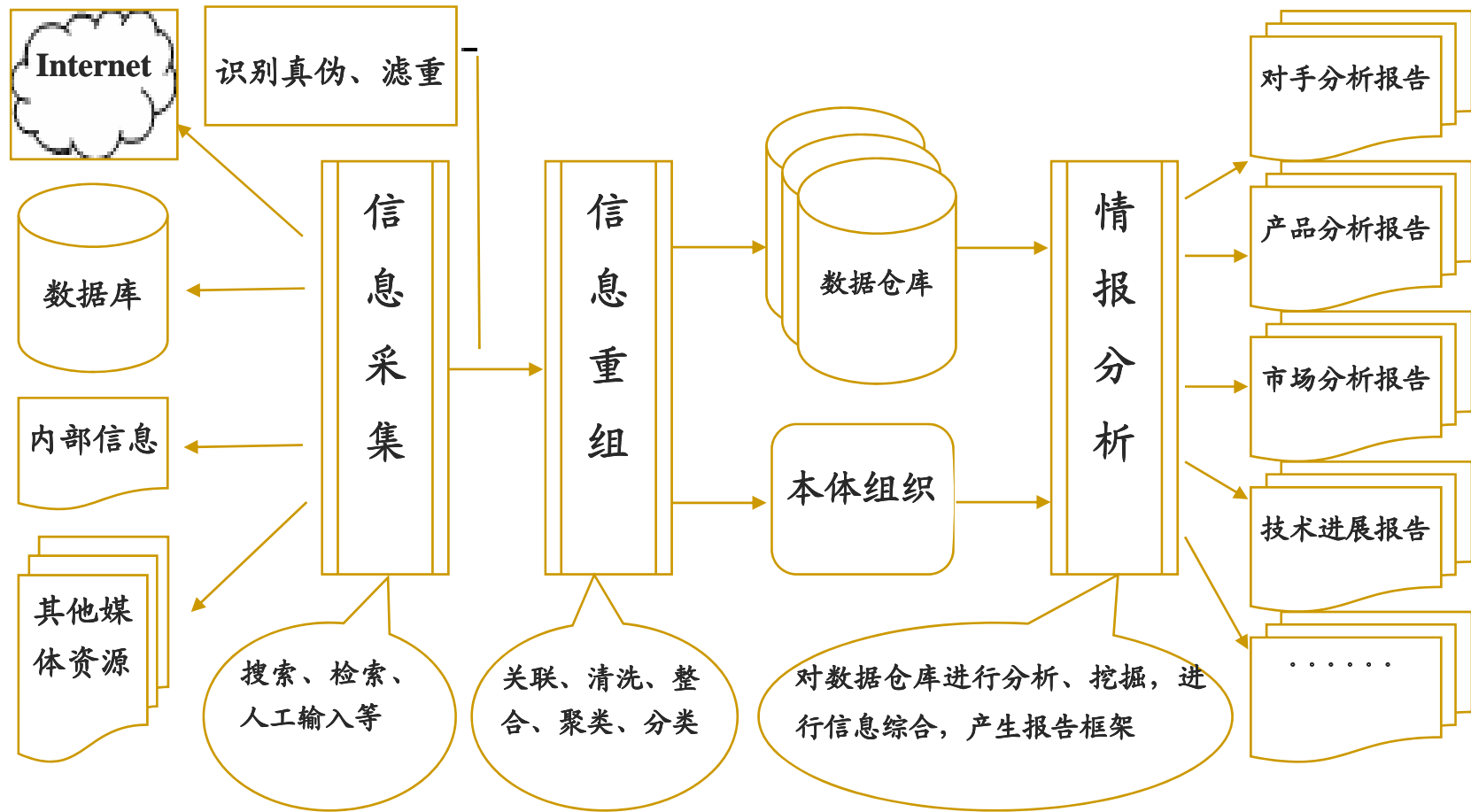
技术的发展使竞争情报系统成为可能



- **数据挖掘为竞争情报奠定了技术基础**
采集、聚类、数据仓库
- **各种数据分析方法为情报分析提供新方法**
聚类分析、关联分析、时序分析.....
- **自动撰写报告提纲**
多文档摘要、数据关联图.....

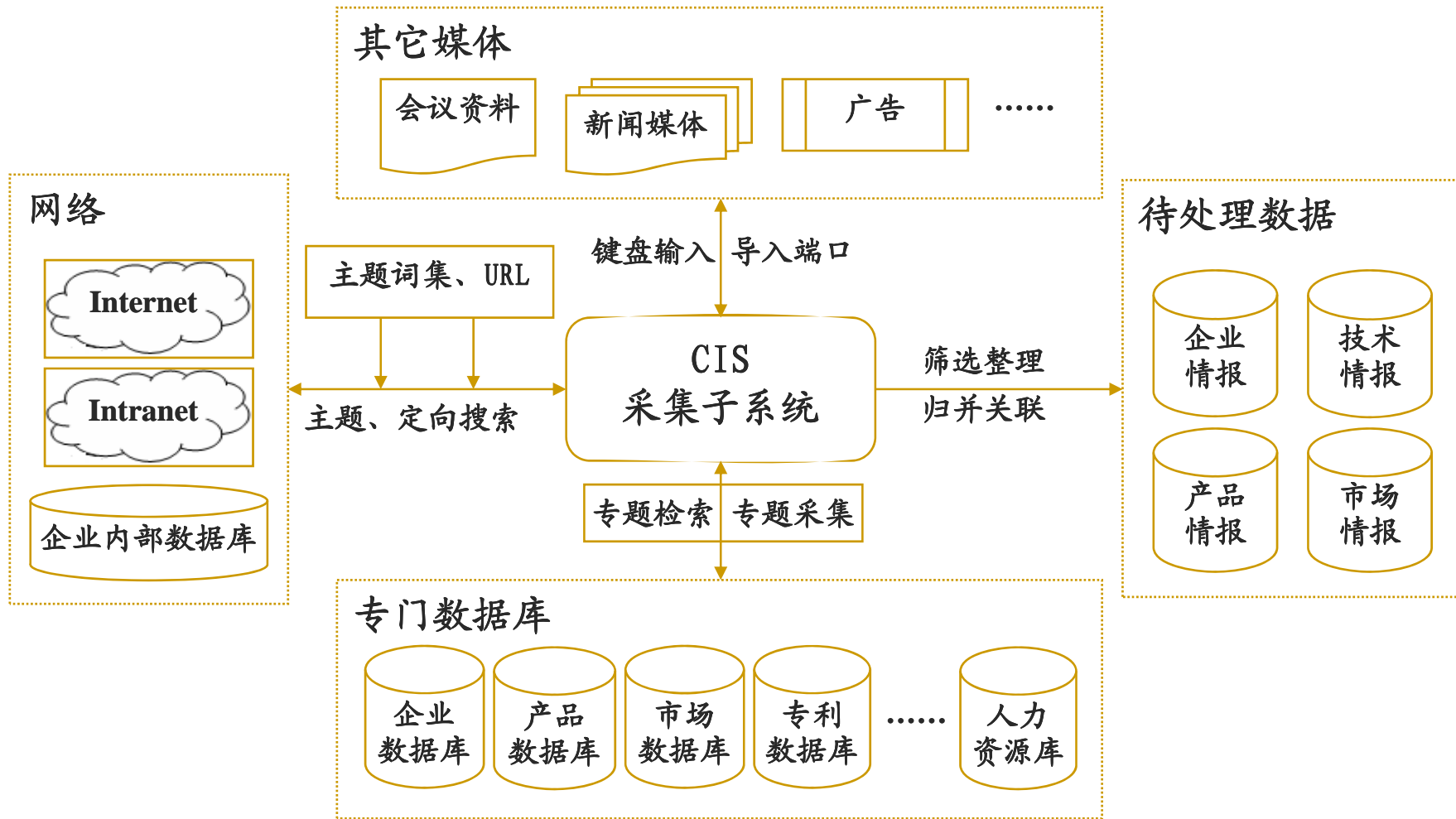


竞争情报系统框架及数据流图



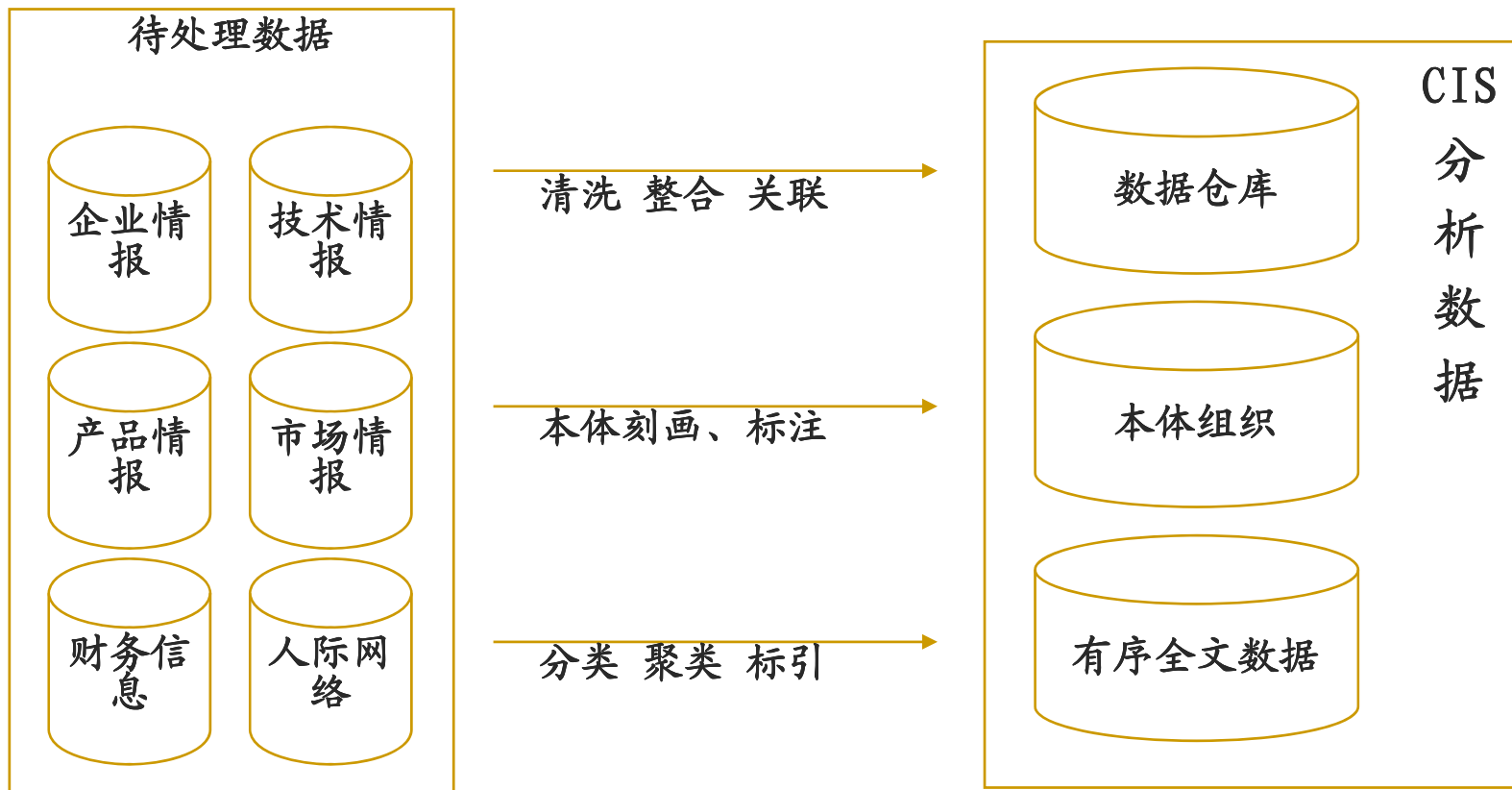


CIS情报采集子系统框架



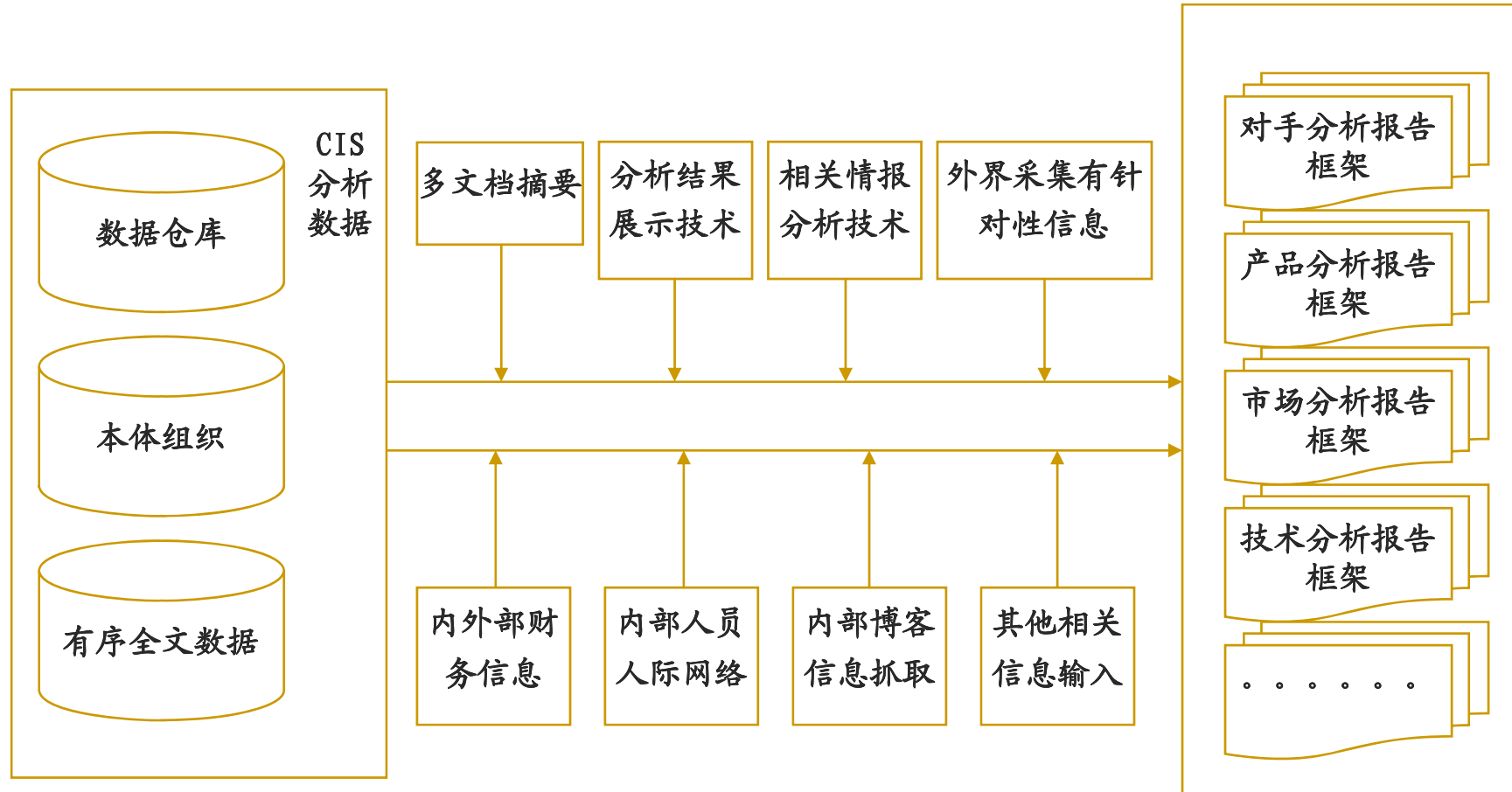


CIS情报处理与组织子系统框架





情报分析子系统框架





结束语



网络技术与计算机技术的发展促进了信息资源管理技术的变革，未来信息资源管理技术的主要动向将围绕战略信息提供、个性化服务、知识服务这些目标，对信息采集、处理、组织、检索、分析和服务等技术进行深入研究，促进信息技术更快的发展。因此，我们应当关注技术的发展，技术的发展将会带来图书情报学研究理论的升华，大量横断科学理论在图书情报学领域得到更深入的运用和阐释，新的理论和方法将会在情报学领域大量出现，给图书情报学学者和工作人员带来了很大挑战，也给图书情报学发展带来更多机遇，这需要我们图书情报学学者共同努力，为更好地为社会提供信息服务、更快的发展图书情报学学科，为图书情报学学科地位的提升，共同努力！



谢谢!

