

基于用户的知识组织与服务

□ 孙卫 / 中国科学技术信息研究所 北京 100038

摘要：对于基于复杂资源类型、海量资源的信息网络化服务，只有针对用户的工作流程、用户的信息牵引条件、用户需求的分类、用户行为习惯的分类建立的知识组织和服务，才能有效地提高信息服务的效率和质量。

关键词：知识组织，知识服务，信息网络化服务，用户行为，数字图书馆

DOI: 10.3772/j.issn.1673—2286.2011.02.008

自从上个世纪九十年代逐步开始使用互联网以来，互联网这种新的传媒方式，极大改变了人们在学习、生活、工作上的习惯。传统图书馆是以进馆读者为主、围绕馆藏资源的（被动式）服务方式，尽管有书目记录、数字资源等网络形态内容组织和信息服务，还是没有从根本上改变图书馆被动服务的局面。

而随着网络、移动通信而成长起来的一代，已经不像上个世纪八十年代以前的人们，把学习图书馆使用与操作课程当作必修课程了，这一代用户今天已经逐步成为推进社会发展的主要群体。如何使图书馆资源组织和服务基本理论与这一代中坚力量用户的需求、行为习惯相关起来，主动为他们的需求提供资源组织和信息服务，是摆在图书馆人面前的现实而急需解决的问题。

传统图书馆尽管增加了很多的

数字资源，依然受到来自互联网搜索引擎对于资源发现的挑战，受到互联网免费资源的挑战，受到数字资源供应商大幅提高价格的挑战。另一方面，传统资源的购置、保存与数字资源的购置、保存的矛盾和竞争也是需要面对的问题。越来越多的图书馆感受到了这些挑战，也有资料预测图书馆未来将在2019年消亡^[1]。所以，图书情报信息服务

体系需要进行思考与选择，建立全新的基于用户的知识组织与服务模式。

1 互联网时代研究者与资源的关系

1.1 数字资源种类多、数量大、复杂性高

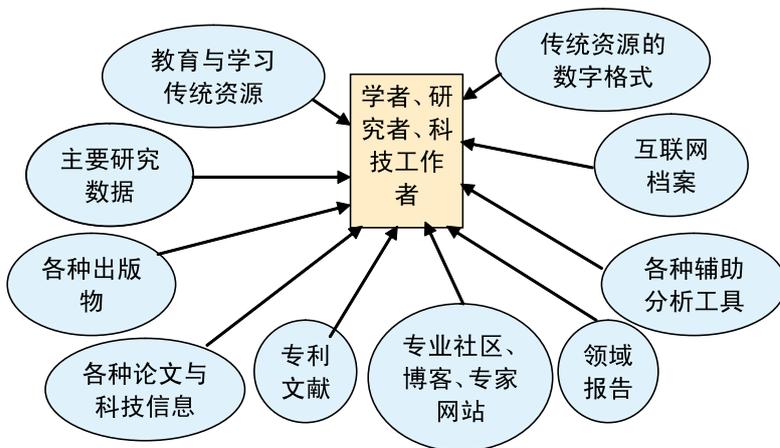


图1 学者、研究者、科技工作者对于资源的依赖

如图1所示, 学者、研究者、科技工作者已经从对于传统图书馆的主要出版物和学术作品的依赖, 转换到在互联网上对于各类出版物、非正式出版物、信息、数据、工具的依赖。IEEE的有关研究也表明^[2], “在学术社区的数字资源中, 有23%电子期刊、20%数据、12%百科、11%社区讨论、7%博客、5%观察、5%预印资源、17%混合资源”。由此可见, 今天研究者所面临的资源种类、复杂性、数量都是前所未有的。而研究者仅仅依靠搜索引擎进行发现、判断、定位、再应用的模式, 是一个很全、很准、效率很低模式。因为搜索到很多杂乱无章的、数量大到几万、几十万符合检索条件的信息; 人们再通过大脑逐一进行分析判断, 浪费了大量的时间; 最后定位到资源, 这些资源可能依然是收费的, 或者是资源无效链接。由于被发现的信息总量巨大、噪声过滤困难、有效资源定位困难, 所以这个模式已经不再适合需要高效有质量保证的研究者对于资源发现和利用的要求了。

1.2 资源组织模式的多样性

大部分学者、研究者、科技工作者是根据自己的知识、专业、行为习惯去发现与利用资源的。在上个世纪九十年代中期, 数字图书馆的研究目标之一是, 利用“图书馆人对于资源的有序化组织和管理的的方法”去实现互联网资源的有序化和服务的高效化。在数字图书馆资源建设过程, 文献服务机构主要利用分类—主题、关键词、名称规范、引文等知识对于资源进行有序

化组织。但是, 大部分使用者不具备图情理论, 而他们自然形成过程中的知识与图情专业对于资源进行组织的知识是不相同的。所以, 数字图书馆研究初期的序化问题解决了, 而服务的高效率并没有实现, 造成了图情专业工作者很大的困扰。

1.2.1 主要的资源事实

图2展现了大家在日常学习和工作中经常会遇到的最主要资源的

形成过程和管理体系。由此可见, 传统的资源组织者是围绕资源管理为目的进行有序化的, 并不是为了使用者按照各自的知识形成体系进行组织和有序化。

在网络化服务时代, 不仅需要熟练掌握图书期刊类型的出版物, 也需要掌握档案资料、专利资料、电子政务资料、产品资料等组织和服务模式。图书馆的信息服务者还没有清晰地意识需要了解与掌握这些传统上不属于图书馆的资源特点。

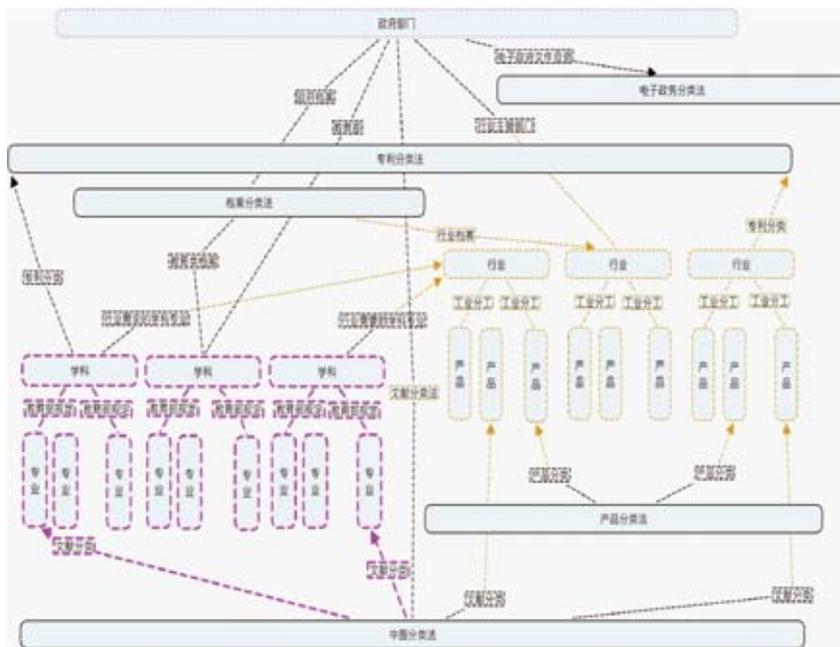


图2 资源形成事实示意

1.2.2 用户知识结构

在图3中可以看到, 只要接受过正规教育的人, 无论是中学、大学还是研究生, 知识体系形成的过程基本上都是一样的模式。在资源组织和资源服务中, 图情工作者需要思考如何匹配这个知识形成过程。

但是, 要注意两个区别, 学

术研究和领域研究是完全不同的含义。学术研究是解决专业理论和基础理论长期应用后可能对理论和基础进行重要发现和补充的阶段。学术研究是一个创新的过程, 但是结果不一定是知识。所以, 学术研究成果, 不能简单称为学术知识, 知识一定要有统计意义上被认同或者普遍的承认、或者可以证明和复用。领域研究主要是解决行业中的

共性的具体问题，阶段性的特点很强，是对实际产业过程中已有的技术、理论、方法、基础等的不足进行的攻关和补充。领域研究一定要被作用于产业，可能会经历很长的时间才能作用于理论形成和基础的知识。

在图3中，专业与学术对于行业专业不是紧耦合，这也说明教育是系统的解决知识学习。进入行业以后，行业中具体特点与专业的结

合、思考才变成行业专业，这是真正能够应用的、产生生产力的专业。同时，由于行业对于产生生产力很关注，而对于产业中预计的问题和研究可能会有缺陷，逐步积压出很多问题。在不同的行业中都有问题，称为共性问题，这就是领域研究的产生基础。所以，领域问题也是阶段性问题，对于行业不是紧耦合的。

1.3 用户对于知识掌握的过程

(1) 数据、信息、知识、工具

在传统的“数据、信息、知识、智慧”的示意中，缺少了重要工具阶段。在上个世纪初期，由于科技不发达，对于人类知识的总结和规范化不够，工具没有被广泛使用，而现在高科技的发展，把人类的经验、方法形成了工具，而且普遍被使用。所以，工具也成为一种不可缺少的知识资源。

对于数据、信息、知识的递推不是绝对的线性关系。对于同一种资源，有些使用者当作工具，有些当作数据，有些当作知识。比如，年鉴整体上属于文献信息（整理后的数据）。但是对于具体的一个栏和内容，可能就是别人进行研究的的数据了。再比如，很多知识受时间、环境的限制而产生，当时间、环境变化以后，知识也可能就是信息或者是数据了。所以，要根据使用的需求来决定资源处于四个部分的哪个部分，或者是混合存在。

工具是把前人的经验、方法、标准、规范等进行总结形成的能提高效率、保证质量、保证规范和一致性的辅助的有形手段。而前人的经验、方法都被验证成为普遍意义的知识了，所以，工具是一种知识的结晶，并作用于数据、信息、知识的各个阶段。例如，微软公司的字处理工具，是最典型的把人类写作与出版排版印刷的经验，规范总结出来，变成可以使用的工具，而且效率、质量得到有效提高。

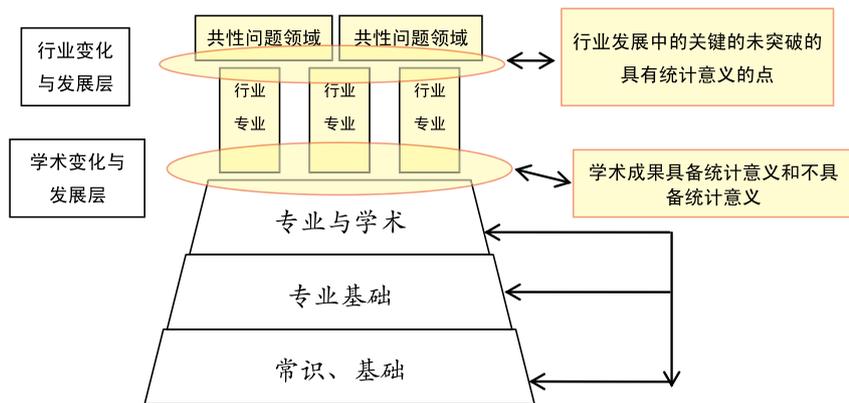


图3 知识形成结构

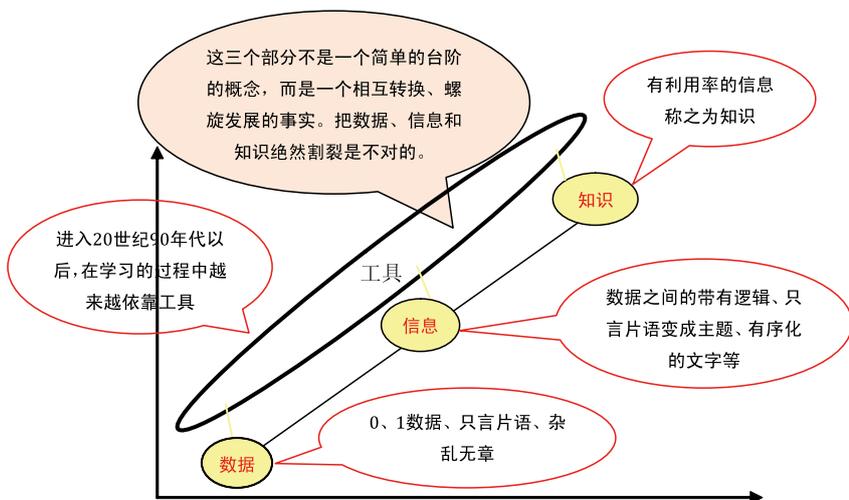


图4 数据、信息、知识、工具关系示意

(2) 知识分类与掌握层次

在一般情况下，把知识分为事实知识、概念知识、元感知知识、程序化知识四个主要的类型。使用者在教育、研究、科技的学习掌握过程中，需要经历记忆、理解、应用、分析、评估和创造等六个阶段。如何把握使用者在这六个阶段中的需求、行为习惯，针对需求和行为习惯进行资源的组织、服务的组织，传统图书馆从被动服务转变

到主动服务就需要解决这个理解用户的问题。

图书馆信息服务人员，首先要对用户进行调研取样，然后给出统计分析的结果。表1给出了知识的四个类型，给出了资源的数据、信息、知识、工具四个形态和用户对于这些知识需要掌握的程度是什么。

握程度的关系。

1.4 用户需求行为习惯与资源组织和服务组织对应

对于图书情报服务机构，资源进行归类整理、有序化后，要面对用户需求进行资源细分组织和针对用户行为习惯的服务组织，把情报资料的主题发现、汇聚、分析和报告能力应用到匹配用户需求、行为的知识组织的方向上来。

(1) 用户需求行为习惯分类

传统的文献管理的分类法是针对文献描述内容 and 应用相关行业进行组织和分配的。而教育、研究、科技工作者对于信息资源的需求主要是按照本身的行业、知识结构来利用的。而文献分类与学科分类的差异是比较大的，并且不是一一对应和匹配的。所以，信息服务机构就需要找到用户针对什么问题需要什么资源的群族关系，把基于文献的组织方法作为基础，在此基础上转变到基于用户主题的需求进行资源组织匹配。

用户使用资源基于很多自己的习惯。抽取共性的习惯，信息服务机构需要去进行用户分析和用户行为习惯的汇聚。

只有把用户按照需求进行分类，把用户行为习惯按照统计规律进行了解和分析，才有可能提高资源的利用效率和服务质量。

图6表明用户分类中聚合用户需求 and 用户行为两个部分。这个需求不是大而泛的，而是主题、时效、针对性很强的需求。这个行为肯定不能是完全个性化的，

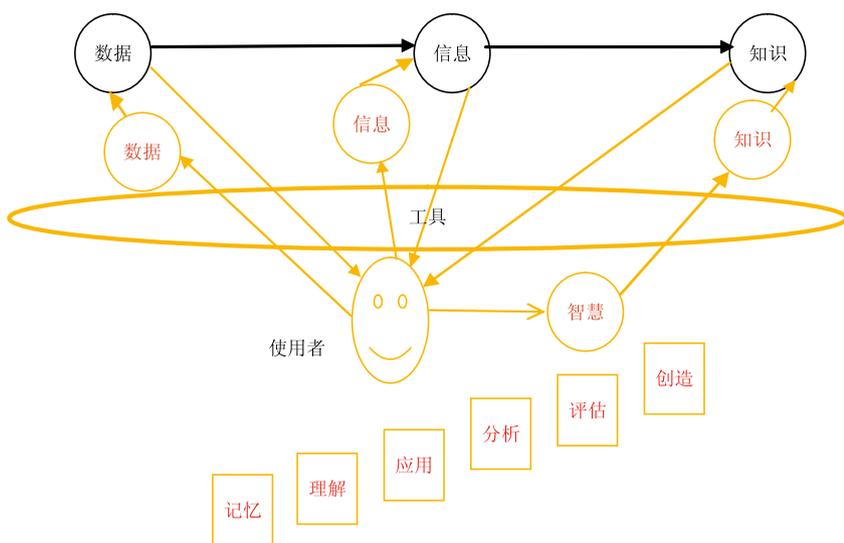


图5 数据、信息、知识、工具与学习掌握的六个阶段示意

表1 调查表 (知识类别、掌握的程度结合)

	事实知识	概念知识	元感知知识	程序化知识
记忆	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具
理解	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具
应用	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具
分析	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具
评估	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具
创造	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具	数据/信息/知识/工具

当然，图书馆调查内容还有其他部分，比如用户的学历、专业、职称、项目、阶段目标等，同时要把自己的资源也进行切分，哪些属

于数据类、信息类、知识类、工具类。把资源结合在一起，取样到一定数量，就可以按照专业、学历、项目等总结出用户的很多需求与掌

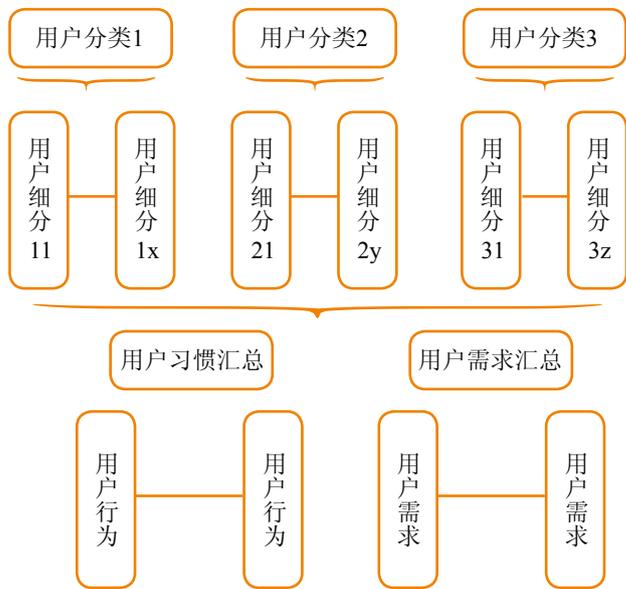


图6 用户分类、需求分类、行为习惯分类

必须按比例归纳，用最合理的服务方式最大限度地符合用户的行为习惯，才能达到提高效率和提高质量的目的。

(2) 匹配用户需求和行为习惯的知识组织与服务组织

传统的文献，一种是按照文献类型进行组织的，比如书籍、学术

期刊、学位论文等；有的资源按照种、册、件、篇进行分类管理。这个分类如何匹配到用户需求呢？怎么可以把主题、时效、针对性深入到种、册、件、篇里面去进行组织呢？如何满足非全文献阅读型的学习、研究、科技工作者的需求呢？

传统的图书馆，主要是查询一定位一阅读模式的，针对全文献阅读需要，提供检索服务（传统

查询服务的计算机化）、定位服务（馆藏位置、数字资源的定位）、获取阅读服务。对于数字图书馆，有很多新的行为习惯，比如研究需要的辅助手段、利用资源的深入、快速等。

要真正把用户类型、用户需求和用户行为搞清楚。图7表明不同的使用者对于不同的使用目的，到底有什么资源使用的习惯和行为；如何对这些习惯和行为进行统计，变成公共的行为（比如检索）、群组行为（比如按字母排序）等。通过不断的细化和总结，逐步做到服务方式匹配用户的行为习惯，这种服务的效率和质量才能提高。

要把服务按照匹配用户的行为习惯进行组织。图8是把行为、需求归纳以后的工具组织和资源组织关系。公共工具服务组织、群组服务方式组织、资源/工具组织，分解到各种对应的服务方式，然后在分类资源的基础上，把资源按照需求进一步细分到不同的服务工具中去，这样才有可能使得服务效率、质量得到提高。

把用户需求分类、行为习惯分类，与服务细分组织、需求资源匹配服务的模式概括到图9，就是一个新型的、符合用户需求和行为习惯的服务组织与资源需求组织的模式。

仅仅把用户按照需求分类、按照行为习惯分类，只是完成了服务组织和资源细分组织的基本任务。了解信息在使用者工作过程中的作用，是更好地理解使用者和匹配使用者、更好地为用户服务的关键。

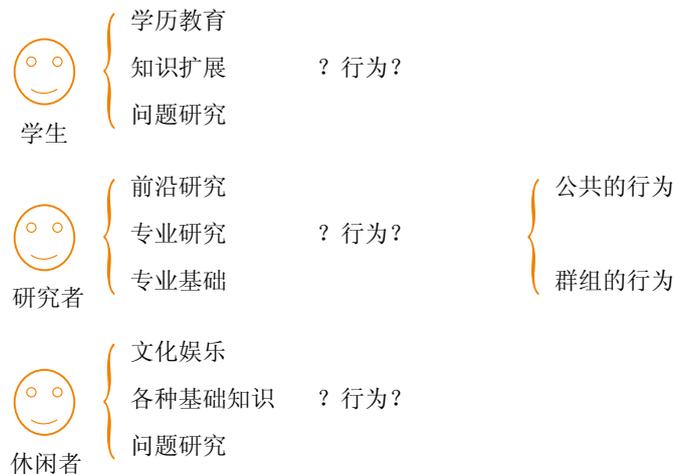


图7 使用者不同、资源需求目的不同的行为聚合示意

(3) 工作流程与资源需求匹配

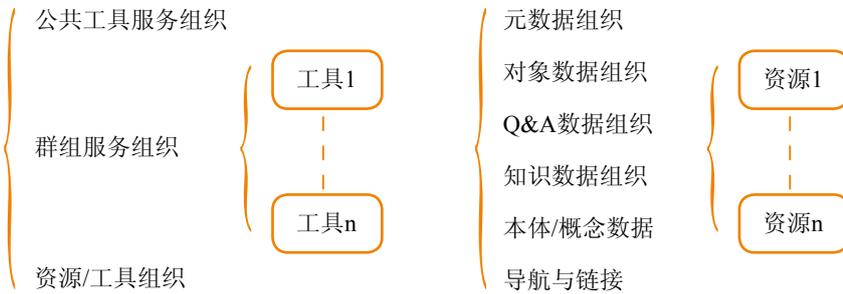


图8 工具分类、服务对应、资源按照服务再组织

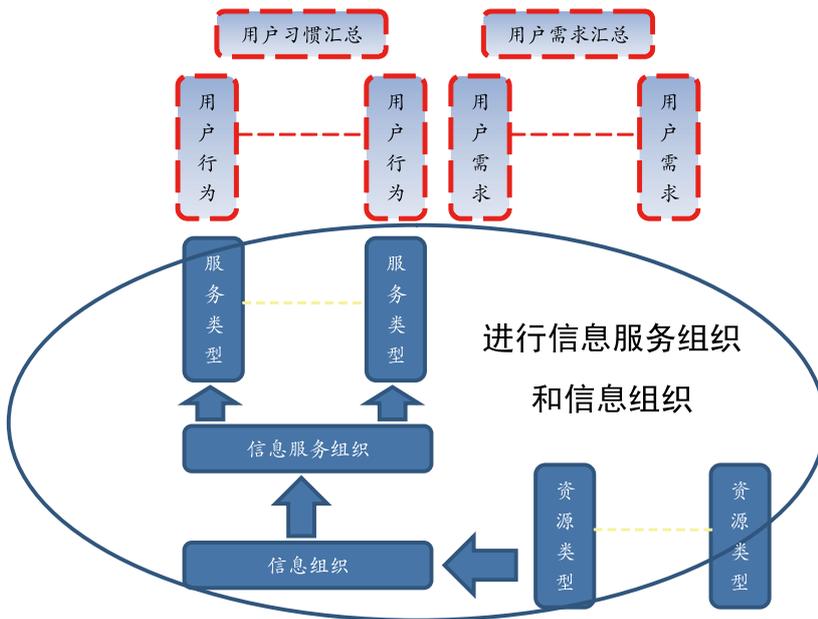


图9 资源的组织满足用户需求、服务的组织满足用户行为示意图

图10是一个基本的产品生产过
程，每个群组和行业中的具体项目，都可以抽象出一个相对准确的工作流程关系图。图11说明了这个产品生产过程对于资源类型的需求是不同的。图10和图11说明产品产生过程和阶段中与需求之间的关系不同。图书馆的信息服务人员，要很好地对各种通用的过程/流程进行总结，用一个总结过的流程图，和使用者讨论资源需求，这样才能做到有的放矢。

同理，图12是一个专利寻求市场、资本的一个过程阶段示意。图13是匹配这个过程和阶段的资源需求示意。

(4) 信息服务牵引条件

用户需求与行为分类、工作过程和资源需求匹配都是建立信息服务模型的基础。而了解到到底是什么激发了用户去寻找资源，是工作中的问题，还是听到一些信息，还是因为要思考问题？了解与建立信息服务牵引条件类型也是非常重要的。

图14是一个产品研发的过程，这个研发过程激发不同的信息需求牵引条件。

主要牵引条件是：

- 消息 (Message)：市场需求 (Market Requirement)、新科技发展 (Science and Technique)、专利 (Patent)、问题 (Problems) 等

- 思考与论证 (Think and Reasoning)：专利 (Patent)、科学技术 (Science and Technique)、核心人物 (Core Leader)、投入产出 (Investment、Cost、Return)、风险 (Risk)、原型 (Model) 等

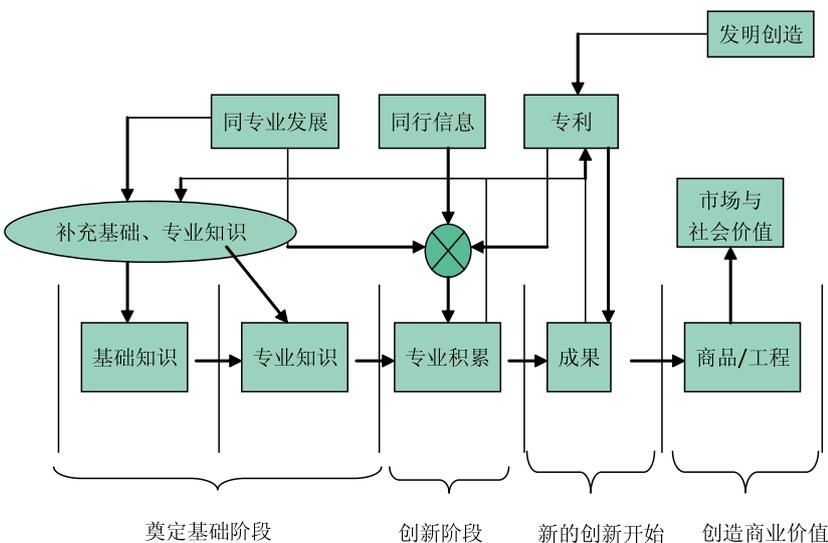


图10 产品过程示意图

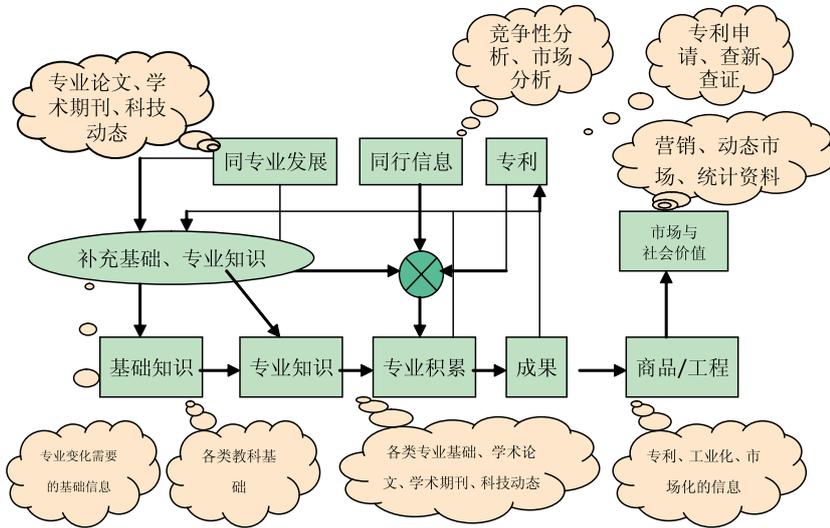


图11 匹配产品生产过程的资源示意

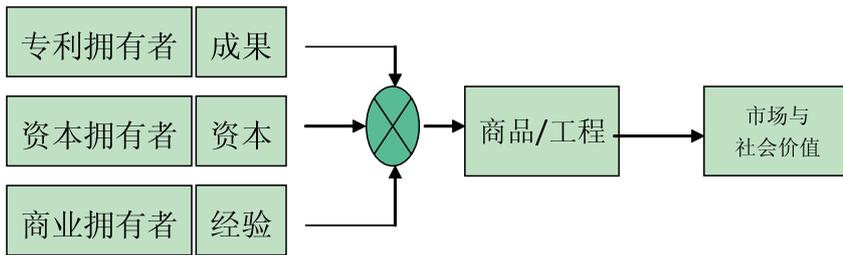


图12 智慧、资本、产业化共同作用下的商品市场

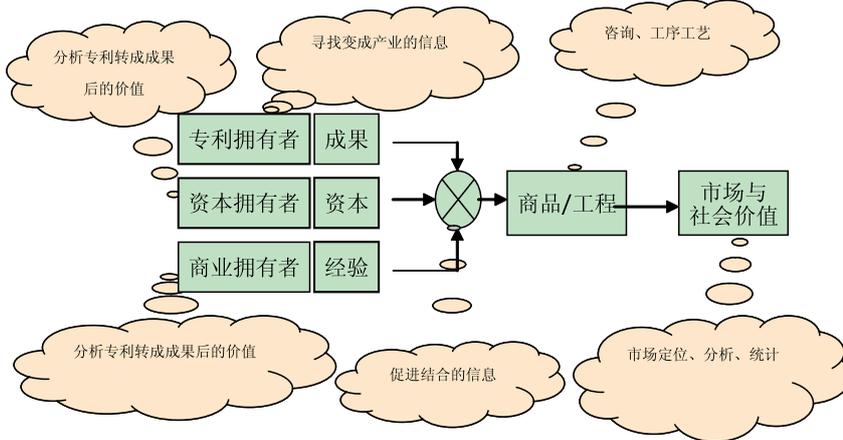


图13 用户需求与信息服务的定位关系示意

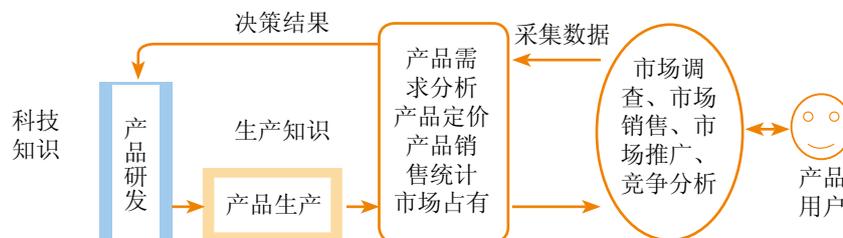


图14 产品研发过程

● 决策 (Decision)：投入产出 (Investment、Cost、Return)、风险 (Risk)、原型 (Model)、核心人物 (Core Leader)、竞争能力 (Competition)

图书馆信息服务人员要不断地把这些因素作为区分资源类型和作用的条件。综合建立适合各自信息服务系统服务对象的信息模型，是图书馆信息分析与处理工作人员需要掌握的信息素养的很重要的部分。理解用户，是真正使得用户喜欢图书馆、离不开图书馆的关键。理解用户、用户的工作，有的放矢地进行知识组织和知识服务，是有效地提高信息服务的效率和质量的的关键。

2 基于用户的知识组织和服务

2.1 基于事实知识的知识组织和服务组织

用户知识结构与文献内容组织中，人名（人物、作者）、机构名（及其演变）、地名（及其演变）、时间名、引文等知识属于事实知识，也是用户和图书馆管理者共性的知识。在传统文献学中，把名称规范的内容隐性化，利用这些事实进行文献连接的服务。而在数字图书馆服务中，专门建立的作者、机构、引文、项目这类事实知识库，可以在这类事实库和元数据关联中，进行人物、机构、项目等的分析服务，利用这些分析服务进行命题、研究方向的跟踪等用户需求的匹配，提供专门的分析工具进行这个服务。

中国科学技术信息研究所承担的国家基础科研工作的科技评价

表2 牵引条件与资源的关系

牵引要素	资源来源	依赖性
市场需求	互联网新闻、用户调查、产品销售数据统计分析、行业协会报告	依赖调查手段、工具、网络环境
科学与技术	书、刊、产品手册、技术档案、互联网新闻、科技项目、科技报告	图书馆、资料馆
专利	专利数据	专利分析工具
问题	期刊、专著、产品信息、科技报告	调查、分析
核心人物	专利、国家各种研究项目	
投资、成本、回报	市场同类产品、行业协会、财务数据、竞争对手、市场占有率	风险分析工具
风险	专利、项目、市场、行业协会、国家政策导向	
原型	书、刊、产品手册、技术档案、专利、用户调查	
竞争能力	评价评估指标	竞争力分析工具

分析项目，就是基于这种共性的事实知识进行组织、统计分析和可视化处理的。它可以帮助使用者在建立项目、申请项目的过程中，进行分析和研究命题。这个系统（<http://168.160.200.61>）已经提供试用，该项目也是第一次把研究与应用工作进行了结合，改变了过去研究成果很难得到应用推广的局面，为教育、研究、科技工作者提供多角度研究的服务。同方知网的知网节和万方数据提出的五要素分析与处理的方法都是基于事实知识的组织和服务的。

由于这类知识多数在各种文献中出现过，可以针对这些文献提取、整理这类知识点，建立独立的知识点数据库。这样就使得使用者的知识点和组织过的知识点是共性的事实知识。这个部分可以在很大程度上匹配用户的自然形成的知识。而对于技术方法的自动处理，目前还没有更为准确的算法进行这类知识点数据库的自动提取与自动

进行相互间的关联，而图书馆的很多资源管理方法中的名称规范已经建立了这类知识点处理的方法。

2.2 基于概念的知识组织和服务

叙词表是一个典型的概念知识范畴。在叙词表中，对于某一类文献有一定使用频率的叙词（不是最高频率，也不是最低频率），称为主题词；而对于在一个具体文献中表述主要含义抽取出来的叙词，称为关键词。所以，叙词里面含有各种专有名词、主题词、关键词。而这个叙词表不同于信息服务商自由提取的各种组合词的词表。

叙词表主要由以下名词术语组成：（1）表示具体事物名称的名词术语；（2）表示事物的状态或者现象的名词术语；（3）表示科学门类的名词术语；（4）表示研究方法、技术方法的名词术语；（5）表示工艺方法、加工技术的

名词术语；（6）表示化学元素、化合物、金属材料与合金的名词术语；（7）表示国家名称、地名、组织机构名称及人名的专有名词以及文献类型、文献载体的名称术语。

由此可见，在事实知识中，人名、机构名、地名等是有特征的（特征、内涵、关系）。而在概念知识中，则只有名词术语而没有特征部分了，这个就是知识组织层级的方式和方法。

在数字图书馆服务中，专门建立分类、主题知识库，利用主题关联、相似建立文献关系。传统的文献服务只能通过词发现和文献链接，无法在知识角度和知识关联中进行转换。

（1）基于词关联分析的复合检索

大部分专业工作者使用谷歌，在输入检索词时，会显示推荐词组合，供输入者思考与选择。谷歌的关键词提示法，是基于已经发过这个词根的人用词的统计而进行推荐的。但是，由于每个人的知识结构不同，从事的工作不同，需要资源的目的不同，科学、技术、生物、医学等专业者对于这些推荐的选择概率并不大，依然是自己按照自己的特点去输入检索词。另外，谷歌的资源组织中，利用自动摘要（多层次语义收敛）、词提取、字提取等技术构成资源元数据结构，对于细致的专业（行业）的分类是没有专门进行知识组织的，这样的自由词检索^[3]的噪声就非常大。只有依靠输入者的经验，合理地进行检索词的组配，才能有效地提高检索的效率，同时具备一定的质量。而大

部分专业人员对于检索词的组合类型并不熟悉。

而对于文献组织者，对于文献进行分类、主题、关键词的建立和管理，这些词汇也没有被用户有机地进行组合和掌握。在被组织过的文献中，对于分类、主题、关键词这些概念知识，依然用简单输入检索词的方式，是无法改变检索的效率和命中率的。

那么，利用使用者输入的检索词，对概念词组织进行关联性分析，在这个分析的基础上，给出资

源组织中概念词同时出现最多的那种组合，就可以把简单输入的检索词（个人的）与资源组织的概念关系关联在一起（组织过），使得利用概念知识组织收敛到对应的资源上来，这样的知识组织和知识服务是可以进行尝试的，也是避免检索效率、检索命中完全个性化的方法之一。

图15就是数字图书馆关键词和其他词一起出现的关联，及其关联系数。

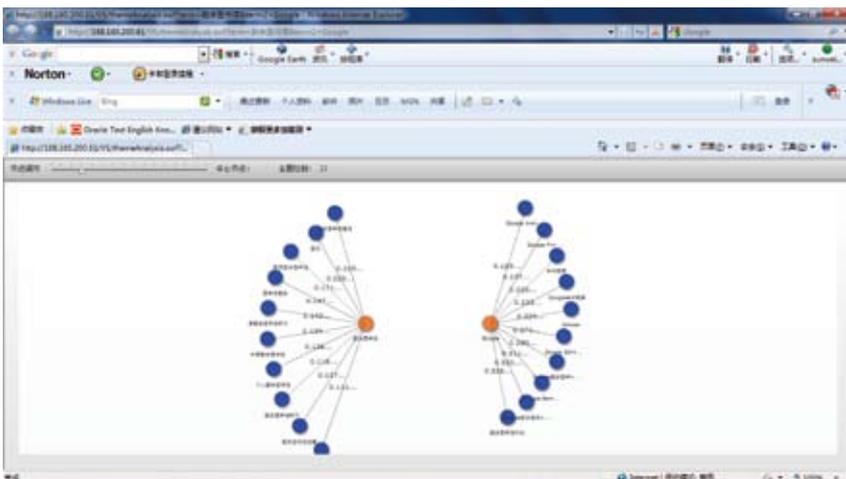


图15 数字图书馆与Google与其他关键词同时出现的概率示意



图16 检索数字图书馆数字资源

(2) 基于主题与资源链接的知识组织和服务

传统的中图分类法是可以有基本的树形结构的，而主题词与分类相关，如何把与分类树的相关主题图树建立起来，比分类树相对的群组化要精确，且复杂度较低。

基于主题图的服务模式，首先是检索关键词，出现命中结果，见图16。

选择关心的结果，然后选择主题分析，见图17。

选择关心的主题，看看和哪些文章相关（见图18）。点击关联的点，直接连到文献，这个是一个初始的主题图模式，方便那些主题明确、直接关心内容的使用者高效率地进入资源。

这个使用方式，可以在分类一主题中，对于每个分类中的主题词建立主题词视图。使用者遍历分主题树，确定分支，进入具体主题，找到对应的资源。这样把检索词检索、过滤、排序收敛的检索模式，变成了分主题树形显示、直接进入资源的知识组织和服务的方式，可以提高效率和保证质量。

2.3 基于程序的知识组织和服务

如何把人们已经形成的过程、方法和经验积累起来？有些过程，是长期经验与方法或者规定下来的，例如查新查证方法。有些方法是某一群人做事能够使用的，例如统计分析方法。还有一些是对于具体的事情积累的经验，当遇到这类事情时，这些经验就是知识。所以，过程、方法、经验统称为程序化知识。就是在前提（环境变量）



图17 出现与这些文章相关的主题示意

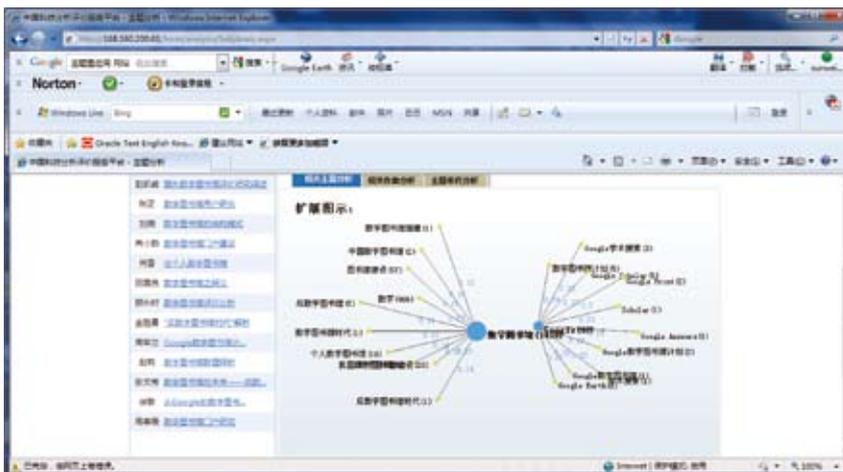


图18 主题—资源图示意 (选择数字图书馆、Google)

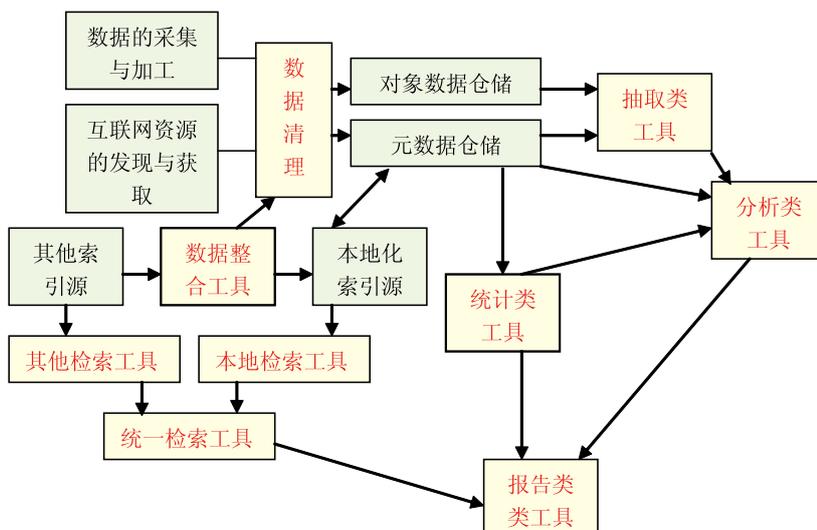


图19 各种工具与信息服务平台部署关系示意

条件一致的时候, 这些知识可以高效地被复用, 提高效率和质量。

(1) 工具类型与作用

最常见的工具类型有: 检索类工具、整合类工具、统计类工具、分析类工具、报告类工具等。

● 检索工具、整合类工具被广泛地在线利用到了在线信息服务中。

● 而大部分的统计类工具是离线应用, 把结果形成报告再在线利用。这样, 不利于使用者利用原始数据, 根据使用者角度思考统计工作。所以, 类似于汤姆森数据分析(TDA)的工具还没有被在线广泛使用。

● 分析类工具, 数据挖掘、关联分析、建模分析等工具更是没有被广泛应用到文献在线研究和发现隐形内容上, 最典型的这类工具是德温特的专利分析工具, 国内开始逐步提供在线的学术不端(相似性)分析服务, 为编辑部、研究生教学机构、科研机构提供科技论文质量把关服务。

● 独立的报告规范类的工具, 在图书馆信息服务中更是很少见到, 以色列艾利贝斯(Exlibris)公司提供的独立的报告软件系统, 专门把图书馆得到的各种数据收集、统计, 从而形成报告。但是这个系统是给图书馆工作人员用的, 不是给信息分析需求者用的。

逐步建立各种工具, 部署到信息服务系统中去, 这样, 才可以由查资源、阅读资源, 逐步转变成使用者研究与工作中需要的环境。到达这个地步, 才是信息组织和服务的最佳状态。

图19给出了不同的工具部署在

信息组织和管理流程的不同位置的示意。

(2) 科技创新辅助决策系统

回顾研究工作中的几个状态，首先是资料与文献的查找。基于主题，对于已经产生的文献资源进行序化与汇聚，可以更快地帮助研究者对于将要研究的主题方向的代表人物、代表机构、是否有项目等进行概括性的了解。对于核心人员、核心单位的系统化的分析，是把握研究方向、研究主体能力必不可少的分析研究工作之一。科技的查新

查证，原来是分散在多个数据库中所进行的。那么，现在多种资源情况下、在海量资源的情况下，可以利用工具帮助科研工作者在立项、选题、查信查证、论文命题、人物、机构、项目上进行统计分析，并且进行报告主要框架自动形成，成为科研工作者最主要的知识组织和服务工具。

图20是科技创新辅助决策支持系统(stads.wanfangdata.com.cn)界面示意图。在选择创新报告以后，该系统可以对机构、人物、主题、基金导航、学科导航等进行统计汇聚报告的自动产生。



图20 科技创新辅助决策支持系统示意——创新报告



图21 科技创新辅助决策支持系统示意——辅助查新功能

图21是选择了辅助查新以后的工作界面，可以进行快速检索，也可以进行专业检索，使得原来的多

库检索查新，变成用一个工具就可以完成查新和查新报告生成的工作。



图22 科技创新辅助决策支持系统示意——相似论文功能

图22是选择了相似论文功能以后，把自己要想写的论文题目、分类号、摘要输入后，这个系统就可以给出是否有人写过相似的论文的报告，保证写论文前就保证论文的质量。

(3) 基于用户行为轨迹分析

利用遗留在服务系统上的各种信息，分析用户的基本信息、关键词与选择资源的关系、使用的次数、时间、热点、主题、栏目利用率等，获得用户分类、阶段需求主题、分类—主题词关系。根据这个分析结果，设计的信息系统要再过滤加权、在服务推介上关联起来，逐步地学习用户行为后，信息服务的效率和质量就能不断地提高，给予满足时间需求的阶段组织、满足热点需求的背景相关组织、满足主题的内部外部的专业性的社区、博客、专家学者网站的汇聚、根据栏目利用状态调整网络服务系统的利用率等。同时，根据文献的特点，综合时间、是否是核心期刊、是否是核心人物、是否是热点等，进行加权分析与过滤关联的算法。

传统的信息服务排序，主要是按字母或者汉字偏旁部首排序。而用户行为中，多以时间、核心、被引多等要素为主要排序原则，所以，在检索系统中，可以改变过滤排序的策略就是一种用户行为控制的体现(见图23)。

图23中的红圈部分，就是除了传统排序以外的三种优先原则：加权相关优先、时间优先、按被引高低优先等原则，帮助用户发现的时候主要靠工具进行优先处理，再进



图23 用户可以选择的排序策略示意

行人工判断。

3 需要解决的相关问题

要做到能对于用户工作流程、流程中的资源配合、资源需求的牵引条件、用户需求分类、用户的行为习惯分类建立信息资源服务模式，最主要的是信息服务人员要提高自己的信息素养。信息素养，不是一个简单的技能型的素养。很多图书馆说检索是信息素养，这种提法是不恰当的。检索更多的是一种技能，并非是一种素养。而素养是需要方法、适合自己的方法来提高效率和质量的。素养不是简单的通过实践就可以提高的技能。信息服务人员要按照情报分析、参考咨询

服务、商业的咨询公司的咨询分析师的要求来要求自己。

最主要的有：

(1) 沟通素养

了解用户需求，归纳用户需求。除了需要专业知识以外，沟通素养是图书馆馆员需要具备的信息素养之一。

(2) 概括与抽象的素养

在《电子时代的图书馆馆员》一书中，特别强调的就是图书馆馆员要学习情报的汇总、分析的技能，而不是传统图书馆的检索一平铺—列出的方式。节省需求者的时间是服务的目的之一，而概括、抽象是一个信息素养。

(3) 使用各种工具的技能

检索只是工具的一种，而统计、分析、报告、整合在信息研究

中是不可缺少的技能，掌握这些工具成为一个必需的要求。大部分图书馆馆员是不会使用这些工具的，要提高掌握、学习、利用各种工具的水平是能否提供更好服务的技能基础。

(4) 图情理论与实际结合

如何把图情理论与实际环境结合，也是图书馆人要普遍具有的参考咨询馆员能力的要求之一。定性、定量、分析、决策理论要与具体服务相结合。

(5) 专业知识的掌握

除了图情理论以外，还要掌握一定的专业知识，这个知识是听懂专业人员问题和探讨问题的基础。

所以，要进行素养的提高，要适应基于用户的知识组织和服务组织是未来信息服务的主要的任务和挑战。

最后，根据以上出现的知识组织和服务的环境，有理由相信，有了完备的信息服务系统，以及有针对性地为教育和科研服务的资源组织、知识组织及其服务，信息服务将具备新的生命力。

参考文献

- [1] Extinction timeline 1950-2050 [EB/OL]. [2010-12-20]. http://rossdawsonblog.com/extinction_timeline.pdf.
- [2] Qing li Digital publishing in the new era of scholarly communication [EB/OL]. [2010-12-20]. http://www.stpi.org.tw/fdb/tr/2010/11/07_QingLi.ppt.
- [3] 孙卫. 用分类主题理论调整检索功能要求[J]. 数字图书馆论坛, 2008(4).

作者简介

孙卫，中国科学技术信息研究所高级顾问，科学出版社有限公司数字出版技术总监，计算机高级工程师，中国计算机学会高级会员，北京通信学会理事，中国图书馆学会会员。目前主要在知识组织、知识挖掘、知识处理技术上进行研究及教学。E-mail: sunwei0125@gmail.com

Knowledge Organization and Services for Users

Sun Wei / Institute of Scientific & Technical Information of China, Beijing, 100038

Abstract: For information network services based on the complexity and mass of resources, only knowledge organization and services based on the user's work flow, the user's information traction conditions, the classification of user needs, user behavior classification can effectively improve the efficiency and quality of information services.

Keywords: Knowledge organization, Knowledge services, Information network services, User behavior, Digital library

(收稿日期: 2010-12-28)