

从数字信息服务到智慧服务 ——以“淘智”为例*

温有奎^{1,2}

(1. 中国科学技术信息研究所, 北京 100038; 2. 北京万方数据股份有限公司, 北京 100038)

摘要: 在智慧地球的构建中, 如何面向知识创新建设智慧服务系统是科技信息服务的下一个发展目标。本文讨论了从数字信息服务到智慧服务的发展问题, 分析了互联网时代信息资源服务业所经过的数字信息服务、个性化知识服务、智慧服务的研究轨迹。结果表明, 随着物联网的到来, 今后信息资源服务业将借助于物联网技术向智慧服务发展, 并以“淘智”服务工具为例探讨了信息资源服务业的智慧服务模型。

关键词: 信息资源服务; 智慧服务; 知识创新

中图分类号: G202

DOI: 10.3772/j.issn.1673-2286.2015.10.001

1 引言

2008年11月IBM提出“智慧地球”(Smarter Planet)概念^[1]; 2009年1月, 美国奥巴马总统公开肯定了IBM“智慧地球”思路; 2009年8月, IBM又发布了《智慧地球赢在中国》计划书, 正式揭开IBM“智慧地球”中国战略的序幕。IBM“智慧地球”战略已经得到了各国的普遍认可。数字化、网络化和智能化, 被公认为是未来社会发展的大趋势, 而与“智慧地球”密切相关的物联网、云计算等, 更成为科技发达国家制定本国发展战略的重点。2008年欧洲信息社会与媒体委员会发布了《2020年的物联网: 未来之路》^[2], 被认为是目前“物联网”(Internet of Things)的一个较为权威的定义。把“物联网”定义为由一些“具有身份标识与虚拟的个性化特征, 可以利用智能化接口在智能空间进行相关操作, 并可以与社会的、环境的、用户的上下文相互连接并进行有效沟通”的物体构成的网络, 认为物联网具有“整合性”, 是未来的因特网, 其中的物体互为连通且扮演积极角色, 具有主体性。

智慧地球概念的提出揭开了自然系统和人工系统融合感知的物联网时代序幕, 为物联网的利用找到了

智慧应用的先例, 进而出现了更为具体的智慧城市^[3]、智慧校园^[4]、智慧图书馆^[5]、智慧服务^[6]等利用物联网实现智慧服务的平台。吴良镛院士起草的《北京宪章》提出21世纪建筑事业发展应该着眼于人居环境建设, 必须统筹兼顾、妥善处理以下五个方面的内容: 生态(正视生态困境, 加强生态意识)、经济(使人居环境建设和经济发展良性互动)、科技(充分利用科学技术, 推进经济发展和社会繁荣)、社会(关怀广大人民群众, 重视社会发展的整体利益)、文化(推动建设文化和艺术的创造、发展和繁荣)^[7]。如何充分利用科技信息资源为知识创新服务是智慧地球建设的一个重要部分。本文分析了互联网时代, 信息资源服务业所经过的数字信息服务、个性化知识服务和物联网服务轨迹, 由此提出随着物联网的到来, 今后信息资源服务业将借助于物联网技术的发展向智慧服务发展, 并以“淘智”服务工具为例探讨了信息资源服务业的智慧服务模型。

2 “淘智”的智慧服务概念模型

“淘智”^[8]是北京万方软件公司开发的一个具有知识发现功能的智慧服务工具。“淘智”根据科技文献摘

* 本研究得到国家科技支撑计划课题“面向科技情报分析的信息服务平台研发与应用示范”(编号: 2015BAH25F02)资助。

要具有概括介绍文章创新点和研究价值的特点,采用统计分析与智能挖掘技术、语义关联技术、本体建模技术、智能推理技术等开发出了一种新知识发现服务理论模型^[9]。“淘智”智慧服务理念的设计有两个主要的新功能(如图1所示):一是给出了某一个研究领域的国内科研人员的主要热门研究轨迹;二是针对用户的提问给出了多个相关的解决方案。“淘智”工具能挖掘出摘要中的“问题-方法-结果”的核心创新点,实现了基于科研兴趣的科学研究轨迹的揭示和知识发现,以及基于用户问题的解决方案的感知分析与推荐的智慧服务功能。这两个功能改变了传统的信息检索方式,使得在大数据环境下能够快速发现解决问题的方案,为实现大数据知识挖掘提供了一种新方法。“淘智”开启了从数字信息服务向智慧服务发展的新探索。

从数字信息服务向智慧服务发展是智慧地球的一个组成部分,它是物联网发展的成果。物联网强调“any things connection”,而“things”不但包括现实世界的物件(object),也包括各种计算设备(computer)与虚拟空间的人工物件(artifact),还包括网络用户(human)。因此,将“淘智”这两个功能与物联网绑定,会对科学研究的知识创新服务带来帮助。基于上述物联网的定义,本文提出一个智慧服务概念模型如图1所示。模型将利用物联网技术把“实体文献资源”与“用户”连接起来,将虚拟的“研究问题”与知识库中“科研解决方案”感知连接,将关联推理的“科学研究轨迹”与用户的“科研兴趣”感知连接,以此建立概念模型平台将用户与知识创新融合在一起。

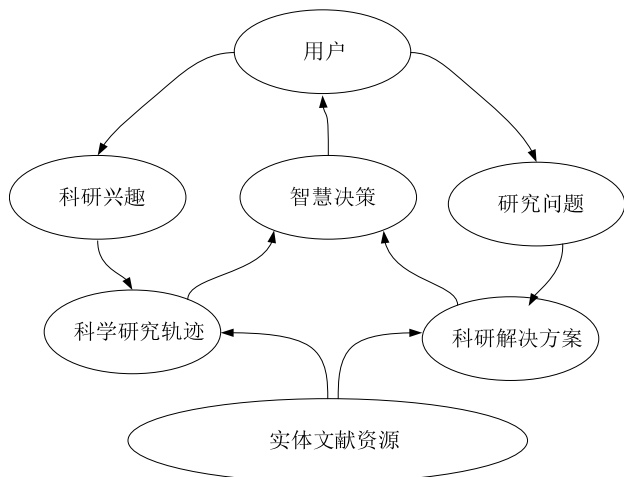


图1“淘智”智慧服务概念模型

图1描述了数字信息服务与智慧服务的主要区别:数字信息服务主要是对实体文献资源的服务,即图1的底层实体文献资源服务方式,包括个性化知识服务。而智慧服务将要对实体文献资源做内容的深度语义挖掘,建立科学研究轨迹库,建立科研解决方案库,对用户实施物联网感知服务。基于智慧服务概念模型建立智慧服务平台,将用户的需求分析为科研兴趣和科研问题,并通过感知、挖掘、关联、聚类、推荐等技术处理提供智慧决策服务。

3 “淘智”的问题解决方案

3.1 挖掘文摘中的创新知识

创新点分布在论文的整体结构中,表现为主题中的创新点、技术背景中的创新点、技术方法中的创新点、论文结论中的创新点和总体创新点。根据写作要求,每种创新点表现出了独特的知识本体结构。建立创新点的知识本体模型,是实现创新点智能识别和动态挖掘的关键理论。文献[9]给出了科技论文创新点的本体模型图,认为一般科技论文对创新点的描述由五大部分组成,既展现出一种层次关系,又表现出一种网状关联关系;并总结了科技论文表达创新点的写作特征,即题目凸显创新点、摘要陈述创新点、绪论论述问题与假设、方法论论证创新点、结果和讨论阐明创新点。

基于上述摘要具有陈述创新点的特征假设,我们以期刊论文的摘要为数据进行统计分析,发现表达论文创新点的要素特征,提出一个基于“问题-方法-结果”三元组结构的创新点知识单元,如图2所示。

图2是用“物联网”提问挖掘出的“问题-方法-结果”三元组结构的创新点知识单元实例。由图2可以看出,“淘智”将科技文献摘要的创新点解析成“问题-方法-结果”的知识单元,为用户提供了比文摘更简练的关注点阅读方式。

3.2 基于问题的解决方案的推荐服务

在挖掘出“问题-方法-结果”的知识单元的基础上,“淘智”将问题分析-聚类,将方法分析-聚类,将结果分析-聚类,建立基于问题的解决方案推荐单元,并由结果聚类作为判定条件,辅助用户自己判定感兴趣的解决方案(如图3所示)。

原文	学科	作者	问题	方法	结果	引文量
[1]	工业技术	杨楠	分析了物联网环境下智能产品原型的特征和意义	结论提出了通过开源平台和物联网应用平台快速、低成本地构建智能产品原型的方法	能够帮助设计师更加有效地对智能产品的设计方案进行表达、探索、测试和评估	0次
[2]	工业技术	周云科	针对传统矿井安全监控系统技术单一等缺点	设计了一种基于物联网技术的矿井安全监控系统	实验结果表明：系统稳定性好、可靠性强	0次
[3]	工业技术	滕志军	针对铁路牵引变电站传统有线监测系统线缆密布、安装维护困难等问题	设计了一种基于物联网技术的高铁电能智能监测系统	实验结果表明-该监测系统实时性强	
[4]	工业技术	马小平	研究了3种技术在煤矿安全生产保障中的作用和地位	提出了3种技术在煤矿生产安全保障中的关系：物联网是煤矿各个子系统建设的框架和路线图	可有效提升煤矿安全生产水平	0次
[5]	工业技术	付贵祥	阐述了各预警模型即瓦斯预警模型、火灾预警模型、水灾预警模型及顶板预警模型的设计	给出了智能决策的业务流程	有效提高煤炭企业的安全生产水平	0次

图2 “淘智”挖出文摘中的创新知识

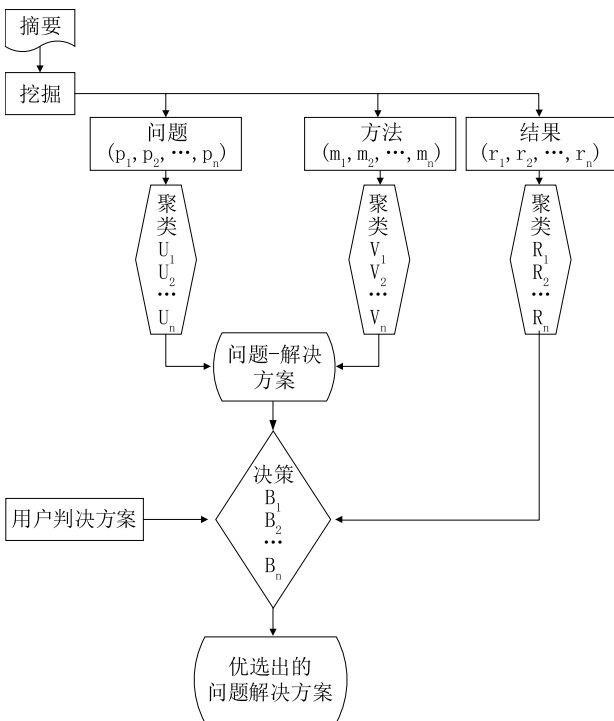


图3 问题解决方案的智慧决策模型

3.3 模糊决策实现

定义1. 模糊集

设U是论域， μ_A 是把任意 $u \in U$ 映射为[0,1]上某个值的函数，即 $\mu_A: U \rightarrow [0,1]$ 或者 $u \rightarrow \mu_A(u)$ ，则称 μ_A 为定义在U上的一个隶属函数，由 $\mu_A(u) (u \in U)$ 所构成的集合A称为U上的一个模糊集， $\mu_A(u)$ 称为 μ 对A的隶属度。

定义2. 模糊关系

A_i 是 $U_i (i=1,2,\dots,n)$ 上的模糊集，则称 A_i 为 A_1, A_2, \dots, A_n 的笛卡儿乘积，它是 $U_1 \times U_2 \times \dots \times U_n$ 上的一个模糊集。 $U_1 \times U_2 \times \dots \times U_n$ 上一个n元模糊关系R是指以 $U_1 \times U_2 \times \dots \times U_n$ 为论域的一个模糊集，记为：

$$R = \int_{U_1 \times U_2 \times \dots \times U_n} \mu_R(\mu_1, \mu_2, \dots, \mu_n) / (\mu_1, \mu_2, \dots, \mu_n)$$

一般地说，当U和V都是有限论域时，其模糊关系R可用一个模糊矩阵表示：

$$U = (U_1, U_2, \dots, U_m)$$

$$V = (V_1, V_2, \dots, V_n)$$

则 $U \times V$ 上的模糊关系为

$$R = \begin{bmatrix} \mu_R(u_1, v_1) & \mu_R(u_1, v_2) & \dots & \mu_R(u_1, v_n) \\ \mu_R(u_2, v_1) & \mu_R(u_2, v_2) & \dots & \mu_R(u_2, v_n) \\ \dots & \vdots & \dots & \dots \\ \mu_R(u_m, v_1) & \mu_R(u_m, v_2) & \dots & \mu_R(u_m, v_n) \end{bmatrix}$$

定义3. 模糊关系的合成

设 R_1 与 R_2 分别是 $U \times V$ 与 $V \times W$ 上的两个模糊关系，则 R_1 与 R_2 的合成是指从U到W的一个模糊关系，记为： $R_1 \circ R_2$

其隶属函数为：

$$\mu_{R_1 \circ R_2}(u, w) = \vee \{ \mu_{R_1}(u, v) \wedge \mu_{R_2}(v, w) \}$$

定义4. 模糊变换

设 $A = \{ \mu_A(u_1), \mu_A(u_2), \dots, \mu_A(u_n) \}$ 是论域U上的模糊集，R是 $U \times V$ 上的模糊关系，则： $A \circ R = B$ ，称为模糊变换。

由此可以利用模糊推理的方法实现解决问题的方法的推荐：

第一步：我们根据3.1节得到的“问题-方法-结果”三个知识元，通过聚类方法将其聚类为问题类、方法类、结果类抽象概念；

第二步：分别将问题类、方法类、结果类通过特征函数的处理，建立三个模糊集U,V,W；

第三步: 建立“问题-方法”构成的模糊关系矩阵R;

第四步: 建立以模糊集W与模糊关系矩阵R的模糊变换, 实现解决问题方法的优化推荐。

3.4 举例

设评价目标集合U={问题, 方法, 结果}; 评语集合V={很好, 较好, 一般, 不好}。

首先对问题、方法、结果分别进行评判, 这样得到模糊评判矩阵R。

此外, 假设用户对挖掘决策的重要性评判主要是方法和结果, 即假设用户对挖掘结论的综合评判权重为: A={0.2,0.5,0.3},

$$R = \begin{bmatrix} 0.2 & 0.7 & 0.1 & 0 \\ 0 & 0.4 & 0.5 & 0.1 \\ 0.2 & 0.3 & 0.4 & 0.1 \end{bmatrix},$$

则得出模糊决策结果为: B=[0.2 0.4 0.5 0.1]。

由实例可以看出得出的决策结果为满足第二和第三个因素的权值较大的方案, 说明研究假设中比较强调第二(方法)和第三个(结果)因素的重要性, 研究结果和假设相符。

4 数字信息服务与智慧服务的研究轨迹

4.1 基于“淘智”的数字信息服务研究轨迹的发现

(1) “互联网”主题下的主要研究对象

利用“淘智”对“互联网”进行检索挖掘, 研究轨迹表明: 数字信息、出版物、国际标准、条形码、object等成为与互联网关联度较高、数量较大的研究对象。

(2) 互联网推进了数字出版发展

由(1)的调查结果发现, 互联网推进了数字信息的发展。数字化技术的发展, 推动了出版物发展, 进而推进了文献信息资源的发展, 为信息资源服务业提供了信息资源远程服务的商机。我们对“数字信息”的研究轨迹做了分析, 发现其研究集中在: 出版物、国际标准、互联网、条形码、object、中华医学会。数字信息主要是出版物的数字信息, 而且中华医学在发表论文和系列杂志方面的研究活动更为突出。

(3) 互联网推进了数字信息服务发展

对“信息资源”进行检索挖掘, 研究轨迹表明: ①“图书馆”是与“信息资源”关联度最大的一个词。图书馆需要建立数字图书馆, 需要信息资源建设, 需要文献信息资源, 需要网络信息资源, 需要网络环境, 需要高校图书馆的信息服务。②“网络信息资源”与“信息资源”关联度较大。网络信息资源的建设需要研究信息组织、信息资源开发、元数据。③“信息资源管理”与“信息资源”关联度较大。它与知识管理、电子商务、电子政务构成了一个研究体系, 其中电子政务的研究成为信息资源管理的主要对象。

(4) 互联网推进了个性化服务发展

对“个性化服务”进行检索挖掘, 研究轨迹表明: 数字图书馆是开展个性化服务的最主要的研究对象, 其次是图书馆、用户需求、数据挖掘、推荐系统、用户模型等对象(如图4所示)。

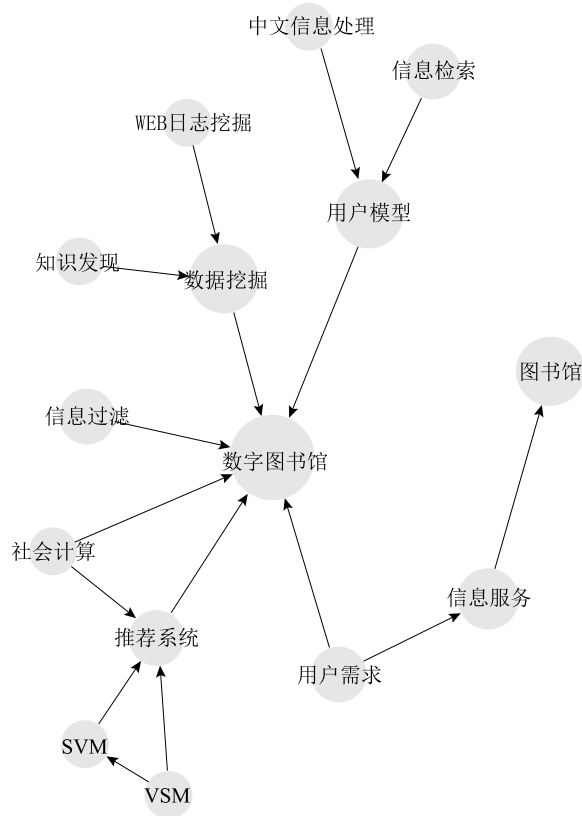


图4 “个性化服务”主题下的主要研究对象

由图4可以看出, “数字图书馆”是与“个性化服务”关联度最高的一个词。“数字图书馆”开展了“用户需求、用户模型、信息过滤、社会计算、推荐系统、数据挖掘”等研究, 向个性化知识服务发展。

(5) “物联网”主题下的智慧图书馆发展

我们用“物联网”这个主题进行文献研究轨迹调查，“云计算”、“无线传感器网络”、“RFID”这三个词是与“物联网”最相关的顶层研究对象。进一步分析发现，“智慧图书馆”这个词通过“大数据”与“云计算”密切相关。即对“物联网”对象的研究包括下面对象和对象之间的关联关系：“智慧图书馆”→“大数据”→“云计算”（如图5所示）。

通过“互联网”和“物联网”这两个研究对象轨迹的调查，我们得到了两个重要的不同研究对象：一是“互联网”时代人们主要研究“数字信息”；“物联网”

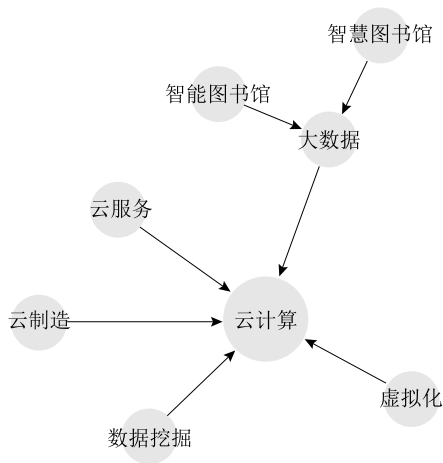


图5 “物联网”主题下对象之间的关系

时代人们关注“智慧图书馆”。

4.2 基于“淘智”的智慧服务研究轨迹的发现

(1) “智慧”的研究轨迹

虽然人们对“智慧”无法给出理想的定义，但李德仁院士在对智慧城市下的定义“智慧城市=物联网+互联网”中表现出了李院士对智慧城市构想的“智慧”，它是由“物联网+互联网”构成，突破了当前的数字城市的概念。文献[6]认为“智慧服务”是一个整体的概念，强调通过动态感知，来实现对各个构成要素的动态管理，以人为本，使用户可以利用技术手段实现智能、互联、协同、共享，实现智能决策、知识生产和自动化控制等。显然人们对“智慧服务”的研究边界仍过于笼统。

图6给出的是国内研究人员对“智慧”这个词的主要研究轨迹，它对我们研究的对象——“智慧”的理解带来了启发。

由图6可以看出：目前我国科研人员对“智慧”的研究有以下几个特点。

①目前对“智慧”的研究主要以“物联网”为核心。如智慧城市、云计算、智慧图书馆、数字城市、智慧地球、RFID、智能感知、三网融合、智能手机、智能图书馆、大数据、智能服务等都以“物联网”为核心。

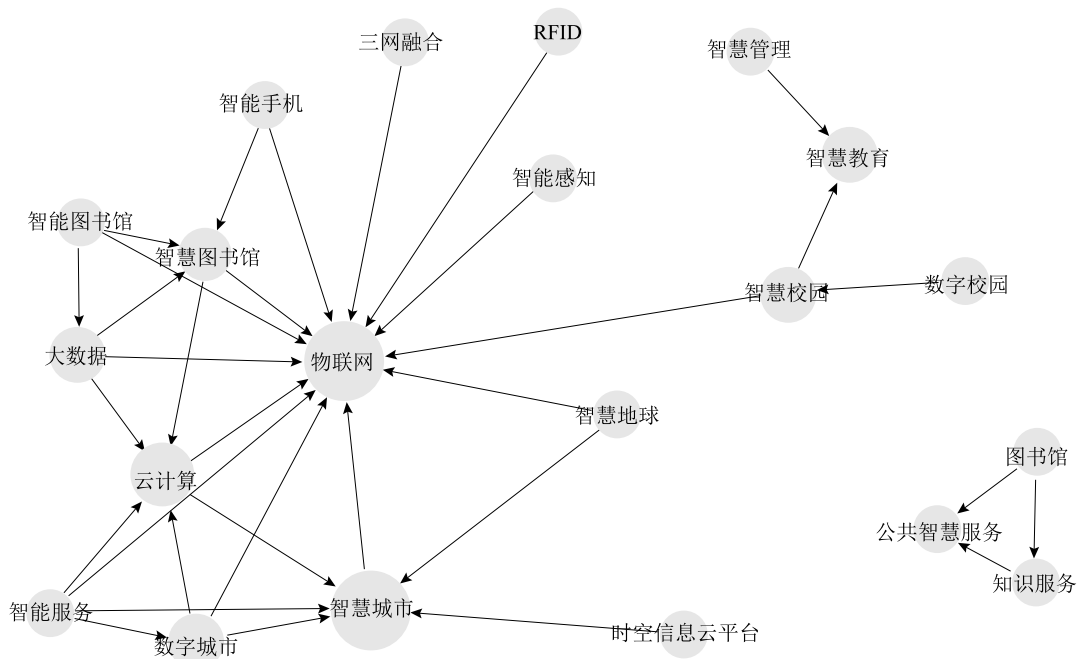


图6 “智慧”的研究轨迹

②目前对“智慧”研究的重要对象是“物联网”、“智慧城市”，其次是“云计算”。

③“智慧XXX”都与“数字XXX”关联，如“智慧校园”与“数字校园”关联，“智慧城市”与“数字城市”关联，“智慧矿山”与“数字矿山”关联。

④“智慧XXX”与“物联网”紧密关联。如“智慧校园”与“物联网”关联，“智慧图书馆”与“物联网”关联。

⑤“智慧XXX”都与“云计算”关联。如“智慧城市”与“云计算”关联，“智慧图书馆”与“云计算”关联。

⑥“物联网”的研究离不开对“云计算”、“大数据”、“RFID”的研究。

⑦当今人们对“智慧”研究的是以“数字”为基础的，对“物联网”的研究又是以“智慧”为基础的。即“数字”→“智慧”→“物联网”。

⑧“公共智慧服务”的研究基于“图书馆”、“知识服务”。

⑨“智能服务”的研究基于“数字城市”、“智慧城市”、“云计算”、“物联网”的研究。

(2) “智慧图书馆”的研究内容

由图6可以看出，以“物联网”、“云计算”为物理基础的“智慧图书馆”的研究正在出现，“智慧图书馆”正借助于“智能手机”、“智能图书馆”、“大数据”成为研究对象。此外也可以看到“公共智慧服务”借助于“图书馆”、“知识服务”为对象，开始了智慧服务的研究与发展。可以预测“智慧图书馆”的发展将会开启智慧服务，数字信息服务将会走向智慧服务。

5 结语

物联网的发展推动了“智慧地球”的发展，必将带动数字信息服务向智慧服务转变。“淘智”这一新型知识服务工具开创了基于问题解决方案的智慧服务模式。本文探索了“淘智”的概念模型和智慧决策模型，并应用“淘智”分析了互联网时代信息资源服务和基于物联网的智慧服务的研究轨迹。事实表明基于物联网的智慧服务将成为数字信息服务的下一个发展目标。

参考文献

- [1] 智慧地球[EB/OL].[2015-10-08].<http://baike.baidu.com/subview/2168958/12266828.htm>.
- [2] Internet of Things in 2020: Roadmap for the Future[R/OL].[2015-10-08].
<http://www.docin.com/p-57876410.html>.
- [3] 李德仁,邵振峰,杨小敏.从数字城市到智慧城市的理论与实践[J].地理空间信息,2011(6): 1-5.
- [4] 黄荣怀,张进宝,胡永斌,杨俊锋.智慧校园:数字校园发展的必然趋势[J].开放教育研究, 2012,18(4):12-16.
- [5] 乌恩.智慧图书馆及其服务模式的构建[J].智慧图书馆及其服务模式的构建,2012(5): 102-104.
- [6] 冯璐,赵佳因,郭乐深.智慧城市智慧服务体系构建[J].北京城市学院学报,2014(1):65-70.
- [7] 吴良镛.21世纪建筑学的展望[J].建筑学报,1998(12):4-12.
- [8] 淘智[EB/OL].[2015-10-08].<http://168.160.200.188:9201/login.aspx>.
- [9] 温有奎,吴广印.碎片化科研创新点动态挖掘研究[J].数字图书馆论坛,2014(7):25-31.

作者简介

温有奎,男,1951年生,管理学博士,教授,研究方向:智能搜索引擎、文本知识挖掘, E-mail: wykui123@126.com。

From Digital Information Resources Services to Smarter Services: A Case of "Taozhi"

- (1. Institute of Scientific and Technical Information of China, Beijing 100038, China;
2. WANFANG Data Co., Ltd, Beijing 100038, China)

Abstract: In order to explore the construction of a smarter planet, how to build the knowledge innovation system of intelligence services is the next target the development of science and technology information resources and services. This paper discusses development issues from the digital information resources and services to the intelligence services, analyzes the trajectory of digital resources services through which information resources sector, track of personalized knowledge service, and researchers track of smarter service. The results show that with the advent of things, Information Resources service will aid the development Internet of Things technology to the smarter services developments. And "Amoy wisdom" service tool as an example to explore the smarter service model of information resource services.

Keywords: Information Resources Services; Smarter Services; Knowledge Innovation

(收稿日期: 2015-10-09; 编辑: 雷雪)