

面向政府决策的知识挖掘^{*}

Government Decision Oriented Knowledge Mining

吴 鹏

苏新宁

(南京理工大学信息管理系 南京 210094) (南京大学信息管理系 南京 210093)

摘 要 在政府决策过程中,需要发掘蕴涵于政府智力资产中的隐性知识,形成决策智慧。将知识挖掘概念引入政府决策中,指出政府决策中需要挖掘的知识资产的类型,并分析知识挖掘对政府决策的作用,最后给出了政府决策过程中知识挖掘的流程。

关键词 知识挖掘 政府决策 隐性知识

知识挖掘是信息处理新技术,又是一门受到多学科领域研究者共同关注的边缘学科,因此“知识挖掘”还有了“数因发现”、“数据开采”、“知识抽取”、“信息发现”、“知识发现”、“智能数据分析”、“探索式数据分析”、“信息收获”和“数据考古”等提法和不同的定义。

知识挖掘可以支持完成从隐性知识到显性知识的转换,这个过程可以从已有的组织知识资产中选择需要的决策知识,形成决策智慧,进行正确的决策。知识挖掘可以基于特定领域,收集、分析和可视化组织知识,提高分析结果的质量和值。在政府决策过程中存在两个层次的隐性知识:一类是存在于组织个人和环境中的可编码性隐性知识,另一类是存在于组织显性知识中而没被发现的相对隐性知识。在政府决策过程中采用知识挖掘,可以发现这两类隐性知识。组织的知识管理团队可以从组织的领域专家、信息系统或文献等非数据性资源那里获取可编码性的隐性知识;运用电子政务数据挖掘等手段,沿着“数据—信息—知识”的价值链,从组织的数据资源中转换和获取有价值的新知识,从而解决相对隐性知识的挖掘问题。

政府知识挖掘是按照既定的目标对大量的数据进行探索,揭示隐含其中的规律并进一步将之模型化的先进、有效的方法。政府知识挖掘的目的是将大量非结构化的多媒体信息融合成有序的、分层次的、易于理解的信息,并进一步转换成可用于干预预测和决策的知识。简言之,这其实是一个智能化、自动化的过程。

1 知识挖掘在政府决策中的作用

电子政务的实施,将迫使我国政府管理职能在加入 WTO 后加速转变。同样,在电子政务环境下,政府也必将在行政管理职能方面,面临创新能力、突发事件快速响应能力等诸多方面的挑战。而知识管理技术的应用将有效地应对这些挑战,并在重塑政府形象、提升政府工作效率等方面起着重要的作用。

1.1 快速响应能力 快速响应能力是政府在应对社会突发事件、自然灾害等各类发生突然、危害极大的重大事件时的预防、响应、处理等的综合能力。它可以通过对历史事件记载文献挖掘出的知识来制定应对措施,增强快速反映能力。突发事件以其不可预知、危害极大的特点,使人们对这类事件的性质、原因、发生时间、地

点、危害程度等均无法预先确定。因此我们可以利用知识挖掘工具,对历史记载下来的突发事件进行挖掘,从中提取、总结、升华前辈们的经验和教训,得到我们今后应对突发事件的各种知识,制定出高效的应对措施,提高政府在处理此类事件时的快速响应能力。

1.2 创新能力 创新是政府职能向知识型政府转变的基本要求,它不仅仅局限于行政方法和政务处理流程层面上的创新,而且也包括政府制定发展战略和公共政策的创新。如,政府职能将更多地面向社会提供公共服务,实施公共管理,为适应这种转变,政府需要对历史的经验和教训、外国政府的职能、社会公众的需求等进行有效地挖掘、吸收、借鉴和利用,依靠挖掘出的公共管理战略知识来指导政府的战略与政策的制定。

1.3 促进公务员素质的提高 我国政府公务员制度虽然已实施多年,但计划经济体制下的管理模式在今天的公务员制度上仍有体现,过去的政务管理过程属程序操作型管理,对政府公务员办事的基本要求就是认真、负责、严格按照规定办事等,公务员只需要按照既定的工作程序开展工作即可。外界对此的印象是呆板、缺乏灵活性、效率较低,因此公务员自身素质也很难得到提高。随着电子政务中的自助式服务、一站式服务的逐步实施,政府公务员事必躬亲的现象将大大减少。但随之而来的是,社会公众有可能提出许多以前不曾提出过的要求或服务,如纳税人提出缓交、减免等要求,甚至针对税收政策提出疑问等。作为政府公务员必须对此予以及时、明确的解释和答复。此类问题公务员可能有的知道,有的并不清楚。对于不清楚的问题,公务员可以利用数据挖掘工具寻找答案。因此,政府各类公务员都必须提高自身挖掘知识、学习知识的能力和办事技能,并能通过网络寻求专家库和知识库的有力帮助。

1.4 高效率的行政管理 对现代政府来说,效率不仅仅意味着规范、快速和程序化,而且从更深层次上理解,它要求的是更智能化和更高品质的工作。如利用数据挖掘,尽可能地获得社会公众在网络虚拟环境中提出的各种服务请求的解决方案或应对措施,建立主题知识库,使社会公众提出的服务都能有较满意的结果。

2 政府知识挖掘的流程

政府知识挖掘利用知识管理、知识发现、元数据分析和数据可

基金项目:南京理工大学经济管理学院青年教师科研扶持基金项目资助。

作者简介:吴 鹏,男,1976 年生,讲师,博士,研究方向为电子政务、知识管理、信息系统;苏新宁,男,1955 年生,教授,博士生导师,研究方向为情报检索理论与技术、知识管理、电子政务。

可视化等方法,可以增强分析结果的质量和值,从而支持决策的正确性。政府决策过程知识挖掘包括知识整理阶段和挖掘阶段。知识整理阶段包括知识分类和知识编码。而挖掘阶段包括面向政府可编码隐性知识和显性知识的挖掘。

知识整理阶段的第一步就是开发一个组织知识分类体系。识别和发现组织的知识资产,离不开对于知识类型的了解。通过一个合适的知识分类系统,将组织内的所有知识按照一定逻辑习惯编排起来,从而便于人们对知识仓库中知识的浏览、查找和使用。例如:著名的网络搜索引擎雅虎首页上的网站分类目录,就提供了一个实用的知识分类系统。经过分类的知识需要进行知识编码,才能进行知识挖掘,知识编码是知识挖掘的重要步骤,其目的是将各种形式存在于不同位置的知识目标抽取出来,转化为计算机可以处理的形式,并集中存放在一个知识缓存区,为下一步的知识处理做准备。

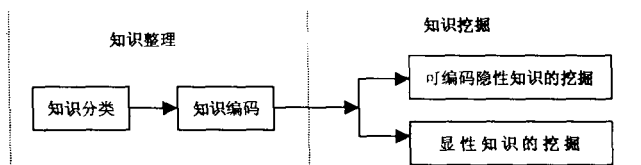


图1 政府知识挖掘流程

2.1 政府知识分类体系 知识分类体系将组织内的所有知识按照一定逻辑习惯编排起来,从而便于人们对组织知识的浏览、查找和使用。为政府组织开发一个适用的知识分类系统有两种主要的途径:第一种途径,是自顶向下的方法,是指从某个标准分类法模板开始,然后根据组织的具体情况加以修改后应用;第二种途径,是自下向上的方法,指的是通过某种技术手段,对数据、信息进行挖掘,从而构造出一个适用的分类法。

无论采用哪种途径来构造政府组织的知识分类系统,都要遵循两个适用性原则。首先,要遵循简单实用的原则。例如,分类的层次不能太深,同一层次的类目不能过多等。简洁明了的分类方案比模糊复杂的方案更容易被用户接受。

其次,分类系统的设计要能够适应组织的当前结构和未来发展。例如,考虑到政府组织的规模,一个庞大而包罗万象的分类法可能并不适用,而为单独的部门或者实践社团创造一个较小的、更个性化的分类法可能是更好的解决方案。

根据各种知识对政府的战略重要性程度、发展的潜力和发展的不同阶段,我们可以将政府知识分成四种类型,即:a. 发展中知识(Promising Knowledge),指那些仍处于萌芽阶段,但却会引发政府重大变革的知识;b. 核心知识(Core Knowledge),是一个政府之所以不同于其他组织的标志,它们对政府的独特地位有重大影响,也称为核心竞争力(Core Competences);c. 基本知识(Basic Knowledge),是指完成政府各种活动所必须和最基本的知识,这类知识在所有相似的政府机构中都是可以获得的;d. 过期知识(Outdated Knowledge),是几乎不再被应用于政府业务处理过程的知识。

以知识的可应用范围和可传递性为标准,我们可将政府知识划分为四个类型:a. 快速存取型知识(Quick Knowledge),指那些容易传递(甚至可以程序化)但却无法广泛应用的知识。b. 宽泛型知识(Broad-based Knowledge),指那些传递起来非常容易,又可以进行广泛应用的知识。c. 个性化知识(One-off Knowledge),是既不易传递又不可能广泛应用的一类知识。d. 复杂知识(Complex Knowledge),可以广泛应用但却不易传递的一类知识。

以上这些分类法有助于我们了解政府组织知识资产的构成、分布和结构,有助于我们确定政府决策过程知识挖掘工作的重点,有

利于我们根据不同类型知识的特点,采取不同的识别和编码策略。至于在具体的操作中应如何分类,则取决于组织的特点以及它所需知识的特性。

2.2 政府组织内的知识编码 由于人类知识的复杂性,知识编码是一个难度很大的过程,尤其是隐性知识的编码。知识目标存在形式的多样化,决定了知识编码需要采用不同的编码技术。根据中国学者周军的研究成果,笔者按照知识的存在形式,表1总结了政府组织的各种存在形式的知识来源及其可以采用的编码技术。这里所说的知识来源,泛指可以从中获取知识的数据、信息和知识资源。

表1 知识编码技术

类别	存在形式	说明	编码方法
隐性	内含于员工头脑内的观念型知识	包括知觉能力和观念化能力,如:领域专家的领域专门知识等	知识获取技术
	内含于员工的经验型知识	通过“边做边学”学习的 Know-how 型知识,如:操作技巧等	数字媒体技术 运动分析技术
	内含于组织文化的知识	未被记录或明确说明的组织的文化、共识、价值观等	知识获取技术
	内含于组织运作的知识	未被记录或明确说明的组织业务流程	面向对象分析技术
显性	支持组织运作的计算机程序	支持组织运作的各类信息系统中的业务逻辑、业务流程、数学模型等	应用挖掘技术
	知识库	专家系统中的知识库,如:某种疾病诊断专家系统中的诊断知识	知识表示的转换
	数据库或数据仓库	支持组织运作的各类信息系统中的数据库,或用于组织联机分析的数据仓库,如:客户信息数据库、员工数据库等	数据挖掘技术
	电子文献	采用电磁、光学等新型媒介存储的文字、图像、声音等文献。如:文档库、Web 文档、E-mail、视听材料、竞争情报报告等	数字媒体技术 文本挖掘技术 Web 挖掘技术 语音识别技术
	纸质文献	采用传统纸质媒介存储的文字、图像等文献。如:档案资料、技术图纸、操作手册、文件等	光学扫描技术 OCR 技术

2.3 政府可编码隐性知识的挖掘 政府组织的隐性知识是知识挖掘中最有价值而相对困难的部分,还没有形成系统的理论,目前研究的重点在于隐性知识的捕获模型和挖掘技术上。

2.3.1 政府可编码隐性知识的捕获模型。政府知识捕获模型包括知识历程图、知识缺陷分析器和知识地图三部分。确定了政府组织知识分类体系后,就可以开发政府知识历程图(Knowledge Storyboard)模型。政府知识历程图的目的是发现和识别政府决策过程的知识需求,它的形式是一份组织所需求的知识实体(Knowledge Entity)的清单,以及知识实体的使用说明。所谓知识实体,是指组织某项活动所需要的一种具体的知识内容,如“行政知识”。政府知识缺陷(Knowledge Gap)分析就是使用知识缺陷分析器模型对每一份知识实体的需求和供给情况给予评估,从而发现政府机构的知识缺陷。在此基础上,可以根据各知识实体重要性、捕获难度等因素制定一个可行的知识捕获策略。最后,是开发描述组织知识供给的政府知识地图(Knowledge Map)模型。政府知识地图将每一份知识实体转化为相应的知识目标(Knowledge Object),并为每个知识目标标示出所存在的场所。所谓知识目标,是指知识实体的具体知识源。通过对政府知识历程图和政府知识地图的扫描和分析,将组织的知识供给和知识需求联系起来,就可以得到政府组织知识市场(Knowledge Market)的全景视图,这样的视图可以使人们对组织知识的供给和需求有更全面的认识。

2.3.2 政府隐性知识客观化技术。政府决策隐性知识客观化

过程离不开技术的支持。这些技术可以从其他学科借鉴而来,如软件开发中的面向对象分析方法、数字媒体技术、基于概念格方法的检索词分析技术、专家系统中的知识获取技术、人工智能中的运动分析技术等。

2.4 政府显性知识的挖掘 政府显性知识的挖掘,是指沿着“数据—信息—知识”的价值链,运用数据挖掘等手段,从组织的数据资源中转换和获取有价值的新知识。数据挖掘(Data Mining)是指从大量的数据(结构化或非结构化)中提取有用的信息和知识的过程。需要说明的是,数据挖掘的数据源是大量的、真实的、可能含有噪音的;所发现的信息和知识是潜在的并隐藏在大量数据背后的,是用户感兴趣的、可理解、可运用的知识。数据挖掘的对象可以是政府的数据库、数据仓库,也可以是非数据库组织的文本数据源、Web 数据源以及复杂的多媒体数据源等。政府通过实施电子政务工程,整合了政府各级各类数据信息资源,这些数据在满足正常政务工作需要方面起到了重要作用,但这些数据中隐含的各种政务知识则由于其原始数据量极大、结构复杂等因素的困扰,而未能有效地进行挖掘,因此急需在电子政务系统中开展政务知识的数据挖掘工作。

电子政务将带来政府职能的转变,作为服务型政府,除日常行政办公外,将更多地面对宏观政策、法规的研究制定,政府决策支持,突发性事件的快速反映及应对措施。为了做好这些工作,政府部门不仅要积累大量的知识,而且要具有在大量的政务数据和信息中挖掘出知识的能力。近年来,在我国电子政务系统的长期建设过程中,政府机构利用信息技术生产和搜集数据的能力大幅度的提高,大量数据库被各级政府、部门开发建设出来,数据的增长呈指数上升、数据类型日益复杂,而且今后这一势头将会随着社会环境、计算机软硬件技术的发展更加迅猛、持续地发展下去。但是,现实却是:一方面政府制造和生产了海量的各种类型的数据,另一方面,仍

然采用传统的手段来处理海量的数据和复杂的数据已变得不切合实际,大量的、历史的政务数据已成为政府的包袱,但其中却又不可避免的隐藏着大量“金矿”(知识)。因此,电子政务期待着新的数据分析处理技术来解决这个问题,这就是近年来引起业界广泛关注的数据挖掘(Data Mining)技术。

3 结 语

网络信息社会环境下的各级政府和部门,都需要从已有的大量历史数据中挖掘隐藏在其中的战略知识,并整理优化这些知识,使之成为政府行政管理的“知识智慧”。知识挖掘作为挖掘、管理和共享政府所拥有的信息、知识、资源的方法,其目标就是快速而方便地找到所需要的信息,使行政管理的主体——行政领导和各级公务员可以快速而方便地访问所需要的信息和知识。使最恰当的知识,在最恰当的时间,传递给最合适的行政人员,以实现最佳的行政决策和高效的行政管理。

参 考 文 献

- 1 Mohamed Khalifa. Integrating Knowledge Management Into Enterprise Environments for the Next Generation Decision Support. Narasimha Bolloju, 2003
- 2 Nikitas Spiros Koutsoukis. Decision Modelling and Information Systems: The Information Value Chain, 2003. 10
- 3 Paul A Golder. Decision Making. Information and Systems, 2003; (3)
- 4 Peter Drucker. Where Tomorrow's Decisions are Being Shaped Today. New York: Truman Talley Books, 1986
- 5 张福学. 知识管理领域的知识分类. <http://www.chinakm.com>, 2002. 9
- 6 周 军. 知识管理中的知识仓库. 南京大学博士学位论文, 2003. 6

(责编:京阳)

(上接第 44 页)第一个层次是业务运营层,是用于启动和记录企业内部组织单元绩效评价活动的最基本的层次。它处于金字塔的底部,所处理的信息量在四个层次中是最大的。它决策的目的是保证有效完成某个具体操作和任务,所解决的问题通常属于结构化决策。

第二个层次是评价控制层,主要由计划、控制、报告和反馈以及功能衡量组成。它面向绩效信息需求者编制各内部组织单元的绩效指标计划,并在这个基础上结合标准指标值和历史指标值生成各内部组织单元的绩效预算,并对各内部组织单元的绩效活动进行控制、报告和反馈。同时,评价控制层还有对企业内部组织单元绩效信息系统进行功能衡量的作用。

第三个层次是决策支持层,它主要解决的是评估未来策略的可选方案并做出选择。它既包括选择使企业绩效最大化的未来经营方案,也包括根据企业具体需要选择未来各内部组织单元的评价方法和评价指标体系。

第四个层次是战略支持层,这个层次要确定企业的远景、使命和价值观,并通过把这些战略转化成企业及其内部组织单元的经营目标,并层层落实,来最终保证企业实现其价值的不断增值和整体绩效的最大化。它处于金字塔的顶部,所处理的信息量在四个层次中是最小的,解决的问题通常属于半结构化决策。

不同的层次所需要的信息的特性也不同,相邻两个层次所需的信息在一定程度上具有一些共性。而层次之间的距离越远,所需信息的差异性就越大。战略支持层和业务运营层所需信息在特性上具有很大的差异,而评价控制层和决策支持层所需信息介

于前两者之间。

参 考 文 献

- 1 侯炳辉. 企业信息化领导手册. 北京:北京出版社, 1999
- 2 Gordon B Davis. Management Information Systems Conceptual Foundations. Structure and Development, New York: Mc Graw - Hill, 1998
- 3 钱 玲. 电算化会计信息系统. 上海:华东师范大学出版社, 2004
- 4 王 勇, 许庆瑞. 以战略为中心的业绩评价理论框架研究. 科研管理, 2001; (9)
- 5 马 璐. 企业战略性绩效评价系统研究. 北京:经济管理出版社, 2004
- 6 Kaplan Robed S Norton, David P. Transforming the Balanced Scorecard from Performance Measurement to Strategic Management: Part I. Accounting Horizons, 2001; (15)
- 7 靳 新, 刘美云. 财务分析学. 北京:经济科学出版社, 2000
- 8 Satty T L. Decision Making for Leaders. Lifetime Learning Publication, Mexic. City, 1982
- 9 Eccles R G. The Performance Measurement Manifesto. Harvard Business Review, 1991 Jan/Feb91
- 10 Prober, Larry M. EVA: A Better Financial Reporting Tool. Pennsylvania CPA Journal, 2000; 71(3)
- 11 Ittner Christopher D, Larcker David F. Innovations in Performance Measurement: Trends and Research Implications. Journal of Management Accounting Research, 1998; (10)
- 12 M Looijen 著, 耿继秀, 张 璇, 周清华等译. 信息系统管理、控制和维护. 北京:电子工业出版社, 2002

(责编:禄钧)