



IBM全球企业咨询服务部

知识管理

走出信息 爆炸困境

信息分类管理：
提升企业信息价值
的有效途径

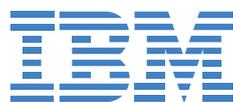


IBM全球企业咨询服务部

通过遍及160多个国家/地区的业务专家，IBM全球企业咨询服务部为跨越17个行业的客户提供深刻的业务流程和行业专业知识服务，并通过创新更快地发现、创造和实现价值。我们以IBM全面雄厚的实力为后盾，为客户提供各种建议，帮助他们实施各种影响深远的、可带来持续业务成果的解决方案。

IBM中国研究院

IBM中国研究院(CRL)成立于1995年9月，是IBM在全球设立的8个研究实验室之一，目前共有190多名研究人员。与IBM全球研究机构一样，IBM中国研究院的使命是创造推动人类前进的世界级的信息技术和基础科学，尤其关注与中国IT产业及其它新兴市场相关的技术和需求。



走出信息爆炸困境 - 信息分类管理: 提升企业信息价值的有效途径

概述

经济的全球化使越来越多的中国企业不得不直面来自国内市场和国外市场的双重压力。为了生存和发展,企业必须保持并不断加强自身的竞争力。竞争力从哪里来?这个问题有多种答案,信息管理是一个较新的切入点。在企业信息化越来越深入,信息积累越来越丰富的情况下,有效的信息管理能够提升信息化的价值,强化企业与客户以及合作伙伴间的信息共享能力,改进企业整体运营水平,并增强企业创新能力和竞争力^{1,2}。正因如此,近年来中国企业对信息管理的关注和投入呈明显增长的趋势。

然而,很多企业并没有意识到,有效的信息管理是建立在科学的信息分类之上的。目前中国普遍存在的情况是:随着企业规模的不断扩大,以及信息化建设的不断发展,信息的总体拥有量越来越大,有效地组织和查找信息变得越来越困难。为了解决不同用户群体和不同业务部门的信息访问的需要,企业在信息管理和信息访问方面的投入不断增加,先后建立了知识管理系统、门户网站、数据管理系统、内容管理系统等等,用来组织、处理、存储和展现信息。可最终往往事与愿违——企业的信息架构越来越复杂,而企业员工定位和查找信息的效率越来越低。

上述种种问题不仅困扰着中国企业,也是全球企业的通病。Gartner通过对全球企业的信息管理状况的研究,指出“到2006年,70%以上的企业在非结构化信息管理方面的投资无法得到预期的回报,主要原因是在信息分类方面的投入不足。”³ Gartner的研究报告还提出,“企业大约每年投资7500亿美元用于查找有用的信息。30~40%的员工的时间花费在管理工作文档。到2007年,60%的信息访问将结合分类(taxonomy)、搜索(search)、本体论(ontology)和可视化(visualization)技术。”⁴

为了解信息分类管理在中国企业中的应用现状,明确信息分类对中国企业的价值和作用,并指导中国企业在信息管理方面的实践,IBM进行了相关专题研究。结果表明,信息分类管理在中国大多数行业中仍处于起步阶段,大多数企业在信息分类管理领域面临挑战,迫切需要行业专家以其科学的方法论、先进的技术和丰富的经验来指导企业进行信息分类管理系统的建设。

本书展示了我们的这项最新研究成果,并在深入分析现状与问题的基础上,为中国企业提供了实施企业信息分类管理项目的基本方针和战略途径,以期帮助中国企业提升信息的价值,实现有效的信息管理,促进企业的成功。

有效的信息分类管理提升企业信息的价值

经过多年在信息化建设方面的不断投入,中国企业在信息系统中积累的信息数量已经相当庞大,它们眼下面临的问题是:如何有效地组织和管理这些信息?如何实现结构化信息和非结构化信息的统一访问?如何进一步挖掘和提升这些信息的价值?

业务发展要求更有效的信息分类管理

企业外部因素 经营环境的改变,更加激烈的市场竞争等外部因素促使企业采用统一的信息分类管理。例如,企业的规模化和全球化发展必然导致组织架构越来越复杂,不同地区之间,不同部门和分公司之间,经常使用不同词汇描述同一种业务,影响了跨部门的协作水平,并导致运营效率下降。5全球化还带来多语种的环境,统一的业务语言(意味着统一的信息分类)更显重要。企业兼并和重组是另外一种情况。兼并后,原来不同公司的业务资料、解决方案、财务报告、工作计划书等保存在不同的IT系统中,信息架构各异,需要有统一的信息分类来指导IT系统的整合。

什么是分类体系?

分类体系(taxonomy)是对给定的对象集合按照一定的分类原则组织成的一个树型的层次结构。在分类体系的最上面的结点是根结点,适用于所有的对象;在根结点下面的结点的分类原则更加具体,分别适用于某些对象。几乎所有的事物,都可以按照一定的分类原则进行分类。例如,对计算机科学进行分类,可以分为软件、硬件、服务和标准;软件可以进一步划分为操作系统、系统软件、工具软件等;硬件也可以细分为台式机、笔记本、服务器等。

分类体系通常可以从多个角度对同一个对象集合进行分类,每个角度称为一个分类维度(facet)。例如,对企业内的大量信息进行描述时,从组织结构这个维度,可定义人事部、市场部、开发部等类别;从客户所在的行业这个维度,可定义为政府行业、金融行业、制造业等;也可以按文档的类型分为产品手册、市场新闻、白皮书、客户信息等。这些不同的维度的分类构成企业统一的分类体系。

什么是企业信息分类管理?

企业信息分类管理是职责明确的组织机构、任务清晰的工作流程与相关的IT技术有效的结合,涵盖企业信息分类的设计、创建、使用和维护整个生命周期。信息分类组织机构的职责是做决策,制定管理流程和管理分类体系。一系列完整有效的流程可以保证信息分类管理得以执行并遵从严格的步骤和规范。而成熟和先进的技术可以帮助企业更快地、更有效地建立分类体系,并保证分类的质量。

企业信息分类通过对企业拥有的各种信息进行合理组织,能够简化信息的复杂性,提高信息的可用性和易用性。在企业信息分类的基础上,可以为企业的员工和客户提供一个统一的信息访问视图,将分散独立的多个系统集中到统一的信息访问平台,有利于很快地定位和查找信息,发现信息之间的关系。从部署角度看,信息分类管理可以作为一个系统独立存在,也可以与其它信息管理系统协同工作。(见图1)

图1. 企业信息分类的典型应用场景



信息来源: IBM全球企业咨询服务部

企业内部因素 管理信息和查找信息越来越困难,这给企业的日常业务带来一系列问题。例如,客服人员由于无法很快地查找到解决问题的相关信息,导致很难立即回答客户的问题,引起客户的不满,甚至丧失了商机;财务和审计部门由于无法及时完成对各类信息的确认和审核,导致无法满足萨班斯法案等内控体系对财务报表和财务报告完整性和准确性的要求;研发部门由于各个研发团队之间缺乏有效的信息和知识共享,有时会发生多个团队重复研究同一课题的情况,不仅浪费了投资,还可能失去研发先机。引入科学高效的信息分类管理成为企业的当务之急。

从IT角度 随着企业信息化的发展,企业的IT环境日渐复杂。门户网站、内容管理、知识管理、数据仓库等系统林立。但是这些系统通常面向不同的领域,以不同的方式组织和管理信息,相互隔离和独立,没有形成企业级的、囊括全部信息的、跨越结构化和非结构化数据的统一的视图。其结果是,尽管这些系统的规模、范围和复杂程度不断地增长,查找信息却越来越困难,为了完成一项工作任务,经常需要到多个系统寻找信息。

企业客服人员对信息查找的意见

国内某制造业公司的客服坐席人员抱怨在答复客户的问题时,经常需要从多个不同的系统界面查找信息,导致客户需要等待过长的时间才能得到回复。例如,需要从客户信息数据库中查找客户基本信息,需要从内容管理系统中获得产品图片和安装说明等资料,需要从ERP系统中获得合同的信息,需要从交易处理系统中获得价格和折扣的信息。

实际上,上述案例并非个别现象。市场研究公司Strategic Focus一项针对全球300家大中型公司的调查结果表明,85%的企业信息是非结构化信息⁶。咨询公司Delphi Group的研究也表明,企业员工花费20%甚至更多的时间查找有用信息,但是只有10%的员工认为查找需要的信息比较容易和快速,高达62%的用户对目前的工作环境中查找信息的方法不满意或非常不满意⁷。

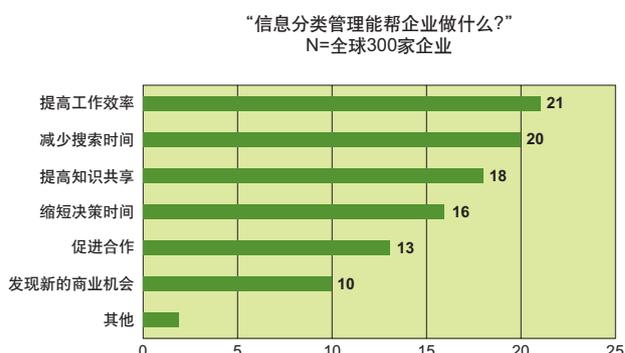
信息分类管理能使企业获益

信息分类给企业带来的价值主要体现在三个方面:

促进信息共享与信息有效访问 企业级信息分类体系不是纯技术领域的工作,而是与业务紧密相关的业务词汇的映像。采用这一分类体系,企业内部的业务描述均采用一致的术语,业务之间的关系也有了统一的表述,企业各部门之间的条块分隔被打破,信息共享真正实现,企业的运营效率得以提高。另外,信息分类浏览与搜索技术的结合大大提高了信息的可理解性,并加快了查找和定位信息的效率。

提高领导决策和员工创新能力 企业级的分类体系可以将各种有用信息集中进行展现,例如,从行业的市场分析、客户的反馈、竞争伙伴的最新动态,可以帮助企业高层领导全面地了解企业的运营和市场动态,有利于提高企业高层领导的决策能力。同时,企业级信息分类使企业一线员工能够从多个角度了解企业,多方位获得知识,有利于提高员工的创新能力。

图2. 信息分类管理为企业带来多方面的价值



信息来源: "Information Intelligence: Content Classification and the Enterprise Taxonomy Practice", June 2004, Delphi Research Report

提高信息化的投资回报率 通过企业信息分类管理,企业一方面可以提高对已有信息资源的利用水平,另外一方面可以减少新建信息系统所带来的成本。实际上,将企业信息分类作为企业所有信息系统建设的标准和规范,可以指导企业信息系统的建设,降低信息重复率和企业IT资源浪费,减少系统开发维护的成本,从而降低信息的管理费用。

中国企业信息分类管理现状与挑战

通过本次专题调研,并结合近几年来与各行各业客户接触的经验,我们发现大多数中国企业的信息分类管理还处于初始阶段。我们的经验涉及企业门户、知识管理、数据管理、内容管理等多种类型的信息管理项目,客户类型包括国有企业和私营企业,客户的业务涉及电信业、金融业、制造业、政府等多个领域。

我们把企业的信息分类管理分为以下三个级别:

初始级 企业内分散地存在一些部门级的、小规模分类体系。分类体系由各个部门独立管理,一般不会考虑与企业的信息架构的关联。中国大多数企业的信息分类管理处于初始级。

重用级 这种类型的企业，一般采用企业级的分类体系，与内容管理、知识管理等系统结合紧密，由企业统一管理。中国少数企业的信息分类管理属于此类。

协作级 企业级的信息分类管理，能够与企业的业务战略保持一致，与业务流程融为一体，并作为企业信息架构的标准之一贯彻执行。只有极少数领先的中国企业能够将信息分类管理运用到这个级别。

尽管对信息分类管理的应用仍处于初始阶段，可喜的是，很多企业已经把知识分类视作知识管理领域的重点并将加大投入。2005年中国一项针对知识管理的调查显示，53%的中国企业将继续增加在知识管理方面的关注和投入⁸。

由于信息分类管理在国内还是一个比较新的概念，中国企业要想成功实施信息分类管理将面临以下主要挑战：

战略关注不够 大多数企业仍然将信息分类看作是纯IT领域的内容，是IT部门的事情。而实际上，需要站在业务战略的高度，分析信息分类与企业组织架构、业务流程之间的紧密联系。战略关注不够也导致在信息分类领域缺乏与业务人员的配合和沟通，导致最终用户接受信息分类的程度比较低。

标准制定滞后 在国内，除了一些特殊行业外，基本上还缺乏面向行业的或者面向企业的信息分类体系。国内企业信息和档案一般由企业的档案室、图书馆(室)或者资料室负责。在信息分类方面，主要参考国家档案馆制定的《档案分类标引规则》以及国家图书馆制定的《中国图书馆分类法》。这些标准都不是根据企业信息管理的要求制定的。

成功经验较少 由于刚刚起步，中国企业在信息分类管理方面的参考案例很少。大多数企业都要在摸索中前进。缺乏经验，缺少科学的方法论的指导和先进技术的帮助，使得部分企业的信息分类体系缺乏权威性与全面性，信息分类管理的作用和价值难以完全发挥。这反过来又阻碍了信息分类管理在企业中的推广和使用。

“信息分类是知识管理系统中最重要的部分。良好的信息分类应该与企业的组织架构、业务流程、客户服务结合起来。可目前国内电信企业的组织架构、业务流程等变化很频繁，缺乏连续性，而信息分类却要求保持一定的稳定性。因此，信息分类在电信企业又是最困难的环节之一。”

—国内省级电信公司CIO⁹

确定企业信息分类管理的基本方针

今天，企业的信息规模十分庞大，覆盖的主题非常广泛，如何将这些内容丰富、形式多样的信息组织在一个科学、全面的分类体系之中，需要系统的考虑和规划。根据全球最佳实践以及IBM的经验，我们建议企业在启动信息分类项目之前必须考虑一个最基本的问题：即如何着手开始进行信息分类体系的建设？

信息分类体系的建设有三种基本的启动模式，即购买现成的分类体系，企业自制分类体系和利用机器自动生成分类体系。

购买现成的分类体系

如果能够找到已经按照行业或最佳实践建立的分类体系，并且主题范围与企业的业务基本一致，企业可以考虑购买这样的信息分类体系，以此为起点进行企业自身的分类体系的设计。

购买创建好的分类体系的好处是可以节约时间和人力，立即部署，快速见效。但是，在采用这种途径时，需要考虑一些重要的问题。首先，需要对现存的分类体系主题的范围、覆盖面以及分类质量进行全面评测。其次，现存的分类体系通常不能直接照搬，因为即使是同一个行业，不同企业的业务模式和内部流程都不尽相同，需要在此基础上根据企业的具体需求，进行针对性地修改和扩展。

企业自己制定分类体系

如果没有已创建好的、适合企业业务的分类体系, 或者, 在企业内部已经存在一些小规模分类体系, 这种情况下, 企业可以考虑自主制定企业级的分类体系。根据企业自身特点制定的分类体系, 可以更好地适应企业的业务流程和组织架构, 并照顾到员工的使用习惯; 可以与企业的信息管理战略目标结合在一起, 更大程度满足企业内外部信息访问需求。

提醒注意的是, 企业自制分类体系需要时间、人力和资金的大量投入, 还需要具备专业的技能。体系制定者需要全面了解企业的运作、业务、管理, 以及过往积累的企业信息资产; 需要深入各个部门进行调查和交流, 了解它们正在使用的一些分类体系, 然后加以整理和总结; 需要从业务和技术的角度来考虑分类体系的科学性、合理性。根据企业的规模、员工的数量以及可使用的资源情况的不同, 一个信息分类体系的创建可能需要几个星期甚至几个月的时间来完成。企业自制分类体系最困难的一点在于, 很难找到既懂得信息分类学科, 又熟悉业务领域的专家。

机器自动生成分类体系

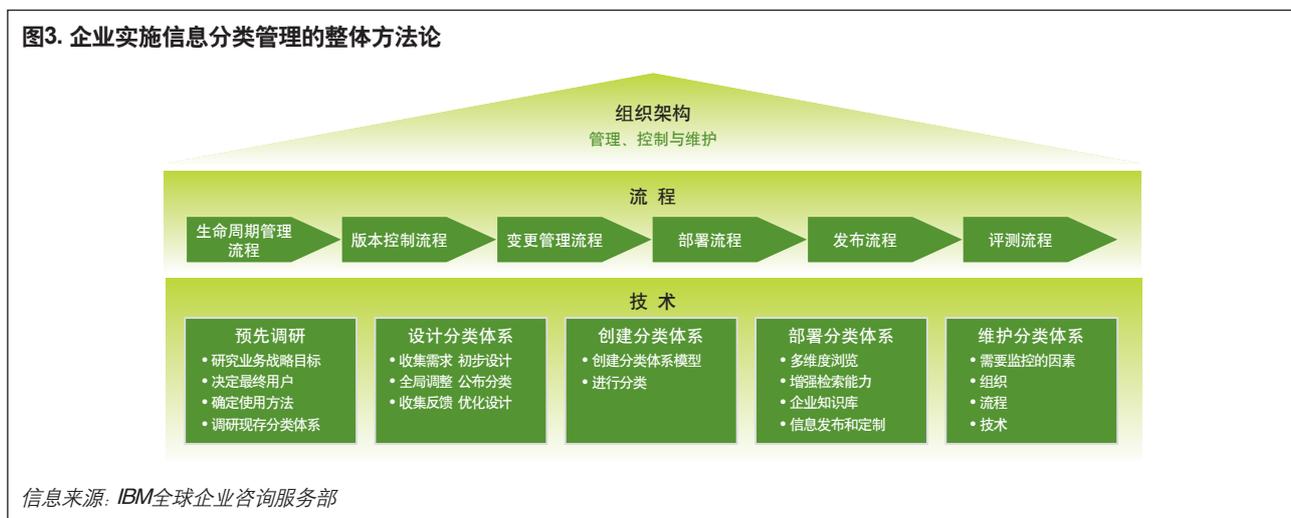
目前, 机器自动生成分类体系的技术基本成熟。这种技术结合自然语言处理和人工智能技术, 通过对信息内容之间的相似性和相关性的分析, 自动生成层次结构化的知识分类体系。

这种方法的优点是“自动化”, 使用合适的自动创建工具, 可以将企业已经积累的信息资源直接输入, 快速处理大量的信息, 在此基础上快速生成分类体系。自动生成的分类体系的缺点是可理解性与直观性不够。每个类别的名称是机器根据统计的方法自动产生的, 抽取经常出现的一些词汇来命名, 例如, 对体育类的文档, 工具可能会用以下几个词来描述: “赛, 中国, 国际, 队, 冠军, 新, 联赛, 场, 全国, 体育”。因此, 企业分类体系的创建不能完全依赖自动工具。

创建分类体系, 没有唯一的最佳途径。一般而言, 手工和自动相结合的方式是最切实可行的途径: 由具备行业知识和分类管理经验的专家来制定分类体系, 同时, 利用分类体系创建工具来加速分类体系的创建速度, 节省人力和时间。无论采用上述哪种模式来创建企业的知识分类体系, 首先要确定企业信息分类体系的基本策略, 并注意结合业务流程和员工的使用需求。

企业信息分类管理的战略途径

针对国内外企业在信息分类管理的实施中可能遇到的各方面的挑战, 我们提出了一套完整的信息分类管理方法论, 指导企业科学地制定切实可行的组织架构和管理流程, 并应用合适的、成熟的技术, 最大程度地获得企业信息分类的收益。(见图3)



建立跨部门的企业信息分类管理机构

企业需要成立一个专门的跨部门的常设机构，负责信息分类体系的制定、应用和发展。其核心管理小组中，需要以下几类关键的角色：

分类体系管理经理 需要具备深厚的行业经验，熟悉企业日常业务，并具有技术背景。一般由信息管理部门经理级别的员工担任。负责制定分类体系管理的政策和工作流程；定期召开分类体系核心小组会议，确定信息分类的需求和流程；与企业高层领导进行沟通，确保分类体系与企业的业务保持一致。

业务专家 由各个业务部门的专家组成。他们需要在分类体系的制定过程中，解释各自部门的业务需求并分析本部门的分类需求。在分类体系的维护过程中，他们负责根据业务的变化和发展，提交信息分类修改需求。

信息分类专家 包括分类学专家，用户体验专家等。他们根据各个部门的需求设计信息分类体系，测试信息分类体系，与最终用户及主要相关人员进行沟通。

技术人员 包括信息分类管理员和系统管理人员。负责信息分类体系的存储、更新和维护方面的技术支持，并将信息分类体系部署到企业的信息系统中。

制定完整有效的管理流程

信息分类管理流程是用于企业信息分类体系的创建、管理与维护的流程。一套完整的信息分类管理流程包括6大类流程，即生命周期管理流程、版本控制流程、变更管理流程、部署流程、发布流程、评测流程等。

生命周期管理流程 信息分类类别不是固定不变的。在企业信息分类体系中，分类体系的类别会经历创建、使用、修改和退出四个不同的阶段。生命周期管理就是管理其演进过程的。

版本控制流程 维护信息分类体系修改的历史，管理企业信息分类体系的不同版本的修改、定稿和发布。需要借助于版本管理工具实现。

变更管理流程 根据业务的变化和技术的发展，当修改需求产生时，启动本流程进行新需求的分析，对变更进行审核，并发布新的信息分类。

部署流程 部署流程规定如何将信息分类贯彻、部署和应用到企业现有的信息系统中去。其中非常重要的一点是评估对现有信息系统的影响，并审核部署方案的可行性。

发布流程 企业信息分类不能停留在纸面上，需要在企业内部进行广泛的宣传、培训和交流，以得到用户的反馈，促进用户的理解和应用。

评测流程 定期和不定期地对企业信息分类的使用状况、投资回报等进行分析评估。

按照科学的方法实施信息分类

实施信息分类管理，需要按照科学的方法论的指导，遵从一套完整的步骤，以选定的启动模式开始，应用先进和成熟的技术来加速信息分类的实施，最终将分类体系部署到企业的信息架构中。下文将详细阐述实施信息分类管理的关键步骤。

成功实施企业信息分类管理的关键步骤

信息分类管理的实施包括预先调研、设计分类体系、创建分类体系、部署分类体系和分类体系的维护五个阶段。需要指出的是，由于企业信息分类在不断发展，这五个阶段又是一个循环往复的过程。

阶段1: 预先调研

在启动企业信息分类管理项目之前，需要对内部的情况和外部的资源进行详细的调研，并给出现状分析报告。

案例: 信息分类管理的好处¹⁰

美国一家大型制药公司成功地将信息分类应用于内容管理系统, 为该公司分布在全美的4万多名员工服务。建立这套系统前, 该公司面临的主要问题是, 企业内部存在大量的格式各异的内容文件, 存储在多个文件服务器中, 并通过多个相互独立的内部网站进行访问, 内容的组织缺乏标准, 查找信息非常不准确。

考虑到企业内容的特性和用户的使用需求, 该公司决定采取自建分类体系的途径。他们一方面从现有的内容资料中分析和抽取分类特性, 另一方面从公司业务层面定义企业级的分类体系, 最终通过和员工进行沟通, 确定下来一套容易理解、并能覆盖全企业的内容分类体系。在应用分类的过程中, 该公司首先将分布式的内容迁移到一个集中的内容管理系统中, 然后采取先机器自动分类, 后手工确认的方法。在部署方式上, 他们将企业的内容按照分类体系的组织方式展示到企业信息门户, 并结合分类体系浏览与基于内容的搜索技术。

该系统的实施不仅为该企业员工提供了准确、快速的一站式信息访问服务, 大大提高了工作效率, 同时还大大节省了在信息内容管理方面的投入。

这一报告的内容应包括六个方面: 首先要了解企业的战略目标, 确保企业信息分类支持企业的业务战略目标, 并与企业的业务流程融为一体; 其次, 要了解最终用户, 综合考虑各种用户的需求, 比如外部用户和内部用户对分类体系的需求就是不同的; 第三, 要确定分类体系的使用方法, 以便与企业的信息管理系统进行集成, 这将关系到分类体系的设计及其最终的部署方案; 第四, 了解现存的分类体系以资借鉴; 第五, 调研现有的结构化和非结构化的信息资源, 有助于更好地理解目前员工使用信息和利用信息的情况; 最后, 了解企业中已经具备哪些信息分类项目需要的专业技能, 以确定需要从外部获取和采购那些尚不具备的技能。

阶段2: 设计信息分类体系

设计信息分类体系的目标就是得到多层次、多维度的信息分类结构。这个过程需要以下6个步骤:

收集需求 对各个部门进行访谈, 了解各自的组织架构、部门职责和业务流程, 以及对分类体系的需求, 并收集目前的知识库和分类体系存在的问题等, 然后综合各部门的调查结果, 得出分类体系设计需求报告。

初步设计分类体系 根据分类体系设计需求, 确定分类体系划分的各个维度, 例如, 按照组织结构、业务类别、文档类型等划分维度。并进一步设计每个维度的下级分类。

全局调整 将各个维度的信息分类体系的初步设计方案综合起来, 审查分类体系中的类别是否需要调整, 例如类别的扩充、合并或移动, 得到企业整体的分类体系设计方案。

公布分类体系 将整体信息分类体系的设计方案在企业内部予以公布, 并与最终用户以及企业信息管理人员沟通信息分类的设计。

收集反馈 从各个部门及员工得到反馈意见, 例如, 分类的名称是否容易理解, 分类体系是否合理, 能否覆盖主要业务、工作需要和查询需求。

优化设计 按照意见反馈, 进一步优化分类体系的设计。

在这个阶段, 需要遵循信息分类管理小组制定的相关的工作流程; 并需要一个集中的管理系统进行信息分类的保存、发布, 以供企业员工访问; 另外, 这个系统还需要能够记录设计过程中收集到的资料、各种调查记录和分析报告, 以方便对分类体系的修改进行跟踪记录, 有利于今后的维护。

国内某电信公司的信息分类体系设计¹¹

根据用户对原有知识分类体系的反馈，并参考行业模型以及相关的标准，该公司采用了四个维度的知识分类体系：组织分类，即按照该公司当前的组织架构划分，定义市场部、客户服务部、财务部等分类；以专业分类，按流程定义调研分析、营销策划、产品定价等分类；按用途分类，如政策法规、标准规范、行业动态等；按内容信息分类，以知识库内容的核心信息点设计分类，例如客户、竞争对手、业务、资费、网络设备、供应商等。

需要指出的是，在每个维度下面可以进一步细分，从而形成多维度、多层次的知识分类，方便信息的访问。实施以上信息分类体系后，该公司员工普遍反映信息访问效率与实施前相比有了很大的提高。例如，市场部的员工在做方案设计和市场活动的策划时，可以首先浏览组织分类下的市场部，得到本部门的近期活动规划；在其基础上，进一步浏览专业分类下的营销策划，了解策划方面的相关知识和注意事项；再通过用途分类的行业动态，了解目前本行业的一些最新动态；然后找到内容信息分类下的客户类，根据客户的类型进行分析。

阶段3: 创建分类体系模型并进行分类

信息分类管理系统的核心任务就是将企业内部的全部信息，按信息分类体系的结构进行组织，也就是说，将企业的信息打上“类别”标记。由于信息量的巨大，这也是最耗时的一个阶段。

针对具体信息进行分类有2种方式：手工分类和自动分类。总体来讲，在分类体系规模较小、文档数目很少的情况下可以选择手工分类。通常情况下，企业内部的文档数量非常庞大，人工对每一篇文档进行浏览然后分配到相关的类别，时间和人力投入非常大，手工分类往往无法满足需求；尤其是在分类体系非常复杂的情况下，手工分类的准确性也是一个问题。因为不同人对分类体系的理解不同，而且单个人的知识不可能覆盖所有的类别，就会造成不同人的手工分类不一致¹²。相比之下，机器自动分类的优点是速度快，成本低，一致性强¹³。但是，机器毕竟是机器，自动分类也达不到100%的准确率。另

外，自动分类的准确率依赖于大规模的高质量训练文档，并且收集训练文档需要花费较大的人力。

我们的建议是采用手工分类与自动分类相结合的方式。例如，自动分类可以对大量的文档进行初次分类，对重要的文档可以经过人工的确认，以保证比较高的准确率。系统可以收集手工纠正的文档，反馈给自动学习算法，用于改进和优化已有的自动分类模型，提高自动分类的准确率。

阶段4: 将分类体系部署到企业信息基础架构中

通常，信息分类体系不是孤立地运行，而是与其它信息系统一起工作，来提高和改进信息访问的效率。可以采用下面的四种方式部署信息分类体系：

多维度浏览 信息分类体系可以应用在内容管理和知识管理系统中。例如，通过浏览器展示多维度的分类体系，并且为用户提供自由的、在多维度之间进行切换的功能，将给用户获得知识带来很大的方便。

增强检索能力 将信息分类体系和全文搜索引擎集成起来，提高信息查找的效率。与搜索引擎的结合方式，可以是先搜索，然后按类别浏览；或者先浏览分类体系，然后在选定的类别下进行全文检索，从而有助于快速定位和发现所需的信息。

企业知识库 企业的知识分类本身就是一个很好的知识库。企业信息分类体系潜在的意义在于，具有层次结构的分类把不同的信息按照父子关系联系在一起，相关的知识在特定的上下文环境下更容易被发现。在浏览知识分类体系的过程中，员工就是在不断的学习和接受培训。

信息发布与定制 企业信息分类可以与企业信息门户，或者商业智能系统进行集成。在门户系统中，信息分类体系可以直接作为信息导航架构使用。不同的员工或客户由于工作职能不同和关注的主题不同，需要不同的信息。一般来讲，这可以通过基于信息分类体系的个性化的定制，以满足个性化的信息需求。

以上只是企业信息分类应用最典型的一些场景。实际上, 企业信息分类体系作为企业信息基础架构的重要组成部分, 可以应用到企业几乎所有的IT系统中, 作为企业的IT规范进行应用。

阶段5: 持续不断地维护企业信息分类

在企业中, 业务战略也可能发生变化, 组织架构和业务流程会不断地调整, 导致不断有信息在产生, 新的信息分类需求也会出现。这些都需要信息分类体系能够保持灵活性, 及时得到更新和维护。

另外在信息分类体系的维护中, 还需要借助长期存在的信息分类管理机构, 通过合适的反馈机制, 及时导入新的成熟技术和工具。无论是管理机构, 还是管理流程、技术和工具, 都需要企业长期不断地进行投入。如果没有长期的、足够的投入, 企业分类体系很快会过时, 这很有可能导致企业信息管理和信息访问又退化到初级阶段, 浪费前期的投资。

结论

随着全球经济的发展, 中国企业面临越来越大的竞争压力, 如何充分管理和利用企业内部的信息, 是企业提高竞争力的一个重要因素。企业信息分类是企业成功实现信息管理, 提高信息使用价值的关键。通过有效的信息分类管理, 可以提高企业的竞争优势, 降低运营成本, 增强员工专业技能, 提高客户服务质量。但是, 大多数中国企业目前还没有认识到信息分类管理的重要性, 在信息分类管理方面还处于起步阶段。企业需要克服一系列困难和挑战, 才能成功实施科学有效的信息分类管理, 提升信息的价值, 增强竞争力。

关于作者

张俐, IBM中国研究院资深研究员

电子邮件地址: lizhang@cn.ibm.com

徐启昌, IBM全球企业咨询服务部资深顾问

电子邮件地址: xuqc@cn.ibm.com

Keh-shin F. Cheng(郑傅可欣), IBM Watson研究中心高级软件工程师

电子邮件地址: kfcheng@us.ibm.com

合作者

苏中, IBM中国研究院资深研究员

Kashyap Avashia, IBM全球企业咨询服务部资深顾问

致谢

本文作者在此感谢IBM中国商业价值研究院主管 Alan Beebe、IBM全球企业咨询服务部合伙人Simon Benjamin和IBM商业价值研究院副合伙人Eric Lesser对本文内容提出的改进建议; 感谢IBM全球企业咨询服务部资深顾问马斌的经验分享; 感谢IBM中国研究院战略合作部蔡菁菁的大力帮助; 并非常感谢IBM商业价值研究院顾问孙爱军对本文内容的建议和调整。

参考文献

- 1 Amy Casher, Eric Lesser, "Gray Matter Matters: Preserving Critical Knowledge in the 21st Century", IBM商业价值研究院白皮书, 2003
- 2 苏咏, Amy Casher, Eric Lesser, "知识无价: 构建现代企业的知识管理体系", IBM商业价值研究院白皮书, 2005
- 3 Rita E. Knox, Debra Logan, "What Taxonomies Do for the Enterprise", Gartner Research Report, September 10, 2003
- 4 "Predicts 2005: Support Improves for Knowledge Workers", Gartner Research Report, November 1, 2004
- 5 Alan Beebe, 丘琪铮, 冯月圻, 石代伦, "中国企业走向全球 - 实践、挑战与对策", IBM商业价值研究院白皮书, 2006
- 6 "2001 Trends in Enterprise Portal & EAI Markets", Strategic Focus Publishes Results of Major Research Study, Strategic Focus Inc.
- 7 "Information Intelligence: Content Classification and the Enterprise Taxonomy Practice", Delphi Research Report, June 2004
- 8 2005知识管理&创新论坛, 《IT经理世界》、《计算机世界》、《首席财务官》杂志、计算机世界网等媒体与深圳蓝凌公司共同主办。
- 9 IBM全球企业咨询服务部
- 10 同上
- 11 同上
- 12 Alan D. Marwick, "Technology and the Future of Knowledge Management", workshop on Knowledge Management and the Global Firm, Sydney, February 6-7, 2003
- 13 Wendi Pohs, Stephen C. Gates, Jochen D rre, "Implementing Enterprise Taxonomies with WebSphere Information Integrator OmniFind Edition", April 2005



© Copyright IBM Corporation 2006

IBM Global Services
Route 100
Somers, NY 10589
U.S.A.

All Rights Reserved

IBM, and the IBM logo are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both.

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products and services do not imply that IBM intends to make them available in all countries in which IBM operates.

北京总公司

北京朝阳区工体北路甲二号
盈科中心IBM大厦25层
邮政编码: 100027
电话: (010)65391188
传真: (010)65391688

上海分公司

上海市淮海中路333号
瑞安广场10楼
邮政编码: 200021
电话: (021)63262288
传真: (021)63261177

广州分公司

广州林和西路161号
中泰国际广场B塔40楼
邮政编码: 510620
电话: (020)85113828
传真: (020)87550182