

● 李晓翔 谢阳群

关于政府信息工厂的若干思考^{*}

摘 要 政府信息化建设的一个基础性内容就是构建政府信息系统。当前政府信息系统大多是由各个政府机构根据局部的信息需求构建起来。它们存在诸如数据转换成本高、缺乏稳定性等问题。建立政府信息工厂(GIF)能够解决这些问题。GIF包括3个主要构件:操作型构件、数据仓库和用于支持决策的多维构件。它有以下优势:在整个政府层面构建了一致性框架,增加了各系统之间的互通性;能够根据需求自主地定义信息生产活动;低成本地保存和处理历史数据;保证政府信息安全;增加信息产品生产的灵活性和可控性;降低使用成本等。表1。图4。参考文献9。

关键词 政府信息工厂 政府信息系统 系统整合 系统架构 信息服务

分类号 G250.76

ABSTRACT A basic work of the government informatization is to construct government information systems. At present, most government information systems are constructed by different government departments according to their local information requirements, and have many problems, such as high cost and low stability. The authors think that to construct government information factories (GIF) can solve these problems. In this paper, the authors analyze GIF's components and advantages, which include consistent frameworks, interoperability, low costs, information security, flexibility and controllability. 1 tab. 4 figs. 9 refs.

KEY WORDS Government information factory. Government information system. System integration. System infrastructure. Information service.

CLASS NUMBER G250.76

政府信息系统是基于网络的符合因特网技术标准的面向政府机关内部、其他政府机构、企业以及社会公众的信息服务和信息处理系统,是一个利用信息和通信技术,有效地实现行政、服务及内部管理等功能,在政府、社会和公众之间建立有机服务系统的集合^[1]。纵观发达国家政府信息化的进程,从20世纪70、80年代的办公自动化,到80年代的管理信息系统,一直到90年代的“电子政府”,政府信息化在社会生活中的地位和作用发生了质的飞跃^[2]。但是,当前的政府信息系统正面临着新的挑战,例如如何从整个政府的层面整合和共享信息资源、如何满足信息消费者各种定制化需求。本文在研究当前政府信息系统所存在问题的基础上,借鉴W. H. Inmon所提出的系统建设逻辑,探讨关于构建政府信息工厂的若干构想。

1 当前政府信息系统存在的问题

当前政府信息系统大多是由各个政府机构根据

局部的信息需求构建起来的。各个政府机构面向特定的行政或管理事务,所建立的信息系统也都面向特定的主题。一旦某个政府机构基于特定的供应商构建了信息系统,该机构就有了很高的转换成本,而且这种成本随着时间的推移不断增加。原因在于,政府的大量数据资料全都是用特定的格式存储的,转移到另一种格式需要大量费用并面临损失数据的风险;而且政府的大批公务人员为熟练使用该系统,也需要花费相当大的费用^[3]。因此,这种系统只能按照既有的模式发展,形成了所谓的“烟囱式”架构(stovepipe architecture),如图1所示。随着政府信息化的发展,各个政府机构的信息系统需要进行集成,以实现信息资源的有效利用。而集成这些“烟囱式”系统,通常采用的模式是在两个系统之间安放一个转换层或者接口,以实现数据或信息的传输。这样形成了图2所示的“蛛网”模型。

通过图1和图2可以看到,开始时政府信息系统

可能只是一些简单的应用。它们都围绕需求而建立,

^{*} 本文系国家社会科学基金资助项目“企业危机信息管理体制研究”部分研究成果,项目编号04BTQ015。

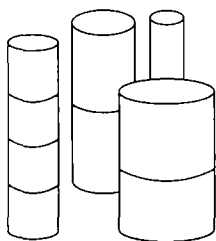


图1 “烟囱式”系统

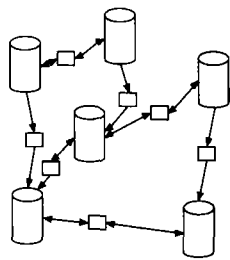


图2 “蛛网”模型

需求决定了内容、结构以及应用的处理过程。随着各个政府部门信息系统的建立,逐步形成了大量的应用。这些应用的大量存在带来了一些问题,例如:①从信息系统整合的角度看,每个政府机构的应用都有自己对于事实独特的理解,并基于此设计构建自己的系统,而不管其他信息系统所包括的数据,从而导致应用之间存在太多的转换界面。各个应用环境不同,转换并不能保证那些数据和信息原料从内容、结构等方面一定满足要求,政府机构很难跨越部门界限看待数据和信息,其跨部门决策的能力将因此被削弱。②从面对变化的角度看,当有变革需求时(例如新技术出现、新的信息需求出现),所有政府机构的信息系统都要相应改变,这需要大量的协调活动,会引起政府信息系统的急剧震荡。③从数据或信息的可获得性看,一旦数据和信息被局限于特定的应用环境,就将被限制使用。它们被相应的政府机构所占有,这些机构为了自身的政治利益等可能会限制对于数据和信息的共享。数据和信息即使跨机构流动,也仅仅以满足对方的请求为标准。④从数据和信息质量的角度看,各个政府机构信息系统的数据库,根据各自所面向的应用进行处理,为了节约支出和提高在线处理的速度,数据和信息往往按照系统的主题被进行不同程度的汇总,在整个政府层面并没有建立一致的数据和信息管理模式,因此很难对数据进行规范的质量管理。

2 政府信息工厂

针对当前政府信息系统面临的问题,我们探讨在政府机构中建立信息系统的一种新的框架结构,是由 W. H. Inmon 所提出的政府信息工厂(Government Information Factory, GIF)。GIF 能够满足所有信息系统的请求。它囊括了数据和处理、活跃型数据和不活跃数据、操作型处理和信息型处理、网络通信、数据粒度、数据的多维处理、数据整合、跨部门的访问数据、元数据、数据安全等^[4]。GIF 将以一致、集成的方

式处理这些事,为政府机构提供整合所有信息资源的新思路。GIF 可以面向一个城市、省、国家等不同层面而构建,它使信息系统不再依附于政府机构,而成为相对独立的信息产品生产者,与政府信息相关的信息活动均在此进行;各个政府机构作为数据原料的供应者和信息产品的消费者,不再直接单独设置专门的信息部门,不需完成所有的信息活动。

2.1 GIF 的构件

GIF 包括 3 个主要构件,即操作型构件、数据仓库和用于支持决策的多维构件(即数据集市)^[5]。操作型构件主要是指应用型软件的集合,我们称之为工具箱,每个政府机构都能够利用其中诸如数据挖掘、联机分析处理等商务智能类软件工具进行决策。数据仓库是整合 GIF 的核心,它用于向整个 GIF 提供数据原料的支持。数据仓库支持海量数据的存储,能够提供粒度最细的数据,并且能够建立整个 GIF 中一致的数据和信息视图。多维构件则是专门面向特定的数据和信息处理,它从数据仓库中获取粒度最细的数据,然后按照政府机构的特定要求对数据进行集成、汇总等不同的操作处理,以支持操作型构件对数据的需求。数据在这 3 个主要的构件之间流转,使得这些构件相互连接。

2.2 GIF 的构建

建立 GIF 是一项系统性工程,首先需要对现有的政府信息系统的状况进行分析。在分析的过程中,需要考虑:目前信息系统之间的互通性如何?信息系统之间的应用以及所存储和处理的数据是否存在重叠?跨越不同机构界限制定决策时,当前信息系统能否满足决策对于数据的要求?

在了解目前信息状况之后,就需要设计 GIF 的蓝图。GIF 蓝图主要包括数据模型、技术架构、系统架构 3 个部分。数据模型决定着 GIF 的成败。GIF 在所处的政府层面内要能够建立一致的多维数据模型,而且其中的数据必须是粒度最低的。只有这样才能支持不同的应用对于数据的不同需求,做到数据的一次性存储,减少 GIF 的运营成本。从新技术引进的角度看,GIF 总能满足新的信息技术对于数据原料的需求,GIF 支持自身的不断优化和升级。

按照整合信息系统的常用逻辑,有两种不同的构建方法可供选择,即“大爆炸”法和进化法。“大爆炸”法类似于 W. H. Inmon 所提出的企业信息工厂(Corporate Information Factory, CIF)架构^[6],即完全打碎目前的信息系统架构,重建一个新的、完全集成

的 GIF,如图3。它包括:①外部环境。包括数据原料提供者和信息产品消费者。GIF 的运转需要不断与外部环境进行数据或信息的交换,它必须要创建有效的信息传递界面,实现积极而持续的输入和输出。②应用。指专门面向特定事务的集成或非集成软件系统,包括专门收集数据和信息原料的应用以及面向决策的操作型软件。成熟的应用应当是基于 GIF 的信息生产流程而集中开发的;但是目前政府机构中大量存在着专门用于特定活动的应用软件或应用系统,我们需要把这些非集成的应用整合到统一的架构中。③转换层。根据数据模型和转换逻辑改造、重组、集成数据原料和初始信息。数据经过转换层后传递给操作型数据存储和数据仓库中。④操作型数据存储。主要用于暂时存储实时数据,并支持对历史数据与实时数据的分析。一般来说,它对转换层输出的数据进行系统集成,并向数据仓库输入数据。⑤数据仓库。⑥数据集市。它是为满足特定信息消费者的信息需求、快速解决某些具体问题而建立的一种灵活的分析型环境。数据集市根据信息活动规则抽取数据,并按照特定逻辑和需求对数据的结构、格式和粒度等进行处理,为政府机构或应用软件提供信息和数据支持。⑦网络。GIF 的网络系统包括互联网和内联网所组成的通信线路及其所承载的数据流。网络将不同的构件连接为一个整体,实现信息生产的流程化。基于网络技术的 GIF 还可以实时传递信息消费者对信息产品的需求定义,并能将相应的信息产品交付给信息消费者。此外,网络(尤其是互联网)也将成为 GIF 最主要的数据来源通道。⑧元数据。元数据是 GIF 的数据和信息目录。元数据通过描述数据仓库中的数据类型、结构、质量,定义、安排数据仓库中数据和信息的输入与输

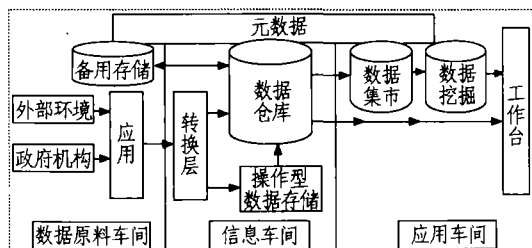


图3 “大爆炸”式的 GIF

出,以及提供基于信息消费者需求的信息,从而支持 GIF 对数据的管理和维护,提高利用数据和信息的效率与效用。这种方法所建立起来的 GIF 实质上是建立了一个信息中心:GIF 从不同的数据源获取数据,数据

经过相应转换然后集中存储在一个巨型的、一致的数据仓库中。各个信息消费者通过客户端(即工作台)发送信息产品的订单,然后 GIF 进行相关数据和信息的抽取、处理等,并通过操作型应用软件进行分析处理,最终支持决策,并将其推至客户端。

进化式的构建方法遵循的是 Ralph Kimball 所提出的构建数据仓库的理念。他认为“数据仓库仅仅是构成它的数据集市的联合”^[7]。基于这种理念构建的 GIF 中(如图4),各个政府机构的客户端通过使用“一致的”维,能够共同看到不同数据集市中的数据,而一系列维数相同的数据集市通过递增、建立相互的关联,最终建立一致的数据仓库。在这种模式的 GIF 中,最为核心的内容仍然是建立数据模型。通过数据模型的建立,各个政府机构的信息系统都能按照一致的标准存储数据,各个系统之间的数据能够自由流动。这种方法所构建的 GIF 能够灵活地将“烟囱式”系统逐步集成,减少新架构的建立给正常的政府工作带来影响。步骤是,首先整合数据集市,等它们集成为一致的数据仓库后,再逐步进行各个应用软件的集成,最终形成图3所示的结构。

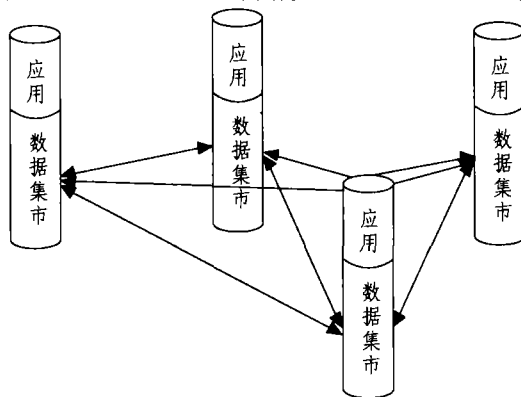


图4 进化式的 GIF

2.3 GIF 的实现

GIF 架构及各组件建立之后,我们可以按照政府活动流程和信息生产的各个阶段将 GIF 分为数据原料车间、信息车间、应用车间及工作台(图3)。

(1)数据原料车间的职责是从各个政府机构及外部环境获取数据原料。随着信息技术的发展,外部数据和信息的获取将逐步通过应用软件而自动实现,例如可以利用内容管理等技术根据自己定义的数据或信息需求而对互联网上的信息自动抓取。备用存储作为数据仓库的扩展组件,负责为其提供数据支持,故也被归为数据原料车间。GIF 在建立之初还需

要面对原有“烟囱式”系统的数据迁移问题,例如现有的信息系统环境的大小和复杂性、所构建 GIF 环境的复杂性及所采用的技术、政府各个机构对它的支持等。

(2)信息车间主要对原料数据进行预处理和存储,对应用系统采集到的各种结构的数据进行抽取、净化、转换并装载至数据仓库,以便应用车间的即时访问。信息车间根据应用车间的信息需求和政府的信息管理战略来调整容量与内容。

(3)应用车间和工作台共同构成了前台工作区。应用车间实质上就是一个包括各种信息技术的工具箱,它在信息车间所提供数据和信息的基础之上,整合各种信息技术 and 知识,既提供一般化的信息处理能力,也提供基于数据挖掘工具、OLAP 工具等商务智能的产出,满足各种定制化的信息产品需求。工作台是连接信息生产和信息消费的界面,负责接收信息消费者的请求和交付信息产品。GIF 的前台工作区应当基于模块化的理念而构建,支持信息消费者对于任意单个或多个信息技术工具的灵活应用。

2.4 GIF 的优势

通过以上对 GIF 的介绍,我们总结出 GIF 与当前政府信息系统之间的显著区别(表 1)。相对于目前的政府信息系统而言,GIF 有如下优势:①它从全局的视角构建数据模型,定义数据整合的规则,并且利用元数据的方式说明这些规则。在生产信息产品的过程中,它从数据源获取数据,并对这些数据进行整合、清洗和集成,最终通过数据仓库提供一致的整体视图,并不管数据最终将如何呈现给信息消费者。它在整个政府层面构建了一致性框架,增加了各信息系统之间的互通性。②它作为相对独立的信息产品生

产者,能够根据信息消费者的需求自主地定义信息生产活动,而并非以单个政府机构的政治利益为导向,因此能够有效地识别各种数据和信息的价值。③它通过数据仓库和备用存储等存储粒度最细的数据,做到数据的一次性保存,减少了冗余数据的存在,并能够低成本地保存和处理历史数据,这些历史数据同时也能够面向各种应用。④它把信息作为政府的一种相对独立的产出,生产信息产品的过程有一定的独立性和流程性,加上其所构建组织环境使得政府进行信息和数据的全面质量管理成为可能。此外,信息活动集中在 GIF 中进行,各个政府机构只是放置客户端,政府信息系统的集中管理便得以实现,政府信息安全得到进一步保证。⑤数据仓库能够提供最细粒度的数据,加上 GIF 的工具箱集成了各种应用、信息处理能力,使得信息消费者在统一的工作台上就能实现对于数据和信息的各种操作处理。信息消费者可以通过工作区定制所需的产品,GIF 随后安排相应的信息产品生产,实现了对于数据和信息的拉动式生产,增加了信息产品生产的灵活性和生产过程的可控性。⑥它提供了一致性的系统构建模式,其扩展性得到了增强。引入新的信息技术只需要更换相应的构件,而其他的构件仍可以继续投入使用,因此它支持不断增加和集成新的应用。而新应用所需的数据和信息原料等也能够从数据仓库中获得,降低引入和使用成本。GIF 本身也是一个信息生态系统,它将在构件不断换代和增加的过程中得到优化。传统的“烟囱式”系统则往往伴随着核心技术生命周期的结束而消亡,新信息技术的引入会引起相应政府机构的震荡,因为“烟囱式”系统的数据存储等都要进行转换以支持新技术。

表 1 GIF 的信息生产与当前政府信息系统的信息活动之间的区别

区别		目前的政府信息系统	GIF
开发	开发方法	按照定义的信息需求以及系统开发人员对于政府信息需求的理解而开发单个政府信息系统。	设计 GIF 蓝图,建立核心构件,然后不断增加、更改或优化其他构件。
	开发视角	面向单个应用,以局部的视角设计系统架构,很少考虑各个系统之间的相关性。	面向所有应用,以前瞻性的、全局的视角和系统的观点开发 GIF。
	开发理念	强调如何将系统整合至政府管理活动中。	支持各个主体对于各种政府信息产品的需求。

续表

区别	目前的政府信息系统	GIF
运作	数据来源	单个信息系统只支持来自特定数据源的、具有特定格式的、结构化的数据。
	数据整合	在不同的平台上建立各自的数据库,然后通过参考数据和转换界面实现整合。
	数据存储	利用孤立的、专门记录历史数据的大容量数据库或数据集市存储数据。数据和信息由信息系统所服务的政府机构控制。
	数据抽取	基于行政活动流程和系统预先设定规则的推动式。
	信息产出	面向即时处理,能够快速满足特定的信息需求,信息在进入信息系统时就已经确定了用途。
	质量管理	数据和信息往往被各个部门占有,难以统一进行数据和信息的质量管理。
	新技术的应用	独立开发的、分散的。
责任	由各个政府机构的信息系统人员负责提供技术支持,数据和信息由各个部门拥有和管理。	信息人员负责各种工具的提供以及数据和管理、审计等工作,进行信息需求的开发和信息产品的营销,拥有更大的责任和独立性;政府信息主管专注于信息产品本身的生产,在政府组织中的地位得到进一步提高。
信息安全	各个政府机构都有自己的信息系统,均需要进行相应的信息安全工作。	集中了政府机构信息系统的核心构件,能够统一进行信息安全管理。

3 美国政府对 GIF 的构建

1996年美国国会通过 Clinger-Cohen 法案,明确指出政府部门缺乏展现以及协调和管理整个信息系统建设、使用、维护的总体框架。针对法案要求,1999年美国各级政府部门 CIO 组建委员会,出版了企业架构的评测标准及实践指导文档,称为 FEAF(Federal Enterprise Architecture Framework, 联邦企业架构框架),为各级联邦政府机构的机构设置提供参考标准。美国各级政府部门随后陆续发布各自的企业架构,如美国财政部的 TEAF(Treasury Enterprise Architecture Framework, 财政部企业架构框架)、美国国防部的 DODAF(Department of Defense Architecture Framework, 国防部企业架构)、美国环保局的 EPAF

(Environmental Protection Agency Architecture Framework, 环境保护局企业架构)等^[8]。这些法案和架构为特定层面的政府机构整合信息系统特供了依据。

在 9·11 恐怖袭击事件之前,美国政府需要的是高效率地、一致性地处理信息,事件之后,美国政府意识到需要更大程度上跨越政府部门的界限共享和整合数据,更多地关注国家安全。因此,美国政府在 2001 年 9 月 15 日着手构建自身的 GIF。但到目前,美国政府联邦层面的 GIF 仍然没有得到完全实现。

4 结束语

对于政府机构而言,信息需求的数量日益增加,政府信息系统的复杂性也急剧增加。政府需要更为精细的信息系统来跨越不同的部门界限整合和管理

数据,以面对多样化的信息需求。GIF 能够整合多种因素以形成综合、集成的信息系统战略,它对于政府信息处理需求提供了最高层次的系统视角。GIF 对于政府而言是一种全面的、详尽的事务处理架构。但是成功实现 GIF 还面临诸多困难,例如数据库管理技术、存储技术、多维技术等需要进一步发展;如何实行供应商所提供技术之间的整合,如何推行数据标准和技术标准以实现各个政府机构应用软件之间的互通性;如何消除单个机构对于数据和信息的独占,其本质就是如何实现各个机构之间政治利益的合并等等。本文对 GIF 政府信息系统建设和整合方案进行探讨,希望能对当今的各级政府信息主管提供一些有益的启示。

参考文献

- 1 刘春年,姜策群.政府信息系统建设承包商的选择问题.情报科学,2004(10)
 - 2 吴新博.政府信息化问题与对策.情报理论与实践,2005(6)
 - 3 王刊良.从多样化、标准化与系统可靠性看中国政府信息化建设.西安交通大学学报(社会科学版),2004(3)
 - 4,5 W. H. Inmon. Slaying The Stovepipe Dragon: Part 1, 2003-01. <http://www.inmongif.com>
 - 6 W. H. Inmon 等著;刑国庆等译.企业信息工厂.北京:电子工业出版社,2004
 - 7 R Kimball, L Reeves, M Ross et al. The Data Warehouse Lifecycle Toolkit. New York: Wiley, 1998;27
 - 8 刘嘉.如何在同质中寻找异质. [2006-04-04]. <http://www.zdnet.com.cn>
- 谢阳群 宁波大学商学院教授,管理学博士。通信地址:浙江宁波大学。邮编 315211。
- 李晓翔 宁波大学商学院 04 级硕士研究生。通信地址同上。(来稿时间:2006-08-12)
-
- (上接第 37 页)息化进程的日渐明朗,人们对信息化指标体系结构理性的探讨,终将能够作为一种理论成果,对现实产生影响。
- #### 参考文献
- 1 [英]威廉·配第著;陈冬野译.政治算术.北京:商务印书馆,1960;8
 - 2 斯拉沃热·齐泽克等著;方杰译.图绘意识形态.南京:南京大学出版社,2002;105
 - 3,20 [美]约翰·希利·布朗,保罗·杜奎德著;王铁生,葛立成译.信息的社会层面.北京:商务印书馆,2003;23
 - 4 姜爱林.中国信息化的涵义与一般特征.图书馆工作,2003(2)
 - 5 秦海菁.知识经济测评论.北京:社会科学文献出版社,2004;112
 - 6 郑建明,王育红.社会信息化进程测度案例及方法分析.图书与情报,2000(2)
 - 7,11 [英]L·弗洛里迪著;刘钢译.什么是信息哲学?世界哲学,2002(4)
 - 8 马克斯·H·布瓦索著;王寅通译.信息空间——认识组织、制度和文化的框架.上海:上海译文出版社,2000;70
 - 9 [美]埃利泽·盖斯勒著;周萍等译.科学技术测度体系.北京:科学技术文献出版社,2004;41~44
 - 10,21 [美]E·拉兹洛著;李吟波等译.决定命运的抉择.北京:三联书店,1997;2~9
 - 12 陈禹,谢康.知识经济的测度理论与方法.北京:中国人民大学出版社,1998;39
 - 13 [美]马克·波拉特著;李必祥等译.信息经济论.长沙:湖南人民出版社,1987;3
 - 14,16,17 [日]伊藤阳一.日本信息化概念与研究的历史.见:李京文等.信息化与经济发展——国际会议论文集萃.北京:社会科学文献出版社,1994;87~92
 - 15 [日]小松崎清介等.信息化的由来及其经济含义.见:李京文等.信息化与经济发展——国际会议论文集萃.北京:社会科学文献出版社,1994;92
 - 18 吕斌,李国秋.信息化的理论反思.情报科学,2005(12)
 - 19 [俄]B. II. 伊诺泽姆采夫著;安启念等译.后工业社会与可持续发展问题研究.北京:中国人民大学出版社,2004;23
 - 22,23 吕斌,李国秋.信息社会测度:信息社会研究的新焦点.中国图书馆学报,2006(1)
 - 24 ECLAC. Road Map Towards an Information Society in Latin America and the Caribbean. [2006-10-23]. <http://www.globalpolicy.org/soecon/inequal/2003/09digdiv.pdf>
 - 25 马克思,恩格斯著;中共中央马克思、恩格斯、列宁、斯大林著作编译局编译.德意志意识形态(节选本).北京:人民出版社,2003;26
- 郑建明 南京大学信息管理系教授,博士,博士生导师。通信地址:南京。邮编 210093。
- 万里鹏 南京大学信息管理系博士研究生。通信地址同上。(来稿时间:2006-11-13)