

OIS 开发实践技术

浙江省财政厅 徐蓉嘉

摘要: 本文通过 OISA 的开发实例, 简明扼要地论述了系统分析和系统总体方案设计的基本思路及其战略内涵, 并密切结合开发实践, 论述了系统分析和方案设计中的若干成功的实践。

一、前言

无论是管理信息系统(MIS), 还是办公室信息系统(OIS)的开发过程, 一般而言, 采取“自顶向下分析, 自底向上实现”的原则。然而, 经验表明, 对一个庞大系统的泛泛而谈, 不如针对系统运行机制的一个部门中的一个业务环节进行比较深入的剖析, 更加具有实用价值。作者以市政府办公室诸多信息事务处理工作中的一个相对独立的业务环节——公文档案检索管理系统(以下简称 OISA)进行剖析, 论述 OIS 开发的系统分析和系统总体方案设计的实践技术。

二、OISA 系统分析

1. 系统现状和可行性分析

市政府办公室公文档案管理的事务处理过程, 系根据中央、国务院、人大、省、市政府颁发的形形式式公文, 按综合、计划、经济、工业、农业、城乡建设、财贸金融、外经、外事、科技、政法、统战、组织人事、宣传等十六个大类及其子类, 进行分类组合成案卷的形式存档, 如: 案卷号、案卷名称、起止日期、卷内张数、保管期限等; 在每一宗案卷内首页是卷内目录, 按规格表明本案卷中所有文件的特征信息, 如: 案卷名称、案卷号、起止日期、保管日期、文件字号、起止张号等; 由此构成了以公文类别、案卷目录和卷内目录为索引的档案分类存放和检索查找的管理形式。该业务工作的主要用途和目的是满足市政府领导和机关日常的大量的随机查询, 及时提供必要的决策依据。鉴于公文信息量庞大, 手工查找往往不能满足需求而延误决策时机。因此, 在客观上产生了“公文管理计算机化”的现实需求和迫切性。根据现状的描述和分析可

见, 该项业务的原始信息事务处理在日积月累的工作中, 已形成了有章可循的操作规范; 用户有计算机化需求的迫切性; 在技术方面, 可充分利用电脑的大容量存储功能和高速运算, 满足大量公文信息的分类存储和转换存储以及快速查询检索的种种需求; 因此, OISA 的开发有比较成熟的现实基础和心理支撑环境, 成功的期望值高。

2. 数据流程图和数据字典

数据流程图是结构化系统分析的主要工具, 也是设计系统总体逻辑模型的有力工具。它不仅能够表达数据在系统内部的逻辑流向, 而且还可以表达系统的逻辑功能和数据的逻辑变换。它有四种基本符号: 外部项, 处理逻辑、数据存储和数据流符号。采用数据流程图对系统进行分析的基本方法是先用少数几个主要的处理逻辑, 高度概括地、抽象地描述整个系统的逻辑功能, 然后再分别对每一个处理逻辑扩展, 即“自顶向下”地逐层扩展, 逐层具体化, 从而形成一个系统总体逻辑模型。在画数据流程图时, 首先要确定系统的外部项, 即系统与外部环境的分界线; 数据流程图不反映执行时间的顺序, 不反映判断和控制条件, 只反映数据的流向、自然的逻辑过程和必要的逻辑数据存储; 第一张数据流程图只集中反映系统主要的逻辑功能, 根据第一张数据流程图, 对其中每一个处理逻辑, 逐层向下扩展, 每一层数据流程图中的处理逻辑一般不超过七至八个, 下一层的数据流程图中的输入和输出至少要上一层相对应。OISA 的主数据流程图如图 1 所示。

鉴于数据流程图难以表达一个系统的全部逻辑特征, 特别是关于数据的详细内容无法表达。数据字典则是对数据流程图的解释, 对图中所有的数据, 包括数据元素、数据结构、数据流、数据存储、处理逻辑、外部项全部

进行定义,它和数据流程图一起,是系统设计和程序设计的必要依据。

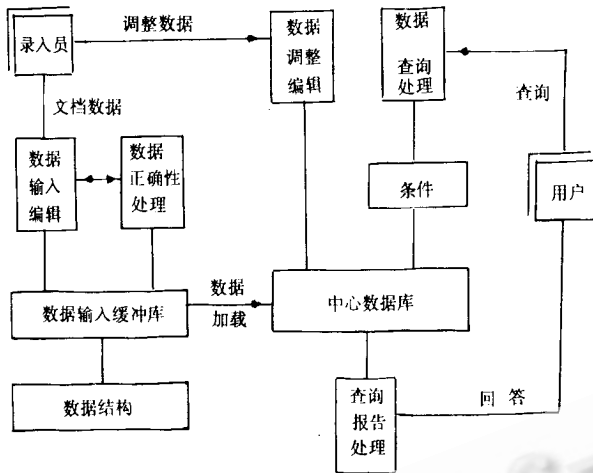


图1 OISA主数据流程图

据单位。其定义内容应包括名称、别名、取值范围和含义、长度等。OISA中卷内的目录的数据元素为“案卷号”:91--永久--36,年份&存期&卷号,长度14;“文号”:国发(1991)56,公文作者(年份)公文编号,长度30等等。数据结构用来表达数据之间的组合关系,是一种逻辑描述。其定义内容包括名称、组成部分、简单描述、与它有关的数据流或数据结构以及有关的处理逻辑。数据流是数据结构在系统内部传输的路径,它的内容包括来源、去处、组成、流量及高峰流量。数据存储的定义内容包括名称、流入/流出的数据流、包含的数据结构、数据存取要求和用户访问权限。处理逻辑描述的内容包括名称及其编号、输入和输出、处理功能简述或基本说明。外部项的定义包括名称和有关的数据流等,外部项应尽可能少,以提高系统的独立性。

3.系统目标和基本逻辑功能

系统目标:提高市政府公文档案检索管理的工作效率与质量;准确、迅速地为市府领导和政府机关提供必要或急需的决策依据和信息。

数据元数是数据的最小组成单位,即不可再分的数

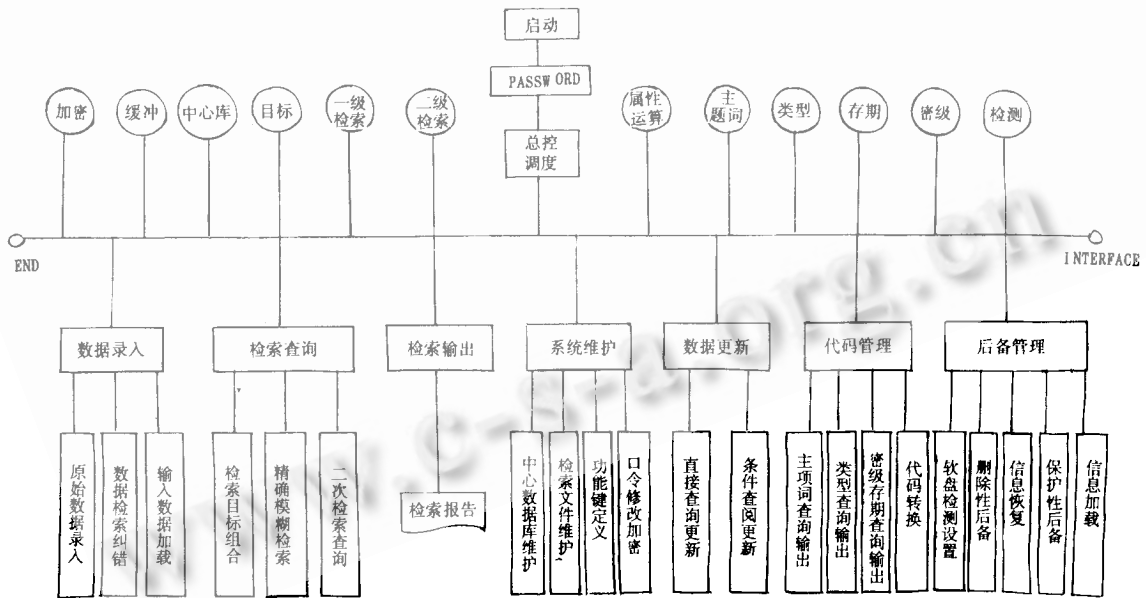


图2 OISA总体逻辑模型

系统基本逻辑功能:(1)公文档案特征信息的全屏幕输入、增加、删除、修改、更新等编辑功能。(2)公文档案特征信息的数据库存储、转换、排序、索引、存取等功能。(3)检索目标的按特征码屏幕提示输入查询信息的自生

成功能。(4)关系数据库的模糊查询和精确查询功能。(5)人机交互菜单技术和HELP功能。(6)代码体制的编辑生成、查询和转换存取功能。(7)检索报告的全屏幕浏览和打印输出及日记功能。(8)载体切换和系统维护功

能。(9)注册控制和系统加密功能等。

4.系统逻辑模型

根据对现存系统的定性、定量分析,根据系统的数据流程图和数据字典,根据用户的实际需求的计算机化的边界范围,根据新系统的目标和基本功能技术指标,从系统的总体出发,将系统划分为有机联系的各个逻辑功能模块及其各功能模块所属的子功能模块,以构成一个系统的总体逻辑设计模型。OISA 的总体逻辑模型如图 2 所示。

三.OISA 系统设计

IS 的开发设计是将一个原有的手工信息处理系统转化为一个新的计算机信息系统。这关键的一步。既包含着继承的成分,又包含着改革的重要成分,后者是 IS 开发的價值所在。因此,IS 的开发,绝对不是简单地重复原系统,而是要对原系统的各种信息进行合理的取舍,对原系统的处理过程进行大刀阔斧的加工改造和精炼提高,以塑造出一个既适合计算机运行机制又能更好地满足实际需求的新的计算机信息系统。因此,IS 总体方案设计阶段的工作是系统开发生命周期中最富有创造性的全局性工作,关系到新系统的总体质量、效果和水平。

1.数据库结构设计和存储方式

计算机处理的对象是数据和信息,数据库技术是处理数据和信息的有效工具,一个 MIS 或 OIS 凭籍数据库技术来实现,首先面临的问题是如何合理有效地组织和存储数据和信息,使系统的各项功能在数据库的基础上得到有效的实施。

表 1 公文档案数据逻辑模式

案卷号	顺序号	文号	日期	文件标题	起止页	类型	密级	主题词	备注
-----	-----	----	----	------	-----	----	----	-----	----

数据库结构—数据库: A: hzz. dbf

数据库中的数据记录个数: 32

数据库的最后更新日期: 03 / 13 / 9

采用 E-R 原理对系统的原始逻辑元素和关系进行综合分析后,将重复有的、派生的、非目标性的诸逻辑元素剔除,并进行新的合理组合,这样做一则提高了数据和信息的质量,二则减少数据和信息的冗余度,提高计算机存储空间的利用率,三则为建立一个合理高效的数据库

结构奠定扎实的基础。经过对公文档案检索管理系统原始逻辑元素和关系分析处理后,获得如下的逻辑模式,如表 1 所示。公文档案数据逻辑模式实质上已将案卷目录与卷内目录中原始逻辑元素按逻辑关系进行合理取舍,并精炼为一体。其相应的数据库物理模式如表 2 所示。

根据上述物理模式。我们就能对新系统即将处理和存储的主数据量作出进一步的定量估算如下:

$$\text{年主信息量} = 368\text{byte} \times 90 \text{ 个公文} / \text{卷} \times 200 \text{ 卷} / \text{年} = 6.6\text{Mbyte} / \text{年}$$

若以三年为一个信息处理周期,则系统的主数据库就需要占用微机的 20M 硬盘容量,加上系统的输入、输出数据库文件以及一系列基础文件和辅助文件,总的数据库信息量将占用更多的硬盘容量。每个信息处理周期采取递推的存储切换方式,按年度为单位进行逐年的存储转换和主数据库的清理调整处理。

表 2 公文档案数据库物理模式

字段	字段名	类型	宽度	小数
1	案卷号	字符型	14	
2	顺序号	字符型	4	
3	文号	字符型	30	
4	日期	字符型	11	
5	文件标题	字符型	200	
6	起止页	字符型	9	
7	类型字	符 型	26	
8	密 级	字符型	1	
9	主题词	字符型	42	
10	备 注	字符型	30	
** 总计 **			368	

2.编码方案

代码是对系统公用的数据元素赋予的一种符号标识,是系统公共信息的识别键。采用代码可以节省存储空间,提高信息处理速度;代码化是实现 MIS 和 OIS 标准化、系统化的重要基础。

代码的类型有数字码(连续码、组码、群码),文字码(单码、缩写码),数字文字组合码等。代码设计原则,要求系统性、结构简洁可记性、通用性、可扩展性和效率性。代码的设计过程如下:

确定编码对象—>确定使用目的和范围—>分析代码化对象的特征—>决定编码方法(类型和位数)—>编码方案。

代码位数的设定是代码设计中的重要环节。首先,分类数 N 的码要大于—组分类对象中的最大值 m,并考

虑一定的富裕度 a , 则 $a = N / m$, 式中 $1 < a < 2$ 。 aN 为对象数, 设代码位数为 d , 设 n 为 n 进制数, 则 $n^d > aN$, 即 $d \log n > \log(aN)$, 所以, 代码的位数估计值:

$$d > \frac{\log(aN)}{\log n} \quad (d \text{ 为整数})$$

OIS 的代码设计对象是公文档案信息中的主题词、类型、存期、密级等, 根据系统分析的要求, 对各组分类对象进行分析, 决定采用的代码类型为数字文字组合码, 代码位数的测算对公文主题词而言为 3 位, OISA 主题词编码方案的部分编码如表 3 所示。

C. 工业:

C01 工业	C02 机械	C03 冶金	C04 建材	C05 电子
C06 化工	C07 纺织	C08 轻工	C09 二轻	C10 食品
C11 医药	C12 乡镇工业	C13 街道工业	C14 节约	C15 设备
C16 材料	C17 产品	C18 企业管理		

P. 城乡建设:

P01 城市规划	P02 村镇规划	P03 环境保护	P04 风景园林	P05 建筑
P06 市政建设	P07 文物	P08 环境卫生	P09 地名	P10 房地产
P11 城市管理	P12 公共交通	P13 煤气	P14 供排水	

表 3 公文主题词代码方案示例

3. 系统硬件和软件环境

计算机系统的正确选择是信息系统建设的关键, 是系统建设中最重要基础工作之一。因为计算机信息系统的全部功能都是在计算机设备所提供的硬件、软件环境的支撑下运行的, 计算机设备选择正确, 将使信息系统的应用和发展建立在一个性能可靠、稳定、健康的环境上, 机器设备的错误选择, 必定造成难以弥补的巨大损失。计算机系统是一种密集型的高技术产品, 因此必须高度重视机型的选择, 必须以市场上计算机系统有选择地进行充分研究、权衡论证、慎重决策。

计算机系统的选择, 需要考虑的因素如下:

- 系统的先进性、存储容量、运算速度、汉字处理能力;
- 计算机硬件设备的质量可靠性、稳定性、可扩性、兼容性;
- 操作系统、数据库、程序设计语言、实用工具软件等软件支撑环境的性能;
- 图形和通讯功能;
- 供应商的信誉、经验、设备维修能力和响应时间、技术培训等支援能力;

• 计算机系统的用户量, 用户的使用历史是否表明经过考虑证实质量可靠、性能稳定。

• 计算机系统的性能 / 价格比。

• 计算机的种类繁多, 根据信息系统的功能需求、发展目标和可能的投资, 选择几种计算机系统, 进行综合比较, 将上述所需考虑的各项因素, 以相对值计分, 并采用“加权求和”的方式进行统计:

$$S = \sum_{i=1}^n A_i X_i$$

式中, S 为总分, X_i 为每项的得分, A_i 为相应项的加权因子。

根据统计结果, 提出“计算机系统选型论证报告”, 慎重决策定型。大批量购机要组织专家论证集体决策。

根据 OISA 的功能需求和系统目标以及上述各项因素的综合分析比较, 并考虑系统建设的发展, 选择了质量可靠的 32 位超级微型计算机系统 (Intel 80386cpu, 主频 25MHZ, 主存 2MRAM, 80MB 硬盘, 3.5" (1.44M)、5.25" (1.2M) 软盘驱动器, 32 位 SHADOW MEMORY, 4 个串行异步通讯端口, 2 个并行打印端口, VGA 彩色显示器, 分辨率: 字符方式 648 × 504, 图形方式 640 × 480, 显示方式: 英文 80 字 × 28 行, 中文 40 字 × 28 行, 101 键标准键盘, 24 针汉字打印机) 为硬件支撑环境; 采用 MS3.3 和 Unix 操作系统, MFOXBASE2.0 关系型数据库管理系统, C 语言以及 XE 多窗口字处理编辑工具等为系统的软件支撑环境。

4. 精确查询和模糊查询

在系统设计阶段需要对信息系统的核心技术路线进行切实的分析和探索, 以确定采用何种技术措施或手段使系统的核心功能得以可靠的实现。

OISA 的核心功能是检索查询。系统要求能向用户提供对公文档案的特征信息, 按文号、文件号、主题词、案卷号、日期、类型、密级以及这些属性的各种逻辑组合、关系组合、关系逻辑组合等多种渠道的精确查询和模糊查询, 这是本系统得以实现的亟待解决的关键技术。经过对计算机系统所提供的软件支撑环境的具体分析, 决定采用 MFOXBASE2.0 数据库管理系统命令集中的宏代换函数。宏代换 (Macro Substitution) 亦称巨集代换, 是一个多功能的变量值传递函数, 与汇编语言中的宏指令类似。基本功能是: 不论在命令或程序中的任何位置, 字符型的内存变量名前, 只要冠以宏代换符 &, 执行时就会被事先赋入该字符变量的值所代换。所以换的值, 可以

是字符串、文件名、命令行、数值、逻辑值,也可以是字符、数值、逻辑符(·OR·,·AND·,·NOT·)、关系符(=,)=,),(<, <=, <>)组合而成的表达式,这就为系统的精确查询,模糊查询提供了可行的实现途径。在理论分析的基础上,进行实际的模拟试验。以 OISA 为例,需要建立模拟数据库、输入实际的模拟数据、进行检索方法的实验。譬如,对公文档案的唯一性关键词:“发文机关&(年份)&文号”,进行精确检索试验,输入“文号=国发(1991)56”的检索目标,其中目标值采用宏代换函数进行变换;然而,在大部分情况下,检索条件往往是不确切的,譬如:要求检索一个有关美国南伊利诺州州立大学的

公文,确切的文件名全称是模糊的,于是建立一个如下的检索目标:“州立大学&文件标题”,在一次检索结果的目标命中范围内,可按需进行二次检索,以进一步提高检索准确率;利用宏代换函数的丰富功能,尚可进行各种逻辑组合、关系逻辑组合的模糊条件检索;通过实验,为系统核心功能的实现提供了可靠的关键技术的成功途径。此外,尚需对其它一些主要功能的技术实现途径和方法进行实际的技术方案设计,换言之,在系统设计阶段,系统开发者对全系统物理实现的技术途径都要有一个确定可靠的实质性的总体把握,为下一步的系统实施奠定扎实的基础。

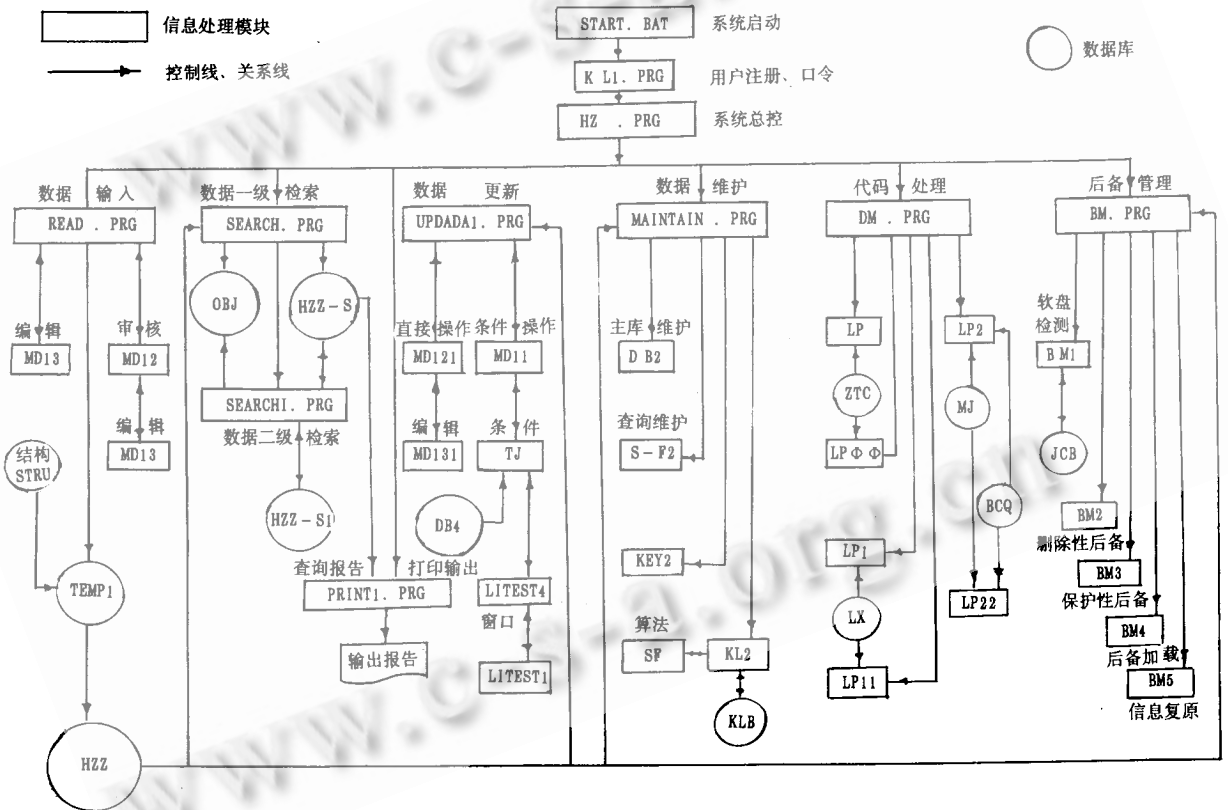


图 3 OISA 总体设计方案

5.系统总体设计方案

系统总体方案设计是建立在系统分析所提供的数据流图、数据字典、处理逻辑基本说明,系统逻辑方案的基础上,从系统的全局出发,对系统进行全面的、概括的总体结构设计。将数据流图转化为系统结构图,首先要抓

住核心数据的流向,以确定系统结构的基本形态;进而确定系统各个环节的处理功能,确定各个环节之间的联系和接口,将各个环节合理组织成一个有机整体。环节,在系统设计中称为模块,系统结构图能清楚地反映出模块的功能、模块间的联系和模块内的联系。数据流图反映

数据流的情况,结构图反映控制流的情况。OISA 的系统结构总体设计方案如图 3 所示。

从系统总体设计方案图中,可以看到构成系统的各功能模块的名称和模块间的联系和通讯关系。模块是一个独立执行某个特定功能的紧凑的子程序。模块应当是自封闭的,一个模块从系统中取走,只是使该模块的功能从系统中消失,别无其它影响,同样地,一个系统若要扩充功能,只要把实现该功能的模块加进去就可以了。评价模块化设计的质量有两个原则:一个是模块间的通信接口关系,称为模块的耦合或联结(Coupling),一个是模块本身的功能,称为模块的紧凑性或内聚性(Cohesion)。

OISA 的主要功能模块:公文档案数据输入模块 READ . PRG;一级检索模块 SEARCH . PRG;二级检索模块 SEARCHL . PRG;数据更新模块 UPDATA . PRG;数据维护模块 MAINTAIN . PRG;代码处理模块 DM . PRG;数据后备管理模块 BM.PRG。OISA 的

基本数据库文件:主数据库 HZZ . DBF;缓冲库 TEMPL . DBF;结构库 STRU . DBF;目标库 OBJ . DBF;检索结果库 HZZ-S . DBF;二级检索库 HZZ-SL .DBF;符号库 DB4 . DBF;加密库 KLB . DBF;主题词库 ZTC . DBF;类型库 LX . DBF;密级库 MJ .DBF;存期库 BCQ . DBF;检测库 JCB . DBF。

在系统设计中,需要将每一个模块的功能和接口设计要求以及每一个数据库的结构和相关存取关系,都要给予具体的技术说明,以“设计任务书”的方式交待清楚。上述系统分析和系统设计过程中产生的技术文件和资料,是程序设计人员进行系统实施的技术依据。

系统总体设计方案从系统全局把握了系统的总体功能模块框架结构关系,数据库的层次调用及其存取关系,揭示了功能模块的内部联系、模块与模块之间的相互关系及其控制流,为 SDLC 方法的系统实施和运行维护优化阶段的开发工作,奠定了坚实的行动方针。