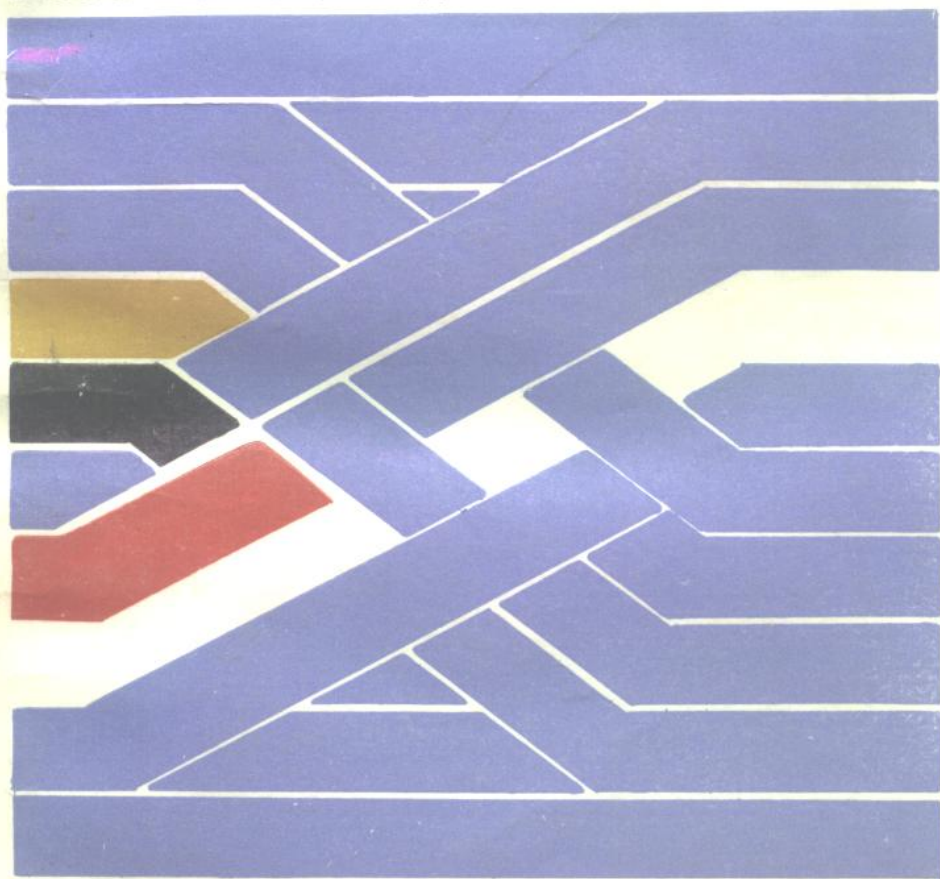


信息工程与 总体数据规划

高复先 吴曙光 等 编译



人民交通出版社

信息工程与总体数据规划

Xinxi Gongcheng yu Zongti
Shuju Guihua

高复先 吴曙光等 编译

人民交通出版社

内 容 提 要

信息工程是80年代初形成的比软件工程更为广泛的新学科，是建立基于当代数据库的计算机化的企业所需要的方法论。总体数据规划是信息工程的基础，是企业实现计算机化的奠基性工作。

本书主要是根据詹姆斯·马丁的三本著作编译而成。在介绍信息工程的一般原理和方法的基础上，重点介绍总体数据规划的一整套方法。本书为管理信息系统和计算机应用人员、企业领导和管理人员，提供了进行企业计算机应用规划的方法和信息系统建设的策略，并介绍了一些富于启发性的思想和见解。

2P65/27

信息工程与总体数据规划

高复先 吴曙光等 编译

人民交通出版社出版发行

(北京和平里东街10号)

各地新华书店经销

人民交通出版社印刷厂印刷

开本：850×1168 $\frac{1}{32}$ 印张：10.625 字数：259千

1989年9月 第1版

1989年9月 第1版 第1次印刷

印数：0001—1450册 定价：8.05元

序 言

詹姆斯·马丁 (JAMES MARTIN) 是国际知名学者, 他有丰富的著作, 仅与信息工程直接有关的著作, 已经见到的就有六本, 这些书从战略到战术, 从总体到局部, 有理论、有实际, 全面、系统地论述了80年代发展起来的, 并还在逐步完善的信息工程这一新兴学科。大连海运学院高复先等同志在研究了马丁的有关著作后, 并结合我国的具体情况和他们的实践、体会编译了本书, 无疑是一件十分有意义的工作和成果。

我国四化建设的开展, 信息将是重要的推动力和支柱, 因此信息系统的建设将是一项基础性的建设。而现代信息系统的建设和发展, 除了要确定正确的方针、策略外, 那就是要研究其方法论, 工具和环境。马丁的著作系统地总结了西方 (主要是美国) 信息系统建设的经验, 并在信息系统开发策略、方法论、工具和环境等多方面进行了系统的阐述, 其中有很多地方对我们都是很有启发的。而马丁著作的另一宝贵之处是书中不但有丰富的理论, 而且有很多活生生的实例, 因而其启发性就更大。尽管书中大量的内容谈的是美国的情况, 与我国现状有所不同, 但其原理, 思想和观点是符合和适应我国目前情况或未来发展的, 因此具有指导意义。

本书编译者对马丁著作的取舍和加工是合适的, 他们系统地就信息工程的原理和研究内容, 以及这一领域中最基础性的内容进行了介绍。目前, 我国对信息工程学科的研究还处于开始阶段, 而我国信息工程建设却又负有繁重的任务, 其中有很多工程规模甚大。因此, 开展与信息工程建设有关的研究工作是十分急需和重要的。这就要求一方面总结我们的成果, 另一方面借鉴国外的经验。在此之际出版本书, 必将产生巨大的效果。

中国科技大学研究生院 罗晓沛

88.5.26于北京中关村

目 录

前 言	1
第一章 信息工程与总体数据规划概述	3
1.1 引言	3
1.2 信息工程简介	3
1.3 总体数据规划简介	9
1.4 挑战与对策	13
第二章 数据处理的危机和转折	18
2.1 引言	18
2.2 数据处理及其发展历史	19
2.3 数据处理的危机	22
2.4 数据处理的新阶段	25
2.5 新的转折	31
2.6 信息系统宣言	32
2.7 变革过程	35
第三章 信息工程的原理、方法和工具	41
3.1 引言	41
3.2 信息工程概貌	42
3.3 信息工程方法论	49
3.4 信息工程的工具模型	63
3.5 ADT 的定义方式	66
3.6 开发工具集成环境	69
第四章 结构化方法与信息工程	78
4.1 引言	78
4.2 结构化系统分析	78
4.3 结构化设计	82

4.4	信息工程的分析设计思想	86
4.5	战略需求规划	89
4.6	信息需求规划	92
4.7	数据结构的规范化	93
4.8	规范说明语言与应用开发自动化	95
第五章	总体数据规划的组织	104
5.1	引言	104
5.2	总体数据规划的必要性	105
5.3	经验教训	106
5.4	三种类型的战略规划	108
5.5	技术策略	110
5.6	高层管理人员参加的必要性	112
5.7	自顶向下数据规划的组织	115
5.8	人员培训	118
5.9	自顶向下规划与自底向上设计的组织	119
5.10	时间规定	120
第六章	企业模型	122
6.1	引言	122
6.2	职能区域	123
6.3	业务过程	125
6.4	业务活动	131
6.5	业务活动分析	134
6.6	企业模型的建立过程	136
6.7	企业模型的特点	137
6.8	企业模型的计算机表示	138
6.9	关键成功因素 (CSF)	138
第七章	主题数据库	143
7.1	引言	143
7.2	数据库与文件	144
7.3	四类数据环境	152

7.4	数据库工程与信息工程	161
7.5	主题数据库的优越性	169
7.6	主题数据库的规划	171
7.7	主题数据库的实例	173
7.8	信息体系结构的设计	178
7.9	四类系统	179
7.10	实施的优先顺序	193
第八章	实体活动分析	197
8.1	引言	197
8.2	企业实体和实体分析	198
8.3	实体间的联系	200
8.4	结构化的实体图	204
8.5	实体的组合	208
8.6	一个实例	212
8.7	实体图的计算机表示	217
8.8	实体分析应注意的问题	217
8.9	实体与活动之间的对应	223
8.10	相关分析与分组算法	224
8.11	主题数据库的调整	227
8.12	企业的改组	228
第九章	数据分布规划	233
9.1	引言	233
9.2	分散存储数据的理由	233
9.3	集中存储数据的理由	235
9.4	多份数据副本	237
9.5	分布式数据的六种形式	239
9.6	同步数据与不同步数据	243
9.7	数据分布的初步规划	244
9.8	数据分布的定性分析	249
9.9	实例研究	253

9.10	数据分布的定量分析.....	263
9.11	分布式数据规划过程.....	270
第十章	规划与开发建议.....	280
10.1	引言.....	280
10.2	BSP 方法的目标.....	281
10.3	BSP 方法的步骤.....	283
10.4	确定业务过程和数据类.....	286
10.5	分析当前业务与系统的关系.....	286
10.6	访问高级管理人员.....	287
10.7	访问资料的整理.....	293
10.8	信息体系结构的设计和优先顺序.....	296
10.9	信息系统管理的复查.....	298
10.10	总体数据规划过程建议	299
10.11	规划工作进度控制	303
10.12	一套完整的方法	306
附录一	事件图与处理过程生成.....	309
附录二	信息工程实例.....	326
文献		331

前 言

本书讨论计算机信息系统总体规划的方法论，重点是总体数据规划。本书的主要取材是詹姆斯·马丁的三本著作：《信息工程》〔1〕、《总体数据规划方法论》〔2〕和《信息系统宣言》〔3〕。其中第一本书被认为是世界上第一本信息工程专著，第二本书是总体数据规划方法论的专著。我们以这三本书为主，少量地选取了其它材料，经编译组织而成稿。

本书的框架和取材结构，是在为研究生授课的基础上形成的，共分十章。

第一章是编者为本书写作的综述，目的是使读者尽快了解信息工程概念和总体数据规划方法的大意，因此是全书的提纲。

第二章到第四章是信息工程产生的背景和方法论的分析。这三章按三个主题展开：数据处理危机与转折；信息工程的原理、方法与工具；信息工程与结构化方法。三章材料的组织既自成单元，又相互联系，阅读时详略可由读者根据兴趣自定。

第五章到第十章比较全面深入地介绍总体数据规划的一整套方法，是本书的主体。编者根据近几年参与中大型企业计算机信息系统总体规划设计工作的实践，深感探讨先进科学的方法论的极端重要性，特别是总体数据规划的内容、方法，以及与后续开发工作的衔接等问题，更是迫切需要解决的。为此，我们较全面地翻译介绍了詹姆斯·马丁所倡导的一整套方法，供有兴趣的读者参考，从而尽快形成适合我国国情的总体规划方法论。

由于信息工程内容十分丰富，而且正处于发展过程，所以《信息工程》一书中的许多内容没有编入本书；但信息工程中处理过程设计所使用的独特方法——事件图/作用图，我们编译在附录一中。此外，《信息工程》的附录是80年代初信息工程中几

个富有启发性的实例，我们译出作为本书的附录二，以便供读者参考。

本书由大连海运学院计算机与自动控制系统信息系统研究室主任高复先主编。书中第一、二、五、六、九、十章及附录由高复先编译。第七、八章由上海机械学院系统工程研究所吴曙光编译，第三、四章（初稿）由大连海运学院研究生修璞、孙咏梅编译。

本书在编译过程中得到了中国科技大学研究生院罗晓沛教授的关注，他为本书写了序言。此外，上海机械学院、大连海运学院、大连理工大学、东北财经大学等院校的许多同行，以及交通部经济信息系统专家组的同志们也给予编者许多关心和鼓励，谨向他们表示衷心的感谢。

由于我们水平有限，对编译工作又缺乏经验，在取材、翻译和编写方面错误与不妥之处一定不少，敬请读者批评指正。

编 者

1988年 5 月

第一章 信息工程与总体 数据规划概述

1.1 引言

信息工程 (Information Engineering) 是一门新兴学科,它是在20多年的数据处理发展过程中,总结了成功的和失败的经验,进行一系列实际考察、理论分析与多种方法综合的基础上,于80年代初创立的。信息工程的目标是建立计算机化的企事业管理系统。信息工程的基本论题之一,是总体数据规划 (Strategic Data-Planning), 这是建立计算机化的企事业管理系统的奠基性工作。

我国的四个现代化进程已将国家经济信息系统及其分系统,各企业管理信息系统的规划建设工作提到议事日程上来。引进、消化、吸收信息工程这一新兴学科的方法论和有关工具,结合我们自己的实际情况,建立我国管理现代化的计算机信息系统的体系、方法论和技术工具,是当前计算机应用战线上的迫切任务。

1.2 信息工程简介

1. 信息工程的产生

信息工程的诞生,也象其它学科的出现一样,有它自己的特殊原因和动力。它是数据处理危机的必然结果。整个60年代和70年代,以美国为代表的一些计算机技术发达的国家,差不多都经历了计算机在数据处理(DP)领域应用的发展时期。开始搞成批

处理 (Batch Processing)，如工资计算，单据汇总，库存盘点等；后来逐步搞日常数据处理，如生产统计，库存控制等等。但是，当人们试图开发综合的信息服务系统，即所谓的管理信息系统 (MIS^①)；和进一步支持决策的信息系统，即所谓的决策支持系统 (DSS^②) 时，就发现原先的开发方法和工具是远不适应的。尽管70年代数据库理论与技术有了很大的发展，以结构化开发方法为主要内容的软件工程开始普遍使用，但事实表明这一时期所形成的一套方法只是在一些较小的系统上取得了成功，对中大型、复杂信息系统的开发，不仅耗资大，成功少，而且还造成许多后来难以克服的隐患。当管理上的要求越来越高，为满足用户的需求，一些老系统要修改或重建，就连一些单项应用也要加以不断的维护；特别是随着计算机设备的不断跌价，个人计算机越来越多地出现在管理人员的办公桌上，要发挥这些设备的效益，必须把它们互连起来，既满足每个管理人员的信息需要，又给高层领导提供及时的决策信息。这时，人们才吃惊地发现，分散的开发所带来的严重后果：修改原先的软件，重新组织数据，连成一个统一的大系统，所耗费的人力和资金比重新建立还要多；甚至，采取维护和修改的办法是根本行不通的。美国80年代初的统计表明，国防部每年支付的软件维护费为20亿美元，估计到80年代末要高达160亿美元；80年代初美国全国每年软件维护费耗资200亿美元。系统维护问题就象病魔似的缠住了数据处理的发展，这就是人们所说的“数据处理危机”。传统的数据处理开发方法所遭到的一些失败，也是这种危机的表现。例如，IBM^③公司为日本的两家报社搞自动化系统，总编辑在终端上如何工作的问题一直搞不清楚，使IBM公司损失200万美元；而这些无畏的开发者们不懈努力，在几年后使美国的新闻管理工作自动化，设计文档资料竟达2400页。这使人们开始怀疑，从需求分

① Management Information System 的缩写。

② Decision Support System 的缩写。

③ International Business Machine Corporation (美国) 国际商业机器公司的缩写。

析开始的传统的生命周期开发方法论，是否符合大型复杂信息系统的开发？

以詹姆斯·马丁 (James Martin) 为代表的美国学者，总结了这一时期数据处理发展的正反两方面经验，在有关数据模型理论和数据实体分析方法的基础上，再加上他发现的企业数据处理中的一个基本原理——数据类和数据之间的内在联系是相对稳定的，而对数据的处理过程和步骤则是经常变化的，于1981年出版了《信息工程》〔1〕一书，提出了信息工程的概念、原理和方法，勾画了一幅建造大型复杂信息系统所需要的一整套方法和工具的宏伟图景。第二年出版了《总体数据规划方法论》〔2〕一书，对信息工程的基础理论和奠基性工作——总体数据规划方法，从理论上到具体作法上详加阐述。经过几年的实践和深入研究，詹姆斯·马丁于80年代中期又出版了《信息系统宣言》〔3〕一书，对信息工程的理论与方法加以补充和发展，特别是关于自动化的自动化思想，关于最终用户与信息中心的关系，以及用户在应用开发中应处于恰当位置的思想，都有充分的发挥；同时加强了关于原型法(Prototyping)、第四代语言和应用开发工具的论述；最后，向与信息工程有关的各类人员——从企业高层领导和最终用户到系统分析员和系统设计员，从软件公司到计算机制造商，提出了各自的“宣言”(Manifesto)，实际上是一系列关于建设高效率高质量的复杂信息系统的经验总结。到此，可以认为信息工程作为一个学科已经形成了，用信息工程方法指导，成功地开发了越来越多的信息系统，逐渐引起了人们的注意。

2. 什么是信息工程

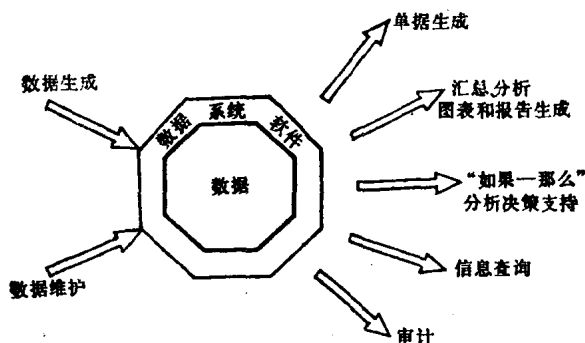
约翰·柯林斯 (John Collins) 在为世界第一本信息工程专著〔1〕所写的序言中说：“信息工程作为一个学科要比软件工程更为广泛，它包括了为建立基于当代数据库系统的计算机化企业所必需的所有相关的学科。”

从这一定义中可以看出这样三个基本点：一、信息工程的基

础是当代的数据库系统；二、信息工程的目标是建立计算机化的企业管理系统；三、信息工程的范围是广泛的，是多种技术、多种学科的综合。这自然要联系到软件工程，马丁认为，软件工程仅仅是关于计算机软件的规范说明、设计和编制程序的学科，实际上是信息工程的一个组成部分。

信息工程的基本原理和前提是：

1)数据位于现代数据处理的中心。如图1.1所示，借助于各种数据系统软件，对数据进行采集建立和维护更新。使用这些数据生成日常事务单据，例如打印发票、收据、运单和工票等。上级部门或专业人员有时要进行信息查询，对这些数据进行汇总或分析，得出一些图表和报告。为帮助管理人员进行决策，要用这些数据来回答“如果怎样，就会怎样”一类问题。审计员检查某些数据，以确信是否有问题。



2)数据是稳定的，处理是多变的。一个企业所使用的数据类很少变化。稍具体一点说，数据实体的类型是不变的，除了偶尔少量地加入几个新的实体外，变化的只是这些实体的属性值。对于一些数据项集合，我们可找到一种最好的方法来表达它们的逻辑结构，即稳定的数据模型。这种模型是企业所固有的，问题是如何把它们提取出来，设计出来。这些模型在其后的开发和长远应

用中很少变化，而且避免了破坏性的变化。在信息工程中，这些模型成为建立计算机化处理的坚实基础。虽然企业的数据模型是相对稳定的，但是应用这些数据的处理过程却是经常变化的。事实上，最好是系统分析员和最终用户可以经常地改变处理过程。只有建立了稳定的数据结构，才能使行政管理上或业务处理上的变化能被计算机信息系统所适应，这正是**面向数据的方法**所具有的灵活性，而**面向过程的方法**往往不能适应管理上的变化需要。

3) 用户必须真正参加开发工作。企业的高层领导和各级管理人员都是计算机应用系统的用户，他们都在计算机终端上存取和利用系统的数据，是最终用户 (End-User)⁶。正是他们最了解业务过程和管理上的信息需求，所以从规划到设计实施，在每一阶段上都应该有用户的参加。在总体规划阶段，有充分理由要求企业高层领导参加：首先，信息是企业的重要资源，对如何发挥信息资源作用的规划工作，高层领导当然要亲自掌握；其次，总体规划要涉及企业长远发展政策和目前的组织机构及管理过程的改革和重新调整，而只有高层领导才能决定这些重大事情。各管理层次上的业务人员对业务过程和信息需求最熟悉，单靠数据处理部门无法搞清用户的需求；要使频繁的业务变化在计算机信息处理上得到及时的反映，满足管理上的变化要求，同样是数据处理部门所不能完全胜任的。这样，用户和数据处理部门的关系应加以改变，用户要参与开发，由被动地使用系统变为积极地开发系统；数据处理部门由独立开发变为培训、组织、联合用户开发，这就是信息中心的重要职能。自然，用户参与开发工作，修改、维护应用系统，决不能象过去数据处理部门那样，使用一套老方法、老的程序设计语言（如COBOL），而应是与用户充分友好的第四代语言和一系列开发工具，提高系统从设计实施到修改维护的自动化程度。

3. 信息工程的组成部分

从上述的基本原理和前提出发，马丁阐述了一整套自顶向下

规划 (Top-Down Planning) 和自底向上设计 (Bottom-Up Design) 的方法论, 指出: 建设计算机化的企业需要该组织的每一成员都为这一共同目标进行一致的努力, 这就包括采用新方法论的总体策略, 并要求每一成员对此应有清楚的理解^[2]。几经修改, 他在《信息系统宣言》一书中提出了信息工程组成的13块构件, 如图1.2所示。这13块构件是相互联系的, 构成一个统一体——信息工程方法论的宏伟大厦。本书作为对马丁三本著作的编译, 将重点介绍三个部分: 企业模型, 实体关系分析和数据模型的建立 (即主题数据库规划), 以及数据分布规划。这三个部分是信息工程的基础。

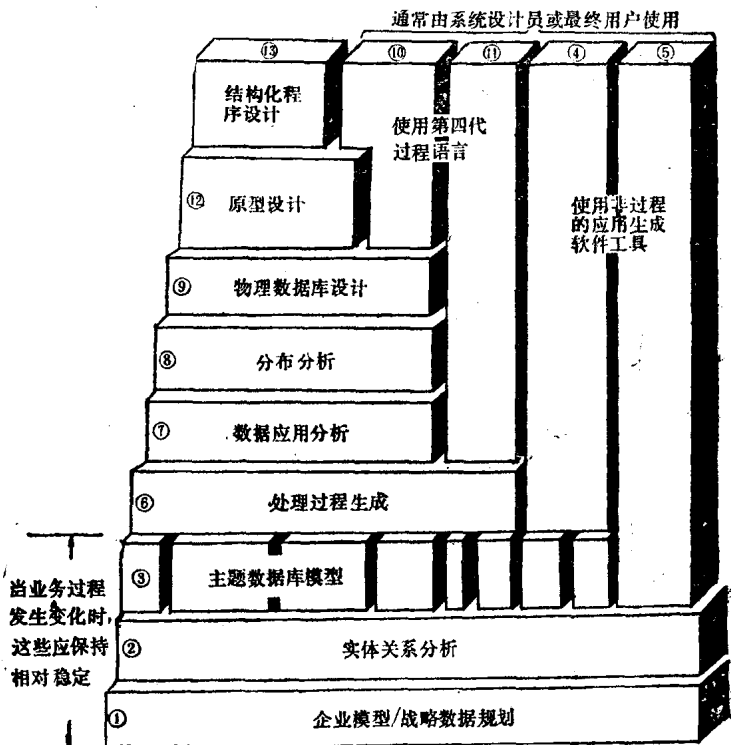


图1.2 信息工程——现代DP管理方法论

1.3 总体数据规划简介

1. 四类数据环境与总体数据规划的目标

马丁在《信息工程》和《总体数据规划方法论》中将计算机的数据环境分为四种类型，并认为清楚地了解它们之间的区别是很重要的，因为它们对不同的管理层次，包括高层管理的作用是不同的。一个高效率的企业应该具有第三类和第四类的数据库环境，以保证具有真正的数据基础。

第一类数据环境：数据文件 (Data Files)。其特征是：没有使用数据库管理系统，根据大多数的应用需要，由系统分析员和程序员分散地设计各种数据文件。其特点是简单，相对容易实现。但随着应用程序增加，数据文件数目剧增，导致很高的维护费用；一小点应用上的变化都将引起连锁反应，使修改又慢又贵，并很难进行。

第二类数据环境：应用数据库 (Application Data Bases)。使用了数据库管理系统，但没达到第三类数据库那种共享程度。分散的数据库为分散的应用而设计。实现起来比第三类数据库环境简单。象数据文件环境一样，随着应用的扩充，应用数据库也在剧增。维护费用高，有时甚至高于第一类数据库环境。还没有发挥使用数据库的主要优越性。

第三类数据库：主题数据库 (Subject Data Bases)。所建立的一些数据库与一些具体的应用有很大的独立性，数据经设计，其存储的结构与使用它的处理过程是独立的。各种业务科目的数据，如顾客数据，产品数据或人事数据，通过一些共享数据库被联系和体现出来。这种主题数据库的特点是：经过严格的数据分析，建立模型，需要花费时间，但其后的维护费用很低。最终（但不是立即）会使应用开发加快，并能使用户直接与这些数据库交互使用数据。需要改变传统的系统分析方法和整个数据处理的管理方法。如果管理不善，也会蜕变成第二类（或者有可