

电力信息化咨询项目 实施与应用分析

李 强 主 编

施明泰 赵建伟 李浩松 许中平 胡全贵 副主编

清华大学出版社
北 京

内 容 简 介

北京国网信通埃森哲信息技术有限公司依托国网信通产业集团丰富的电力行业服务经验,组织并完成国家电网公司系统 SG186 和 SG-ERP 等重大信息化工程设计、项目建设及信息技术咨询与服务。本书在总结成功实践经验的基础上,系统地讲解了电力信息化咨询项目实施与应用的基本理论以及作者的最新研究成果。

全书共 11 章,主要介绍电力信息化咨询工程的基本理论和基础知识、国内外电力信息化现状与发展趋势、我国电力信息化咨询工程技术发展特点、电力大数据技术、人工智能工程技术、电力物联网工程技术、电力内存计算技术、电力信息安全技术、电力企业 ERP 工程、工业控制与两化融合工程、电力信息等级保护工程等项目的咨询与应用分析。本书理论联系实际,应用分析针对性、系统性强。

本书可作为国家能源行业中电力信息化咨询工程技术的培训用书,也可作为企业领导、信息化与工程技术管理人员指导信息系统建设、管理与应用工作的参考资料,还可作为大学教师及科研人员教学、研究和指导信息化工程的参考用书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

电力信息化咨询项目实施与应用分析 / 李强主编. —北京:清华大学出版社, 2020.2

ISBN 978-7-302-54859-1

I. ①电… II. ①李… III. ①电力工业—企业管理咨询—企业信息化—研究 IV. ①F407.61

中国版本图书馆 CIP 数据核字(2020)第 025567 号

责任编辑:杨如林

封面设计:杨玉兰

版式设计:方加青

责任校对:胡伟民

责任印制:

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:

经 销:全国新华书店

开 本:185mm×260mm 印 张:24 字 数:588 千字

版 次:2020 年 5 月第 1 版 印 次:2020 年 5 月第 1 次印刷

定 价:79.00 元

产品编号:081022-01

编委会名单

主 编	李 强			
副主编	施明泰	赵建伟	李浩松	许中平 胡全贵
前 言	潘明惠			
第 1 章	程 瑶	魏志丰	龙长贵	刘才华
第 2 章	郭 翔	苏 斌	李 檀	杨永鑫
第 3 章	赵恩来	夏岳红	沈 磊	杨 超
第 4 章	于希永	梁 昆	韩雨彤	黄婉书
第 5 章	潘 霄	张 瑜	刘永昌	王泽宁
第 6 章	袁东州	杨雅麒	陈其祥	姜玉琳
第 7 章	高秉强	晏荣煜	刘 凯	祝榕岭
第 8 章	杨海峰	江丽娜	刘 坤	曹国强
第 9 章	李向阳	苏彦龙	杨大威	冉定国
第10章	贾伟昭	田俊丽	刘永昌	唐瑗瓊
第11章	吴 菲	潘 邈	张 冉	徐莲荫

前言

以互联网为代表的信息技术日新月异，引领了社会生产新变革，创造了人类生活新空间，拓展了国家治理新领域，加速驱动了人类社会的发展和进步，并不断改变着社会生产、经营管理和人们的日常生活方式，极大地提高了人类认识世界、改造世界的能力。信息革命则增强了人类脑力，从蒸汽时代的工业1.0，电气时代的工业2.0，自动化时代的工业3.0，到正在进入网络化和智能化时代的工业4.0，为人类带来了生产力的一次又一次质的飞跃，对国际政治、经济、文化、社会、生态、军事等领域产生了深刻影响。信息化是一个规模庞大和因素复杂的社会、经济、科学、技术相互作用的系统工程。信息化是现代人类社会发展的必然过程，没有信息化就没有当今的现代化。只要大力推进信息化，以信息化带动工业现代化和农业现代化，充分发挥后发优势，就能够实现社会生产力的跨越式发展，把工业（农业）社会推进到高度发达的信息社会。

电力行业是关系国计民生的基础行业，随着信息化和网络技术的发展，电力系统的信息化进一步快速发展。从早期的生产过程自动控制，到电力系统信息综合管理，再到“智能电网”的建设热潮，通过结合计算机技术、网络技术、自动化控制技术以及管理技术，电力系统在物理结构、物理性能、人员状况、经济管理等方面的信息采集与控制技术日趋完善。电力行业信息化发展规划在不同规划期之间、相关行业间信息化发展规划和标准的衔接程度、不同专业之间的融合程度、电力信息化项目咨询研究规范程度等方面存在明显差距，距离全面实现企业信息化还有诸多问题需要解决。同时，信息化工程管理与应用技术研究不论是在理论上，还是在实践上，远远不能适应企业信息化发展的要求。特别是需要及时解决先进的信息化技术与我国企业实际结合上的大量应用技术问题，更是当务之急。

国网信息通信产业集团有限公司是国家电网有限公司（以下简称为国家电网公司）为了进一步优化信息通信产业资源，提升集约化和专业化管理水平，更好地服务信息通信建设和智能电网发展，立足智能电网，面向社会重要领域，以推动工业化与信息化融合为使命，加强信息通信技术与智能电网的深度融合，打造的国内领先、国际一流的信息通信产业集团。北京国网信通埃森哲信息技术有限公司以国际化视野引入人才培养模式，坚持

“诚信、责任、创新、奉献”的企业文化，打造了一支学习型、创新型、开拓型的咨询和研发团队。依托国网信息通信产业集团有限公司丰富的电力行业服务经验、信息技术积累以及埃森哲全球最佳业务实践知识库，以服务国家电网公司为根本，同时，向其他电力企业、大型集团及政府机构提供信息技术咨询与服务。依托电力行业经验积累，整合集团咨询服务力量，为政府及大型企事业单位开展各类咨询服务：开展IT规划、两化融合体系贯标咨询、智慧城市顶层设计；开展业务流程优化提升、企业项目管理体系、企业标准化管理咨询、企业知识管理体系、企业员工素质提升；开展专业领域整体解决方案设计、企业架构（EA）管理咨询、新技术应用咨询、企业数据分析与预测、信息化标准体系、信息系统监理服务、企业资源管理、信息安全评估测评、“红蓝队”培养与建设、安全技术与解决方案咨询、多领域信息安全咨询等。

本书作者参考了国家电网公司总部、31个区域和省级电力公司、25个直属单位、304个地市级电力公司以及1924个县级供电公司的信息化SG186与SG-ERP重大工程取得的成功实践经验，运用国际先进信息化与ERP工程理论以及电力信息化咨询工程理论指导的大量电力信息化咨询及大数据、人工智能等最新信息技术工程实践取得的成功经验。

全书共11章：第1章 绪论；第2章 电力信息化咨询项目基本理论；第3章 电力信息化咨询工程基础知识；第4章 电力大数据工程技术项目咨询与应用分析；第5章 人工智能工程咨询项目实施与应用分析；第6章 电力信息安全工程技术咨询与应用分析；第7章 电力物联网工程技术咨询与应用分析；第8章 电力企业ERP工程咨询与应用分析；第9章 工业控制与两化融合咨询与应用分析；第10章 电力内存计算技术项目咨询与应用分析；第11章 电力信息安全等级保护咨询与应用分析。

本书的突出特点是作者结合实践经验和理论研究成果，全面分析了国内外信息化与咨询工程技术研究现状与发展趋势，针对我国信息化与咨询工程技术的发展特点，通过分析电力信息化与咨询工程技术的实际案例，提出了相应的解决方案。读者通过本书既可以学习信息化与咨询工程技术的理论和基础知识，也可以通过大量实例掌握信息化与咨询工程组织、管理和技术实现的方法。本书是一本信息化与咨询工程技术的工具书，可作为高等院校、信息化与咨询工程技术培训班的教材，也可以作为企事业单位从事电力信息化工程人员的参考用书。

衷心感谢国家电网公司、信通部、国网信息通信产业集团有限公司领导和专家的指导与帮助。衷心感谢全国各网、省公司有关领导和信通战线工程技术人员给予的大力支持，衷心感谢北京国网信通埃森哲信息技术有限公司有关工程技术人员对本书的编撰、出版所付出的辛勤劳动。

由于时间仓促，作者水平有限，书中内容难免有误或不妥之处，敬请读者批评与指正。

潘明惠

目 录

第1章 绪论	1
1.1 背景与意义	2
1.2 我国电力信息化咨询项目发展历程及展望	3
1.2.1 我国电力行业信息化发展历程	3
1.2.2 电网“十三五”信息化规划主要目标	5
1.2.3 电力信息通信产业“十三五”发展目标与任务	7
1.3 电力信息化咨询项目面临的新挑战	24
1.3.1 电力信息化咨询项目发展环境分析	24
1.3.2 电力信息化咨询的主要优势与面临的挑战	29
1.3.3 电力信息化咨询工程的主要规划方向	31
第2章 电力信息化咨询项目基本理论	33
2.1 信息化系统工程基本理论	34
2.1.1 信息与数据基本定义	34
2.1.2 信息的基本性质及分类	35
2.1.3 信息资源与信息技术	38
2.1.4 信息化的现代化基本概念	41
2.2 电力企业信息化咨询工程规划	44
2.2.1 制定企业发展目标与经营战略的基本方法	44
2.2.2 制定企业信息化战略规划的步骤	47
2.2.3 企业信息化评价及管理制度	51
2.2.4 企业信息化标准规范的设计方法	53

2.3 电力信息化咨询工程基本理论	55
2.3.1 企业信息化咨询工程概述	55
2.3.2 企业战略管理咨询基本概念	58
2.3.3 信息业务咨询框架及项目管理方法论	59
2.3.4 国际咨询项目管理十大行业趋势	64
第3章 电力信息化咨询工程基础知识	67
3.1 信息化咨询工程基本原理	68
3.1.1 信息工程的基本原理	68
3.1.2 数据管理基础标准与数据字典	70
3.1.3 信息化咨询项目实施方法论	78
3.1.4 信息系统项目组织及实施方法	80
3.2 信息化工程咨询分析模型	82
3.2.1 PEST分析模型	82
3.2.2 SWOT分析模型	85
3.2.3 波士顿矩阵模型	87
3.2.4 关键成功因素分析法	90
3.3 数据分析处理及软件工程开发方法	92
3.3.1 数据仓库的基本概念	92
3.3.2 数据仓库联机分析处理	96
3.3.3 结构化生命周期法开发方法	99
3.3.4 原型法及其主要特点	101
3.3.5 公共信息模型基础知识	103
第4章 电力大数据工程技术项目咨询与应用分析	107
4.1 大数据工程项目咨询基础知识	108
4.1.1 电力大数据应用工程基础知识	108
4.1.2 国内外大数据技术应用历程及发展趋势	111
4.1.3 电力行业大数据技术应用历程及发展趋势	114
4.1.4 大数据产业发展需求与面临的形势	116
4.2 电力大数据项目咨询及实用方法	119
4.2.1 大数据时代新思维基本方法	119
4.2.2 数据处理与数据分析方法	123
4.2.3 电力业务大数据融合应用方法	128

4.2.4	大数据技术创新发展全球能源互联网	131
4.3	电力大数据咨询项目实施与应用分析	133
4.3.1	大型电力企业信息化需求分析	133
4.3.2	大型集团企业信息化建设总体目标	136
4.3.3	电力人财物数据集约化设计	137
4.3.4	电力安全生产综合数据管理	138
第5章	人工智能工程咨询项目实施与应用分析	141
5.1	人工智能工程技术基础知识	142
5.1.1	人工智能基本概念及发展历程	142
5.1.2	新一代人工智能技术发展战略态势	144
5.1.3	我国电力企业人工智能应用工程概述	148
5.2	电力人工智能咨询工程实用方法	151
5.2.1	输变电系统人工智能工程方法	151
5.2.2	配用电人工智能工程应用方法	153
5.2.3	电网安全与新能源人工智能工程应用方法	156
5.2.4	企业经营管理人工智能应用工程方法	158
5.3	电力企业人工智能技术应用案例分析	160
5.3.1	电网智能运检人工智能技术应用	160
5.3.2	电网运行人工智能工程应用分析	166
5.3.3	企业管理人工智能工程应用分析	168
5.3.4	供用电服务人工智能工程应用分析	170
第6章	电力信息安全工程技术咨询与应用分析	174
6.1	电力信息安全工程技术基础知识	175
6.1.1	电力信息安全工程技术基本概念	175
6.1.2	信息安全工程技术发展主要趋势	177
6.1.3	我国电力信息安全工程技术发展历程	179
6.1.4	信息安全工程技术发展新特点	182
6.2	国内外信息安全事件主要案例分析	185
6.2.1	棱镜门事件	185
6.2.2	震网病毒	187
6.2.3	超级火焰	188
6.2.4	乌克兰电力系统遭受攻击事件	189

6.3	电力信息安全工程技术咨询案例分析	191
6.3.1	电力信息安全风险评估方法应用分析	191
6.3.2	电力信息安全防护工程技术主要实施方法	195
6.3.3	电力信息网络PKI-CA身份认证系统	198
6.3.4	数据存储备份与灾难恢复基本原理	201
第7章	电力物联网工程技术咨询与应用分析	205
7.1	电力物联网工程技术咨询基础知识	206
7.1.1	物联网工程技术基本概念	206
7.1.2	电力系统基本概念及主要特点	209
7.1.3	电力物联网技术基本概念	211
7.1.4	美国电网智能化与物联网的本质特征	214
7.2	智能电网应用物联网技术工程方法	217
7.2.1	智能电网具备物联网技术基本特性	217
7.2.2	特高压输电网络概念及主要特点	220
7.2.3	电力物联网技术在智能电网中的应用	222
7.3	电力物联网工程技术咨询与应用分析	225
7.3.1	微电网技术与能源互联网应用	225
7.3.2	智能变电站高可靠性保护控制系统集成技术应用	228
7.3.3	智能配用电通信架构体系技术应用	230
7.3.4	输电线路运检故障分析与防治方法	232
第8章	电力企业ERP工程咨询与应用分析	237
8.1	电力企业ERP工程咨询基础知识	238
8.1.1	ERP基本概念及发展历程	238
8.1.2	业务流程重组的基本概念	241
8.1.3	SAP ERP系统的功能特点	245
8.2	ERP项目咨询工程实施主要方法	248
8.2.1	ERP项目实施方法论阶段任务及成果	248
8.2.2	项目实施细节过程管理与控制方法	251
8.2.3	ERP项目实施障碍与失败原因分析	254
8.2.4	ERP项目实施成功的关键因素	256
8.3	国家电网辽宁省电力有限公司ERP项目咨询实施案例分析	260
8.3.1	前期准备与试点阶段实施历程分析	260

8.3.2 供电公司全面推广阶段实施历程分析	264
8.3.3 县（郊区）供电分公司推广阶段（第一批、第二批）	267
8.3.4 辽宁电力ERP全覆盖推广与划转阶段	269
8.3.5 辽宁电力ERP项目实施的特色与成果	272

第9章 工业控制与两化融合咨询与应用分析

275

9.1 工业自动控制系统基础知识	276
9.1.1 工业控制系统基本概念及安全新趋势	276
9.1.2 世界各国工业控制网络信息安全政策	277
9.1.3 我国工业控制系统信息安全存在的问题与风险	282
9.1.4 电力工业控制系统测评目的和意义	285
9.2 信息化与工业化融合基础理论	286
9.2.1 工业化与信息化互为促进	286
9.2.2 信息化与工业化融合的基本含义	287
9.2.3 信息化与工业化深度融合的目标	288
9.2.4 物联网技术发展与制造业两化融合	292
9.3 工业控制与两化融合应用案例分析	294
9.3.1 电力信息安全管理综合平台	294
9.3.2 输电线路智能安全监测系统应用案例分析	297
9.3.3 面向智能电网的电力通信网络应用分析	302

第10章 电力内存计算技术项目咨询与应用分析

306

10.1 电力内存计算技术项目咨询基础知识	307
10.1.1 内存计算的概念	307
10.1.2 内存数据库系统	309
10.1.3 内存数据库的分类	312
10.2 SAP HANA内存计算技术典型应用设计	314
10.2.1 内存计算中按列存储数据库的设计原则	315
10.2.2 内存计算技术中保存数据及并行化处理	318
10.2.3 电网企业内存计算技术成功应用参数	320
10.2.4 内存计算技术对应用程序设计的影响	329
10.3 SAP HANA内存计算技术验证及应用分析	332
10.3.1 技术验证目标与实施范围	333
10.3.2 技术验证方法及实施历程	334

10.3.3 验证系统功能及结构	336
10.3.4 SAP HANA技术验证项目主要成果	338

第11章 电力信息安全等级保护咨询与应用分析.....343

11.1 电力信息安全等级保护基础知识	344
11.1.1 电力信息安全等级保护基本概念	344
11.1.2 我国信息安全等级保护标准体系主要内容	347
11.1.3 对不同保护等级信息系统的基本保护要求	349
11.1.4 国家对电力行业信息安全等级保护的新要求	350
11.2 电力信息安全等级保护技术基础	352
11.2.1 电力信息安全等级保护纵深防御体系总体架构	352
11.2.2 电力信息安全等级保护纵深防御体系	353
11.2.3 提高安全产品测评与事件调查取证能力的方法	354
11.2.4 电力信息内外网逻辑强隔离装置	356
11.3 电力信息安全等级保护应用案例分析	357
11.3.1 电力信息系统安全等级保护实施方案	357
11.3.2 电力企业全面应用信息安全等级保护标准	360
11.3.3 加强电力信息安全等级保护管理措施	363
11.3.4 统一部署电力两级信息安全技术督查体系	364
11.3.5 健全电力智能型移动存储介质管理系统	366
11.3.6 构建电力统一管理信息系统调运体系	368

参考文献.....370

第3章

电力信息化咨询工程基础知识

信息工程的基本原理、数据管理基础标准与数据字典、信息化咨询项目实施方法论、信息化工程咨询典型分析模型、数据仓库联机分析处理（OLAP）、结构化生命周期法开发方法、原型法及其主要特点，以及公共信息模型基础知识是本章重点介绍的主要内容。



3.1 信息化咨询工程基本原理

信息工程的基本原理、数据管理基础标准与数据字典、信息化咨询项目实施方法论、信息系统项目组织及实施方法是本节介绍的主要内容。

3.1.1 信息工程的基本原理

约翰·柯林斯（John Collins）在为世界第一本信息工程专著所写的序言中说：“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：信息工程的基础是当代的数据库系统；信息工程的目标是建立计算机化的企业管理系统；信息工程的范围是广泛的，是多种技术、多种学科的综合。

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。

1. 数据位于现代数据处理系统的中心

在现代数据处理系统中，借助各种数据系统软件，对数据进行采集建立和维护更新，如图3-1所示。使用这些数据生成日常事务单据，如打印发票、收据、运单和工票等。上级部门或专业人员只要进行信息查询，对这些数据进行汇总或分析，就可得出图表和报告。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”之类的问题。数据库管理人员检查某些数据，以确信是否有问题。

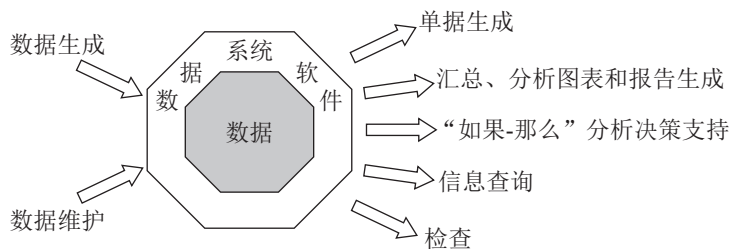


图3-1 数据位于现代数据处理系统的中心

2. 数据是稳定的，处理是多变的

一家企业所使用的数据类型很少变化。具体来说，数据实体的类型是不变的，除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，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。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被计算机信息系统适应，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. 最终用户必须真正参与开发工作

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，他们是最终用户（End User）。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，在每一阶段都应该有用户的参与。在总体规划阶段，有充分理由要求企业高层领导参与。首先，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。然后，要使频繁的业务变化在计算机信息处理中得到及时的反映，满足管理的变化要求，也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。

《信息系统宣言》的作者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 Designing）的方法论。他指出，建设计算机化的企业需要该组织的每个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，并要求每个成员对此应有清楚的理解。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3-2所示。

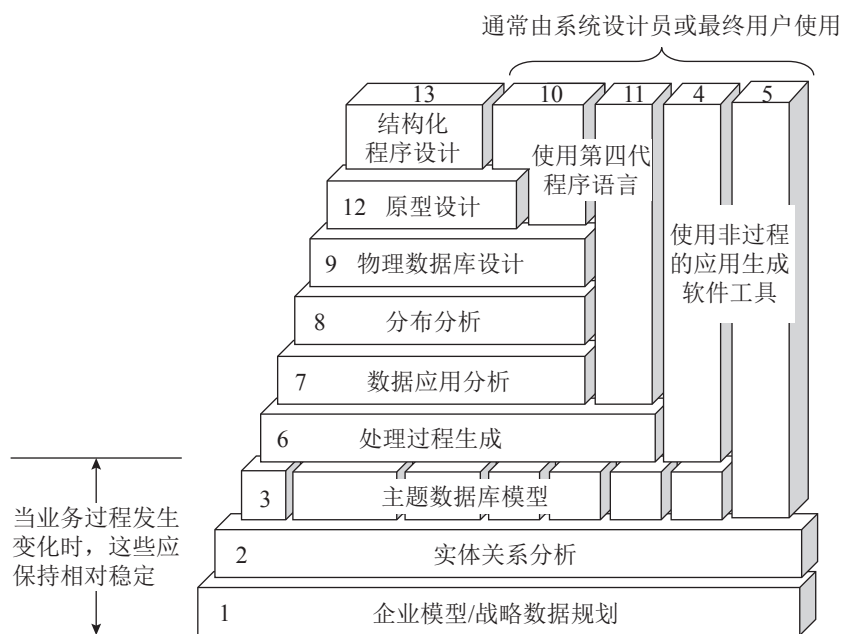


图3-2 信息工程方法论的组成

3.1.2 数据管理基础标准与数据字典

威廉·德雷尔1985年出版的专著《数据管理》（*Data Administration: A Practical Guide to Successful Data Management*）总结了数据管理标准化方面的经验。著名的论点是：没有卓有成效的数据管理，就没有成功高效的数据处理，更建立不起整个企业的计算机信息系统。他认为，早期的计算机信息系统开发缺乏数据结构的设计和管理方面的科学方法，直到20世纪80年代，才对这些问题加以认真地考虑。信息系统设计人员应该了解和掌握数据管理的标准，否则，是设计不好信息系统的。为了有效地制定和实施这些标准，威廉提出了一些重要的原则。

- （1）不能把例外当成正规。任何原则都有例外的情况，没有适用于所有情况的标准。但是，数据管理人员决不允许把例外当成正规。
- （2）管理部门必须支持并乐于帮助执行标准。如果违背了标准，管理部门必须帮助确保那些违背标准的行为得以纠正。
- （3）标准必须是从实际出发的、有生命力的、切实可行的。标准必须以共同看法为基础，标准中复杂难懂的东西越少，就越好执行，要保持标准的简明性。
- （4）标准不是绝对的，必须有某种灵活的余地。尽管有些标准必须严格遵守，但是大多数标准不应该严格到严重束缚数据设计人员灵活性的程度。
- （5）标准不应该迁就落后。标准要控制和管理当前和未来的活动，而不是恢复和重演过去的做法。在大多数情况下，今天制定的标准是几个月前数据设计所未曾采用的。

(6) 标准必须是容易执行的。要达到这一点, 必须容易发现违反标准的情况。能自动检查标准符合情况的方法愈多, 标准本身愈加有效。

(7) 标准必须加以宣传推广, 而不是靠强迫命令。即使上级主管部门完全支持数据管理标准, 也要向各级业务人员宣传这些标准。数据管理人员必须热情地向所有职员宣传这些标准, 向他们讲明为什么需要这些标准。数据管理标准要求程序员和分析员改变他们的数据设计方法。任何持久的、有意义的变化必须来自员工自己的认识。

(8) 关于标准的细节本身并不是重要的, 重要的是有某些标准。数据管理人员必须善于综合考虑和商讨所要制定的标准细节。

(9) 标准应该逐渐制定出来, 不要企图把所有的数据管理标准一次搞完。一旦标准制定出来, 就要开始执行, 但执行标准是渐进的、有节奏的。允许非数据管理人员有充足的时间对新的标准做出反应和适应。标准的实现必须是渐进过程, 而不是突变过程。

(10) 数据管理的最重要标准是一致性标准, 即数据命名、数据属性、数据设计和数据使用的一致性。

1. 数据管理基础标准

所谓“数据管理基础标准”, 是指那些决定信息系统质量的标准, 因而也是进行数据管理的最基本的标准。数据管理标准有数据元素标准、信息分类编码标准、用户视图标准、概念数据库标准和逻辑数据库标准。

1) 数据元素标准

(1) 数据元素(Data Elements)是最小的不可再分的信息单位, 是一类数据的总称。例如, 从电厂资料中的厂名“清河厂”“沈海厂”等, 可以抽象出“电厂名称”这个数据元素; 每一座电厂都有一个编号, 可以概括出“电厂编号”这个数据元素。通常, 职工档案中的“简历”“获奖情况”等不是数据元素。因为“简历”至少包括时间、地点等信息, 是可以继续分解的信息; “获奖情况”也是可以继续分解的信息。

数据处理系统中所使用的大部分数据名, 是少数数据元素的同义词或别名。这些同义词或别名的产生, 是因为缺少数据元素命名标准, 或者缺少考虑数据元素的创建和使用的规划。它们是由那些只管标记自己拥有的数据, 而不顾及别人(其他的程序员和分析员)已建立并命名数据的人所建立的重复数据。实际上, 在数据处理系统中所使用的数千个数据名, 可以减少到或合并成为数不多的“核心”数据元素。

(2) 数据元素命名有既定的标准。数据元素命名的原则是用一个简明的词组来描述一个数据元素的意义和用途。这个词组的一般结构是:

修饰词—基本词—类别词

其中, 类别词和基本词都只有一个, 修饰词可以有一个或多个; 一般类别词居后, 修饰词和基本词居前, 但按汉语或英语习惯, 顺序可以灵活调整。

类别词是数据元素命名中的一个最重要的名词, 用来识别和描述数据元素的一般用途

或功能，一般不具有行业特征，条目比较少。

基本词是类别词的最重要的修饰词，它对一大类数据对象进一步分类（反映一小类数据对象），一般具有行业特征，条目比较多。

（3）为数据元素命名，就能控制数据元素的定义准确性和总体数目。运用类别词表与基本词表分析数据元素集，可进一步标准化数据元素的命名和标识，辅助数据元素的一致性控制，并可大大压缩数据元素的总数。

（4）“数据元素命名”和“数据元素定义”是有区别的。前者是给一个数据元素起名字，这个名字一般来说是一个词组；后者是给一个数据元素下定义，即用一句话来确切界定。如果数据元素命名的词组结构和含义清晰，就不必再作定义；否则，需要作定义。例如，“员工总数”“机组编号”和“合同签订日期”等数据元素名不需要再作解释，已经很清楚了。但“发电设备容量”的定义是“统计期末发电机铭牌额定容量的总和”，“原煤耗用量”的定义是“统计期内用正平衡算法计算的发电（供热）耗用的原煤数量”。如果只有数据元素的命名，而无进一步的定义，就不能满足管理的需要；而且，数据元素的定义常常与专业技术术语解释相结合，甚至包括数据之间的关系或计算方法，这些都是在数据元素创建时应该注意的。

（5）数据元素标识有既定的标准。数据元素标识即数据元素的编码，是计算机和管理人员共同使用的标识。数据元素标识用限定长度的大写字母字符串表达，字母字符可按数据元素名称的汉语拼音抽取首音字母，也可按英文词首字母或缩写规则得出。

（6）数据元素一致性标准。数据元素命名和数据元素标识要在全企业中保持一致，或者说不允许有“同名异义”的数据元素，也不允许有“同义异名”的数据元素。这里的“名”是指数据元素的标识，“义”是指数据元素的命名或定义。

例如，会出现“同名异义”的数据元素，数据元素标识（汉语）DC和DC对应的数据元素名称可以是电厂编码或电车编码。这是因没有遵循数据元素标识规则而用两个字符来标识数据元素（过于简单）的结果。

也会出现“同义异名”的数据元素，如“职工姓名”和“员工姓名”的标识分别是ZGXM和YGXM，“单位编码”和“单位号码”的标识分别是DWBM和DWHM等，实际上是对同一个数据元素用了不同的标识。

2) 信息分类编码标准

信息分类编码（Information Classifying and Coding）是标准化的一个领域，有自身的研究对象、研究内容和研究方法。

（1）信息分类就是根据信息内容的属性或特征，将信息按一定的原则和方法进行区分和归类，并建立起一定的分类系统和排列顺序，以便管理和使用信息。信息编码就是在信息分类的基础上，将信息对象（编码对象）赋予有一定规律性的、易于计算机和人识别与处理的符号。具有分类编码意义的数据元素是最重要的一类数据元素。按照“国际/国家标准—行业标准—企业标准”的顺序原则，引用或建立企业的信息分类编码标准。

（2）一般可将信息分类编码对象划分为A、B、C三种类型。

①A类编码对象：在信息系统中不单设编码库表，代码表寓于主题数据库表之中的信息分类编码对象称为A类编码对象。这类编码对象具有一定的分类方法和编码规则，其编码表内容一般随信息的增加而逐步扩充，很难一次完成。虽然不单设编码库表，但其编码表可以从数据库表中抽取出来作为一个虚表（是数据库表的一个投影）在信息系统中使用。这类编码对象一般在具体的应用系统中使用较多。如身份证号码（国家标准）、客户编码、职工编码、设备编码（企业标准）等，都是A类编码。

②B类编码对象：在信息系统中单独设立编码库表信息分类的编码对象称为B类编码对象。这类码表内容具有相对稳定性，可以组织力量一次编制出来。这类编码表一般都较大，像一些数据库表一样，在应用系统中往往被多个模块共享，作为一些单独的库表管理是方便的。如国家行政区划编码、职称编码（国家标准）、生产统计项目编码（行业标准）、设备配件编码（企业标准）等，都是B类编码。

③C类编码对象：在应用系统中有一些编码表短小而使用频度很大的编码对象，如人的性别代码、文化程度代码和婚姻状况代码等，如果都设立编码库表，不仅系统运行时资源开销大（或内外存交换编码信息频繁），还给系统管理带来一系列的问题，所以把这类对象统一设一个编码库来管理。这类对象称为C类编码对象。

（3）信息分类编码的标准化。首先要分析识别企业生产经营所需要的信息分类编码对象，并规定将其归属为A、B、C的哪种类型。然后，对每种编码对象制定出相应的编码规则，编制代码表。

编码规则一般采用层级分类法，即按生物学的“门、纲、目、科、属、种”的分类思想，将信息对象从上到下逐层划分。为方便企业内部信息处理并满足系统外的信息交换，可以在企业内制定简化编码规则标准，同时制定与上级标准的自动换码规则或对照表。

（4）信息分类编码标准的建立过程。在企业信息系统建设之初，就应该注意信息分类编码工作。按“诺兰模型”在由控制阶段向集成阶段发展的总体数据规划期间，更应该设置专门的信息分类编码工作小组来负责信息分类编码工作。在总体数据规划过程中，通过对全企业的信息需求分析，建立起全企业的信息资源管理标准和稳定的数据模型，这是系统建设成功的关键。用户分析员和系统分析员在建立数据元素标准时，就要识别出哪些数据对象具有分类编码意义，按该对象的什么属性或特征进行分类编码，并把这些成果和资料提交给信息分类编码小组，由信息分类编码小组继续完成后续部分的工作。信息分类编码小组首先要调查这些编码对象是否有国际、国家或行业标准，以及在本组织和外单位的应用情况。信息分类编码的使用原则，一般是“自上而下”的，即有上级标准的就要使用上级标准，在没有上级标准的情况下，可自行组织力量，根据编码对象的某些属性或特征，依照信息分类编码的原理和方法，按一定的规律搞好信息分类编码工作。

在系统设计和建造阶段，信息分类编码工作要确定每个编码对象的编码规则、编码表结构和代码表，支持含有信息分类编码的数据库逻辑设计，并建成物理的数据库。

在系统运行维护阶段，要做好代码表的更新维护工作。随着应用的不断发展，信息分类编码也要做一些相应的调整和更新维护，如编码规则的调整、代码表的扩充和修改，以

便随时满足系统运行的需要。

3) 用户视图标准

用户视图 (User View) 是一些数据元素的集合, 它反映了最终用户对数据实体的看法。用户视图是数据在系统外部 (而不是内部) 的样子, 是系统的输入或输出的媒介或手段。

(1) 用户视图的分类编码。我们将用户视图分为三大类和四小类: “输入” 大类代码为1, “存储” 大类代码为2, “输出” 大类代码为3; “单证” 小类代码为1, “账册” 小类代码为2, “报表” 小类代码为3, “其他” (屏幕表单、电话记录等) 小类代码为4; 为区别不同的职能域的用户视图, 需要在编码的最前面标记职能域的代码。

(2) 用户视图组成的规范化。用户视图组成是指顺序描述其所含的数据元素或数据项, 一般格式是:

序号	数据元素/项标识	数据元素/项标识
----	----------	----------

对于用户视图的组成的表述, 不是简单地照抄现有报表的栏目, 而是要做一定的分析和规范化工作。一般来说, 存储类用户视图在表述其组成时要规范化到一范式, 标出其主关键字。

4) 概念数据库标准

概念数据库 (Conceptual DataBase) 是最终用户对数据存储的看法, 是对用户信息需求的综合概括。简单说, 概念数据就是主题数据库的概要信息。概念数据库一般用数据库名称及其内容的描述来表达。

5) 逻辑数据库标准

逻辑数据库 (Logical DataBase) 是系统分析设计人员的观点, 是对概念数据库的进一步分解和细化, 一个逻辑主题数据库由一组规范化的基本表 (Base Table) 构成。基本表是按规范化的理论与方法建立起来的数据结构, 一般要达到三范式 (3-NF)。

由概念数据库演化为逻辑数据库, 主要工作是采用数据结构规范化的理论与方法, 将每个概念数据库分解、规范化成三范式 (3-NF) 的一组基本表。企业的逻辑数据库标准是指以基本表为基本单元, 列出企业全部的逻辑数据库。

2. 数据字典

“有关数据的信息” 也称元数据 (Meta Data)。例如, “职工姓名” 是职工张大光、李小惠等姓名的抽象, 这些具体的姓名是业务数据, 而抽象的“职工姓名” 即为元数据。其实, 与“职工姓名” 相关的元数据还有“职工代码” “出生日期” “基本工资” “住址” 等。信息系统中代表“职工姓名” 的标识 (字母字符串) 的数据类型、长度等信息也属于元数据。

企业所有的元数据可以存储在一个数据库中, 这个数据库就叫元数据库 (Meta DataBase) 或中心元库 (Central Repository)。

1) 数据字典的基本内容

数据字典提供关于数据元素、元素组（记录或片段）、记录组（文件或数据库表）信息的定义和使用机制，以及这些数据实体之间的联系。可以定义其他对象，如输入格式、报表、屏幕界面、处理程序模块等。但是，所有数据实体定义是建立在数据元素定义的基础之上的。

在企业信息系统开发的不同阶段，对数据元素在数据字典中描述的详略程度是不同的。在总体规划阶段，只需要界定数据元素的标识和名称（“数据元素标准”），称为概念数据字典；在系统设计阶段，要界定数据元素的数据类型、长度等，称为逻辑数据字典；在系统建造实现阶段，要说明数据元素在具体应用系统内是什么样的以及如何被使用的（数据元素内容），称为物理数据字典。基于数据元素的记录、库表等数据实体的元数据，以及与数据实体关联的程序模块等元数据，也将分别存储在概念数据字典、逻辑数据字典和物理数据字典之中。

2) 通过数据字典进行数据管理

数据库、数据字典是数据管理人员可以使用的资源和工具，但这仅仅是整个目标的一个方面。数据管理的整个目标是规划、管理和控制全企业的信息资源。数据字典和数据库帮助数据管理人员达到这一目的，但是两者本身并不是最终目的。数据管理的任务并不是维护个别的数据库和数据字典。数据管理的任务是通过使用数据字典和设计良好的数据结构来集中管理全企业范围的信息资源。

通常，数据库管理员（DataBase Administrator，DBA）仅仅负责物理数据库（Physical DataBase）的设计、实现、安全性和维护工作。数据管理员（Data Administrator，DA）的职责在于确定每个数据库的内容和范围。数据管理首先建立数据库的逻辑模型（Logical Model），随后由DBA来实现。这类似于系统分析员和系统设计员之间的区别。尽管在数据管理之前，DBA就可以设计单个的逻辑的和物理的数据库，而数据管理则应该致力于规划和协调整个组织的所有数据库的建设。表3-1说明了DA和DBA的职责区别。

表3-1 数据管理员（DA）和数据库管理员（DBA）的职责区别

项目	数据管理员（DA）	数据库管理员（DBA）
主要职责	管理上	技术上
范围	所有数据库	特定数据库
数据设计	逻辑的	物理的
主要联系人	管理人员	程序员、系统分析员
关注问题	长期数据规划	更关心短期数据库的开发和使用
主要倾向	元数据、数据字典、数据分析、DBMS 无关性	数据、数据库、数据设计、特定的 DBMS

（注：DBMS是“数据库管理系统”的英文缩写）

分析员和程序员所管理的仅仅是在他个人控制下的系统和程序中的数据结构。这种分散的管理已经导致程序与系统中的冗余数据和不一致数据激增。

企业建立的数据字典用来存储信息资源管理基础标准的全部内容，以支持各个应用项目的开发，多个项目的互连，全系统的建设、运行、维护和数据信息的使用。数据字典之所以需要计算机化，不仅体现在上述标准的执行上，更重要的是这些标准的形成过程需要计算机的辅助。这是因为大量调查数据的整理与分析、多个小组工作结果的合成与一致性保障、多次修改的执行和新版本的发布等工作，单靠手工是无法完成的。概念层的数据字典系统主要支持总体数据规划，建立概念数据字典从数据元素定义到概念数据库定义；而逻辑层和物理层的数据字典系统主要支持应用项目的系统分析设计和实施工作中的数据管理。

数据字典的建立和使用会加强企业内各部门之间关于数据和系统知识的沟通。图3-3所示的交叉阴影部分代表了数据字典的通信含义，是共享信息的集中存放地。数据字典是一个中心元库，它所存储的资料可以被公司所有部门存取。不像传统的借助备忘录或其他纸张形式的通信手段，数据字典在更新流通方面是不受限制的。任何人只要有一台终端，并且具有数据字典方面的知识，就可以存取所需要的全部资料。

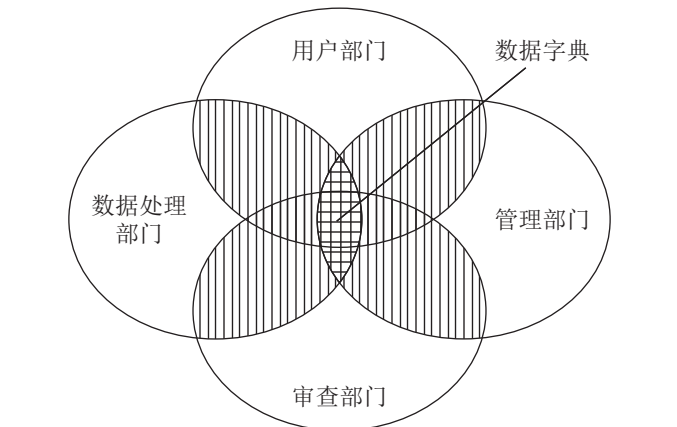


图3-3 数据字典作为沟通工具

数据字典的另一作用是作为一个术语汇编。在许多方面，数据字典类似于大词典，其中包含公司所使用的术语。数据字典可支持最终用户参与应用系统开发。图3-4所示为使用数据字典的系统开发过程。

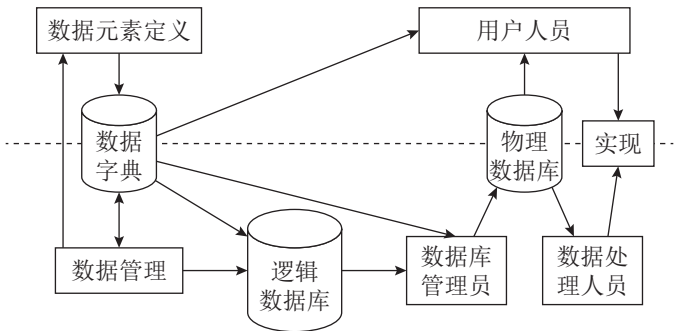


图3-4 使用数据字典的系统开发过程

(1) 在数据管理人员的帮助下, 用户定义用于进行日常业务工作的数据元素。这些数据元素是从最终用户的日常经营活动所有输入方式和报表中收集的。实际上, 这些数据元素应该构成所有数据处理系统中数据元素的大部分。其他数据元素仅是定义系统和程序控制所必需的, 这些数据元素可在其后的设计阶段由数据处理工作人员来定义。

(2) 利用数据字典中定义的数据性质和关系, 建立逻辑数据库模型。

(3) 利用逻辑数据库模型和硬件的物理约束、数据库管理系统、操作系统的知识, 由数据库管理人员建立物理数据库。

(4) 根据处理过程的复杂程度, 以及处理过程与现行系统的接口情况, 每个过程的编码或者由用户, 或者由数据处理工作人员来完成。编码过程可以使用传统的语言, 也可以使用对用户友好的非过程语言。

3) 数据字典的使用: 新系统与旧系统

参与新系统开发的数据管理工作所获得的效益, 要比把现有系统的元数据送入数据字典的收益大得多。然而, 很少有新系统的开发与现有系统是无关的。大多数情况下, 一些新系统的开发是对现有系统的若干部分进行有步骤的替换或改造。采取这种方法, 一些系统可以分阶段实现或交付使用。这样, 项目开发的时间进度和费用消耗更加容易控制, 还会使培训和对最终用户的转换影响减至最小。

图3-5反映了新系统开发生命周期中, 数据字典内容变化的一条重要原则, 即在一个项目开发的稍后阶段, 数据元素的数量可直接依赖于数据设计的质量。假如在详细设计和程序设计期间, 数据元素的数目还呈现一直增长的趋势, 这就意味着初步设计期间的数据设计工作是很不完善的; 这也可能意味着项目后面阶段数据管理工作对冗余数据元素的引进缺乏控制。图中的曲线反映了经验丰富的数据管理人员对数据定义主动的最佳控制。

图3-6反映了新系统开发期间数据元素的变化情况。在最佳情况下, 图3-6中的曲线应该是图3-5中曲线的相反情形。详细设计阶段的数据元素变化数目应该明显地少于初步设计阶段的变化数目。这反映了初步设计阶段, 数据处理和用户全体人员数据元素设计进行综合复查的情况。程序设计阶段的少量变化意味着前两个阶段的数据设计与复查工作的彻底性。

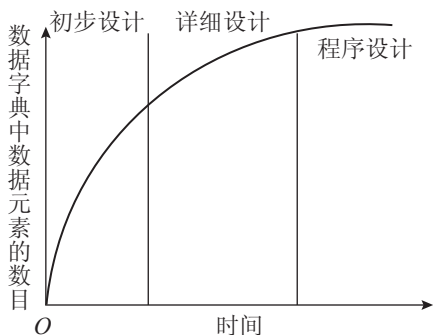


图3-5 系统开发期间数据元素的变化情况

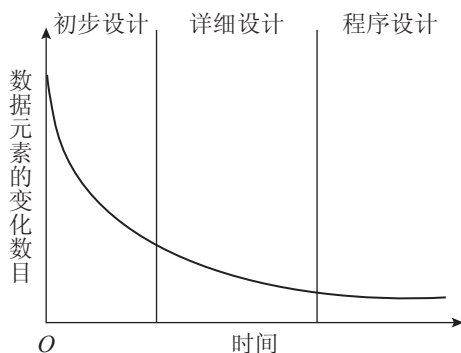


图3-6 新系统开发期间数据元素的增长情况

4) 数据字典的设置

对于具有若干部门或子公司的大公司，所建数据字典的个数取决于该公司各个领域所使用的公用数据的情况。

数据字典系统的用户可分为两类：一类为系统用户，即数据处理部门的系统开发、维护人员；另一类为最终用户，即信息系统的使用人员。虽然一个公司可能有几个分散在很大的地理区域的分部，但是该公司所有的信息资源应该由一个中心数据字典来管理。整个中心数据字典或其中一些部分可以下载装入或传输到几个远处地点，形成分布式数据字典。这就为大批的远处地点的数据字典用户提供了元数据和各种数据定义。但是，最终用户所做的任何更新或修改，应该与数据管理部门沟通。

3.1.3 信息化咨询项目实施方法论

优秀的咨询公司和ERP厂商几乎都有自己的ERP实施方法论，能够全面提供面向客户的、步骤清晰的项目实施过程指导。比如，埃森哲、SAP、IBM对ERP项目实施都有自己独特的、在实践中得到检验的ERP项目实施方法论。下面对埃森哲公司和SAP公司的ERP实施方法论做简要介绍。

1. 埃森哲ADM项目实施方法论

埃森哲是全球领先的管理咨询、信息技术及外包服务机构，凭借在各个行业领域积累的丰富经验、广泛能力以及对全球最成功企业的深入研究，致力于帮助客户成为卓越绩效的企业和政府。埃森哲公司实施ERP项目采用其ADM项目实施方法论（模型如图3-7所示），其ERP项目实施过程包括计划（Plan）、分析（Analyze）、设计（Design）、建置（Build）、测试（Test）、上线（Deploy）。

在ADM项目实施方法论的每个阶段，如同项目的工作分解结构（Work Breakdown Structure, WBS），又有下一级具体的工作包支持，细分每个阶段的任务、方法、模板和成果。比如，在埃森哲ADM项目实施方法论在计划阶段，根据客户的愿景和商业目标，明确管理和业务的需求和最能满足这些需求的软件实施范围；对客户可利用资源、实施成

本、实施进度、主要业务情景、潜在风险等因素进行评估；在此基础上选择合适的软件、模块，定义解决方案蓝图，并确定项目实施的范围。

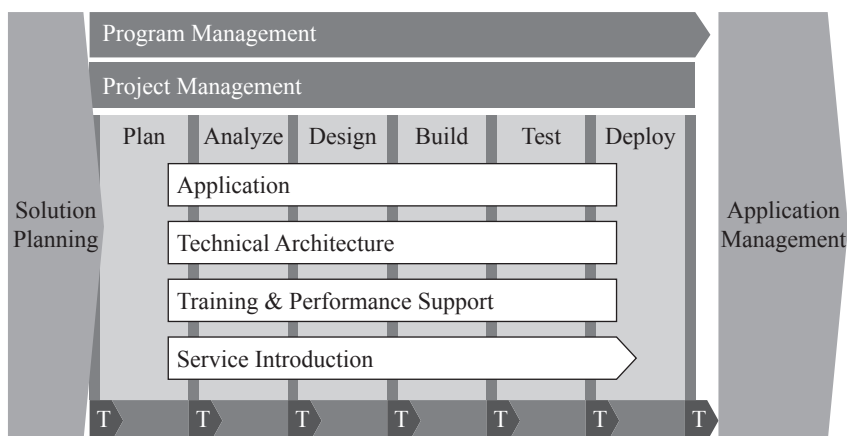


图3-7 ADM方法论模型

2. SAP公司ASAP实施方法论

ASAP是SAP公司在项目实施经验的基础上总结出来的一整套ERP项目实施方法论以及辅助工具体系，用于对项目实施时间、质量和资源的使用等方面进行有效的控制，保障项目成功实施。SAP的ASAP实施方法论包含五部分内容，分别是ASAP路线图、SAP工具包、SAP技术支持和服务、SAP培训和SAP参考模型。其中路线图标明ERP项目实施流程分为五个阶段，分别为项目准备阶段（Project Preparation Phase）、企业蓝图设计阶段（Business Blueprint Phase）、实现阶段（Realization Phase）、上线准备阶段（Final Preparation Phase）、上线与支持阶段（Go Live & Support Phase），如图3-8所示。



图3-8 ASAP方法论模型

SAP工具包包含项目实施过程中用到的所有工具，如MS-Project用于项目计划制订。

ASAP的“估算师”（Best Estimator）工具测算实施中所需的资源、成本和时间；在这个工具箱中包含了建模、实施、改进和建立技术文件等工具和模板。

SAP的技术支持和服务指SAP公司为客户在项目实施或上线后使用过程中提供的远程或现场问题解答。比如，在系统环境准备上线前，SAP提供对系统硬件、参数配置等进行“上线检查”（Go-Live Check）或“早期预警”（Early Watch），在系统上线应用支持服

务上可以提供24小时在线和现场的技术支持。

SAP的培训策略和参考模型包含了对项目小组和最终用户的培训方法，参考模型是SAP开发的以商业术语描述ERP系统所支持的标准应用功能与业务过程，帮助企业识别应用中的不同过程以及各应用之间的集成关系。

3. 两个项目实施方法论的比较

对于ERP项目实施方法论，每个咨询企业或ERP软件厂商都有不同的特点，但以结果为导向的目标管理思想是一致的，都以项目管理的基本思想为理论基础，保障按时、高质量地成功实施和交付项目。不同的实施方法论之间内容彼此相通，往往在阶段划分、过程控制和应用模板上各有特点。在实施方法论中都应有明确详细的项目规划向导、明确的阶段划分和过程及交付管理，过程中每个子项都有明确的目标、输入、时间、人员安排和成果输出；各项活动如何开展的指导性描述；各类文档都有模板和交付成果要求；质量管理、风险管理、项目变更管理等，进而项目实施团队可利用方法论定义的标准规范实施的流程，控制项目风险，在保证实施质量的前提下快速实现项目的预定实施目标，提高项目实施成功率。通过图3-9可以比较埃森哲和SAP公司ERP项目实施方法论的共性之处。

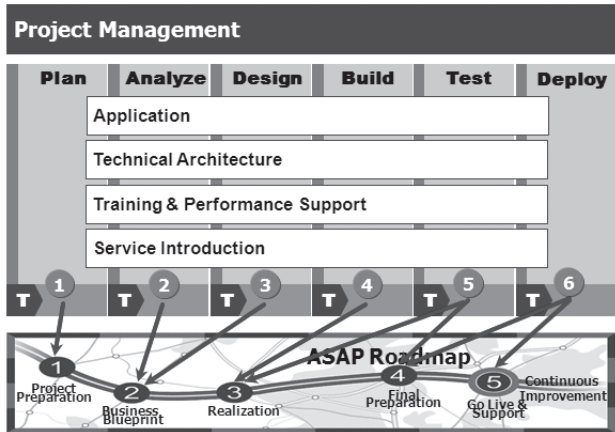


图3-9 两个项目实施方法论的比较

3.1.4 信息系统项目组织及实施方法

1. 瀑布式（Waterfall）方法

该方法在20世纪60年代由美国海军开发，用于开发复杂的军事软件。使用瀑布式模型进行项目开发，需要按照一定的步骤进行。每个步骤结束，项目小组都要进行复核或签字移交。客户满意后，才能继续进行下一步的开发。若发生修改，则开发小组需要重复项目中的各个步骤，这个过程既困难又烦琐。瀑布式方法有两个主要特点：一是文档驱动非

常正规，有大量文档式的可交付的产品；第二是高度重视规划，减少项目进展过程中的不断规划。需求相对固定和可预见的用户，适合使用瀑布式方法，如用于航天飞机的导航系统。然而，依赖于文档来描述客户和用户的需求可能会引发问题。交流方面常会出现问题，导致软件质量低劣。开发过程中编写的文档可能无可挑剔，但实际的产品可能存在缺陷、不完整或不可用。

瀑布式方法是所有软件开发方法的鼻祖。它所归纳的开发活动是其他所有方法的基本组成单元。瀑布式方法是软件开发方法的自然总结，符合人们思维的一般规律。在软件开发行业中已经和正在被不同程度地实施着，所以在中小型软件项目中实施起来会比较自然，变动较小。从某种意义上说，其他方法基本都是瀑布式方法的变种，差别在于速度、可交付产品的类型以及更高的灵活性。

2. 进化式原型化方法

进化式原型化是一种RAD方法，它偏向于项目进展过程中的系统概念开发。这种模型依赖于使用可视化原型和最终系统的模型。这些模型可能是简单的白板略图、图形图像，甚至是期望系统的完整的HTML副本。通过这种可视化方法，可以降低最终结果的不确定性，真正的开发工作并非要等到原型确定后才开始。当需求发生变化时，EP方法非常灵活，但它仍需要确定和控制。其缺点在于难以进行高效的规划。开发工作开始后，还原为编码-修复周期。对开发工作的规划可能是种挑战，因为小组并不能真正确定工作需要花多长时间。

3. 分布交付方法

分布交付方法是瀑布式方法的另一种变体。在每一个流程结束，也都需要进行签字移交。分布交付方法与瀑布式方法的区别就是将客户的整个业务需求划分为大型组件，并在各个独立的阶段交付这些组件。分布交付方法把主题和特性组作为重点，让客户能够首先开发最重要的需求。然而，这种分布交付并不要求进行详细的规划，因此在处理需求变化时仍然会不方便。因为没有进行阶段规划，开发人员需要了解所要考虑的组件和函数之间的互依赖性。使用分布式交付方法时，开发小组必须通过构建字节块来降低风险，同时要控制和监视变化。

4. RUP方法

RUP是一种过程，它使用一套标准的工具、模板和可交付产品提供了一种严格有序的开发方法。由厂商Rational Software所拥有，是RUP与前面介绍的其他方法之间的最大不同之处。这种标准化方法对于需要在公司内部使用某种通用的语言或者工具的大型组织很有吸引力。RUP使用统一建模语言（UML）来表达需求、架构和设计。Rational Software公司首次开发了UML，现在由标准组织OMG（Object Management Group）维护。RUP是另一种迭代式开发方法，重点关注降低项目风险。

RUP对需要维护外部交流以及开发小组内部交流标准的组织是极具价值的。其缺点在于需要更多的文档和客户必须对UML有所了解。此外,这种方法归一个软件厂商所有。开发人员开始更多地结合使用RUP和EP,不足是EP必须按已有的指导原则进行开发工作,但EP在开发速度和构建流程方面显然优于RUP。

5. MSF方法

微软解决方案框架(Microsoft Solutions Framework, MSF)使用模型来解释开发过程,从某种意义上讲,它并非一种方法。每个使用MSF进行开发的小组都将根据它们的需要对其进行修改。与RUP不同,MSF只能在某个范围内使用,因为MSF并没有阐释如何完成真正的开发工作。然而,这种灵活性可能成为MSF积极的一面,可以按照现有的开发风格来定制其实现模型。另一方面,它将导致框架的用户艰难地想象核心开发工作是如何被管理的。

6. 敏捷方法

敏捷方法在两个方面解决了传统的重量级软件开发难以解决的问题:一是需求的变更;二是软件开发中软件设计与软件编程实现分离。为此,敏捷软件开发方法都采用迭代式开发过程,强调软件开发过程的自适应性。因为低成本和灵活高效,中小型软件企业非常青睐敏捷软件开发方法,特别适合在要求高、工期短的项目中进行推广。极限编程(Extreme Programming, EP)因为对测试的极度重视成为近年来最为流行的敏捷软件开发方法。在国际上还成立了对应的极限编程组织,每两年召开一次敏捷软件工程及极限编程国际会议,由工业界和学术界双方共同探讨解决极限编程及其他敏捷软件开发过程中的问题,分享成功的经验。

3.2 信息化工程咨询分析模型

信息化工程咨询典型分析模型、PEST分析模型、SWOT分析模型、波士顿矩阵模型、KSF-关键成功因素分析法是本节介绍的主要内容。

3.2.1 PEST分析模型

PEST分析是战略咨询顾问用来帮助企业检阅其外部宏观环境的一种方法。宏观环境又称一般环境,是指影响一切行业和各种企业的各种宏观力量。对宏观环境因素作分析时,不同行业和企业根据自身特点和经营需要,分析的具体内容会有差异,但一般都应对政治(Political)、经济(Economic)、社会(Society)和技术(Technological)这四大类影响

企业的主要外部环境因素进行分析。简单而言，称之为PEST分析法，如图3-10所示。表3-2所示为一个典型的PEST分析。

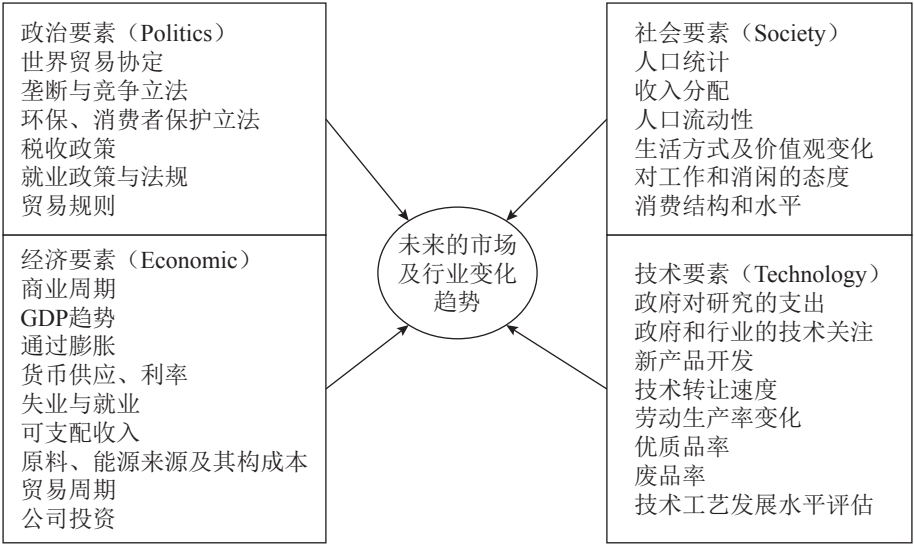


图3-10 典型的PEST分析

表3-2 一个典型的PEST分析

政治（包括法律）	经济	社会	技术
环保制度	经济增长	收入分布	政府研究开支
税收政策	利率与货币政策	人口统计、人口增长率与年龄分布	产业技术关注
国际贸易章程与限制	政府开支	劳动力与社会流动性	新型发明与技术发展
合同执行法 消费者保护法	失业政策	生活方式变革	技术转让率
雇用法律	征税	职业与休闲态度、企业家精神	技术更新速度与生命周期
政府组织/态度	汇率	教育	能源利用与成本
竞争规则	通货膨胀率	潮流与风尚	信息技术变革
政治稳定性	商业周期的阶段	健康意识、社会福利及安全感	互联网的变革
安全规定	消费者信心	生活条件	移动技术变革

1. PEST分析的内容

1) 政治法律因素

政治法律因素（Political Factors）包括一个国家的社会制度，执政党的性质，政府的方针、政策、法令等。不同的国家有不同的社会性质，不同的社会制度对组织活动有不同的限制和要求。即使社会制度不变的同一国家，在不同时期，由于执政党的不同，其政府的方针特点、政策倾向对组织活动的态度和影响也是不断变化的。

重要的政治法律变量包括执政党性质，政治体制，经济体制，政府的管制，税法的改变，各种政治行动委员会，专利数量，专程法的修改，环境保护法，产业政策，投资政

策,国防开支水平,政府补贴水平,反垄断法规,与重要大国的关系,地区关系,对政府进行抗议的活动的数量、严重性及地点,民众参与政治的行为,政局稳定状况,各政治利益集团。

2) 经济因素

经济因素(Economic Factors)主要包括宏观和微观两个方面的内容。宏观经济因素主要指一个国家的人口数量及其增长趋势,国民收入、国民生产总值及其变化情况以及通过这些指标能够反映的国民经济发展水平和发展速度。微观经济因素主要指企业所在地区或所服务地区的消费者的收入水平、消费偏好、储蓄情况、就业程度等。这些因素决定着企业目前及未来的市场大小。

重要关键经济变量包括:GDP及其增长率,国家向工业经济转变,贷款的可得性,可支配收入水平,居民消费(储蓄)倾向,利率,通货膨胀率,规模经济,政府预算赤字,消费模式,失业趋势,劳动生产率水平,汇率,证券市场状况,外国经济状况,进出口因素,不同地区和消费群体间的收入差别,价格波动,货币与财政政策。

3) 社会文化因素

社会文化因素(Sociocultural Factors)包括一个国家或地区的居民教育程度和文化水平、宗教信仰、风俗习惯、审美观点、价值观念等。文化水平会影响居民的需求层次,宗教信仰和风俗习惯会禁止或抵制某些活动的进行,价值观念会影响居民对组织目标、组织活动以及组织存在本身的认可与否,审美观点则会影响人们对组织活动内容、活动方式以及活动成果的态度。

关键的社会文化因素包括妇女生育率,人口结构比例,性别比例,特殊利益集团数量,结婚率、离婚率,人口出生、死亡率,人口移进移出率,社会保障计划,人口预期寿命,人均收入,生活方式,平均可支配收入,对政府的信任度,对政府的态度,对工作的态度,购买习惯,对道德的关切,储蓄倾向,性别角色,投资倾向,种族平等状况,节育措施状况,平均教育状况,对退休的态度,对质量的态度,对闲暇的态度;对服务的态度,对老外的态度,污染控制,对能源的节约,社会活动项目,社会责任,对职业的态度,对权威的态度,城市、城镇和农村的人口变化,宗教信仰状况。

4) 技术因素

技术因素(Technological Factors)除了要考察与企业所处领域的活动直接相关的技术手段的发展变化外,还应及时了解国家对科技开发的投资和支持重点、该领域技术发展动态和研究开发费用总额、技术转移和技术商品化速度、专利及其保护情况等。

2. PEST分析的应用

PEST分析相对简单,并可通过头脑风暴来完成。PEST分析的运用领域有公司战略规划、市场规划、产品经营发展、研究报告撰写。

有时会用到PEST分析的扩展变形形式,如SLEPT分析、STEEPLE分析。STEEPLE是以下因素英文单词的缩写,社会/人口(Social/Demographic)、技术(Technological)、

经济（Economic）、环境/自然（Environmental/Naturals）、政治（Political）、法律（Legal）、道德（Ethical）。此外，地理因素（Geographical Factors）有时也可能会有显著影响。

3.2.2 SWOT分析模型

SWOT分析模型（SWOT Analysis）即态势分析法，也称TOWS分析法、道斯矩阵，分析内容包括：优势（Strengths）、劣势（Weaknesses）、机会（Opportunities）和威胁（Threats）。因此，SWOT分析实际上是对企业内外部条件各方面内容进行综合和概括，进而分析组织的优劣势、面临的机会和威胁的一种方法。通过SWOT分析，可以帮助企业把资源和行动聚集在自己的强项和有最多机会的地方，并让企业的战略变得明朗。

1. 优劣势分析

优劣势分析主要着眼于企业自身的实力及其与竞争对手的比较，而机会和威胁分析将注意力放在外部环境的变化及对企业的可能影响上。在分析时，应把所有的内部因素（即优劣势）集中在一起，然后用外部的力量来对这些因素进行评估。

在适应性分析过程中，企业高层管理人员应在确定内外部各种变量的基础上，采用杠杆效应、抑制性、脆弱性和问题性四个基本概念进行这一模式的分析。

1) 杠杆效应（优势+机会）

杠杆效应产生于内部优势与外部机会相互一致和适应时。在这种情形下，企业可以用自身内部优势撬起外部机会，使机会与优势充分结合。然而，机会往往是稍瞬即逝的，因此企业必须敏锐地捕捉机会，把握时机，以寻求更大的发展。

2) 抑制性（劣势+机会）

抑制性意味着妨碍、阻止、影响与控制。当环境提供的机会与企业内部资源优势不相适合，或者不能相互重叠时，企业的优势再大也将得不到发挥。在这种情形下，企业就需要提供和追加某种资源，以促进内部资源劣势向优势方面转化，从而迎合或适应外部机会。

3) 脆弱性（优势+威胁）

脆弱性意味着优势的强度或程度的降低、减少。当环境状况对公司优势构成威胁时，优势得不到充分发挥，出现优势不优的脆弱局面。在这种情形下，企业必须克服威胁，以发挥优势。

4) 问题性（劣势+威胁）

当企业内部劣势与企业外部威胁相遇时，企业就面临着严峻挑战，如果处理不当，可能直接威胁到企业的生存。

2. SWOT分析步骤

（1）确认当前的战略是什么？

- (2) 确认企业外部环境的变化（五力分析模型或者PEST）。
- (3) 根据企业资源组合情况，确认企业的关键能力和关键限制。
- (4) 按照通用矩阵或类似的方式打分评价。

把识别出的所有优势分成两组，分组的时候以两个原则为基础：它们是与行业中潜在的机会有关，还是与潜在的威胁有关。用同样的办法把所有的劣势分成两组，一组与机会有关，另一组与威胁有关。

(5) 将结果在SWOT分析图上定位或者用SWOT分析表将优势和劣势按机会和威胁分别填入表格，如图3-11和表3-3所示。

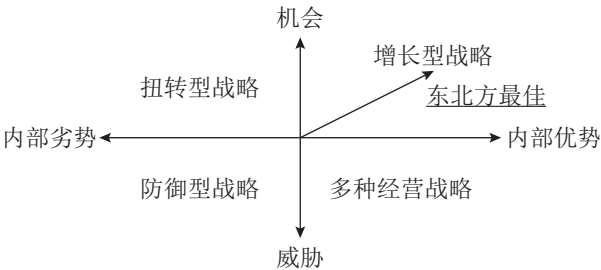


图3-11 SWOT分析图

表3-3 SWOT分析表

内部因素			
外部因素	2 利用这些	3 改进这些	机会
	监视这些	1 消除这些	威胁
优势		劣势	

3. 成功应用SWOT分析法的简单规则

规则如下。

- (1) 进行SWOT分析的时候必须对公司的优势与劣势有客观的认识。
- (2) 进行SWOT分析的时候必须区分公司的现状与前景。
- (3) 进行SWOT分析的时候必须考虑全面。
- (4) 进行SWOT分析的时候必须与竞争对手进行比较，比如优于或是劣于竞争对手。
- (5) 保持SWOT分析法的简洁化，避免复杂化与过度分析。
- (6) SWOT分析法因人而异。

一旦使用SWOT分析法决定了关键问题，也就确定了市场营销的目标。SWOT分析法可与PEST分析法、五力分析模型等工具一起使用。市场营销人士热衷于SWOT分析法是因为它的易学性与易用性。运用SWOT分析法的时候，要将不同的要素列入相关的表格中，很容易操作。

3.2.3 波士顿矩阵模型

波士顿矩阵（BCG Matrix）又称为市场增长率—相对市场份额矩阵、波士顿咨询集团法、四象限分析法、产品系列结构管理法等。波士顿矩阵认为，一般决定产品结构的基本因素有两个，即市场引力与企业实力。市场引力包括整个市场的销售量（额）增长率、竞争对手强弱及利润高低等。其中最主要的是反映市场引力的综合指标——销售增长率，这是决定企业产品结构是否合理的外在因素。

企业实力包括市场占有率，技术、设备、资金的利用能力等，其中市场占有率是决定企业产品结构的内在要素，它直接显示出企业竞争实力。销售增长率与市场占有率既相互影响，又互为条件。市场引力大，市场占有率高，可以显示产品发展的良好前景，企业也具备相应的适应能力，实力较强；如果仅仅是市场引力大，而没有相应的高市场占有率，则说明企业尚无足够实力，则该种产品也无法顺利发展。相反，企业实力强，而市场引力小的产品也预示了该产品的市场前景不佳。

通过以上两个因素相互作用，会出现四种不同性质的产品类型，形成不同的产品发展前景：①销售增长率和市场占有率“双高”的产品群（明星类产品）；②销售增长率和市场占有率“双低”的产品群（瘦狗类产品）；③销售增长率高、市场占有率低的产品群（问题类产品）；④销售增长率低、市场占有率高的产品群（现金牛类产品）。

1. 基本原理

本法将企业所有产品从销售增长率和市场占有率角度进行再组合。在坐标图上，以纵轴表示企业销售增长率，横轴表示市场占有率，各以10%和20%作为区分高、低的中点，将坐标图划分为四个象限，依次为“明星类产品（★）”“问题类产品（？）”“现金牛类产品（¥）”“瘦狗类产品（×）”。其目的在于通过对产品所处不同象限的划分，使企业采取不同决策，以保证其不断地淘汰无发展前景的产品，保持“问号”“明星”“现金牛”产品的合理组合，实现产品及资源分配结构的良性循环。

2. 基本步骤

1) 核算企业各种产品的销售增长率和市场占有率

销售增长率可以用本企业的产品销售额或销售量增长率。时间可以是一年或是三年以至更长时间。市场占有率，可以用相对市场占有率或绝对市场占有率，但是用最新资料。

基本计算公式为：

本企业某种产品绝对市场占有率=该产品本企业销售量/该产品市场销售总量

本企业某种产品相对市场占有率=该产品本企业市场占有率/该产品市场占有率最大者
(或特定的竞争对手)的市场占有率

2) 绘制四象限图

以10%的销售增长率和20%的市场占有率为高低标准分界线，将坐标图划分为四个象限；然后把企业全部产品按其销售增长率和市场占有率的大小，在坐标图上标出其相应位置（圆心）。定位后，按每种产品当年销售额的多少，绘成面积不等的圆圈，顺序标上不同的数字代号以示区别。定位的结果将产品划分为四种类型，如图3-12所示。

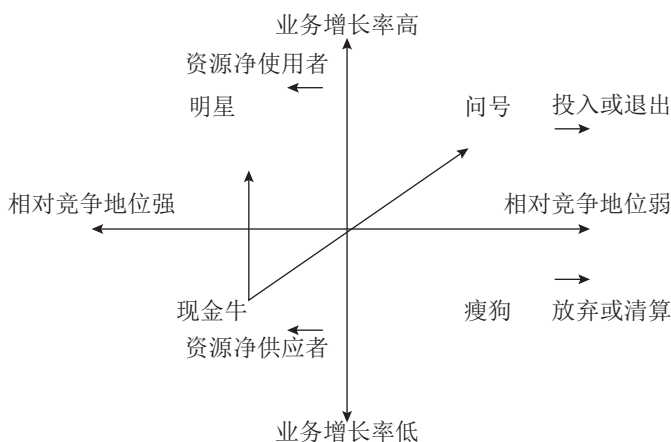


图3-12 波士顿矩阵分析模型

波士顿矩阵对于企业产品所处的四个象限具有不同的定义和相应的战略对策。

(1) 明星产品 (Stars)。它是指处于高增长率、高市场占有率象限内的产品群，这类产品可能成为企业的现金牛产品，需要加大投资以支持其迅速发展。采用的发展战略是：积极扩大经济规模和市场机会，以长远利益为目标，提高市场占有率，加强竞争地位。发展战略以及明星产品的管理与组织最好采用事业部形式，由对生产技术和销售两方面都很内行的经营者负责。

(2) 现金牛产品 (Cash Cow)，又称厚利产品。它是指处于低增长率、高市场占有率象限内的产品群，已进入成熟期。其财务特点是销售量大，产品利润率高，负债比率低，可以为企业的资金，而且由于增长率低，也无须增大投资。因而成为企业回收资金，支持其他产品，尤其明星产品投资的后盾。采用的战略是：把设备投资和其他投资尽量压缩；采用榨油式方法，争取在短时期内获取更多利润，为其他产品提供资金。对于这一象限内的销售增长率仍有所增长的产品，应进一步进行市场细分，维持现存市场增长率或延缓其下降速度。对于现金牛产品，适合于用事业部制进行管理，其经营者最好是市场营销型人物。

(3) 问题产品 (Question Marks)。它是处于高增长率、低市场占有率象限内的产品群。前者说明市场机会大，前景好，而后者则说明在市场营销上存在问题。其财务特点是利润率较低，所需资金不足，负债比率高。例如，在产品生命周期中处于引进期、因种种

原因未能开拓市场局面的新产品属于问题产品。对问题产品应采取选择性投资战略。因此,对问题产品的改进与扶持方案一般均列入企业长期计划中。对问题产品的管理组织,最好是采取智囊团或项目组织等形式,选拔有规划能力,敢于冒风险、有才干的人负责。

(4) 瘦狗产品(Dogs),也称衰退类产品。它是处在低增长率、低市场占有率象限内的产品群。其财务特点是利润率低,处于保本或亏损状态,负债比率高,无法为企业带来收益。对这类产品应采用撤退战略:首先应减少批量,逐渐撤退,对那些销售增长率和市场占有率均极低的产品应立即淘汰;其次是将剩余资源向其他产品转移;第三是整顿产品系列,最好将瘦狗产品与其他事业部合并,统一管理。

3. 基本应用法则

按照波士顿矩阵的原理,产品市场占有率越高,创造利润的能力越大;另一方面,销售增长率越高,为了维持其增长及扩大市场占有率所需的资金亦越多。这样可以使企业的产品结构实现互相支持,资金良性循环的局面。按照产品在象限内的位置及移动趋势的划分,形成了波士顿矩阵的基本应用法则。

第一法则:成功的月牙环。在企业所从事的事业领域内,各种产品的分布若显示月牙环形,这是成功企业的象征,因为盈利大的产品不止一个,而且这些产品的销售收入都比较大,还有不少明星产品。问题产品和瘦狗产品的销售量都很少。若产品结构显示为散乱分布,说明其事业内的产品结构未规划好,企业业绩必然较差。这时就应区别不同产品,采取不同策略。

第二法则:黑球失败法则。如果在第三象限内一个产品都没有,或者即使有,其销售收入也几乎近于零,可用一个大黑球表示。该种状况显示企业没有任何盈利大的产品,说明应当对现有产品结构进行撤退、缩小的战略调整,考虑向其他事业渗透,开发新的事业。

第三法则:东北方向大吉。一个企业的产品在四个象限中的分布越是集中于东北方向,则显示该企业的产品结构中明星产品越多,越有发展潜力;相反,产品的分布越是集中在西南角,说明瘦狗类产品数量越大,说明该企业产品结构衰退,经营不成功。

第四法则:踊跃移动速度法则。从每个产品的发展过程及趋势看,产品的销售增长率越高,为维持其持续增长所需资金量也相对越高;而市场占有率越大,创造利润的能力也越大,持续时间也相对长一些。按正常趋势,问题产品经明星产品最后进入现金牛产品阶段,标志了该产品从纯资金耗费到为企业提供效益的发展过程,但是这一趋势移动速度的快慢也影响到其所能提供的收益的大小。

如果某一产品从问题产品(包括从瘦狗产品)变成现金牛产品的移动速度太快,说明其在高投资与高利润率的明星区域的时间很短,因此对企业提供利润的可能性及持续时间都不会太长,总的贡献也不会大;但是相反,如果产品发展速度太慢,在某一象限内停留时间过长,则该产品也会很快被淘汰。

这种方法假定一个组织有两个以上的经营单位组成,每个单位产品又有明显的差异,

并具有不同的细分市场。在拟定每个产品发展战略时，主要考虑它的相对竞争地位（市场占有率）和业务增长率。以前者为横坐标，后者为纵坐标，然后分为四个象限，各经营单位的产品按其市场占有率和业务增长率高低填入相应的位置。

在本方法的应用中，企业经营者的任务是通过四象限法的分析，掌握产品结构的现状及预测未来市场的变化，进而有效地、合理地分配企业经营资源。在产品结构调整中，企业的经营者不是在产品到了“瘦狗”阶段才考虑如何撤退，而应在“现金牛”阶段时就考虑如何使产品造成的损失最小而收益最大。

3.2.4 关键成功因素分析法

关键成功因素法（Key Success Factors, KSF），是以关键因素为依据来确定系统信息需求的一种MIS总体规划的方法。在现行系统中，总存在多个变量影响系统目标的实现，其中若干个因素是关键的和主要的（即成功变量）。通过对关键成功因素的识别，找出实现目标所需的关键信息集合，从而确定系统开发的优先次序。

关键成功因素指的是对企业成功起关键作用的因素。关键成功因素法就是通过分析找出使得企业成功的关键因素，然后再围绕这些关键因素来确定系统的需求，并进行规划。

1. 关键因素

关键成功因素的重要性置于企业其他所有目标、策略和目的之上，寻求管理决策阶层所需的信息层级，并指出管理者应特别注意的范围。若能掌握少数几项重要因素（一般关键成功因素有5~9个），便能确保相当的竞争力，它是一组能力的组合。如果企业想要持续成长，就必须对这些少数的关键领域加以管理，否则将无法达到预期的目标。即使同一个产业中的个别企业会存在不同的关键成功因素，关键成功因素有4个主要的来源。

（1）个别产业的结构：不同产业因产业本身特质及结构不同，而有不同的关键成功因素，此因素由产业本身的经营特性决定，该产业内的每一公司都必须注意这些因素。

（2）竞争策略、产业中的地位及地理位置：企业的产业地位是由过去的历史与现在的竞争策略决定，在产业中每个公司因其竞争地位的不同，而关键成功因素也会有所不同。对于由一两家大公司主导的产业而言，领导厂商的行动常为产业内小公司带来重大的问题，所以对小公司而言，大公司竞争者的策略，可能就是其生存的竞争的关键成功因素。

（3）环境因素：外在因素（总体环境）的变动都会影响每个公司的关键成功因素。如在地震需求波动大时，存货控制可能就会被高阶主管视为关键成功因素之一。

（4）暂时因素：大部分是由组织内特殊的理由而来，这些是在某一特定时期对组织的成功产生重大影响的活动领域。

2. 确认方法

(1) 环境分析法：包括将要影响或正在影响产业或企业绩效的政治、经济、社会等外在环境的力量，换句话说，即重视外在环境的未来变化，比公司或产业的总体变化更重要，唯实际应用到产业或公司业务会产生困难。

(2) 产业结构分析法：应用五力分析模型，作为此项分析的基础。此架构由五个要素构成。要素和要素间关系的评估可提供分析者客观的数据，以确认及检验产业的关键成功因素。产业结构分析的另一个优点是此架构提供一个很完整的分类，另一项优点就是以图形的方式找出产业结构要素及要素间的主要关系。

(3) 产业/企业专家法：向产业专家、企业专家或具有知识与经验的专家请教，除可获得专家累积的智慧外，还可获得客观数据中无法获得的信息，唯因缺乏客观的数据导致实证或验证上的困难。

(4) 竞争分析法：分析公司在产业中应该如何竞争，以了解公司面临的竞争环境和态势，研究焦点的集中可以提供更详细的资料，且深度的分析能够有更好的验证性，但其发展受到特定的限制。

(5) 产业领导厂商分析法：经由该产业领导厂商的行为模式，可当作产业关键成功因素重要的信息来源。因此对领导厂商进行分析，有助于确认关键成功因素，唯对于其成功的解释仍会受到限制。

(6) 企业本体分析法：此项技术乃针对特定企业，对某些构面进行分析，如优劣势评、资源组合、优势稽核及策略能力评估等。由于透过各功能的扫描，确实有助于关键成功因素的发展，但实在耗费时间且数据相当有限。

(7) 突发因素分析法：此项技术亦是针对特定企业，需要对企业相当熟悉的专家协助。虽然较主观，却常能揭露一些其他传统客观技术无法察觉到的关键成功因素，且不受别的功能限制，甚至可以获得一些短期的关键成功因素，唯难以验证这些短期的关键成功因素。

(8) 市场策略对获利影响的分析法（PIMS Results）：针对特定企业，以PIMS（Profit Impact of Market Strategy）研究报告的结果进行分析。此技术的主要优点为其实验性基础，而缺点在于“一般性的本质”，即无法指出这些数据是否可直接应用于某一公司或某一产业，也无法得知这些因素的相对重要性。

3. 基本步骤

关键成功因素法主要包含以下步骤。（一个完整的KSF分析方法主要有5个步骤：公司定位；识别KSF；收集KSF情报；比较评估KSF；制订行动计划。）

(1) 确定企业或MIS的战略目标。

(2) 识别所有的成功因素：主要是分析影响战略目标的各种因素和影响这些因素的子因素。

(3) 确定关键成功因素。不同行业的关键成功因素各不相同。即使是同一个行业的

组织，由于各自所处的外部环境的差异和内部条件的不同，其关键成功因素也不尽相同。

（4）明确各关键成功因素的性能，指标和评估标准。

关键成功因素法的优点是能够使所开发的系统具有很强的针对性，能够较快地取得收益。应用关键成功因素法需要注意的是，当关键成功因素解决后，又会出现新的关键成功因素，就必须重新开发系统。

3.3 数据分析处理及软件工程施工方法

数据分析处理及软件工程施工方法、数据仓库联机分析处理（OLAP）、结构化生命周期法开发方法、原型法及其主要特点、公共信息模型基础知识是本节介绍的主要内容。

3.3.1 数据仓库的基本概念

1. 数据仓库的定义

目前，数据仓库尚没有统一的定义。著名的数据仓库专家W. H. Inmon在其著作《数据仓库》（*Building the Data Warehouse*）一书中给予如下描述：数据仓库（Data Warehouse）是一个面向主题的（Subject Oriented）、集成的（Integrate）、相对稳定的（Non-Volatile）、反映历史变化（Time Variant）的数据集合，用于支持管理决策。对于数据仓库的概念，可以从两个层次予以理解。首先，数据仓库用于支持决策，面向分析型数据处理，它不同于企业现有的操作型数据库；其次，数据仓库是对多个异构的数据源有效集成，集成后按照主题进行了重组，并包含历史数据，而且存放在数据仓库中的数据一般不再修改。

数据仓库可以理解为企业数据资源的合理集合，数据仓库拥有以下四个特点。

1) 面向主题

操作型数据库的数据组织面向事务处理任务，各个业务系统之间各自分离，而数据仓库中的数据是按照一定的主题域进行组织。主题是一个抽象的概念，是指用户使用数据仓库进行决策时所关心的重点方面，一个主题通常与多个操作型信息系统相关。

2) 集成的

面向事务处理的操作型数据库通常与某些特定的应用相关，数据库之间相互独立，并且往往是异构的。数据仓库中的数据是在对原有分散的数据库数据进行抽取、清理的基础上经过系统加工、汇总和整理得到的，必须消除源数据中的不一致性，以保证数据仓库内的信息是关于整个企业的一致性的全局信息。

3) 相对稳定的

操作型数据库中的数据通常实时更新，数据根据需要及时发生变化。数据仓库的数据

主要供企业决策分析之用，所涉及的数据操作主要是数据查询，一旦某个数据进入数据仓库以后，一般情况下将被长期保留，也就是数据仓库中一般有大量的查询操作，但修改和删除操作很少，通常只需要定期加载和刷新。

4) 反映历史变化

操作型数据库主要关心当前某一个时间段内的数据，而数据仓库中的数据通常包含历史信息。系统记录了企业从过去某一时点（如开始应用数据仓库的时点）到目前的各个阶段的信息，通过这些信息，可以对企业的发展历程和未来趋势做出定量分析和预测。

2. 数据仓库的建设

企业数据仓库的建设以现有企业业务系统和大量业务数据的积累为基础。数据仓库不是静态的概念，只有把信息及时交给需要这些信息的使用者，供他们做出改善其业务经营的决策，信息才能发挥作用，信息才有意义。把信息加以整理归纳和重组，并及时提供给相应的管理决策人员，是数据仓库的根本任务。整个数据仓库系统是一个包含四个层次的体系结构，如图3-13所示。

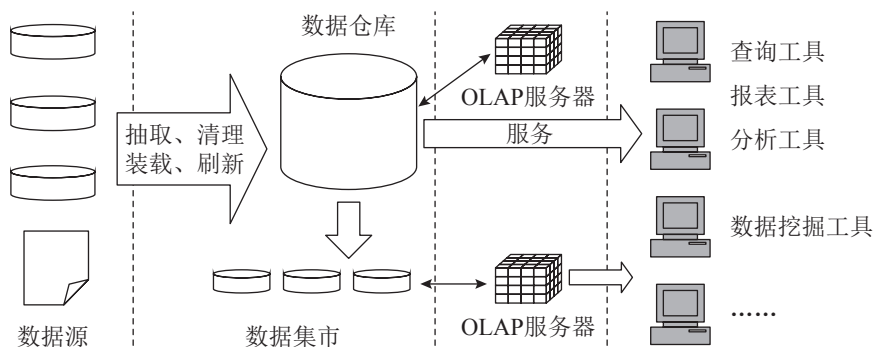


图3-13 数据仓库系统

1) 数据源

这是数据仓库系统的基础，是整个系统的数据源泉，通常包括企业内部信息和外部信息。内部信息包括存放于RDBMS中的各种业务处理数据和各类文档数据。外部信息包括各类法律法规、市场信息和竞争对手的信息等。

2) 数据的存储与管理

这是整个数据仓库系统的核心。数据仓库的真正关键是数据的存储和管理。数据仓库的组织管理方式决定了它有别于传统数据库，同时也决定了其对外部数据的表现形式。要决定采用什么产品和技术来建立数据仓库的核心，则需要从数据仓库的技术特点着手分析。针对现有各业务系统的数据，进行抽取、清理，并有效集成，按照主题进行组织。数据仓库按照数据的覆盖范围可以分为企业级数据仓库和部门级数据仓库（通常称为数据集市）。

3) OLAP服务器

对分析需要的数据进行有效集成，按多维模型予以组织，以便进行多角度、多层次的

分析, 并发现趋势。其具体实现可以分为ROLAP、MOLAP和HOLAP。ROLAP基本数据和聚合数据均存放在RDBMS之中; MOLAP基本数据和聚合数据均存放于多维数据库中; HOLAP基本数据存放于RDBMS之中, 聚合数据存放于多维数据库中。

4) 前端工具

主要包括各种报表工具、查询工具、数据分析工具、数据挖掘工具以及各种基于数据仓库或数据集市的应用开发工具。其中数据分析工具主要针对OLAP服务器, 报表工具、数据挖掘工具主要针对数据仓库。

3. 知识发现(KDD)与数据挖掘

数据挖掘就是从大量不完全的、有噪声的、模糊的或者随机的数据中提取人们事先不知道的但又是有用的信息和知识。数据挖掘其实是知识发现的核心部分, 而知识发现是在积累了大量数据后, 从中识别出有效的、新颖的、潜在的、有用的及最终可以理解的知识, 人们利用这些知识改进工作, 提高效率和效益。数据挖掘是信息发展到一定程度的必然产物, 是利用积累数据的一个高级阶段。

用数据库管理系统来存储数据, 用机器学习的方法来分析数据, 挖掘大量数据背后的知识, 这两者的结合促成了数据库中的知识发现(KDD)的产生。

1) 知识发现(KDD)与数据挖掘(DM)

KDD是从数据集中识别出有效的、新颖的、潜在有用的以及最终可理解模式的高级处理过程。从定义中可以看出, KDD是一个高级的处理过程, 它从数据集中识别出以模式来表示的知识。高级的处理过程是指一个多步骤的处理过程, 多步骤之间相互影响、反复调整, 形成一种螺旋式的上升过程。数据挖掘是KDD的核心部分, 是采用机器学习、统计等方法进行知识学习的阶段。KDD是一门交叉学科, 涉及人工智能、机器学习、模式识别、统计学、智能数据库、知识获取、数据可视化、专家系统等多个领域。数据挖掘算法的好坏将直接影响到所发现知识的好坏。数据挖掘的任务是从数据中发现模式。

2) 数据挖掘模式的种类

(1) 分类模式: 是一种分类器, 能够把数据集中的数据映射到某个给定的类上, 从而可以应用与数据预测。它常表现为一棵分类树, 根据数据的值从树根开始搜索, 沿着数据满足的分支往上走, 走到树叶就能确定类别。

(2) 回归模式: 与分类模式相似, 其差别在于分类模式的预测值是离散的, 回归模式的预测值是连续的。

(3) 时间序列模式: 根据数据随时间变化的趋势预测将来的值。其中要考虑时间的特殊性质, 只有充分考虑时间因素, 利用现有的数据随时间变化的一系列的值, 才能更好地预测将来的值。

(4) 聚类模式: 把数据划分到不同的组中, 组之间的差别尽可能大, 组内的差别尽可能小。与分类模式不同, 进行聚类前并不知道将要划分成几个组和什么样的组, 也不知道根据哪些数据项来定义组。

(5) 关联模式：是数据项之间的关联规则。关联规则是描述事物之间同时出现的规律的知识模式。在关联规则的挖掘中要充分理解数据，目标明确，数据准备工作要做好，选取恰当的最小支持度和最小可信度，很好地理解关联规则等。

(6) 序列模式：与关联模式相似，它把数据之间的关联性与时间联系起来。为了发现序列模式，不仅需要知道事件是否发生，而且需要确定事件发生的时间。

在解决实际问题时，经常要同时使用多种模式。分类模式和回归模式使用最为普遍。

在数据挖掘技术日益发展的同时，出现了许多数据挖掘工具，如何选择满足需要的数据挖掘工具已成为一个问题。常用的模式有以上六种。

3) 解决复杂问题的能力

数据量的增大，对模式精细度、准确度要求的增高都会导致问题复杂性的增大。数据挖掘系统可以提供下列方法解决复杂问题。

(1) 多种模式：多种类别模式的结合使用有助于发现有用的模式，降低问题的复杂性。

(2) 多种算法：很多模式，特别是与分类有关的模式，可以用不同的算法来实现，以适应不同的需求环境。数据挖掘系统提供多种途径产生同种模式，将更有能力解决复杂问题。

(3) 验证方法：在评估模式时采用多种校验方法，从而达到最大的准确度。

(4) 可视化：可视化工具提供了直观、简洁的方法，方便了用户，更有助于定位重要的数据，评价模式的质量，从而减少建模的复杂性。

(5) 数据选择和转换：模式通常被大量的数据项隐藏。有些数据是冗余的，有些数据是完全无关的，而这些数据项的存在会影响有价值的模式的发现。数据挖掘系统的一个很重要的功能就是能够处理数据复杂性，提供工具，选择正确的数据项和转换数据值。

(6) 扩展性：为了更有效地提高处理大量数据的效率，数据挖掘系统的扩展性十分重要。要了解数据挖掘系统能否充分利用硬件资源？是否支持并行？支持哪种并行计算机？当处理器的数量增加时，计算规模是否相应增长？是否支持数据并行存储？为单处理器的计算机编写的数据挖掘算法不会在并行计算机上自动以更快的速度运行。为充分发挥并行计算的优点，需要编写支持并行计算的算法。

(7) 操作性能：操作性能的好坏是一个至关重要的因素。图形界面友好的工具可以方便用户操作，引导用户执行任务，为用户节省时间。提供嵌入技术的工具更是它的可取之处，通过嵌入到应用程序中，缩短了开发时间。既可以将模式运用到已存在或新增加的数据上，也可以把模式导出到程序或数据库中。

(8) 数据存取能力：好的数据挖掘工具可以使用SQL语句直接从DBMS中读取数据。这样可以简化数据准备工作，并且可以充分利用数据库的优点。没有一种工具可以直接支持大量的DBMS，但可以通过通用的接口来连接大多数流行的DBMS。在考虑数据仓库应用时，重点考虑与OLAP数据访问的结合能力。

(9) 与其他产品的接口：传统的查询工具、可视化工具可以帮助用户理解数据和

结果。数据挖掘工具能否提供与这些工具集成的简易途径是衡量数据挖掘工具好坏的标准。

3.3.2 数据仓库联机分析处理

1. 联机分析处理（OLAP）与联机事务处理（OLTP）

联机分析处理（OLAP）的概念最早是由关系数据库之父E. F. Codd于1993年提出的，他同时提出了关于OLAP的12条准则。OLAP的提出引起了很大的反响，OLAP作为一类产品同联机事务处理（OLTP）明显区分开来。

1) 联机分析处理（OLAP）与联机事务处理（OLTP）的比较

当今的数据处理大致可以分成两大类：联机事务处理OLTP（On-Line Transaction Processing）、联机分析处理OLAP（On-Line Analytical Processing）。OLTP是传统的关系型数据库的主要应用，主要是基本的、日常的事务处理，如银行交易。OLAP是数据仓库系统的主要应用，支持复杂的分析操作，侧重决策支持，并且提供直观易懂的查询结果。表3-4所示为OLTP与OLAP之间的比较。

表3-4 OLTP与OLAP之间的比较

比较项目	OLTP	OLAP
用户	操作人员，低层管理人员	决策人员，高级管理人员
功能	日常操作处理	分析决策
DB设计	面向应用	面向主题
数据	当前的，最新的细节的，二维的分立的	历史的，聚集的，多维的集成的，统一的
存取	读/写数十条记录	读上百万条记录
工作单位	简单的事务	复杂的查询
用户数	上千个	上百个
DB大小	100MB~100GB	100GB~100TB

2) 联机分析处理（OLAP）的多维分析

OLAP是使分析人员、管理人员或执行人员能够从多角度对信息进行快速、一致、交互的存取，从而获得对数据的更深入了解的一类软件技术。OLAP的目标是满足决策支持或者满足在多维环境下特定的查询和报表需求，它的技术核心是“维”这个概念。

“维”是人们观察客观世界的角度，是一种高层次的类型划分。“维”一般包含层次关系，这种层次关系有时会相当复杂。通过把一个实体的多项重要的属性定义为多个维（Dimension），使用户能对不同维的数据进行比较。因此OLAP也可以说是多维数据分析工具的集合。OLAP的基本多维分析（如图3-14所示）操作有钻取（Roll Up和Drill Down）、切片（Slice）和切块（Dice），以及旋转（Pivot）、钻透（Drill Across）、钻取细节（Drill Through）等。

钻取是改变维的层次，变换分析的粒度。它包括向上钻取（Roll Up）和向下钻取（Drill Down）。向上钻取是在某一维上将低层次的细节数据概括到高层次的汇总数据，或者减少维数；而向下钻取则相反，它从汇总数据深入到细节数据进行观察或增加新维。

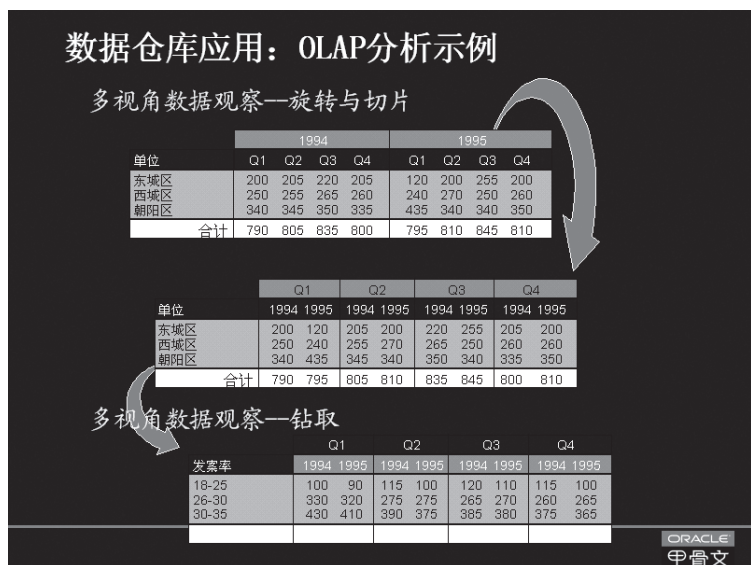


图3-14 OLAP的基本多维分析

切片和切块是在一部分维上选定值后，关心度量数据在剩余维上的分布。如果剩余的维只有两个，则是切片；如果有三个，则是切块。

旋转是变换维的方向，即在表格中重新安排维的放置（如行列互换）。

3) 联机分析处理（OLAP）的技术实现方法

OLAP有多种实现方法，根据存储数据的方式不同可以分为ROLAP、MOLAP、HOLAP。

ROLAP表示基于关系数据库的OLAP实现（Relational OLAP）。以关系数据库为核心，以关系型结构进行多维数据的表示和存储。ROLAP将多维数据库的多维结构划分为两类表：一类是事实表，用来存储数据和维关键字；另一类是维表，即对每个维至少使用一个表来存放维的层次、成员类别等维的描述信息。维表和事实表通过主关键字和外关键字联系在一起，形成了“星形模式”。对于层次复杂的维，为避免冗余数据占用过大的存储空间，可以使用多个表来描述，这种星形模式的扩展称为“雪花模式”。

MOLAP表示基于多维数据组织的OLAP实现（Multidimensional OLAP）。以多维数据组织方式为核心，也就是说，MOLAP使用多维数组存储数据。多维数据在存储中将形成“立方块（Cube）”的结构，在MOLAP中对“立方块”的“旋转”“切块”“切片”是产生多维数据报表的主要技术。

HOLAP表示基于混合数据组织的OLAP实现（Hybrid OLAP）。如低层是关系型的，高层是多维矩阵型的。这种方式具有更好的灵活性。

还有其他的一些实现OLAP的方法，如提供一个专用的SQL Server，对某些存储模式

（如星形、雪片形）提供对SQL查询的特殊支持。

4) 联机分析处理（OLAP）前端工具功能

OLAP工具是针对特定问题的联机数据访问与分析。它通过多维的方式对数据进行分析、查询和报表。

多维分析是指对以多维形式组织起来的数据采取切片（Slice）、切块（Dice）、钻取（Drill-Down和Roll-Up）、旋转（Pivot）等各种分析动作，以求剖析数据，使用户能从多个角度、多侧面地观察数据库中的数据，从而深入理解包含在数据中的信息。

在数据仓库应用中，OLAP应用一般是数据仓库应用的前端工具，同时OLAP工具还可以同数据挖掘工具、统计分析工具配合使用，增强决策分析功能。

2. 联机分析处理（OLAP）与数据挖掘（DM）

1) 联机分析处理（OLAP）与数据挖掘（DM）的关系

OLAP（联机分析处理）与DM（数据挖掘）都是数据仓库（数据库）的分析工具，在实际应用中各有侧重。前者是验证型的，后者是挖掘型的。前者建立在多维视图的基础之上，强调执行效率和对用户命令的及时响应，而且其直接数据源一般是数据仓库；后者建立在各种数据源的基础上，重在发现隐藏在数据深层次的对人们有用的模式（Pattern），一般并不过多考虑执行效率和响应速度。联机分析挖掘（On-Line Analytical Mining, OLAM）是两者相结合的产物，又称为联机分析处理挖掘（OLAP Mining）。

2) 联机分析处理（OLAP）与数据挖掘（DM）的特点

OLAP主要有两个特点，一是在线性（On-Line），体现为对用户请求的快速响应和交互式操作，它的实现是由Client/Server体系结构来完成的；二是多维分析（Multi-Analysis），这也是OLAP技术的核心所在。

数据挖掘（Data Mining, DM），或者说知识发现（KDD），是从大量原始数据中抽取模式的一个处理过程，抽取出来的模式就是所谓的知识，必须具备可信、新颖、有效和易于理解的特点。

3) 联机分析处理（OLAP）与数据挖掘（DM）的结合

从以上的分析中我们可以发现，OLAP和DM虽然都是数据仓库（数据库）的分析工具，但其应用范围和侧重点是不同的。

OLAP的在线性体现在与用户的交互和快速响应，多维性则体现在它建立在多维视图的基础上。用户积极参与分析过程，动态地提出分析要求、选择分析算法，对数据进行由浅及深的分析。

DM与OLAP不同，主要体现在它分析数据的深入和分析过程的自动化，其分析过程不需要用户的参与。这是它的优点，也正是它的不足，因为在实际中，用户也希望参与到挖掘中来，如只想对数据的某一子集进行挖掘，以及对不同抽取、集成水平的数据进行挖掘，还想根据自己的需要动态选择挖掘算法等。

由此可见，OLAP与DM各有所长，如果能将两者结合起来，发展一种建立在OLAP和

数据仓库基础上的新的挖掘技术，将更能适应实际的需要。

4) 技术结合的优势和缘由

DM工具需要的数据是一些经过净化、集成处理的数据，通常这种处理过程也是昂贵的；而DW（Data Warehouse，数据仓库）作为OLAP的数据源，存储的就是这样的数据。它能为OLAP提供数据，当然也可以为DM提供数据。

DW是一项崭新的技术，很多人在研究它。围绕着它有许多工具或体系结构。而DM作为数据分析工具的一种，不是孤立的，也必然要与其他工具发生联系。因此，考虑到如何最大限度地利用这些现成的工具，也是OLAM发展之初所关心的问题。

成功的数据挖掘需要对数据进行钻探性（Exporatory）分析，比如，挖掘所需的数据可能只是一部分、一定范围的数据。因此，对多维数据模型的切片、切块、下钻等操作，同样可以应用于DM。也就是说，可以将DM建立在多维模型（或说超级立方体）的基础之上。

用户的参与对DM的重要性体现在动态地提出挖掘要求、选择挖掘算法。故可以将OLAP的Client/Server结构或三层结构应用于DM中来。

通过以上分析可以得出结论：OLAP与数据挖掘相结合，兼有OLAP多维分析的在线性、灵活性和数据挖掘对数据处理的深入性，是数据仓库（数据库）应用工具未来发展的方向。在具体的数据仓库项目中，采用的根本技术思想也正是此先进思想。这为项目实施过程也提供了特有的扩展性。

3.3.3 结构化生命周期法开发方法

结构化开发方法（Structured System Analysis and Design，SSA&D）亦称结构化生命周期法，是指用系统工程的思想 and 工程化的方法，按照用户至上的原则，自顶向下整体性分析设计和自底向上逐步实施的系统开发过程。

结构化开发方法是系统工程思想和工程化方法在系统开发领域的运用。它先将整个信息系统开发过程划分出若干个相对独立的阶段，如系统规划、系统分析、系统设计、系统实施等，再严格规定每个阶段的任务和工作步骤，同时提供便于理解和交流的开发工具与方法（图表）。在系统分析时，采用自顶向下、逐层分解，由抽象到具体的逐步认识过程；在系统设计时，先考虑系统整体的优化，再考虑局部的优化问题；在系统实施时，则自底向上，先局部后整体，通过标准化模块的链接形成完整的系统。

使用结构化开发方法开发一个系统，可将整个开发过程划分为若干个首尾相连接的阶段，每个阶段内部又包含若干前后关联的工作步骤，一般称之为系统开发生命周期。生命周期各阶段的划分不尽相同，如三阶段、四阶段、五阶段等，但人们对整个开发过程完成的主要工作有一致的认同。管理信息系统生命周期的五阶段划分方法如图3-15所示。

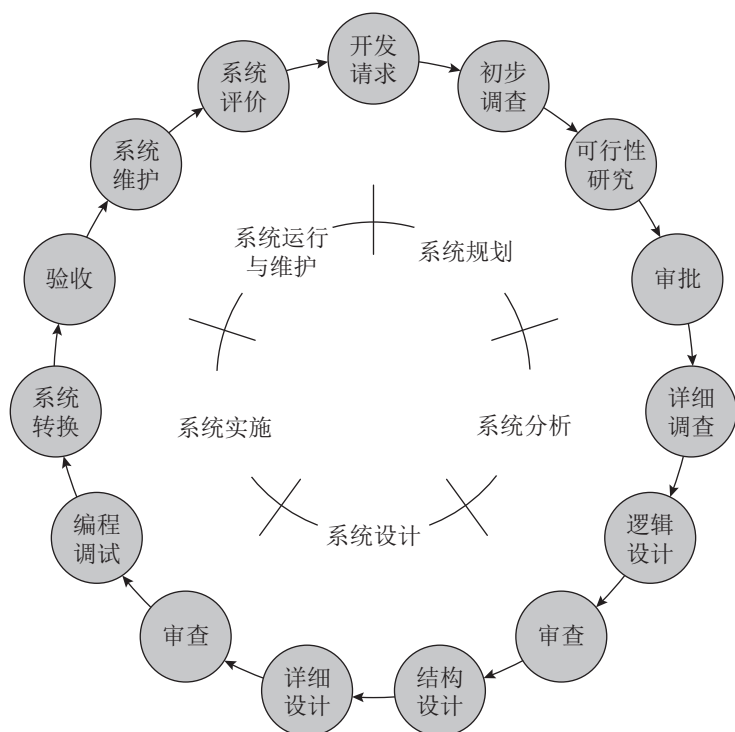


图3-15 管理信息系统的生命周期

1) 系统规划阶段

这一阶段的任务是对企业的环境、目标、现行系统的状况进行初步调查，明确问题，确定信息系统的发展战略，对建设新系统的需求做出分析和预测，分析建设新系统所受的各种约束，研究建设新系统的必要性和可能性。根据需要和可能，给出拟建系统的备选方案，从技术和经济角度对方案进行可行性分析，写出可行性分析报告，提交用户批准后，将系统建议方案及实施计划编写成系统开发任务书，进入系统分析阶段。

2) 系统分析阶段

该阶段的主要内容是对现行系统进行详细调查，在此基础上进行组织机构功能分析、管理业务流程分析、数据与数据流程分析、功能与数据之间的关系分析，再建立新系统的逻辑模型，形成综合性的系统分析报告，并提交用户讨论审核，然后转入系统设计阶段。

3) 系统设计阶段

在系统分析工作基础上，以系统分析报告为依据，进行总体结构设计，然后分别进行代码设计、数据库/文件设计、输入/输出设计、模块结构与功能设计。与此同时，根据总体设计的要求，购置有关设备，并且安装调试。最终给出系统的物理模型和系统设计报告，提交用户讨论审核。批准确认后，转入系统实施阶段。

4) 系统实施阶段

系统实施是将设计的系统付诸实现的阶段。这一阶段工作内容包括程序设计，程序调试（单调、分调、总调），计算机等设备的购置、安装与调试，人员培训，数据准备和初

始化,系统调试与转化,最后投入试运行并且进行完善性维护。

5) 系统运行与维护阶段

该阶段是系统开发成功后,交付用户正式使用、发挥效益的时期。其主要工作内容包括系统的日常运行管理与维护,系统综合评价及系统开发项目的监理审计等。在系统运行过程中,可能由于环境变化导致系统功能不足,或者开发过程中未能发现或无法解决的功能要求,需要对系统进行修改、维护或者局部调整。当系统运行若干年之后,系统运行的环境可能发生了根本性的变化,出现一些不可调和的大问题。此时,用户将会进一步提出开发新系统的要求,也就标志着老系统生命的结束,新系统的诞生。

强调规范化、标准化的结构化系统开发方法逐渐地暴露出一些缺点和不足,主要表现在以下四个方面。

(1) 开发周期过长。严格的阶段划分和文档要求费时费事,致使开发周期延长。

(2) 难以适应迅速变化的环境。由于开发周期长,不能随便变更前一阶段的工作成果,可能导致后一阶段工作无法及时把环境变化的要求反映到开发方案中来,致使最终开发出来的系统可能脱离现实。

(3) 使用的工具落后。大量的分析设计图表难以采用计算机完成,而必须通过手工绘制,费事费时。

(4) 有违认识事物的规律性。要求系统开发者在调查中充分掌握用户需求、管理状况以及预见可能发生的变化,这不太符合人们循序渐进地认识事物的规律性。

3.3.4 原型法及其主要特点

原型法(Prototyping Approach, PA)是20世纪80年代随着计算机软件技术的发展,特别是在关系数据库系统、第四代程序设计语言和各种功能强大的辅助系统开发工具产生的基础之上,提出的一种具有全新的设计思想和开发工具的系统开发方法。

1. 原型与原型法概念

所谓原型,是指由系统分析设计人员与用户合作,在短期内定义用户基本需求的基础上,开发出来的一个只具备基本功能、实验性的、简易的应用软件。它不同于只是逻辑意义上的、不可运行的“模型”。实际上,原型就是一个实实在在的、可运行的管理信息系统软件,只不过由于对用户需求把握尚不全面和准确,软件的功能并不十分完善而已。

原型法是指借助于功能强大的辅助系统开发工具,按照不断寻优的设计思想,通过反复的完善性实验而最终开发出来符合用户要求的管理信息系统的过程和方法。首先快速开发一个原型,然后运行这个原型,再通过对原型的不断评价和改进,使之逐步完善,直至用户满意为止。

原型法的工作流程如图3-16所示。原型法的开发必须满足四个条件,即开发周期必须短且成本低,要求用户参与评价原型,原型必须是可运行的,原型易于修改。

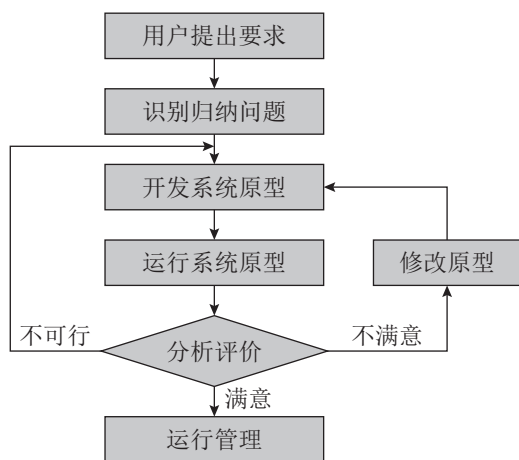


图3-16 原型法的工作流程

2. 原型法的软件支持环境

原型法的运用必须有一个前提，那就是要有一个强有力的软件支持环境。没有功能强大的辅助系统开发工具的支持，原型法的优势难以变成现实。

3. 原型法的特点

原型法的原理和流程都十分简单，并无任何高深的理论和技术，因而备受推崇，并且在实践中获得巨大的成功。这主要是由于原型法更多地遵循了人们认识事物的规律，更容易为人们所普遍接受。

1) 认识论上的突破

从认识论的角度来看，人们认识事物不可能一次就完全了解。开发过程是一个循环往复的反馈过程，它符合用户对计算机应用的认识逐步发展、循序渐进的过程。人们对于事物的描述，往往都受环境的启发而不断完善，开始时用户和设计者对于系统的功能要求的认识是不完整的、粗糙的，通过建立原型、演示或使用原型、修改原型的循环过程，设计者以原型为媒介，及时取得来自用户的反馈信息，不断发现问题，反复修改、完善系统，确保用户的要求得到较好的满足。

2) 改进了用户和系统设计者的信息交流方式

原型法将模拟手段引入系统分析的初期阶段，有用户的直接参与，所有问题的讨论都是围绕某个确定原型而进行的，这样可以直接及时地发现问题，并且进行修正。能够及早暴露系统实现后存在的一些问题，可以减少信息系统的设计错误，开发人员在开发过程中就能对系统做出改进。

3) 更加贴近实际，提高用户满意程度

借助于原型系统，为用户建立正确的信息模型和功能模型，由用户和系统设计者、编程人员共同制订合理的解决方案。由原型法得出的最终系统的需求来自原型系统的运行经

验，为用户建立了正确的信息模型，完全符合用户的需求。

4) 降低了开发风险及开发成本

由于使用原型系统测试开发思想及方案，只有当风险程度通过原型使用户和开发人员意见一致时，才能继续开发最终系统，因而减少了开发失败的可能性。在原型法的应用中无须多余的文档资料，还可充分利用最新的软件工具，摆脱了老一套的工作方法，使系统开发的时间、费用大大减少，效率、技术等方面都大大地提高了。

4. 原型法的局限性

原型法比起结构化系统开发方法固然有其时代的优越性，但它不能取代结构化开发方法。它的应用有一定的局限性，主要表现在以下四个方面。

1) 开发工具要求高

原型法需要现代化的开发工具支持，否则开发工作量太大，成本过高，就失去了采用原型法的优势。应该说，开发工具的水平是原型法能否顺利实现的第一要素。

2) 解决复杂系统和大系统问题很困难

对于大型的系统，如果不经系统分析来进行整体性划分，想要直接用屏幕一个一个地模拟是很困难的。对于复杂系统（如复杂的控制系统），功能种类多，技术复杂，实现困难，与性能仿真模拟工具和应用业务领域知识密切相关，进入实用阶段的很少，所以原型法很难解决大系统和复杂系统问题。

3) 管理水平要求高

如果原系统的基础管理不善，信息处理过程混乱，将使原型法的运用产生困难。首先是由于对象工作过程不清，构造原型有一定困难；其次是由于基础管理不好，没有科学合理的方法可依，系统开发容易走上机械地模拟原来手工系统的做法。

4) 系统的交互方式必须简单明了

对于大量运算的、逻辑性较强的程序模块，原型法很难构造模型供人评价，因为这类问题没有那么多的交互方式，也不是三言两语就可以把问题说清楚的。对于有大量批处理的系统，由于交互方式问题，使用原型法也会遇到某些困难。

3.3.5 公共信息模型基础知识

1. 公共信息模型（CIM）基本概念

公共信息模型（CIM）是一个抽象模型，描述电力企业的所有主要对象，特别是与电力运行有关的对象。它分为CIM规范（CIM Specification）和CIM模式（CIM Schema）两个部分，包含15种逻辑包，且CIM的类包含唯一标识对象实例的关键字。通过提供一种用对象类和属性及它们之间关系来表示电力系统资源的标准方法，CIM方便了实现不同卖方独立开发的能量管理系统（EMS）应用的集成，多个独立开发的完整EMS系统之间的集

成, 以及EMS系统和其他涉及电力系统运行的不同方面的系统, 如发电或配电系统之间的集成。这是通过定义一种基于CIM的公共语言(即语法和语义), 使得这些应用或系统能够不依赖于信息的内部表示而访问公共数据和交换信息来实现的。

CIM中描述的对象类本质上是抽象的, 可以用于各种应用。CIM的使用远远超出了它在EMS中应用的范围。应当把本标准理解作为一种能够在任何领域实行集成的工具, 只要该领域需要一种公共电力系统, 使得应用和系统之间能够实现互操作和插入兼容性, 而与任何具体实现无关。

由于完整的CIM的规模较大, 所以将包含在CIM中的对象分成了几个逻辑包, 每个逻辑包代表整个电力系统模型的某一部分。这些包的集合发展成为独立的标准。DL/T890的本部分规定了包的基本集合, 提供了电力企业内部各应用共享的EMS信息的物理方面的逻辑视图。其他标准规定了某些特定应用所需的模型的特殊部分。

(1) 核心包(Core): 包含所有应用共享的核心的命名(Naming)、电力系统资源(Power System Resource)、设备容器(Equipment Container)和导电设备(Conducting Equipment)实体, 以及这些实体的常见的组合。并不是所有的应用都需要所有的核心包实体。这个包不依赖于任何其他包, 而其他包中的大部分都具有依赖于本包的关联和普遍化。

(2) 拓扑包(Topology): 这个包是核心包的扩展, 它与Terminal类一起建立连接性(Connectivity)的模型, 而连接性是设备怎样连接在一起的物理定义。另外, 它还建立了拓扑(Topology)的模型, 拓扑是设备怎样通过闭合开关连接在一起的逻辑定义。拓扑的定义与其他的电气特性无关。

(3) 电线包(Wires): 这个包是核心包和拓扑包的扩展, 它建立了输电(Transmission)和配电(Distribution)网络的电气特性的信息模型。这个包用于网络应用, 如状态估计(State Estimation)、潮流(Load Flow)及最优潮流(Optimal Power Flow)。

(4) 停运包(Outage): 这个包是核心包和电线包的扩展, 它建立了当前及计划网络结构的信息模型。这些实体在典型的网络应用中是可选的。

(5) 保护包(Protection): 这个包是核心包和电线包的扩展, 它建立了保护设备, 例如继电器的信息模型。这些实体用于培训模拟和配电网故障定位应用。

(6) 测量包(Meas): 这个包包含描述各应用之间交换的动态测量数据的实体。

(7) 负荷模型包(Load Model): 这个包以曲线及相关的曲线数据的形式为能量用户及系统负荷提供模型。这里还包括影响负荷的特殊情况, 如季节与日类型。这一信息由负荷预测(Load Forecasting)和负荷管理(Load Management)使用。

(8) 发电包(Generation): 这个包分成两个子包, 即电力生产包(Production)和发电动态包(Generation Dynamics)。

①电力生产包(Production): 这个包提供了各种类型发电机的模型。它还建立了生产成本信息模型, 用于发电机间进行经济需求分配及计算备用量大小。这一信息用于机组

组合（Unit Commitment）、水力和火力发电机组的经济调度（Economic Dispatch）、负荷预测及自动发电控制（Automatic Generation Control）等应用。

②发电动态包（Generation Dynamics）：这个包提供原动机，如汽轮机和锅炉的模型，这些模型在模拟和培训应用中需要用到。这一信息用于动态培训仿真器（Dynamic Training Simulator）应用的机组建模。

（9）域包（Domain）：域包是量与单位的数据字典，定义了可能被其他任何包中的任何类使用的属性的数据类型。此包包含原始数据类型的定义，包括量测的单位和允许的值。每一种数据类型包含一个值（Value）属性和一个可选的量测单位（Unit），这个单位指定为一个被初始化为该量测单位文字描述的静态变量。枚举型数据的允许值在该属性的文档（documentation）中用UML约束句法在大括号（{}）内列出。字符串长度在文档中列出，并被指定为长度属性。

（10）财务包（Financial）：财务包与结算和会计有关。这些类表达了参与正式和非正式协议的法律实体。

（11）能量计划包（Energy Scheduling）：能量计划包提供了对公司之间的电力交易进行计划和考核的能力。它包括电力产生、消费、损失、输送、出售和采购的交易。这些类应用在电能的考核结算、发电容量、电能传输、辅助服务中。

（12）备用包（Reservation）：备用包包含了用于电能交易计划、发电容量、电能传输、辅助服务中的信息。

（13）SCADA包（SCADA）：SCADA包描述了用于数据采集（SCADA）和控制应用的模型信息，涉及量测、PT、CT、RTU、扫描装置、通信电路等设备。控制应用支持对设备的控制操作，如断开/合上断路器；数据采集应用从多个来源采集遥测数据，遥测实体的子类型要遵照IEC61850标准的定义。SCADA包也支持报警的表达，但是不希望被其他应用使用。

2. CIM类和关系

每一个CIM包的类图展示了该包中的所有的类及它们的关系。在与其他包中的类存在关系时，这些类也展示出来，而且标以表明其所属的包的类别。

类与对象所建的模型正是电力系统中需要一种对多种EMS应用通用的方法来描绘的东西，一个类是对现实世界中发现的一种对象的表示，如在EMS中需要表示为整个电力系统模型的一部分的变压器、发电机或负荷。其他类型的对象包括EMS应用需要处理、分析与储存的计划与量测。这些对象需要一种通用的表示，以达到EMS-API标准的插入兼容和互操作的目的。在电力系统中具有唯一身份的一个具体对象则被建模成它所属类的一个实例。

还应该注意，定义CIM是为了方便数据交换。CIM实体除了默认主人的生成、删除、更新和读出外，没有其他行为。为了使CIM尽可能地通用，非常希望对具体实现于配置。一般来说，改变属性的值或域比改变类定义更为容易。这些原则暗示CIM应当避免定

义类的太多的具体子类型。相反，CIM定义一些通用的类，由属性来给定类型名。因而应用可以根据需要使用这一信息去实例化具体的对象类型。应用可能需要其他信息去定义有效类型与关系的集合。

类具有描述对象特性的属性。CIM中的每一个类包含描述和识别具体实例的属性。只有各个EMS应用共同感兴趣的那些属性才包括在类的描述中。

每一个属性都有一个类型，它识别该属性是哪一种类型的属性。典型的属性类型有整型、浮点型、布尔型、字符串型及枚举型，它们被称为原始类型。然而，许多其他类型也被定义为CIM规范的一部分。例如，补偿器（Compensator）有一个类型为Voltage的Maximumkv属性。数据类型的定义包含在Domain包中。

类之间的关系揭示了它们相互之间是怎样构造的，即普遍化、简单关联、聚集。

1) 普遍化

普遍化是一个被称为普遍的类与一个较具体的类之间的一种关系。较具体的类只能包含附加的信息。例如，一台电力变压器是电力系统资源的一种具体型。普遍化使具体的类可以从它上层的所有更普遍的类中继承属性和关系。

2) 简单关联

简单关联是类之间的概念上的联系。每一个关联具有两个角色。每个角色表示了关联中的一种方向，表示目标类与源类有关系，角色给定为目标类的名称，具有或不具有动词短语。每一个角色也具有多样性/汇集性的属性，以确定有多少对象可以参与指定的关系。在CIM中，关联关系是不需要命名的。

3) 聚集

聚集是关联关系的一个特例。聚集关系指明类之间的关系是整体和局部的关系，整体的类由局部的类“组成”，或整体的类“包含”局部的类，局部的类是整体的类的一部分。局部的类并不是像泛化关系中从整体的类继承而来。

然而，与面向对象的类不同，CIM的类包含唯一标识对象实例的关键字。纯面向对象的设计不用关键字标识实例，虽然一些对象技术（如Enterprise JavaBeans）的确有关键字定义。另外，CIM包含一个很像数据库连接的特别的关联类。这个关联涉及的实例，就像磁盘外壳和外壳里的物理磁盘。这种关联不是创建外壳和磁盘类之间的关系，而是定义两个类之间的关系，类本身没有这样的关系。通过关联而不是通过显式查询受管资源类实例来为实例定位，使用CIM对象实例中的关键字才可使关联定位成为可能。

CIM和纯面向对象设计的区别是由于它们在各自环境中有不同的使用意向。然而面向对象技术已经发展成服务于创建应用程序编程环境的需求，CIM则明确地适用于描述，编目录和与受管资源交互。为进一步理解CIM是如何工作的，应该看创建和构建它的类层次和各种模型的方式。

第5章

人工智能工程咨询项目实施与应用分析

人工智能基本概念及发展历程、新一代人工智能技术发展战略态势、输变电系统人工智能工程方法、配用电人工智能工程应用方法、电网智能运检人工智能技术应用、电网运行人工智能工程应用分析、供用电服务人工智能工程应用分析是本章重点介绍的主要内容。



5.1 人工智能工程技术基础知识

人工智能基本概念及发展历程、新一代人工智能技术发展战略态势、我国电力企业人工智能应用工程概述是本节介绍的主要内容。

5.1.1 人工智能基本概念及发展历程

1. 人工智能的基本概念

人工智能（Artificial Intelligence，AI）是研究开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术学科。人工智能是计算机科学的一个分支，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器，该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。人工智能从诞生以来，理论和技术日益成熟，应用领域也不断扩大，可以设想，未来人工智能带来的科技产品，将会是人类智慧的“容器”。人工智能不是人的智能，但能像人那样思考，也可能超过人的智能。

著名的美国斯坦福大学人工智能研究中心的尼尔逊教授对人工智能下了这样一个定义：“人工智能是关于知识的学科——怎样表示知识以及怎样获得知识并使用知识的学科。”而美国麻省理工学院的温斯顿教授认为：“人工智能就是研究如何使计算机去做过去只有人才能做的智能工作。”这些说法反映了人工智能的基本思想和基本内容，即人工智能是研究人类智能活动的规律，构造具有一定智能的人工系统，研究如何让计算机去完成以往需要人的智力才能胜任的工作，也就是研究如何应用计算机的软硬件来模拟人类某些智能行为的基本理论、方法和技术。

人工智能是计算机科学的一个分支，20世纪70年代以来被称为世界三大尖端技术之一（空间技术、能源技术、人工智能）。也被认为是21世纪三大尖端技术（基因工程、纳米科学、人工智能）之一。这是因为近30年来它获得了迅速的发展，在很多学科领域都获得了广泛应用，并取得了丰硕的成果，人工智能已逐步成为一个独立的分支，无论在理论和实践上都已自成一个系统。

人工智能是研究使计算机模拟人的某些思维过程和智能行为（如学习、推理、思考、规划等）的学科，主要包括计算机实现智能的原理、制造类似于人类智能的计算机，使计算机能实现更高层次的应用。人工智能涉及计算机科学、心理学、哲学和语言学等学科，可以说几乎是自然科学和社会科学的所有学科，已远远超出了计算机科学的范畴。人工智能与思维科学的关系是实践和理论的关系，人工智能处于思维科学的技术应用层次，是它的一个应用分支。从思维观点看，人工智能不仅限于逻辑思维，还要考虑形象思维、灵感思维，才能促进人工智能的突破性发展。数学常被认为是多种学科的基础科学，数学也进入了语言、思维领域，人工智能也必须借用数学工具。数学不仅在标准逻辑、模糊数学等范畴发挥作用，数学也进入了人工智能领域，它们将互相促进，从而更快地发展。

2. 人工智能的发展历程

从诞生至今，人工智能已有60多年的发展历史，大致经历了三次浪潮。第一次浪潮为20世纪50年代末至20世纪80年代初，第二次浪潮为20世纪80年代初至20世纪末，第三次浪潮为21世纪初至今。在人工智能的前两次浪潮中，由于技术未能实现突破性进展，相关应用始终难以达到预期效果，无法支撑起大规模商业化应用，最终在经历过两次高潮之后，人工智能归于沉寂。随着信息技术快速发展和互联网快速普及，以2006年深度学习模型的提出为标志，人工智能迎来第三次高速成长。

1) 第一次浪潮：人工智能诞生并快速发展，但技术瓶颈难以突破

符号主义盛行，人工智能快速发展。1956—1974年是人工智能发展的第一个黄金时期。科学家将符号方法引入统计方法中进行语义处理，出现了基于知识的方法，人机交互开始成为可能。科学家发明了多种具有重大影响的算法，如深度学习模型的雏形贝尔曼公式。除在算法和方法论方面取得了新进展，科学家们还制作出具有初步智能的机器，如能证明应用题的机器STUDENT（1964年）、可以实现简单人机对话的机器ELIZA（1966年）。人工智能发展速度迅猛，以至于研究者普遍认为人工智能代替人类只是时间问题。

模型存在局限，人工智能步入低谷。1974—1980年，人工智能的瓶颈逐渐显现，逻辑证明器、感知器、增强学习只能完成指定的工作，对于超出范围的任务则无法应对，智能水平较为低级，局限性较为突出。造成这种局限的原因主要体现在两个方面：一是人工智能基于的数学模型和数学手段被发现具有一定的缺陷；二是很多计算的复杂度呈指数级增长，依据现有算法无法完成计算任务。先天的缺陷是人工智能在早期发展过程中遇到的瓶颈，研发机构对人工智能的热情逐渐冷却，对人工智能的资助也相应被缩减或取消，人工智能第一次步入低谷。

2) 第二次浪潮：模型突破带动初步产业化，但推广应用存在成本障碍

数学模型实现重大突破，专家系统得以应用。进入20世纪80年代，人工智能再次回到了公众的视野当中。人工智能相关的数学模型取得了一系列重大发明成果，其中包括著名的多层神经网络（1986年）和BP反向传播算法（1986年）等，这进一步催生了能与人类下象棋的高度智能机器（1989年）。其他成果包括通过人工智能网络来实现能自动识别信

封上邮政编码的机器，精度可达99%以上，已经超过普通人的水平。与此同时，美国的卡耐基·梅隆大学为DEC公司制造出了专家系统（1980年），这个专家系统可帮助DEC公司每年节约4000万美元的费用，特别是在决策方面能提供有价值的内容。受此鼓励，很多国家包括日本、美国都再次投入巨资开发所谓第5代计算机（1982年），当时叫作人工智能计算机。

成本高且难维护，人工智能再次步入低谷。为推动人工智能的发展，研究者设计了LISP语言，并针对该语言研制了Lisp计算机。该机型指令执行效率比通用型计算机更高，但价格昂贵且难以维护，始终难以大范围推广普及。与此同时，在1987—1993年，苹果和IBM公司开始推广第一代台式机，随着性能不断提升和销售价格的不断降低，这些个人电脑逐渐在消费市场上占据了优势，越来越多的计算机走入家庭，价格昂贵的Lisp计算机由于古老陈旧且难以维护逐渐被市场淘汰，专家系统也逐渐淡出人们的视野，人工智能硬件市场出现明显萎缩。同时，政府经费开始下降，人工智能又一次步入低谷。

3) 第三次浪潮：信息时代催生新一代人工智能，但未来发展存在诸多隐忧

新兴技术快速涌现，人工智能发展进入新阶段。随着互联网的普及、传感器的泛在、大数据的涌现、电子商务的发展、信息社区的兴起，数据和知识在人类社会、物理空间和信息空间之间交叉融合、相互作用，人工智能发展所处信息环境和数据基础发生了巨大而深刻的变化，这些变化构成了驱动人工智能走向新阶段的外在动力。与此同时，人工智能的目标和理念出现重要调整，科学基础和实现载体取得新的突破，类脑计算、深度学习、强化学习等一系列的技术萌芽也预示着内在动力的成长，人工智能的发展已经进入一个新的阶段。

人工智能水平快速提升，得益于数据量的快速增长、计算能力的大幅提升以及机器学习算法的持续优化，新一代人工智能在某些给定任务中已经展现出达到或超越人类的工作能力，并逐渐从专用型智能向通用型智能过渡，有望发展为抽象型智能。随着应用范围的不断拓展，人工智能与人类生产生活联系的愈发紧密，一方面给人们带来诸多便利，另一方面也产生了一些潜在问题。一是加速机器换人，结构性失业可能更为严重；二是隐私保护成为难点，数据拥有权、隐私权、许可权等界定存在困难。

5.1.2 新一代人工智能技术发展战略态势

1. 战略态势

1) 人工智能发展进入新阶段

经过60多年的演进，特别是在移动互联网、大数据、超级计算、传感网、脑科学等新理论新技术以及经济社会发展强烈需求的共同驱动下，人工智能加速发展，呈现出深度学习、跨界融合、人机协同、群智开放、自主操控等新特征。大数据驱动知识学习、跨媒体协同处理、人机协同增强智能、群体集成智能、自主智能系统成为人工智能的发展重点，

受脑科学研究成果启发的类脑智能蓄势待发，芯片化硬件化平台化趋势更加明显，人工智能发展进入新阶段。当前，新一代人工智能相关学科在理论建模、技术创新、软硬件升级等整体推进，正在引发链式突破，推动经济社会各领域从数字化、网络化向智能化加速跃升。

2) 人工智能成为国际竞争的新焦点

人工智能是引领未来的战略性技术，世界主要发达国家都把发展人工智能作为提升国家竞争力、维护国家安全的重大战略，加紧出台规划和政策，围绕核心技术、顶尖人才、标准规范等强化部署，力图在新一轮国际科技竞争中掌握主导权。当前，我国国家安全和国际竞争形势更加复杂，必须放眼全球，把人工智能发展放在国家战略层面系统布局、主动谋划，牢牢把握人工智能发展新阶段国际竞争的战略主动，打造竞争新优势、开拓发展新空间，有效保障国家安全。

3) 人工智能成为经济发展的新引擎

人工智能作为新一轮产业变革的核心驱动力，将进一步释放历次科技革命和产业变革积蓄的巨大能量，并创造新的强大引擎，重构生产、分配、交换、消费等经济活动各环节，形成从宏观到微观各领域的智能化新需求，催生新技术、新产品、新产业、新业态、新模式，引发经济结构重大变革，深刻改变人类生产生活方式和思维模式，实现社会生产力的整体跃升。我国经济发展进入新常态，深化供给侧结构性改革任务非常艰巨，必须加快人工智能深度应用，培育壮大人工智能产业，为我国经济发展注入新动能。

4) 人工智能带来社会建设的新机遇

我国正处于全面建成小康社会的决胜阶段，人口老龄化、资源环境约束等挑战依然严峻，人工智能在教育、医疗、养老、环境保护、城市运行、司法服务等领域广泛应用，将极大提高公共服务精准化水平，全面提升人民生活品质。人工智能技术可准确感知、预测、预警基础设施和社会安全运行的重大态势，及时把握群体认知及心理变化，主动决策反应，将显著提高社会治理的能力和水平，对有效维护社会稳定具有不可替代的作用。

5) 人工智能发展的不确定性带来新挑战

人工智能是影响面广的颠覆性技术，可能带来改变就业结构、冲击法律与社会伦理、侵犯个人隐私、挑战国际关系准则等问题，将对政府管理、经济安全和社会稳定乃至全球治理产生深远影响。在大力发展人工智能的同时，必须高度重视可能带来的安全风险挑战，加强前瞻预防与约束引导，最大限度降低风险，确保人工智能安全、可靠、可控发展。

6) 我国发展人工智能具有良好基础

国家部署了智能制造等国家重点研发计划重点专项，印发了“互联网+”人工智能三年行动实施方案，从科技研发、应用推广和产业发展等方面提出了一系列措施。经过多年的持续积累，我国在人工智能领域取得重要进展，国际科技论文发表量和发明专利授权量已居世界第二，部分领域核心关键技术实现重要突破。语音识别、视觉识别技术世界领先，自适应自主学习、直觉感知、综合推理、混合智能和群体智能等初步具备跨越发展的

能力,中文信息处理、智能监控、生物特征识别、工业机器人、服务机器人、无人驾驶逐步进入实际应用,人工智能创新创业日益活跃,一批龙头骨干企业加速成长,在国际上获得广泛关注和认可。加速积累的技术能力与海量的数据资源、巨大的应用需求、开放的市场环境有机结合,形成了我国人工智能发展的独特优势。

同时,要清醒地看到,我国人工智能整体发展水平与发达国家相比仍存在差距,缺少重大原创成果,在基础理论、核心算法以及关键设备、高端芯片、重大产品与系统、基础材料、元器件、软件与接口等方面差距较大;科研机构和企业尚未形成具有国际影响力的生态圈和产业链,缺乏系统的超前研发布局;人工智能尖端人才远远不能满足需求;适应人工智能发展的基础设施、政策法规、标准体系亟待完善。

面对新形势新需求,必须主动求变应变,牢牢把握人工智能发展的重大历史机遇,紧扣发展、研判大势、主动谋划、把握方向、抢占先机,引领世界人工智能发展新潮流,服务经济社会发展和支撑国家安全,带动国家竞争力整体跃升和跨越式发展。

2. 新一代人工智能的主要驱动因素

当前,随着移动互联网、大数据、云计算等新一代信息技术的加速迭代演进,人类社会与物理世界的二元结构正在进阶到人类社会、信息空间和物理世界的三元结构,人与人、机器与机器、人与机器的交流互动愈加频繁。人工智能发展所处的信息环境和数据基础发生了深刻变化,愈加海量化的数据,持续提升的运算力,不断优化的算法模型,结合多种场景的新应用已构成相对完整的闭环,成为推动新一代人工智能发展的四大要素。

1) 人机物互联互通成趋势,数据量呈现爆炸性增长

近年来,得益于互联网、社交媒体、移动设备和传感器的大量普及,全球产生并存储的数据量急剧增加,为通过深度学习的方法来训练人工智能提供了良好的土壤。目前,全球数据总量每年都以倍增的速度增长,预计到2020年将达到44万亿GB,中国产生的数据量将占全球数据总量的近20%。海量的数据将为人工智能算法模型提供源源不断的素材,人工智能正从监督式学习向无监督学习演进升级,从各行业、各领域海量数据中积累经验,发现规律,持续提升。

2) 数据处理技术加速演进,运算能力实现大幅提升

人工智能领域富集了海量数据,传统的数据处理技术难以满足高强度、高频次的处理需求。人工智能芯片的出现加速了深层神经网络的训练迭代速度,让大规模的数据处理效率显著提升,极大地促进了人工智能行业的发展。目前,出现了GPU、NPU、FPGA和各种各样的AI-PU专用芯片。相比传统的CPU只能同时做一两个加减法运算,NPU等专用芯片多采用“数据驱动并行计算”的架构,特别擅长处理视频、图像类的海量多媒体数据。在具有更高线性代数运算效率的同时,只产生比CPU更低的功耗。

3) 深度学习研究成果卓著,带动算法模型持续优化

2006年,加拿大多伦多大学教授杰弗里·辛顿提出了深度学习的概念,极大地发展了人工神经网络算法,提高了机器自学习的能力。例如,谷歌大脑团队在2012年通过使用深

度学习技术，成功让计算机从视频中“认出”了猫。随着算法模型的重要性进一步凸显，全球科技巨头纷纷加大了这方面的布局力度和投入，通过成立实验室，开源算法框架，打造生态体系等方式推动算法模型的优化和创新。目前，深度学习等算法已经广泛应用在自然语言处理、语音处理以及计算机视觉等领域，并在某些特定领域取得了突破性进展，从有监督式学习演化为半监督式、无监督式学习。

4) 资本与技术深度耦合，助推行业应用快速兴起

当前，在技术突破和应用需求的双重驱动下，人工智能技术已走出实验室，加速向产业各个领域渗透，产业化水平大幅提升。在此过程中，资本作为产业发展的加速器发挥了重要的作用。一方面，跨国科技巨头以资本为杠杆，展开投资并购活动，得以不断完善产业链布局；另一方面，各类资本对初创型企业的支持，使得优秀的技术型公司迅速脱颖而出。据美国技术研究公司Venture Scanner的调查报告显示，到2017年12月，全球范围内总计2075家与人工智能技术有关公司的融资总额达到65亿美元。同时，美国行业研究公司CB Insight公布了对美国人工智能初创企业的调查结果，这类企业的融资金额约是2012年的10倍。目前，人工智能已在智能机器人、无人机、金融、医疗、安防、驾驶、搜索、教育等领域得到了较为广泛的应用。

3. 新一代人工智能的主要发展特征

在数据、运算能力、算法模型、多元应用的共同驱动下，人工智能的定义正从用计算机模拟人类智能演进到协助引导提升人类智能，通过推动机器、人与网络相互连接融合，更为密切地融入人类生产生活，从辅助性设备和工具进化为协同互动的助手和伙伴。

1) 大数据成为人工智能持续快速发展的基石

随着新一代信息技术的快速发展，计算能力、数据处理能力和处理速度实现了大幅提升，机器学习算法快速演进，大数据的价值得以展现。与早期基于推理的人工智能不同，新一代人工智能是由大数据驱动的，通过给定的学习框架，不断根据当前设置及环境信息修改、更新参数，具有高度的自主性。例如，在输入30万张人类对弈棋谱并经过3000万次的自我对弈后，人工智能AlphaGo具备了媲美顶尖棋手的棋力。随着智能终端和传感器的快速普及，海量数据快速累积，基于大数据的人工智能也因此获得了持续快速发展的动力来源。

2) 文本、图像、语音等信息实现跨媒体交互

当前，计算机图像识别、语音识别和自然语言处理等技术在准确率及效率方面取得了明显进步，并成功应用在无人驾驶、智能搜索等垂直行业。与此同时，随着互联网、智能终端的不断发展，多媒体数据呈现爆炸式增长，并以网络为载体在用户之间实时、动态传播，文本、图像、语音、视频等信息突破了各自属性的局限，实现跨媒体交互，智能化搜索、个性化推荐的需求进一步释放。未来人工智能将逐步向人类智能靠近，模仿人类综合利用视觉、语言、听觉等感知信息，实现识别、推理、设计、创作、预测等功能。

3) 基于网络的群体智能技术开始萌芽

随着互联网、云计算、物联网等新一代信息技术的快速应用及普及,大数据不断累积,深度学习及强化学习等算法不断优化,人工智能研究的焦点,已从单纯用计算机模拟人类智能打造具有感知智能及认知智能的单个智能体,向打造多智能体协同的群体智能转变。群体智能充分体现了“通盘考虑、统筹优化”的思想,具有去中心化、自愈性强和信息共享高效等优点,相关的群体智能技术已经开始萌芽并成为研究热点。例如,我国研究开发了固定翼无人机智能集群系统,并于2017年6月实现了119架无人机的集群飞行。

4) 自主智能系统成为新兴发展方向

在长期以来的人工智能发展历程中,对仿生学的结合和关注始终是其研究的重要方向。例如,美国军方曾经研制了机器骡,各国科研机构研制了一系列人形机器人,但均受技术水平的制约和应用场景的局限,没有在大规模应用推广方面获得显著突破。当前,随着生产制造智能化改造升级的需求日益凸显,通过嵌入智能系统对现有的机械设备进行改造升级成为更加务实的选择,也是中国“制造2025”、德国“工业4.0”、美国“工业互联网”等国家级战略的核心举措。在此引导下,自主智能系统正成为人工智能的重要发展及应用方向。例如,沈阳机床以i5智能机床为核心,打造了若干智能工厂,实现了“设备互联、数据互换、过程互动、产业互融”的智能制造模式。

5) 人机协同正在催生新型混合智能形态

人类智能在感知、推理、归纳和学习等方面具有机器智能无法比拟的优势,机器智能则在搜索、计算、存储、优化等方面领先于人类智能,两种智能具有很强的互补性。人与计算机协同,互相取长补短将形成一种新的 $1+1>2$ 的增强型智能,也就是混合智能,这种智能是一种双向闭环系统,既包含人,又包含机器组件。其中人可以接收机器的信息,机器也可以读取人的信号,两者相互作用,互相促进。在此背景下,人工智能的根本目标已经演进为提高人类智力活动能力,更智能地陪伴人类完成复杂多变的任务。

5.1.3 我国电力企业人工智能应用工程概述

1. 国内人工智能应用的发展现状

我国人工智能研究起步较晚,研究水平与发达国家存在一定差距,但近年来,国内人工智能产业得到了快速发展,形成了以基础技术支撑、人工智能技术及人工智能应用为主体的生态圈。

(1) 基础技术支撑层主要由硬件、数据库和运算平台组成。国内处于基础技术支撑层的企业主要有百度、阿里巴巴等IT巨头及少量创业公司,业务范围主要集中在数据工厂和超算平台建设方面。

(2) 人工智能技术层旨在建立基于不同算法的模型,形成可供应用的有效技术。目前,国内人工智能技术层的优势领域主要集中在智能识别领域,这一领域主要以创业公司

为主，FACE++、小i机器人、科大讯飞、格灵深瞳等企业的视觉、语音识别技术处于国际领先水平。

(3) 人工智能应用层利用技术层输出的人工智能技术为用户提供智能化的服务和产品。目前，此领域参与企业众多，以百度为首的IT巨头依托现有海量数据资源和强大数据处理能力，加大深度学习等核心算法和其他人工智能技术的研究力度，企图开拓人工智能应用市场；以科大讯飞、FACE++为首的创业新贵，凭借自身技术优势，从技术层发力，开拓认知智能应用市场；以海尔、格力为首的传统制造业龙头企业加快人工智能技术研发和引进，力争实现从制造到“智造”的转变。

2. 电力行业人工智能的发展现状

人工智能在国内电力行业的应用，总体来讲可以归纳为传统方式的智能化改进和关键技术的延展与创新，主要体现为管理方式的升级和业务模式的创新。

1) 专家系统

人工智能在国内电力行业的应用较多的是专家系统，主要用于故障诊断、规划调度与运行控制等领域，多面向系统而非装置。

2) 负载和可再生能源预测

电力行业需要重点关注的是每时每刻保持电力生产和消费的平衡。大量风能和太阳能等可再生能源极不稳定，对电力生产调度提出了极大的考验，因此精准预测地区负载以及可再生能源发电量对如何调度发电至关重要。利用人工智能技术对大量历史数据进行建模，可以有效提高预测的准确度，保持电力生产和消费的平衡。

3) 智能设备维护

电力行业大量的电表、变压器及变电站设备需要维护，此类设备的稳定可靠对电网安全运行相当重要。一旦失效，轻则造成停电，重则可能引发火灾。基于大量变压器历史维护记录和其他相关数据，通过人工智能技术训练设备健康度回归模型，可有效监测发现设备运行异常情况，及时预警，减少电网重大事故的发生概率。

4) 核电机机器人

在核电领域，为了降低工作人员受辐射剂量和劳动强度，采用机器人进行设备检修、放射性废物处理、应急响应等工作。目前，核电机机器人大致分为观察型和作业型。“观察型”，指携带摄像头、温度和压力传感器以及辐射强度检测仪等进入后传回现场数据。“作业型”的工作内容包括切割、搬运放射物质、关阀门、喷水等。

3. 集团企业人工智能的发展现状

重视语音识别、自然语言理解、计算机视觉、深度学习等人工智能技术在电网中的科研和应用，目前已在智能电网输、变、配、用和公司经营管理领域研发和应用了多款产品，涉及智能运检、智能配网运维、智慧客服、智能营业厅、智能家居以及可信身份认证等领域。

智能运检以机器人、智能穿戴设备、无人机为系统核心,整合计算机视觉、自然语言处理、机器学习等前沿人工智能技术,实现变电站、电气设备、输电线路全天候、全方位和全自主运检作业、关键数据采集、微气象数据收集、数据智能分析及展现、视频监控及设备安全隐患预警等服务,目前已用于天津市检修公司数字化变电站运维管理系统建设和天津城西供电公司变电站运检智能化管控平台建设。

智能配网运维平台融合应用“互联网+”技术、移动4G通信技术、人工智能技术和便携式设备,构筑完整的作业现场安全、质量监督、任务调度、路径规划、身份识别智能管控体系,目前已用于福建三明市等地供电公司。

应用智慧客户服务方案的客服中心智能质检分析系统,目前已在国网客服中心上线使用,提升了座席人员的工作效率,减少了客户的等待时间。

智能营业厅机器人系统及设备已在国网晋江供电公司智能营业厅和国网石狮供电公司智能营业厅投入使用,有效减少了营业大厅工作人员的工作量,节约了人力成本,并有效缓解了营业大厅的排队现象。

互联网可信身份认证签到系统在国网客服中心南方分中心“精准服务,精彩互动”会议上首次应用成功,参会人员可通过生物识别、人证比对的方式到会登记,使会议组织高效便捷。

同时,近年来,集团在大数据分析处理领域进展显著,自主研发了多项业界领先的大数据平台和产品,并完成了一些业务应用。

4. 问题与分析

近年来,随着新能源持续高比例运行、电力电子装置大量应用、电力市场化水平不断提高,电力系统的动态非线性、多时间尺度、不确定性和难预测性表现得更加突出,基于物理模型的分析方法,难以精准地刻画电力系统特征,以人工智能为核心的数据驱动方法将发挥越来越重要的作用。同时,为了应对上述变化,集团为国网公司部署了众多用于监测、控制、管理的信息通信系统,产生着大量的数据,在很多应用场景下表现出大维度、小样本、非结构化的数据特性,常规的统计分析、数据挖掘、机器学习,难以满足需求。

针对上述问题,新一代人工智能的应用,将为集团提供新的解决方案,以应对电网发展带来的挑战。目前,集团在人工智能产业处于初步发展阶段,存在着技术协同创新不足、产业技术基础薄弱、行业应用不够深入等问题和矛盾,导致技术与市场的“孤岛”和“碎片化”情况严重。

1) 产业基础薄弱,缺乏核心技术

集团在人工智能方面的高端人才比较欠缺,人才引进培育难度较大,导致在基础理论方面研究较为薄弱,缺乏具有自主知识产权的核心算法模型。在配用电、电网安全与控制、企业经营管理等领域,如智能客服、可信认证、舆情分析等,业务研究不深,很多核心技术极其依赖外部单位,甚至还未拥有核心技术产品。

2) 行业应用不深入, 效果提示不明显

集团人工智能技术产品在电力系统中的应用不够广泛和深入, 尚未涉及电力核心领域, 难以对电力主营业务和经营管理产生明显和深远的影响。比如输变电领域, 关于线损的监测、分析、反制等场景, 若能基于新一代人工智能技术研发出有所突破的产品和应用, 将带来巨大收益。

3) 产品线较单一, 未能形成规模

在配用电领域, 集团主要以智能客服机器人、智能家居、电动汽车等应用为主。产品线较为单一, 未能形成一定规模, 未能覆盖配用电全业务领域。集团可与高校等科研院所合作, 共同研发核心业务技术和规模化的产品。

5.2 电力人工智能咨询工程实用方法

输变电系统人工智能工程方法、配用电人工智能工程应用方法、电网安全与新能源人工智能工程应用方法、企业经营管理人工智能应用工程方法是本节介绍的主要内容。

5.2.1 输变电系统人工智能工程方法

1. 输电线路智能巡视

1) 目标

通过人工智能技术的研发与应用, 对生导线断股、绝缘子破损、销钉缺失等缺陷进行智能识别, 对树障、鸟害、金具浮放电、绝缘子污闪等安全隐患进行智能检测。快速实现输电线路的故障定位和故障类型识别, 以便及时准确地消除故障, 确保电力系统安全稳定运行, 提高巡检自动化水平, 降低运维成本。

2) 内容

在本解决方案里, 将重点发展四类硬件产品, 即架空线路智能巡检无人机、架空输电线路巡检机器人、电力隧道巡检机器人、智能眼镜/智能安全帽/智能腕表。

架空线路智能巡检无人机: 研究开发架空线路智能巡检无人机, 实现输电杆塔、导地线、通道走廊各部位的可见光图像、红外图像数据采集, 在无人机上直接完成对巡检目标的智能图像分析及故障诊断, 给出巡检结论, 而无需将视频、图像数据回传至服务器再分析, 大幅提升巡检工作效率。

架空输电线路巡检机器人: 开展机器人接触式巡检作业研究, 打造架空输电线路巡检机器人, 利用携带的传感仪器对杆塔、导线及避雷线、绝缘子、线路金具、线路通道等实施接近检测, 辅助智能带电检修, 提高输电线路巡检效率和检修质量。

电力隧道巡检机器人：基于精确定位技术、自动避障、路径规划等相关技术，打造电力隧道智能巡检机器人，通过机器人携带的可见光、红外线摄像机采集图像视频，结合人工智能技术，实现线路定点巡检、故障缺陷识别；对机器人采集的红外图谱数据进行智能分析与诊断，对隧道巡检过程中发现的异常数据进行实时定位报警。

用于智能巡检的可穿戴设备：开展服务于巡检业务的智能可穿戴设备的研发与应用，打造智能眼镜、智能安全帽、智能腕表等产品，实现智能语音交互、设备自动识别、设备台账自动分析等功能，提高巡检效率。

2. 变电站智能巡检

1) 目标

通过机器人、在线检测装置、带电检测仪器等智能设备在变电站的应用，实现复杂设备的部件部位识别、故障智能诊断等应用，将一线巡检人员从设备巡检层面解脱出来，缓解变电运维绝对缺员及结构性缺员带来的运维压力，切实提高工作效率和质量，真正起到减员增效的作用，更快地推进变电站无人值守的进程。

2) 内容

在本解决方案里，将重点发展两款硬件产品，即变电站巡检机器人和带电作业机器人视觉伺服控制模块。

变电站巡检机器人：以智能巡检机器人为系统核心，开展变电站智能巡检应用研发，整合前沿智能技术，实现变电站全天候、全方位和全自主的巡检作业、关键数据采集、微气象数据收集、数据智能分析及展现、视频监控及设备安全隐患预警等服务。

带电作业机器人视觉伺服控制模块：掌握双目立体视觉技术，结合人工智能算法，利用双目视觉得到设备的准确位置信息；机器人控制单元接收此位置信息，实现高压带电作业机器人的视觉伺服控制。

3. 电网智能规划、研判与故障诊断

1) 目标

通过人工智能手段，实现对于电网规划的智能决策支撑，实现自动化的电气设备运行状况分析和热故障诊断，实现对电气设备的热状态进行监控，并根据热状态进行分析和诊断，保障设备和电网可靠运行；为电网运检过程管控、故障图像研判等提供深层次的价值信息，为运检决策提供重要的平台、技术及数据支撑；为专业技术人员的深度分析和处置提供强有力的技术手段。

2) 内容

在本解决方案里，将重点发展4个系统，即电网规划智能决策支撑系统、输变电设备智能综合诊断系统、电网灾害智能预报预警系统和输变电智能视频识别监控系统。

电网规划智能决策支撑系统：采用深度学习技术、迁移学习技术，基于电网历史规划数据及电网管理、运行大数据，打造电网规划智能决策支撑系统，对电网规划进行全面、

深入的学习与训练，实现对于电网规划的智能决策支撑，有效优化能源布局和电网运行。

输变电设备智能综合诊断系统：针对不同设备运行特点，掌握不同数据收集机制，研发基于电气设备可见光、红外、紫外图像的设备故障识别、问题研判技术；掌握基于多源异构数据的输变电设备故障智能诊断技术；结合电气设备历史维护记录，对设备运行状况及健康状况进行分析，预测设备可能发生的故障点及故障类型；对变电设备历史维修记录分析、汇总，推送故障处理建议方案，辅助运维人员对故障设备进行维修；对设备故障原因统计分析汇总，进行故障规律关联分析，对变电设备重复故障进行预防维修，为运检决策提供重要的技术及数据支撑；结合变电设备历史运行工况，诊断设备寿命，合理安排设备维修、更换策略；掌握基于知识图谱构建和推理的线路通道隐患智能识别技术和通道环境风险评估技术。

电网灾害智能预报预警系统：基于多类传感数据、跨媒体分析推理技术及深度学习网络，综合考虑气象参数（覆冰、降水、风速、风向），还有地形地貌特征、线路结构等，掌握影响电网主要灾害特征的主导因素及灾害发生机理；基于小样本深度学习技术，提升电网主要灾害特征分析及精确预警能力；完善基于气象—监测—线路结构—灾害发生—破坏程度等环节的一体化灾害预警模式。

输变电智能视频识别监控系统：利用深度学习、多媒体处理技术，打造输变电智能视频识别监控系统，实现对输变电场景监控视图的智能分析，建立可持续扩展的视频图像库、标签库，对视频图像中变压器、刀闸、杆塔、绝缘子、杆塔鸟巢、金具锈蚀、线路树障等典型设备及隐患进行自动标签分类，对视频图像中大型物体靠近、人员进出等典型事件进行智能识别。对重要输电通道运行状态进行实时监测、监视预警、分析诊断和评估预测，提升输电专业生产运行管理精益化水平。

5.2.2 配用电人工智能工程应用方法

1. 智能营业厅

1) 目标

智能供电营业厅是在营业厅标准化建设的基础上，利用人工智能技术优化营业厅的软硬件系统，改变传统营业厅的单向被动服务模式，以主动服务、互动服务、智能服务和全时服务为设计原则，为用电客户提供一个智能化、互动化、人性化、24小时全天候的新型用电服务营业厅。

2) 内容

智能营业厅能够承载的业务可以包括：智能服务机器人，负责辅助业务办理以及业务引导；业务智能预办理，基于影像识别和签名识别技术实现证件、资质材料、审批材料等资料的采集提交，以及填单无纸化；大客户智能交互，通过智能茶几、互动终端等为大客户提供智能供电方案、业务查询和用能建议；智能互动展厅，基于增强现实为客户提供互

动创意体验，如新技术、新产品的展示和试用。在本解决方案里，将重点发展一款硬件产品，即营业厅服务机器人。

营业厅服务机器人：营业厅服务机器人内部集成语音识别、语音合成技术，实现机器人与客户直接的语音交互，利用人脸识别技术，读取客户面部信息，实现对来客的识别，并自动关联到其用电记录等多类信息。通过机器人的交互能力，实现宾客迎接、业务查询，并协助用户进行事务处理和业务办理。

2. 智能化配电

1) 目标

通过深度学习、强化学习、智能控制等人工智能技术手段，实现对配电网运行状态在线分析，对配网故障做出有效预警，对配网故障进行智能分析和研判，并根据故障问题进行及时的自动带电作业。

2) 内容

在本解决方案里，将重点发展一款软件产品及一款硬件产品，分别是配网运行状态在线分析及故障预警系统和配网带电作业机器人。

配网运行状态在线分析及故障预警系统：掌握基于智能算法的配电网运行状态分析技术，并结合故障特征、外部环境数据与运维经验的关联特性，掌握故障原因自动分析和推送的关联规则技术，掌握包含故障录波数据、故障原因、故障位置、故障处理方案等信息的配电网故障样本库知识建模方法；突破配电网故障原因知识推理技术，掌握利用故障样本库的知识推理实现在线故障原因快速分析的方法；构建配电网故障线路和设备信息库，建立面向故障预测的深度神经网络及在线增强学习模型；掌握基于历史数据、实时波形、高频模型交叉关联性分析并结合长短期记忆网络的配电网故障智能化预测技术；掌握基于深度学习的配电网运行态势构建技术，通过拓扑、量测、停电、电气参数等电网运行数据的大规模训练，构建配电网运行可靠性、供电能力、转供能力、负荷需求等综合态势。

配网带电作业机器人：掌握基于图像分割、图像分析的配网故障分析技术，掌握基于智能推理的配网故障研判技术，掌握基于智能控制的机器人控制技术，并基于上述技术研发成果研制配网带电作业机器人，实现对于配网故障的智能识别、问题研判，并依据不同故障问题进行有针对性的带电作业，将故障及时解决。

3. 智慧客服

1) 目标

实现智能互动服务与人工客服无缝衔接，为客户提供更加人性化的交互体验，提升服务质量与竞争力，提升服务效率。

2) 内容

在本解决方案里，将重点发展两款软件产品，即人工客服应答智能辅助系统和多渠道客服智能化应答系统。

人工客服应答智能辅助系统：利用语音识别技术实时转写座席和客户对话，为座席及其班组长实时展示客服双方对话内容。同时，利用自然语言理解技术提取客户通话过程中的关键信息，结合知识库，实时给出座席回复话术建议。在整个客户服务过程中，对座席的服务进行实时监督与检查。发现问题后，及时通知座席人员修正，并通知客服班组长注意。

多渠道客服智能化应答系统：通过自然语言理解、语音识别、语音合成等多项人工智能技术的整合，构建全局知识共享的智能知识库，实现知识管理流程规范化、知识处理能力一体化、知识应用服务互动化，为客服专员提供精准、高效的知识检索服务，场景化的知识推送服务，有效提升服务响应速度及服务质量，并能基于海量客服数据分析客户热点诉求。

4. 智能化用电服务

1) 目标

利用移动互联、物联网、云计算等多种技术手段，打造智能用电互动服务平台。通过智能家居、智能小区建设和运营，形成规模化用户资源，借助国家相关部委、电网企业在需求响应方面的政策和服务支持，引导居民用户积极参与电网调峰运行，实现电网的移峰填谷和有序用电，促进社会资源的合理利用。

2) 内容

在本解决方案里，将重点发展一款硬件产品，即家居服务机器人，以及两款软件产品，即用电能效智能管理系统、需求侧智能管理系统。

家居服务机器人：打造家居服务机器人，采用语音识别技术和自然语言处理技术，与用户进行友好的互动，实现对家用电器的智能控制、对于家庭电器的巡视、对于日常生活的辅助等。

用电能效智能管理系统：掌握自学习、自感知的家居、楼宇、园区智能用电服务技术，打造智能用电服务系统，通过收集不同用户的用电信息和电器的使用情况，挖掘用户的用电行为和最优用电模型，为用户提供多种用电服务和增值服务。

需求侧智能管理系统：掌握需求侧错峰资源优化管理知识库、集群行为特征分析技术，以及基于递归神经网络算法的资源集群优化决策算法；掌握基于贝叶斯网络随机学习算法的需求侧资源潜力预测模型和方法；掌握电动汽车行为分析方法，形成用户、车辆、充电设施及电网间复杂时变拓扑及交互的多智能体模型；综合气象条件、空间信息、社会活动、用户偏好等大数据，掌握含电动汽车、分布式新能源与移动式储能等源荷设备的供需策略与风险管控技术，建立兼顾交通状况、充电用户体验和电网均衡运行的智能充电协调控制机制；打造并应用需求侧智能管理系统。

5. 智能线损监测与防窃电

1) 目标

利用物联网、移动互联等技术对配用电领域的线损进行实时监测，并运用人工智能技

术，对配用电领域的电能线损进行实时分析，及时判断线损产生的原因。一旦判定线损是由窃电行为造成的，及时发出警告，提醒相关部门及时处理窃电事故，同时对窃电处进行智能断电。

2) 内容

在本解决方案里，将重点研发一款软硬件结合的系统，即配电线损智能监控系统。

配电线损智能监控系统：打造并推广配用电线损监控系统，采用非侵入式方法实现对线损的实时监测；同时，运用深度学习等人工智能技术，对测得的线损数据进行智能分析，研判是否为窃电行为。当系统判定线损由窃电行为造成时，将及时向相关管理部门发出警报提醒，同时实施智能断电，及时阻止窃电行为。

6. 充电桩选址智能决策

1) 目标

通过新能源汽车的行驶轨迹以及时域信息，采用机器学习方法，预测新能源汽车高利用率的充电桩位置，在减少建设成本的同时，为新能源汽车用户提供更加有效、便捷的充电服务。

2) 内容

在本解决方案里，将重点发展一款软件产品，即充电桩选址智能决策系统。

充电桩选址智能决策系统：打造充电桩选址智能决策系统，通过和新能源车企合作，获取新能源汽车的行驶轨迹以及对应的时间分布信息，结合车辆里程信息，使用深度学习技术，训练优化的充电桩充电位置预测模型，利用该模型预测服务新能源汽车的高利用率的充电桩优选位置。

5.2.3 电网安全与新能源人工智能工程应用方法

1. 电网生产与调度智能研判

1) 目标

通过开展基于人工智能的故障分析与研判技术研究，实现对电网运行状态的智能研判与智能决策支撑，实现电网故障的原因查找、快速定位和智能研判，提升电网运行管控质量和水平，有力保障电网安全。

2) 内容

在本解决方案里，将重点发展三款软件产品，即电网运行控制智能决策支撑系统、配网多源信息故障诊断系统和电网设备及元件故障智能研判系统。

电网运行控制智能决策支撑系统：掌握大型交直流混联电网故障特征深度学习和智能识别及控制技术；掌握基于电网运行数据的调控知识发现总体架构和数理分析智能算法，以及覆盖全业务全过程全周期的调控知识模型构建技术；掌握电网运行操作、计划调整、

自动控制等智能化技术，以及电力系统故障及薄弱点智能识别技术；掌握电力系统源网荷一体化智能协调控制技术；打造并应用电网运行控制智能决策支撑系统。

配网多源信息故障诊断系统：基于电力生产管理系统、调度管理系统以及地调能量管理系统海量数据，以事故分闸为主线获取关联的遥信变位、遥测、保护、遥控、操作票等辅助数据，基于这些数据训练深层神经网络模型，及时给出故障预判提示，实现配网故障的高效准确处理。同时，结合线路过负荷、运行操作、检修试验、雷电、气象、重大活动和节假日等内外部数据进行关联分析和推理，为配网运维检修策略的制定提供参考依据。

电网设备及元件故障智能研判系统：打造电网设备及元件故障智能研判系统，对于重合闸不成功的电网故障，定位故障设备，对于重合闸成功的电网故障，识别和定位电压暂降源并分析电压暂降规律，掌握基于母线故障的电气量、线路故障的电气量、保护/断路器和PMS/OMS为判据的电网故障设备定位以及电网故障影响范围研判技术，分析不同类型暂降事件的波形特征，对电压暂降海量数据进行波形特征匹配判断。

2. 新能源发电预测

1) 目标

通过引入人工智能预测技术，利用风能、太阳能时域、空域分布数据以及用户用电量时域分布信息有效预测风能、太阳能发电状况，提升新能源发电效率，降低发电成本。

2) 内容

在本解决方案里，将重点发展二款软件产品，即风能发电预测系统和太阳能发电预测系统。

风能发电预测系统：基于电厂过往的风能时域分布信息，以及对应的用户用电量信息，借助深度神经网络等机器学习方法，训练风能效率预测模型，并对小区域气象变化进行准确预测，打造高精度的预测系统，对风能发电效能进行预判，指导风能发电效能分配，提高发用电效率。

太阳能发电预测系统：根据历史发电信息、光照情况、天气情况等，采用深度学习等智能技术，建立针对电源具体情况的太阳能发电情况预测模型，并采用图像增强、图像识别技术基于云图进行光伏功率快速波动的准确预测，打造太阳能发电智能预测系统，对太阳能发电出力状况进行全面的预测。

3. 新能源消纳分析及研判

1) 目标

采用深度学习、主因子分析等人工智能技术，实现对新能源消纳能力的分析与评价，并依据态势进行分级预警，提升新能源接纳能力智能化在线评估水平，全方位掌握新能源消纳状况。

2) 内容

在本解决方案里，将重点发展一款软件产品，即新能源消纳能力评估及预警系统。

新能源消纳能力评估及预警系统：基于大量历史运行数据和复杂场景消纳能力算例结果，掌握不同运行场景的新能源消纳特征提取技术，挖掘新能源消纳的关键影响因子；建立新能源消纳智能评估模型，掌握基于深度学习算法和区域新能源消纳规律的在线新能源消纳评估技术；基于深度学习挖掘新能源消纳复合影响因子，打造新能源消纳能力评估及预警系统。

5.2.4 企业经营管理人工智能应用工程方法

1. 信息安全预警

1) 目标

通过过往大量的安全攻击记录数据，利用机器学习方法训练攻击预测模型，防范第三方人工智能的破解，实现自我学习，持续优化提升信息系统的智能安全化程度，降低安全隐患。

2) 内容

在本解决方案里，将重点发展一款软件产品，即信息安全智能预警系统。

信息安全智能预警系统：通过获取包括用户数据、攻击数据和各种威胁情报数据，使用深度神经网络等机器学习方法，训练网络攻击分析与研判模型，基于这些模型打造信息安全智能预警系统，对各类安全威胁进行有效监测和预警。

2. 人员身份认证

1) 目标

基于生物识别认证技术的用户认证可同时解决身份可靠认证、访问控制和直观追溯抗抵赖的安全防范技术，既可增强业务系统身份信息安全的主动防御能力，又可提升国网公司整体信息安全水平和安全管控；同时实现对办公大楼出入口人脸抓拍、布控、识别，增强国网公司安防能力。

2) 内容

在本解决方案里，将重点发展一款软件产品，即面向电网业务的可信身份认证系统。

面向电网业务的可信身份认证系统：利用人脸识别、生物特征识别技术，实现与企业门户、综合运维审计、安全系统的用户身份认证集成应用；实现作业现场人员身份认证、信息机房人脸识别门禁、外协人员人脸识别考勤管理等应用。

3. 人才智能招聘

1) 目标

利用人工智能技术，搭建智能化招聘平台，实现人才自动筛选、测评，并进行精准的简历推荐服务。辅以大数据分析，避免现有“人工”招聘模式的缺点，减少招聘HR

劳力费时的简历筛选和初面环节，提高招聘效率，更快、更好地为企业找到所需的合适“人才”。

2) 内容

在本解决方案里，将重点发展一款软件产品，即智能人才筛选及人岗匹配系统。

智能人才筛选及人岗匹配系统：借助互联网海量的简历，利用深度神经网络学习算法和个体画像技术，构建智能人才筛选及人岗匹配系统，借助专业招聘网络和社交网站，获取公开可见的人员简历信息，建立人员筛选模型，对人才进行智能筛选，并建立人岗匹配模型，从多个维度对人才和岗位进行自动匹配和智能推荐。

4. 舆情智能管理

1) 目标

通过人工智能关键技术，使舆情分析得以智能化、自动化，而不是只是单靠人工来完成，从而让舆情工作人员的工作效率大幅提升。利用舆情分析产品，进行企业形象分析，话题用户参与度情况分析以及风险管控。

2) 内容

在本解决方案里，将重点发展一款软件产品，即智能舆情管理系统。

智能舆情管理系统：利用先进的爬虫技术，爬取以新闻、论坛、微博为主的与国家电网相关的互联网文章，将其作为舆情分析数据。通过“文本摘要提取”技术，得到国网相关的主题结构。建立文档主题生成模型（LDA）、浅层语义分析（LSA）、概率浅层语义分析（PLSA）等关键技术主题模型，从而得到公司相关的关键主题词。通过卷积神经网络等深度学习技术，对每个主题的网民情感倾向性进行分析。通过社交网络的图计算关键技术，发现舆情的传播途径以及关键的意见领袖，准确把握传播核心和关键用户。

5. 企业智能经营支撑

1) 目标

利用人工智能技术，实现对企业经营状况的智能研判及预测，对企业各项投资的智能研判，有效支撑企业的经营决策和活动。

2) 内容

在本解决方案里，将重点发展两款软件产品，即企业经营状况智能分析与预测系统和企业投资智能决策支撑系统。

企业经营状况智能分析与预测系统：掌握基于神经网络的企业物资管理、廉政建设、法律事务、企业文化等内部管理风险智能识别技术，实现企业经营管理风险的全方位跟踪预警，打造采用人工智能算法、基于企业生产状况、营销状况、财务状况等多方面信息的企业经营状况智能分析与预测系统，实现对企业各方面状况的全面、深入研判和准确预测。

企业投资智能决策支撑系统：打造企业投资智能决策支撑系统，掌握基于会计税务政

策、人工成本、利率汇率、政府规划等复杂环境下的投资目标知识图谱技术；掌握跨公司跨行业的数据动态关联技术，增强企业对投资目标未来基本面的量化见解能力；掌握基于深度学习的高精度产品/业务收入预测技术，实现从信息获取到投资研判的自动跨域，解决传统人工投研的干扰因素多、主观性大、效率较低、准确性不足等问题。

5.3 电力企业人工智能技术应用案例分析

电网智能运检人工智能技术应用、电网运行人工智能工程应用分析、企业管理人工智能工程应用分析、供电服务人工智能工程应用分析是本节介绍的主要内容。

5.3.1 电网智能运检人工智能技术应用

1. 变电设备智能识别与故障诊断

1) 应用简介

目前，通过机器人巡检、变电统一监控获取表计、开关/刀闸开合状态、油位计等可见光图像、红外图谱、声音检测数据，代替了传统的人工现场巡视相关工作。基于采集的图像、图谱、声音数据，应用人工智能技术，先后实现了刀闸和开关状态检测、仪表自动读数、异物悬挂及设备外观异常检测、变压器声音异常检测、温度异常检测等变电站设备状态检测应用场景。

在变电站巡检中，通过监控视频、巡检机器人对设备状态感知及感知数据的智能化识别分析，及时发现电网运行的事故隐患与故障，提高了变电站的自动化及智能化应用水平。

2) 设计实现

通过分析变电站内采集的红外图像、可见光图像、声音数据，建立不同设备、不同状态的样本库，并根据样本库中不同状态设备的图像或声音特征建立变电站设备故障监测模型，对变电站设备温度情况、外观情况、刀闸及开关状态、仪表示数等通过图像预处理、图像识别、特征提取及机器学习等算法进行自动分析，实现设备状态的实时监测。

3) 基于红外图像的变电站设备故障诊断

红外图像包含设备的温度信息，通过图像的分割和特征提取，建立红外图像特征数据库，并对采集的红外图像中的设备进行分类识别。

(1) 基于图像的设备状态识别。通过可见光图像中不同设备的特征分析，进行刀闸和开关的状态识别、仪表的自动读数、异物及悬挂物的检测，对异常状态的设备进行报警。

(2) 基于声音的变压器状态识别。变压器运行的声音信号很大程度上反映变压器的运行状况,借助高灵敏度的声音采集器和现代数字信号处理技术,通过分析声音信号中的多类特征,设计合理的分类器进行判别,可以实现比人耳更加客观可靠的在线检测和故障诊断。

4) 应用成效

应用统一监控视频、变电站室内/外智能巡检机器人巡检及人工智能识别技术,大大减轻运维人员日常巡视、红外测温、设备工况检查等简单重复性工作,实现简单劳动自动化,避免了运维站至各变电站之间的往返交通占用的大量时间,特巡和事故后巡视的及时性不够。同时,解决了受运维人员技术水平等因素影响,巡视质量参差不齐,容易发生漏查、误判等问题,大大提高了变电运维工作效率与智能化水平。

2. 输电线路智能识别与缺陷诊断

1) 应用简介

目前,应用无人机与巡线机器人实现了输电杆塔、导地线、通道走廊的可见光图像数据采集,在一定程度上减轻了传统的人工巡视工作量。基于采集的可见光图像数据,将人工智能技术应用于电网智能巡检,有效提高了公司对海量巡检数据处理能力,实现了输电线路部件、部位智能识别和输电线路缺陷智能诊断。集成了巡检数据后处理管理平台,对识别出的设备及缺陷进行管理,并形成相应的缺陷诊断报告,从而及时发现输电线路设备的外观异常和热缺陷等隐患,在降低运维成本和提高巡检自动化水平方面有巨大的应用价值。

2) 设计实现

(1) 提高图像数据处理能力。深度学习是图像智能处理的常用手段,训练数据规模庞大,因此计算能力是智能图像分析的基础。首先需在视频图像智能分析基础支撑技术层面做出重大革新,形成满足算法和模型需要的高性能基础支撑平台,进一步提高电网海量巡检图像处理能力。基础支撑平台的建设包括设计面向电网的图像智能分析技术框架,搭建异构计算架构,有效整合现有软件和计算资源,以及数据和标签管理。此平台可为试验和实践最新的人工智能理论和算法提供整套环境。

(2) 巡检图像智能理解。对有人机、无人机与巡线机器人采集的输电线路巡检图像进行预处理和特征提取,在此基础上采用深度学习和传统图像处理技术实现巡检图像智能理解。

(3) 输电线路巡检图像智能分析系统。立足于航拍巡检图像智能分析典型业务应用场景,开发输电线路巡检图像智能分析系统,相当于一个数据后处理管理平台。其中集成了图像智能理解算法核心模块,在实际应用中可与现有的数据采集系统结合,从而实现在直升机/无人机巡检、巡线机器人所采集图像中,对电网设备典型缺陷的智能识别。

3) 应用成效

(1) 输电线路巡检图像场景分类。多角度不同距离下拍摄的图片差异巨大,因此针对单张巡检航拍图像,从整体角度理解拍摄内容,并进行图片拍摄内容的分类,降低图片

的无序性，为缺陷识别提供先验知识。

(2) 输电线路部件、部位智能识别。在输电线路巡检时，判断输电线路中的杆塔金具、绝缘子、导地线在图像上的位置有着极其重要的意义。输电线路缺陷智能诊断当前已实现杆塔异物缺陷检测、导线异物缺陷检测、绝缘子掉片缺陷检测、防震锤移位缺陷检测等缺陷诊断。

3. 基于可穿戴装备的变电站智能巡检专家系统

1) 应用简介

采用自主研发的可穿戴智能巡检主机、高清可见光/红外热成像双光头盔、超小型便携式紫外成像仪、便携式局放检测仪等智能化巡检装备，获取现场巡检高清可见光、红外、紫外图片/图像及局放图谱数据，通过电力4G无线专网（或VPN网）实时传送至电力内网后台服务器，通过基于人工智能技术开发的智能巡检专家诊断系统对巡检数据进行处理、分析及预测，从而有效降低变电站巡检人员的工作量，大幅提升巡检效率和故障检出率，全面支撑设备运行维护和状态检修。

2) 设计实现

系统通过基于智能图像处理的关键变电设备的识别技术体系，实现对可见光、红外、紫外图像中的设备准确识别、定位及融合；通过典型故障诊断专家系统，利用融合方法实现多重故障信息的综合表达，利用深度学习技术实现典型故障的研判及综合诊断。

(1) 变电站主设备识别。基于智能图像处理的变电站主设备识别技术主要分为两部分。首先是通过特征提取技术提取对象的几何、纹理、颜色、空间关系等特征，形成特征表达。比如，通过稀疏编码或视觉单词模型等，对特征进行训练特征子集选择，也就是对特征样本进行特征降维，再训练特征分类器，分类器完成对输入的图像表达特征进行分类，从而达到对设备的识别。

特征提取采用HOG特征或不变矩通过计算经过预处理后剩余图像区域的7个不同特征来构成一组特征；卷积神经网络（CNN）通过卷积层和池化层来提取特征；决策分类采用支持向量机（Support Vector Machine, SVM）或者卷积神经网络的全连接神经网络等。

(2) 基于深度学习的专家诊断模型。基于深度学习的专家诊断模型分为两部分。首先通过智能图像处理技术实现主设备的识别，通过主设备的识别获取设备工况信息，以卷积神经网络（CNN）为诊断核心算法，对获取的设备图像进行状态分类，再融合设备状态信息，调用层级实时记忆（HTM）算法及推理知识，最终得到设备的状态情况和可能的故障诊断信息。

(3) 基于可穿戴装备和深度学习的专家诊断系统。通过智能可穿戴装备完成巡检数据的实时采集、预处理和传输，通过卷积神经网络（CNN）人工智能算法对可见光、红外、紫外图像及局放图谱进行特征提取，通过层级实时记忆（HTM）人工智能算法对带时间标签的特征参数数据进行深度学习，形成设备故障诊断信息和趋势预测，为专业技术

人员的深度分析和处置提供强有力的技术手段。

3) 应用成效

目前,可穿戴巡检装备(可穿戴智能巡检主机、高清可见光/红外热成像双光头盔、超小型便携式紫外成像仪、便携式局放检测仪等)已完成样机研制,并已经过近一年的应用测试,其中可穿戴巡检主机已达量产水平。基于深度学习的专家诊断系统部分功能已在实验室进行测试,但目前还未投入现场实际应用。

2017年1月,已完成500kV南京东善桥变、无锡梅里变、苏州木渎变TD-LTE 4G无线专网基站部署,构建了无线专网络环境,为下一步开展可穿戴巡检装备的现场应用和基于人工智能的专家诊断系统的建立创造了条件。2017年10月,基于深度学习的人工智能专家诊断系统投入实际应用测试,着力通过人工智能技术为变电站主设备的故障诊断和趋势预测开辟一条新思路。

4. 智能变电站安全远程监控

1) 应用简介

目前,国网公司110kV及以上变电站约有1.5万个,大部分虽然已经实现了基于场区安全防护需要的视频监控,但还是完全使用人工远程查看的方式。这种方式从实际应用的价值上来说,仅仅能起到一个辅助查看的作用,应用价值不高。其实并非视频监控的应用方式只能满足人工查看,根本原因还是缺乏人工智能的监控手段,业务价值的发挥被技术支撑能力的短板限制了。

2) 设计实现

针对无人值守变电站远程安全监控业务需求,面向已建变电站摄像头终端设备升级换代及新建变电站摄像头终端设备布点,通过运用高清复眼摄像头产品,保证变电站各个运行部件清晰展示,为监控人员提供清晰、真实、立体化的变电站实时运行场景。

3) 应用成效

根据变电站视频监控数据本地化存储的特征,使用公司非结构化数据存储系列产品,并基于视频浓缩摘要服务,突破原有30天的视频数据存储时限,实现安全事件管理可追溯周期的大幅延伸。采用三维可视化、视频智能识别与分析及物联传感技术,实现设备运行状态与场区环境状态的全时联动,打造智能变电站一体化全景监控防护体系。

5. 基于智能眼镜的智能巡检

1) 应用简介

智慧单兵平台对接生产管理系统,可从生产管理系统数据库获取待办事项及设备台账等数据。智能眼镜与智慧单兵平台实时连接,包含待办事项提醒、设备台账获取、红外测温、调焦拍照、巡视轨迹记录、远程专家等功能,智能眼镜的相关功能都是独立的,用户可随时调用,当用户接收待办事项以后,在使用过程中产生的数据都将与该待办事项绑定,在后台形成记录,供管理人员进行查看。

2) 设计实现

待办事项：智慧单兵平台对接生产管理系统数据，从生产管理系统里获取人员近期的待办事项，并实时下发至智能眼镜上，当到达待办事项时间点时，自动进行提醒。

设备台账获取：智能眼镜可以根据现场情况实时获取设备的台账数据，并通过AR技术实时叠加在智能眼镜的屏幕上，下面介绍智能眼镜获取台账的方式。

(1) 语音识别：用户通过语音输入台账名称，智能眼镜自动将用户的语音转化成设备台账名称文字并传到智慧单兵平台，智慧单兵平台通过设备名称到生产管理系统后台数据库检索设备对应的台账信息并下发至智能眼镜。

(2) 标牌文字识别：现场设备都有对应的标牌信息，标明了设备的名称，智能眼镜通过摄像头智能识别标牌上的文字信息，得到设备名称传到智慧单兵平台，智慧单兵平台通过设备名称到生产管理系统后台数据库检索设备对应的台账信息并下发至智能眼镜。

(3) 设备实物识别：智慧单兵平台对现场设备实物进行建模，智能眼镜通过摄像头拍摄设备实物传到智慧单兵平台，智慧单兵平台根据得到的设备实物照片进行匹配，得到设备名称，通过设备名称到生产管理系统后台数据库检索设备对应的台账信息并下发至智能眼镜。

(4) 红外测温：智能眼镜集成红外成像测温功能，用户可实时获取设备的温度数据，并实时传送到后台，后台设定的温度阈值数据，当温度超出设备的阈值时，及时产生报警并实时下发给智能眼镜。

(5) MR拍照：巡视过程中用户可以调用拍摄功能，拍摄现场照片或者视频并上传至云平台。其中，拍摄照片时，可以通过语音输入缺陷内容并叠加在照片上。

(6) 巡视轨迹记录：智慧单兵平台下发待办事项时自动形成一条巡视轨迹路线，智能眼镜结合GPS信息实时记录人员的位置信息，并上传至智慧单兵平台，智慧单兵平台形成实时巡视轨迹，智慧单兵平台对用户的实时位置信息及巡视轨迹进行监控匹配，当用户位置或者巡视轨迹偏移时产生报警信息并发送给相关管理人员，用户通过智能眼镜拍摄的照片同时结合位置信息叠加于巡视轨迹上。

(7) 远程专家：抢修过程中用户可以发起远程协助请求，专家通过前端第一人称视角对现场做出准确判断，并给出切实可行的指导意见。

3) 应用成效

(1) 2016年在龙岩局试点应用，主要利用智慧单兵平台及智能眼镜进行现场巡视工作以及安全稽查，效果显著，工作效率提高了30%以上。

(2) 2017年在厦门城市共用电检修有限公司进行试点应用，利用智能眼镜进行现场作业，实时回传现场音视频，方便了现场工作监督以及现场工作调配。

6. 基于国网统一视频平台的输变电监控分析

1) 应用简介

基于国网统一视频平台，利用多媒体处理、大数据、深度学习技术，实现对输变电场景监控视图的接入管理及分析，建立了可持续扩展的视频图像库、标签库，对视频图像中

变压器、刀闸、杆塔、绝缘子、杆塔鸟巢、金具锈蚀、线路树障等典型设备及隐患进行自动标签分类,对视频图像中大型物体靠近、人员进出等典型事件进行智能识别,为电网运检过程管控、故障图像研判等提供深层次的价值信息,为运检决策提供重要的平台、技术及数据支撑。

2) 设计实现

视频平台从运检业务层面出发,梳理输变电巡检内容,针对设备、人员、环境等重点关注信息,建立视频图像库、标签库,实现对视频图像及结构化信息的配置管理,基于大数据存储技术实现分布式视频图像的存取,并结合大数据计算、视频图像处理、深度学习技术,搭建分布式视频图像文件处理框架,实现分布式的编解码、图像分析,完成输变电视频图像文件的典型标签自动分类与智能识别,为输变电视频图像深度智能分析提供了重要的平台框架及技术基础,在提供输变电可视化监控的同时,可以为运检管控及分析决策提供重要视频图像价值内容信息。

(1) 存储管理。通过存储管理,可以实现大数据存储环境配置以及与集中存储服务的关联,保证视频图像数据的分布式存储。

(2) 算法管理。通过算法管理功能,实现对智能分析算法包的添加、修改和删除,支持各类分析算法的灵活接入和快速扩展。

(3) 标签管理。通过标签管理功能,可以为输变电视频图像不同业务应用场景提供灵活的标签管理。

(4) 算法与标签关联。通过算法与标签关联功能,实现算法与标签目录中标签类别的关联,实现视图标签的自动标注。

(5) 智能分析任务管理。通过智能分析任务管理功能,实现智能分析任务的管理。输变电图像智能分析任务不仅包括标签分类(变压器、刀闸、杆塔、绝缘子、杆塔鸟巢、金具锈蚀、线路树障、线路违建),还包括人员进出检测、大型物体靠近分析,实现了对典型视频图像文件的分布式结构化分析。

(6) 视图库。通过视图库功能,可以设置查询条件检索大数据存储中的视频图像,对视频图片进行手动标签标注和修改。

3) 应用成效

智能型统一视频平台继承前期建设成果,进一步提升平台实用化水平,利用大数据、深度学习技术,打造智能型统一视频平台,重点实现了对输变电监控视频图像价值数据的高效存储管理、分析挖掘及快速检索。

(1) 通过高效存储管理,实现对电力视频图像数据的集中存储。针对重要视频资源未有效存储的情况,通过建立大数据分布式存储机制,实现对平台重要视频图像的集中存储,建立可持续扩展的电力场景视频图像库和标签库,妥善保存电力视频图像数据资源。

(2) 通过快速检索分析,实现由人工分析检索到智能分析检索的转变。针对视频图像处理中数据、算法、计算三大资源需要,结合大数据、深度学习等技术,对视频图像文件进行结构化分析,实现输变电图像标签分类(变压器、刀闸、杆塔、绝缘子、杆塔鸟

巢、金具锈蚀、线路树障、线路违建），还有人员进出检测、大型物体靠近分析，提升平台对输变电视频图像的快速检索分析能力，保证工程应用成效。

（3）通过提供价值内容，实现对视频图像业务应用支撑能力的全面提升。针对各业务场景智能分析算法的多样性情况，形成的平台存储及分析能力，能够便于后期集成优势分析算法，支撑运检视频图像深度智能应用。

5.3.2 电网运行人工智能工程应用分析

1. 基于多源数据融合的电网故障分析及应用

1) 应用简介

在当前互联大电网下当故障发生时，大量的报警信息会在短时间内涌入调度中心，这远远超过运行人员的处理能力，往往使调度员误判、漏判。为了适应各种简单和复杂事故情况下对故障的快速而准确的识别，需要电网故障诊断系统进行决策参考。同时由于电力系统调度自动化水平不断提高，越来越丰富的报警信息通过各变电所的远方终端，传送到各级电网调度中心，使得利用采集的实时信息进行电力系统故障诊断成为可能。另外，对于电力系统故障的仿真分析和模拟培训也可以通过电力系统故障诊断系统来提升调度员的经验和水平。

2) 设计实现

针对电网实际情况，侧重电网运行信息、设备状态信息、环境监测信息深度融合与综合应用，开展基于多源数据融合的电网故障辅助分析系统的研制。研究内容主要包括三个方面。

（1）多源数据融合处理技术。多线程通用数据采集框架整合了数据采集程序常用的各种基础功能，包含线程管理、业务逻辑接口和公共服务三个模块，如图5-1所示。

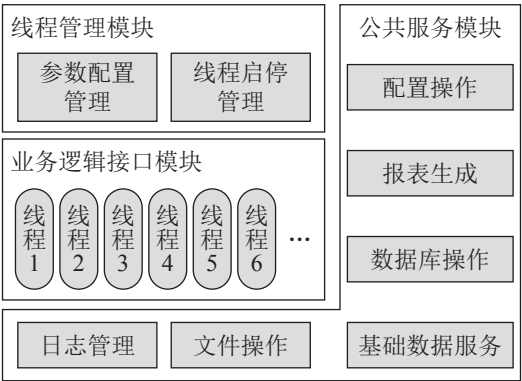


图5-1 多线程通用数据采集框架

（2）利用多源数据的电网故障智能诊断技术。以一种基于时间序列匹配的电力系统故障诊断方法，通过计算警报假说时间序列与实际警报时间序列之间的距离，得出时间序列假说置信度和元件故障置信度，识别出故障元件；以一种综合利用电气量和时序信

息的电力系统故障诊断方法，建立元件的加权模糊时序Petri网模型，综合SCADA信息、WAMS的电气量信息以及这些信息所包含的时序特性，形成警报信息判据、库所延时约束和电气量判据；采用高斯函数，结合时序推理，为Petri网模型置信度矩阵赋初值，通过迭代进行模型求解，得出元件故障置信度，识别出故障元件，提高诊断结果准确性；以一种考虑气象等影响要素的电网故障原因推理分析方法，通过比较故障前后的电力系统拓扑结构，识别停电区域，确定候选故障元件，再根据继电保护和断路器状态，确定故障元件，最后分析导致故障的气象等影响因素，以帮助系统运行人员快速定位和排除故障，尽快恢复系统正常运行。利用多源数据的电网故障智能诊断框架如图5-2所示。

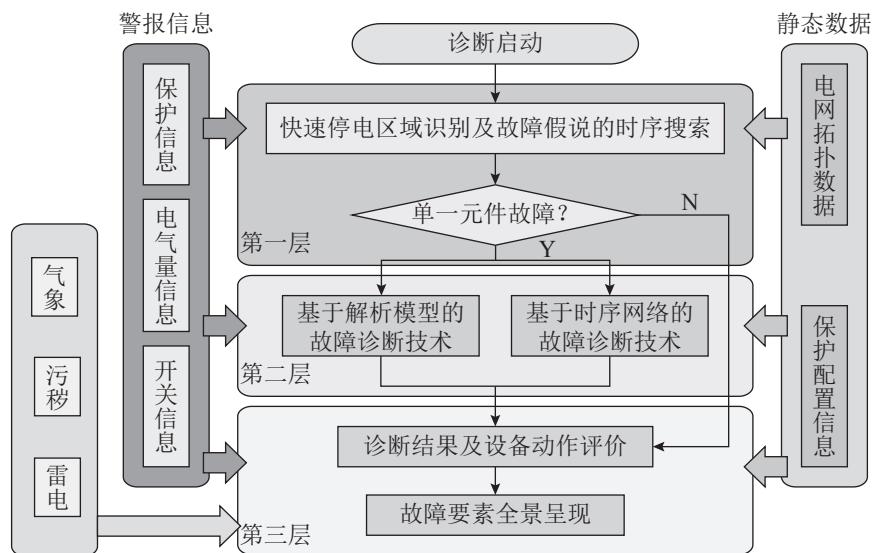


图5-2 利用多源数据的电网故障智能诊断框架

（3）基于规则推理的配网多源信息故障诊断技术。该模块集成了电力生产管理系统（PMS）、调度管理系统（OMS）以及辽宁省13个地调能量管理系统（EMS）海量数据，以事故分闸为主线获取关联的遥信变位、遥测、保护、遥控、操作票等辅助数据，并用基于时序信息的规则给出预判提示，简化了配网故障分析的流程，实现了配网故障的高效准确处理。此外，在本系统积累的配网故障信息基础之上，结合线路过负荷、运行操作、检修试验、雷电、气象、重大活动和节假日等内外部数据开展大数据关联分析研究，为配网运维检修策略的制定提供参考依据。

3) 应用成效

针对电网故障诊断预警应用特点，实现了电网运行信息、设备检修信息、气象环境信息的深度融合与综合应用，为电网企业生产、运行部门高效开展电网故障诊断、配网事故跳闸在线分析统计、基于机器学习分析的电网故障预警、故障应急处理以及专业闭环管理工作提供了重要的技术支撑手段。项目研究成果直接服务于电网生产，有助于提升电网安全稳定运行水平，提高电网供电可靠性，具有显著的经济和社会效益。

2. 配网站房智能综合监控

1) 应用简介

针对配电网的开关站、配电房、环网柜、分支箱、柱上开关、箱式变压器等配网运行相关环节,运用视频监控、环境监测、门禁管理、智能控制、红外热成像测温、智能分析、通信传输等技术,实现对配电网现场图像及各种环境信息采集、分析、监控、控制管理等功能。

2) 设计实现

通过与配电自动化系统的信息交互,实现配电遥控可视化联动。通过智能化管理平台,结合前端视频采集设备,实现配电站点远程智能巡视。采用先进的图像识别技术,基于丰富的标准模型库,实现图像采集终端与图像识别分析单元的智能交互,为视频监控装上永不疲倦的“大脑”,既可作为独立系统部署应用,也可作为图像智能分析核心功能模块与输电线路远程巡检云平台等系统集成,可提供目标图像识别、行为分析、边界防护、流量统计、视频质量自诊断等功能。

3) 应用成效

通过智能化管理平台,结合前端视频采集设备,实现配电站点远程智能巡视。

5.3.3 企业管理人工智能工程应用分析

1. 国网公司文档自动归类技术

1) 应用简介

应用深度神经网络以及统计学模型,实现了一种可用于在线实时文档自动归类的功能模块,主要基于人工对文档的标注,应用循环神经网络、LDA主题模型以及TF-IDF词频信息等进行模型特征提取,通过线下模型学习实现文档自动分类的功能。对输入文档通过模型预测可给出其在各个类别上的概率分布,并对最高概率结果进行输出,对概率分布均匀,即分类结果不明确的样本进行单独输出,后续由人工进行判断。单样本的预测在毫秒级别,精确度在85%以上,整个过程减少了大量的人力,释放了大量的人力资源,同时提升了工作效率。

2) 设计实现

(1) 标注数据。针对文档管理系统设定的文档各个类别,人工针对每个类别的数据进行批量标注,为提升模型准确度以及学习时间,确保各个类别的标注文档的个数几近相等。另外,为确保后续模型学习过程顺利进展,对标注数据进行分词处理,并将结果进行保存。

(2) 应用统计学模型以及深度神经网络模型进行文档类别分布分析。根据文档固有的结构特征以及文档内容,提出一种集合了统计学特征、深度特征以及规则判断三种方法

的预测模型。该模型方案综合考虑了文档词频、题和章节题目信息以及文档内容所具有的深层次的隐含特征等三方面信息，模型学习通过最小化模型目标函数（及模型预测结果与实际结果的误差），求得各个特征的权重，最后进行权重求和，进行文档的类别概率分布预测。

（3）不确定文档单独输出。该功能模块所用模型功能适用于分类，模型实现算法为回归算法，因此可用于判断输入文档属于每个类别的概率（经过归一化处理）。经过信息熵判断，对于结果超过一定阈值的文档我们认为其分类结果不明确，将其进行单独输出，为保证文档归类的准确度，对不明确归类的文档进行后续的人工处理。

3) 应用成效

（1）文档分类效率的提高。人工对文档进行分类的过程，包括打开文档、查看文档、关闭文档三个过程。针对每个文档该过程是一致的，对专业人员每个文档归类的完成大概需要15秒，而且随着重复工作的进行，这个时间会有所提升。

然而，我们的模型对每个文档的归类耗时是在毫秒级别的，其准确度在85%以上，针对不确定文档会进行单独处理，所以该功能模块的引入总体给工作效率带来大幅度提升。

（2）保证文档归类的精确度。人工对文档分类是一项重复的工作，长时间重复工作往往会造成员工的浮躁以及一定程度的误判率提升。该功能模块的引入，可以让相关人员至少减轻85%的工作量，只需要定期对模型不确定文档的输出进行判断即可，其他时间可进行其他工作，保证了相关人员的工作多样性，同时保证了文档分类的精确度。

2. 基于人脸识别的用户身份可靠认证

1) 应用简介

传统用户身份认证使用密码、数字证书等数字认证技术，由于数字认证技术与身份主体的可分离性，使其存在不少问题。

（1）存在防伪性差，易泄露、伪造、盗用、破译等安全隐患问题，同时复杂的密码易忘记。

（2）擅自授意他人利用个人身份信息登录业务系统，私自替代他人执行业务操作等现象难以杜绝。

（3）在盗用、伪造等用户身份信息非法侵入业务系统的情况下，侵入者真实身份难以查询和追溯，造成公司蒙受的物质和经济损失因缺乏强有力的证据而难以有效追回。

（4）在公司重要场所出入管控方面，传统的身份证、工作牌、门禁卡等通行或出入认证方式，存在私自窃取或借用他人身份证件冒名出入机房重地、重要办公场所、核心涉密区域等安全隐患，对公司的人、财、物安全带来严重威胁。

基于人脸识别的用户身份可靠认证作为一种先进的生物识别认证技术手段，是当前同时解决身份可靠认证、访问控制和直观追溯抗抵赖的安全防范技术，对于增强业务系统身份信息安全的主动防御能力，提升公司整体信息安全水平和安全管控具有重要的意义。同时实现对办公大楼出入口人脸抓拍、布控、识别，增强公司安防能力。

2) 设计实现

针对当前身份认证方式中存在的问题,结合先进、科学的安防理念和先进的人脸识别技术,为业务系统提供统一人脸识别身份认证服务。

公共组件主要包括人脸注册组件和人脸识别认证组件。

3) 应用成效

系统在福建省电力公司上线运行,实现了与企业门户、综合运维审计系统的用户身份认证集成应用;实现了信息机房人脸识别门禁应用、外协人员人脸识别考勤应用管理。通过人脸识别身份认证登录企业门户,有效地解决了现有的用户身份认证方式中存在的安全隐患问题;通过与综合运维审计系统的集成,对系统运维人员身份信息和可访问维护信息进行严格验证管控,可有效杜绝私自替代他人执行非法运维操作的现象;通过机房人脸识别门禁应用,对机房出入人员进行识别认证和准入授权管控,有效杜绝私自窃取他人身份证件冒名进入机房重地的现象,消除非法潜入或未经授权许可擅自闯入所构成的潜在安全威胁;通过外协人员人脸识别考勤管理应用,实现对信息化建设厂商人员进行考勤管理,杜绝厂商人员迟到、旷工、早退及随意更换人员等现象,为评价厂商人员用工费用评估及厂商评价提供数据参考依据。

5.3.4 供用电服务人工智能工程应用分析

1. 潜在电子化缴费用户自动化识别

1) 应用简介

通过对低压居民用户的行为分析,建立用户实名制认证的用户(人)与用电户(户)的动态数据映射库,透过用户画像建立人户间的用电行为和缴费行为关系,为后期开展业务精准推荐、信息精准推送、诉求精准预测的延伸性营销服务,唤醒对现有用户数据的使用和价值挖潜,真正实现主动感知、主动服务,解决以往对于服务谁、服务什么、如何服务的业务瓶颈。

2) 设计实现

通过分析低压居民用电信息的基本数据、属性数据、用电数据和外部数据(房价和天气等),建立电子化渠道适应度模型和渠道偏好度模型,对低压居民的用电量、缴费方式、居住区域和立户年限等因素进行分析,构建分类规则,通过数据挖掘算法,创建基于客户体验的电子渠道运营评估体系,建立闭环的优化机制,持续提升电子渠道客户满意度及运营能力,促进电子渠道用户数的快速增长。

(1) 场景验证思路。制订详细的需求调研计划,收集用户业务问题,掌握客户核心应用需求,落实以客户需求为核心的工作思路和方法,完成分析体系搭建和建模,形成电子渠道业务验证方案。

(2) 缴费渠道模型建模思路。制订详细的需求调研计划,收集用户业务问题,掌

握客户核心应用需求，落实以客户需求为核心的工作思路和方法，完成分析体系搭建和建模。

3) 应用成效

在大数据环境下，一方面通过构建偏好度模型，将偏好程度进行量化，从而进行用户偏好度划分；另一方面通过构建渠道适应度模型，将电子缴费场景的预测结果与客户的渠道偏好度相结合，创建基于客户体验的电子渠道运营评估体系，建立闭环的优化机制，持续提升电子渠道客户满意度及运营能力，进而促进电子渠道用户数的快速增长。

2. 基于文本相似度的重复来电分析

1) 应用简介

利用文本挖掘技术，结合客服业务需求，对来电信息中隐含的语义特征进行挖掘，重点关注重复来电次数多的客户，对来电出现的热点问题及时发现与跟踪，以便能把握处理问题的最佳时机，从而提高处理热点问题的能力和监测能力。

2) 设计实现

在提取重复来电工单的过程中，首先进行数据清理，即对95598工单文本清除异常来电号码，比如无故挂断、内部拨测、12345等异常来电。其次进行数据预处理，即根据关联规则来提取来电相同的工单。再次进行文本语义分析，包括中文分词、特征向量矩阵，相似度计算，聚类步骤，对来电相同的工单的受理内容进行文本聚类，提取同一用户和来电内容相似的重复来电工单，并对语义特征进行分析。

3) 应用成效

首先，从效率上来讲，重复来电专题分析替代了人工查找重复来电工单，能提高工作速率。一个月工单量大约24万条，仅凭人工借助关联工单或电话号码、户号筛查的方式统计，不仅耗时巨大，要基于全量工单进行重复来电分析的可操作性也不高，也较难彻底反映用户用电重复关心的问题原因。通过该系统，利用大数据挖掘、语义分析技术，计算时间短，时效性更强，更有助于及时决策。其次，从质量上来讲，利用重复来电模型对数据进行处理，经实验验证，准确率能达到90%，数据质量较优。

3. 利用智能语音技术实现话务录音自动质检

1) 应用简介

利用智能语音技术，将95598话务录音数据精准转译为结构化文本数据，实现海量话务录音数据的全量自动质检，并对话务服务过程中存在的问题进行及时判别和分析，辅助人工质检，提高人工质检效能。智能语音质检能自动提取业务关键数据信息，有效辅助管理人员开展来电原因、客户行为、疑似投诉等专项分析工作，提升95598服务监督力度和质量管理工作质效。

2) 设计实现

采用目前国际主流的DNN（深度神经网络）+HMM（隐马尔可夫模型）的方法训练

中文语音识别技术，能够适应不同年龄、不同地域、不同人群、不同信道、不同终端和不同噪声环境的应用环境，利用中心海量语音、文本语料进行模型定制化训练，构建高可用、高识别率的语音转写及分析平台。按照公司统一接口规范与第三方业务系统对接，提供标准统一的集成方案，便于后续的升级和扩展。此外，该平台还可提供国际通用的媒体服务协议接口，与业务应用系统进行无缝集成。

智能语音识别核心为智能语音转写技术，主要包括语言模型训练、前端语音处理、后端识别处理三部分。

3) 应用成效

通过智能语音识别技术与呼叫行业质量管理工作深度融合，实现工单全量自动质检和分级分类，并提供相关功能辅助人工质检，实现质量管理模式由“随机抽取”质检向“先分类、后质检”的精准化、智能化模式转变。

一是对电力客户，开展规范、全面的客户服务，切实解决客户实际问题，更准确地识别客户需求、评估服务能力、识别服务风险，提高客户满意度，不断推进供电服务品质的提升。

二是对国网客服中心，为持续优化业务流程、服务话术、系统运行提供依据，全面提升客服专员服务能力，实现中心服务质量持续改进、风险防控能力持续加强的目标。

三是对省（市）公司，实现95598业务协同和全过程监督，为省（市）公司经营管理提供以客户为导向的服务质量改善建议。

四是对国网公司，提高公司智能化水平，打造公司品牌效应，通过对智能语音识别技术的研究和应用，引入深度学习、神经网络技术等人工智能核心技术。

五是对呼叫中心行业，通过所积累的研究和应用经验，为其他呼叫中心开展人工智能技术研究提供技术支撑、理论指导和应用经验。

4. 电子渠道智能客服机器人

1) 应用简介

智能客服机器人借助自然语言处理和深度学习等核心技术，构建基于结构化数据和非结构化数据的专业知识库，提供拟人化的场景式自动应答能力，提供智能检索、智能应答、智能分析等能力，有效改善用户体验和提高用户满意度，辅助企业提升电子渠道业务发展和营销客服业务能力。

2) 设计实现

智能客服机器人系统的核心功能主要通过自然语言处理技术和深度学习技术来实现，系统主要功能模块包括统一客服工作台、机器人客户端、智能应答知识库管理、统计分析等功能模块。

客服工作台：包含服务会话窗口、客服资料展示、智能回复、快捷回复、客服转接、客服交接单、交办单、忽略词处理、近义词管理等。

客户端：多渠道接入、自动应答、转人工、客户留言、满意度评价等。

客服管理后台：智能应答知识库管理、知识审批、用户管理、机器人设置、系统报表、自动化处理功能等。

接口能力：接入各渠道（微信公众号、网站、95598、微博、APP等）客服数据的标准化接口和集成能力。

3) 应用成效

亿榕智能客服机器人在福建电力省客服中心开展了试点应用，系统集成了福建电力微信公众号及营销业务系统，实现了机器人能够自动回答用户通过公众号的留言，提升了福建省客服工作的智能化、自动化应用水平，有效解决了电费、电价、停电、故障等覆盖客服72%工作量的咨询工作，极大提升了客服人员的工作效率和客户满意度。