



面向信息服务的 信息技术

MIANXIANG XINXI FUWU DE XINXI JISHU



陈 豫 主编



国防工业出版社

National Defense Industry Press

014034203

E87
25

总装部队军事训练“十一五”统编教材

面向信息服务的信息技术

陈 豫 主编



国防工业出版社

· 北京 ·



北航

C1722474

E87
25

805280210

图书在版编目 (CIP) 数据

面向信息服务的信息技术 / 陈豫主编. — 北京:
国防工业出版社, 2013. 12

总装部队军事训练“十一五”统编教材

ISBN 978 - 7 - 118 - 09259 - 2

I. ①面... II. ①陈... III. ①军事情报 - 情报服务 -
信息技术 - 教材 IV. ①E87 - 39

中国版本图书馆 CIP 数据核字 (2013) 第 301441 号

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100048)

国防工业出版社印刷厂印刷

新华书店经售

*

开本 880 × 1230 1/32 印张 11 1/4 字数 334 千字

2013 年 12 月第 1 版第 1 次印刷 印数 1—3000 册 定价 45.00 元

(本书如有印装错误, 我社负责调换)

国防书店: (010) 88540777

发行邮购: (010) 88540776

发行传真: (010) 88540755

发行业务: (010) 88540717

总装备部军事训练统编教材 编审委员会

(2009)

主任委员	夏长法		
副主任委员	乐嘉陵	赵洪利	
委 员	于俊民	李 钢	黄水潮
	安敏建	陈长贵	姚志军
	杨开忠	白凤凯	裴承新
	姜世忠	张 渊	陈正兴
秘 书	石根柱	余敬春	

面向信息服务的信息技术

主 编 陈 豫

编写人员 陈 豫 曾 铮

于 洋 王三梅

主 审 真 溱

前 言

《面向信息服务的信息技术》是为从事信息服务工作的专业技术人员编写的教材,是《总装部队军事训练“十一五”统编教材》中的一本。本书由两个部分组成:第一部分试图从通用的信息技术中,选取对科技信息工作具有基础支撑作用的主流技术,进行概貌性介绍,使读者比较系统地掌握此方面技术;第二部分力图全面梳理科技信息专业领域关键技术,使读者能够把握开展信息服务的基础知识和发展趋势。

本书特点为:一是为给读者提供一个合理的知识结构,选材经过仔细推敲,既重点考虑当前有价值特别是有实用价值的内容,又考虑一些与科技信息工作相关的技术发展前沿;二是全书两个部分结构清晰、合理,概念清楚,层次分明,系统性好。每章的编写体例,采用了以专题学习为中心、知识技能为核心的模式,详细介绍了各项技术的基本概念、研究现状、应用案例和影响作用等。

本书由集体讨论大纲,分头编写。陈豫编写第1、2、5、6、7、13章;曾铮编写第3、4、12章;于洋编写第8、10章;王三梅编写第9、11章。陈豫统稿,中国国防科技信息中心真溱研究员审阅了全部书稿。

本书的编写出版工作一直在总装备部军事训练统编教材编审委员会领导下进行,并得到中国国防科技信息中心各级领导、机关的支持,在此表示感谢。由于信息技术发展极为迅速,知识面宽、内容广,虽经我们艰苦努力,但错漏之处在所难免,欢迎专家、读者批评指正。

编 者

2013年7月1日

目 录

第1章 科技信息工作与信息技术	1
1.1 信息技术相关概念	1
1.1.1 信息技术的定义	1
1.1.2 信息技术的分类	3
1.1.3 信息技术的发展历程	3
1.2 科技信息工作中信息技术	5
1.3 我国科技信息工作中信息技术发展历程	6
1.3.1 科技情报现代化的早期探索	6
1.3.2 国内联机检索中心系统的广泛创建	8
1.3.3 先进信息技术的实用化	9
1.3.4 技术应用水平全面提升	10
1.3.5 网络信息服务的快速发展	12
1.3.6 数字化资源加工与服务深入开展	13
1.3.7 知识服务模式的探索和发展	14
1.4 大数据与科技信息工作	16
1.4.1 大数据的定义	16
1.4.2 大数据产生原因	16
1.4.3 大数据的特点	17
1.4.4 大数据对科技信息工作的影响	18
第2章 网络管理与安全技术	21
2.1 网络管理技术	21
2.1.1 基础知识	21
2.1.2 主要任务	22

2.1.3	网络管理功能	24
2.1.4	网络管理协议	26
2.1.5	网络管理模式	27
2.1.6	影响和作用	28
2.1.7	应用案例	28
2.2	信息安全基础知识	30
2.2.1	基本概念	30
2.2.2	发展阶段	32
2.2.3	影响和作用	33
2.3	网络安全技术	33
2.3.1	基本概念	33
2.3.2	网络安全问题	34
2.3.3	主要技术	37
2.3.4	影响和作用	44
2.3.5	应用案例	45
第3章	服务器与存储技术	47
3.1	服务器基础知识	47
3.1.1	服务器概念	47
3.1.2	服务器分类	48
3.1.3	服务器介绍	49
3.2	集群技术	50
3.2.1	基本概念	50
3.2.2	主要分类	51
3.2.3	影响和作用	53
3.2.4	应用案例	54
3.3	服务器虚拟化	56
3.3.1	基本概念	56
3.3.2	技术架构和产品	57
3.3.3	影响和作用	59

3.3.4	应用案例	59
3.4	网络存储技术	62
3.4.1	基本概念	62
3.4.2	研究热点	62
3.4.3	常用网络存储系统	65
3.4.4	影响和作用	75
3.4.5	应用案例	77
3.5	存储虚拟化	79
3.5.1	基本概念	79
3.5.2	技术架构	80
3.5.3	影响和作用	81
3.5.4	应用案例	81
第4章	平台架构技术	85
4.1	云计算	85
4.1.1	基本概念	85
4.1.2	关键技术	87
4.1.3	关键应用	89
4.1.4	影响和作用	90
4.1.5	应用案例	91
4.2	面向服务的体系结构	95
4.2.1	基础概念	95
4.2.2	系统架构	97
4.2.3	实施步骤	101
4.2.4	开源 SOA 平台	102
4.2.5	影响和作用	105
4.2.6	应用案例	106
4.3	Web2.0	110
4.3.1	基本概念	110
4.3.2	相关理论	113

4.3.3	相关技术	114
4.3.4	相关应用	115
4.3.5	基本特点	117
4.3.6	影响和作用	118
第5章	数据仓库与机构仓储技术	121
5.1	数据仓库	121
5.1.1	基本概念	121
5.1.2	基本特点	122
5.1.3	基本构架	124
5.1.4	应用模式	125
5.1.5	影响和作用	127
5.1.6	应用案例	129
5.2	机构仓储	130
5.2.1	基本概念	130
5.2.2	体系架构	132
5.2.3	系统功能	133
5.2.4	机构仓储软件 DSpace	134
5.2.5	机构仓储软件 Fedora	136
5.2.6	影响和作用	138
5.2.7	应用案例	138
第6章	信息获取技术	140
6.1	网络资源的自动采集	140
6.1.1	基本概念	140
6.1.2	关键技术	141
6.1.3	系统类型	142
6.1.4	影响和作用	145
6.1.5	应用案例	147
6.2	信息资源数字化加工	151
6.2.1	基本概念	151

6.2.2	资源类型	152
6.2.3	数字化加工方法	153
6.2.4	影响和作用	157
6.2.5	应用案例	158
6.3	网络信息集成技术	161
6.3.1	基本概念	161
6.3.2	主要模式	163
6.3.3	基本方法	165
6.3.4	技术实现	165
6.3.5	影响和作用	167
6.3.6	应用案例	168
第7章	信息组织技术	172
7.1	元数据	172
7.1.1	基本概念	172
7.1.2	主要类型	173
7.1.3	机读目录	174
7.1.4	都柏林核心元数据	174
7.1.5	资源描述框架	177
7.1.6	元数据互操作	178
7.1.7	影响和作用	180
7.1.8	应用案例	182
7.2	书目记录功能需求	186
7.2.1	基本概念	186
7.2.2	主要内容	187
7.2.3	影响和作用	189
7.2.4	应用案例	190
7.3	数字对象标识与开放链接	194
7.3.1	基本概念	194
7.3.2	数字对象标识	195

7.3.3	OpenURL	197
7.3.4	影响和作用	199
7.3.5	应用案例	200
7.4	自动标引技术	202
7.4.1	相关概念	202
7.4.2	研究历程	203
7.4.3	自动抽词标引	204
7.4.4	自动赋词标引	205
7.4.5	问题和展望	206
7.4.6	应用案例	208
第8章	知识组织技术	212
8.1	词汇知识	212
8.1.1	概念词发现	212
8.1.2	概念关系表示方法	214
8.1.3	概念关系构建	215
8.1.4	应用案例	216
8.2	语义网	218
8.2.1	基本概念	218
8.2.2	研究重点	220
8.2.3	语义网应用	222
8.2.4	语义网工具	223
8.2.5	影响和作用	223
8.3	本体	224
8.3.1	基本概念	225
8.3.2	关键技术	227
8.3.3	影响和作用	230
8.3.4	应用案例	231
8.4	关联数据	231
8.4.1	基本描述	232

8.4.2	主要做法	232
8.4.3	影响和作用	233
8.4.4	应用案例	233
8.5	知识组织体系	234
8.5.1	基本概念	234
8.5.2	主要类型	235
8.5.3	常见形态	236
8.5.4	研究热点	241
8.5.5	影响和作用	243
8.5.6	应用案例	244
第9章	信息检索技术	246
9.1	文本检索	246
9.1.1	技术核心	246
9.1.2	工具介绍	247
9.2	多媒体检索	248
9.2.1	基本概念	249
9.2.2	应用系统介绍	250
9.2.3	影响和作用	251
9.3	语义检索	251
9.3.1	基本概念	252
9.3.2	技术架构	252
9.3.3	研究方向	253
9.3.4	影响和作用	255
9.3.5	应用案例	255
9.4	文献相关性技术	256
9.4.1	基本概念	257
9.4.2	文献相关的特性	260
9.4.3	实现技术	262
9.4.4	影响和作用	265

9.4.5 应用案例	266
第 10 章 信息推荐技术	270
10.1 用户建模	270
10.1.1 基本概念	270
10.1.2 建模方法	271
10.1.3 研究方向	273
10.1.4 影响和作用	274
10.1.5 应用案例	275
10.2 日志挖掘	275
10.2.1 基本概念	275
10.2.2 技术特点	276
10.2.3 影响和作用	276
10.2.4 应用案例	277
10.3 信息推荐	278
10.3.1 基本概念	279
10.3.2 技术实现	279
10.3.3 应用案例	281
第 11 章 信息导航技术	283
11.1 信息导航	283
11.1.1 基本概念	283
11.1.2 技术实现	284
11.1.3 研究热点	287
11.1.4 影响和作用	288
11.1.5 应用案例	288
11.2 信息可视化	292
11.2.1 基本概念	293
11.2.2 技术实现	294
11.2.3 研究热点	295
11.2.4 文献信息可视化	297

11.2.5	叙词表可视化	298
11.2.6	影响和作用	299
11.2.7	应用案例	300
第 12 章 数据挖掘技术		307
12.1	数据挖掘的基础知识	307
12.1.1	基本概念	307
12.1.2	基本流程	308
12.1.3	主要类型	309
12.1.4	工具与方法	311
12.1.5	影响和作用	312
12.1.6	应用案例	314
12.2	文本自动分类	316
12.2.1	基本概念	317
12.2.2	数学化描述	317
12.2.3	主要方法	318
12.2.4	影响和作用	323
12.3	文本聚类技术	324
12.3.1	基本概念	324
12.3.2	数学化描述	325
12.3.3	主要方法	325
12.3.4	影响和作用	328
12.4	信息抽取技术	329
12.4.1	基本概念	329
12.4.2	主要类型	330
12.4.3	主要方法	332
12.4.4	影响和作用	333
12.4.5	应用案例	334
第 13 章 情报分析技术		336
13.1	开源情报分析	336

13.1.1	基本概念	336
13.1.2	关键技术	336
13.1.3	影响和作用	339
13.1.4	应用案例	340
13.2	信息资源计算模式	342
13.2.1	基本描述	342
13.2.2	典型应用	343
13.2.3	实现技术	344
13.2.4	应用案例	345
参考文献		347

第1章 科技信息工作与信息技术

人类社会活动过程,就是人类交流信息和应用信息的过程,人类社会的发展和进步,必然伴随着信息技术的进步与发展。

信息技术作为很宽泛的概念,覆盖领域广泛。本书将针对科技信息工作的实际,着重介绍近年内与科技信息工作密切相关、可能带来重大影响、具有代表性的新技术。

1.1 信息技术相关概念

1.1.1 信息技术的定义

信息技术(Information Technology, IT)是计算机学、图书馆学、情报学等多个学科领域中非常重要的研究对象,相关行业中存在着多种定义。

1) 基于计算机学科视角的定义

信息技术是用于信息自动处理与检索的方法与技术,主要指计算机、通信设备及办公系统^[1]。信息技术包括所有与信息系统及应用软件的设计、开发、安装和使用相关的技术^[1]。

信息技术被描述为用于自动获取、存储、处理、管理、控制、显示、交换、传输、接收数据和信息的设备、互联系统或其子系统,包括计算机及附属设备、软件、固件及类似的程序、服务和相关资源^[2]。

2) 基于信息科学视角的定义

信息技术是以信息科学的原理与方法为基础,以计算机技术、通信技术和网络技术等为手段和工具,对各种传感信号和多种媒体的信息(数