

Broadview[®]
www.broadview.com.cn

IBM

China Development Laboratories Series
IBM中国开发中心系列



全球化软件开发

最佳实践

童春杰 周皓峰 杨普 舒芳蕊 等编著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

IBM

China Development Laboratories Series
IBM中国开发中心系列

《全球化软件开发最佳实践》

全球化软件开发

最佳实践

童春杰 周皓峰 杨普 舒芳蕊 等编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书是介绍全球化基本体系结构、技术和方法的经典力作。全书共 18 章，介绍了全球化的背景、Java 开发中的国际化技术、全球化测试、常见问题的解决方法、DB2 数据库等知识，并以一个完整的全球化开发实例，指导读者利用该用例中的方法和实现，自行实施一个精巧、完整的全球化开发项目。

本书适合具有一定基础的软件开发人员、计算机或相关专业的学生，以及需要为企业决策者提供技术和信息支持的架构设计师。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目 (CIP) 数据

全球化软件开发最佳实践 / 童春杰等编著. —北京: 电子工业出版社, 2008.6

(IBM 中国开发中心系列)

ISBN 978-7-121-06315-2

I. 全… II. 童… III. 软件开发 IV. TP311.52

中国版本图书馆 CIP 数据核字 (2008) 第 043661 号

责任编辑: 高洪霞

印 刷: 北京智力达印刷有限公司

装 订: 北京中新伟业印刷有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×980 1/16 印张: 26.5 字数: 504 千字

印 次: 2008 年 6 月第 1 次印刷

印 数: 4000 册 定价: 55.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

丛书序

致力于软件业务在中国的长期发展,IBM公司于1999年在中国投资成立了中国开发中心(IBM China Development Lab, CDL)。在为IBM全球客户提供满足需求的软件、硬件产品,以及技术和解决方案的同时,作为IBM全球软件资源在中国的窗口,CDL把全球先进技术引入中国,为中国软件产业与世界的交流搭建桥梁。

八年来,CDL不断吸引全球卓越的科技和管理人才,以及中国各大学府的顶尖学生、优秀工程师加入,并积极致力于同本土合作伙伴共同建设团队,研发队伍从八年前的100多位软件工程师增至今天的3000多位,年平均成长速度超过50%。

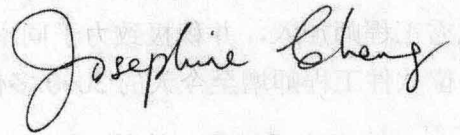
CDL高度重视员工培养,除各种短期或长期的培训,员工还被派往IBM在美国或其他地区的实验室,与世界各地同仁一起工作,在工作中学习先进的技术和管理方式。CDL更是将IBM全球实验室中拥有丰富经验的技术和管理人才请到中国工作,向中国员工传授经验。庞大的资金注入,人性化的管理方式,以及对人才方面的巨大投资,对人力资本的极度重视,使中国开发中心得以与IBM全球实验室共同成长,最终拥有一支经验丰富,训练有素的团队。

目前,CDL与全世界同步发展多项领域产品,正在为包括Information Management、WebSphere、Lotus、Tivoli、Rational在内的所有IBM软件核心产品的研究和开发做出卓越贡献,并在SOA、数据库、WebSphere产品系列、普及运算、Lotus Workplace Client技术及Linux系统方面取得非凡的成就,被视为IBM全球产品的开发重心之一。

在CDL高速发展的同时,为将信息产业的最新技术尽快地转化为对中国用户有价值的解决方案,帮助用户更有成效地开展业务,增强竞争优势,我们恪守为中国软件业与世界交流搭建桥梁的承诺,希望将IBM全球公司几十年的技术积淀和我们的心得与大家共同分享,于是,我们选择了实力非凡、专业创新的电子工业出版社博文视点公司作为合作伙伴,推出这一由IBM中国开发中心(CDL)的架构师、资深软件工程师们编写的系列丛书,范围涵盖了从开发实践、测试方法、项目实践、最新技术标准和发展趋势探讨,到先进解决方案构建、面向服务的架构的提供等诸多方面。

我希望这套丛书能把我们一线专家宝贵的经验，以及我们的见解呈现给读者，并希望无论是企业 IT 经理、程序设计和开发人员、软件工程师、软件架构师，还是在校学生，或者是对计算机领域有兴趣的人员，都能从中获取知识或者得到启发。

在同业界分享经验和世界最新技术及趋势的同时，我们希望能为推动中国软件产业的加速发展贡献微薄之力。IBM 中国开发中心将一如既往地同业界同仁一起，共铸中国信息产业的辉煌明天！



郑如勤

IBM 院士、总经理 IBM 中国开发中心

本书编委会

童春杰，毕业于浙江大学，部门经理，在多个全球化项目中担任项目管理和具体开发工作，对 J2EE 和 Web 应用开发有浓厚的兴趣。

周皓峰，毕业于复旦大学，博士，软件工程师，目前从事软件测试工作，主要对包括数据库、智能信息和全球化在内的各方面有兴趣。

杨普，毕业于清华大学自动化系，现专注于开放源代码和 Java 运行时技术开发，对软件国际化开发也有浓厚的兴趣。

舒芳蕊，毕业于上海交通大学计算机系，软件工程师，现专注于 Web 2.0 应用开发，对嵌入式设备上的软件应用和 Java 企业应用也有浓厚的兴趣。

其他作者，按照章节顺序

朱晓蕙，毕业于上海交通大学计算机系，软件工程师。在 IBM 先后参与了本地化翻译，双字节测试，全球化测试，全球化解决方案开发等项目。

王珏，毕业于上海交通大学计算机系，软件开发高级经理，主要从事全球化及用户技术的开发。

连理，毕业于复旦大学计算机系，现任部门经理，带领软件团队设计与开发信息开发工具。曾参与多个全球化相关的项目。

朱廷勇，毕业于华东师范大学计算机科学与技术系，软件工程师，主要从事 Web 应用和全球化应用的设计和开发。

张元鹏，毕业于上海交通大学，软件工程师，从事 DITA 以及相关工具的开发。

蔡俊杰，毕业于浙江大学，部门经理，曾参与开发多个全球化项目，对软件的用户体验领域有浓厚的兴趣。

尹俊，毕业于南京大学，软件工程师，专注于软件全球化前沿技术的研究和开发。

费伽，毕业于浙江大学计算机系，软件工程师，在多个 J2EE 平台项目中参与开发与项

目管理，对 Web 应用开发和 SOA 软件架构有浓厚的兴趣。

李扬，毕业于上海交通大学计算机系，软件工程师，擅长 Java Web 开发，热爱生活，业余喜欢电脑游戏。

黄玮，毕业于浙江大学，软件工程师，从事全球化技术开发工作。

王华红，毕业于南京大学电子科学与工程系，软件工程师。在多个全球化项目中担任设计开发工作，对 Portal 开发有丰富的经验。现从事软件集成性与易用性测试等工作。

周建新，毕业于上海交通大学电子工程系，软件工程师，在多个项目中担任全球化测试工作，对测试自动化和性能测试也有浓厚的兴趣。

夏铭，毕业于华东师范大学计算机科学与技术系，软件工程师。先后从事 GB18030 符合性测试，Rational 软件配置管理系统维护和技术支持，软件功能性测试等工作。现专注于软件自动化测试的研究与实施。

单炜，毕业于南京大学计算机系，软件工程师，从事软件测试工作。

戚馨文，毕业于浙江大学计算机系，软件工程师。先后曾从事过软件功能系统测试，GB18030 符合性测试等工作。

曹智清，毕业于浙江大学，软件工程师，从事软件测试工作。

李茂雨，毕业于上海同济大学计算机系，软件工程师，在多个全球化项目中参与开发与测试，对软件全球化开发与测试有浓厚兴趣。

瞿菲，毕业于上海工程技术大学，美术设计师。在多个全球化项目中担任美术设计工作。

赵磊，毕业于复旦大学，软件工程师，曾参与多个全球化项目的开发和设计。

施尉霖，毕业于浙江大学，软件工程师，现从事全球化技术和工具开发相关工作。

黄省江，毕业于南京邮电学院，软件开发经理，主要从事全球化技术和工具的开发。

前 言

随着全球经济逐渐淡化地区经济的模式，而发展为跨越国家地区的真正意义上的世界经济，软件产品也出现了一个不可逆转的趋势，这就是全球化——使同一软件产品能同时服务于不同地区、不同语言和不同文化的技术革命。

过去我们常谈到的汉化是指将英文版的软件产品进行二次开发，加入中文支持，使软件拥有中文界面、中文输入/输出、中文处理能力等，这样开发出来的中文版软件产品我们已经不陌生了。但是这种汉化的开发技术已经不能适应今天的全球化经济了。当今世界的经济发展，尤其是因特网的广泛使用，要求业务能够跨越国家地区的限制，让来自全球不同地区、说着不同语言、有着不同文化习俗的客户能够自如地进行业务沟通和交易。

IBM 是全球 IT 业的蓝色巨人，作为业界的领袖，他们很早就认识到这种趋势。长期以来，全球化在 IBM 得到充分的体现，同时 IBM 也一直引领着业界为客户提供全球化服务应用的潮流。今天，很多企业希望 IBM 能帮助他们将业务拓展到全球，IBM 为他们提供的解决方案不仅能满足其业务需求，而且能够使资源的使用最大化。IBM 在技术体系结构方面建立了和全球化应用程序设计开发相对应的原则，以适应不断发展的全球化需求。全球化开发技术的快速形成，以及对新技术提供全球化支持，都有助于客户实施全球化的解决方案。本书正是基于 IBM 在这个领域的经验累积，着重介绍全球化的基本体系结构、技术和方法。

本书特色

本书是国内为数不多的介绍全球化基本体系结构、技术和方法的经典力作。不同于通常的 IT 领域技术书籍，本书的独到之处在于带给读者的丰富实战经验和最优解决之道很多都是直接来自于作者团队——IBM 中国软件开发实验室上海全球化实验室。该团队为 IBM 设计全球化开发流程，研制全球化开发工具，负责中文版本的翻译及测试工作，其在全球化技术方面的建树领先于业界，很多新颖创新的方法为业界首创，在中国和美国拥有技术专利。

本书既有全球化经典体系结构的介绍，又有大量的程序样例，使读者能够快速理解不同方法的实现原理和具体的应用。本书还包括了一个完整的全球化用例，读者可以利用该用例中的方法和实现，自行实施一个精巧而完整的全球化开发项目，从而真正掌握这种方

法。无论是程序员还是架构设计师，都可以从中受益。

本书的读者对象

本书的读者对象是具有一定基础的软件开发人员、计算机或相关专业的大学生，以及那些可能需要为企业决策者提供技术和信息支持的架构设计师。

本书组织结构

全书共分为 18 章。

第 1 章“全球化重要性”，由朱晓蕙、连理、王珏撰写，介绍了全球化的背景，以及全球化和本地化的概念。

第 2 章“字符与编码”和第 3 章“Unicode”，由周皓峰撰写，介绍了字符编码。第 2 章介绍了字符编码的基本知识，包括字符的基本特点、编码的定义和中文编码等。第 3 章介绍了 Unicode，这种行业标准旨在实现能够统一表示和操作世界上所有书写系统中的文字和符号。

第 4 章“全球化软件体系结构”，由朱廷勇和张元鹏撰写，介绍国际化软件体系结构的特点及组成要素，包括单一可执行体、区域模型、多语言输入/输出和本地化等。

接下来的三章集中介绍了 Java 开发中的国际化技术。

第 5 章“Java SE 国际化开发”，由杨普、蔡俊杰和尹俊撰写，介绍了 Java 编程语言和 Java 核心类库对国际化编程的支持，以及使用 Java 开发桌面应用程序所采用的技术，并简要介绍了 ICU4J 对 Java 标准库国际化支持的增强。

第 6 章“Java Web 国际化开发”，由费伽、李扬和黄玮撰写，介绍了 Java 国际化开发在 Web 应用程序领域的相关技术。涉及的技术包括 HTTP 协议、HTML/JSP/Servlet 编码设置、“资源包”和“语言目录”的实现策略、标记库 (Tag Lib)，以及 JavaScript 的国际化开发等。

第 7 章“Java ME 国际化开发”，由舒芳蕊撰写，从多语言的支持及和文化习俗相关的本地化功能支持这两个方面介绍了 Java ME 的国际化开发。

第 8 章，由周建新、曹智清、夏铭、单炜和戚馨文撰写，介绍了全球化测试。

第 9 章由瞿菲撰写，介绍了美工。

第 10 章由李茂雨撰写，介绍了一些常见问题的解决方法。

第 11 章介绍了 WebSphere Application Server，由费伽、童春杰和赵磊撰写。第 12 章介绍了 Websphere Portal，由王华红撰写。第 13 章介绍了 DB2 数据库，由周皓峰撰写。第 14 章介绍了如何使用 RAD 进行全球化开发，由尹俊和李茂雨撰写。

第 15 章~第 18 章介绍了一个完整的全球化开发实例——全球快递解决方案，由童春杰、施尉霁、王华红、舒芳蕊、黄省江、曹智清和单炜撰写，读者可以此为范例来规范自己项目的全球化设计和开发。

最后，朱晓蕙对全书进行了审阅和修改。

致 谢

本书是集体创作的结晶，写作占用了作者大量的休息、娱乐，以及和家人在一起的时间，在此感谢作者家人的理解和支持。同时，成书的过程与许多人的关怀、鼓励和支持密不可分，其中包括 IBM 中国软件开发实验室用户技术实验室的总监 Sally Lam 女士，上海全球化开发实验室高级经理 Joyce Wang 女士等，在此表示感谢。最后，特别感谢出版社的有关领导、协调人员及编辑，有了他们的支持和参与，本书得以顺利出版。

由于时间仓促、水平有限，书中难免有许多不妥、甚至错误之处，在此敬请读者不吝指出，我们将愿意与读者共同探讨，并不胜感激。

本文提供的建议或最佳实践只作为一般的经验分享，只在作者的特定环境下验证过，作者不保证这些建议或最佳实践在任何情况下都有效。

本书为了内容介绍的方便，提到了一些测试数据、人名、网站名称或公司名称，这些人名、网站名称或公司名称纯属虚构，如果与任何实际的名称相同，则纯属巧合。

联系方式

关于本书的内容，如果读者有任何问题和建议，可以通过电子邮件 utdev@cn.ibm.com 联系我们。

编 者

目 录

第 1 章 全球化重要性	1
1.1 全球化背景	2
1.2 全球化与本地化	3
1.3 对全球化技术的认识	3
1.4 全球化技术和我	6
1.5 小结	7
第 2 章 字符与编码	9
2.1 字符 (Character)	10
2.2 字符编码	10
2.3 中文编码	13
2.3.1 GB 2312	13
2.3.2 GBK	14
2.3.3 GB 18030	17
2.3.4 Big5	20
2.3.5 中国香港增补字符集 (HKSCS)	21
2.4 小结	22
第 3 章 Unicode	23
3.1 统一字汇 (Universal Repertoire)	24
3.2 编码方式	29
3.2.1 UTF-32	30
3.2.2 UTF-16	30
3.2.3 UTF-8	31
3.2.4 字节顺序	32
3.2.5 Unicode 编码模式 (Encoding Scheme)	33

3.2.6	三种 UTF 编码方式的选择	34
3.3	字符的语义属性	34
3.4	字符的分解与组合	35
3.5	转换	37
3.6	Unicode 标准	37
3.7	Unicode 算法	38
3.7.1	叠合 (Folding)	40
3.7.2	大小写转换 (Case Mapping / Folding)	41
3.7.3	规范化	42
3.7.4	文本边界 (Text Boundary)	44
3.7.5	排序和查找	46
3.8	双向文字	47
3.8.1	双向文字的外在表现	47
3.8.2	双向文字的图形界面显示	51
3.8.3	双向文字的处理	52
3.9	小结	55
第 4 章	全球化软件架构	56
4.1	全球化软件的特点	57
4.2	全球化软件体系结构的组成要素	58
4.2.1	单一可执行体	58
4.2.2	区域模型	60
4.2.3	多语言数据的输入/输出	65
4.3	与文化习俗相关的本地化功能	67
4.3.1	与应用程序相关的本地化	67
4.3.2	与应用程序无关的本地化	67
4.3.3	文档的本地化	67
4.4	小结	78
第 5 章	Java SE 国际化开发	79
5.1	区域 (Locale) 和本地化信息	80

5.1.1	区域管理	81
5.1.2	本地化的信息	84
5.1.3	资源束	89
5.2	字符和编码转换	94
5.2.1	字符	94
5.2.2	字符编码转换	98
5.3	操作字符串	107
5.3.1	字符及字符串操作	107
5.3.2	枚举	108
5.3.3	转换	109
5.3.4	识别	113
5.3.5	分割	114
5.3.6	比较	114
5.4	日期和时间	115
5.4.1	日期和时间的表示	115
5.4.2	时区的表示	116
5.4.3	日期和时间的计算	118
5.5	格式处理	124
5.5.1	日期格式处理	124
5.5.2	数字格式处理	127
5.5.3	货币格式处理	129
5.5.4	消息格式处理	131
5.6	界面布局	133
5.6.1	指定界面布局	133
5.6.2	镜像	134
5.7	字体渲染	134
5.8	IBM ICU4J 和国际化开发的类库选择	137
5.8.1	IBM ICU4J	137
5.8.2	如何选择合适的国际化类库	138
5.9	小结	140

第 6 章	Java Web 国际化开发	141
6.1	HTTP 协议及浏览器编码行为	142
6.1.1	HTTP 协议	142
6.1.2	浏览器行为分析	144
6.1.3	简单总结	149
6.2	HTML/JSP/Servlet 的编码设置	149
6.2.1	HTML 的编码设置	149
6.2.2	JSP 的编码设置	152
6.2.3	Servlet 的编码设置	155
6.2.4	主流 Web 框架的编码设置	158
6.3	多语言支持的实现	161
6.3.1	主流多语言支持的实现策略	161
6.3.2	在服务器端获取用户的语言偏好	164
6.3.3	HTML 页面的多语言实现	167
6.3.4	使用标记库格式化多语言信息	171
6.4	JavaScript 国际化开发	177
6.4.1	设计原则	177
6.4.2	常用的 JavaScript 国际化 API	180
6.4.3	使用 JavaScript 实现资源包	182
6.4.4	Ajax 应用程序中的国际化	187
6.4.5	Dojo 国际化 API 介绍	190
6.4.6	Dojo 国际化 API 使用样例	193
6.5	小结	200
第 7 章	Java ME 国际化开发	201
7.1	Java ME 技术平台对国际化支持的概述	202
7.1.1	Java ME (Java Platform, Micro Edition) 技术平台介绍	202
7.1.2	CDC 对国际化支持的概述	205
7.1.3	MIDP/CLDC 对国际化支持的概述	206
7.2	Java ME 对多语言支持的实现	211

7.2.1	国际化 Java ME 应用简介	211
7.2.2	连接设备配置 (CDC) 及其相关简表的多语言支持	211
7.2.3	MIDP/CLDC 的多语言支持	212
7.2.4	使用 JSR 238 提供 MIDP/CLDC 的多语言支持	220
7.3	Java ME 对文化习俗相关的本地化支持的实现	226
7.3.1	连接设备配置 (CDC) 及其相关简表的本地化支持的实现	226
7.3.2	MIDP/CLDC 的文化习俗相关的本地化支持的实现	227
7.3.3	使用 JSR 238 提供 MIDP/CLDC 的文化习俗相关的本地化支持	229
7.4	Java ME 软件国际化设计的一些最佳实践	232
7.4.1	Java ME 国际化应用的体系结构设计	232
7.4.2	Java ME 国际化应用程序的用户界面设计	233
7.4.3	Java ME 国际化应用程序的资源管理	236
7.5	本章小结	237
第 8 章	全球化测试	239
8.1	全球化验证测试	240
8.1.1	全球化验证测试的概念和特点	240
8.1.2	全球化验证测试的流程	240
8.1.3	全球化验证测试案例	241
8.2	翻译验证测试	243
8.2.1	翻译验证测试的概念和特点	243
8.2.2	翻译验证测试的流程	243
8.3	GB 18030 测试简介	246
8.3.1	什么是 GB 18030	246
8.3.2	GB 18030 符合性测试及常见问题	248
8.4	小结	250
第 9 章	CSS 和全球化美术设计	251
9.1	什么是 CSS	252
9.2	CSS 的优势	252
9.3	如何使 CSS 变得 Single Executable	253

9.3.1	去除区域相关限制并利用文化偏好	253
9.3.2	双向数据显示的更多思考	255
9.4	网站的结构	256
9.5	新的趋势	257
9.6	小结	257
第 10 章	其他常见问题解答及技巧	259
10.1	在使用 Tomcat 时遇到中文乱码	260
10.1.1	Tomcat 无法解析使用 Post 方法提交的表单中的中文	260
10.1.2	Tomcat 无法解析使用 Get 方法提交的表单中的中文	262
10.2	在使用 WebSphere 时遇到中文乱码	263
10.3	在使用 Struts 时遇到中文乱码	263
10.4	在 JavaScript 及 Ajax 编程中遇到中文乱码	264
10.4.1	Javascript 文件本身的编码和 Html/JSP 页面所指定的输出编码不同	264
10.4.2	Ajax 动态装入的中文内容未包含正确的编码信息	266
10.5	Java 中的压缩解压类无法读取 Zip 文件中以中文命名的文件	269
10.6	在使用 MySQL 数据库进行开发时遇到中文乱码	270
10.7	Java 中的中文字体显示为一个方块	271
10.8	Java 中的中文字体很难看	272
10.9	如何实时查看浏览器与 Web 服务器的 HTTP 头信息	272
10.10	小结	273
第 11 章	WAS 应用服务器	274
11.1	WebSphere Application Server 编码设置	275
11.1.1	请求编码	275
11.1.2	响应编码	275
11.2	国际化服务	280
11.2.1	国际化服务的概念	280
11.2.2	国际化上下文 API	282
11.2.3	开发一个国际化 Web 服务	285
11.2.4	为国际化的 Web 服务开发一个客户机	289

11.3	小结	294
第 12 章	WebSphere Portal 开发	295
12.1	编码和区域设置	296
12.1.1	Portlet 中的编码和区域设置	296
12.1.2	JSP 中的编码	296
12.1.3	资源包的编码	297
12.1.4	Portal 数据库编码	297
12.1.5	浏览器的编码和区域设置	297
12.2	如何在 Portal 中显示多语言文字	298
12.2.1	资源包的收集和翻译	298
12.2.2	Portal 项目中的资源包	299
12.2.3	实现 Portal 标记	300
12.2.4	在 JSP 中使用标记	303
12.2.5	验证 Portal 页面上的多语言	304
12.2.6	支持复杂的字符串	304
12.3	选择语言	305
12.3.1	选择语言	305
12.3.2	Portal 确定的语言	305
12.3.3	在会话期间动态更改语言	306
12.3.4	验证选择语言	306
12.4	在 Portal 中实现 BiDi 镜像	308
12.4.1	在 JSP 中引用 BiDi 标记	308
12.4.2	在 JSP 中定义和使用 BiDi 变量	308
12.4.3	Portlet 类的 BiDi	309
12.4.4	样式表中的 BiDi	310
12.4.5	BiDi 在 IE 浏览器和 Firefox 浏览器中的不同表现	310
12.5	小结	311
第 13 章	DB2 全球化支持	312
13.1	DB2 的全球化特性	313