

国家自然科学基金课题

定向移动对象基于有偏随机走的概率模型研究
基于差分式更新的数字地图制图模型研究

项目支持

Web 2.0 GIS 原理与方法教程

尹章才 李霖 主编



武汉理工大学出版社

国家自然科学基金课题

定向移动对象基于有偏随机走的概率模型研究 项目支持
基于差分式更新的数字地图制图模型研究

Web 2.0 GIS 原理与方法教程

主 编 尹章才 李 霖
副主编 袁艳斌 王 红

武汉理工大学出版社
· 武汉 ·

内 容 提 要

本教材针对 Web 2.0 环境下的地理信息技术,从基础语言、关键技术和实践应用等三个层面进行了总结和阐述。其中,语言篇主要介绍了 Web 2.0 GIS 的基础性语言,即 XML 及其扩展的标记语言,如 GML、XSLT、SVG 等,并集成了基于 XML 的地图表达系统;技术篇主要介绍了 Web 2.0 GIS 的关键技术,如 Web 地图服务、Ajax、GeoRSS、Mashup、KML、Web 地图 API、参与型 GIS 等,这些技术之间不是孤立的,而是相互联系的;应用篇介绍了几个典型的 Web 2.0 GIS 应用。本教程不需要特别的基础性知识,如果已学习了 Web (1.0)GIS 或 Web 知识,将有助于本教程的学习。

本书适应于地理信息系统、资源管理和城乡规划等专业的本科生、研究生作为专业课的教材使用。

图书在版编目(CIP)数据

Web 2.0 GIS 原理与方法教程/尹章才,李霖主编. —武汉:武汉理工大学出版社,2011.9
ISBN 978-7-5629-3585-8

I. ①W… II. ①尹…②李… III. 地理信息系统-高等学校-教材 IV. ①P208

中国版本图书馆 CIP 数据核字(2011)第 187068 号

项目负责人:陈军东 陈 硕

责任编辑:陈 硕

责任校对:陈军东

装帧设计:董君承

出版发行:武汉理工大学出版社

社 址:武汉市洪山区珞狮路 122 号

邮 编:430070

网 址:<http://www.techbook.com.cn>

经 销:各地新华书店

印 刷:武汉兴和彩色印务有限公司

开 本:787×1092 1/16

印 张:12.5

字 数:321 千字

版 次:2011 年 9 月第 1 版

印 次:2011 年 9 月第 1 次印刷

印 数:0—1000 册

定 价:25.00 元

凡购本书,如有缺页、倒页、脱页等印装质量问题,请向出版社发行部调换。

本社购书热线电话:027-87394412 87383695 87384729 87397097(传真)

• 版权所有 盗版必究 •

前言

相对于传统的纸质地图、单机版 GIS 而言,Web 地图/GIS 的出现极大地促进了地理信息的共享,但这种共享主要表现为地理信息的发布,对地理信息的采集、处理、用户体验以及免费的、简单的二次开发应用等还未引起关注。Web 2.0 的出现,改变了传统 Web 地图/GIS 的 Web 环境,随之也产生了 Web 2.0 GIS 技术,如 KML、OpenStreetMap、PPGIS 等,这些技术不仅继承了传统 Web 地图/GIS 单项发布地理信息功能,而且还能让用户参与到地理信息的采集、管理、加工、处理、分析等环节中来,也使得普通公众使用地理信息具有更好的用户体验,并通过免费的 API 共享和应用地理信息,从而实现真正意义上的地理信息大众化。

由于 Web 2.0 GIS 的开放性,使得虚拟的 Web 2.0 地理信息与现实的地理环境相互映射,相互构建。传统的纸质地图、Web GIS 的功用主要是地理信息的发布,而 Web 2.0 地图/GIS 则不仅仅是地理信息,而是地理信息媒介——人类活动 80% 的信息与地理位置有关,所以可以说——在这种地理信息媒介上承载人类 80% 的信息。

本书遵循理论与实践相结合的原则,在介绍基础性语言、技术的同时总是尽量结合实际应用实例,让读者不仅掌握理论、技术的意义,而且还初步掌握应用的方法,为 Web 2.0 GIS 的发展起到了一定的阶段性总结的作用。

从 Web 2.0 技术出发,结合当前 Web 2.0 GIS 的范例,本书阐述了 Web 2.0 GIS 的相关语言,如 XML、GML、XSL、SVG 等,分析了 Web 2.0 GIS 的相关技术,如 Ajax、GeoRSS、Mashup、KML 等,最后探索了 Web 2.0 GIS 产品的应用。因此,本书将分为三篇:语言篇;技术篇和应用篇。

一、语言篇

本篇主要介绍了 Web 2.0 广泛运用的基础语言 XML,及其与 Web 2.0 GIS 应用密切相关的语言,如 SVG、XSLT 等。这些语言具有互补的关系,因而它们之间的联系也比较紧密。本篇将从相互联系的角度,以地理信息的 Web 为基础,分别介绍 XML、SVG、XSLT 等语言,并通过地理信息可视化将它们联系在一起。

第 1 章:绪论,介绍了 Web 2.0 基础及 Web 2.0 地理信息技术。

第 2 章:基于 Web 的图示表达,介绍了 XML 在地理信息图示表达中的意义与模型,并从概念模型的角度将 XML、SVG、XSLT 等联系起来,从而在总体上把握 XML 及其扩展标记语言在地理信息地图表达中的功能和角色。

第 3 章:地图制图标记语言,介绍 XML 基础以及作为元标记语言如何构建地图制图标记语言的过程。让学生不仅了解数据结构类型、XML 的基本原理,而且初步掌握如何利用 XML 构建自定义的文档结构。

第 4 章:基于 XSL 的表达规则,介绍 XSL 语言基础,以及基于 XSL 的地理信息图示表达规则的描述方法。让学生不仅了解 XSL 语言规范,还将 XSL 应用于图示表达规则的描述中,同时也介绍了 GML 通过 XSLT 转换到 SVG 格式的方式。这样,学生可以进一步把握 XML 将内容与样式分离的意义以及内容与样式结合的方法。

第5章:基于 SVG 的表达规范,介绍了 SVG 图形表示规范,以及应用于地图符号表示的方法。

第6章:Web 地图表达系统集成,本章是与第2章相呼应的。通过第2章概念模型的介绍,以及第3、4、5章具体内容的实现,从而达到第6章系统的集成。此外,也延伸了用户参与地理信息技术的功能和工具。

上述各章内容可以分为两个部分,一是基础语言,另一则是语言的应用,因此后者可以作为选修内容。

二、技术篇

本篇主要介绍 Web 2.0 GIS 的关键技术,主要是介绍 Web 服务及其二次开发应用技术,包括 Web 服务、Ajax、Google 地图服务、Web 脚本语言、混搭和聚合等,这些技术大多涉及 XML 技术或者直接建立在 XML 基础上。因此,从这个角度上讲,中篇的内容是 XML 在 Web 2.0 GIS 中的深入应用,并在应用过程中能较全面体现 XML 的特点,包括优点和不足。

第1章:Web 地理信息服务,是中篇的基石,主要介绍 Web 服务以及主要的和基本的 Web 地理信息服务。

第2章:Ajax 地图发布,主要介绍 Web 2.0 的重要特征之一的 Ajax 技术及其在 Web 地图发布中的应用与特点。

第3章:Google 地图服务技术,主要介绍 Google 地图服务基础及其开放的 API、KML 等。

第4章:Web 脚本与 GeoJSON。介绍了 Web 脚本语言与 GeoJSON 技术。

第5章:Web 地理信息混搭与聚合。主要介绍 Mashup 技术和地理信息聚合标准。

上述各章是按照技术进行的划分,但其内容相互联系,都是 Web 2.0 地理信息技术组成中不可或缺的成分。

三、应用篇

本篇介绍了部分 Web 2.0 地理信息技术的应用,包括基于 Google 地图的协同设计、公众参与地理信息收集和处理、Google 地图的混搭等。

本书在总体结构上将纷繁复杂的 Web 2.0 地理信息技术与实践进行了分类。虽然语言、技术和应用密不可分,但我们在语言篇仅仅介绍 Web 2.0 的基础性语言;在技术篇介绍的技术,也包括一定的专有或专业性的语言,如 KML 等。这意味着,技术篇并不是不包含语言的介绍,这也是传统分类方法的不足。如果从 Web 2.0 的分众分类的角度,KML 的标签有:Web 2.0 语言和 Web 2.0 技术。本书的三篇内容相互联系,逐级推进,内容覆盖从基础性语言、技术到实际的应用领域。这种联系,在语言篇表现尤甚,各章内容共同构成了地理信息的 Web 地图表达系统。同时,每篇内容又自成体系,在每章,不是仅仅介绍语言或技术,而是先介绍语言或技术再深入介绍语言在地理信息技术中的应用方法与实践。

本书是作者在多年从事 Web GIS 教学工作、参与国家“863”课题《网络空间信息标准与共享服务关键技术(子课题〈地理信息可视化符号描述规则〉)》(编号:2002AA131030,2002 年)和主持国家自然科学基金课题《定向移动对象基于有偏随机走的概率模型研究》(41071283)及《基于差分式更新的数字地图制图模型研究》(40871178)基础上编写而成的。本书由尹章才、李霖担任主编,袁艳斌、王红任副主编,参编人员有何晓蓉、黎华、夏琳、袁梅、张晓盼。在编写过程中,得到了武汉理工大学资源与环境工程学院规划系的领导、老师的大力支持,尤其是崔

魏教授、黄解军、詹云军、吴浩、彭华老师等的鼓励,还有参与本科毕业设计的学生王世民、赵晶辉、张世伟、李彦鹤、周沂、吴迎霞、张林、陈有超、张辉、蔡育秋、农耘、裴璐颖,在此一并表示感谢!我们还参阅、引用了其他书籍和论文的部分内容或思想,在此对相关作者表示衷心感谢。

作为一本新书,必然会有学术观点的不同,也可能存在学识上的片面甚至错误。我们衷心地欢迎与期待同行及广大读者提出批评与建议。

作 者

2011 年 6 月于鉴湖

目 录

语言篇	(1)
1 绪论	(2)
1.1 Web 2.0 基础	(2)
1.1.1 Web 2.0 发展简介	(2)
1.1.2 Web 2.0 的主要特点	(3)
1.1.3 Web 2.0 技术	(3)
1.1.4 六度关系理论	(4)
1.1.5 Web 2.0 的基本原则	(5)
1.1.6 Web 2.0 公司的核心竞争力	(7)
1.2 Web 2.0 地理信息技术	(7)
1.2.1 Web GIS 特点	(8)
1.2.2 Web GIS 关键技术	(8)
1.2.3 Web GIS 功能	(9)
1.2.4 Web GIS 实现方法	(9)
1.2.4 Web 2.0 GIS	(11)
2 基于 Web 的图示表达	(12)
2.1 图示表达概念模型	(12)
2.2 图示表达逻辑模型	(13)
2.2.1 图示表达机制的扩展	(13)
2.2.2 地图表达机制	(15)
2.3 图示表达物理模型	(16)
2.3.1 基于 XML 的图示表达	(16)
2.3.2 基于 XML 的地图表达	(18)
练习题	(19)
3 地图制图标记语言	(20)
3.1 半结构化数据	(20)
3.1.1 数据的结构化类型	(20)
3.1.2 半结构化数据模型	(21)
3.1.3 HTML 与 XHTML	(22)
3.2 可扩展标记语言 XML	(23)
3.2.1 XML 语法基础	(24)
3.2.2 XML 基础标准与特征	(26)
3.2.3 XML 解析器与 DOM	(27)

3.2.4 DTD 和 Schema	(29)
3.3 地图制图标记语言	(32)
3.3.1 地图标记语言基础	(32)
3.3.2 MapMML 模型	(33)
3.3.3 MapMML 模式	(35)
练习题	(39)
4 基于 XSL 的表达规则	(40)
4.1 可扩展样式表语言 XSL	(40)
4.1.1 XSL 概述	(40)
4.1.2 XSLT 与 CSS 语言基础	(41)
4.1.3 XSLT 语法结构	(46)
4.1.4 基于 XSL 的规则表达	(47)
4.2 图示表达规则	(48)
4.2.1 我国地形图图式规范	(48)
4.2.2 图式规范中的表达规则	(50)
4.3 基于 XSL 的表达规则	(51)
4.3.1 表达规则的概念模型	(51)
4.3.2 基于 XSL 的表达规则逻辑模型	(52)
4.3.3 表达规则的物理模型	(53)
练习题	(56)
5 基于 SVG 的表达规范	(57)
5.1 可升级矢量标记语言	(57)
5.1.1 SVG 语法基础	(57)
5.1.2 SVG 与地图表达	(58)
5.2 基于 SVG 的地图符号描述	(60)
5.2.1 基于 SVG 的地图符号描述模型	(60)
5.2.2 基于 SVG 的图元描述	(62)
5.3 基于 SVG 的地图符号设计	(64)
5.3.1 基于 SVG 的符号库结构	(64)
5.3.2 基于 SVG 的地图符号设计	(65)
5.3.3 基于 SVG 的地图注记设计	(69)
5.3.5 基于 SVG 的地图交互与多重表达	(70)
练习题	(72)
6 Web 地图表达系统集成	(73)
6.1 原型系统概况	(73)
6.2 表达规则—测试过程	(76)
6.2.1 地理数据标记语言	(76)
6.2.2 基于 XSL 的表达规则	(76)
6.2.3 地图制图数据的生成	(76)

6.2.4 表达规则的验证	(77)
6.3 表达规范—符号库建立	(77)
6.3.1 基于 SVG 的地图符号库	(78)
6.3.2 表达规范的验证	(81)
6.3.3 个性化地图符号制作	(82)
6.4 表达服务—符号化	(83)
6.4.1 表达服务的结构与功能	(83)
6.4.2 表达服务的方法	(85)
6.5 Web 地图制图编辑	(86)
练习题	(88)
参考文献	(89)
技术篇	(92)
1 Web 地理信息服务	(93)
1.1 Web 服务	(93)
1.1.1 核心定义	(93)
1.1.2 使用方式	(95)
1.2 基础 Web 地理信息服务	(97)
1.2.1 网络地图服务	(98)
1.2.2 网络要素服务	(100)
1.2.3 网络覆盖服务	(102)
1.2.4 Web 地理信息服务的发展	(102)
练习题	(103)
2 Ajax 地图发布	(104)
2.1 互联网应用程序架构	(104)
2.1.1 一般互联网架构	(104)
2.1.2 丰富互联网应用程序	(105)
2.2 Ajax 技术基础	(106)
2.2.1 Ajax 结构	(106)
2.2.2 Ajax 工作原理	(107)
2.2.3 Ajax 的工作流程	(108)
2.2.4 Ajax 的特点	(108)
2.3 基于 Ajax 的栅格地图发布	(109)
2.4 基于 Ajax 的矢量地图制图	(111)
2.4.1 同步交互地图表达模型	(111)
2.4.2 异步交互地图表达	(113)
2.4.3 异步协同矢量制图	(113)
练习题	(114)
3 Google 地图服务技术	(115)

3.1	Google 地图服务应用基础	(115)
3.1.1	Google Earth	(115)
3.1.2	Google Maps	(117)
3.2	KML	(119)
3.2.1	KML 语法	(119)
3.2.2	KML 格式转换	(121)
3.3	Google Earth API	(124)
3.3.1	API 接口	(124)
3.3.2	COM 组件	(125)
3.3.3	Google Earth API 基础	(126)
3.4	Google Maps API	(129)
3.4.1	Google Maps API 简介	(129)
3.4.2	Google Maps API 地图操作	(130)
3.4.3	Google Maps API 地理译码	(132)
	练习题	(135)
4	Web 脚本与 Geo JSON	(136)
4.1	脚本语言	(136)
4.1.1	脚本语言起源和目的	(136)
4.1.2	脚本语言定义	(136)
4.1.3	脚本语言的分类	(137)
4.1.4	脚本语言的特点	(137)
4.2	JavaScript	(138)
4.2.1	ECMAScript	(138)
4.2.2	JavaScript 概述	(138)
4.2.3	JavaScript 语法	(139)
4.3	Geo JSON 技术	(139)
4.3.1	JSON	(139)
4.3.2	Geo JSON	(141)
	练习题	(142)
5	Web 地理信息混搭与聚合	(143)
5.1	Mashup	(143)
5.1.1	Mashup 关键问题及其集成特点	(143)
5.1.2	系统构架和关键技术	(144)
5.1.3	一个简单实例	(146)
5.2	地理信息聚合标准	(147)
5.2.1	RSS	(147)
5.2.2	Atom 标准	(151)
5.2.3	Geo RSS	(152)
	练习题	(154)

6 参与型地理信息技术	(155)
6.1 Wiki	(155)
6.1.1 Wiki 简介	(155)
6.1.2 Wiki 与 Blog 的区别	(155)
6.1.3 Wiki 的技术和规范	(156)
6.2 几种典型技术	(156)
6.2.1 Geo URL	(156)
6.2.2 GeoNames	(156)
6.2.3 GeoCommons	(157)
6.2.4 Mapufacture	(157)
6.3 新地理	(158)
6.3.1 新地理简介	(158)
6.3.2 Neogeography 与 ArcGIS 10	(158)
6.4 参与型 GIS	(159)
6.4.1 参与型 GIS 相关概念	(159)
6.4.2 Web 2.0 背景下参与型 GIS 发展趋势	(160)
练习题	(160)
参考文献	(161)
应用篇	(164)
1 Web 地图协同路线规划	(165)
1.1 Google Maps 规划平台	(165)
1.2 出行路线设计	(166)
1.3 行程安排与返程路线设计	(169)
1.3.1 行程安排	(169)
1.3.2 返程路线设计	(169)
练习题	(172)
2 OSM 地图制图	(173)
2.1 OSM 简介	(173)
2.2 制图分析	(173)
2.2.1 OSM 数据采集及预处理	(173)
2.2.2 OSM 数据编辑	(174)
练习题	(178)
3 空间气象信息服务	(179)
练习题	(180)
参考文献	(181)
附录 常用术语及缩写汇编	(182)
后记	(185)

语言篇

XML 是 HTML 的一种升级换代,是针对网络的深层次应用而提出的国际标准网络语言。如果说 HTML 是 Web 1.0 GIS 的主要网络语言,那么 XML 就是 Web 2.0 GIS 的核心语言。本书将在第一篇介绍与 Web 2.0 GIS 相关的 XML 语言及其扩展语言,为后续篇章提供基础。首先,介绍 XML 并构建制图标记语言 MapMML;然后,利用 XSL 描述规则,并将 GML 转换为 MapMML;其次,利用 SVG 描述地图符号;最后,集成实现地理信息基于 XML 的地图表达系统。

1 绪 论

1.1 Web 2.0 基础

1.1.1 Web 2.0 发展简介

Web 2.0 是相对 Web 1.0(2003 年以前的互联网模式)的新的一类互联网应用的统称,是一次从核心内容到外部应用的革命。由 Web 1.0 单纯通过网络浏览器浏览 html 网页模式向内容更丰富、联系性更强、工具性更强的 Web 2.0 互联网模式的发展已经成为互联网新的发展趋势。表 1.1 枚举了 Web 2.0 与 Web 1.0 的部分对应实例。

表 1.1 Web 2.0 与 Web 1.0 的部分对应实例

Web 1.0	Web 2.0
大英百科全书在线	维基百科全书(Wikipedia)
个人网站	博客(Blog)
屏幕抓取(screen scraping)	网络服务(web services)
发布	参与
内容管理系统	维基
目录(分类)	标签(“分众分类”等)
粘性	聚合

Web 1.0 到 Web 2.0 的转变,具体地说,从模式上是单纯的“读”向“写”、“共同建设”发展;由被动地接收互联网信息向主动创造互联网信息迈进!从基本构成单元上,是由“网页”向“发表/记录的信息”发展;从工具上,是由互联网浏览器向各类浏览器、RSS 阅读器等内容发展;运行机制上,由“Client Server”向“Web Services”转变;作者由程序员等专业人士向全部普通用户发展。Web 2.0 实际上是对 Web 1.0 的信息源进行扩展,使其多样化和个性化。

总之,Web 2.0 是以 Flickr、Craigslist、Linkedin、Tribes、Ryze、Friendster、Del.icio.us、43Things.com 等网站为代表,以 Blog、TAG、SNS、RSS、Wiki 等应用为核心,依据六度分隔、XML、Ajax 等新理论和技术实现的新一代互联网模式。Web 2.0 是信息技术发展引发社会变革所带来的面向未来、以人为本的创新 2.0 模式在互联网领域的典型体现,是由专业人员织网到所有用户参与织网的创新民主化进程的生动注释。

1.1.2 Web 2.0 的主要特点

(1) 用户参与网站内容制造

与 Web 1.0 网站单项信息发布的模式不同, Web 2.0 网站的内容通常是用户发布的, 使得用户既是网站内容的浏览者也是网站内容的制造者, 这也就意味着 Web 2.0 网站为用户提供了更多参与的机会, 例如博客网站和 Wiki 就是典型的以用户创造内容为指导思想, 而 tag 技术(用户设置标签)则将传统网站中的信息分类工作直接交给用户来完成。

(2) Web 2.0 更加注重交互性

在发布内容过程中, Web 2.0 实现了用户与网络服务器之间的交互, 而且, 也实现了同一网站不同用户之间的交互, 以及不同网站之间信息的交互。其中, 网站之间信息的交互在有的 Web 类型划分中被独立为 Web 3.0。

(3) Web 2.0 网站与 Web 1.0 没有绝对的界限

Web 2.0 符合 Web 标准的网站设计。Web 2.0 技术可以成为 Web 1.0 网站的工具, 一些在 Web 2.0 概念之前诞生的网站本身也具有 Web 2.0 特性, 例如 B2B 电子商务网站的免费信息发布和网络社区类网站的内容也来源于用户。

(4) Web 2.0 的核心不是技术而在于指导思想

Web 2.0 有一些典型的技术, 这些技术是为了达到某种目的所采取的手段。Web 2.0 技术本身不是 Web 2.0 网站的核心, 但典型的 Web 2.0 技术体现了具有 Web 2.0 特征的应用模式。因此, 与其说 Web 2.0 是互联网技术的创新, 不如说是互联网应用指导思想的革命。

1.1.3 Web 2.0 技术

对于 Web2.0 概念的说明, 通常采用典型 Web 2.0 网站案例介绍, 加上对部分 Web 2.0 相关技术的解释, 这些 Web 2.0 技术主要包括: 博客(Blog)、RSS、Wiki、网摘、社会网络(SNS)、P2P、即时信息(IM)等。

(1) Blog——博客/网志

Blog 的全名应该是 Web log, 后来缩写为 Blog。Blog 是一个易于使用的网站, 用户可以在其中迅速发布想法、与他人交流以及从事其他活动, 所有这一切都是免费的。

Blog 是个人或群体以时间顺序所作的一种记录, 并且不断更新。Blog 之间的交流主要是通过回溯引用(TrackBack)和回响/留言/评论(comment)的方式来进行的。Blog 大量采用了 RSS 技术, 所有的 RSS 文件都必须符合由 W3C 发布的 XML 1.0 规范。对读者来说, 可以通过 RSS 订阅一个 Blog, 确知该 Blog 作者最近的更新。对 Blog 作者来说, RSS 可以使自己发布的文章易于被计算机程序理解并摘要。

(2) RSS

RSS(也叫聚合内容)是站点用来和其他站点之间共享内容的一种简易的技术和数据交换规范, 它起源于网景通讯公司的推(Push)技术, 是将用户订阅的内容传送给他们的通讯协同格式(Protocol)。

(3) Wiki

Wiki 是一种多人协作的写作工具。Wiki 站点可以由多人(甚至任何访问者)维护, 每个人都可以发表自己的意见, 或者对共同的主题进行扩展或者探讨。

(4)网摘

“网摘”又名“网页书签”,起源于一家叫做 del.icio.us 的美国网站自 2003 年开始提供的一项叫做“社会化书签”(Social Bookmarks)的网络服务,网友们称之为“美味书签”(Delicious Bookmarks,delicious),在英文中的意思就是“美味的;有趣的”。

(5)SNS——社会网络

Social Network Software,社会性网络软件,依据六度理论,以认识朋友的朋友为基础,扩展自己的人脉。SNS 帮你运营朋友圈的朋友,无限扩张自己的人脉,在需要的时候,可以随时获取一点,得到该人脉的帮助。

(6)P2P——对等联网

P2P 是 peer-to-peer 的缩写,peer 在英语里有“(地位、能力等)同等者”、“同事”和“伙伴”等意义。这样一来,P2P 也就可以理解为“伙伴对伙伴”、“点对点”的意思,或称为对等联网。目前人们认为其在加强网络上人的交流、文件交换、分布计算等方面大有前途。

(7)IM——即时通讯

即时通讯(Instant Messenger,IM)软件可以说是目前我国上网用户使用率最高的软件。聊天一直是网民们上网的主要活动之一,网上聊天的主要工具已经从初期的聊天室、论坛变为以 MSN、QQ 为代表的即时通讯软件。大部分人只要上网就会开着自己的 MSN 或 QQ。

作为使用频率最高的网络软件,即时聊天已经突破了作为技术工具的极限,被认为是现代交流方式的象征,并构建起一种新的社会关系。它是迄今为止对人类社会生活改变最为深刻的一种网络新形态,没有极限的沟通将带来没有极限的生活。

(8)Tag

Tag(标签)是一种更为灵活、有趣的日志分类方式,用户可以为每篇日志添加一个或多个 Tag,然后可以看到 BlogBus 上所有和该用户使用了相同 Tag 的日志,并且由此和其他用户产生更多的联系和沟通。不仅如此,通过与 Technorati 的合作,还可以把用户的 Tag 发送到全球 Blog 空间,和全世界的人们共同分享。

Tag 体现了群体的力量,使得日志之间的相关性和用户之间的交互性大大增强,可以让用户看到一个更加多样化的世界,一个关联度更大的 Blog 空间,一个热点实时播报的新闻台。Tag 为用户提供前所未有的网络新体验……

当然,也可以简单地把一个 Tag 理解为一个日志分类,但是 Tag 和分类的不同之处也很明显:首先,分类是在写日志之前就定好的,而 Tag 是在写完日志之后再添加的;其次,可以同时为一篇日志贴上好几个 Tag,方便随时查找,而原先一篇日志只能有一个分类;再次,当积累了一定数量的 Tag 之后,可以看看用户自己在 Blog 中最经常写的是哪些话题;最后,可以看到有哪些人和用户自己使用了一样的 Tag,进而找到和自己志趣相投的 Blogger。

(9)屏幕抓取

屏幕抓取是使用软件工具从 HTTP Response 的 HTML 页面中提取出可以通过编程进行使用和操作的信息及其语义结构(李峰 等,2009)。

1.1.4 六度关系理论

目前流行的“六度分隔”理论是 20 世纪 60 年代由美国的心理学家米格兰姆(Stanley Milgram)提出的,这个理论可以通俗地阐述为:最多通过六个人你就能够认识任何一个陌生人。

“六度分隔”成为人际关系世界中无可否认而又令人震惊的特征,许多社会学上的深入研究也给出令人信服的证据,说明这一特征不只是特例,在一般情形下也存在。最近,美国哥伦比亚大学社会学系的瓦茨教授领导的 EMAIL 试验也再次证明了这一人际关系世界中惊人的规律。然而,在现实世界中,六十多亿人怎么可能真的构成如此紧密的相互关联呢?是互联网使一切成为现实。

1.1.5 Web 2.0 的基本原则

正如许多重要的理念一样,Web 2.0 没有一个明确的界限,而是一个重力核心。不妨将 Web 2.0 视作一组原则和实践,由此来把距离核心或远或近的网站组成为一个类似太阳系的网络系统,这些网站或多或少地体现着 Web 2.0 的原则(图 1.1)。该图是在名为“O'Reilly 的朋友”(Friend of O'reilly, FOO)的会议的一个研讨会上产生的。这个图基本上仍处于演化阶段,但已经描绘出了从 Web 2.0 核心理念中衍生出的许多概念。

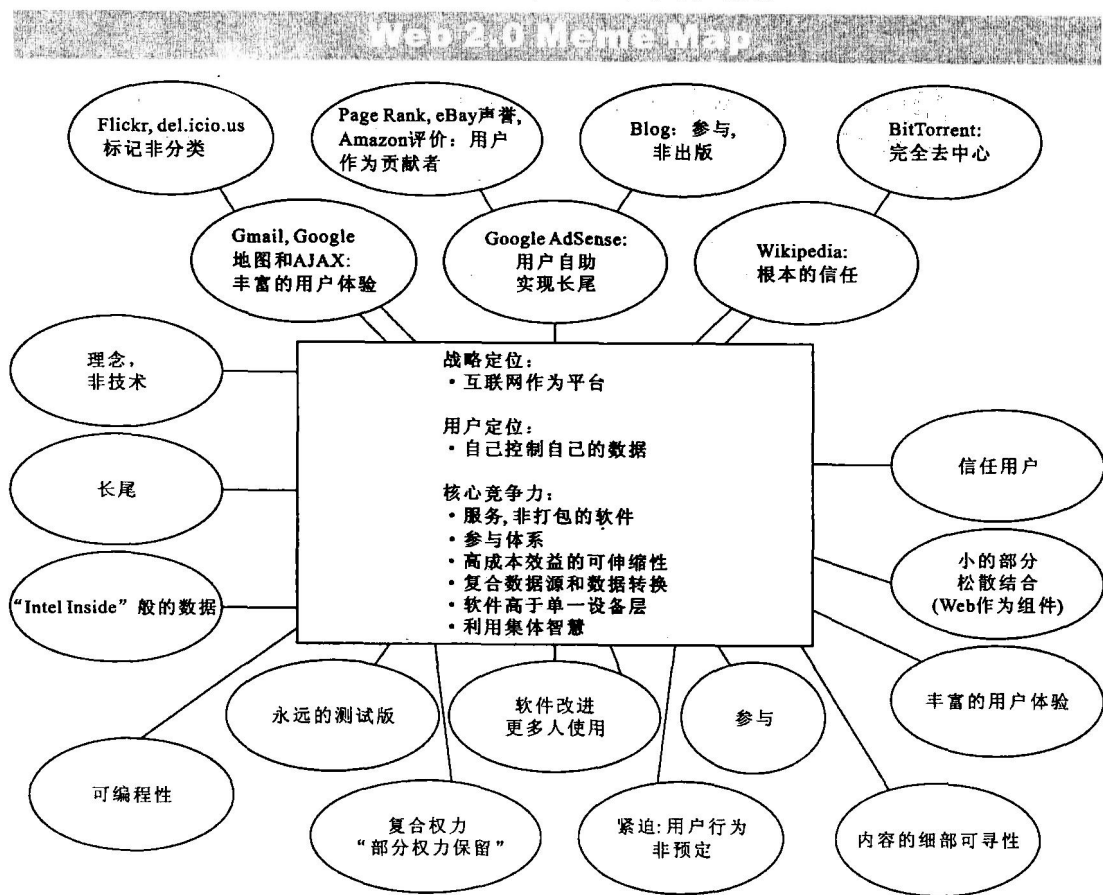


图 1.1 Web 2.0 的“模拟图”

(1) 互联网作为平台

我们早先的 Web 1.0 的楷模中的两个,DoubleClick 和 Akamai 公司,皆是将网络当做平台的先驱。人们往往不认为这是一种网络服务,但事实上,广告服务是第一个被广泛应用的网络服务,同时也是第一个被广泛应用的混合处理(Mashup)。如果用另一个近来流行的词来说

的话,每个旗帜广告(banner AD)都是用来在两个网站之前无缝合作,向位于另外一台计算机上的读者传递页面信息。

如果 Netscape 可以称为 Web 1.0 的旗手,那么 Google 可以说是 Web 2.0 的旗手。

(2) 利用集体智慧

维基百科全书(Wikipedia)是一种在线百科全书,其实现基于一种看似不可能的观念。该观念认为一个条目可以被任何互联网用户所添加,同时可以被其他任何人编辑。无疑,这是对信任的一种极端的实验,将埃里克雷·蒙德(Eric Raymond)的格言:“(源自开放源码软件的背景之下)有足够的眼球,所有的程序缺陷都是肤浅的(with enough eyeballs, all bugs are shallow)”运用到了内容的创建之中。本书中的大量内容来源于维基百科和百度百科,这意味着本书也是基于 Web 2.0 的。

“分众分类”是一种使用用户自由选择的关键词对网站进行协作分类的方式,而这些关键词一般称为标签。标签化运用了像大脑本身所使用的那种多重的、重叠的关联,而不是死板的分类。举一个经典的例子,在 Flickr 网站上,一幅小狗照片可能被加上“小狗”和“可爱”这样的标签,从而允许系统依照用户行为所产生的自然的方式来进行检索。

协作式垃圾信息过滤产品,例如 Cloudmark,就聚集了电子邮件用户们对于“一封邮件是或者不是垃圾邮件”的众多相互独立的决策,从而胜过了依赖于分析邮件本身的那些系统。

伟大的互联网成功者并不主动地到处推销其产品,这几乎成为公理。他们采用“病毒式营销”(viral marketing)的方式,也就是说,一些推介会直接从一个用户传播到另外一个用户。如果一个网站或产品依赖广告来进行宣传,几乎可以断定它不是 Web 2.0。

许多互联网基础设施本身,包括在大多数网络服务器中用到的 Linux, Apache, MySQL, 以及 Perl, PHP 或 Python 代码,都依靠开放源码的对等生产(peer-production)的方式。其中包含了一种集体的、网络赋予的智慧。在 SourceForge.net 网站上列有至少 10 万种开放源码软件项目,任何人都可以添加一个项目,任何人都可以下载并使用项目代码。

(3) 数据是下一个 Intel Inside

数据库管理是 Web 2.0 的核心竞争力,其重要性使得我们有时候称这些程序为“讯件”(infoware)而不仅仅是软件。该事实也引出了一个关键问题:谁拥有数据?

注意一下由 MapQuest、maps.yahoo.com、maps.msn.com 或者 maps.google.com 等网站提供的每张地图下面的版权声明,就会发现这样一行字“地图版权 NavTeq, TeleAtlas”,或者如果使用的是新的卫星图像服务,则会看到“图像版权 Digital Globe”的字样。这些公司对其数据库进行了大量的投资(仅 NavTeq 一家,就公布投资 7.5 亿美元用于创建其街道地址和路线数据库。Digital Globe 则投资 5 亿美元来启动其自有卫星工程,来对政府提供的图像进行改进)。NavTeq 竟然已做了很多模仿 Intel 的耳熟能详的 Intel Inside 标识的事:例如带有导航系统的汽车就带有“NavTeq Onboard”的印记。数据是许多此类程序事实上的 Intel Inside,是一些系统的唯一的信息源组件,这些系统的软件体系多数是开放源码的,也有商业化的。

当前竞争火热的网络地图(Web Mapping)领域显示着,对拥有软件核心数据的重要性的疏忽,将最终削弱其竞争地位。MapQuest 在 1995 年率先进入地图领域,随后是 Yahoo!,再后来是 Microsoft,而最近 Google 也挺进这一市场,他们可以轻松地通过对同一数据的授权来提供一个具有竞争力的程序。