

Contents

Preface xi

1 A Brief History of the Web 1

1.1	A new breed of applications: the rise of the Web	2
1.1.1	The arrival of the browser	2
1.1.2	The flattening of the world	7
1.1.3	From linking to searching	10
1.1.4	Commercialization of the Web	22
1.1.5	Peer-to-peer networks and free file sharing	34
1.2	Technological developments in a nutshell	38
1.2.1	IP networking	40
1.2.2	HTML and XML	42
1.2.3	Web services and RSS	46
1.3	User participation and contribution: socialization of the Web	49
1.3.1	Blogs and wikis	49
1.3.2	Social networks	58
1.4	Merging the streams: the arrival of “Web 2.0”	64

2 A Review of the Technological Stream 69

2.1	Developments in Web technology	69
2.1.1	HTML	71
2.1.2	XML	76
2.1.3	CSS	80
2.1.4	Scripting technologies	87
2.2	Web applications of XML	102
2.2.1	Web services	102
2.2.2	Web feed formats	115

2.3 P2P	124
2.3.1	P2P file-sharing networks
2.3.2	Other P2P applications
2.4 Summary	132

3 Enabling Techniques and Technologies 135

3.1 Rich Internet Applications	136
3.1.1	Sample RIAs: e-mail applications
3.1.2	XMLHttpRequest, the link needed for Ajax
3.1.3	More RIAs: Office and map applications
3.2 APIs, WPCs, and their mash-ups	161
3.2.1	The situation in Web 1.0
3.2.2	Content syndication with Web feeds
3.2.3	Mash-ups based on WPCs
3.3 Tagging	183
3.3.1	Flickr and tagging
3.3.2	Social bookmarking
3.3.3	Folksonomies
3.4 Summary	203

4 Sample Frameworks for Web Application Development 205

4.1 Development methodologies	206
4.2 Client-side Ajax frameworks	211
4.2.1	Kabuki Ajax toolkit
4.2.2	Development of a Zimlet using AjaxTK
4.3 Server-side frameworks	222
4.3.1	Ruby on Rails
4.3.2	Creating a Web application with Rails
4.4 Frameworks for other RIA technologies	232
4.4.1	RIA development with OpenLaszlo
4.4.2	Flash versus Ajax
4.5 Summary	238

5 Impacts of the Next Generation of the Web 241

5.1 Business models for Internet and Web	242
5.1.1	Commission-based brokerage and merchants
5.1.2	Advertising
5.1.3	Information intermediaries
5.1.4	The community and the subscription models
5.1.5	Summary

5.2	Data ownership	251
5.3	Software as a Service (SaaS)	254
5.3.1	A look back: the ASP model	256
5.3.2	The provider-oriented view	257
5.3.3	The consumer-oriented view and service customization	263
5.4	Socialization and cocreation of content	266
5.4.1	Social search	266
5.4.2	Social aspects of software	271
5.4.3	Impacts of online social networks	273
5.4.4	User-generated content in advertising	275
5.4.5	Second Life	276
5.5	Summary	279

6 The Semantic Web and Web 2.0 281

6.1	Basics	283
6.1.1	Search revisited	283
6.1.2	Data and information integration	287
6.1.3	The Semantic Web idea	289
6.1.4	The structure of the Semantic Web	292
6.2	Languages of the Semantic Web	295
6.2.1	The Resource Description Framework (RDF)	295
6.2.2	RDF Schema (RDFS)	298
6.2.3	Web Ontology Language (OWL)	303
6.3	Ontologies	308
6.3.1	Introduction	309
6.3.2	Design of an ontology	310
6.3.3	OntoMedia: an ontology-based personal entertainment system	320
6.4	From tagging to ontologies and back	323
6.4.1	Micro-formats	325
6.4.2	Collaborative tagging and folksonomies	328
6.5	Summary	334

References 337

Index 345

Index

Action Controller, Ruby on Rails, 229–231
Action View, Ruby on Rails, 227–229
Active Record, Ruby on Rails, 224–227
AdSense
 principles, 32, 247
 social aspects, 271
Advertising
 AdSense, 247
 AdWords, 246–247
 banner ad, 245
 e-commerce, 31–32
 micropayments, 248
 paid clicks, 245–246
 traditional features, 245
 user-generated content, 275–276
AdWords, principles, 32, 246–247
Agent-based software, e-commerce, 26
Ajax, *see also* Kabuki Ajax Toolkit
 asynchronous page requests, 146, 148
 evolution, 70
 map applications, 159
 Open Ajax Initiative, 213
 overview, 46
 popularity, 206
 synchronous page requests and refreshes,
 146–147
 Web interaction model, 146, 148
 XMLHttpRequest, 141–142, 146, 149
AllPeers, network overview, 128–129
Amazon Mechanical Turk, features, 263–264
Amazon Web Services, overview, 260–261

API, *see* Application programming interface
Application development, *see also* Rich Internet
 Applications
 agile software methods, 208–209
 client-side Ajax frameworks, 211–222
 server-side frameworks, 222–232
 traditional software methods, 207–208
 Web 2.0 overview, 209–211
Application programming interface (API)
 functions, 167
 Google Maps, 175–178
 mash-ups, *see* Mash-up
 service calls, 66–67
Application Service Provisioning (ASP), over-
view, 256–257
Application stream
 definition, 1
 Web 2.0, 64
Artificial intelligence, Semantic Web, 291, 294
ASP, *see* Application Service Provisioning
Asynchronous JavaScript and XML, *see* Ajax
ATOM
 Blogger feed creation, 118
 OpenSearch namespace, 120–121
 popularity, 118
 sample code, 119–120
Automated reasoning, Semantic Web, 294
B2B, *see* Business-to-business
B2C, *see* Business-to-customer
Banner ad, overview, 245

- BitTorrent, network overview, 126–128
- Blog
- Blogger feed creation, 118
 - code of ethics, 53
 - definition, 50
 - e-commerce items, 30–31
 - examples, 50–51
 - growth, 52
 - misuse, 53–54
 - resources, 52
 - searching, 63
 - social aspects, 272
 - trust, 53
- Bookmark, browser, 163
- Broadband, advantages and growth, 41–42
- Broker, business models, 243–244
- Browser
- bookmarks, 163
 - historical perspective, 2–7
- Business models
- advertising, 245–248
 - classification, 242–243
 - commission-based brokerage and merchants, 243–245
 - community model, 249–250
 - components, 242
 - information intermediaries, 248–249
 - subscription model, 250
- Business-to-business (B2B), overview, 27–28
- Business-to-customer (B2C), overview, 27–28
- C2C, *see* Customer-to-customer
- Cascading style sheet (CSS)
- Extensible Markup Language implementation, 81–84
 - Hypertext Markup Language implementation, 84–86
 - presentation, 43
- CDnow, historical perspective, 23–24
- CGI, *see* Common Gateway Interface
- ChaCha, search interface, 284–285
- Client/server principle, 3–4
- CMS, *see* Content management systems
- Collective knowledge, definition, 329
- Commercialization, *see* E-commerce
- Commission, business models, 243
- Common Gateway Interface (CGI), server-side scripting, 88
- Common Request Broker Architecture (CORBA), Web service, 102, 104
- Community model, overview, 249–250
- Conceptual design, ontology, 314–316
- Content management systems (CMS), overview, 34
- Content syndication
- acquiring, 168–169
 - content attribution, 172
 - permalinks, 172
 - Pipes, 170–171
 - transformation, 168–169
- CORBA, *see* Common Request Broker Architecture
- CSS, *see* Cascading style sheet
- Customer-to-customer (C2C), overview, 27–29
- DAML, OWL development, 307
- Dashboard, gadgets, 182
- Data
- collection, 252–253
 - machine-generated data, 251–252
 - ownership, 251–254
 - responsibilities of owners, 253–254
- Data stream, Web 2.0, 67
- DCMI, *see* Dublin Core Metadata Initiative
- DCOM, *see* Distributed Component Object Model
- del.icio.us
- social bookmarking, 196–198
 - structured tagging, 331
- DHTML, *see* Dynamic Hypertext Markup Language
- Digg, social networking, 59
- Dilbert blog, overview, 50–51
- Directory
- overview, 19
 - search engine competition, 20
- Distributed Component Object Model (DCOM), Web service, 104
- DNS, *see* Domain Name System
- Document Type Definition (DTD), Extensible Markup Language, 79–80
- Domain Name System (DNS), function, 41

- Dot-com bubble, overview, 29
DTD, *see* Document Type Definition
Dublin Core Metadata Initiative (DCMI), elements, 301–302
Dynamic Hypertext Markup Language (DHTML)
client-side scripting, 93–95
evolution, 70
- eBay, feedback, 271
- E-commerce, *see also* Business models
advertising and marketing, 31–32
classification, 27–29
components, 24–26
customer feedback and recommendations, 29–31
growth, 33–34
historical perspective, 22–23
payment and trusted third parties, 32–33
software, 26
- E-mail, Rich Internet Applications
comparison of services, 145–146
Outlook, 138–139
Outlook Web Access, 139–141
search engine mail clients, 142–144
Springdoo video e-mail, 144–145
XMLHttpRequest, 141–142, 146, 149
Zimbra, 144
- Enfinity Suite 6, architecture, 26–27
- Enterprise resource planning (ERP), e-commerce, 27
- Entropia Universe, features, 278
- ERP, *see* Enterprise resource planning
- Eurekster, features, 270
- Extensible Hypertext Markup Language (XHTML), overview, 80
- Extensible Markup Language (XML)
application- versus technology-oriented languages, 76–77
base technology, 76–77
Document Type Definition, 79–80
evolution, 70
frameworks, 77
language development, 43–44
namespace, 44
overview, 43
- RSS applications, 48
sample code, 78–79
schema definition, 43–44, 79
tags and elements, 77
Web feed formats, 115–124
Web services, 102–115
Zimlet declaration, 218
- Extensible Markup Language, cascading style sheet implementation, 81–84
- Extreme Programming (XP), application development, 208–209
- Facebook, social networking, 61
- FastTrack, network overview, 126
- Flash, Rich Internet Application development, 236–238
- Flickr
social aspects, 272–273
tags, 185–191, 332–333
- Folksonomy
collaborative tagging and folksonomies in Semantic Web
collective knowledge, 329
overview, 328–329
serendipity, 330
structured tagging, 331–334
tag qualitative classes, 329–330
definition, 201
social search, 267–268
tagging, 201–203
- Friedman, Thomas, world flatteners, 7–9
- Friendster, social networking, 61
- Gliffy, features, 156–157
- Gnutella, principles, 35–36
- Gold Box, Amazon.com, 30
- Google Apps, overview, 258
- Google Mail, features, 142–144
- Google Maps
application programming interface, 175–178
features, 158–160
mash-up development, 178–181
- Grid computing, applications, 39
- hCard, features, 325–327
- HIT, *see* Human Intelligence Task

- HTML, *see* Hypertext Markup Language
HTTP, *see* Hypertext Transfer Protocol
Human Intelligence Task (HIT), Amazon Mechanical Turk, 263–264
Hyperlink
navigation, 11
overview, 10
Hypertext Markup Language (HTML)
cascading style sheet implementation, 84–86
document type definitions, 72–73
evolution, 70
history and advantages, 4–5, 42–43
links as anchors, 10–11
sample code, 72
structure versus presentation, 74–75
tags
overview, 43, 71, 73–75
standards and development, 75–76
Hypertext Transfer Protocol (HTTP), historical perspective, 4–5
- ICANN, *see* Internet Corporation for Assigned Names and Numbers
Information broker, function, 12–13
Information intermediary, business models, 248–249
Internet, development stages, 40–41
Internet Corporation for Assigned Names and Numbers (ICANN), [V1]origins and functions, 41
Internet Protocol television (IPTV), prospects, 42
IPTV, *see* Internet Protocol television
- JavaScript
client-side scripting, 45, 88–93
overview, 45
- Kabuki Ajax Toolkit
components
DHTML Widget Toolkit, 215
event model, 214
network communication, 215
overview, 213–214
Open Ajax Initiative, 213
Zimbra, 212
- Zimlet development, 216–222
Knowledge management, social networking, 60
- Link, *see* Hyperlink
Lixto Suite, features, 164–166
Logical design, ontology, 316–319
- Map applications
Google Maps, 158–160
Yahoo! Maps, 157
- Mash-up
definition, 160–161
Google Maps
application programming interface, 175–178
mash-up development, 178–181
start page personalization, 181–183
tags, 191–195
Web procedure call mash-ups
overview, 172–173
styles, 173–175
wrappers as precursors, 164
- MeCommerce, definition, 265
- Merchant model, overview, 244
- Metadata, Semantic Web, 290–291
- Meta-tag, overview, 183–194
- Micro-formats, ontology, 325–328
- Mindset, overview, 16–17
- Moore's Law, 38–39
- Mosaic, historical perspective, 2–3
- MoveOn, social networking, 63
- MySpace, social networking, 59, 61
- Namespace
Extensible Markup Language, 44
Resource Description Framework Schema, 298
SOAP, 113
Web Services Description Language, 109–110
- Napster, network overview, 125
- Netscape
dot-com bubble, 29
historical perspective, 5–7
initial public offering, 6–7, 9
- Nextlinks, features, 284
- Ning, social networking, 62

- Office applications
Gliffy, 156–157
ThinkFree Office, 155–156
- OIL, *see* Ontology Interface Layer
- Ontology
design
approach, 311–312
conceptual design, 314–316
logical design, 316–319
multimedia data hierarchies, 314–315
physical design, 319
requirements analysis, 312–313
requirements specification, 313
running example, 311
software, 310
overview, 308–310
- Ontology Interface Layer (OIL), OWL development, 307
- OntoMedia, features, 309, 320–323
- Open access, file sharing, 37
- OpenLaszlo
Ajax support, 238
Flash utilization, 236–238
overview, 232–233
Rich Internet Applications
deployment, 235–236
development process, 233–235
server architecture, 236
- Outlook, features, 138–139
- Outlook Web Access (OWA), features, 139–141
- OWA, *see* Outlook Web Access
- OWL, *see* Web Ontology Language
- P2P networks, *see* Peer-to-peer networks
- Page scraping, overview, 165, 167
- Pageflakes, start page, 181–182
- PageRank, origins and function, 15–16
- Paid clicks
AdWords, 246–247
fraud, 247
principles, 245–246
- Pandora, user interface, 237
- PayPal, payment and trusted third parties, 32–33
- PC, *see* Personal computer
- Peer-to-peer (P2P) networks
- examples
AllPeers, 128–129
BitTorrent, 126–128
FastTrack, 126
Napster, 125
Skype, 130–132
- file sharing, 36–37
- Gnutella, 35–36
- open access, 37
- overview, 34–35
- Permalink, definition, 172
- Personal computer (PC), evolution, 38–39
- PHP
applications, 45
server-side scripting, 95–101
- Physical design, ontology, 319
- Pipes, content syndication, 170–171
- Portal
definition, 19
search engine competition, 20–21
- Prefound, features, 270
- Rails, *see* Ruby on Rails
- RDF, *see* Resource Description Framework
- RDFS, *see* Resource Description Framework Schema
- RealTravel, structured tagging, 333–334
- Requirements analysis, ontology design, 312–313
- Requirements specification, ontology design, 313
- Resource Description Framework (RDF)
property, 295–296
resource, 295
Semantic Web, 291, 293, 295–298
- Resource Description Framework Schema (RDFS)
class hierarchy, 299–300
- Dublin Core Metadata Initiative, 301–302
- instance, 300–301
- namespace, 298
- property, 299
- Semantic Web, 291, 298–302
- Web Ontology Language class relationships, 304–305
- RIA, *see* Rich Internet Application

- Rich Internet Application (RIA)
desktop application comparisons, 160–161
Flash versus Ajax, 236–238
 OpenLaszlo, 232–236
e-mail applications
 comparison of services, 145–146
 minimum requirements, 137–138
 Outlook, 138–139
 Outlook Web Access, 139–141
 search engine mail clients, 142–144
 Springdoo video e-mail, 144–145
 XMLHttpRequest, 141–142, 146, 149
 Zimbra, 144
map applications
 Google Maps, 158–160
 Yahoo! Maps, 157
office applications
 Gliffy, 156–157
 ThinkFree Office, 155–156
overview, 136–137
Web procedure call samples, 149–155
Rollyo, features, 268–270
RSS
 applications, 48
 Blogger feed creation, 118
 overview, 48, 116
 popularity, 118
 promoting feeds, 122–123
 RSS 2.0 sample code, 121–122
 syndication feed elements, 116–118
Ruby on Rails
 application development
 Action Controller, 229–231
 Action View, 227–229
 Active Record, 224–227
 agility, 231–232
 model view controller and Rails components, 223–224
 overview, 222–223
- SaaS, *see* Software as a Service
SCM, *see* Supply-chain management
Scripting
 client side versus server side, 45, 87–88
Common Gateway Interface, 88
Dynamic Hypertext Markup Language, 93–95
JavaScript, 45, 88–93
languages, 87
PHP, 95–101
synchronous versus asynchronous data processing, 46
Search engine
 blogs, 53
 components, 14–15
 historical perspective, 13–14
 long tail of queries, 18
 PageRank, 15–16
 personalization, 267
 popular searches, 17
 prospects, 16–17
 Semantic Web augmentation, 283–287
 social search engines, 266–271
 tagging, 18
Second Life
 applications, 277–278
 features, 276–277
 growth, 276
 origins, 276
Semantic Web
 collaborative tagging and folksonomies
 collective knowledge, 329
 overview, 328–329
 serendipity, 330
 structured tagging, 331–334
 tag qualitative classes, 329–330
 integration
 data, 287–288
 information, 288
 materialized integration, 288–289
 virtual integration, 288–289
languages
 Resource Description Framework, 295–298
 Resource Description Framework Schema, 298–303
 Web Ontology Language, 303–308
metadata, 290–291
micro-formats, 325–328
ontology design
 approach, 311–312
 conceptual design, 314–316
 logical design, 316–319

- multimedia data hierarchies, 314–315
overview, 308–310
physical design, 319
requirements analysis, 312–313
requirements specification, 313
running example, 311
software, 310
overview, 281–283
prospects, 334–336
search augmentation, 283–387
structure, 292–295
tag clouds, 324
vision, 289–290
Web 2.0 convergence, 335–336
- Serendipity, definition, 330
- SeRQL, Semantic Web, 296
- SGML, *see* Standard Generalized Markup Language
- Skype
network overview, 40, 130–132
social aspects, 272
- Slashdot, historical perspective, 50
- Snap, features, 284
- SOAP
e-commerce samples, 112, 114–115
namespaces, 113
overview, 103, 106
structure, 113
- Social bookmarking
del.icio.us, 196–198
overview, 195–196
Yahoo! MyWeb 2.0, 199–201
- Social network
analysis, 63–64
collective knowledge, 329
community types, 59
functions, 60
impacts, 273–274
MySpace, 61
popularity, 59
social search engines, 266–271
software
overview, 62
social aspects, 271–272
- Software as a Service (SaaS)
Application Service Provisioning, 256–257
- consumer-oriented view and service customization, 263–266
costs, 254
forms, 255
provider-oriented view
application services, 258–260
infrastructure services, 260–262
service lock-in, 262
- SongBird, features, 259–260
- SPARQL, Semantic Web, 293
- Springdoo, video e-mail, 144–145
- Standard Generalized Markup Language (SGML), overview, 43
- Structured tag, overview and application, 331–334
- Subscription model, overview, 250
- Supply-chain management (SCM),
e-commerce, 27
- Tagging
collaborative tagging and folksonomies in Semantic Web
collective knowledge, 329
overview, 328–329
serendipity, 330
structured tagging, 331–334
tag qualitative classes, 329–330
e-commerce items, 30
examples, 184–185
Flickr, 185–191
folksonomic tagging, 201–203
historical perspective, 183–184
mash-ups, 192–195
tag clouds, 324
- TCP/IP, *see* Transmission Control Protocol/Internet Protocol
- ThinkFree Office, features, 155–156
- Touchgraph, features, 284
- Transmission Control Protocol/Internet Protocol (TCP/IP), packets, 40
- Tribe, social networking, 62
- UDDI, *see* Universal Description, Discovery, and Integration
- Unicode, Semantic Web, 292

- Uniform Resource Identifier (URI), Extensible Markup Language, 44
- Universal Description, Discovery, and Integration (UDDI), Web Services Description Language, 103
- Universal resource locator (URL), definition, 11
- URI, *see* Uniform Resource Identifier
- URL, *see* Universal resource locator
- User-generated ads, features, 275–276
- User participation and contribution
 blogs, 49–54
 definition, 1
 social networking, 58–64
 wikis, 54–58
- Virtual communities, e-commerce, 30
- Voice over Internet Protocol (VoIP), principles, 40
- VoIP, *see* Voice over Internet Protocol
- W3C, *see* World Wide Web Consortium
- WDL, *see* Web Services Description Language
- Web 3.0, prospects, 68
- Web feed, *see also* RSS
 ATOM, 118–121
 content syndication, *see* Content syndication
 overview, 48
 prospects, 123–124
 specification and history, 115–116
 syndication feed elements, 116–118
- WebFountain, features, 284–285
- Web Ontology Language (OWL)
 classes, 304–306
 inferences, 307–308
 precursor languages, 307
 properties
 data type, 305
 description, 306
 object, 305
- Resource Description Framework Schema
 class relationships, 304–305
- Semantic Web, 291, 303–308
 sublanguages, 303–304
- Web procedure call (WPC)
 mash-ups, *see* Mash-up
- overview, 105–107
- Rich Internet Applications, 149–155
- software as a service, 255
- Web service
 decomposition, 105
 definition, 102
 design, 105
 infrastructure, 106
 paradigm, 103
 service-oriented architectures, 104
 SOAP, 103, 106, 112–115
 standards stack, 104
 structure, 110–112
 Web procedure calls, 105–107
 Web Services Description Language, 103, 108–112
- Web Services Description Language (WSDL)
 e-commerce sample, 108–109
 namespaces, 109–110
 overview, 103
- Widget, definition, 181
- Wiki
 definition, 54
 editing, 54–56
 historical perspective, 54
 Wikia, 56
- Wikipedia
 accuracy, 57
 bibliographic details, 56
 code of ethics, 57
 historical perspective, 56
 social aspects, 271–272
- Wish list, e-commerce items, 31
- World Wide Web Consortium (W3C), origins and functions, 41
- WPC, *see* Web procedure call
- Wrapper
 features, 164, 166
 XMLHttpRequest, 205–206
- XHR, *see* XMLHttpRequest
- XHTML, *see* Extensible Hypertext Markup Language
- XML, *see* Extensible Markup Language
- XMLHttpRequest (XHR)
 Ajax, 146, 149

overview, 141–142
Web procedure call samples, 149–155
wrappers, 205–206
XP, *see* Extreme Programming

Yahoo! Mail, features, 142–144
Yahoo! Maps, features, 157
Yahoo! MyWeb 2.0, social bookmarking,
 199–201
Yahoo!Search MyWeb2, features, 270
YouTube
 social impact, 274
 social networking, 63

Zimbra
 features, 144
 Kabuki Ajax Toolkit, 212

Zimlet
 content objects and applications, 216–218
 declaration in XML, 218
 Kabuki Ajax Toolkit for development,
 218–222
 panel items, 216