Unleashing Web 2.0
From Concepts to Creativity
Unleashing Web 2.0
From Concepts to Creativity

Gottfried Vossen
Stephan Hagemann
Dedication

To Martina, in remembrance of our Aotearoa year – Gottfried Vossen
To Julia – Stephan Hagemann
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
CONTENTS

2.3 P2P  124
  2.3.1 P2P file-sharing networks  124
  2.3.2 Other P2P applications  128

2.4 Summary  132

3 Enabling Techniques and Technologies  135

3.1 Rich Internet Applications  136
  3.1.1 Sample RIAs: e-mail applications  137
  3.1.2 XMLHttpRequest, the link needed for Ajax  146
  3.1.3 More RIAs: Office and map applications  155

3.2 APIs, WPCs, and their mash-ups  161
  3.2.1 The situation in Web 1.0  162
  3.2.2 Content syndication with Web feeds  167
  3.2.3 Mash-ups based on WPCs  172

3.3 Tagging  183
  3.3.1 Flickr and tagging  185
  3.3.2 Social bookmarking  195
  3.3.3 Folksonomies  201

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  212
  4.2.2 Development of a Zimlet using AjaxTK  216

4.3 Server-side frameworks  222
  4.3.1 Ruby on Rails  222
  4.3.2 Creating a Web application with Rails  224

4.4 Frameworks for other RIA technologies  232
  4.4.1 RIA development with OpenLaszlo  232
  4.4.2 Flash versus Ajax  236

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  243
  5.1.2 Advertising  245
  5.1.3 Information intermediaries  248
  5.1.4 The community and the subscription models  249
  5.1.5 Summary  250
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
During the year 2006, everything seemed to have come out in Version 2, Release 0 (commonly abbreviated as “2.0”): The IEEE Spectrum journal reported on Silicon Valley 2.0 (beta) in its August 2006 issue, German tech news ticker heise.de stated in September that “technology keeps the Family 2.0 together,” the Australian edition of the T3 tech magazine described Gadgets 2.0 in its October 2006 issue, the San Jose Mercury News had a story on India 2.0 in early December. There are many other such examples, including Pub 2.0, Jobs 2.0, Health 2.0, Entertainment 2.0, Business 2.0, or Music 2.0. All of this can be traced back to O’Reilly Media, where the term was coined in late 2004. As can be read at www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html.

The concept of “Web 2.0” began with a conference brainstorming session between O’Reilly and MediaLive International. Dale Dougherty, Web pioneer and O’Reilly vice president, noted that far from having “crashed,” the Web was more important than ever, with exciting new applications and sites popping up with surprising regularity. What’s more, the companies that had survived the collapse seemed to have some things in common. Could it be that the dot-com collapse marked some kind of turning point for the Web, such that a call to action such as “Web 2.0” might make sense? We agreed that it did, and so the Web 2.0 Conference was born.

“Web 2.0” has rapidly become a label that everybody seems to be able to relate to: Hardware manufacturer Intel suddenly entered the software market and created an Internet Business Suite called “SuiteTwo,” an integrated package of Web 2.0 software. U.S. marketing research giant Gartner recognized a major driver for the IT industry in Web 2.0 technologies and awarded it the “Hype Oscar” of 2006. SEOmoz, a Seattle-based search engine optimization company even rated and ranked more than 300 “Web 2.0 companies” in thirty-eight categories and gave a “Web 2.0 Award” 2006 in twenty-one of them (http://www.seomoz.org/web2.0/). Finally, Time magazine made “You” the Person of the Year.
In its December 2006 issue, acknowledging the fact that the Web meanwhile heavily benefits from user contributions in a variety of media forms. On the other hand, critics are already warning of a “Bubble 2.0,” while the New York Times jumped to a “Web 3.0” in November 2006.

But what actually is “2.0” in the context of the Web? What is new and what is not? The media have tried to answer this question in the recent past. An exploding number of blogs, showcases, magazine and journal special issues, the famous O’Reilly “summit” conferences, and news messages have offered numerous “definitions” and explanations of Web 2.0, what its features are, what can be expected from it in the very near future, and how important it is to join the bandwagon. Among other effects, this has made many people and companies nervous. They now ask themselves questions such as, “Are we Web 2.0 ready?” or “Do we comply with Web 2.0?” or “What will happen to us if we do not convert to this movement?” Although various attempts have been made to provide a precise definition, we are looking at a moving target characterized by a certain amount of fuzziness that lacks a simple and concise description and impact analysis.

The aim of this book is to remedy this situation by taking a critical look at developments commonly attributed to Web 2.0, by putting them in perspective, in particular with respect to the “historic” development of the Web, and by describing the current state of things from that perspective. The approach we take does not try to bring together each and every “definition” that we could find on the subject. Instead, we take a look at the subject from a technical and from a business point of view. The former allows us to draw several lines of development whose current end point is Web 2.0, but whose roots can be clearly recognized and identified. Moreover, several services and sites have been around for a while already and turn out not to be a result of Web 2.0, but one of its prerequisites. The business perspective lets us identify various ways of monetizing the Web, and we again try to analyze what is new and what is not. One of the findings resulting from this is that the term “Web 2.0” should actually have been “Internet 2.0.” However, we stick to the former term, being aware of the discussion it provokes.

We start out in Chapter 1 with a historical perspective that looks at how the Web has emerged since its inception in 1993. We look at technological advances that have paved the way for the ubiquitous spread of the Web we have experienced in recent years, and at how the perception and usage of the Web by people and their interaction with this medium have changed over the years, in particular when looking at the younger generation. In Chapter 2, we take a closer look at the advances in Web technology, with an emphasis on languages, formats, and standards that have been developed in recent years and that see a confluence today in various ways. Chapter 3 presents the core Web 2.0 technical mechanisms, which are APIs, Web procedure calls, and their mash-ups, rich Internet applications, and tagging. In Chapter 4, we complement this with sample frameworks for Web development, including AjaxTK, OpenLaszlo, and
Ruby on Rails. In Chapter 5, we study impacts of the next generation of the Web. The major ones we see are data ownership, software as a service, and the socialization and cocreation of content, which includes topics such as social networks and social software. In Chapter 6, we look at another important Web development that has been around for quite a number of years already, the Semantic Web. We contrast Web 2.0 with the Semantic Web and try to elaborate on the aspects that will make them fit together.

With a topic selection such as this, we should say a word to the intended readership for the book. The techie maintaining one of the many blogs covering the topic on an almost daily basis is certainly not our target audience. Instead, the book is intended for readers who have a basic understanding of what the Web is about and who are interested in finding the various facets of the present-day Web in one place. Depending on how deep a reader’s knowledge about HTML, scripting languages, or P2P protocols actually is, he or she may skip Chapter 2 and jump right into the newer technologies dealt with in Chapter 3. By a similar token, a reader mostly technically, but less conceptually interested, may decide to skip Chapter 6. Another type of reader may take our detailed exposition of a single running example in Chapter 2 as an introduction into the various technologies underlying the Web and as a starting point for further studies. We are aware that the audience for our book will be diverse, but we hope that the brief overview of the contents given here can serve as a guideline for what to read and what to skip.

Since we are from academia, we have been thinking for some time about how to integrate Web 2.0 into Computer Science (CS) and Information Systems (IS) educational programs. We often approach new topics through seminars in which students have to prepare presentations and write papers. We have done the same with the Web 2.0 topic in the summer term of 2006, yet the material produced by our graduate IS students at the University of Münster, Germany, was so interesting that we were motivated to develop it further. We are thus grateful to Björn Böhmer, Florian Christ, Nicolas Dillmann, Henning Eiben, Christian Heitmann, Alexander Rajko, Julian Reimann, Jens Sieberg, Björn Skall, Steffani Ungerath and Heiko Zeus for granting us access to their material. The seminar and therefore the book have also benefited from discussions with students of Professor Wojciech Cellary from the Economical University of Poznan, Poland, with whom we met in Potsdam near Berlin during May 2006 to listen to some of the presentations the students had to prepare.

When it came to finding a title for the book, one of the greatest challenges arose. We are aware that having “Web 2.0” in the title may help sell the book today, but what if tomorrow the term is replaced by “Web 3.0”? Other books have had this problem, yet we are willing to take a risk here. We are convinced that the term will stick for a while, and we will modify the title in the future if necessary.

We need to say a word about the use of URLs in this book. It is clear that Web addresses are all over the text, as they represent the numerous examples
we have explored, looked at, and found worth mentioning. However, we have decided not to state URLs explicitly each time, and the rule we apply is the following. If a company by name xxx can simply be found on the Web by going to www.xxx.com, we do not mention the URL. However, we do as soon as the URL deviates in the slightest way from this generic format.

Finally, the book was designed and written during the sabbatical stay of the first author at the University of Waikato in Hamilton, New Zealand. We are grateful to Professor Jim Corner, chairman of the Department of Management Systems in the University of Waikato Management School, for his hospitality, in particular during the time when both authors stayed in New Zealand. Clearly, working in a country where November and December, known to be dark and wet months in Central Europe, fall into early summer is particularly fun, and so we were able to take the efforts of writing this book out into nature at times, for example into the famous Polynesian Spa in Rotorua (www.polynesianspa.co.nz/).

As is often the case with books, this has been an effort in which many people were involved, some in the design of the structure, and many in the development of the various chapters. In particular, we are indebted to Jim Anderton, Doug Barry, Jim Corner, Bodo Hüsemann, Catherine Legg, Andrew Otwell, Joachim Schwieren, Dennis Shasha, Andreas Simon, Todd Stephens, Gunnar Thies, and Ian Witten for all the feedback they have provided, ranging from comments on specific sections to in-depth reviews of the entire manuscript. The book has benefited considerably from these interactions, and of course it remains our responsibility what we have made of it.

We are grateful to our publisher, Diane Cerra, who believed in the project from the moment we suggested it to her, and who has once more accompanied us very professionally through an intense writing and production process. We are also grateful to Mary James (assistant editor), Monica Mendoza (production) and Misty Bergeron (marketing manager). Last, but not least, the artwork on the cover was painted by Laura, Gottfried’s oldest daughter, following a suggestion from Jens Sieberg. It is based on the idea that the Web is currently getting a substantial overhaul during which old, static stuff gets replaced by fresh, dynamic new elements and colors, and the tools that are used to this end include Ajax, tags, and a few other ingredients. The image is also reminiscent of the New Zealand fern as well as Maori ornaments common all over Aotearoa, which both symbolize very well the connection between established traditions and evolving nature that we also find on the Web today.

Gottfried Vossen and Stephan Hagemann
Hamilton, New Zealand, and Münster, Germany, March 2007

We plan to make additional material available over time at
www.unleashingweb20.com