The Web Was Done by Amateurs

Marco Aiello

The Web Was Done by Amateurs

A Reflection on One of the Largest Collective Systems Ever Engineered

Springer

Marco Aiello
University of Stuttgart
Stuttgart, Germany

ISBN 978-3-319-90007-0 ISBN 978-3-319-90008-7 (eBook)
https://doi.org/10.1007/978-3-319-90008-7

Library of Congress Control Number: 2018939304

© Springer International Publishing AG, part of Springer Nature 2018, corrected publication 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The original version of this book was revised. An correction to this book can be found at https://doi.org/10.1007/978-3-319-90008-7_12

Preface

The field of computer science is so young that sometimes we think of it as history-less, as a set of cutting-edge technologies without a past. This is a crucial mistake. The field might be relatively young, especially when compared with other traditional exact sciences, such as mathematics and physics, but it has a very dense history. A fitting comparison is the life expectancy of a dog vs. that of a human: a year in computer science is equivalent to seven years in other scientific fields. On the one hand, such speed of innovation is exciting and one of computer science's characterizing features; on the other hand, it too often prevents us from reflecting on the history, and consequently we reinvent the wheel.

In my 20 years as a lecturer of computer science, I have noticed that students are often incredibly skilled in the latest technologies but are not able to place them into their historical and societal context. Something like the Web is taken for granted. Occasionally, a student will place the Web's birth in the 1950s. The problem becomes even more evident when they start designing a system for their final project. The intuitions and ideas may be very worthwhile, but often they have been proposed before, unbeknownst to the student. My feeling is that they lack heroes and role models. They lack an Einstein or Fermi to look up to, a Freud or a Jung to place at the origin of their field. This gap is not due to the absence of exceptional computer science founding fathers—and mothers. It is rather that most ignore the origins of a model, an idea, a technique, or a technology. Who invented the Web? When? Who proposed object-oriented programming? Why? Who coined the term Artificial Intelligence? How is it defined? These are questions that Web engineers, software engineers, and Artificial Intelligence students—not to mention the general public—too often cannot answer.

The present book was born with the desire to systematize and fix on paper historical facts about the Web. No, the Web was not born in the 1950s; it is not even 30 years old. Undoubtedly, it has changed our lives, but it has done so in just a few decades. So, how did it manage to become such a central infrastructure of modern society, such a necessary component of our economic and social interactions? How did it evolve from its origin to today? Which competitors, if any, did it have to win over? Who are the heroes behind it? These are some of the questions that the present book addresses. The book also covers the prehistory of the Web so as to better understand its evolution.

Even if it is perhaps obvious, it is still worthwhile to remark that there is an important difference between the Web and the Internet. The Web is an application built over the Internet. It is a system that needs a communication infrastructure to allow users to navigate it and follow a link structure distributed among millions of Web servers. The Internet is such an infrastructure, allowing computers to communicate with each other. The confusion sometimes arises due to the fact that the Web and its companion email are the most successful applications over the Internet. Nevertheless, the Web and the Internet are two distinct systems. The present book is about the Web. It will often refer to the Internet, as the relation between the two is very close indeed, but the book focuses only on the Web.

The book is organized into four parts. *Part I: The Origins* covers the prehistory of the Web. It looks at the technology that preexisted the Web and fostered its birth. It also covers earlier hypertextual systems that have preceded the emergence of the Web. The narrative is historical in nature with many references and quotations from the field's pioneers.

Part II: The Web describes the original Web proposal as defined in 1989 by Tim Berners-Lee and the most relevant technologies associated with it. The presentation is mostly historical in nature.

Part III: The Patches combines the historical reconstruction of the evolution of the Web with a more critical analysis of the original definition and of the necessary changes to the initial design. The presentation has both an historical and an engineering flavor.

Finally, *Part IV: System Engineering* looks at the Web as an engineered infrastructure and reflects on its technical and societal success. The narrative here predominantly takes a system's engineering view, considering the Web as a unique, gigantic case study. There are occasional historical elements and a few considerations with a philosophy of science twist to them.

The book was written with the technological-engaged and knowledge-thirsty reader in mind, ranging from the curious daily Web user to the computer science and engineering student. People with diverse backgrounds

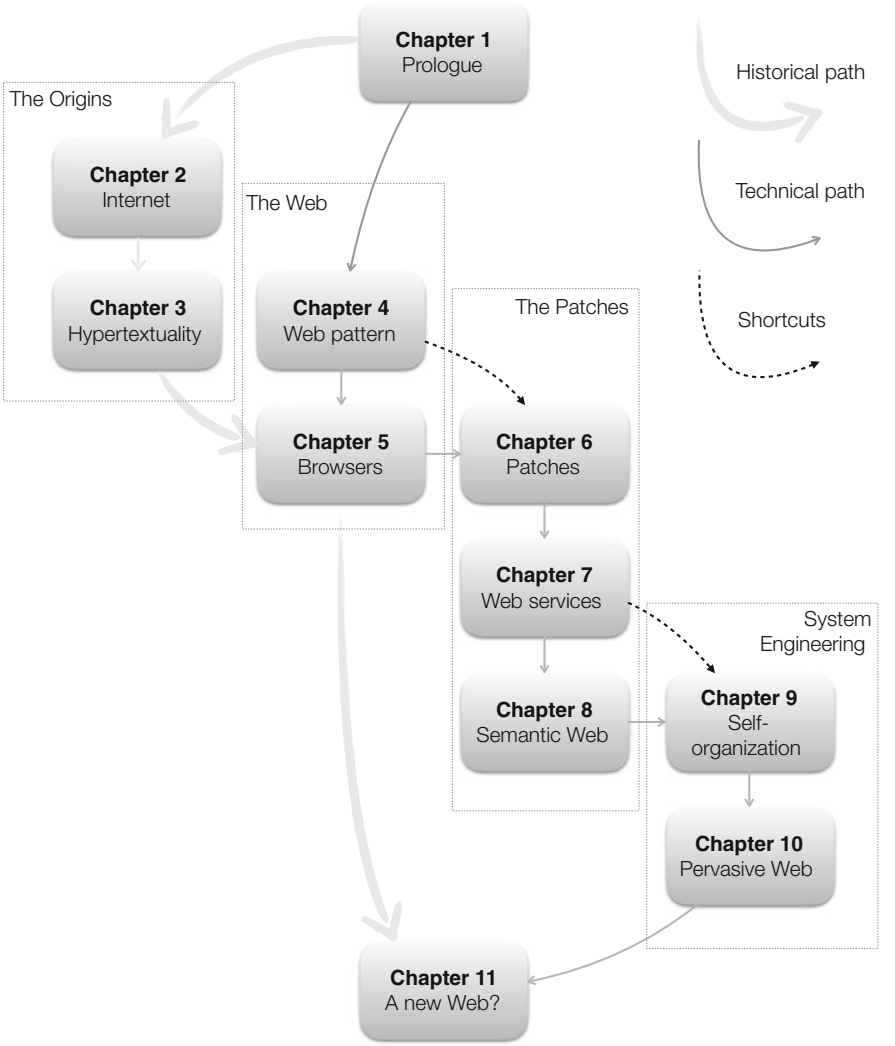

Fig. 1 Possible reading paths

might want to personalize their reading experience. The more historically oriented reader who has less background and interest in computer science should follow the thick, gray arrow on Fig. 1, most notably skipping Part III and optionally going through Part IV. Similarly, those already familiar with the history of the Internet and of the prehistory of the Web can follow the thin, gray line in Fig. 1 and go for the more technical chapters. Two chapters can be considered optional: Chap. 8 on the Semantic Web is slightly more technical

than the rest and can be safely skipped. Chapter 5 on Web browsers and their wars has a vintage taste that will appeal to the baby boomers, but may be less relevant to the millennials.

In looking at the history and evolution of the Web, we will encounter many interesting characters and pioneers. A few recur throughout the history and will be often present. The most notable three are Tim Berners-Lee, who invented the Web; Alan Kay, who is one of the founding fathers of computer science and has a strong feeling about the Web (he also inspired the title of the present book); and Ted Nelson, who defined the field of hypertextuality with his pioneering Xanadu system. Could these be the heroes that computer science generations need? For sure they are visionaries to look up to and who will be remembered.

I have based the historical reconstruction presented here on many books, papers, and Web pages. These are all cited throughout the book. I have also employed facts from my personal experience or directly communicated to me by prominent colleagues. Wikipedia has often been the starting point for my research. I did not put references to the Wikipedia entries, though, as they are quite straightforward and I can imagine anybody being able to just input the keywords in the Wikipedia search box. As a sign of my appreciation, I did regularly donate to the Wikipedia Foundation, and I plan to do so again in the future. If you have downloaded this book for free from the Internet, you know, kind of illegally, I do suggest that at least you make a donation to the Wikipedia Foundation, too.

Writing this book has been great fun, and it helped me to reflect on the history of the Web, at times reconstructing facts that were vaguely stored in the back of my mind. I took the liberty of the occasional personal and subjective consideration, based on my understanding of science and technology. Being used to writing objective and unbiased scientific papers, such freedom was new to me and at times inebriating. While the fumes of freedom might have made my style looser than usual, it has never been my intention to offend anyone or put down the hard work of respectable individuals. In fact, there are only good, heroic visionaries in this book, no traces of bad guys—at most, some people who might have misjudged the ugly effects of specific design decisions or who have simply behaved like *amateurs* by ignoring the history of the field to which they were contributing.

Sydney, June 2017 Marco Aiello

Acknowledgements

"The Web Was Done by Amateurs" could not exist without the help of many people. I take this occasion to thank the prominent ones and apologize if I have unintentionally forgotten anyone. First and foremost, I thank Alan Kay for being who he is and for his contributions to our field. Second, I thank Tim Berners-Lee for creating the Web, bringing it to success, and defending its openness ever since. I also thank him for being a physicist.

My internship at Apple's Advanced Technology Group in 1995 was eye opening in many ways. I thank Jim Spohrer for the opportunity and Martin Haeberli for his mentoring while there. Martin is also the one who first pointed me to the "As We May Think" paper of Vannevar Bush cited in Chap. 3, and the "End-to-End Arguments" paper cited in Chap. 9.

After my introduction to the Web, my journey continued with Web services thanks to a suggestion of Fausto Giunchiglia and the introduction to the theme by Mike Papazoglou. I owe them both for this.

Alexander Lazovik has been my first PhD student and the person who has given body, concreteness, and theoretical foundations to many of my intuitions. He has been my most valuable colleague since we first met in 2002. I also thank the many members of the Distributed Systems group at the University of Groningen with whom I collaborated over the years to obtain some of the results mentioned throughout the book.

Matt McEwen has done an incredible job at analyzing the story behind my book and helping me better present the material. I also received many precious suggestions from: Frank Blaauw, Talko Dijkhuis, Laura Fiorini, Heerko Groefsema, Massimo Mecella, Andrea and Gaetano Pagani, Jorge Perez, Azkario Pratama, Rolf Schwitter, and Brian Setz. Any remaining error can only be ascribed to myself.

I am indebted to Alfred Hofmann and Ralf Gerstner from Springer who enthusiastically embraced this book project, not being intimidated by the controversial title. Their professional and dedicated help gave great support and improved the value proposition of the present book.

Hannah Sandoval of PurpleInkPen has acted as my copy editor and has done a wonderful job over the various evolutions of the manuscript. She knows the art of making otherwise convoluted sentences flow.

I have written the present book while on sabbatical leave from the University of Groningen at the Macquarie University of Sydney. I thank both institutions for making this possible and supporting my visit Down Under.

I thank Andrew Binstock and UBM for granting permission to reproduce the entire 2012 interview of Alan Kay.

My parents, Mario and Gigina Aiello, have been two pioneers of computer science and artificial intelligence. This led them to first meet Alan Kay in 1974, and they have had regular contact since. I thank them for having provided genes, inspiration, foundations, and love. Additionally, my mother endured in reading many early drafts of the book. Serves her right for having given birth to yet another computer scientist.

I thank my family for supporting and bearing with me during the book writing process: my children, Maurizio and Aurelia, for being the biggest source of laughter and smiles I have and will ever encounter; my wife, Heike, for supporting all my ideas, no matter how crazy, putting up with my unadjusted sleeping patterns, and being a source of love, tenderness, and many great suggestions on how to make my text more crisp and accessible. The book would have not been possible nor readable without her presence in my life.

Contents

1 The Web Was Done by Amateurs 1
1.1 Text Versus Objects 3
1.2 The Birth of the Web 4

Part I The Origins 7

2 The Pacific-Ocean Internet 9
2.1 ARPANET 10
2.2 Comments, Please! 12
2.3 The Internet 16
2.4 Why Is It a Natural Resource? 17

3 Hypermedia Until the Web 21
3.1 How We May Think 25
3.2 The oN-Line System (NLS) 28
3.3 Xanadu 29
3.4 HyperCard 32
3.5 Enquire and the Birth of the Web 33

Part II The Web 39

4 The Original Web Pattern 41
4.1 Uniform Resource Locator 45
4.2 HyperText Transfer Protocol 46
4.3 HyperText Markup Language 48

5 The Browser Lament 51
5.1 The Browser Is an Application 52
5.2 Early Browsers 53
5.3 Netscape 56
5.4 Microsoft's Internet Explorer 57
5.5 Google's Chrome 60

Part III The Patches 63

6 Patching the Web 65
6.1 Patch I: Cookies 66
6.2 Patch II: Run Anywhere 69
6.3 Patch III: Scripting 73

7 Patch IV: Web Services 79
7.1 Securing the Intranet 79
7.2 Corba and IDLs 80
7.3 The Magic Triangle 82
7.4 Service-Oriented Computing 87
7.5 A Personal Tale: Compose It Now! 89
7.6 The Patch 92
7.7 Who's Serving? 93
7.8 XML Web Services Versus the REST 94
7.9 Vaporizing Services 95

8 The Unexploited Patch 101
8.1 The Semantic Gap 102
8.2 Subsumptions Subsumed by Subsumptions 106
8.3 The Patch 108

Part IV System Engineering 113

9 The Self-Organizing Web 115
9.1 The Size and Shape of the Web 116
9.2 Self-Organization and Complex Networks 119
9.3 Searching the Web 122
9.4 Self-Organization, Patching, and the Role of Amateurs 126

10 The Pervasive Future 129
10.1 Apps 130
10.2 Web (on the Internet) of Things 133
10.3 The Web and Artificial Intelligence 136

11 Should a New Web Be Designed? 141

Correction to: The Web Was Done by Amateurs E1

A Dr. Dobb's Interview with Alan Kay 145
A.1 A Note About Dr. Dobb's Journal 145
A.2 The Interview 146

References 153

Index 161