WEB TECHNOLOGY

Theory and Practice

Akshi Kumar

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Web Technology

Web Technology

Theory and Practice

Dr. Akshi Kumar

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper

International Standard Book Number-13: 978-1-1385-5043-8 (Hardback)

Library of Congress Cataloging-in-Publication Data

Names: Kumar, Akshi, author.
Title: Web technology : theory and practice / Akshi Kumar.
Description: Boca Raton : Taylor & Francis, a CRC title, part of the Taylor & Francis imprint, a member of the Taylor & Francis Group, the academic division of T&F Informa, plc, 2019. | Includes bibliographical references and index.
Identifiers: LCCN 2018014613| ISBN 9781138550438 (hardback : acid-free paper) |
ISBN 9781351029902 (ebook)
Subjects: LCSH: World Wide Web.
Classification: LCC TK5105.888 .K86 2019 | DDC 025.042--dc23
LC record available at https://lccn.loc.gov/2018014613

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

I dedicate this book to my parents, Shri Prem Parkash Kumar and Shrimati Krishna Kumar, for their infinite love, understanding, and support. Without them none of my success would have been possible and I would not be who I am today.

Thank you Mom and Dad!

Contents

Section II Web Theory

Section IV Web Research

List of Figures

List of Tables

List of Acronyms

AC	Accuracy
ACID	Atomicity, Consistency, Isolation, Durability.
AJAX	Asynchronous JavaScript and XML
API	Application Program Interface
APRANET	Advanced Research Projects Agency Network
AS Number	Assigned Number
ASP	Active Server Pages
AQE	Automatic Query Expansion
BOOTP	Bootstrap Protocol
CERN	European Organization for Nuclear Research
CGI	Common Gateway Interface
CIR	Contextual Information Retrieval
CMS	Content Management System
CRUD	Create, Retrieve, Update, Delete
CSS	Cascading Style Sheets
CSV	Comma Separated Values
DBMS	Database Management System
DDP	Distributed Data Protocol
DOM	Document Object Model
DRY	Don't Repeat Yourself
DSL	Digital Subscriber Line
DSS	Digital Satellite Systems
DTD	Document Type Definition
DTL	Django Template Language
ERb	Embedded Ruby
FN	False Negative
FP	False Positive
FTP	File Transfer Protocol
HITS	Hyperlink-Induced Topic Search
HTML	Hypertext Markup Language
HTTP	Hyper Text Transfer Protocol
HTTPS	Hyper Text Transfer Protocol Secure
IAB	Internet Architecture Board
IANA	Internet Assigned Numbers Authority
ICANN	Internet Corporation for Assigned Names and Numbers
ICMP	Internet Control Message Protocol
IE	Information Extraction
IESG	Internet Engineering Steering Group
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol

IMAP	Internet Message Access Protocol
IP	Internet Protocol
IR	Information Retrieval
IRC	Internet Relay Chat
IRSG	Internet Research Steering Group
IRTF	Internet Research Task Force
ISDN	Integrated Services Digital Network
ISOC	Internet Society
ISP	Internet Service Provider
IQE	Interactive Query Expansion
JDBC	Java Database Connectivity
JS	JavaScript
JSON	JavaScript Object Notation
JSP	Java Server Pages
KDD	Knowledge Discovery in Databases
KDT	Knowledge Discovery in Textual Database
LAMP	Linux, Apache, MySQL, Px (Perl, PHP, Python)
LAN	Local Area Network
MAN	Metropolitan Area Network
MDA	Mail Delivery Agent
MIME	Multipurpose Internet Mail Extensions
MTA	Message Transfer Agent
MVC	Model View Controller
MVT	Model View Template
MVW	Model View Whatever
NAP	Network Access Point
NER	Named Entity Recognizer
NLP	Natural Language Processing
NNTP	Network News Transfer Protocol
NSP	Network Service Provider
NVT	Network Virtual Terminal
ODBC	Open Database Connectivity
OGM	Neo4j Object Graph Mapping
ORM	Object-Relational Mapping
OSI	Open System Interconnection
OWL	Web Ontology Language
P	Precision
P2P	Peer-to-Peer
PC	Personal Computer
PDA	Personal Digital Assistant
PGP	Pretty Good Privacy
PHP	Personal Home Pages/Hypertext Pre-processor
POP	Post Office Protocol
PPP	Point-to-Point Protocol

PR	PageRank™
PTP	Point-to-Point
QoS	Quality of Service
R	Recall
RDBMS	Relational Database Management System
RDF/RDFS	Resource Description Framework/RDF Schema
RegEx	Regular Expression
REST	Representational State Transfer
RIR	Regional Internet Registries
ROR	Ruby on Rails
RSS	Rich Site Summary/RDF Site Summary/Really Simple Syndication
RTP	Real-time Transport Protocol
RTT	Round Trip Time
SMTP	Simple Mail Transfer Protocol
SNMP	Simple Network Management Protocol
SONET	Synchronous Optical Networking
SPARQL	Simple Protocol and RDF Query Language
SQL	Structured Query Language
SSL/TLS	Secure Socket Layer/Transport Layer Security
TCP	Transmission Control Protocol
TDM	Text Data Mining
tf-idf	Term Frequency-Inverse Document Frequency
TFTP	Trivial File Transfer Protocol
TLD	Top Level Domain
TN	True Negative
TP	True Positive
TTT	Total Transmission Time
UDP	User Datagram Protocol
UI	User Interface
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
VoIP	Voice over Internet Protocol
VSM	Vector Space Model
WAMP	Windows, Apache, MySQL, Px (Perl, PHP, Python)
WAN	Wide Area Network
WWW	World Wide Web
XAMPP	Cross-Platform (X), Apache (A), MariaDB (M), PHP (P), Perl (P)
XML	eXtensible Markup Language

Preface

With the transformation of the Web into a ubiquitous tool for e-activities, such as e-commerce, e-learning, e-governance, and e-science, its use has pervaded to the realms of day-to-day work, information retrieval, and business management. Moreover, much progress has been made in terms of technological innovations to match the tech-savvy, socially mindful Gen-Y. This growing popularity has fostered the need to include studies in this field as a part of the core curriculum. Many universities are now offering a full course on web technology for undergraduate, postgraduate, and doctoral students in the disciplines of computer science engineering, information technology, and computer applications. This book is intended to be a complete reference book for a course on Web Technology. It includes a clear description of relevant concepts of the Web. The significant feature of this book is that it explicates the developments and approaches in the field of the Web from three perspectives: Web in Theory, Web in Practice, and Web in Research.

This book is organized into 12 chapters. The topics are intelligibly explained with examples and real-life case studies, making it suitable for a course on Web Technology. At the end of each chapter, self-review questions have been given to reinforce the concepts covered within the text.

This book content has been divided into four sections: Internet Computing, Web Theory, Web Development, and Web Research. Section I (Chapters 1 and 2) focuses on the Internet Computing for highlighting the difference and relationship between the terms "Internet" and "Web." They are two discrete but correlated concepts. This section of the book proffers a pre-requisite primer on the What, How, Who, and Why of the Internet. It expounds the details pertaining to the architecture and applications of the Internet.

Section II (Chapters 3 through 6) of this book offers the theoretical concepts of the World Wide Web with a detailed discussion on the underlying application layer protocol, HTTP, which makes the Web work. Much progress has been made regarding the Web and related technologies in the past two decades. Thus, the evolution of the Web is witnessed to understand the state-of-the-art for Web generations from its advent. Search engines have assumed a central role in the World Wide Web's infrastructure as its scale and impact have escalated and so a complete illustration on their architecture and working is presented in this section.

Section III (Chapters 7 through 11) expounds on web development. The development of the Web involves multiple concepts, tools, and technologies that help developers build complex yet fascinating web sites that are dynamic and interactive. It explains browsers, client-side and server-side programming languages, frameworks, and databases. The discussion includes foundational technologies like HTML, CSS, JavaScript, and PHP and is extended to the current industry trends Bootstrap, AngularJS, Node.js, Django, and Ruby on Rails.

Section IV (Chapter 12) describes web research. It is a primer on recent research trends, such as web mining, contextual retrieval, and sentiment analysis. With the rapid development of the Web, it is imperative to provide users with tools for efficient and effective resource and knowledge discovery, thus making research studies in this field highly active and dynamic.

Lastly, Appendixes A and B have been added for reference and hands-on practice, which includes examples using HTML and a case study project using Django and Neo4j, respectively.

The end of this book includes several bibliographic notes for further reading.

It is my pleasure in presenting this book to you and hope that it will serve as a course textbook to undergraduate and postgraduate students, as well as a reference guide for faculty, research scholars, and professionals. Readers are welcome to send suggestions and constructive criticism, which may help to improve the contents of the book and motivate me to continue working hard.

As the adage says: "You are your only limit"

Akshi Kumar
Delhi Technological University, India.