

XML for Bioinformatics

ETHAN CERAMI

XML, or Extensible Markup Language, is rapidly becoming a critical tool in bioinformatics and biological data exchange. XML is currently used to represent a diverse set of biological data, from nucleotide and protein sequences to protein-protein interactions and signal transduction pathways. XML is also used in a wide array of bioinformatics applications, including stand-alone applications, federated database systems, distributed applications, and web services.


The goal of *XML for Bioinformatics* is to provide a solid introduction to the emerging use of XML in the field of bioinformatics. It assumes no prior knowledge of XML, and illustrates all core concepts with specific bioinformatics examples and case studies. Core XML concepts include: fundamentals of XML, Document Type Definitions (DTDs), XML Namespaces, XML Schema, XML parsing in Perl and Java, web services and SOAP. Examples and case studies are drawn from a wide range of bioinformatics applications, including the Bioinformatic Sequence Markup Language (BSML), NCBI E-Fetch, the Distributed Annotation System (DAS), and the National Cancer Institute Cancer Bioinformatics Infrastructure Objects (caBIO) project.

A companion web site, available at: <http://www.xmlbio.org/>, provides complete access to all examples in the book.

About the Author:

Ethan Cerami is a Senior Software Engineer at the Computational Biology Center at Memorial Sloan-Kettering Cancer Center. He is the author of several books, including *Web Services Essentials: Distributed Applications with XML-RPC, SOAP, UDDI & WSDL* (O'Reilly & Associates, 2002), and *Building XML Applications*, co-written with Simon St. Laurent (McGraw-Hill, 1999).

Cover design by Susanne Van Duyne (Trade Design Group).
Cover photo by Getty Images Inc.

 Springer

springeronline.com

ISBN 0-387-23028-9



9 780387 230283