GEOINFORMATICS

Cyberinfrastructure for the Solid Earth Sciences

Advanced information technology infrastructure is being employed increasingly in the Earth sciences to provide researchers with efficient access to massive databases and an ability to integrate diversely formatted information from a variety of sources.

A range of geoinformatics initiatives are enabling manipulation, modeling, and visualization of Earth Science data and are helping to develop integrated Earth models at various scales, and from the near surface to the deep interior.

This book provides a series of case studies that demonstrate the use of cyberinfrastructure across the Earth Sciences. Chapters are grouped thematically into sections that cover data collection and management; modeling and community computational codes; visualization and data representation; knowledge management and data integration; web services and scientific workflows.

Geoinformatics is a fascinating and accessible introduction to this emerging field for readers across the solid Earth sciences and is an invaluable reference for researchers interested in initiating new cyberinfrastructure projects of their own.

G. RANDY KELLER is a Professor of Geophysics and McCullogh Chair at the Mewbourne College of Earth and Energy, University of Oklahoma.

CHAITANYA BARU is a Distinguished Scientist at the San Diego Supercomputer Center, University of California-San Diego.

The two editors have collaborated since the emergence of geoinformatics as an organized scientific initiative in the USA in the late 1990s – helping to lead and organize the US Geoinformatics initiative and communicate its potential to colleagues around the world, both informally and through many appointments to advisory committees. Both Professor Keller and Dr. Baru are also Principal Investigators on the GEON (Geoscience Network) project that is a major effort funded by the National Science Foundation. While early geoinformatics programmes focused on database creation and on the development of highly functional software tools, these have since been merged with other efforts, such as high-performance computing and integrated earth-system modeling, to create a more extensive cyberinfrastructure for the geosciences. Dr. Baru's work at the San Diego Supercomputer Center has involved cyberinfrastructure activities across a range of scientific subject areas, while Professor Keller's research has focused on applications specific to the geosciences.

GEOINFORMATICS

Cyberinfrastructure for the Solid Earth Sciences

Edited by

G. RANDY KELLER
University of Oklahoma

CHAITANYA BARU
University of California-San Diego

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge, CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521897150

First published 2011

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this publication is available from the British Library

ISBN 978-0-521-89715-0 Hardback

Contents

The color plate section can be found between pages 114 and 115.

Contributors

Arun Agarwal, University Center for Earth & Space Sciences, University of Hyderabad, Central University P.O., Hyderabad – 500 046, India

Timothy K. Ahern, IRIS Data Management Center, 1408 NE 45 St., Suite 201, Seattle, WA 98105, USA

Michael Aivazis, Center for Advanced Computer Research, California Institute of Technology, Pasadena, CA 91125, USA

Raed Aldouri, Regional Geospatial Service Center, University of Texas at El Paso, El Paso, TX 79968, USA

M. Lee Allison, Arizona Geological Survey, 416 W. Congress St., Suite 100, Tucson, AZ 85701, USA

Ilkay Altintas, San Diego Supercomputer Center, University of California-San Diego, La Jolla, CA 92093-0505, USA

Luis Armendariz, Computational Infrastructure for Geodynamics, California Institute of Technology, Pasadena, CA 91125, USA

J Ramón Arrowsmith, School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404, USA

Hassan Babaie, Department of Geosciences, Georgia State University, P.O. Box 4105, Atlanta, GA 30302-4105, USA

Chaitanya Baru, San Diego Supercomputer Center, University of California-San Diego, La Jolla, CA 92093-0505, USA

Ralf Bill, University of Rostock, Geodesy and Geoinformatics 18051 Rostock, Germany

H. P. Bunge, Department of Earth and Environmental Sciences, Munich University, Theresienstr. 41, 80333 Munich, Germany

James A. Boyden, University of Sydney, Madsen Building (F09), Room 410, Sydney, NSW 2006, Australia

John S. Cannon, University of Sydney, School of Mathematics and Statistics (F07), Sydney, NSW 2006, Australia

Robert Casey, IRIS Data Management Center, 1408 NE 45 St., Suite 201, Seattle, WA 98105, USA

Cinzia Cervato, Iowa State University, Department of Geological and Atmospheric Sciences, 253 Science I, Ames, Iowa 50011-3212, USA

Amit Chourasia, San Diego Supercomputer Center, University of California-San Diego, La Jolla, CA 92093-0505, USA

James A. Clark, University of Sydney, Madsen Building (F09), Room 410, Sydney, NSW 2006, Australia

Ryan J. Clark, Arizona Geological Survey, 416 W. Congress St., Suite 100, Tucson, AZ 85701, USA

Sierd Cloetingh, Department of Tectonics, Faculty of Earth and Life Sciences, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands

Peter Cornillon, Graduate School of Oceanography, University of Rhode Island, Narragansett, RI 02882, USA

Daniel Crawl, San Diego Supercomputer Center, University of California-San Diego, La Jolla, CA 92093-0505, USA

Christopher J. Crosby, San Diego Supercomputer Center, University of California-San Diego, La Jolla, CA 92093-0505, USA

Nicholas Del Rio, Department of Computer Sciences, University of Texas at El Paso, El Paso, TX 79968, USA

Peter Fox, Tetherless World Constellation, Rensselaer Polytechnic Institute, 110 8th Street, 2nd floor, Troy, NY 12180, USA

Aida Gandara, Department of Computer Sciences, University of Texas at El Paso, El Paso, TX 79968, USA

Ann Gates, Cyber-ShARE Center, Department of Computer Sciences, University of Texas at El Paso, El Paso, TX 79968, USA

Wolfgang Grunberg, Arizona Geological Survey, 416 W. Congress St., Suite 100, Tucson, Arizona 85701, USA

Linda C. Gundersen, U.S. Geological Survey, 12201 Sunrise Valley Drive, MS 911, Reston, VA 20192, USA

Michael Gurnis, Seismological Laboratory, California Institute of Technology, Pasadena, CA 91125, USA

Richard Hughes, British Geological Survey, Kingsley Dunham Centre, Keyworth, Nottingham, NG12 5GG, UK

Hamish Ivey-Law, Institut de Mathématiques de Luminy, Université de la Méditerannée Aix-Marseille II, Marseille, France

Ian Jackson, British Geological Survey, Kingsley Dunham Centre, Keyworth, Nottingham, NG12 5GG, UK

G. Randy Keller, University of Oklahoma, School of Geology and Geophysics, 100 E. Boyd Norman, OK 73019, USA

Jens Klump, Helmholz Centre Potsdam German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

Sriram Krishnan, San Diego Supercomputer Center, University of California-San Diego, MC0505, San Diego, CA 92093-0505, USA

Walter Landry, Computational Infrastructure for Geodynamics, California Institute of Technology, Pasadena, CA 91125, USA

Matthias Lendholt, Helmholtz Centre, Potsdam German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

Qingsong Li, Lunar and Planetary Institute, Houston, TX 77058, USA

Mian Liu, Dept. of Geological Sciences, University of Missouri, Columbia, MO 65211, USA

Peter Löwe, Helmholz Centre Potsdam German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

Gang Luo, Dept. of Geological Sciences, University of Missouri, Columbia, MO 65211, USA

Ravi Madduri, Argonne National Lab, 9700 South Cass Avenue, Building 221, Argonne, IL 60439-4844, USA

David R. Maidment, Center for Research in Water Resources, University of Texas at Austin, Austin, TX 78712, USA

Deborah L. McGuinness, Tetherless World Constellation, Rensselaer Polytechnic Institute, 110 8th Street, 3rd floor, Troy, NY 12180, USA

Charles Meertens, UNAVCO Inc., 6350 Nautilus Drive, Boulder, CO 80301-5554, USA

R. Dietmar Müller, University of Sydney, Madsen Building (F09), Room 406, Sydney, NSW 2006, Australia

Viswanath Nandigam, San Diego Supercomputer Center, University of California-San Diego, La Jolla, CA 92093-0505, USA

Omar Ochoa, Department of Computer Sciences, University of Texas at El Paso, El Paso, TX 79968, USA

Paulo Pinheiro da Silva, Cyber-ShARE Center, Department of Computer Sciences, University of Texas at El Paso, El Paso, TX 79968, USA

Stephen M. Richard, Arizona Geological Survey, 416 W. Congress St., Suite 100, Tucson, AZ 85701, USA

Peter M. Sadler, Department of Earth Sciences, University of California, Riverside, 900 University Ave., Riverside, CA 92521, USA

Leonardo Salayandia, Cyber-ShARE Center, University of Texas at El Paso, El Paso, TX 79968, USA

Leif Strand, Computational Infrastructure for Geodynamics, California Institute of Technology, Pasadena, CA 91125, USA

K. V. Subbarao, University Center for Earth & Space Sciences, University of Hyderabad, Central University P.O., Hyderabad – 500 046, India

Eh Tan, Computational Infrastructure for Geodynamics, California Institute of Technology, Pasadena, CA 91125, USA

Trond H. Torsvik, PGP, University of Oslo, Physics Building, Sem Selands vei 24, Oslo NO-0316, Norway

Mark Turner, Tectonics Observatory, California Institute of Technology, Pasadena, CA 91125 USA

Joachim Wächter, Helmholtz Centre, Potsdam German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

Rajeev Wankar, Department of Computer and Information Sciences, University of Hyderabad, Central University P.O., Hyderabad – 500 046, India

Robin J. Watson, Norwegian Geological Survey, NGU, Postboks 6315 Sluppen, Trondheim NO-7491, Norway

Stuart Wier, UNAVC Inc., 6350 Nautilus Drive, Boulder, CO 80301-5554, USA

Youqing Yang, Dept. of Geological Sciences, University of Missouri, Columbia, MO 65211, USA

Ilya Zaslavsky, San Diego Supercomputer Center, University of California-San Diego, La Jolla, CA 92093-0505, USA

Huai Zhang, Laboratory of Computational Geodynamics, Graduate University of the Chinese Academy of Sciences, Beijing 100049, P. R. China

Preface

The idea for this book arose out of the development of Geoinformatics as a research emphasis in the Division of Earth Sciences at the U. S. National Science Foundation (NSF) that was fostered by Dr. Herman Zimmerman over a decade ago. This development was coincident with the creation of the Information Technology Research (ITR) for National Priorities program at NSF. Thus, the content of this book features several papers based on research that was inspired and/or funded as a result of these developments. Geoinformatics is certainly not a term or concept that is the invention of the U. S. earth sciences community, and this book also features chapters by authors from nine different countries. In addition, most chapters are the result of research in which geoscientists and computer scientists work together to solve complex scientific questions. This research involves topics such as data systems and models, data integration, advanced computer simulations, visualization, ontologies, workflows, webservices, and international cooperation.

Over most of the past decade, we have been supported by the National Science Foundation to pursue research in Geoinformatics primarily through the GEON (Geosciences Network) project, and we gratefully acknowledge this support. GEON was not created in a vacuum. A number of distributed and grid computing-based projects were in early stages at the time that GEON was originally conceived, including the Grid Physics Network (GriPhyN), funded by the US National Science Foundation and other agencies; the Biomedical Informatics Research Network (BIRN), funded by the National Center for Research Resources (NCRR) at the US National Institutes for Health; and the Southern California Earthquake Consortium's Common Modeling Environment (CME), which was also funded by an NSF ITR grant. From its inception, GEON recognized the need for "cross training" between earth and computer scientists. Dr. Margaret Leinen, then Assistant Director for Geosciences at NSF, proposed that one way to facilitate such cross training was by organizing summer institutes that would attract students, researchers, and faculty from both groups. This vision was realized, and the 7th

Cyberinfrastructure Summer Institute for Geoscientists (CSIG) was held in 2010. In addition, a series of meetings on Geoinformatics have been organized, and the most recent one was held in Potsdam, Germany in 2008; its proceedings are available at http://pubs.usgs.gov/sir/2008/5172.

The activities mentioned above have set the stage for programs such as the U.S. Geoscience Information Network (GIN) and OneGeology, which are represented in this book and which are now tackling the organizational issues (as opposed to purely technical ones) surrounding building metadata-based discovery and search across many organizations around the world.

Finally, we want to express our appreciation to the 58 authors who contributed to the 24 chapters in this book. Geoinformatics is a vibrant and dynamic field. It has taken over 2 years to finalize the contents of this book, and their patience is greatly appreciated. We also want to acknowledge the staff at Cambridge University Press who were helpful, knowledgeable, and effective at every step along the path that led to publication of this book.