Lecture Notes in Artificial Intelligence 2705

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Steve Renals   Gregory Grefenstette (Eds.)

Text- and Speech-Triggered Information Access

8th ELSNET Summer School
Chios Island, Greece, July 15-30 2000
Revised Lectures

Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Steve Renals
University of Sheffield, Department of Computer Science
Sheffield S1 4DP, UK
E-mail: s.renals@dcs.shef.ac.uk

Gregory Grefenstette
Clairvoyance Corp.
Pittsburg PA 15213-1854, USA
E-mail: g.grefenstette@clairvoyancecorp.com

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

CR Subject Classification (1998): I.2.7, I.2, F.4.2-3, H.5.5, H.3, H.2, H.4

ISSN 0302-9743
ISBN 3-540-40635-2 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de

Typesetting: Camera-ready by author, data conversion by Boller Mediendesign
Printed on acid-free paper SPIN: 10928707 06/3142 5 4 3 2 1 0

Preface

This book originated from the 8th ELSNET Summer School on Language and Speech Communication that was held in the summer of 2000 on the island of Chios in Greece. ELSNET is the European Network in Human Language Technologies, a network of some 140 academic institutions and private companies from all over Europe, all active in language and speech technology, covering the whole range from basic research to industrial development and integration. It was created in 1991 with the objective to bring together the language and speech technology communities on the one hand, and the academic and industrial communities on the other.

The ELSNET Summer Schools have now become a tradition. They are different from other summer schools in that they are always dedicated to one specific topic area, always on the borderline of language and speech technology. They bring together a mixed audience of academic and industrial researchers, with backgrounds in language processing, speech processing, software engineering, and many other fields.

The topic selected for the 2000 Summer School was "Text- and Speech-Triggered Information Access." The underlying problem is well-known to all of us: we are continuously producing and storing enormous amounts of data in many different forms (such as text, speech, images, mixed modalities) – how can we ensure that these vast data collections remain accessible in an efficient and natural way?

The problem has of course many facets other than just accessibility, such as physical storage, data integrity, data acquisition, intellectual property rights and data management, but this Summer School focused on how language and speech technology can be deployed to get access to information.

We have found our Summer Schools to be an effective and efficient instrument for the transfer of knowledge. At the same time it should be noted that they can only accommodate a limited number of people, in a specific location, at a specific moment in time. In order to make the same knowledge accessible to a wider audience than just the happy few who were able to attend the School, we have adopted the policy of trying to transform the material taught at the School into a book. This book is the fifth ELSNET Summer School book, and the first one to be published in Springer-Verlag's LNCS Tutorials series.

The book attempts to give newcomers in the field a clear overview of the main technologies and problems and to give existing practitioners a concise review of the technologies used in state-of-the-art deployment of language and speech technology in information access.

We would like to thank the editors, Steve Renals and Gregory Grefenstette, for their efforts to make this book happen, all the authors for their contributions to the 2000 Summer School and to this book, and the European Commission for their financial support.

Steven Krauwer
ELSNET Coordinator
http://www.elsnet.org

Table of Contents