

Editor-in-Chief

Kai Rannenber, Goethe University Frankfurt, Germany

Editorial Board

TC 1 – Foundations of Computer Science

Jacques Sakarovitch, Télécom ParisTech, France

TC 2 – Software: Theory and Practice

Michael Goedicke, University of Duisburg-Essen, Germany

TC 3 – Education

Arthur Tatnall, Victoria University, Melbourne, Australia

TC 5 – Information Technology Applications

Erich J. Neuhold, University of Vienna, Austria

TC 6 – Communication Systems

Aiko Pras, University of Twente, Enschede, The Netherlands

TC 7 – System Modeling and Optimization

Fredi Tröltzsch, TU Berlin, Germany

TC 8 – Information Systems

Jan Pries-Heje, Roskilde University, Denmark

TC 9 – ICT and Society

Diane Whitehouse, The Castlegate Consultancy, Malton, UK

TC 10 – Computer Systems Technology

Ricardo Reis, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

TC 11 – Security and Privacy Protection in Information Processing Systems

Steven Furnell, Plymouth University, UK

TC 12 – Artificial Intelligence

Ulrich Furbach, University of Koblenz-Landau, Germany

TC 13 – Human-Computer Interaction

Jan Gulliksen, KTH Royal Institute of Technology, Stockholm, Sweden

TC 14 – Entertainment Computing

Matthias Rauterberg, Eindhoven University of Technology, The Netherlands

IFIP – The International Federation for Information Processing

IFIP was founded in 1960 under the auspices of UNESCO, following the first World Computer Congress held in Paris the previous year. A federation for societies working in information processing, IFIP's aim is two-fold: to support information processing in the countries of its members and to encourage technology transfer to developing nations. As its mission statement clearly states:

IFIP is the global non-profit federation of societies of ICT professionals that aims at achieving a worldwide professional and socially responsible development and application of information and communication technologies.

IFIP is a non-profit-making organization, run almost solely by 2500 volunteers. It operates through a number of technical committees and working groups, which organize events and publications. IFIP's events range from large international open conferences to working conferences and local seminars.

The flagship event is the IFIP World Computer Congress, at which both invited and contributed papers are presented. Contributed papers are rigorously refereed and the rejection rate is high.

As with the Congress, participation in the open conferences is open to all and papers may be invited or submitted. Again, submitted papers are stringently refereed.

The working conferences are structured differently. They are usually run by a working group and attendance is generally smaller and occasionally by invitation only. Their purpose is to create an atmosphere conducive to innovation and development. Refereeing is also rigorous and papers are subjected to extensive group discussion.

Publications arising from IFIP events vary. The papers presented at the IFIP World Computer Congress and at open conferences are published as conference proceedings, while the results of the working conferences are often published as collections of selected and edited papers.

IFIP distinguishes three types of institutional membership: Country Representative Members, Members at Large, and Associate Members. The type of organization that can apply for membership is a wide variety and includes national or international societies of individual computer scientists/ICT professionals, associations or federations of such societies, government institutions/government related organizations, national or international research institutes or consortia, universities, academies of sciences, companies, national or international associations or federations of companies.

More information about this series at <http://www.springer.com/series/6102>

Bojan Bozic · Gavin Mendel-Gleason
Christophe Debruyne · Declan O'Sullivan (Eds.)

Computational History and Data-Driven Humanities

Second IFIP WG 12.7 International Workshop
CHDDH 2016

Dublin, Ireland, May 25, 2016

Revised Selected Papers

Editors

Bojan Bozic
Trinity College Dublin
Dublin
Ireland

Gavin Mendel-Gleason
Trinity College Dublin
Dublin
Ireland

Christophe Debruyne
Trinity College Dublin
Dublin
Ireland

Declan O'Sullivan
Trinity College Dublin
Dublin
Ireland

ISSN 1868-4238

ISSN 1868-422X (electronic)

IFIP Advances in Information and Communication Technology

ISBN 978-3-319-46223-3

ISBN 978-3-319-46224-0 (eBook)

DOI 10.1007/978-3-319-46224-0

Library of Congress Control Number: 2016954945

© IFIP International Federation for Information Processing 2016

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.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The Second International Workshop on Computational History and Data-Driven Humanities held on May 25, 2016, in Dublin, Ireland, was the second in a series of workshops focusing on the challenges and opportunities of data-driven humanities, bringing together world-leading scientists and scholars at the forefront of this emerging field, at the interface between computer science, social science, humanities, and mathematics.

The first edition, held on June 27, 2014, in Dublin, Ireland, consisted of a series of invited talks that were divided into two main sessions. The first session focused on “teaching computers history” where computer scientists and digital humanities experts introduced technologies and tools they have applied in order to extract knowledge from historical records in a form that can be processed by computers without losing its meaningfulness. The theme of the second session was “calculating the unknown.” A collection of leading international scientists, working at the forefront of mathematical and theoretical analysis of historical data, described what is possible with current tools.

For the second edition of this workshop, we decided to solicit papers for a single blind review to be presented at the event. We received 14 submissions, of which seven (50 %) were accepted for presentation and inclusion as a paper in the proceedings. The authors of some of the rejected papers were invited to submit an extended abstract of at most four pages and present their work in a lightning talk at the event.¹

We are proud to announce that, this year, we had keynotes and invited talks by:

- Prof. Harvey Whitehouse, University of Oxford, UK
- Dr. Kevin Feeney, Trinity College Dublin, Ireland
- Prof. Peter Turchin, University of Connecticut, USA

The keynote of Prof. Whitehouse and Dr. Feeney’s invited talk are both included in the proceedings. The event, which was free thanks to the support of the Irish Resource Council, attracted 60 participants of whom 35 travelled to Ireland, clearly demonstrating the international character of the workshop. The keynote was followed by the first peer-reviewed paper sessions before lunch. The program of the afternoon consisted of the invited talks, the second peer-reviewed paper session, and a lightning talk session. Ample time between sessions was foreseen to cater for networking. Many of the participants engaged in dialogue after each presentation and in between sessions,

The original version of the book frontmatter was revised: For detailed information please see the Erratum available at DOI: [10.1007/978-3-319-46224-0_12](https://doi.org/10.1007/978-3-319-46224-0_12)

¹ Unfortunately by mistake, the lightning talk “The levelled approach - Using and evaluating text mining tools AVResearcherXL and Texcavator for historical research on public perceptions of drugs” by Berrie van der Molen, Lars Buitinck, and Toine Pieters was not included in the volume. It can be accessed at [http://arxiv.org/abs/1701.00487](https://arxiv.org/abs/1701.00487). The conference chairs and team acknowledge that this error was not a result of any actions by the authors.

with insightful comments from scholars with different backgrounds such as digital humanities, library and archives, computer science, history, and, of course, cliodynamics.

We would like to congratulate Siobhán Grayson, Karen Wade, Gerardine Meaney, and Derek Greene for winning the best paper award for their contribution entitled “The Sense and Sensibility of Different Sliding Windows in Constructing Co-occurrence Networks from Literature,” which received the highest scores in the peer-reviewing process.

We thank all the authors who submitted valuable papers to the workshop. We would like to express our gratitude to the members of the Program Committee for their constructive and insightful comments. We are also indebted to the many individuals who contributed to this event, and we would like to thank Springer and Prof. Ulrich Furbach, current chair of IFIP TC12 “Artificial Intelligence,” for allowing this event to be organized under the umbrella of IFIP and to be published in Springer’s *IFIP Advances in Information and Communication Technology* series.

Finally, this event would not have been possible without the financial support of the Irish Research Council via the New Foundations Scheme, the Science Foundation Ireland-funded ADAPT Centre for Digital Content Technology (Grant 13/RC/2106). This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement number 644055 (ALIGNED, <http://www.aligned-project.eu>)². We furthermore express our gratitude to Trinity College Dublin for providing us with a venue and infrastructure.



July 2016

Declan O’Sullivan
Bojan Bozic
Gavin Mendel-Gleason
Christophe Debruyne

² This communication reflects only the author’s view and the commission is not responsible for any use that may be made of the information it contains.

Organization

Executive Committee

Organizing Committee

Bojan Bozic	Trinity College Dublin, Ireland
Gavin Mendel-Gleason	Trinity College Dublin, Ireland
Christophe Debruyne	Trinity College Dublin, Ireland

Steering Committee

Declan O'Sullivan	Trinity College Dublin, Ireland
Peter Turchin	University of Connecticut, USA
Kevin Feeney	Trinity College Dublin, Ireland
Rob Brennan	Trinity College Dublin, Ireland

Logistics Team

Odhran Gavin	Trinity College Dublin, Ireland
--------------	---------------------------------

Program Committee

Arkadiusz Marciniak	AMU Poznan, Poland
Daniel Hoyer	University of Hertfordshire, UK
Daniel Mullins	University of Hertfordshire, UK
Douglas R. White	University of California, USA
Enrico Spolaore	University of Exeter, UK
Harvey Whitehouse	University of Oxford, UK
Peter N. Peregrine	Lawrence University, USA
Peter Turchin	University of Connecticut, USA
Pieter François	University of Hertfordshire, UK
Thomas E. Currie	University of Exeter, UK
Jodi Schneider	University of Pittsburgh, USA
Ian Milligan	University of Waterloo, Canada
Justin Tonra	NUI Galway, Ireland
Georgeta Bordea	NUI Galway, Ireland
Brian Davis	NUI Galway, Ireland
Marieke van Erp	Vrije Universiteit Amsterdam, The Netherlands
Leif Isaksen	University of Southampton, UK
Aidan Kane	NUI Galway, Ireland
Antske Fokkens	Vrije Universiteit Amsterdam, The Netherlands
Sharon Webb	University of Sussex, UK
Serge ter Braake	University of Amsterdam, The Netherlands

Isa Maks

Paul Buitelaar

Dimitris Kontokostas

Monika Solanki

VU University Amsterdam, The Netherlands

NUI Galway, Ireland

Univeristy of Leipzig, Germany

University of Oxford, UK

Contents

Keynotes

Ritual and Social Evolution: Understanding Social Complexity Through Data	3
<i>Harvey Whitehouse</i>	
The Dacura Data Curation System.	15
<i>Kevin Feeney</i>	

Full Papers

Digital History: Towards New Methodologies.	23
<i>Serge ter Braake, Antske Fokkens, Niels Ockeloën, and Chantal van Son</i>	
Epidemics in Venice: On the Small or Large Nature of the Pre-modern World	33
<i>Giovanni Colavizza</i>	
Neoclassica - A Multilingual Domain Ontology: Representing Material Culture from the Era of Classicism in the Semantic Web	41
<i>Simon Donig, Maria Christoforaki, and Siegfried Handschuh</i>	
Improving Part-of-Speech Tagging of Historical Text by First Translating to Modern Text.	54
<i>Erik Tjong Kim Sang</i>	
The Sense and Sensibility of Different Sliding Windows in Constructing Co-occurrence Networks from Literature	65
<i>Siobhán Grayson, Karen Wade, Gerardine Meaney, and Derek Greene</i>	
Storyteller: Visualizing Perspectives in Digital Humanities Projects	78
<i>Janneke M. van der Zwaan, Maarten van Meersbergen, Antske Fokkens, Serge ter Braake, Inger Leemans, Erika Kuijpers, Piek Vossen, and Isa Maks</i>	
Will Historians Ever Have Big Data? Theoretical and Infrastructural Perspectives	91
<i>Jennifer Edmond</i>	

Lightning Talks

Object Classification in Images of Neoclassical Furniture Using Deep Learning.	109
<i>Bernhard Bermeitinger, André Freitas, Simon Donig, and Siegfried Handschuh</i>	
Publishing and Consuming Irish Administrative Boundaries as Linked Data . . .	113
<i>Christophe Debruyne, Atul Nautiyal, and Declan O’Sullivan</i>	
Erratum to: Computational History and Data-Driven Humanities	E1
<i>Bojan Bozic, Gavin Mendel-Gleason, Christophe Debruyne, and Declan O’Sullivan</i>	
Automated Monitoring of Data Quality in Linked Data Systems	121
<i>Kevin Feeney, Rajan Verma, Max Brunner, Andre Stern, Odhran Gavin, Declan O’Sullivan, and Rob Brennan</i>	
SWIMing: Supporting Use Case Data Requirements Modelling for Managing Data Across the Building Life Cycle	124
<i>Kris McGlinn and Matthias Weise</i>	
Author Index	127