

# Lecture Notes in Artificial Intelligence

12733

Subseries of Lecture Notes in Computer Science

Series Editors

Randy Goebel

*University of Alberta, Edmonton, Canada*

Yuzuru Tanaka

*Hokkaido University, Sapporo, Japan*

Wolfgang Wahlster

*DFKI and Saarland University, Saarbrücken, Germany*

Founding Editor

Jörg Siekmann

*DFKI and Saarland University, Saarbrücken, Germany*


More information about this subseries at <http://www.springer.com/series/1244>


Agnès Braud · Aleksey Buzmakov ·  
Tom Hanika · Florence Le Ber (Eds.)

# Formal Concept Analysis

16th International Conference, ICFCA 2021  
Strasbourg, France, June 29 – July 2, 2021  
Proceedings

*Editors*

Agnès Braud   
ICube  
University of Strasbourg  
Illkirch, France

Tom Hanika   
University of Kassel  
Kassel, Germany

Aleksey Buzmakov   
NRU Higher School of Economics  
Perm, Russia

Florence Le Ber   
ICube  
University of Strasbourg  
Illkirch, France

ISSN 0302-9743                      ISSN 1611-3349 (electronic)  
Lecture Notes in Artificial Intelligence  
ISBN 978-3-030-77866-8              ISBN 978-3-030-77867-5 (eBook)  
<https://doi.org/10.1007/978-3-030-77867-5>

LNCS Sublibrary: SL7 – Artificial Intelligence

© Springer Nature Switzerland AG 2021

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, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

This volume features the contributions accepted for the 16th International Conference on Formal Concept Analysis (ICFCA 2021), held during June 29 – July 2, 2021, at Université de Strasbourg, France. Formal concept analysis (FCA) is a mathematical field that emerged about 40 years ago and is rooted in lattice and order theory. Although being of a theoretical nature, FCA proved to be of interest to various applied fields such as knowledge discovery, machine learning, database theory, information visualization, and many others.

The goal of the International Conference on Formal Concept Analysis is to offer researchers from FCA and related backgrounds the possibility to present and discuss their research. Since its first edition in 2003 in Darmstadt, Germany, ICFCA has been held annually in several countries in Europe, Africa, America, and Australia. In 2015, ICFCA became biennial to alternate with the Conference on Concept Lattices and Their Applications (CLA).

The field of FCA originated in the 1980s in Darmstadt as a subfield of mathematical order theory, with prior developments in other research groups. Its original motivation was to consider complete lattices as lattices of concepts, drawing motivation from philosophy and mathematics alike. FCA has since then developed into a wide research area with applications far beyond its original motivation, for example, in logic, knowledge representation, unsupervised learning, data mining, human learning, and psychology.

There were 32 papers submitted to this year’s ICFCA by authors in fifteen countries. Each submission was reviewed by at least two different members of the Program Committee and at least one member of the Editorial Board. The review process was single blind. Fourteen high-quality papers were chosen for publication in this volume, amounting to an acceptance rate of 43%. In addition, five papers were deemed mature enough to be discussed at the conference and were therefore included as short papers in this volume.

The research part of this volume is divided in five different sections. First, in “Theory” we compiled works that discuss advances on theoretical aspects of FCA. Second, the section “Rules” consists of contributions devoted to implications and association rules. The third section “Methods and Applications” is composed of results that are concerned with new algorithms and their applications. “Exploration and Visualisation” introduces different approaches to data exploration and in the final section we collected the five accepted short works.

In addition to the regular contributions, this volume also contains the abstracts of the four invited talks by outstanding researchers we were delighted to have at ICFCA 2021. In detail, we were pleased to host the following talks:

- “What do the Sources Say? Exploring Heterogeneous Journalistic Data as a Graph” by Ioana Manolescu (France)

- “Ontologies for On-Demand Design of Data-Centric Systems” by Magdalena Ortiz (Austria)
- “Towards human-guided rule learning” by Matthijs van Leeuwen (Netherlands)
- “Sustainable AI – What does it take for continued success in deployed applications?” by Stefan Wrobel (Germany)

We are deeply thankful to all authors who submitted their contributions to ICFCA 2021 as a platform for discussing their work. Our strong gratitude goes to the members of the Editorial Board and Program Committee, as well as to all the additional reviewers whose timely and thorough reviews made the fruitful discussions of the high-quality papers during the conference possible. Furthermore, we would like to express our sincere thanks to the local organizers who were always quick to solve any questions or problems that arose, and their hard work made for a great event.

We are very grateful to Springer for supporting the International Conference on Formal Concept Analysis as well as to the Université de Strasbourg and the Centre National de la Recherche Scientifique (CNRS) for hosting the event. Finally, we would like to highlight the great help of easychair.org for organizing the review process for ICFCA 2021 and of sciencesconf.org for hosting the website and their technical support, especially collecting the camera-ready papers.

June 2021

Agnès Braud  
Aleksy Buzmakov  
Tom Hanika  
Florence Le Ber

# Organization

## Executive Committee

### Conference Chair

Florence Le Ber

Université de Strasbourg, France

Agnès Braud

Université de Strasbourg, France

### Local Organization

Xavier Dolques

Université de Strasbourg, France

Nicolas Lachiche

Université de Strasbourg, France

Aurélien Leborgne

Université de Strasbourg, France

Stella Marc-Zwecker

Université de Strasbourg, France

Peggy Rupp

CNRS, Strasbourg, France

## Program and Conference Proceedings

### Program Chairs

Aleksey Buzmakov

Higher School of Economics, Russia

Tom Hanika

Universität Kassel, Germany

### Editorial Board

Jaume Baixeries

Polytechnic University of Catalonia, Spain

Karell Bertet

Université de La Rochelle, France

Peggy Cellier

IRISA, Rennes, France

Sebastien Ferré

Université de Rennes 1, France

Bernhard Ganter

Technische Universität Dresden, Germany

Dmitry Ignatov

Higher School of Economics, Russia

Mehdi Kaytoue

INSA Lyon, France

Sergei Kuznetsov

Higher School of Economics, Russia

Leonard Kwuida

Bern University of Applied Sciences, Switzerland

Rokia Missaoui

Université du Québec en Outaouais, Canada

Amedeo Napoli

LORIA, Nancy, France

Sergei Obiedkov

Higher School of Economics, Russia

Manuel Ojeda-Aciego

University of Malaga, Spain

Uta Priss

Ostfalia University of Applied Sciences, Germany

Sebastian Rudolph

Technische Universität Dresden, Germany

Christian Sacarea

Babes-Bolyai University of Cluj-Napoca, Romania

Stefan E. Schmidt

Technische Universität Dresden, Germany

Bariş Sertkaya

Frankfurt University of Applied Sciences, Germany

Gerd Stumme

Universität Kassel, Germany

Petko Valtchev  
Karl Erich Wolff

Université du Québec à Montréal, Canada  
Darmstadt University of Applied Sciences, Germany

## Program Committee

Pablo Cordero	University of Málaga, Spain
Diana Cristea	Babes-Bolyai University Cluj-Napoca, Romania
Christophe Demko	Université de La Rochelle, France
Jean Diatta	Université de la Réunion, France
Stephan Doerfel	MICROMATA GmbH, Germany
Xavier Dolques	Université de Strasbourg, France
Marianne Huchard	Université Montpellier, France
Robert Jäschke	Humboldt-Universität zu Berlin, Germany
Jan Konecny	Palacky University Olomouc, Czech Republic
Michal Krupka	Palacky University Olomouc, Czech Republic
Marzena Kryszkiewicz	Warsaw University of Technology, Poland
Tatiana Makhalova	LORIA-Inria, France
Jesús Medina	Universidad de Cádiz, Spain
Engelbert Mephu Nguifo	Université Blaise Pascal Clermont-Ferrand, France
Lhouari Nourine	Université Blaise Pascal Clermont-Ferrand, France
Jan Outrata	Palacky University Olomouc, Czech Republic
Jean-Marc Petit	INSA Lyon, France
Pascal Poncelet	Université Montpellier, France
Sandor Radeleczki	University of Miskolc, Hungary
Henry Soldano	Laboratoire d'Informatique de Paris Nord, France
Andreja Tepavčević	University of Novi Sad, Serbia
Martin Trnecka	Palacky University Olomouc, Czech Republic

## Additional Reviewers

Simon Vilmin  
David Lobo  
Lauraine Tiogning  
Issam Falih  
Philippe Fournier-Viger  
Fernando Chacón  
Blaise Bleriot Koguep Njionou



## **Abstracts of Invited Talks**

# What do the Sources Say? Exploring Heterogeneous Journalistic Data as a Graph

Ioana Manolescu

Inria Saclay, Île-de-France, France

**Abstract.** Professional journalism is of utmost importance nowadays. It is a main feature distinguishing dictatorships from democracies, and a mirror sorely needed by society to look upon itself and understand its functioning. In turn, understanding is necessary for making informed decisions, such as political choices.

With the world turning increasingly digital, journalists need to analyze very large amounts of data, while having no control over the structure, organization, and format of the data. Since 2013, my team has been working to understand data journalism and computational fact-checking use cases, to identify and develop tools adapted for this challenging setting. I will describe our SourcesSay project (2020–2024), in which extremely heterogeneous data sources are integrated as graphs, on top of which journalistic applications can be supported through flexible graph queries. I will explain the data source integration module, the role played by Information Extraction and Entity Disambiguation, as well as novel techniques to explore and simplify these graphs.

# Ontologies for On-Demand Design of Data-Centric Systems

Magdalena Ortiz 

Faculty of Informatics, TU Wien  
ortiz@kr.tuwien.ac.at

Over the last decade, ontologies have found impactful applications in data management, where they help improve access to data that is incomplete, heterogeneous, or poorly structured [3, 4]. An ontology can act as a mediator between users and data sources to facilitate query formulation, and allow us to obtain more complete query answers by leveraging domain knowledge to infer implicit facts. Huge research efforts have been put into understanding the problems associated to query evaluation leveraging diverse ontology languages, many algorithms have been developed, and off-the-shelf engines for knowledge-enriched query answering exist, see e.g., [1] and its references.

We present our work advocating a novel use of ontologies [2], not only to access the data stored in systems, but also to facilitate the correct organization of data at design time. We propose a process called *focusing* to harness existing ontologies for the on-demand design of the schema of knowledge-enriched databases. Focusing solutions specify which terms of an ontology are relevant to a specific application, and explicate desired assumptions about their completeness and dynamicity. We present automated inferences services for obtaining and validating focusing solutions, and for answering queries the resulting knowledge-enriched databases. The definitions admit different ontology and query languages for specifying the scope of the system, and are accompanied by concrete decidability and complexity results for selected representative combinations.

**Acknowledgements.** Based on joint work with M. Šimkus, F. Murlak, Y. Ibáñez-García V. Gutiérrez-Basulto, and T. Gogacz. Supported by the Austrian Science Fund (FWF) projects P30360 and P30873.

## References

1. Bienvenu, M., Ortiz, M.: Ontology-mediated query answering with data-tractable description logics. In: ReasoningWeb. LNCS, vol. 9203, pp. 218–307. Springer (2015)
2. Gogacz, T., Gutiérrez-Basulto, V., Ibáñez-García, Y., Murlak, F., Ortiz, M., Simkus, M.: Ontology focusing: knowledge-enriched databases on demand. In: ECAI. Frontiers in AI and Applications, vol. 325, pp. 745–752. IOS Press (2020)
3. Schneider, T., Simkus, M.: Ontologies and data management: A Brief Survey Künstliche Intell. **34**(3), 329–353 (2020)
4. Xiao, G., et al.: Ontology-based data access: A survey. In: IJCAI, pp. 5511–5519 (2018). [ijcai.org](http://ijcai.org)

# Towards Human-Guided Rule Learning

Matthijs van Leeuwen

Leiden University, the Netherlands

**Abstract.** Interpretable machine learning approaches such as predictive rule learning have recently witnessed a strong increase in attention, both within and outside the scientific community. Within the field of data mining, the discovery of descriptive rules has long been studied under the name of subgroup discovery. Although predictive and descriptive rule learning have subtle yet important differences, they both suffer from two drawbacks that make them unsuitable for use in many real-world scenarios. First, hyperparameter optimisation is typically cumbersome and/or requires large amounts of data, and second, results obtained by purely data-driven approaches are often unsatisfactory to domain experts.

In this talk I will argue that domain experts often have relevant knowledge not present in the data, which suggests a need for human-guided rule learning that integrates knowledge-driven and data-driven modelling. A first step in this direction is to eliminate the need for extensive hyperparameter tuning. To this end we propose a model selection framework for rule learning that 1) allows for virtually parameter-free learning, naturally trading off model complexity with goodness of fit; and 2) unifies predictive and descriptive rule learning, more specifically (multi-class) classification, regression, and subgroup discovery. The framework we propose is based on the minimum description length (MDL) principle, we consider both (non-overlapping) rule lists and (overlapping) rule sets as models, and we introduce heuristic algorithms for finding good models.

In the last part of the talk, I will give a glimpse of the next steps towards human-guided rule learning, which concern exploiting expert knowledge to further improve rule learning. Specifically, I will describe initial results obtained within the iMODEL project, in which we develop theory and algorithms for interactive model selection, involving the human in the loop to obtain results that are more relevant to domain experts.

# **Sustainable AI – What Does It Take for Continued Success in Deployed Applications?**

Stefan Wrobel<sup>1,2</sup>

<sup>1</sup> University of Bonn, Germany

<sup>2</sup> Fraunhofer IAIS, Germany

**Abstract.** Advances in machine learning research have been so impressive that one would be tempted to believe that today most practical problems could easily be solved purely with data and machine learning. However, in the real world, the requirements demanded from a deployed application go far beyond achieving an acceptably low error for a trained model. Deployed applications must guarantee sustained success with respect to their functionality, their business viability, and their ethical acceptability. In this talk, we will analyze the challenges faced in practice with respect to these three dimensions of sustainability, pointing out risks and common misunderstandings and highlighting the role of hybrid modeling. We will then discuss our lessons learned from a number of real world projects for companies with respect to approaches for engineering and operating ML systems. The talk will conclude with a perspective on the demands placed on AI systems by customers and society, presenting our methodology for testing and ultimately certifying such systems.

# Contents

## Theory

Representing Partition Lattices Through FCA . . . . .	3
<i>Mike Behrisch, Alain Chavarri Villarelo, and Edith Vargas-García</i>	
Fixed-Point Semantics for Barebone Relational Concept Analysis . . . . .	20
<i>Jérôme Euzenat</i>	
Boolean Substructures in Formal Concept Analysis . . . . .	38
<i>Maren Koyda and Gerd Stumme</i>	

## Rules

Enumerating Maximal Consistent Closed Sets in Closure Systems. . . . .	57
<i>Lhouari Nourine and Simon Vilmin</i>	
A New Kind of Implication to Reason with Unknown Information . . . . .	74
<i>Francisco Pérez-Gámez, Pablo Cordero, Manuel Enciso, and Ángel Mora</i>	
Pruning Techniques in LinCbO for Computation of the Duquenne-Guigues Basis . . . . .	91
<i>Radek Janostik, Jan Konecny, and Petr Krajča</i>	
Approximate Computation of Exact Association Rules. . . . .	107
<i>Saurabh Bansal, Sriram Kailasam, and Sergei Obiedkov</i>	

## Methods and Applications

An Incremental Recomputation of From-Below Boolean Matrix Factorization . . . . .	125
<i>Martin Trnecka and Marketa Trneckova</i>	
Clustering and Identification of Core Implications . . . . .	138
<i>Domingo López-Rodríguez, Pablo Cordero, Manuel Enciso, and Ángel Mora</i>	
Extracting Relations in Texts with Concepts of Neighbours . . . . .	155
<i>Hugo Ayats, Peggy Cellier, and Sébastien Ferré</i>	

**Exploration and Visualisation**

Triadic Exploration and Exploration with Multiple Experts. . . . . 175  
*Maximilian Felde and Gerd Stumme*

Towards Interactive Transition from AOC Poset to Concept Lattice. . . . . 192  
*Tim Pattison and Aaron Ceglar*

Visualization of Statistical Information in Concept Lattice Diagrams . . . . . 208  
*Jana Klompke and Sebastian Rudolph*

Force-Directed Layout of Order Diagrams Using Dimensional Reduction. . . . 224  
*Dominik Dürschnabel and Gerd Stumme*

**Short Papers**

Sandwich: An Algorithm for Discovering Relevant Link Keys in an LKPS  
Concept Lattice. . . . . 243  
*Nacira Abbas, Alexandre Bazin, Jérôme David, and Amedeo Napoli*

Decision Concept Lattice vs. Decision Trees and Random Forests. . . . . 252  
*Egor Dudyrev and Sergei O. Kuznetsov*

Exploring Scale-Measures of Data Sets . . . . . 261  
*Tom Hanika and Johannes Hirth*

Filters, Ideals and Congruences on Double Boolean Algebras . . . . . 270  
*Tenkeu Jeufack Yannick Léa, Etienne Romuald Alomo Temgoua,  
and Léonard Kwuida*

Diagrammatic Representation of Conceptual Structures . . . . . 281  
*Uta Priss*

**Author Index** . . . . . 291