

# **2013 1st International Workshop on Data Analysis Patterns in Software Engineering**

**(DAPSE 2013)**

**San Francisco, California, USA  
21 May 2013**



**IEEE Catalog Number: CFP13ZAN-POD  
ISBN: 978-1-4673-6295-5**

# Contents

Foreword . . . . .	iii
Building Statistical Language Models of Code Peter Schulam, Roni Rosenfeld, and Premkumar Devanbu — <i>CMU, USA; UC Davis, USA</i> . . . . .	1
Commit Graphs Maximilian Steff and Barbara Russo — <i>Free University of Bolzano, Italy</i> . . . . .	4
Concept to Commit: A Pattern Designed to Trace Code Changes from User Requests to Change Implementation by Analyzing Mailing Lists and Code Repositories Scott McGrath, Kiran Bastola, and Harvey Siy — <i>University of Nebraska at Omaha, USA</i> . . . . .	6
Data Analysis Anti-patterns in Empirical Software Engineering Sandro Morasca — <i>University of Insubria, Italy</i> . . . . .	9
Effect Size Analysis Emanuel Giger and Harald C. Gall — <i>University of Zurich, Switzerland</i> . . . . .	11
Exploring Software Engineering Data with Formal Concept Analysis Xiaobing Sun, Ying Chen, Bin Li, and Bixin Li — <i>Yangzhou University, China; Southeast University, China</i> . . . . .	14
Extracting Artifact Lifecycle Models from Metadata History Olga Baysal, Oleksii Kononenko, Reid Holmes, and Michael W. Godfrey — <i>University of Waterloo, Canada</i> . . . . .	17
Measure What Counts: An Evaluation Pattern for Software Data Analysis Emmanuel Letier and Camilo Fitzgerald — <i>University College London, UK</i> . . . . .	20
Parametric Classification over Multiple Samples Barbara Russo — <i>Free University of Bolzano, Italy</i> . . . . .	23
Patterns for Cleaning Up Bug Data Rodrigo Souza, Christina Chavez, and Roberto Bittencourt — <i>UFBA, Brazil; UEFS, Brazil</i> . . . . .	26
Patterns for Extracting High Level Information from Bug Reports Rodrigo Souza, Christina Chavez, and Roberto Bittencourt — <i>UFBA, Brazil; UEFS, Brazil</i> . . . . .	29
Structural and Temporal Patterns-Based Features Venkatesh-Prasad Ranganath and Jithin Thomas — <i>Microsoft Research, India</i> . . . . .	32
The Chunking Pattern David M. Weiss and Audris Mockus — <i>Iowa State University, USA; Avaya Labs Research, USA</i> . . . . .	35