## Version Control and Infrastructure

1. **The Promises and Perils of Mining Git**  
   *(Christian Bird, Peter C. Rigby, Earl T. Barr, David J. Hamilton, Daniel M. German, Prem Devanbu)*

11. **Amassing and Indexing a Large Sample of Version Control Systems: Towards the Census of Public Source Code History**  
    *(Audris Mockus)*

21. **MapReduce as a General Framework to Support Research in Mining Software Repositories (MSR)**  
    *(Weiyi Shang, Zhen Ming Jiang, Bram Adams, Ahmed E. Hassan)*

31. **A Platform for Software Engineering Research**  
    *(Georgios Gousios, Diomidis Spinellis)*

## Defect Prediction

41. **Evaluating the Relation Between Coding Standard Violations and Faults Within and Across Software Versions**  
    *(Cathal Boogerd, Leon Moonen)*

51. **Tracking Concept Drift of Software Projects Using Defect Prediction Quality**  
    *(Jayalath Ekanayake, Jonas Tappolet, Harald C. Gall, Abraham Bernstein)*

61. **Does Calling Structure Information Improve the Accuracy of Fault Prediction?**  
    *(Yonghee Shin, Robert Bell, Thomas Ostrand, Elaine Weyuker)*
Text Analysis

71  Mining Source Code to Automatically Split Identifiers for Software Analysis
    (Eric Enslen, Emily Hill, Lori Pollock, K. Vijay-Shanker)

81  Code Siblings: Technical and Legal Implications of Copying Code Between Applications
    (Daniel M. German, Massimiliano Di Penta, Yann-Gaël Guéhéneuc, Giuliano Antoniol)

MSR Challenge

91  Author Entropy vs File Size in the GNOME Suite of Applications
    (Jason R. Casebolt, Jonathan L. Krein, Alexander C. MacLean, Charles D. Knutson, Daniel P. Delorey)

95  Evaluating Process Quality in GNOME Based on Change Request Data
    (Holger Schackmann, Horst Lichter)

99  Mining the Coherence of GNOME Bug Reports with Statistical Topic Models
    (Erik Linstead, Pierre Baldi)

103 Visualizing Gnome with the Small Project Observatory
    (Mircea Lungu, Jacopo Malnati, Michele Lanza)

107 On the Use of Internet Relay Chat (IRC) Meetings by Developers of the GNOME GTK+
    Project
    (Emad Shihab, Zhen Ming Jiang, Ahmed E. Hassan)
<table>
<thead>
<tr>
<th>Topic Mining</th>
</tr>
</thead>
</table>
| 111 | Mining Search Topics from a Code Search Engine Usage Log  
(Sushil Bajracharya, Cristina Lopes) |
| Developers |
| 121 | From Work to Word: How Do Software Developers Describe Their Work?  
(Walid Maalej, Hans-Jörg Happel) |
| 131 | Assigning Bug Reports Using a Vocabulary-Based Expertise Model of Developers  
(Dominique Matter, Adrian Kuhn, Oscar Nierstrasz) |
| 141 | Mining the History of Synchronous Changes to Refine Code Ownership  
(Lile Hattori, Michele Lanza) |
Poster Papers

151 Using Association Rules to Study the Co-Evolution of Production & Test Code  
(Zeeger Lubsen, Andy Zaidman, Martin Pinzger)

155 On What Basis to Recommend: Changesets or Interactions?  
(Sarah Rastkar, Gail C. Murphy)

159 Mining the Jazz Repository: Challenges and Opportunities  
(Kim Herzig, Andreas Zeller)

163 Using Latent Dirichlet Allocation for Automatic Categorization of Software  
(Kai Tian, Meghan Revelle, Denys Poshyvanyk)

167 Evolution of the Core Team of Developers in Libre Software Projects  
(Gregorio Robles, Jesus M. Gonzalez-Barahona, Israel Herrera)

171 On Mining Data Across Software Repositories  
(Prasanth Anbalagan, Mladen Vouk)

175 Automatic Labeling of Software Components and Their Evolution Using Log-Likelihood Ratio of Word Frequencies in Source Code  
(Adrian Kuhn)

179 Learning from Defect Removers  
(Nathaniel Ayewah, William Pugh)

183 SourcererDB: An Aggregated Repository of Statically Analyzed and Cross-Linked Open Source Java Projects  
(Joel Ossher, Sushil Bajracharya, Erik Linstead, Pierre Baldi, Cristina Lopes)

187 On the Transfer of Evolutionary Couplings to Industry  
(Pierre van de Laar)