

**THANKS FOR COMING!**

Find us in @msrconf.

Use the hashtag #MSR17 to discuss during the conference.



**MAY 20, Saturday**

9:00-10.30 **Welcome, 2007 Most Influential Paper Award and Plenary Discussion 1**  
Auditorio 1. Ground floor

Welcome

2007 Most Influential Paper Award

Plenary Discussion 1

10.30-11.00 **Coffee Break**

11:00-12:30 **Research 1: Mobile**  
Auditorio 1. Ground floor  
Session Chair: Gregorio Robles

An Empirical Study on Android-related Vulnerabilities [preprint]  
Mario Linares-Vásquez, Gabriele Bavota and Camilo Escobar-Velasquez  
Universidad de los Andes, Università della Svizzera Italiana (USI)

Understanding the Origins of Mobile App Vulnerabilities: A Large-scale Measurement Study of Free and Paid Apps [preprint]  
Takuya Watanabe, Mitsuaki Akiyama, Fumihiro Kanai, Eitaro Shioji, Yuta Takata, Bo Sun, Yuta Ishii, Toshiki Shibahara, Takeshi Yagi and Tatsuya Mori  
Waseda University/NTT Secure Platform Laboratories

Developer mistakes in writing Android manifests: An empirical study of configuration errors  
Ajay Kumar Jha, Sunghee Lee and Woo Jin Lee  
Kyungpook National University

How do Apps Evolve in Their Permission Requests? A Preliminary Study (short)  
Paolo Calciati and Alessandra Gorla  
IMDEA Software Institute

A Study on the Energy Consumption of Android App Development Approaches [preprint]  
Wellington Oliveira, Fernando Castor and Renato Oliveira  
Federal University of Pernambuco

Candoia: A Platform for Building and Sharing Mining Software Repositories Tools as Apps  
Nitin Mukesh Tiwari, Ganesh Upadhyaya, Hoan Anh Nguyen and Hridesh Rajan

11:00-12:30 **Research 2: Dependencies**  
Cine. Ground floor  
Session Chair: Shane McIntosh

Analyzing Program Dependencies in Java EE Applications  
Anas Shatnawi, Hafedh Mili, Ghizlane El Boussaidi, Anis Boubaker, Yann-Gaël Guéhéneuc, Naouel Moha, Jean Privat and Manel Abdellatif  
University of Québec at Montréal, École de technologie supérieure, University of Montréal

Mining Social Web Service repositories for social relationships to aid service discovery (short)  
Alejandro Corbellini, Daniela Godoy, Cristian Mateos, Alejandro Zunino and Ignacio Lizarralde  
ISISTAN-CONICET, UNICEN

Who you gonna call? Analyzing Web Requests in Android Applications  
Marianna Rapoport, Philippe Suter, Erik Wittemer, Ondřej Lhoták and Julian Dolby  
IBM T.J. Watson Research Center, University of Waterloo

Extracting Code Segments and Their Descriptions from Research Articles [preprint]  
Preetha Chatterjee, Benjamin Gause, Hunter Hedinger and Lori Pollock  
University of Delaware

Structure and Evolution of Package Dependency Networks [preprint]  
Rivo Kikas, Georgios Gousios, Marlon Dumas and Dietmar Pfahl  
University of Tartu, Delft University of Technology

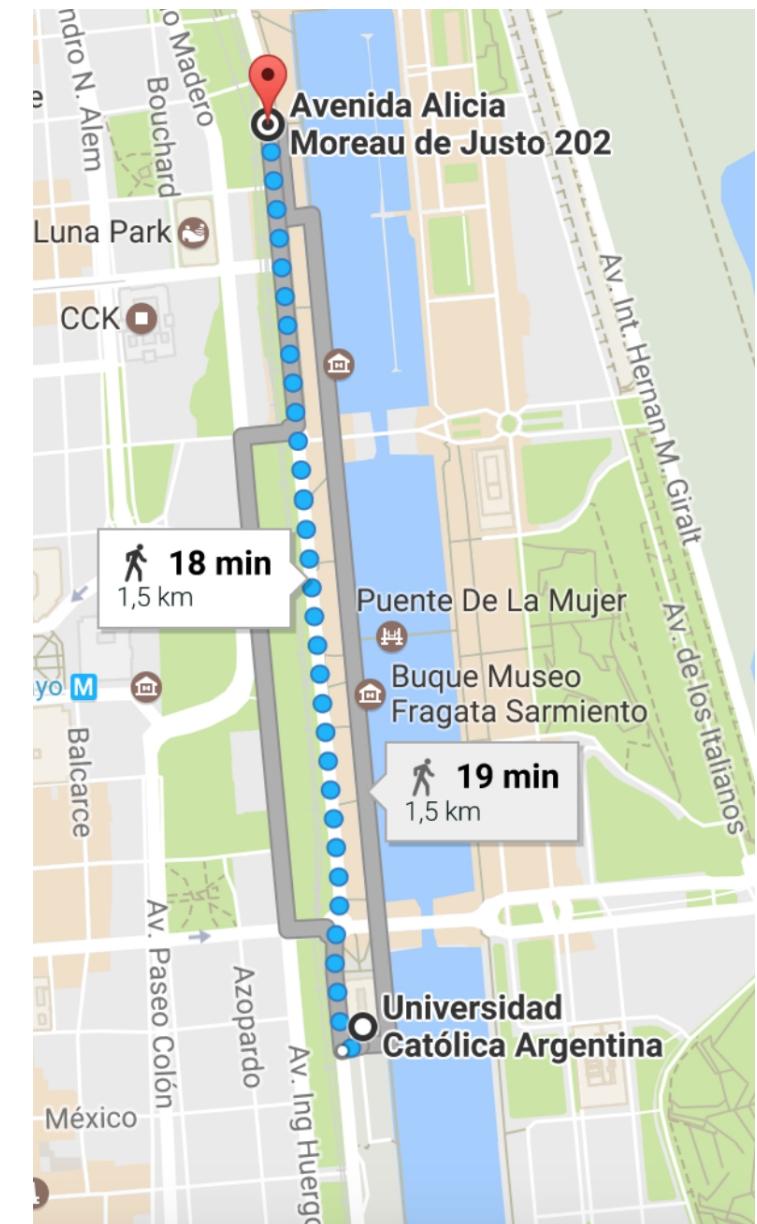
Spencer: Interactive Heap Analysis for the Masses [preprint]  
Stephan Brandauer and Tobias Wrigstad  
Uppsala University

12.30-14.00 **Lunch**

## SOCIAL EVENT

### OnTap

Marcelo T de Alvear 834, Buenos Aires



14.00-15.30  
Auditorio 1. Ground floor

### Research 3: Modelling and Prediction

Session Chair: Yasutaka Kamei

Predicting Likelihood of Requirement Implementation within the Planned Iteration: An Empirical Study at IBM [preprint]

Ali Dehghan, Adam Neal, Kelly Blincoe, Johan Linåker and Daniela Damian

Persistent Systems, University of Victoria, Lund University, University of Auckland

The Impact Of Using Regression Models to Build Defect Classifiers [preprint]

Gopi Krishnan Rajbahadur, Shaowei Wang, Yasutaka Kamei and Ahmed E. Hassan

Queen's University, Kyushu University

A Large-Scale Study of the Impact of Feature Selection Techniques on Defect Classification Models [preprint]

Baljinder Ghora, Shane McIntosh and Ahmed E. Hassan

Queen's University

SpreadCluster: Recovering Versioned Spreadsheets through Similarity-Based Clustering [preprint]

Liang Xu, Wensheng Dou, Chushu Gao, Jie Wang, Jun Wei, Hua Zhong and Tao Huang

Institute of Software, Chinese Academy of Sciences (ISCAS)

Who Will Leave the Company?: A Large-Scale Industry Study of Developer Turnover by Mining Monthly Work Report [preprint]

Lingfeng Bao, Zhenchang Xing, Xin Xia, David Lo and Shaping Li

Singapore Management University, Zhejiang University, University of British Columbia, Australian National University

Concept-based Classification of Software Defect Reports (short)

Sangameshwar Patil

Tata Research Development and Design Centre (TRDDC), TCS Pune

14.00-15.30  
Cine. Ground floor

### Research 4: NLP and Code Review

Session Chair: Christian Bird

Choosing an NLP Library for Analyzing Software Documentation: A Systematic Literature Review and a Series of Experiments [preprint]

Fouad Nasser A. Al Omran and Christoph Treude

University of Adelaide

Bootstrapping a Lexicon for Emotional Arousal in Software Engineering (short) [preprint]

Mika Mäntylä, Nicole Novelli, Filippo Lanubile, Maëlick Claes and Miikka Kuutila

University of Bari, University of Oulu

Leveraging Automated Sentiment Analysis in Software Engineering

Md Rakibul Islam and Minhz F. Zibrani

University of New Orleans

Predicting Usefulness of Code Review Comments using Textual Features and Developer Experience

Mohammad Masudur Rahman, Chanchal K. Roy and Raula Gaikovina Kula

University of Saskatchewan, Osaka University

Classifying code comments in Java open-source software systems [preprint]

Luca Pascarella and Alberto Bacchelli

Delft University of Technology

Using Q&A Websites as a Method for Assessing Systematic Reviews (short) [preprint]

Bruno Cartaxo, Gustavo Pinto, Danilo Monteiro Ribeiro, Fernando Kenji Kamei, Ronnie E. S. Santos, Sérgio Soares and Fabio Q. B. Da Silva

UFPE, Universidade Federal de Pernambuco, IFAL, IFPA

Abnormal Working Hours: Effect of Rapid Releases and Implications to Work Content (short) [preprint]

Maëlick Claes, Miikka Kuutila, Mika Mäntylä and Bram Adams

Polytechnique Montreal, University of Oulu

15.30-16.00

**Coffee Break**16.00-17.30  
Auditorio 1. Ground floor**Mining Challenge (14 papers)**Session Chair: Moritz Beller, Georgios Gousios and Andy Zaidman  
  
On the Differences between Unit and Integration Testing in the TravisTorrent Dataset [preprint]Manuel Gerardo Orellana Cordero, Gulsher Laghari, Alessandro Murgia and Serge Demeyer  
University of Antwerp

Cost-effective Build Outcome Prediction Using Cascaded Classifiers

Ansong Ni and Ming Li  
Nanjing University

Sentiment Analysis of Travis CI Builds [preprint]

Rodrigo Souza and Bruno Silva  
Salvador University - UNIFACS, Federal University of Bahia

A Time Series Analysis of TravisTorrent: To Everything There is a Season

Abigail Atchison, Christina Berardi, Natalie Best, Elizabeth Stevens and Erik Linstead  
Chapman University

Insights into Continuous Integration Build Failures

Md Rakibul Islam and Minhz Zibran  
University of New Orleans

An Empirical Study of the Personnel Overhead of Continuous Integration

Marco Manglaviti, Eduardo Coronado-Montoya, Keheliya Gallaba and Shane McIntosh  
McGill University

How Does Contributors' Involvement Influences the Build Status of an Open-Source Software Project? [preprint]

Marcel Rebouças, Renato Oliveira Dos Santos, Gustavo Pinto and Fernando Castor  
Federal University of Pernambuco, Federal Institute of Pará

On the Interplay between Non-Functional Requirements and Builds on Continuous Integration [preprint]

Klérisson Paixão, Crícia Z. Felício, Fernanda Delfim and Marcelo Maia  
Instituto Federal do Triângulo Mineiro, Universidade Federal de Uberlândia, UFU

Analyzing the impact of social attributes on commit integration success

Mauricio Soto, Zack Coker and Claire Le Goues  
Carnegie Mellon University

Built to Last or Built Too Fast? Evaluating Prediction Models for Build Times

Ekaba Bisong, Eric Tran and Olga Baydal  
Carleton University

The Impact of the Adoption of Continuous Integration on Developer Attraction and Retention

Yusaira Khan, Yash Gupta, Keheliya Gallaba and Shane McIntosh  
McGill University

An Empirical Study of Activity, Popularity, Size, Testing, and Stability in Continuous Integration

Akash Gautam, Saket Vishwasrao and Francisco Servant  
Virginia Tech

Impact of Continuous Integration on Code Reviews

Mohammad Masudur Rahman and Chanchal K. Roy  
University of Saskatchewan

Prevalence of Botched Code Integrations [preprint]

Ward Muylaert and Coen De Roover  
Vrije Universiteit Brussel, Software Languages Lab, Vrije Universiteit Brussel

17:30-18.15, Auditorio 1. Ground floor

**Unofficial BoF 1: Tools for MSR**

17:30-18.15, Cine. Ground floor

**Unofficial BoF 2**

19:30

**Dinner**9.00-10.30  
Auditorio 1. Ground floor**Keynote and Plenary Discussion 2**Keynote: Half Century of Unix: History, Preservation, and Lessons Learned  
Diomidis SpinellisPlenary Discussion 2: Double-blind reviewing: lessons learned and paths ahead  
Abram Hindle and Lin Tan  
Moderator: Tom Zimmerman

10:30-11.00

**Coffee Break**11:00-12.30  
Auditorio 1. Ground floor**Research 5: Clones and Edits**

Session Chair: Chanchal Roy

Mining Change Histories for Unknown Systematic Edits [preprint]

Tim Molerez, Reinout Stevens and Coen De Roover  
Vrije Universiteit Brussel

Source File Set Search for Clone-and-Own Reuse Analysis [preprint]

Takashi Ishio, Yusuke Sakaguchi, Kaoru Ito and Katsuro Inoue  
Osaka University

RefDiff: Detecting Refactorings in Version Histories [preprint]

Danilo Silva and Marco Tulio Valente  
Federal University of Minas Gerais, UFMG

Stack Overflow in Github: Any Snippets There? [preprint]

Di Yang, Pedro Martins, Vaibhav Saini and Cristina Lopes  
University of California, Irvine

Some From Here, Some From There: Cross-Project Code Reuse in GitHub

Mohammad Gherehyazie, Baishakhi Ray and Vladimir Filkov  
University of Virginia, University of California Davis

Exception Evolution in Long-lived Java Systems [preprint]

11:00-12.30  
Cine. Ground floor**Research 6: Continuous Integration and Build**

Session Chair: Diomidis Spinellis

Do Not Trust Build Results at Face Value — An Empirical Study of 30 Million CPAN Builds [preprint]  
Mahdis Zolfagharinia, Bram Adams and Yann-Gaël Guéhéneuc  
Polytechnique Montréal

An Empirical Analysis of the Docker Container Ecosystem on GitHub [preprint]

Jürgen Cito, Gerald Schermann, Erik Wittern, Philipp Leitner, Sali Zumberi and Harald Gall  
IBM T.J. Watson Research Center, University of Zurich

How Open Source Projects use Static Code Analysis Tools in Continuous Integration Pipelines [preprint]

Fiorella Zampetti, Simone Scalabrinio, Rocco Oliveto, Massimiliano Di Penta and Gerardo Canfora  
University of Sannio, STAKE Lab - University of Molise

An Empirical Analysis of Build Failures in the Continuous Integration Workflows of Java-Based Open-Source Software [preprint]

Thomas Rausch, Waldemar Hummer, Philipp Leitner and Stefan Schulte  
University Of Zürich, Vienna University of Technology, TU Wien, Atlassian

Oops, My Tests Broke the Build: An Explorative Analysis of Travis CI with GitHub [preprint]

Moritz Beller, Georgios Gousios and Andy Zaidman  
Delft University of Technology

Extracting Build Changes with BUILDDIFF [preprint]

Christian Macho, Shane McIntosh and Martin Pinzger  
University of Klagenfurt

12:30-14.00

**Lunch**14:00-15.30  
Auditorio 1. Ground floor**Data Showcase (7 papers)**

Session Chair: Megan Squire

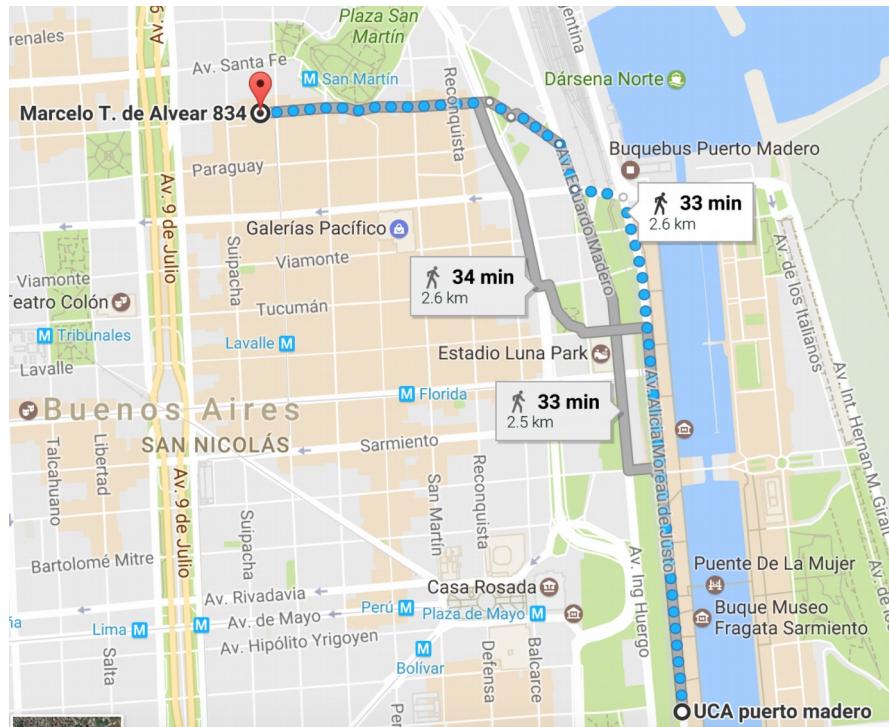
Software Evolution and Defects from Controlled, Multiple Industrial Case Studies

Aiko Yamashita, S. Amirhossein Abtahizadeh, Fousse Khomh, Yann-Gaël Guéhéneuc  
Centrum Wiskunde & Informatica, Polytechnique Montréal

## DINNER

**MSR dinner, for MSR attendees (Saturday evening)**

**in El Mirasol at Puerto Madero**



A Dataset of Scratch Programs: Scraped, Shaped and Scored

Efthimia Aivaloglou, Felienne Hermans, Jesús Moreno-León, Gregorio Robles

Delft University of Technology, Programamos.es & Universidad Rey Juan Carlos, Universidad Rey Juan Carlos

**Continuous Defect Prediction: The Idea and a Related Dataset [preprint]**

Lech Madeyski, Marcin Kawalerowicz

Wroclaw University of Science and Technology, Opole University of Technology

An extensive dataset of UML models in GitHub

Gregorio Robles, Truong Ho Quang, Regina Hebig, Michel Chaudron, Miguel Ángel Fernández

Universidad Rey Juan Carlos, Chalmers & Gothenborg University

**A Dataset for Dynamic Discovery of Semantic Changes in Version Controlled Software Histories [preprint]**

Chenguang Zhu, Yi Li, Julia Rubin, Marsha Chechik

University of Toronto, University of British Columbia

**Rediscovery Datasets: Connecting Duplicate Reports [preprint]**

Mefta Sadat, Ayse Bener, Andriy Miransky

Ryerson University

**A Data Set of OCL Expressions on GitHub [preprint]**

Jeroen F.H. Noten, Josh G.M. Mengerink, Alexander Serebrenik

Eindhoven University of Technology

14:00-15:30

### Research 7: Testing and Bugs

Session Chair: John Wittern

An Exploratory Study on Assessing the Impact of Environment Variations on the Results of Load Tests [preprint]

Ruoyu Gao and Zhen Ming Jack Jiang

York University

A Large-Scale Study on the Usage of Testing Patterns That Address Maintainability Attributes: Patterns for Ease of Modification, Diagnoses, and Comprehension

Danielle Gonzalez, Joanna Cecilia Da Silva Santos, Andrew Popovich, Mehdi Mirakhori and Melyappan Nagappan  
Rochester Institute of Technology, University of Waterloo

To Mock or Not To Mock? An Empirical Study on Mocking Practices [preprint]

Davide Spadini, Mauricio Aniche, Magiel Bruntink and Alberto Bacchelli  
Delft University of Technology, Software Improvement Group

Bug Characteristics in Blockchain Systems: A Large-Scale Empirical Study [preprint]

Zhiyuan Wan, David Lo, Xin Xia and Liang Cai  
Singapore Management University, Zhejiang University, University of British Columbia

Euphony: Harmonious Unification of Cacophonous Anti-Virus Vendor Labels for Android Malware

Médéric Hurier, Tegawendé Bissyandé, Yves Le Traon, Jacques Klein, Guillermo Suarez-Tangil, Santanu Kumar Dash and Lorenzo Cavallaro  
Royal Holloway, University of London, University of Luxembourg

Rationale in Development Chat Messages: An Exploratory Study [preprint]

Rana Alkadhi, Teodora Lata, Emilia Guzman and Bernd Bruegge  
Technische Universität München, University of Zurich

15:30-16:00

### Coffee Break

16:00-17:30

Auditorio 1. Ground floor

### Awards Session, Foundational Contribution Talk, Plenary Discussion 3 and Closing Session

Awards Session

Foundational Contribution Talk: Software Science: What? Why? What next?

Tim Menzies

Plenary Discussion 3

Closing Session