

# Alberto Bacchelli

University of Zurich



## about

I am SNSF professor of **Empirical Software Engineering** at University of Zurich, Switzerland, where I lead the Zurich Empirical Software Engineering Team (ZEST).

My broader research vision is to move software engineering away from decisions based on intuition, or development activities painstakingly conducted manually, into solutions created using data-driven mathematical models, which make use of the large amount of information available during the software engineering process.

The techniques and methods I use and develop are at the intersection of software engineering, machine learning, and social science.

## contacts

Dept. of Informatics  
University of Zurich  
Binzmühlestrasse 14  
8050 Zurich  
Switzerland

+41.44.6357539  
bacchelli@ifi.uzh.ch  
<http://sback.it>

## orcid

0000-0003-0193-6823

## publication indexes

<http://Google Scholar>  
<http://DBLP>

## languages

Italian: native speaker  
English: fluent  
French: conversational

## professional memberships

ACM

## education

- 04.2009–06.2013 **Ph.D.** in Informatics Università della Svizzera Italiana, Switzerland  
Thesis: *Mining Unstructured Software Data*  
Supervisor: *Prof. Dr. Michele Lanza*
- 04.2006–03.2008 **M.Sc.** in Computer Science University of Bologna, Italy  
Final grade: *110/110 cum laude*
- 09.2006–06.2007 **M.Sc. year** within EU Erasmus program Université Libre de Bruxelles, Belgium  
Main subjects: Software Engineering and Probabilistic Models
- 09.2002–03.2006 **B.Sc.** in Computer Science University of Bologna, Italy  
Final grade: *110/110 cum laude*

## professional appointments

- 08.2017–now **SNSF Professor** University of Zurich, Switzerland  
Research and teaching about empirical software engineering.
- 09.2013–08.2017 **Assistant professor, tenured after tenure-track** TU Delft, The Netherlands  
Research in software engineering with focus on peer code review and mining unstructured software data. Teaching 3 courses (B.Sc. and M.Sc.) and supervising M.Sc. as well as Ph.D. students.
- 04–08.2013 **Research intern** Microsoft Research, Redmond, USA  
Work on two projects: (1) visualization of code reviews to support understanding and (2) investigating and implementing a system for reviewer recommendation.
- 05–08.2012 **Research intern** Microsoft Research, Redmond, USA  
Work on two projects: (1) data mining approach to link code reviews and code commits and (2) investigating tool-based code review at Microsoft.
- 04.2008–04.2009 **Software Engineer** CINECA, Italy  
Design and implementation of an automatic test suite for producing, collecting, and analyzing data to test a large-scale web application to manage universities.

# research

## awards and honors

2018	<b>Best Paper Honorable Mention</b> “Information Needs in Contemporary Code Review”	CSCW 2018
2017	<b>ACM SIGSOFT Distinguished Paper Award</b> “Classifying code comments in Java open-source software systems”	MSR 2017
2016	<b>ACM SIGSOFT Distinguished Paper Award</b> “Work Practices and Challenges in Pull-Based Development: The Contributor’s Perspective”	ICSE 2016
2015	<b>Best Paper Award</b> “Supporting Developers’ Coordination in The IDE”	CSCW 2015
2014	<b>Google Faculty Research Award</b> The program is focused on funding world-class technical research in Computer Science, Engineering, and related fields.	Google 2014
2009	<b>Best Paper Award</b> “Benchmarking Lightweight Techniques to Link E-Mails and Source Code”	WCRE 2009
2008	<b>Best Paper Award</b> “On the Effectiveness of Manual and Automatic Unit Test Generation”	ICSEA 2008

## nominations for awards and honors

2018	<b>Nomination for ACM SIGSOFT Distinguished Paper Award</b> “What Makes A Code Change Easier To Review: An Empirical Investigation On Code Change Reviewability”	ESEC/FSE 2018
2016	<b>Nomination for Best Paper Award</b> “On the reaction to deprecation of 25,357 clients of 4+1 popular Java APIs”	ICSME 2016
2015	<b>Nomination for Best Paper Award</b> “Untangling Fine-Grained Code Changes”	SANER 2015
2014	<b>Finalist for the Cor Baayen Award</b> ERCIM awards promising young researchers in computer science and applied mathematics.	ERCIM 2014
2013	<b>Nomination for ACM SIGSOFT Distinguished Paper Award</b> “Expectations, Outcomes, and Challenges of Modern Code Review”	ICSE 2013

## grants

2017	<b>SNSF Professorship - Data-driven Contemporary Code Review</b> On transforming code review into an engineering approach.	PI – approx. 1.4M EUR
2016	<b>STW Take Off - PReview Code</b> On an approach to improve the understandability of changes under review.	PI – 40k EUR
2015	<b>NWO Top-1 grant - Persistent Code Reviewing</b> On investigating techniques to persist and generalize code review efforts.	Co-PI – approx. 650k EUR
2014	<b>H2020 MSCA-ITN-2014-EID - SENECA</b> On data science for software engineering in the cloud.	Co-PI – 255k EUR of 2.2M EUR
	<b>INRIA Associate Team - Improving code review tools</b> On creating novel code review tools for dynamically typed languages.	Co-PI – 10K EUR
	<b>Google Faculty Research Award</b> On investigating the code review process.	PI – approx. 50k EUR

## invited talks

- 2018 **What Do Code Reviews at Microsoft and in OSS Have in Common?** Zurich, Switzerland  
adesso Schweiz
- 2017 **Evidence-based Code Review** London, United Kingdom  
56th CREST Open Workshop, University College London
- Contemporary Peer Code Review** Zurich, Switzerland  
Software Evolution and Architecture Lab, University of Zurich
- 2016 **Advances in Code Review** Amsterdam, The Netherlands  
BESTSELLER
- What Do Code Reviews at Microsoft and in OSS Have in Common?** Bruxelles, Belgium  
FOSDEM 2016
- Data-driven Contemporary Code Review** Zurich, Switzerland  
University of Zurich, Switzerland
- Data-driven Contemporary Code Review** Lugano, Switzerland  
Università della Svizzera Italiana, Switzerland
- 2015 **Code review needs the right approach** Bodegraven, The Netherlands  
Stabiplan BV
- Expectations, Outcomes, and Challenges of Modern Code Review** Berlin, Germany  
Software engineering research group (AG SE), Freie Universität Berlin
- Software Analytics needs the right data and the right questions** London, England  
Software Systems Engineering Group, University College London
- Mining Unstructured Software Data** Nijmegen, The Netherlands  
Digital security group, Radboud Universiteit Nijmegen
- Supporting Developers' Teamwork in The IDE** Vancouver, BC, Canada  
Software Analysis and Testing (SALT) lab, University of British Columbia
- 2014 **Mining Unstructured Software Data** Namur, Belgium  
PReCISE Research Center, Université de Namur
- 2012 **Pearls In The MUD, Mining Structured Data in NL Artifacts** Delft, The Netherlands  
Software Engineering Research Group (SERG), TU Delft
- Mining Development Email Archives** Victoria, Canada  
Computer Human Interaction & Software Engineering Lab, University of Victoria
- Mining Development Email Archives** Montreal, QC, Canada  
SOCCER Lab, École Polytechnique de Montréal
- Mining Development Email Archives** Kingston, ON, Canada  
Software Analysis and Intelligence Lab, Queen's University

# professional service

## conferences

- 2019 ASE (PC member), ESEC/FSE (PC member), ICPC (PC member), ICSME (PC member), MSR (PC member, SC member)
- 2018 ASE (PC member), ESEC/FSE (PC member), ICPC (PC member), MSR (PC member, SC member)
- 2017 ESEC/FSE (PC member), ICSE (PC member)
- 2016 ICSE (PC member), ICPC (PC member), ICSME (Tool demo track, co-chair), IWESEP (PC member), MSR (PC member), MUD (SC member), SANER (ERA track, co-chair)
- 2015 ASE (Tool demo track, PC member), FSE (Tool demo track, PC member), ICPC (PC member), ICSE (social networking co-chair), ICSME (PC member), ISEC (PC member), MSR (PC member), MUD (SC member), SANER (PC member)
- 2014 ACM SRC (PC member), CSEE&T (PC member), CSMR-WCRE (PC member), ECOOP (artifact evaluation, PC member), ICSME (PC member), MSR (PC member), MUD (co-organizer)
- 2013 CSS&T (PC member), CSMR (ERA track, PC member) ICPC (ERA track, PC member), ICSM (Tool demo track, PC member), MSR (Mining challenge chair), MUD (co-organizer), NaturalISE (co-organizer), WCRE (PC member)
- 2012 CSEE&T (PC member), CSMR (publicity chair), ICSM (publicity chair), MSR (Mining challenge, PC member), MUD (co-organizer)

## journal referee

- TOSEM** ACM Transactions on Software Engineering and Methodology, ACM.
- TSE** IEEE Transactions on Software Engineering, IEEE Computer Society.
- EMSE** Empirical Software Engineering: An International Journal, Springer.
- JSEP** Journal of Software: Evolution and Process (formerly known as Journal on Software Maintenance and Evolution: Research and Practice, JSME), John Wiley & Sons.
- JSS** Journal of Systems and Software, Elsevier.
- PeerJ** PeerJ Computer Science.

# teaching

## instructor

- 2018–2019 **Data Science for Software Engineering (B.Sc. and M.Sc.)** instructor – University of Zurich  
This edition continued the structure of the previous one.
- Software Testing (B. Sc. and M.Sc.)** co-instructor – University of Zurich  
Teaching theoretical and practical aspects of software testing.
- 2017–2018 **Data Science for Software Engineering (B.Sc. and M.Sc.)** instructor – University of Zurich  
Students explore the topic and work on creating a research proposal on a novel idea.
- Mining Software Repositories (M.Sc.)** instructor – TU Delft  
This edition of the course has seen an increase in the involvement of my supervised Ph.D. students, in order to make them gain experience in advising Master students on topics related to their Ph.D. tracks.
- Software Engineering Methods (B.Sc.)** co-instructor – TU Delft  
This edition was taught together with Dr. Palomba. A significant addition to the course regarded aspects of software security and defensive programming.
- 2016–2017 **Context Project (B.Sc.)** co-coordinator, SE instructor, and context teacher – TU Delft  
This edition continued the structure of the previous one.
- Mining Software Repositories (M.Sc.)** instructor – TU Delft  
This edition continued the structure of the previous one.
- Software Engineering Methods (B.Sc.)** instructor – TU Delft  
This edition continued the structure of the previous one.
- 2015–2016 **Context Project (B.Sc.)** co-coordinator, SE instructor, and context teacher – TU Delft  
In this edition of the course, I also cover the role of context teacher and lead 5 groups of students in the development of “tools for software engineering.”
- Mining Software Repositories (M.Sc.)** instructor – TU Delft  
This edition focused on using data from past Mining Software Conferences to either replicate previous work or conduct novel research in software analytics.
- Software Engineering Methods (B.Sc.)** instructor – TU Delft  
In this edition, I kept the content and format from the previous edition, since results were satisfactory. I streamlined the logistics of the group project updating the software infrastructure and grading rubrics.
- 2014–2015 **Context Project (B.Sc.)** co-coordinator and SE instructor – TU Delft  
In this edition of the course, I also introduced the use of pull-based development model (on GitHub), static analysis tools, and code reviews.
- Mining Software Repositories (M.Sc.)** instructor – TU Delft  
Students learn about recent research on software analytics and implement a full data mining approach, from data collection and analysis scripts to final report writing.
- Software Engineering Methods (B.Sc.)** instructor – TU Delft  
In this edition, I introduced a group project. Students develop a fully working game and iterate over it using SCRUM and guided by weekly assignments.
- 2013–2014 **Context Project (B.Sc.)** co-coordinator and SE instructor – TU Delft  
Students learn to develop, implement, validate, present, and demonstrate a software product that satisfies the needs of an external party in a given non-ICT context. I co-coordinate the 8 teachers involved and I am responsible for supervising and evaluating the software engineering aspects of the course.
- Software Evolution Seminars (M.Sc.)** co-instructor – TU Delft  
Students form groups of two and present and debate in class recent papers about software evolution. The teachers guide the discussion and evaluate students’ presentations and participation in debates.
- Software Engineering Methods (B.Sc.)** instructor – TU Delft  
Students learn the basics of software engineering. Topics include requirements engineering, advanced OO programming, design patterns, UML, agile methodologies, software visualization.

## student supervision (since being a faculty member)

in progress	<b>Larissa Barbosa</b> <i>Code review and security</i>	since Nov 2018, <u>Ph.D. student</u> , University of Zurich
	<b>Enrico Fregnan</b> <i>Software Coupling</i>	since Aug 2018, <u>Ph.D. student</u> , University of Zurich
	<b>Pavlna Wurzelová</b> <i>Social Aspects Of Code Review</i>	since Jun 2018, <u>Ph.D. student</u> , University of Zurich
	<b>Vladimir Kovalenko</b> <i>Usefulness Of Developers' Tools</i>	since Oct 2016, <u>Ph.D. student</u> , TU Delft
	<b>Davide Spadini</b> <i>Collaborative Software Testing</i>	co-supervision with SIG, since Jul 2016, <u>Ph.D. student</u> , TU Delft
	<b>Luca Pascarella</b> <i>Software Quality</i>	since Jan 2016, <u>Ph.D. student</u> , TU Delft
	<b>Anand Sawant</b> <i>Mining fine-grained large-scale API usage</i>	since Jan 2016, <u>Ph.D. student</u> , TU Delft
	<b>Lorenzo Gasparini</b> <i>Software Visualization for Code Review</i>	<u>M.Sc. Thesis</u> , TU Delft
2018	<b>Enrico Fregnan</b> <i>Auto-ordering Code Changes for Review</i>	<u>M.Sc. Thesis</u> , TU Delft
	<b>Moritz Eck</b> <i>Automatic Categorization Of Flaky Tests</i>	<u>B.Sc. Thesis</u> , University of Zurich
2017	<b>Menno Oudshoorn and Bart de Jonge</b> <i>Unobtrusive automated support for software developers</i>	<u>Honors program</u> , TU Delft
	<b>Jorden van Breemen</b> <i>On the Effect of Code Quality on Agile Effort Estimations: The Case of Shell</i>	<u>M.Sc. Thesis</u> , TU Delft
2016	<b>Eva Anker and Tim van der Lippe and Thomas Smith</b> <i>Enhanced GitHub code review</i>	<u>B.Sc. Project</u> , TU Delft
2015	<b>Reinier Hartog</b> <i>Octopull: Supporting reviews of pull-requests</i>	<u>M.Sc. Thesis</u> , TU Delft
	<b>Ahmad Yandriansyah Reza</b> <i>Requirements Engineering Practices in Global Software Engineering Organization</i>	co-supervision with D.M. van Solingen, <u>M.Sc. Thesis</u> , TU Delft
	<b>Anand Sawant</b> <i>Mining fine-grained API usage from GitHub</i>	<u>M.Sc. Thesis</u> , TU Delft
	<b>Vincent Hellendoorn</b> <i>Empirical Software Linguistics: An Investigation of Code Reviews, Recommendations and Faults</i>	<u>M.Sc. Thesis</u> , TU Delft

**Best Master Thesis of the EEMCS Faculty 2015**

## references (alphabetically ordered)

### **dr. Christian Bird**

Role: Researcher  
Affiliation: Microsoft Research  
Address: One Microsoft Way, Redmond, WA 98052 USA  
Email: [Christian.Bird@microsoft.com](mailto:Christian.Bird@microsoft.com)  
Web: <https://www.microsoft.com/en-us/research/people/cbird/>  
Phone: +1 425 703 8450

### **prof. dr. Premkumar T. Devanbu**

Role: Professor  
Affiliation: University of California, Davis  
Address: Department of Computer Science, Engineering II, Davis CA 95616 (USA)  
Email: [devanbu@cs.ucdavis.edu](mailto:devanbu@cs.ucdavis.edu)  
Web: <http://www.cs.ucdavis.edu/~devanbu>  
Phone: +1 530 752-7324

### **prof. dr. Massimiliano Di Penta**

Role: Associate Professor  
Affiliation: University of Sannio  
Address: Palazzo ex Poste, Via Traiano, I-82100 Benevento (Italy)  
Email: [dipenta@unisannio.it](mailto:dipenta@unisannio.it)  
Web: <http://www.ing.unisannio.it/mdipenta/>  
Phone: +39 0824 305536

### **prof. dr. Harald Gall**

Role: Professor and Dean  
Affiliation: University of Zurich  
Address: Binzmühlestrasse 14, 8050 Zurich  
Email: [gall@ifi.uzh.ch](mailto:gall@ifi.uzh.ch)  
Web: <http://www.ifi.uzh.ch/en/seal/people/gall.html>  
Phone: +41 44 635 43 21

### **prof. dr. Michele Lanza**

Role: Professor and prorector  
Affiliation: Università della Svizzera Italiana  
Address: Via G. Buffi 13, CH-6904 Lugano, Switzerland  
Email: [michele.lanza@usi.ch](mailto:michele.lanza@usi.ch)  
Web: <http://www.inf.usi.ch/lanza/>  
Phone: +41 58 666 4659

### **prof. dr. Margaret-Anne Storey**

Role: Professor  
Affiliation: University of Victoria  
Address: Department of Computer Science, University of Victoria, PO Box 3055,  
STN CSC Victoria, B.C. Canada, V8W 3P6  
Email: [mstorey@uvic.ca](mailto:mstorey@uvic.ca)  
Web: <http://margaretstorey.com>  
Phone: +1 250 472 5713

# complete list of publications

Peer-reviewed conferences are accepted in software engineering as high-quality scholarly articles. This point is expressed in the article “Research Evaluation for Computer Science” in *Communications of the ACM* (April 2009).<sup>1</sup>

The ACM/IEEE International Conference on Software Engineering (ICSE) and the ACM SIGSOFT International Symposium on the Foundations of Software Engineering (ESEC/FSE) are considered the most prestigious conferences, with the highest standards in the software engineering field.<sup>2</sup> I have published 10 papers in the main tracks at ICSE and FSE, of which two have been nominated for an ACM SIGSOFT Distinguished Paper Award (ICSE 2013, FSE 2018) and one has received an ACM SIGSOFT Distinguished Paper Award (ICSE 2016).

According to Google scholar (Dec 23, 2018), my h-index is 23 with 1,896 citations showing an upward trend.

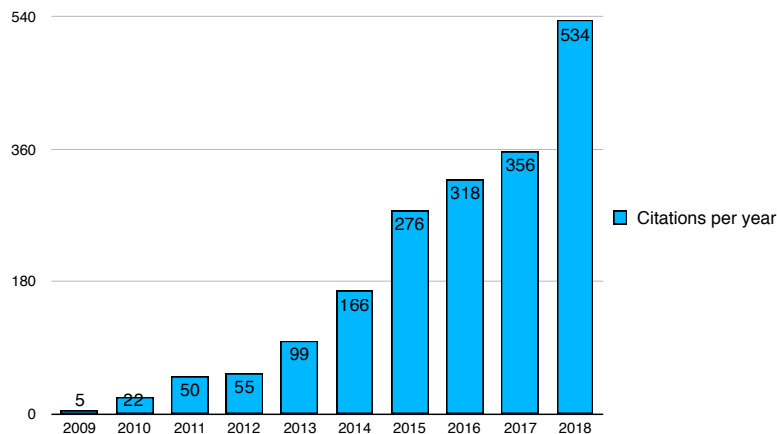


Figure 1: Citations per year, according to Google Scholar on Dec 23, 2018.

## international peer-reviewed conference papers

[ICSE 2019] When Code Completion Fails: A Case Study On Real-World Completions

Vincent Hellendoorn, Sebastian Proksch, Harald C. Gall, Alberto Bacchelli

*41st ACM/IEEE International Conference on Software Engineering*, in press, 2019

[ICSE 2019] Test-Driven Code Review: An Empirical Study

Davide Spadini, Fabio Palomba, Tobias Baum, Stefan Hanenberg, Magiel Bruntink, Alberto Bacchelli

*41st ACM/IEEE International Conference on Software Engineering*, in press, 2019

[CSCW 2018] Information Needs in Contemporary Code Review

Luca Pascarella, Davide Spadini, Fabio Palomba, Magiel Bruntink, Alberto Bacchelli

*21st ACM conference on Computer-Supported Cooperative Work and Social Computing*, in press, 2018

**Best Paper Honorable Mention**

[MOBILESoft 2018] Self-reported activities of Android developers

Luca Pascarella, Fabio Palomba, Franz-Xaver Geiger, Dario Di Nucci, Ivano Malavolta, Alberto Bacchelli

*5th International Conference on Mobile Software Engineering and Systems*, pp. 144–155, 2018

[ICSE 2018 - SEIP] Modern Code Review: A Case Study at Google

Caitlin Sadowski, Emma Söderberg, Luke Church, Michal Sipko, Alberto Bacchelli

*40th ACM/IEEE International Conference on Software Engineering, Software Engineering in Practice Track*, pp. 181–190, 2018

[ICSE 2018] Understanding Developers’ Needs on Deprecation as a Language Feature

Anand Ashok Sawant, Mauricio Aniche, Arie van Deursen, Alberto Bacchelli

*40th ACM/IEEE International Conference on Software Engineering*, pp. 561–571, 2018

<sup>1</sup>This point is also underlined by Michael D. Ernst (Professor at the University of Washington, and previously researcher at Microsoft Research and tenured professor at MIT): “conference papers are arguably more prestigious than journal publications: oftentimes, conferences have higher standards and lower acceptance rates.” For a copy: <http://sback.it/other/ernst-conf-vs-journal-uscis.pdf>.

<sup>2</sup>This is also expressed in the article: “Automatic and versatile publications ranking for research institutions and scholars” in *Comm. of the ACM* (June 2007)



- [ICSE 2018] When Testing Meets Code Review: Why and How Developers Review Tests  
Davide Spadini, Mauricio Aniche, Margaret Storey, Magiel Bruntink, [Alberto Bacchelli](#)  
*40th ACM/IEEE International Conference on Software Engineering*, pp. 677–687, 2018
- [ICSME 2018] On The Relation of Test Smells to Software Code Quality  
Davide Spadini, Fabio Palomba, Andy Zaidman, Magiel Bruntink, [Alberto Bacchelli](#)  
*34th IEEE International Conference on Software Maintenance and Evolution*, pp. 1–12, 2018
- [ICSME 2018] Why are features deprecated? An investigation into the motivation behind deprecation  
Anand Ashok Sawant, Guangzhe Huanag, Gabriel Vilen, Stefan Stojkovski, [Alberto Bacchelli](#)  
*34th IEEE International Conference on Software Maintenance and Evolution*, pp. 13–24, 2018
- [ASE 2018] Mining File Histories: Should We Consider Branches?  
Vladimir Kovalenko, Fabio Palomba, [Alberto Bacchelli](#)  
*33rd IEEE/ACM International Conference on Automated Software Engineering*, pp. 202–213, 2018
- [ASE 2018 - New Ideas] Continuous Code Quality: Are We (Really) Doing That?  
Carmine Vassallo, Fabio Palomba, [Alberto Bacchelli](#), Harald C. Gall  
*33rd IEEE/ACM International Conference on Automated Software Engineering, New Ideas Track*, pp. 790–795, 2018
- [MSR 2018 - Data] A graph-based dataset of commit history of real-world Android apps  
Franz-Xaver Geiger, Ivano Malavolta, Luca Pascarella, Fabio Palomba, Dario Di Nucci, [Alberto Bacchelli](#)  
*15th International Conference on Mining Software Repositories, Data Track*, pp. 30–33, 2018
- [MSR 2018] How Is Video Game Development Different from Software Development in Open Source?  
Luca Pascarella, Fabio Palomba, Massimiliano Di Penta, [Alberto Bacchelli](#)  
*15th International Conference on Mining Software Repositories*, pp. 392–402, 2018
- [ESEC/FSE 2018] What Makes A Code Change Easier To Review: An Empirical Investigation On Code Change Reviewability  
Achyudh Ram, Anand Ashok Sawant, Marco Castelluccio, [Alberto Bacchelli](#)  
*26th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pp. 201–212, 2018  
**Nominated for ACM SIGSOFT Distinguished Paper Award**
- [ESEC/FSE 2018 - Tool Demo] PyDriller: Python framework for mining software repositories  
Davide Spadini, Maurício Finavaro Aniche, [Alberto Bacchelli](#)  
*26th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pp. 908–911, 2018
- [SANER 2018] Re-evaluating method-level bug prediction  
Luca Pascarella, Fabio Palomba, [Alberto Bacchelli](#)  
*25th International Conference on Software Analysis, Evolution and Reengineering, RENE track*, pp. 592–601, 2018
- [ICSE 2017 - Introspection] Double-blind review in software engineering venues: the community’s perspective  
[Alberto Bacchelli](#), Moritz Beller  
*39th ACM/IEEE International Conference on Software Engineering (Companion Volume)*, pp. 385–396, 2017
- [ICSME 2017] On the Optimal Order of Reading Source Code Changes for Review  
Tobias Baum, Kurt Schneider, [Alberto Bacchelli](#)  
*33rd IEEE International Conference on Software Maintenance and Evolution*, pp. 329–340, 2017
- [MSR 2017] Classifying code comments in Java open-source software systems  
Luca Pascarella, [Alberto Bacchelli](#)  
*14th International Conference on Mining Software Repositories*, pp. 227–237, 2017  
**ACM SIGSOFT Distinguished Paper Award**
- [MSR 2017] To mock or not to mock? An empirical study on mocking practices  
Davide Spadini, Mauricio Aniche, Magiel Bruntink, [Alberto Bacchelli](#)  
*14th International Conference on Mining Software Repositories*, pp. 402–412, 2017
- [ESEM 2016] Social Diversity and Activity Levels of Open Source Software Projects on GitHub  
Joop Aué, Michiel Haisma, Kristin Fjola Tomasdottir, [Alberto Bacchelli](#)  
*10th International Symposium on Empirical Software Engineering and Measurement*, 41:1–41:6, 2016
- [GECCO 2016] A Search-based Training Algorithm for Cost-aware Defect Prediction  
Annibale Panichella, Carol V. Alexandru, Sebastiano Panichella, [Alberto Bacchelli](#), Harald C. Gall  
*Genetic and Evolutionary Computation Conference*, pp. 1077–1084, 2016
- [ICSE 2016] Work Practices and Challenges in Pull-Based Development: The Contributor’s Perspective  
Georgios Gousios, Margaret-Anne Storey, [Alberto Bacchelli](#)

38th ACM/IEEE International Conference on Software Engineering, pp. 285–296, 2016  
**ACM SIGSOFT Distinguished Paper Award**

[ICSE 2016] On the “Naturalness” of Buggy Code

Baishakhi Ray, Vincent Hellendoorn, Saheel Godhane, Zhaopeng Tu, Alberto Bacchelli, Premkumar Devanbu  
38th ACM/IEEE International Conference on Software Engineering, pp. 428–439, 2016

[ICSME 2016] On the reaction to deprecation of 25,357 clients of 4+1 popular Java APIs

Anand Sawant, Romain Robbes, Alberto Bacchelli  
32nd International Conference on Software Maintenance and Evolution, pp. 400–410, 2016  
**Nominated for Best Paper Award**

[SCAM 2016] A Security Perspective on Code Review: The Case of Chromium.

Marco Biase, Magiel Bruntink, Alberto Bacchelli  
16th IEEE International Working Conference on Source Code Analysis and Manipulation, pp. 21–30, 2016

[FSE 2016 - Tool Demo] Visualizing code and coverage changes for code review

Sebastiaan Oosterwaal, Arie Deursen, Roberta Coelho, Anand Ashok Sawant, Alberto Bacchelli  
24th ACM SIGSOFT International Symposium on Foundations of Software Engineering, pp. 1038–1041, 2016

[CSCW 2015] Supporting Developers’ Coordination in the IDE

Anja Guzzi, Alberto Bacchelli, Yann Riche, Arie van Deursen  
18th ACM conference on Computer-Supported Cooperative Work and Social Computing, pp. 518–532, 2015  
**Best Paper Award**

[MSR 2015] Will they like this? Evaluating Code Contributions With Language Models

Vincent Hellendoorn, Premkumar Devanbu, Alberto Bacchelli  
12th Working Conference on Mining Software Repositories, pp. 157–167, 2015

[MSR 2015 - Data] A Dataset For API Usage

Anand Sawant, Alberto Bacchelli  
12th Working Conference on Mining Software Repositories, Data Track, pp. 506–509, 2015

[SANER 2015] Untangling Fine-Grained Code Changes

Martin Dias, Alberto Bacchelli, Georgios Gousios, Damien Cassou, Stéphane Ducasse  
22nd IEEE International Conference on Software Analysis, Evolution, and Reengineering, pp. 341–350, 2015  
**Nominated for Best Paper Award**

[ICSME 2014] Improving Low Quality Stack Overflow Post Detection

Luca Ponzanelli, Andrea Mocchi, Alberto Bacchelli, Michele Lanza, David Fullerton  
30th International Conference on Software Maintenance and Evolution, Industrial Track, pp. 541–544, 2014

[MSR 2014] Modern Code Reviews in Open-Source Projects: Which Problems Do They fix?

Moritz Beller, Alberto Bacchelli, Andy Zaidman, Elmar Juergens  
11th Working Conference on Mining Software Repositories, pp. 202–211, 2014

[QSIC 2014] Quantitatively Exploring Non-code Software Artifacts

Luca Bigliardi, Michele Lanza, Alberto Bacchelli, Marco D’Ambros  
14th International Conference on Quality Software, pp. 286–295, 2014

[QSIC 2014] Understanding and Classifying the Quality of Technical Forum Questions

Luca Ponzanelli, Andrea Mocchi, Alberto Bacchelli, Michele Lanza  
14th International Conference on Quality Software, pp. 343–352, 2014

[CSMR 2013] Leveraging Crowd Knowledge for Software Comprehension and Development

Luca Ponzanelli, Alberto Bacchelli, Michele Lanza  
17th European Conference on Software Maintenance and Reengineering, pp. 57–66, 2013

[ICSE 2013] Expectations, Outcomes, and Challenges of Modern Code Review

Alberto Bacchelli, Christian Bird  
35th ACM/IEEE International Conference on Software Engineering, pp. 710–719, 2013  
**Nominated for ACM SIGSOFT Distinguished Paper Award**

[ICSE 2013 - Tool Demo] Seahawk: Stack Overflow in the IDE

Luca Ponzanelli, Alberto Bacchelli, Michele Lanza  
35th International Conference on Software Engineering, Tool Demo Track, pp. 1295–1298, 2013

[ICSM 2013 - ERA] Which Feature Location Technique is Better?

Emily Hill, Alberto Bacchelli, Dave W. Binkley, Bogdan Dit, Dawn J. Lawrie, Rocco Oliveto

29th IEEE International Conference on Software Maintenance, pp. 408–411, 2013

[ICPC 2013] Manhattan: Supporting Real-Time Visual Team Activity Awareness

Michele Lanza, Marco D'Ambros, Alberto Bacchelli, Lile Hattori, Francesco Rigotti

21st IEEE International Conference on Program Comprehension, ERA Track, pp. 207–210, 2013

[MSR 2013] Communication in Open Source Software Development Mailing Lists

Anja Guzzi, Alberto Bacchelli, Michele Lanza, Martin Pinzger, Arie Deursen

10th IEEE Working Conference on Mining Software Repositories, pp. 277–286, 2013

[ICSE 2012] Content Classification of Development Emails

Alberto Bacchelli, Tommaso Sasso, Marco D'Ambros, Michele Lanza

34th ACM/IEEE International Conference on Software Engineering, pp. 375–385, 2012

[CSMR 2011] RTFM (Read The Factual Mails) –Augmenting Program Comprehension with REmail

Alberto Bacchelli, Michele Lanza, Vitezslav Humpa

15th IEEE European Conference on Software Maintenance and Reengineering, pp. 15–24, 2011

[ICSE 2011 - Tool Demo] Miler: A Toolset for Exploring Email Data

Alberto Bacchelli, Michele Lanza, Marco D'Ambros

33rd ACM/IEEE International Conference on Software Engineering, pp. 1025–1027, 2011

[ICSE 2011 - DocSym] Exploring, exposing, and exploiting emails to include human factors in software engineering

Alberto Bacchelli

33rd ACM/IEEE International Conference on Software Engineering, pp. 1074–1077, 2011

[ASE 2011] Extracting Structured Data from Natural Language Documents with Island Parsing

Alberto Bacchelli, Anthony Cleve, Michele Lanza, Andrea Mocchi

26th IEEE/ACM International Conference On Automated Software Engineering, pp. 476–479, 2011

[FASE 2010] Are Popular Classes More Defect Prone?

Alberto Bacchelli, Marco D'Ambros, Michele Lanza

13th International Conference on Fundamental Approaches to Software Engineering, pp. 59–73, 2010

[ICSE 2010] Linking E-Mails and Source Code Artifacts

Alberto Bacchelli, Michele Lanza, Romain Robbes

32nd International Conference on Software Engineering, pp. 375–384, 2010

[ICPC 2010] Extracting Source Code from E-Mails

Alberto Bacchelli, Marco D'Ambros, Michele Lanza

18th IEEE International Conference on Program Comprehension, pp. 24–33, 2010

[QSIC 2010] On the Impact of Design Flaws on Software Defects

Marco D'Ambros, Alberto Bacchelli, Michele Lanza

10th International Conference on Quality Software, pp. 23–31, 2010

[WCRE 2009] Benchmarking Lightweight Techniques to Link E-Mails and Source Code

Alberto Bacchelli, Marco D'Ambros, Michele Lanza, Romain Robbes

16th IEEE Working Conference on Reverse Engineering, pp. 205–214, 2009

**Best Paper Award**

[ICSEA 2008] On the Effectiveness of Manual and Automatic Unit Test Generation

Alberto Bacchelli, Paolo Ciancarini, Davide Rossi

3rd International Conference on Software Engineering Advances, pp. 252–257, 2008

**Best Paper Award**

## peer-reviewed journal articles

[EMSE 2018] Associating Working Memory Capacity and Code Change Ordering with Code Review Performance

Tobias Baum, Kurt Schneider, Alberto Bacchelli

Empirical Software Engineering. *EMSE 2018 (2018) in press. Springer, 2018*

[IST 2018] A Survey on Software Engineering Coupling Relations and Tools

Enrico Fregnan, Tobias Baum, Fabio Palomba, Alberto Bacchelli

Information and Software Technology. *IST 2018 (2018) in press. Springer, 2018*

[JSS 2018] Fine-Grained Just-In-Time Defect Prediction

Luca Pascarella, Fabio Palomba, Alberto Bacchelli

[EMSE 2018] Mock objects for testing Java systems: Why and how developers use them, and how they evolve

Davide Spadini, Mauricio Aniche, Magiel Bruntink, [Alberto Bacchelli](#)  
Empirical Software Engineering. *EMSE 2018 (2018) in press. Springer, 2018*

[TSE 2018] Does reviewer recommendation help developers?

Vladimir Kovalenko, Nava Tintarev, Evgeny Pasyukov, Christian Bird, [Alberto Bacchelli](#)  
IEEE Transactions on Software Engineering. *TSE 2018 (2018) in press. IEEE, 2018*

[EMSE 2018] On the reaction to deprecation of clients of 4+1 popular Java APIs and the JDK

Anand Sawant, Romain Robbes, [Alberto Bacchelli](#)  
Empirical Software Engineering. *EMSE 2018 23.4 (2018) pp. 2158–2197. Springer, 2018*

[SCNTDX 2017] Does single blind peer review hinder newcomers?

Marco Seeber, [Alberto Bacchelli](#)  
Scientometrics. *SCNTDX 2017 113.1 (2017) pp. 567–585. Springer, 2017*

[SCP 2017] Mining Structured Data in Natural Language Artifacts with Island Parsing

[Alberto Bacchelli](#), Andrea Mocchi, Anthony Cleve, Michele Lanza  
Science of Computer Programming. *SCP 2017 (2017) pp. 31–55. Elsevier, 2017*

[EMSE 2017] fine-GRAPe: fine-Grained API usage Extractor – An Approach and Dataset to Investigate API Usage

Anand Sawant, [Alberto Bacchelli](#)  
Empirical Software Engineering. *EMSE 2017 22.3 (2017) pp. 1348–1371. Springer, 2017*

[SCP 2015] IRISH: A Hidden Markov Model to Detect Coded Information Islands in Free Text

Luigi Cerulo, Max Di Penta, [Alberto Bacchelli](#), Michele Ceccarelli, Gerardo Canfora  
Science of Computer Programming. *SCP 2015 105 (July 2015) pp. 26–43. Elsevier, 2015*

[ETRD 2015] Team design communication patterns in e-learning design and development

Chrysi Rapanta, Marcelo Maina, Nicole Lotz, [Alberto Bacchelli](#)  
Educational Technology Research and Development. *ETRD 2015 61.4 (2013) pp. 581–605. Springer, 2013*

[IJAS 2009] How to compare and exploit different techniques for unit-test generation

[Alberto Bacchelli](#), Paolo Ciancarini, Davide Rossi  
International Journal On Advances in Software. *IJAS 2009 2.1 (2009) pp. 131–146. Iaria, 2009*

## international peer-reviewed workshop papers

[CHASE 2018] Code review for newcomers: is it different?

Vladimir Kovalenko, [Alberto Bacchelli](#)  
1st IEEE/ACM International Workshop on API Usage and Evolution, pp. 29–32, 2018

[MaLTeSQuE 2018] Investigating type declaration mismatches in Python

Luca Pascarella, Achyudh Ram, Azqa Nadeem, Dinesh Bisesser, Norman Knyazev, [Alberto Bacchelli](#)  
Workshop on Machine Learning Techniques for Software Quality Evaluation, pp. 43–48, 2018

[WAPI 2017] Mining Motivated Trends of Usage of Haskell Libraries

Mark Juchli, Lars Krombeen, Shashank Rao, Chak Shun Yu, Anand Ashok Sawant, [Alberto Bacchelli](#)  
11th International Workshop on Cooperative and Human Aspects of Software Engineering, pp. 11–14, 2017

[RSSE 2012] Harnessing Stack Overflow for the IDE

[Alberto Bacchelli](#), Luca Ponzanelli, Michele Lanza  
3rd International Workshop on Recommendation Systems for Software Engineering, pp. 26–30, 2012

## other peer-reviewed publications

[MSR 2015 - Mining Challenge] Quality questions need quality code: Classifying code fragments on StackOverflow

Maarten Duijn, Adam Kucera, [Alberto Bacchelli](#)  
12th Working Conference on Mining Software Repositories, Mining Challenge Track, 2015

[MSR 2015 - Mining Challenge] ETA: Estimated Time of Answer, Predicting Response Time in Stack Overflow

Jeffrey Goderie, Brynjolfur Mar Georgsson, Bastiaan Graafeiland, [Alberto Bacchelli](#)  
12th Working Conference on Mining Software Repositories, Mining Challenge Track, 2015

[MSR 2015 - Mining Challenge] Automatic Assessments of Code Explanations: Predicting answering times on Stack Overflow

Selman Ercan, Quinten Stokkink, [Alberto Bacchelli](#)

12th Working Conference on Mining Software Repositories, Mining Challenge Track, 2015

[MSR 2015 - Mining Challenge] One-day flies on StackOverflow - Why the vast majority of StackOverflow users only posts once

Rogier Slag, Mike Waard, [Alberto Bacchelli](#)

12th Working Conference on Mining Software Repositories, Mining Challenge Track, 2015

[NaturaLiSE@ICSE 2013] 1st International workshop on Natural Language analysis in Software Engineering

Lori L. Pollock, David W. Binkley, Dawn J. Lawrie, Emily Hill, Rocco Oliveto, Gabriele Bavota, [Alberto Bacchelli](#)

35th International Conference on Software Engineering, 2013

[MUD@WCRE 2013] Workshop on Mining Unstructured Data (MUD) ... because “mining unstructured data is like fishing in muddy waters”!

[Alberto Bacchelli](#), Nicolas Bettenburg, Latifa Guerrouj, Sonia Haiduc

19th Working Conference on Reverse Engineering, 2012

[MUD@WCRE 2012] Workshop on Mining Unstructured Data (MUD) ... because “mining unstructured data is like fishing in muddy waters”!

[Alberto Bacchelli](#), Nicolas Bettenburg, Latifa Guerrouj

19th Working Conference on Reverse Engineering, 2012

[CSEET 2011 - Poster] Erase and Rewind - Learning by Replaying Examples

Lile Hattori, [Alberto Bacchelli](#), Mircea Lungu, Michele Lanza

24th International Conference on Software Engineering Education and Training, 2011

## phd thesis

Mining Unstructured Software Data

[Alberto Bacchelli](#)

*Università della Svizzera Italiana, Switzerland, June 2013*