



# ISR Research Forum

## Where Research Meets the Real World

**Friday June 2, 2017**

**9:00 am – 5:30 pm**

**University of California, Irvine ♦ 6011 Donald Bren Hall**

### Welcome!

8:30

#### Registration

Donald Bren Hall, 6th Floor Foyer

9:00 – 9:10

#### Welcome

Cristina Videira Lopes, ISR Associate Director and Professor, Informatics

9:10 – 10:10

#### Keynote I



*“Lessons from the Jungle of Open Source Big Data Development”*

**Owen O'Malley**

Co-Founder and Technical Fellow  
Hortonworks

10:10 – 10:40

#### Break

10:40 – 12:00

#### Presentations I

Session Chair: David Redmiles, Professor, Informatics

*“Research at Scale”*

Cristina Videira Lopes, ISR Associate Director and Professor, Informatics

*“Understanding and Analyzing Software Execution Behavior”*

James A. Jones, Associate Professor, Informatics

*“Modeling the Workplace Experience Through Precision-Tracking of Behavior”*

Gloria Mark, Professor, Informatics

*“Privacy by Design in an IoT Environment”*

Alfred Kobsa, Professor, Informatics

12:00 – 1:45 pm

#### Open House + Lunch

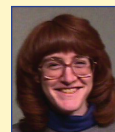
**Lunch** – Lunch served in DBH 6011 foyer at 12:00 pm. Eat at location of your choice.

**Open House** – **begins at 12:30**. See demos and posters! Located on Fifth floor: 5011 and foyer.

2:00 – 3:00

#### Keynote II

Session Chair: Prof. André van der Hoek, Professor and Chair, Informatics



*“Cognitive Software Engineering: Helping Developers Work Successfully in a Cognitive World”*

**Peri Tarr**

Principal Research Staff Member  
IBM T.J. Watson Research Center

3:00 – 3:15

#### Short Break

3:15 – 4:15

#### Presentations II

Session Chair: Sam Malek, Associate Professor, Informatics

*“Who Hit Me and Why Does It Matter: Legal and Policy Challenges in Attributing Cyber Attacks”*

Bryan Cunningham, Executive Director, UCI Cybersecurity Policy & Research Institute (CPRI)

*“Automatic Generation of Inter-Component Communication Exploits for Android Applications”*

Joshua Garcia, Associate Project Scientist, ISR

*“Privacy Compliance by Design: Enhancing Industry Software Practices For Compliance with Privacy Laws and Regulations”*

Sameer Patil, Assistant Professor, Indiana University Bloomington

4:15 – 5:30 pm

#### Reception with Posters

## Keynote I



### *“Lessons from the Jungle of Open Source Big Data Development”*

**Owen O'Malley**

Co-Founder and Technical Fellow

Hortonworks

**ABSTRACT.** The most exciting part of working as a software engineer is that the field evolves so rapidly that we are always learning new skills. Because the field is so dynamic, researchers are well positioned to help practitioners with their problems. My talk will use stories from my 11 years of working in Apache's big data projects (mostly Hadoop, Hive, and ORC) and what kinds of tools would have made a big impact for us. For example, when we originally developed the ORC file format as part of Hive, it was only lightly integrated. However, over the next few years, it became deeply entangled with the rest of Hive. We decided to factor it out to a separate project, but it took a lot of effort because the tool support was not very good. Another example is that while open source is great for allowing small teams to accomplish a lot efficiently, there are also dangers involved if the project owners have a significantly different development model. In particular, Google has released some great open source software like ProtoBuf and Guava. However, Google's release engineering process rebuilds the entire world every night. That means that they have very different requirements for compatibility and create a lot of pain for the rest of us. My talk will include these examples and others to encourage collaboration between the open source community and researchers.

**BIO.** Owen O'Malley is a co-founder and technical fellow at Hortonworks, a rapidly growing company (25 to 750 employees in 4 years), which develops the completely open source Hortonworks Data Platform (HDP). HDP includes Hadoop and the large ecosystem of big data tools that enterprises need for their data analytics. Owen has been working on Hadoop since the beginning of 2006 at Yahoo, was the first committer added to the project, and used Hadoop to set the Gray sort benchmark in 2008 and 2009. In the last 8 years, he has been the architect of MapReduce, Security, and now Hive. Recently he has been driving the development of the ORC file format and adding ACID transactions to Hive. Before working on Hadoop, he worked on Yahoo Search's WebMap project, which was the original motivation for Yahoo to work on Hadoop. Prior to Yahoo, he wandered between testing (UCI), static analysis (Reasoning), configuration management (Sun), and software model checking (NASA). He received his PhD in Software Engineering from University of California, Irvine.

## Keynote II



### *“Cognitive Software Engineering: Helping Developers Work Successfully in a Cognitive World”*

**Peri Tarr**

Principal Research Staff Member

IBM T.J. Watson Research Center

**ABSTRACT.** Cognitive software--which uses machine learning and data science to enable the software to “partner” with humans and augment human abilities--is here. Cognitive software engineering poses some unique challenges to the software practitioner and research communities. One of these is embodied in the cognitive components themselves: they must be trained and tuned appropriately, so that they can understand and respond correctly to users of the cognitive application. It is, however, very difficult to train these components so that they produce good results in your application; to diagnose issues that are impeding their performance; and even to know when they are “good enough” for use. Traditionally, only machine learning and data science experts have been able to produce well-performing cognitive components successfully. Most developers do not have the required knowledge and expertise, but they are the ones creating and evolving cognitive software.

In this talk, I will describe some of the challenges that cognitive software engineering poses for the software engineering research and practitioner communities. I will also demonstrate FARCAST, a framework and tool that aims to help software engineering experts to produce well-performing cognitive components for cognitive applications. FARCAST leverages expert knowledge and insights to help reduce the time and expertise required to produce an appropriately trained model for your cognitive application.

**BIO.** Peri Tarr received her BS in Zoology from the University of Massachusetts Amherst in 1986, and her MS and PhD in Computer Science from the University of Massachusetts Amherst (1992 and 1996, respectively). Between her BS and MS/PhD, she worked full-time at the University of Massachusetts Physical Plant, attempting to introduce an automated system to help with the Plant's operations. After receiving her PhD, she joined the IBM Thomas J. Watson Research Center as a Research Staff Member in 1996, where she worked on and led various projects relating to issues of software composition, morphogenic software, and aspect-oriented software development.

Her work on multi-dimensional separation of concerns was recognized as the Most Influential Paper at the 2009 International Conference on Software Engineering (ICSE). She currently serves as chief architect for Governance of Software Development, an IBM Research initiative that ties together the tools for teams of developers with the planning and financial management aspects required by enterprises.

Dr. Tarr was the 2005 program chair of the Aspect-Oriented Software Development conference and was the 2006 general chair of ACM SIG-PLAN's OOPSLA 2006 Conference.

**Location: DBH 5th floor, in room 5011 and the 5th floor foyer.**

## **Spider Lab: Software Analysis and Visualization for Debugging, Comprehension, and Maintenance**

Prof. **James A. Jones**



The Spider Lab creates techniques for offering automatic recommendations for common software maintenance tasks.

- “Phase Detection to Assist Software Comprehension” (Poster)  
Yang Feng
- “Cerebro: Revealing Runtime Behaviors in Software Executions” (Poster and Demo)  
J.Y. Ku
- “Summarizing Method Invocations for Runtime Data and Control Flow Analysis” (Poster)  
Vijay Krishna Palepu

## **Personalization and Privacy Lab**

Prof. **Alfred Kobsa**

The personalization and privacy lab conducts research on tailoring human-computer interaction to the needs of each individual user, and on reconciling the benefits that personalization provides with the privacy concerns that it evokes.

- “The Effect of Sensory Stimuli on the Performance of Security-Critical Tasks” (Poster)  
Bruce Berg (Cognitive Science), Tyler Michael Kaczmarek (Computer Science), Alfred Kobsa (Informatics and Computer Science), and Gene Tsudik (Computer Science)
- “Cross-Cultural Privacy Prediction” (Poster)  
Yao Li, Bart P. Knijnenburg (Clemson University), Alfred Kobsa, and M-H. Carolyn Nguyen (Microsoft Corporation)

## **The Mondego Group**

Prof. **Cristina Videira Lopes**



The Mondego group conducts research in large systems and large data.

- “Relational Types: Summoning the Power of Views” (Poster)  
Rohan Achar and Cristina V. Lopes
- “File Fragment Clones: A Language Independent Clone Detection With Variable Granularity” (Poster)  
Yixian Chen and Cristina V. Lopes
- “End-to-End Regression Testing of Distributed Systems” (Poster and Demo)  
Eugenia Gabrielov and Cristina V. Lopes
- “DéjàVu: A Map of Code Duplicates on GitHub” (Poster)  
Cristina V. Lopes, Petr Maj (Czech Technical University), Pedro Martins, Vaibhav Saini, Hitesh Sajani (Microsoft Research), Di Yang, Jakub Zitny (Czech Technical University), and Jan Vitek (Northeastern University)
- “MetaCC: Harnessing Source Code Metadata for Faster Clone Detection” (Poster)  
Vaibhav Saini, Farima Farahani, Pedro Martins, Di Yang, and Cristina V. Lopes
- “Stack Overflow in Github: Any Snippets There?” (Poster)  
Di Yang, Pedro Martins, Vaibhav Saini, and Cristina V. Lopes

## **Software Engineering and Analysis Lab (SEAL)**

Prof. **Sam Malek**



The Software Engineering and Analysis Lab (SEAL) is broadly engaged in research to automate the software engineering activities, thereby improving the developer productivity as well as the quality of the resulting software.

- “Automatic Generation of Inter-Component Communication Exploits for Android Applications” (Poster and Demo)  
Joshua Garcia, Mahmoud Hammad, Negar Ghorbani, and Sam Malek
- “DELDroid: Determination and Enforcement of Least-Privilege Architecture in Android” (Poster)  
Mahmoud Hammad, Hamid Bagheri (Univ. of Nebraska-Lincoln), and Sam Malek
- “Advancing Energy Testing in Mobile Applications” (Poster)  
Reyhaneh Jabbarvand and Sam Malek
- “A Taxonomy and Qualitative Comparison of Program Analysis Techniques for Security Assessment of Android Software” (Poster)  
Alireza Sadeghi, Hamid Bagheri (Univ. of Nebraska-Lincoln), Joshua Garcia, and Sam Malek

### Prof. Gloria Mark's Research Group

Prof. Mark's group studies what is known as social computing: studying how individuals, groups, society and technology mutually influence each other. They are particularly interested in studying how information technology use affects multi-tasking, attention, mood, and above all, stress.

- “Digital Footprints: Predicting Personality through Temporal Patterns of Technology Use” (Poster)  
Ted Grover and Gloria Mark

### Software Architecture Research Group (SoftArch) at USC

Prof. Nenad Medvidović

Software Architecture  
Research Group

Prof. Medvidović's Software Architecture Research Group (SoftArch) is a collaborative research team that focuses on cutting edge research in architectural modeling and analysis, component-based development, architecture-based development for distributed, heterogeneous, and resource constrained devices, architecture-based self-adaptation, and event-based middleware technologies.

- “Toward Client-Centric Approaches for Latency Minimization in Mobile Applications” (Poster)  
Yixue Zhao (USC) and Nenad Medvidović (USC)

### Collaboration Research in Action, Design, and Learning Laboratory (CRADL)

Prof. David Redmiles



The Collaboration Research in Action, Design, and Learning Laboratory (CRADL) employs an interdisciplinary approach to research phenomena in human collaborative activity. We primarily study collaborative work, and, particularly, software engineering.

- “End-user Development for the Internet of Things OR How can a light bulb be so complicated?” (Poster)  
Bruno A. Chagas (PUC-Rio and UCI), David F. Redmiles, and Clarisse S. de Souza (PUC-Rio)
- “Designing Effective Synthesized Spatial Audio Cues For The Web” (Poster)  
Tao Wang, Donald Patterson (Westmont College), and David Redmiles
- “Visual Resume: Hiring in the Global Stage Using Card-based Developer Profiles of Online Contribution” (Poster)  
Zhendong Wang and Anita Sarma (Oregon State University)
- “PLAYFUL DRAWING AT WORK: Supporting Affective Expression and Sharing Through Online Drawing to Build Up Distributed Teams” (Poster)  
Mengyao Zhao, Yi Wang (Rochester Institute of Technology), and David Redmiles

### Prof. Walt Scacchi's Research Group

Prof. Scacchi's group studies software engineering and computer games. His two students participating in the Open House are undergraduates whose projects received funding through the UCI Undergraduate Research Opportunities Program (UROP).

- “A New Augmented Reality Interface for Game Based Stroke TeleRehabilitation” (Poster and Demo)  
Arzang Kasiri, Dr. Walt Scacchi (Faculty Mentor)
- “Just-In-Time AR-Based Learning in the Advanced Manufacturing Context” (Poster and Demo)  
Bryce Joe-Kun Tham, Dr. Walt Scacchi (Faculty Mentor)

### Software Design and Collaboration Laboratory

Prof. André van der Hoek

SDCL

The Software Design and Collaboration Laboratory focuses on understanding and advancing the roles of design, collaboration, and education in software development.

- “Understanding the Impact of Support for Iteration on Code Search” (Poster and Demo)  
Lee Martie, Thomas Kwak, and André van der Hoek
- “Geographical Bias in GitHub: Perceptions and Reality” (Poster)  
Ayushi Rastogi, Nachiappan Nagappan (Microsoft Research), André van der Hoek, and Georgios Gousios (Delft University of Technology)

ISR warmly thanks our sponsors:

