

November 10, 2019, New York

Data: Acquisition to Analysis

A SenSys/BuildSys 2019 Workshop

WELCOME!

Thank you for joining us!

A very brief history of DATA

(It's only been around two years... 😊)

- The IPSN'18 business meeting discussed open access & reproducibility...
- Shijia Pan, Pei Zhang, Jie Gao, and Chien-Chun Ni created DATA
 - First run at SenSys 2018
 - Most popular workshop!
 - 14 datasets collected and promoted
 - Conclusion was to do it again
 - Refine criterion and filter for dataset submissions
 - Grow beyond 'just datasets' to understand data issues broadly

The origins and goals of DATA'19

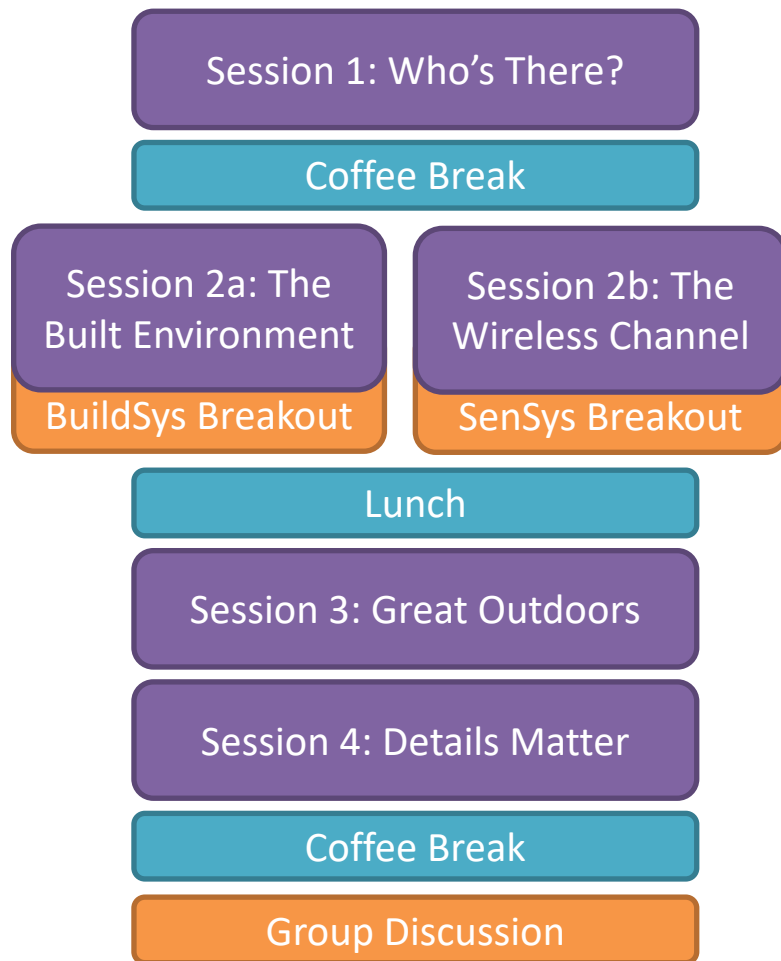
- Introducing the joint SenSys / BuildSys DATA'19 workshop
 - SenSys: Say lots of positive things about a workshop and you end up running it
 - BuildSys: Flora Salim and Mikkel Baun Kjærsgaard were starting a similar one!
 - Long history of these communities working together: join forces!
- The dual goals of DATA:
 - Collect and promote interesting, useful, high fidelity datasets
 - Became Dataset papers (1-2 page short papers)
 - Explore questions around the collection & use of data
 - Became Full papers (3-5 page papers)

DATA'19 by the numbers

- Submissions received 2 reviews, then a 3rd if there was not consensus
 - (Average PC load 2.75 reviews / member)
- 22 submissions total
 - 9 full papers
 - 13 datasets
- Accepted 16 submissions (73%)
 - 5 full papers (56%)
 - 11 datasets (85%)
- Nice growth from DATA'18:
 - Accepted 14/15 (93%) datasets

Today's Agenda

- DATA is at its heart a workshop
 - Discussion is highly valued & encouraged
 - *What brought you to DATA?*
- Session structure
 - Paper presentations
 - Mixture of full papers and datasets
 - “End” discussion
- Breakouts & discussion sections



Early seeding of breakout & group discussion ideas...

- What makes a good dataset?
 - How to communicate, enforce(?) these properties
- How to effectively promote datasets
 - What can DATA do
 - What can we as a community do outside DATA?
- What should the future of the DATA workshop be?
- What datasets are missing that are in demand?
 - How do we promote / reward their collection?
- What is different, the same, about SenSys/BuildSys DATA?

Promotion of DATA and datasets at SenSys!

- SenSys chairs have agreed to give us a little time in the main program
- I will use this time to promote your cool datasets
 - One slide
 - With one “cool graph”
 - With optional “cool phrase”
 - Goal is to hook people, get them to look more
- <https://j2x.us/data19>
- (open to other uses of this time as a breakout topic!)

Thank you to all of our wonderful organizers, committee-members, and volunteers!

- DATA'19 Chairs
 - Shijia Pan, Flora Salim, Mikkel Baun Kjærgaard, and Pat Pannuto
- DATA'19 Advising Committee
 - Jie Gao, Pei Zhang, Prabal Dutta, Jie Liu

Thank you to our wonderful Program Committee!

- Jorge Ortiz, Rutgers University
- Wen Hu, University of New South Wales (UNSW)
- Olga Saukh, TU Graz
- Jun Han, National University of Singapore
- Brano Kusy, CSIRO
- Chris Xiaoxuan, Lu University of Oxford
- Wan Du, University of California, Merced
- Arun Vishwanath, IBM Research Australia
- Clayton Miller, National University of Singapore
- Salvatore Carlucci, NTNU
- Rachel Cardell-Oliver, The University of Western Australia
- Zoltan Nagy, UT Austin
- Mohammad Saiedur Rahaman, RMIT University
- Fisayo Caleb Sangogboye, University of Southern Denmark
- Yongli Ren, RMIT University
- Jason Koh, University of California, San Diego


An enormous thank you to our Dataset Management Chair

- Chien-Chun Ni, Yahoo! Research



Quick Logistical Things

- There's no food or drink allowed in this building
 - So coffee breaks are in the quad in the next building over
 - We'll find them together?
- Restrooms
 - Women's down the hallway through the doors
 - Men's are on floors 1, 3, and 7 (we're on 5)



November 10, 2019, New York

Data: Acquisition to Analysis

A SenSys/BuildSys 2019 Workshop