

Lightning Talk: HydroShare – A Case Study in Software Engineering Best Practices and Culture Change for Developing Sustainable Community Software

Ray Idaszak¹, David G. Tarboton (PI)², Hong Yi¹, Michael Stealey¹, Pabitra Dash², Alva Couch³, Daniel P. Ames⁴, Jeffery S. Horsburgh², Tony Castronova², Jon Goodall⁵, Mohamed Morsy⁵, Venkatesh Merwade⁶, Mauriel Ramirez², Tian Gan², Drew (Zhiyu) Li⁴, Jeff Sadler⁴, Shawn Crawley⁴, Zhaokun Xue³, Lan Zhao⁶, Carol Song⁶, Christina Bandaragoda⁷

¹RENCI, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA, Email: {rayi, hongyi, stealey}@renci.org

²Utah State University, Logan, Utah, USA, Email: {dtarb, pabitra.dash, jeff.horsburgh, tony.castronova, mauriel.ramirez, tian.gan}@usu.edu

³Tufts University, Medford, Massachusetts, USA, Email: {alva.couch, zhaokun.xue}@tufts.edu

⁴Brigham Young University, Provo, Utah, USA, Email: {dan.ames, zhiyu.li, jeffrey.sadler, shawn.crawley}@byu.edu

⁵University of Virginia, Charlottesville, Virginia, USA, Email: {goodall, mmm4dh}@virginia.edu

⁶Purdue University, West Lafayette, Indiana, USA, Email: {venvmerwade, lanzhao, cxsong}@purdue.edu

⁷University of Washington, Seattle, Washington, USA, Email: cband@uw.edu

*4th Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE4)
September 12-14, 2016, University of Manchester, Manchester, U.K.*

About HydroShare

- An online, collaborative system supporting the open sharing of hydrologic data, analytical tools, and computer models (<https://www.hydroshare.org/>)
- Funded under NSF Software Infrastructure for Sustained Innovation program* (<http://nsf.gov/si2>)
- In our 5th year of 2012-2017 five-year award
- Has 10 collaborating organizations
- CUAHSI – the primary U.S. hydrology consortium with 130 member universities and international organizations – positioned to sustain the project

HydroShare Research Software Engineering

- The 10 collaborating organizations put out formally reviewed and tested community code every 2-3 weeks
- The Software Engineering (SWE) HydroShare uses includes
 - Collaborative Development Environment Infrastructure
 - Code Infrastructure
 - Containers; Virtualization (as applied expressly for SWE)
 - Iterative Software Development
 - Testing and Test-Driven Development (TDD)
 - Code Reviews
 - Continuous Integration (CI)
 - Development Operations (DevOps)

Observations

- In fifth year of award, we've had opportunity to see several undergrad, grad students rotate through project
 - As well as see several early career faculty and postdocs onboard to the project
- As a large community project, they accept HydroShare software engineering as the new norm
 - They've not been exposed to any other norm for developing large distributed multi-team academic software
 - They use these SWE best practices for usually no-less than 9 months
 - In stride with their day job or coursework

Recommendations

- We made some missteps along the way and learned from them; here are some of our conclusions that may inform and accelerate other efforts:
 1. Exceptionally committed hands-on PI
 - Lead by example in both science *and* software engineering
 2. “Evidenced-based” Software Engineering for Science
 - E.g. not enough to say you did SWE in industry and now work in academia
 3. Continuous Integration: The good side of technology barriers
 - A good SWE environment blocks SWE poor practices
 4. DevOps!
 - Dial out SWE complexities unless you insist domains scientists become professional software engineers (i.e. SWE best practices is what you’re really after)
 - Large team = large cross-team accountability in uptake of SWE best practices
 5. Not just best of breed technology decisions, but also best for career
 - A great incentive is to choose best technologies that also overlap with career paths