



The government seeks individual input; attendees/participants may provide individual advice only.

Middleware and Grid Interagency Coordination (MAGIC) Meeting Minutes¹

June 2, 2020, 12-2 pm ET

Virtual

Participants

2U Committees/Ministries	Lisa Arafune (CASC)
Alison Derbenwick Miller (Oracle)	Mallory Hinks (NCO)
Arjun Shankar (ORNL)	Miron Livny (OSG)
Dan Gunter (LBL)	Richard Carlson (DOE)
Daniel S Katz (U of Illinois)	Saswata Hier-Majumder (DOE)
David Martin (ANL)	Sharon Broude Geva (U of Michigan)
Donald Petravick (U of Illinois)	Stefan Robila (Montclair State U)
H Birali Runesha (U of Chicago)	Terrill Frantz (Harrisburg University)
Hal Finkel (ANL)	Tevfik Kosar (NSF)
Kate Evans (ONL)	William Pentland (GenBright)
Keith Beattie (LBL)	

Introductions: This meeting was chaired by Richard Carlson (DOE/SC) and Tevfik Kosar (NSF)

Sustainable Software Speaker Series Planning

- Rich and Tevfik described the idea of sustainable software as the next topic for a speaker series.
- Dan Katz gave a summary of recent workshop on this topic
 - He mentioned that the workshop was a little different from the discussion we will have in MAGIC. It was created by a group of people that wrote “Karlskrona Manifesto” on sustainability. That looks at sustainability of software itself as well as how software works in a larger sustainability context. They are looking at a broader question.
 - David Martin asked if there were people from the XSL computing project involved in that?
 - Dan said that there weren’t. He posted a link to the workshop website in the chat: <https://bokss.github.io/bokss2021/>
 - Dan Gunter said that one thing that comes up over and over again in this is that there’s so many different dimensions to software. So, you have to try to scope your discussion to make sure you’re all talking about the same set of metrics and dimensions.
 - Dan Katz said the person who gave the keynote, Pierre Bourque, had an interesting view. Suggested Pierre as a possible MAGIC speaker for this series.

¹ Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program.

- Round Robin: Where do you think the sustainability issue is most in need of discussion?
 - Kate Evans said that coming from the earth sciences, one thing that is near and dear to her heart is reproducibility and confidence in models. We need to maintain models over many years. She talked about staffing difficulties since it isn't a glamorous job. You're not publishing in Nature to say that the model still runs accurately. But it's an important job. We should incentive it better.
 - Keith Beattie echoed what Kate said. He said that establishing the career paths for people who are doing software rather than science. Creating and nurturing a community would go a long way.
 - Sash Hier-Majumder said that Keith's point is really important. He asked: for those interested in scientific software. What defines a good or robust delivery mechanism for your software in your science stream?
 - Keith Beattie: if the software is being used, if there is a growing community in terms of users and developers
 - Dan Katz – if the software is being contributed to by others rather than having others write competing software.
 - Sash said that there is a fundamental challenge in quantifying that impact for software with an emphasis on the scientific side of things. In the computational geophysicist community, there are several different software, and some are embraced by other communities, and some are not. We would like to get a better idea of what's the best mechanism for sustainable delivery of those software that are more widely used.
 - Keith asked if he had any sense as to why certain software products are more embraced in their community? Ease of use? Download?
 - Sash said he thinks it's a combination.
 - Stefan Robila said he one issue is how do you get the community to use different solutions? He also commented that he agrees with Kate. It's difficult to specific career paths for scientific software sustainability.
 - Kate – Oakridge having a discussion on how to evaluate scientists and the organization that's not just writing papers about software. Is it the number of users? How many downloads? How open?
 - Dan Katz shared a list of names in this field – Jim Herbsleb, Kevin Crowston, James Howison, R. Stuart Geiger
 - Dan Katz – Shared a slide based on Kate's was initial comment
 - Arjun – Software for many communities is delivered as large interdependent systems. He has seen firsthand that in the past decade as the architectures have switched, there are large communities that stay entrenched with their CPU code basis and find switching from CPU to GPU hard. "Too big to change". Structuring software in way that it can be changed systematically to accommodate new architectures of using AI that the community should start moving towards.
 - David – There's certainly "too big to fail", but there's also "too small to improve". There may be a code that's important to a community, but there's only 2 or 3 developers. They may be able to maintain but may not have capability or the knowledge to move to a new architecture.
 - Rich asked Alison for the industry perspective.

- Alison – In working with researchers, we encounter this problem of sustainability all the time. This is a huge issue and she thinks it's intertwined with currency of data and being able to use modern data sets with modern software packages and making sure everything is interoperable in a systemic and planned way as opposed to everybody gluing pieces together as they need to.
- Rich asked if there was a consensus that training for workforce would be a valuable topic?
- Dan Katz said that there is certainly training aspect, but it also goes beyond that. He thinks there is a bunch of different workforce aspects: rewarding people, evaluating people, hiring people
- Rich asked for Miron and Terrill's perspective on training students.
- Miron said in his program they don't emphasize this. Most of the faculty don't have experience in it and most of the students are not interested in it because they want to get a job at Oracle and work on the latest, greatest database technology.
- Miron mentioned how programming languages come and go into "fashion".
- Terrill said in the quantum computing community they're really focusing on high school education and engagement. He said that quantum is really about retraining people (engineers, biologists, chemists, etc.) to the quantum programming way of viewing the world.
- Birali said that in his area he sees two aspects.
 - We need to rethink training and education of students and he commented that this new generation learns in different ways than we were trained.
 - There are fewer people interested in the traditional C/C++ type of coding and more towards using different approaches (coding libraries). So, we are losing a core group of skills that are becoming hard to find.
- Birali said that we need to start thinking outside the traditional way of teaching and developing other ways that will supplement what will be expected to gain from the curriculum.
- Keith commented that education needs to be ongoing. He also commented that when we're looking for software developers and engineers, we should move away from requiring a PhD because what you learn in school can be irrelevant within 5 years anyway.
- Alison said that there is a tension between some of the principles with everything being open and free and the lack of "sexiness" in maintaining a code base and making something sustainable. Wondering if one of the challenges to making this an attractive career path for students is that it is perceived as being low value of a career maintaining
- Dan Katz said that there is probably something to Alison's comment about perceived value, but may not be the whole story
- Miron said that at universities there is a fairly flat career structure. The industrial world has more promotions. Academia doesn't have a lot of room for upward mobility.
- Alison said ACM curriculum guidelines factor into a lot of curriculum pathways at universities. She said she didn't think there was anything that addresses research software in any of the ACM curriculum guidelines. Would that be something that would be helpful?
- Dan Katz said that Alison's suggestion would be useful. He did question how much different universities WANTED to be vocational.
- Rich asked if there was a different role for different academic institutions.

- Dan Katz said that Cardiff University has a software academy where they're really trying to train people for jobs like this. He also said that Ed Seidel and Gabrielle Allan (U of Wyoming) are working with students from native reservations in Wyoming to develop software skills.
- Terrill commented that students looking for training for specific jobs has been growing instead of students going to community colleges.
- Stefan said that if ACM puts a stamp on it, it will allow computer science departments to argue for potentially considering a concentration or degree or certificate. It would also allow them to compare what they have in their degrees with what is needed for research compute professional.
- Rich summarized the major topics to have follow on discussions
 - Sustainability of software needs to deal with the fact that hardware changes. Moving to accelerators, AI, quantum, neuromorphic. Dealing with sustainability as the underlying platform underneath you changes.
 - Best practices for maintaining the software, developing it, writing it so that it is sustainable and understandable by someone that you hand it off to.
 - Software languages changing – How to deal with that?
 - Workforce development – Association that would recognize skill set with a certificate or something like that?
 - Specialized software. Writing something for a specific domain vs writing something that's more generic for a database that can be used for many different purposes. What is the sustainability issue? Is it the underlying software or the application use of that?
- Dan Katz said that Academic Data Science Alliance is another group that we could have someone talk to us from. Possible invite the Director (Micaela Parker) as a speaker if we look at career issues.
- Keith said that another topic is funding models. Funding for specific projects may not be the best way to fund sustainable software.
- Donald said he thinks it would be good to find someone running an organization that has a tiered set of careers for software engineers and had them talk about all the other aspects that they had to bring together.
- Miron commented that we didn't talk about software that becomes successful then the community has to pay a fee to use. How do people feel about that? Miron mentioned sustainability in terms of cost. Price goes up over time
- Sash said it might be interesting to bring in someone from ACP to get their thoughts on sustainability. He also suggested bringing in someone from AAU (Tobin Smith) to give insight into how academia thinks of handling the career development issue from a high-level perspective.
- Dan Katz asked about the outcome from the meetings. What will we use their talks for?

Rich suggested that this speaker series would last 4 – 6 months.

Links Shared in the Chat

Links shared:

- <https://bokss.github.io/bokss2021/>
- <https://collegeville.github.io/Workshops/>
- <http://wssspe.researchcomputing.org.uk>
- <https://wosss.org>
- <https://intersect-training.github.io>
- <https://us-rse.org> as a community working in this direction
- <https://danielskatzblog.wordpress.com/2019/07/12/super-rses-combining-research-and-service-in-three-dimensions-of-research-software-engineering/> for more about RSEs
- <https://researchsoftware.org/2020-workshop.html>
- <https://society-rse.org/community/rse-groups/>
- <https://www.cardiff.ac.uk/software-academy>
- <https://academicdatascience.org>
- <https://www.rit.edu/news/rit-professor-lead-international-task-force-shaping-future-computer-science-curriculum>
- <http://urssi.us/blog/2019/02/25/software-incubator-workshop-a-synthesis/>

Speaker suggestions:

- Dan Katz suggested Pierre Bourque.
- Dan Katz suggested Jim Herbsleb, Kevin Crowston, James Howison, R. Stuart Geiger.
- Dan Katz suggested Micaela Parker from ADSA.
- Sash suggested Tobin Smith, AAU, to give insight into how academia thinks of handling the career development issue from a high-level perspective.
- Alison suggested Chris Stephenson at Google chairs the ACM Education Board and would be a good person to speak about developing curriculum guidelines. Mehran Sahami at Stanford would be good as well.
- Dan Katz suggested Rajendra Raj (<https://www.rit.edu/news/rit-professor-lead-international-task-force-shaping-future-computer-science-curriculum>) Alison agreed with the suggestion. She mentioned Raj was appointed to that role by Chris and Mehran
- Dan suggested Chris Holdgraf from UC Berkeley would be a good speaker on this. He's spun up <http://2i2c.org> to address this.
- Dan suggested Mike Heroux would be a good ECP person in this area. They're now trying to figure out what material should be taught to potential RSEs and what already exists and what needs to be developed
- Kate agreed about Heroux, also suggested Lorena Barba
- Dan commented that Lorena's not part of ECP, but she's a good speaker on this topic and reproducibility

Next Meeting

July 7 (12 pm ET)