



DASH 2023 MEETING REPORT

Jon Vandegriff¹, Alex Antunes¹, Brian Thomas², Julie Barnum³, Jack Ireland², Arnaud Masson⁴, R. T. James McAteer⁵

1: Johns Hopkins Applied Physics Lab, 2: Goddard Space Flight Center, 3: Laboratory for Atmospheric and Space Physics, 4: European Space Astronomy Center, 5: New Mexico State University

DOI: 10.5281/zenodo.10668517

EXECUTIVE SUMMARY

The Data, Analysis and Software in Heliophysics (DASH) conference is the first meeting dedicated to software and data practitioners in the arena of space physics and Heliophysics. Strong attendance at the inaugural 2023 meeting indicates that there is indeed a desire in the community for more collaboration and connections. 150 people from 22 countries participated in a 2.5-day meeting and shared 85 presentations. An unconference session allowed for extended discussion on topics decided by attendees at the meeting. The original intent of the meeting to provide a more interactive conference experience was largely met, and there was substantial involvement from early career participants. Plans are underway for a 2024 DASH meeting in Madrid, Spain, which helps maintain and emphasize the international intent of this meeting.

MEETING PURPOSE

The Data, Analysis and Software in Heliophysics meeting was initiated as a way to bring together people in Heliophysics and the wider space physics community for a conference dedicated to software, algorithms, data science, metadata and data management. While the term Heliophysics may refer to a specific science division in NASA, this meeting is intended to cover not just NASA and not just space-based measurements, but ground-based studies as well as models and model outputs. It is also intended to be international, and was purposely held adjacent in time to the International Heliophysics Data Environment Alliance (IHDEA) meeting.

Prior to DASH, there has not been a dedicated meeting for this community. Some meetings have an informatics component, such as the AGU, which has a very broad science focus. Earth science and planetary science have some informatics-themed meetings (such as the Planetary Data Workshop and the Planetary Science Informatics and Data Analytics meetings). DASH was patterned most closely after the well-established Astronomical Data Analysis Software & Systems (ADASS) meeting, which covers informatics topics in astronomy and has been an annual event since 1991.

Because the Heliophysics community has never had its own meeting, many people in the field are not trained in running conferences or conference sessions. Therefore, one focus of the meeting was to engage early career

practitioners to help establish a pool of people who are capable of carrying forward future meetings. The call for sessions specifically asked for contributions from early career scientists and engineers.

We also desired for the meeting to be highly interactive, and to avoid the tendency for presentations to just continue as one slide after the other, with little audience participation. Session chairs were encouraged to utilize a format for their session to encourage interaction, and de-emphasize passive presentation of slides.

EVENT PLANNING

The planning committee for the event consisted of nine people, the co-authors of this report. We created a purpose statement for the meeting, a code of conduct, and a set of topics deemed relevant for the meeting. For specific session topics, we solicited input from the community, with the idea that a person who proposes a session topic becomes the session chair if that topic is accepted for the meeting. We created a simple web site and a flier, with instructions on how to submit a session topic. This setup work was completed just before the Fall 2022 AGU, so that the fliers and the session submission instructions were distributed at the AGU meeting.

The location for the event was set at the Johns Hopkins Applied Physics Lab (APL), and the local hosts for the event applied for and received NASA funding to support the planning and execution activities. The Laboratory for Atmospheric and Space Physics (LASP) applied for and received National Science Foundation (NSF) funds to support travel grants for early career and student attendees.

The non-APL organizing committee was volunteer. Organizers met roughly weekly, starting in November 2022, with breaks after milestones, and when we were waiting for community submissions and registrations. There was a burst of organizing activity for session proposal review and also for coordinating session layout and abstracts placement with the session chairs. Session chairs were responsible for selecting talks or moving a submission to a poster.

The organizing committee utilized a Google drive folder for tracking tasks, planning the schedule, and for writing shared documents. Session chairs also were then given write access to this folder. A public sub-folder was created for participant input during the meeting. APL used conference software called EventLink for registration.

From session proposals received from the community, nine topics were selected and placed in the schedule, which also included an opening slot for three plenary talks. On Day 2, there was an un-conference session, where the topic for that session was to be decided at the meeting.

EVENT EXECUTION

The meeting was held October 9-11, 2023 as a hybrid (in-person and online) event, using the Zoom meeting capability. The audio-visual (AV) setup for the room and for the Zoom session was provided by professional AV staff at APL. The meeting room is equipped with cameras for showing the speaker at the podium, and also side cameras that the AV staff can direct at individuals in the audience when they are asking a question. One conference organizer was designated to monitor the Zoom chat and inject online questions into the discussion or offer online speakers a chance to unmute and ask questions directly.

Here is the resulting agenda as executed:

Day 1 (Monday) - Focus on "Data"

- Opening Plenary talks- Perspectives on the DASH Domain
- Cloud Computing for Heliophysics
- Modern Data Pipeline Strategies
- Data Access Interfaces and Services

Day 2 (Tuesday) - Focus on "Analysis"

- Optimizing ML SWx Modeling and Forecasting through Improved Data Sets and Metrics
- Computational notebooks in heliophysics

- Interoperability between Python Packages (PyHC)
- Unconference Topics (chosen at the meeting)

Day 3 (Wednesday) - Focus on "Software"

- Moving to all open-source in Heliophysics
- Advancing Data Portals for Open Heliophysics Research
- Methods and Software Tools for Forward Synthesis and Inverting Plasma and Magnetic Information in the Solar Corona

We also sponsored an informal outing to the National Cryptologic Museum, a free museum on the history of cryptography in the US that is maintained by and adjacent to the National Security Agency in Fort Meade, MD, which is about 20 minutes from APL.

The three plenary talks were:

Navigating Through the Maze of Heliophysics Data Diversity, **Laura Hayes**
 VSO / DKIST Perspectives on the DASH Environment, **Alisdair Davey**
 Open Science Challenges in Heliophysics, **Baptiste Cecconi**

The specific talk titles within each session are not listed here, but are available on the web site for DASH:

<https://dash.heliophysics.net/2023/> (the main site being <https://dash.heliophysics.net/>)

This site is intended to be used to hold persistent DASH information, and future event hosts can create their own event-specific registration and meeting management website.

DASH participants submitted additional session topics (using an online tool called Slido) for un-conference Sessions. Submissions to this started a few days before the meeting, and then voting concluded on the end of the first day of the meeting. These topics were selected and held as concurrent sessions in separate breakout rooms:

- **Broadening Participation in Open-Source Projects** — building community software carpentry skills & documentation practices
- **Event List Sharing (is caring)** – common ways to advertise, find, format and use list of events
- **What's Out There?** Ideas/methods for collecting links to all open-source helio-related data tools, codes, databases, event lists, etc., so tools and data are easily accessible? Representing information in Heliophysics, linking datasets, knowledge graphs

We also took poster submissions on any topic, not just those related to a specific session topic. Posters were up for the duration of the meeting at the sides of the main meeting room. Each break was at least 30 minutes and was also designed to be a time for viewing the poster presentations. There was no online component to the posters, since the organizers' own experiences with online poster sessions has shown them to not be useful.

ATTENDANCE

A total of 211 people registered for the meeting, with 73 registering for online-only participation. The meeting room holds up to 150 people, with 120 being a more comfortable, realistic limit for a multi-day meeting.

Actual attendance was about 100 people in-person (day 1: 101, day 2: 100, day 3: 81), and about 40 online each day. The online participant count was harder to track, since some in-person people also logged into the Zoom session. Actual unique Zoom attended each day was 89, 89, 63, with a total of 137 unique names logged into Zoom over the three days.

There was significant international participation, with 22 countries represented by participant address and 34 by nationality.

Figures 1 and 2 show in-person and virtual participants at the meeting.

OUTCOMES

There were 55 oral presentations and about 30 posters presented. There was plenty of poster space, and oral presenters were also given the option of also putting up a poster, although most did not since that takes extra time.

A public Google document was made available to attendees during the meeting to take collective notes, and people generated 46 pages of comments covering every session and nearly every talk at the meeting. This document was switched to read-only after the event, and is available here:

https://docs.google.com/document/d/1m2onbgd-BpJyQMxh4e3cw8yMi_GWiTq9zdsu6TcxZfM/edit?usp=sharing

A ChatGPT summary of this document was generated by a participant and published at Zenodo:

<https://doi.org/10.5281/zenodo.8436109>



Figure 1 - The conference photo shows many of the in-person attendees at DASH 2023 at APL.

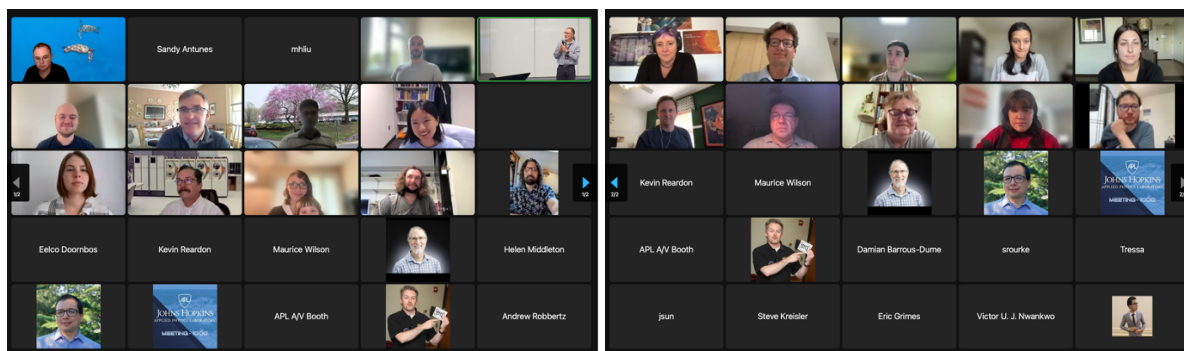


Figure 2 – This screen capture shows some of the online participants from DASH 2023.

Participants were required to submit their materials to a Zenodo collection the week before the meeting. The entirety of the collection was made public after the meeting.

<https://zenodo.org/communities/dash2023>

The Zoom meeting was recorded, and Ar. Antunes on the organizing committee split up the video files by presentation and created a table with links to each presentation, both the Zenodo entry and the video file. These are available on the permanent DASH website: <http://dash.heliophysics.net/>

A post-event survey was sent to participants. The response was light with just 18 people answering. The survey results are presented in Appendix A. The responses were generally very positive and indicate that people thought the meeting was worthwhile.

A presentation about DASH was given at the 2023 Fall AGU Meeting in San Francisco.
<https://zenodo.org/doi/10.5281/zenodo.10602362>

LESSONS LEARNED

Most of the conference organizers were new to the process of organizing a meeting of this scale, so there are many points where we learned as we were doing. Here are some points about things that worked well and some things that could be improved in terms of organizing and running the meeting.

Because the number of registrants exceeded expectations, the system we used for collecting abstracts and presentations exceeded our simple spreadsheet-based approach. APL has conference attendance software, but it is very manual in terms of setting it up for abstract submissions. WE also ended up having people email in presentation materials ahead of time, so those all had to be manually received and put into the right folders. We did require presentations to be given ahead of time, so that the AV staff could manage the slides on the screen

The interactive nature of the meeting worked very well. There was plenty of time for discussion. With 100 in-person attendees, there was usually a chance for people to ask their question or air their opinion. The usual occurrence of more forceful people (usually more senior career-wise) getting more air time did happen, but there was at least a chance for people to speak up if they wanted to.

Having the AV staff operate the slides (which then had to be submitted beforehand) was very helpful in freeing session chairs to ignore logistics and focus on the session content.

Rather than have fixed session lengths in the schedule, it would have been better to wait until we saw the number of submissions to the sessions before assigning times to each session. This is obvious in hindsight, but worth noting for future organizers.

With a free Slido account, there is a limit on the length of time a poll can be open. We would have liked to open the unconference topic poll much earlier – maybe a week or even two before the meeting started. Paid accounts are not that much, so we should have just gotten one. APL has a Slido account, but getting authority to use it takes a while, and we did not realize that in time to take advantage of it.

APL requires any visitors, even those “attending” via Zoom, to be registered ahead of time due to sponsor requirements for vetting of non-US citizens. There were many people who registered very late, and this caused a significant strain on event planning staff, who tried very hard to accommodate people’s request. There needs to be a hard cutoff on registration, or else event staff need to be prepared for a lot of last-minute registrations, or else the venue needs to accept on-site registration.

It may have been better to offer a dedicated poster time, rather than just merge the poster sessions with the breaks.

Also, in terms of timing for the abstract submissions, it would have been better to have an earlier deadline to give session chairs more time to sort out their sessions. Also, conferencing software that allows for the management of abstracts (acceptance, assignment to sessions, transfer to other sessions) would have been useful. Communications to participants about their submissions became very time consuming. Session chairs had some responsibility for this, but they did not have access to the full list of all submissions.

Requiring people to submit their materials first to a Zenodo collection, and then pulling things from there ensured that a large fraction of the presentations did end up archived at Zenodo. It would have been good to require a specific

file naming convention to simplify the task of organizing the submissions by session and author. This had to be done manually by one of the organizers.

For future meetings, the organizing committee will consider offering a way to get submissions published as official proceedings. For the 2024 meeting, we are likely to continue with the Zenodo collection.

PLANS FOR FUTURE DASH MEETINGS

The intent is for DASH to become an annual event, and in order to emphasize the need for international collaboration across Heliophysics software and science, the meetings science and software.

Planning for the next event is underway, and it will be held in Madrid, Spain, at the European Space Astronomy Center (ESAC) on October 14-16, 2024. In the following year, there is a strong desire to hold DASH in San Antonio, TX, USA at Southwest Research Institute. The 2024 meeting will again be adjacent to an IHDEA meeting, and the intent is to continue with that tradition.

The DASH website will be modified to present content from each year's meeting. Persistent information about DASH will be available there, as well as pointers to registration and submission details for the upcoming DASH meeting. We are looking forward to many years of sustaining and connecting the Heliophysics informatics community via DASH.

APPENDIX A

The following pages are the survey questions and results for the post-attendee survey, which were gathered using Google forms.

DASH 2023 Attendee Survey

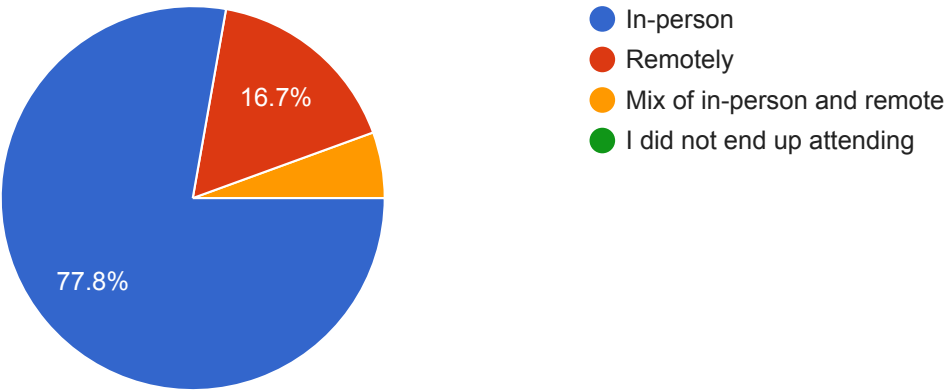
18 responses

[Publish analytics](#)

I attended the DASH meeting

 [Copy](#)

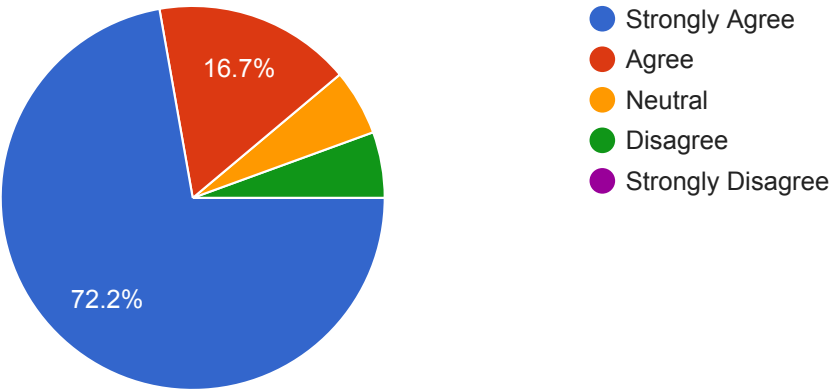
18 responses



I heard about the DASH meeting with plenty of time to consider and plan attending

 [Copy](#)

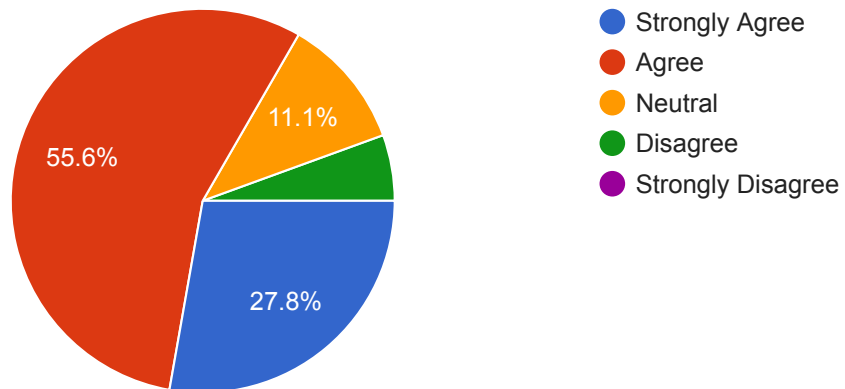
18 responses



I found out about the DASH detailed schedule in a reasonable time.

 Copy

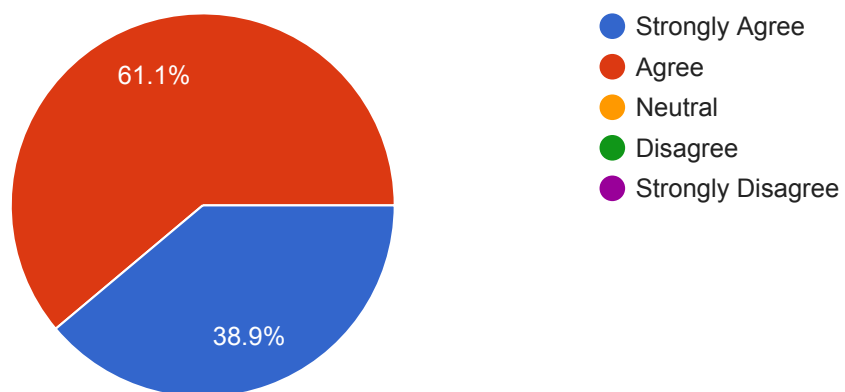
18 responses



The registration process was straightforward.

 Copy

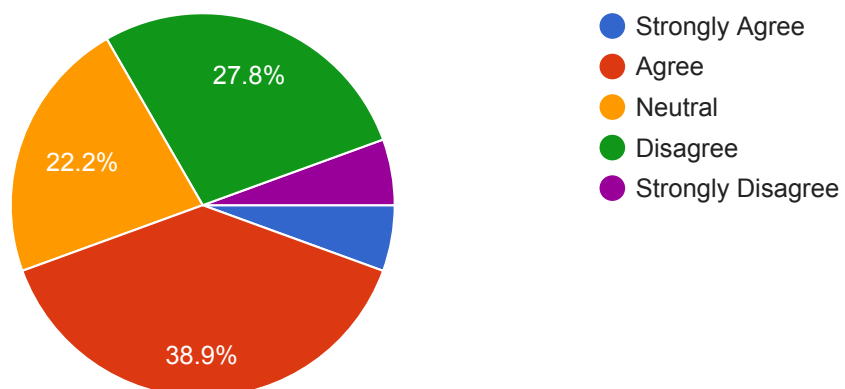
18 responses



For me the attend this year, it was important that there was no registration fee.

 Copy

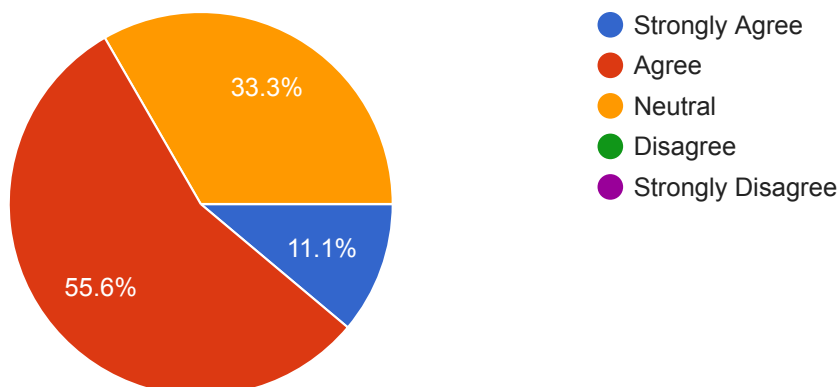
18 responses



In order to attend future DASH meetings, I would be willing to pay a typical conference registration fee.

 Copy

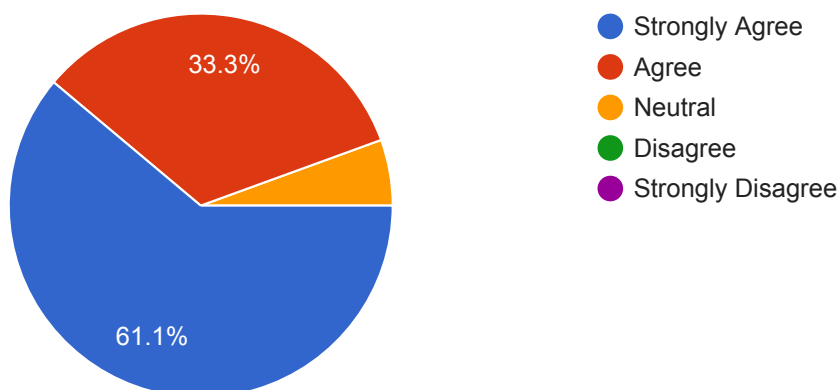
18 responses



The material presented at the meeting was relevant for me.

 Copy

18 responses



I would also like to have heard about these topics:

3 responses

logistical and technical project management issues like GitHub project/Kanban

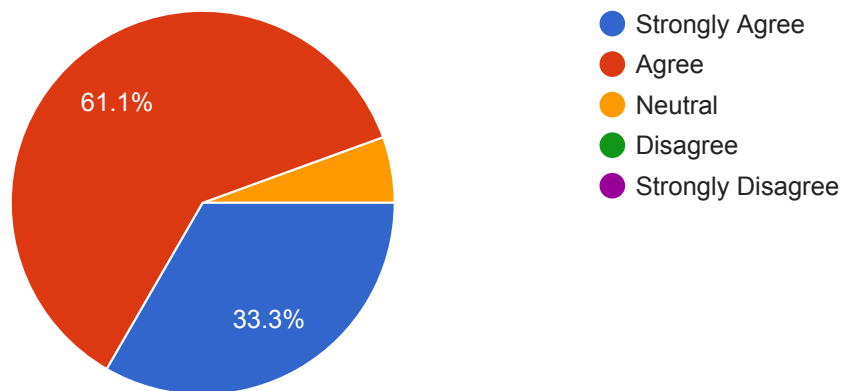
Spack and other package management tools used in HPC and DOE environments.

software engineering practices

The material presented at the meeting was at the right level for me.

 Copy

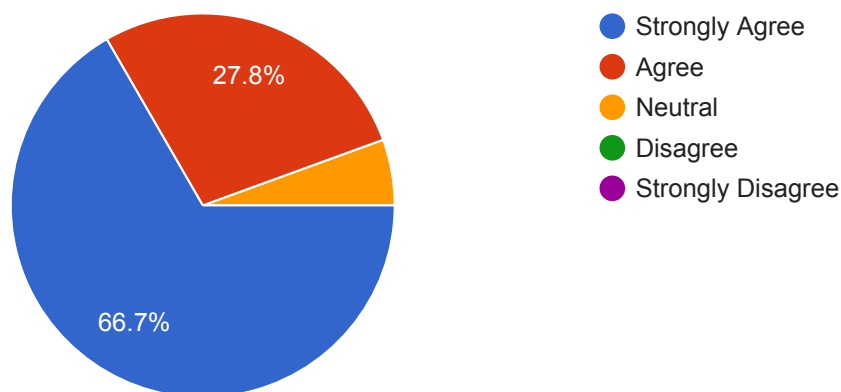
18 responses



The audio visual quality of the meeting was effective.

 Copy

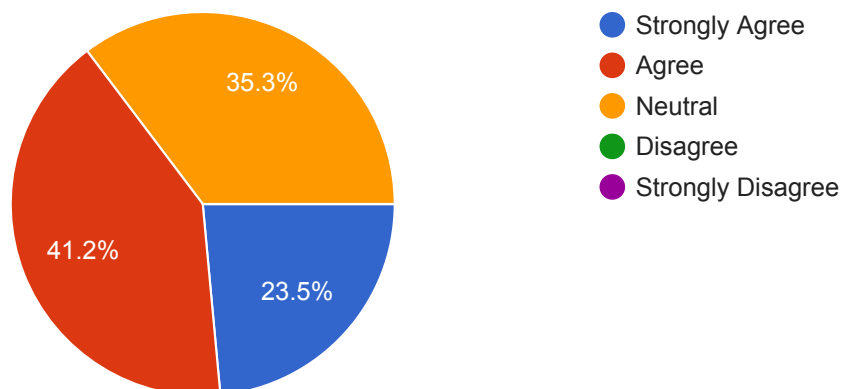
18 responses



Having the posters available during every break was effective.

 Copy

17 responses



I have these comments about the poster sessions:

5 responses

I wish there were even more.

It was good to have the posters available throughout all the breaks. But, it might be nice to also have at least one segment of time called out for all the poster presenters to be at their posters.

It didn't seem like many people went to the posters

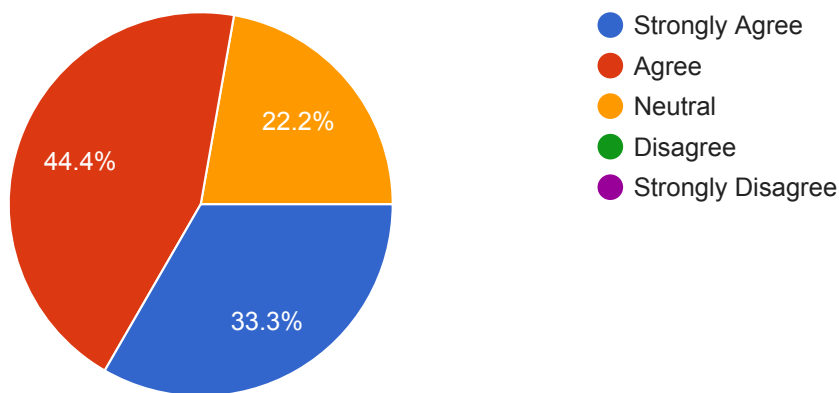
There could be a dedicated time for posters, apart from the break.

Although it's nice to have the posters hanging all the time, I feel that explicit poster attending times (similar as at AGU) might be helpful.

I liked having un-conference sessions (informal sessions whose topics we selected at the meeting).

 Copy

18 responses



I have these comments about the un-conference sessions:

2 responses

I found the unconference session one of the most impactful for me.

I think at least two of our topics were overly broad and somewhat overlapping. I feel like our session started with a roomful of people who didn't necessarily come into the room with the same understanding of what the main thrust was going to be. The discussion was interesting, but, I think we could have benefitted from doing the unconference polling in the week leading up to the meeting and also allowing some of the lower voted topics to simply not get lumped into other topics.

I have these comments about the regular sessions:

2 responses

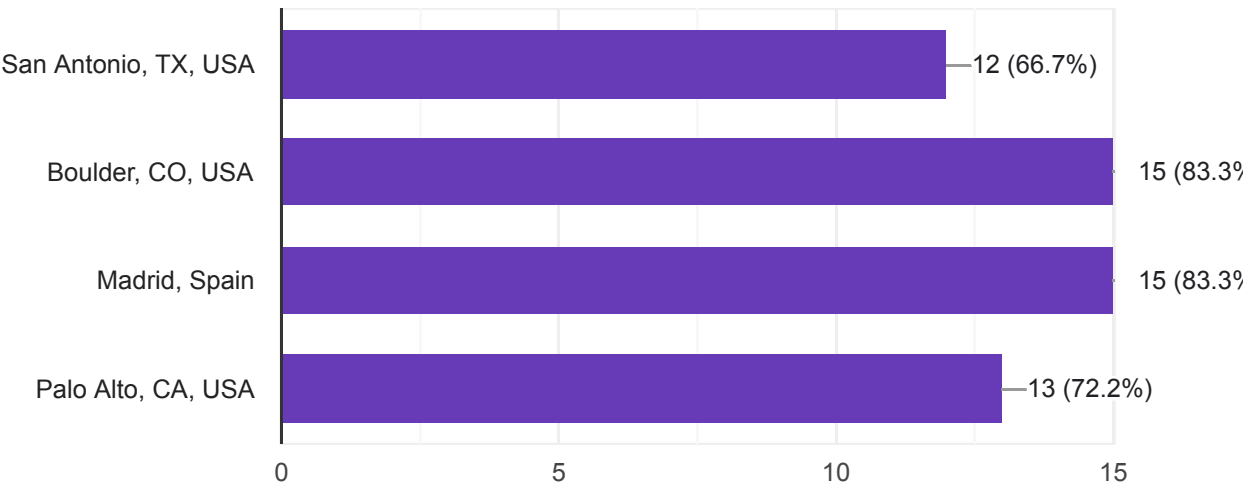
Keynote file support!

I'm not entirely sure that the distinction between the "Analysis" and "Software" day made sense. Also, as a software developer, I would have appreciated the "Software" day not be entirely split sessions.

I would attend DASH next year if it were held in (check all that apply; note that no one has definitively agreed to any of these yet!)



18 responses



Overall, I enjoyed the meeting.

 Copy

18 responses

