ALL CONFERENCE TALKS HAD CAPTIONS: THE 2021 DPS CONFERENCE.

J. R. Masiero[1], J. L. Piatek[2], J. H. Roberts[3], D. Takir[4]; [1]Caltech/IPAC (MR-100, 1200 E. California Blvd, Pasadena, CA 91125, USA, jmasiero@ipac.caltech.edu), [2]Dept. of Geological Sciences, Central Connecticut State University, New Britain CT 06050, USA, [3]Johns Hopkins Applied Physics Laboratory, 11100 Johns Hopkins Rd., M/S 200-W230, Laurel, MD 20723. [4]Jacobs, NASA Johnson Space Center, Houston TX 77058, USA.

Introduction: The decision in May 2021 to make the Oct 2021 Division of Planetary Science Meeting entirely virtual allowed for enough time for the Virtual Organizing Committee (VOC) to consider accessibility concerns for the various components of the meeting. For the 2020 meeting, captioning was optional, and relied on the talk author to create and upload the caption files with their talk. Because of this added burden on authors, combined with the shorter time to prepare due to the impact and evolution of the pandemic, there were only a few talks in the end that included captions. One of the main points of feedback from the virtual 2020 DPS meeting was appreciation for those pre-recorded talks that take advantage of the option to include captions. To expand the accessibility of the talks for the 2021 meeting, and following the recommendations of Piatek, et al. (2021), the VOC decided that ADA-compliant captioning would be provided for all pre-recorded talks, and that live AI-generated captioning would be provided for all live talks, Q&A, and discussion sessions.

Experience: The VOC determined that the best way of ensuring that all pre-recorded talks would receive captions was to use an external captioning service instead of relying on author contributions. This also had the added benefit of simplifying the process of moving the talks from the AAS upload system into the video repository, eliminating the need to match the closed caption files with the associated videos and check them for accuracy, substantially reducing the workload. After reviewing a number of vendors, for the 2021 DPS the VOC selected rev.com to provide captions for the pre-recorded talks. Rev.com uses a two-stage captioning method, first generating auto-captions with an AI program, and then having human reviewers edit that output for accuracy to bring the captions to FCC, ADA, and Section 508 compliance.

One of the key benefits of choosing rev.com was that this vendor is a partner with Vimeo, which was the video platform being used by the AAS for hosting the talks. This meant that once the talks were uploaded, all that was needed was for the two accounts to be linked and then rev.com could access the videos, generate captions, and push the captions back to Vimeo with minimal user interaction. This was a major cost-saving step in this process.

The cost of captioning was based on the length of the videos, with each captioning process from the vendor costing $1.25 per minute of video. Charges were only applied to videos where the full captioning process was successful. Additional costs related to conference IT support (e.g. setting up accounts, linking talks, validating captions were pushed back to Vimeo, etc.) increased the costs by a few dollars per video. By estimating the number of submitted videos and the allotted length for contributed talks and invited talks, the VOC was able to estimate a captioning budget that wound up being very close to the final cost to the conference.

Captioning for live sessions was also provided by rev.com using a Zoom captioning plug-in license. These captions had lower accuracy and a slight delay from the words being spoken, but captured the general conversation with a reasonable accuracy. The cost for this was $20 per Zoom session, representing a minimal impact on the overall conference budget. The final cost for all captioning-related efforts wound up being les than $10 per registered attendee.

The timeline for captioning the talks was a few days. Rev.com uses a roster of independent contractors to do the caption editing, and they select jobs as they are able to. For the DPS conference, however, the majority of talks were around 7 mins long, and so these were quickly processed. Longer talks (~1 hour+) required longer lead times as fewer editors choose these jobs. Submission deadline for the pre-recorded talks was set at t3 weeks before the conference by the AAS, and this was adequate lead time to ingest, upload, and caption the talks (including time for caption corrections), while providing sufficient time to check that the captioning was correct.

Not all captioning for the pre-recorded talks was successful. Approximately 1% of the submitted videos failed to be fully captioned, and were bounced back to the VOC with only the AI-generated auto-captions. Reasons for failure included heavily accented speech, mumbled dictation, unusual jargon, or all of the above. These videos had to have the auto-captions manually edited by the VOC to bring them in line with the captions being provided for the other videos, which was an unanticipated level of effort. Rev.com does offer a caption editing tool which simplified this process greatly. The timeline for editing these videos by a non-professional caption editor was of order a half hour per 7-minute talk, depending on the quality of the AI captions.

Suggested Improvements for Future Captioning: Overall the process of captioning was very smooth, and produced an important product for increasing talk accessibility. Costs for captioning worked out to be a few dollars per registration, and so did not have a significant impact on meeting cost. The only real problem encountered was the videos that failed to be captioned, which put an unexpected work load on the VOC. One possible mitigation of this is to use the optional dictionary upload that pre-seeds the AI with likely words. A list of terms drawn from the written abstracts, or from papers in the field, would likely contain sufficient information to mitigate some of the failures. This listing would either need to be a single broad document lowering overall effectiveness, or individual documents for each talk, though that would increase the overall effort needed to match documents to video files when sending in the files for captioning, and drive up associated IT costs.

For other talks that failed captioning, however, a combination of accent, speaking style, and poor microphone limited the captioning effectiveness. Instructions to speakers that include "best practices" for improving clarity of their recordings (e.g. using an external microphone or headset, avoiding spaces with background noise or that generate echoes) and encouraging them to speak slowly and clearly might alleviate some of this, but it is unlikely to solve the entire problem. Instead, VOCs should plan on a small number of talks needing individual attention.

Conclusions: Our experience with the 2021 DPS Fall Meeting demonstrates that while captions for pre-recorded talks are not trivial to generate, they can be added with minimal financial outlay and effort. Captions provide important support to a range of members of our community, increasing the scientific utility and accessibility of talks. The benefits of captioning to our community outweigh these costs, and we continue to recommend their use at future conferences.

References:
Piatek, J.L., Brooks, S.M., Masiero, J.R., et al., 2021, "All Conference Talks Need Captions", 52nd LPSC, id.2723.