Study of multimedia Input/Output documents and AI tools in Ecma
(this is for info and discussion in TG5, but the Ecma ExeCom and GA are the decision takers on such subjects)

Istvàn Sebestyén
July 2024

1) Pre-planed (usually several months apart) Face-to-Face standards meetings in at the location of the members (included travel)

2) Documentation of the meeting:
   Input documentation: Text – prepared by members and distributed by SDO in advance
   Output documentation: Text – prepared by body and SDO officials after the meeting and distributed by SDO

3) Distribution of input/output: via Mail

- Uninterrupted daily standardization working via email groups / reflectors between meetings is typical. Use of many additional tools (e.g. ES-wiki, ES-discuss, TC39 GitHub besides classical Ecma TC39 Email and File Server). Fragmented input / output from TC39 work. Archival difficulties.

- Meetings still dominantly Face-to-Face with travels and full use of WLAN and Internet/Web.

- Starting use of Web computer conferencing based meetings (like Webex, Zoom, etc.) either "remote only" or with "mixed" (local and remote) participation

- External pressure of saving time, costs, environmental conservation / climate change issues

- Many SDO policies/procedures were/are still lagging behind. All procedures and policies behaved like in the old paper based standardization environment (neglecting the new type of working). Very few policies emerged to reflect new reality

- In 2020 Face-to-Face meetings stopped completely for about 3 years. Slow recovery afterwards but it will never come back fully.

- Generally, homeworking….Therefore, for many groups, like TC39, uninterrupted daily standardization working via email groups / reflectors between meetings continued and went well. Several different Web computer conferencing based meetings were used (like Webex, Zoom, Teams, Google Meets etc.) all with "remote only" participation.

- External pressure of saving time, costs, environmental conservation / climate change issues became even stronger

- In major SDOs (like ISO, ITU, office work completely stopped). Many SDO policies/procedures were/are still lagging behind. All procedures and policies behaved like in the old paper based standardization environment (neglecting the new type of working).

- Starting in 2023 Face-to-Face meetings started. For TC39 out of 6 meetings, 3 "mixed" and 3 "remote". Quality of "mixed" better (audio, better transparency over local participation).

- Work: Mixed homeworking and local….

- For many groups, like TC39, uninterrupted daily standardization working via email groups / reflectors between meetings continued and went well. Several different Web computer conferencing based meetings were used (like Webex, Zoom, Teams, Google Meets etc.) all with "remote only" participation.

- In major SDOs (like ISO, ITU, office work came back). Slow recovery. Many SDO policies/procedures were/are still lagging behind. Few procedures and policies started to emergence on remote and mixed meetings.

- **Besides "classical" text also Multimedia (audio, video, text, data) based contributions both for input (e.g. contributions) and output (e.g. reports)**

- As tools we can expect the strong use of AI based tools, such as:

  - *transcription (multimedia -→ text)*

  - *Summary of contributions / discussions*

  - *programming / coding*

- Standardization will be a continuous effort using various parallel tools, like GitHub type platforms, ICQ, conferencing platforms. Challenge for a clean archival of significant standardization steps among all these.

In several Ecma TCs (e.g. TC39, TC31, TC55…), the ExeCom and the GA we have it already for several years:

- The earlier full "text" based contributions are substituted by "slide only" contributions. Those are pre-published before the meeting. The narrative audio-video comes just at the meeting, but it is not captured. So Ecma does not have a fully captured "multimedia presentation", it has no contribution number, not archived (e.g. as .MP4).

- This is a problem for an SDO, that has archival requirements (a WTO criteria…)

- Remedy with "mixed" results: e.g. TC39 "Technical Notes"….

- Real-time captioning (transcription) of "multimedia" (slides, program code, audio, video) presentation and conversion into "text" based Ecma document. Very difficult, imperfect, heroic work.

- Alternative would be better to recognize and introduce an Ecma multimedia format document (e.g. TC39/2024/xyz.mp4). That could be either pre-published for a meeting, or after the meeting with screen capturing it is easy to publish it as TC39 document without much delay.

- Significant advantage: The emerging AI tools (e.g. whisper, GPT, etc.) work already now quite well on multimedia files.

Many things, but of course further experiments would be needed:

- Automatic transcription from audio presentation to text

- Automatic insertion of slides into the transcribed text

- Generation of summaries of the multimedia presentation

- Summary of discussions

- Summary of discussion conclusions as next steps

Of course AI based tools are "tools" that human must use with knowledge and care, they are not autonomic.

- Start testing true multimedia based input; like preparation of multimedia contributions in video in MP4 e.g. 2 weeks before the meeting and distribution. (So, not only 4 slides just before the meeting and the multimedia presentation – single time – during the meeting, and 3 people try to capture and transpose that into text during the meeting.). Actually, Typical MP4 files with multimedia presentation (using e.g. H.265 video compression are not that large, much is static video information, storage is not a problem)

- Start testing using AI tools to generate from the multimedia mp4 summary and transcript for the meeting (so not after the meeting…)

- Start testing use output documents from the meeting using the above elements.