Brainstorming session 1 Group 2

Transcript

00:00:03 Speaker 1

So, we were one of the largest groups here and we described these datasets in certain terms, that these are usually some niche music genres and styles, like traditional music, or different sub-styles of electronic music, experimental music and ambient music.

They are performed by traditional instruments, or some self-made instruments collected by researchers for archives in expeditions or sometimes in live performances. And from this comes also data quality issues, because, in archives, sometimes, for instance, the recording can be preserved on a scroll or on a back scroll and they can be great in style or if it's live performance collected by someone and then uploaded online again, there can be noise in the recording, clapping and whatever; they can be submitted by artists themselves.

00:01:04 Speaker 1

And, another issue that was commented on: with sharing these data sets comes the requirement that, it is very important to have consent of the person that the recording was made of; and someone suggested that this consent could also be retractable, so you could have a person giving consent and then giving them possibility to also retract the consent depending on how the data set is used.

00:01:37 Speaker 1

Then the qualities of these data sets that they can use: non-western instruments and non-western music systems such as non-diatonic scales, microtonality, and even non-rhythmic music, like some electronic music that not have any beat to it or any repeatable pattern, just the texture and this ambient sort of soundscape-type environment.

00:02:10 Speaker 1

The dataset could include other metadata, as was already mentioned by Rebecca as well, like gesture and dance.

00:02:20 Speaker 1

And yeah, the challenges and recommendations that people suggested were that retractable consent and that the output from this dataset could be much more interesting than the homogeneous large pop music data set, because it is a small, curated dataset

which can have better annotations as well, because if it's collected for specific purpose then often times it has these better annotations.

00:03:12 Speaker 1

And that sometimes with a smaller data set you might not be able to use it for end-to-end composition or end-to-end generation, but you can use it with specific techniques like timber transport or diffusion techniques, where you take some part of that recording, like, for instance, the timber of the traditional instrument that you recorded, and you apply it to a model that can compose music, or maybe you put together some little chunks of music using diffusion.

00:03:51 Speaker 1

Did I forget something?

00:03:51 Speaker 2

Great. Thank you very much.