UE MLPC 2025: DATA EXPLORATION TASK

Abdalaziz Adil Awad Ayoub

Abdulkarim Al Jamal

Beibarys Abissatov

Jeronim Bašić

JɤU
JOHANNES KEPLER
UNIVERSITY LINZ

Annotation Quality
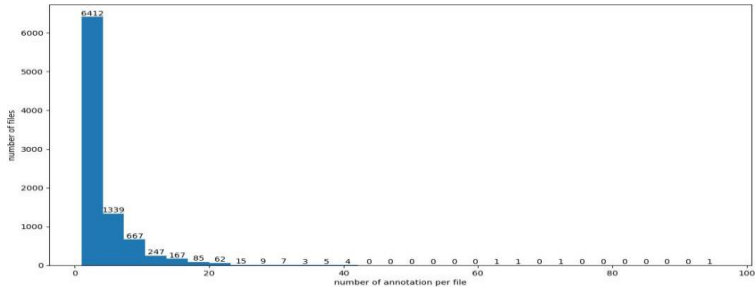
For audio recordings annotated by multiple people:

■ *How precise are the temporal annotations?*

61.5% for the 725 files with 2 annotators (446 files had min. 1 overlapping region within 100 ms threshold). File 560530.mp3 had 20 overlapping regions
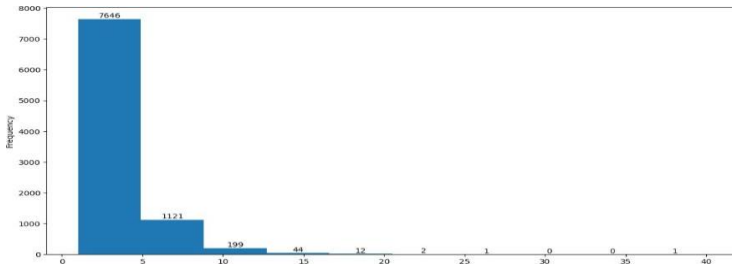
■ *How similar are the text annotations that correspond to the same region?*

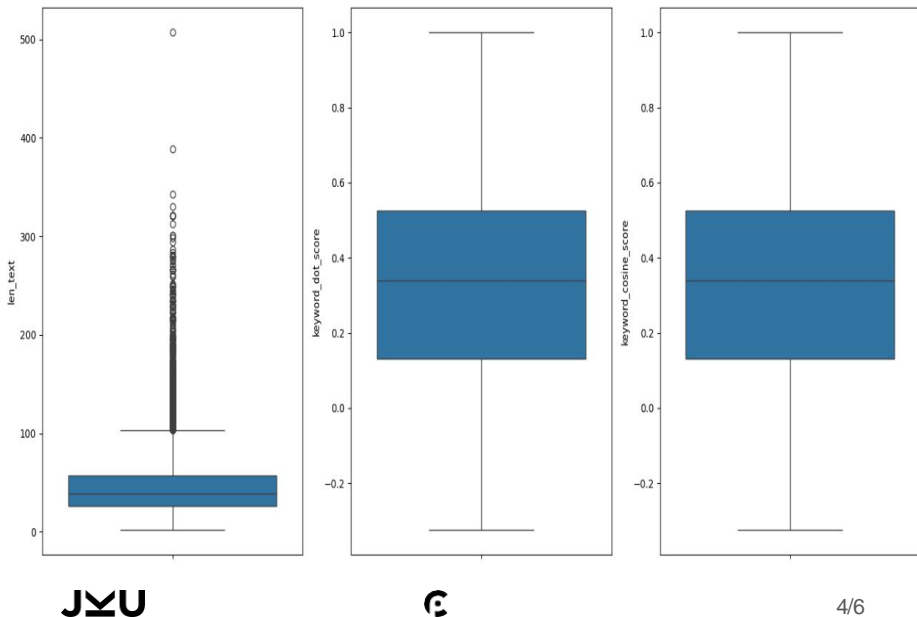**Mean cosine similarity of 42.7% (for two annotators) and 34.5% (for three annotators)**

JⵉU

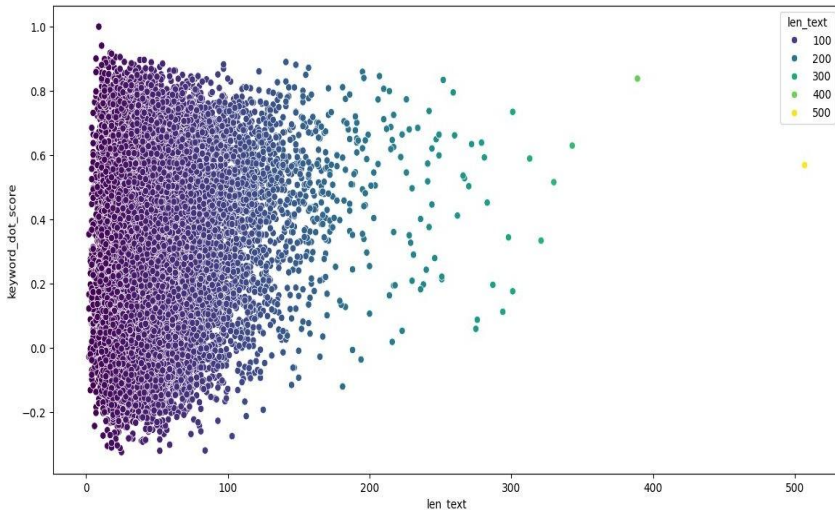- *How many annotations did we collect per file?*

- *How many distinct sound events per file?*
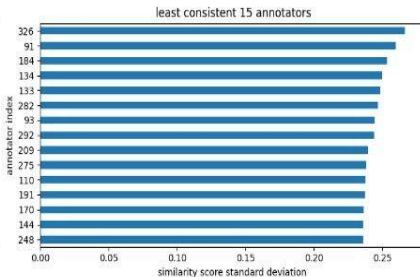
■ *How detailed are the text annotations?*
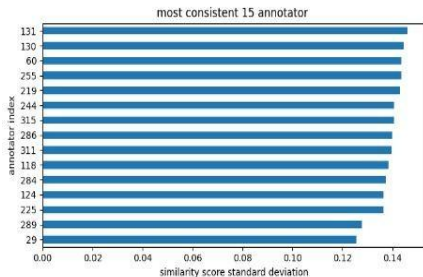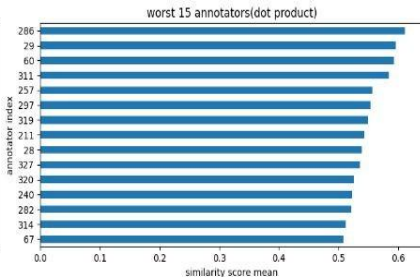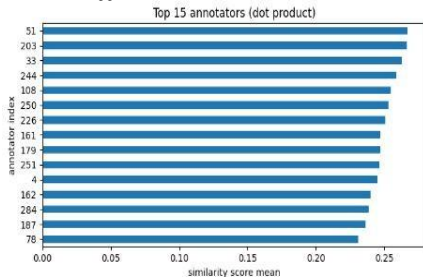
■ *How detailed are the text annotations? (cont.)*

■ *Propose a simple method to filter or fix incorrect or poor-quality annotations.*

■ *How much does the quality of annotations vary between different annotators?*

Thank you!
Any Questions?