



User



Adversary

Speaker
Inference



**Pseudo
Identity**



Human Listener

*Natural, Intelligible
Audios*



Text



ASR

Hello...

*Accurate
Transcriptions*