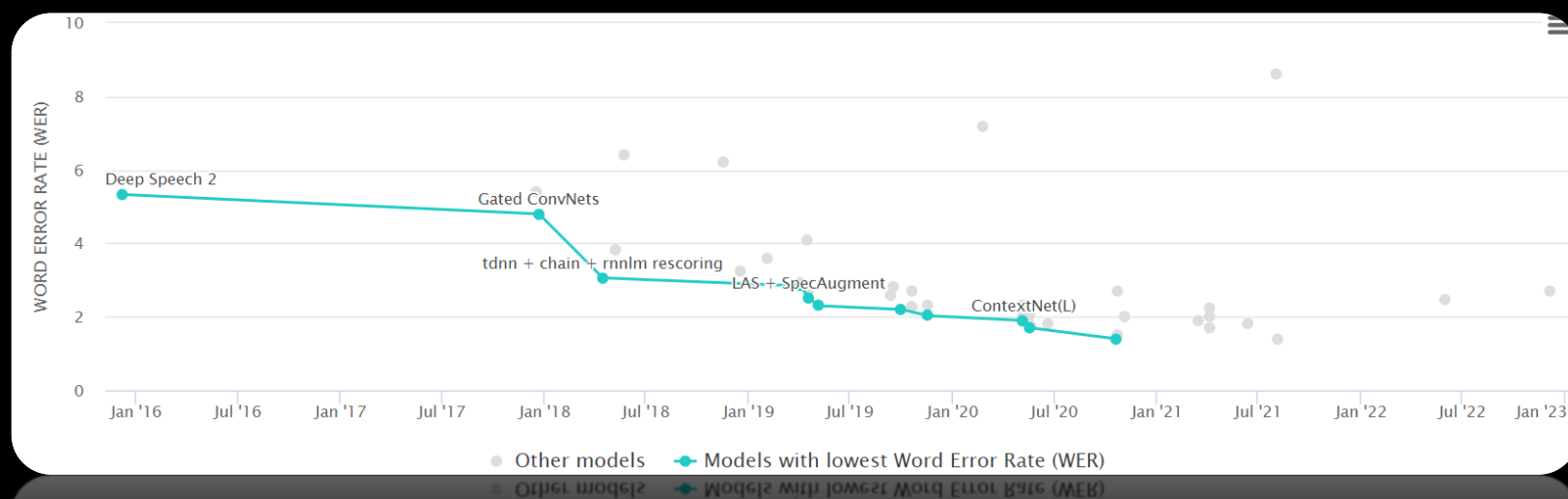# Word Error Rate (WER)



**Prabhjot Kaur**

# What is WER?

- A commonly used performance metric for evaluating Automatic Speech Recognition (ASR) systems.
- Compares number of errors in the transcribed speech to the reference text (usually transcribed by humans)
  - Comparison is done at the word level
- Derived from *Levenshtein distance*
  - Comparison done at the phoneme level for comparing two strings
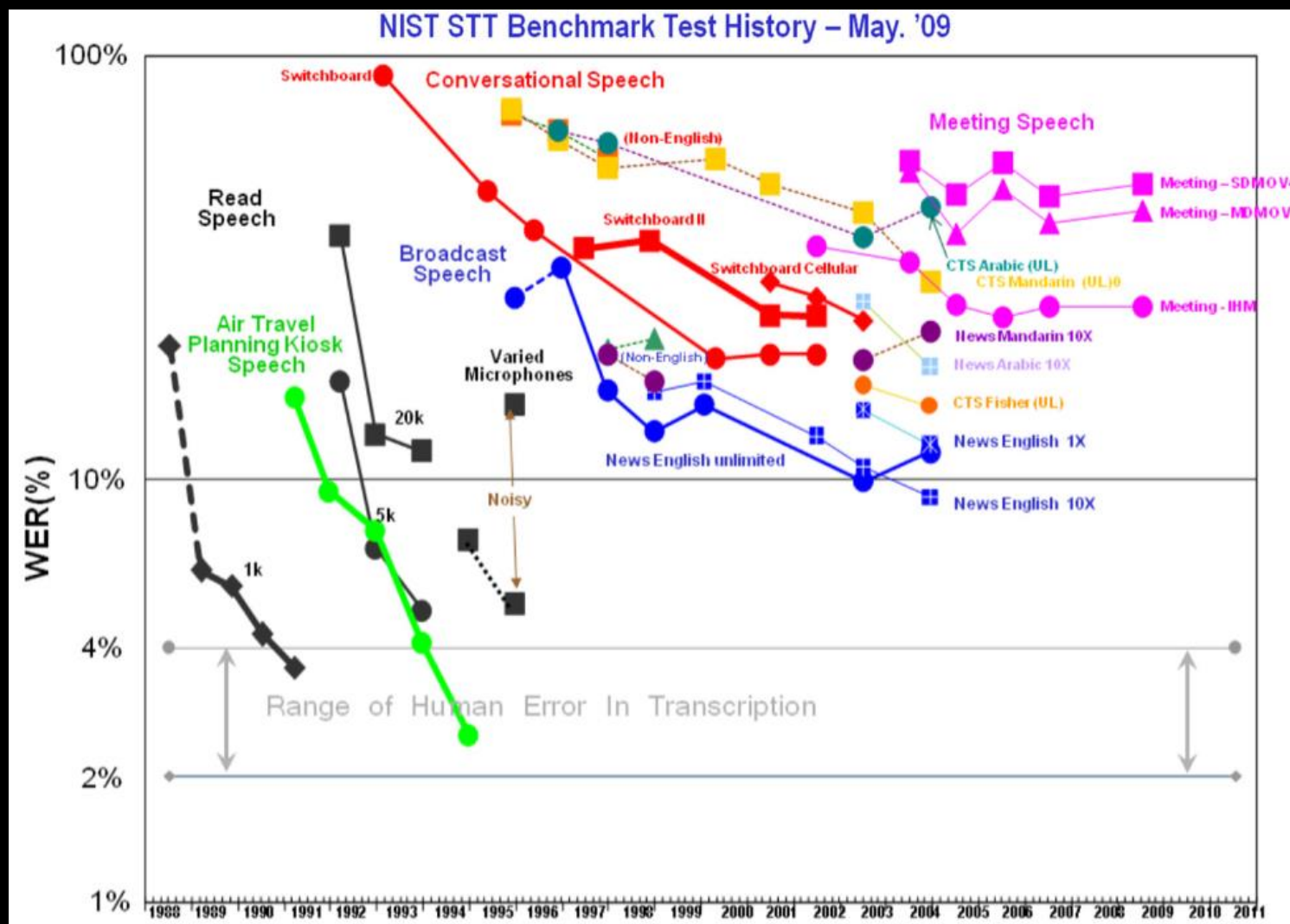    - Ex: Steel vs Steal. The levenshtein distance is (1/5 = 0.2)

# How to calculate the WER?

- Number of errors divided by total number of words in the reference sentence
  - Substitutions, Insertions, Deletions
  - WER does not account for the reason why errors occur

$$Word\ Error\ Rate = \frac{Substitutions + Insertions + Deletions}{TotalSpokenWords}$$

# Examples and Limitations

- Example 1:
  - Spoken: could you get me <span style="color:red">a towel</span>
  - Pred:    could you get me at all
  - WER: (2/6) = 0.333
- Example 2:
  - Spoken: hey tom it is mom will you come
  - Pred:      hey tom it is mom <span style="color:red">what do</span> you <span style="color:red">call</span>
  - WER:  (3/8) = 0.375

Thirty years of progress in speech recognition

# References

- https://smartaction.ai/blog/does-word-error-rate-matter/
- https://huggingface.co/spaces/evaluate-metric/wer

# Thank you!