Compare your performance over time

Analyse the evaluation result and identify the most common error patterns

Upload your annotations in CSV format

Are the results of the approach correct? How to evaluate the quality of hundreds of annotations?

1. Upload a Gold Standard to compare your annotations
2. Upload your annotations in CSV format
3. Analyse the evaluation result and identify the most common error patterns
4. Compare your performance over time

In case you are wondering, the annotations in the example are wrong. They are Pokémon!