

Deepgram Testing Guide



Setting Criteria

Before you even think about testing Deepgram's speech-to-text or any other solution, you should already have a list of criteria that you should evaluate for. Some of which should be based on testing and some based on the business or infrastructure needs. Listed below are some criteria to be considered.

TECHNICAL SELECTION CRITERIA

- Base model WER/Accuracy
- Accuracy on keywords
- Scalability
- Deployment needs

BUSINESS SELECTION CRITERIA

- Pricing
- Contract terms
- Data use and privacy terms

OTHER STAKEHOLDER CRITERIA

- Finance, IT, Sales, Operations, Security

We also have more detailed information in a quick [“How to Vet an ASR Provider”](#) brief.

Self-Testing or Guided-Testing

Depending on the ASR provider, you may have a choice of self-testing or guided testing. Self-testing is the ability to set up your own audio, APIs, and test it with the ASR solution, without any ASR provider involvement. Guided testing means you and the ASR provider discuss and agree on a testing plan that meets all your testing criteria and is tailored to your use case and speech recognition needs. Let's compare the two different approaches.



SELF-TEST ON API:

- + Easy to self-test your audio data
- + Can test on base models
- Need to manually calculate Word Error Rate (WER)
- Cannot test accuracy improvements on tailored models



GUIDED TESTING

- + Testing tailored to your needs
- + Quickly and easily test your audio data
- + Can use your audio data to improve accuracy
- + Provide Word Error Rate (WER) and Word Recognition Rate (WRR)
- + Can estimate how much more accuracy can be achieved
- + Test scalability

As you can see, one major difference is that with self-testing, you would need to create a ground-truth transcript and calculate Word Error Rate (WER) yourself and calculate any other metrics like alphanumeric error rate, keyword (product names, company names, jargon) error rate and Word Recognition Rate (WRR). You also could not perform speech model training with your audio data and see the accuracy improvements. This is why we recommend guided testing, even after you have done some self-testing.

Testing Goal and Expectations

As you work with your ASR provider, you should set up some testing goals and expectations with them. For Deepgram, our guided testing goals and expectations are:

GOALS

- ✓ Measure WER/WRR performance on base vs. tailored model
- ✓ Test accuracy on specific terminology, alphanumeric, names
- ✓ Deploy in the cloud and/or on-premise
- ✓ Select post-processing features to tailor specific use case
- ✓ Test different languages and/or regional dialects
- ✓ Others: scalability, pressure testing, etc.

EXPECTATIONS

- ✓ Higher throughput and lower latency than any other provider
- ✓ Base speech model WER is comparable to other STT providers
- ✓ Tailored speech model WER will normally exceed other STT providers and can continually improve

Guided-Testing Process

For Deepgram, our guided testing process will take from 1-2 weeks, but the data from the process will provide you with much more information and insights on how your technical and business criteria can be met.

DEEPGRAM'S TESTING PROCESS (1-2 WEEKS)

1. Outline use case, abstain representative audio samples
2. Test audio on the base models. Test with/without additional API parameters
3. Train a tailored speech model that improves overall accuracy and meets all business criteria
4. Verify the tailored model is performing well on new, unused audio within the use case
5. Deploy Deepgram to test scalability and minimize latency

Guided Testing Results

From any guided testing, you should be able to have the following standard results to compare with other speech recognition providers.

- ✓ WER/WRR results
- ✓ Throughput/Speed on cloud or on-premise
- ✓ Example transcripts of base model vs trained model
- ✓ Estimated accuracy improvements with tailored model
- ✓ Important keywords results (jargon, terminology, numbers, category names)

For more information, you can [view our webinar “Demystifying Testing”](#) on demand.

Want to get started with our guided testing, [contact us](#) today.

ABOUT DEEPGRAM

Better voice experiences start with better speech-to-text. Deepgram is the first and only End-to-End AI Speech Recognition platform that delivers insanely fast, actually usable transcriptions, with practically zero lag.

Deepgram—Every voice. Heard *and* Understood.
Learn more at deepgram.com.

