

How to Vet an Automatic Speech Recognition Provider



The Opus Research 2021 State of Automatic Speech Recognition report provided great industry insight but also offered recommendations on vetting ASR providers. We listed the top criteria for vetting ASR solutions along with questions you should ask.

Top Vetting Criteria

- 1 Deployment Flexibility**
 - How flexible are your deployment options?
 - Do we have to use your cloud storage or processing?
 - Can we use your solution on-premise or in our own private cloud?
 - Who owns that data?
 - Is the data in a proprietary format?

- 2 Accuracy**
 - What is the general, out of the box, accuracy?
 - Can you improve the accuracy?
 - With custom coding, addition of keywords, or speech model training, how much improvement can we expect?
 - Are there any tradeoffs between accuracy and other variables like costs, speed, or scalability?

- 3 Real-Time Capabilities**
 - Do you have real-time capabilities?
 - What is your transcription lag time; i.e. the period of time from when a speaker stops talking to the last words being transcribed?
 - Are there any tradeoffs between real-time speed and accuracy, costs or scalability?

- 4 Pricing**
 - What is your pricing for real-time streaming audio (price/hour of audio)?
 - What is your pricing for transcription of batch audio (price/hour of audio)?
 - What are the optional add-on pricing features?
 - What are the implementation costs?

5 Custom Training

- Do you offer custom training for your ASR solution?
- What is your custom training?
- What can we gain from custom training?
- What is the price of custom training?

6 Scalability

- What is the cost to scale from X real-time streams to Y real-time streams?
- What is the cost to scale from X hours of audio per day to Y hours of audio per day?
- What are the hardware costs for scaling from X to Y?
- Are there any tradeoffs with scaling from X to Y; slower speed or lower accuracy?

The report also noted two improvements the respondents would like to see from ASR solution providers that were not listed above. You may want to consider these criteria as well.

7 Custom Vocabulary or Language Model

- Can we add custom keywords or phrases to improve the model; i.e. company and product names, terminology, mixed language words (Spanglish, Chinglish, Indish)?
- Are there any tradeoffs with adding these words; i.e. slower speed, lower accuracy of other words, poor scalability, higher costs?

8 Speed to Deploy

- How long does it take to fully implement in production?
- How long to code or customize the solution to my use case?

Get the full report and listen to the webinar at <https://deepgram.com/state-of-asr-report/>.
If you are interested in a comparison of ASR providers [contact us](#).