



Next Gen Speech Recognition
for Enterprise

Scott Stephenson
scott@deepgram.com
415-202-4473

Speech recognition is in a rut. We're the physicist outsiders fixing it.

Starting with the largest dataset in the world:

Enterprise Speech Recognition.

We sell to these leaders in Enterprise & Medical

Product Leaders

Voice Platform

Strong Engineering Team

Forward Thinking

Data Science Teams

Large Call Centers

Established Data Science Team

Voice Data Goldmine

Problem

Speech recognition is **not solved** for phone calls or meetings, i.e. doesn't work for enterprise.

CONSUMER MARKET

93%
ACCURACY

Commands

ENTERPRISE MARKET

71%
ACCURACY

Phone Calls

65%
ACCURACY

Meetings

Enterprise companies can't adopt voice understanding at scale yet.

Solutions from Legacy Competitors don't work. Big Tech solutions are just placeholders.

The current experience:

60-75%

ACCURACY

???

UNRELIABLE

...

MISSING FEATURES

Solution

Deepgram provides fast, reliable speech recognition that **learns from customer audio** to get the best accuracy.

SOLUTION

Data Driven DL API

Fast, reliable, scalable, low cost

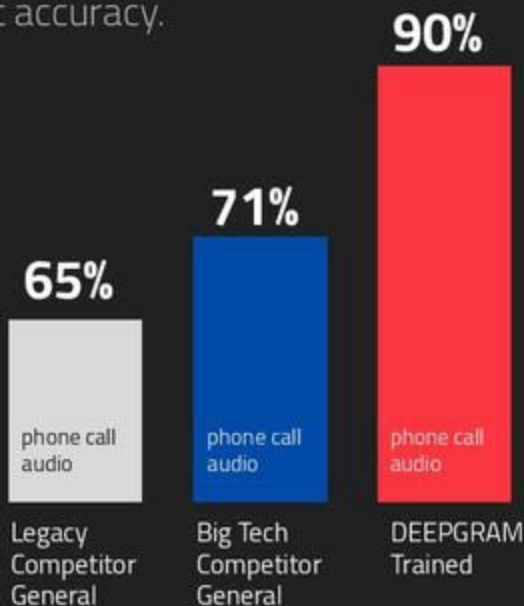
Data Flywheel

We train to get the best accuracy

raw data → label → train → deploy

Patented Architecture

Lean, low cost, truly innovative



actual customer result

How do we do it?

SOLUTION

The tech is groundbreaking, unmatched, and patented.

DATA we get raw data from customers and label in house

+

MODEL we use end to end deep learning to increase speed of model building and reduce the problem to be purely data-driven

+

COMPUTE we host models and infer entirely on GPUs to provide massive throughput (which reduces system complexity and cost)

It's universal.

Our architecture allows models to learn anything a human can do.

- Learn new languages at 10X speed, detect them too
- Figure out who is speaking
- Determine demographic information about the speaker (a/s/I?)
- Gauge their mood or sentiment
- Learn new vernacular, jargon, topics

All with the same data-driven model architecture

B2B Market

B2B MARKET

Enterprises want speech transcription.
They need accuracy and affordability at scale.

Big. It's Only Nice to Have
CONSUMER

95% ACCURACY NEEDED

NO WILLINGNESS TO PAY

This is where we focus

Big & Need It Now

ENTERPRISE

Sales & Support

85% ACCURACY NEEDED

\$\$ WILLINGNESS TO PAY

Big & Have Always Needed It

MEDICAL

99% ACCURACY NEEDED

\$\$ WILLINGNESS TO PAY

What will win enterprise?

One-Size-Fits-All

Tries to serve all, but serves none

Stuck with **65-75% accuracy**

Customer **can't improve accuracy**

- Catch-all model stretched beyond capability
- Poor on calls and meetings
- Lacks enterprise features
- Slow: 2x speedup
- 1 hour transcript done in 1,800 seconds
- Unreliable: 30% of uploads fail

Legacy & Big Tech Competitors

VS.

Trained Expert

Focused on the customer

Achieves **80-95% accuracy**

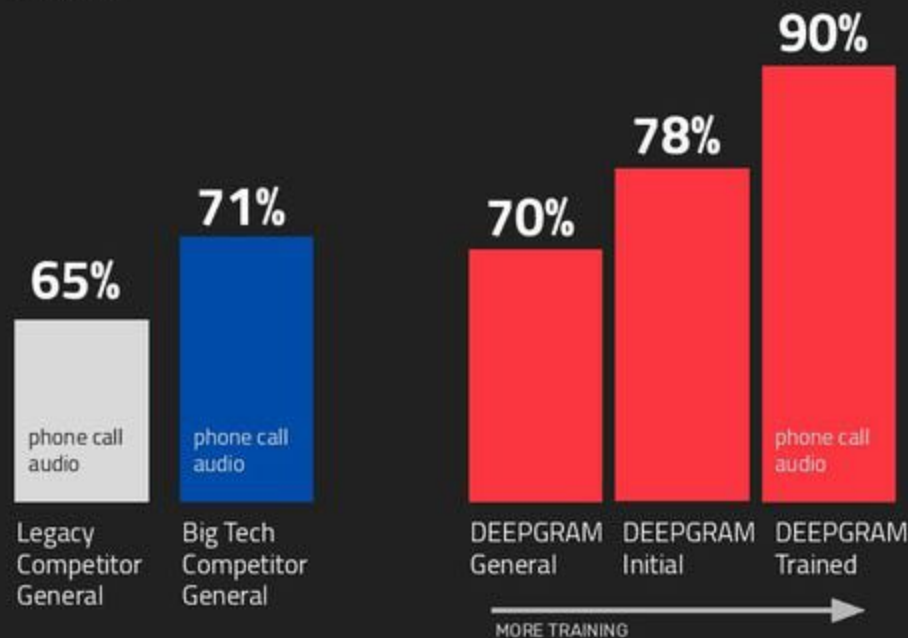
Customer **CAN improve accuracy**

- Models trained for each customer
- Great on calls and meetings
- Loaded with enterprise features
- Fastest in the world: 120x speedup
- 1 hour transcript done in 30 seconds
- Rock solid availability: 99.9% SLA



Our trained model accuracy is second to none

DIFFERENTIATION



65%

REDUCTION IN ERRORS

Just turn the dial to
get more accuracy.

This is unprecedented and
patented.

FEATURES	ENABLES YOU TO	GOOGLE	AMAZON	NUANCE	IBM	DEEPGRAM
Speech-To-Text		Great	Average	Poor	Poor	Great
Realtime Streaming		Poor	None	Average	Average	Great
Expert Custom Model		None	None	None	None	Great
Deep Search		Poor	Poor	Poor	Poor	Great
Language Support		Great	Average	Average	Average	Beta
Custom Vocabulary		Poor	Poor	Poor	Poor	Great
Multi-Channel		None	None	Average	Average	Great
Timestamps		Average	Average	Average	Average	Great
Alignment		None	None	None	None	Great
Punctuation		Average	Average	Average	Average	Great
Phonemes		None	None	None	None	Great
Dedicated Cloud VPC		None	None	Poor	Poor	Great
On-Premises		None	None	Poor	None	Great
API		Average	Average	Average	Average	Great

B2B MARKET

All of these
matter to
Enterprise.

DEEPGRAM
supports
them.

Realtime Streaming
Expert Custom Model
Deep Search
Language Support
Custom Vocabulary
Multi-Channel
Timestamps
Alignment
Punctuation
Phonemes
Dedicated Cloud VPC
On-Premises
API
Price
Understanding



**THEY CARE ABOUT
ACCURACY FIRST,**
then all of these.

Our customers love it

TESTIMONIAL

Randall Reilly

"We tried Google's Cloud Speech API and Nuance Dragon, and investigated several other products from companies including Amazon and Prosodica.

Deepgram had the best accuracy and program by far."

90%

ACCURACY

4M

MINUTES TRANSCRIBED

Brett Evanson

CTO at Smart Rhino Labs Division at Randall-Reilly

True AI Flywheel

We've built the right way to do transcription

HOW WE DO IT

Label Data

to build a unique,
valuable dataset

Train Model

to get world leading
accuracy

Data & Model
Progress

Attract More Data

with accurate, fast, affordable
products

LABELING & TRAINING BOTH DONE BY DEEPGRAM
IS EXTREMELY IMPORTANT.

END-TO-END DEEP LEARNING ON GPUS MAKES IT
FEASIBLE AND FAST.

WE HAVE PATENTS FOR THIS.

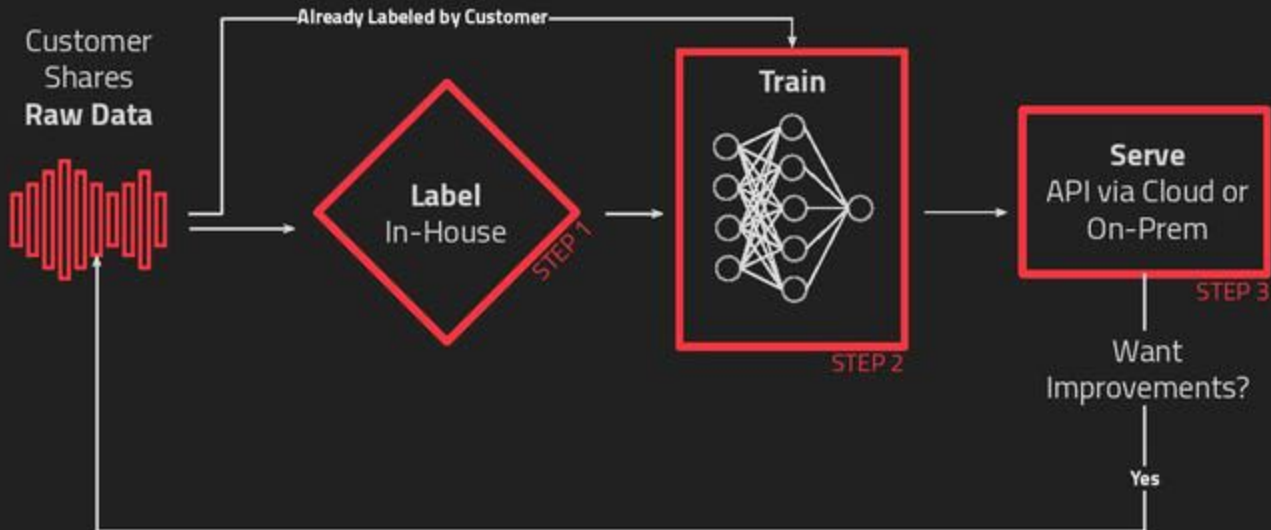
70%

Accuracy
DEEPGRAM
General

TRAINING WORKFLOW

90%

Accuracy
DEEPGRAM
Trained



The right team for next-gen Speech AI

Leadership



Scott Stephenson

CEO

- Particle Physics PhD
- Underground Dark Matter Research
- Applied AI Systems Expert



Adam Sypniewski

CTO

- Dark Energy PhD
- Built next-gen AI tech for DARPA
- Designed AI systems for autonomous vehicles

Team

- 30+ years AI experience
- 3 Physics PhDs
- Built with 5 engineers capital efficient



We've derisked tech, product, & sales.
Now we're going to market.

What's next:

Speech Recognition

Higher Accuracy

Languages

Lower Cost

Speech Understanding

Speaker Recognition

Sentiment Analysis

Topic Classification

Gender Detection

Age Identification

Lead Qualification

Compliance Scoring

Raise A

\$3.6M Seed Raised: YCombinator, Compound, NVIDIA, Slack

Grow Team, Infra, & Data

Marketing, Sales, Engineering, Servers, Data Labeling

\$10M ARR

In 18 months

desktop → laptop → mobile → voice

MARKET 11

Voice is the next interface

fast, natural, easy

MARKET SIZE

\$12.5B

TODAY

*Grand View

We'll take the voice market by winning enterprise first.

\$1T

Voice TAM

- 7B people,
- 3 mins processed by AI/day/person,
- at 1 cent/hr (200x less than G00G today)

Even with 200x cost reduction, it's a \$1.2T market