



rbifold AI

Make Efficient Data

Our team

20+ years of experience crafting AI solutions | **1 billion** users served | **Hundreds** of enterprise clients



Yongding (Enoch) Zhu

Founder & CEO

- Former executive at **Alibaba Group** and **Google**
- Business leader of data & AI, specializing in AI data solutions for LLMs of various sizes
- PhD in Computer Vision research



Mike Cheng

Co-founder & CTO

- Tech lead for **Gemini-powered Google Assistant**
- Specializes in large-scale model training and inference
- PhD in Face and fingerprint recognition

The Problem

Data is costly and noisy



LLM-ready data | 1-100 GB

Can be used for model training, testing, evaluation, tuning, RAG etc.



?



Enterprise Data | 1-100 PB

Including privately held data



Unstructured | Biased | Sensitive | Outdated | Siloed ...

Our Solution

We make data efficient for models by building FLAG datasets

Fresh

Updated timely to avoid the pitfalls of outdated information

Large

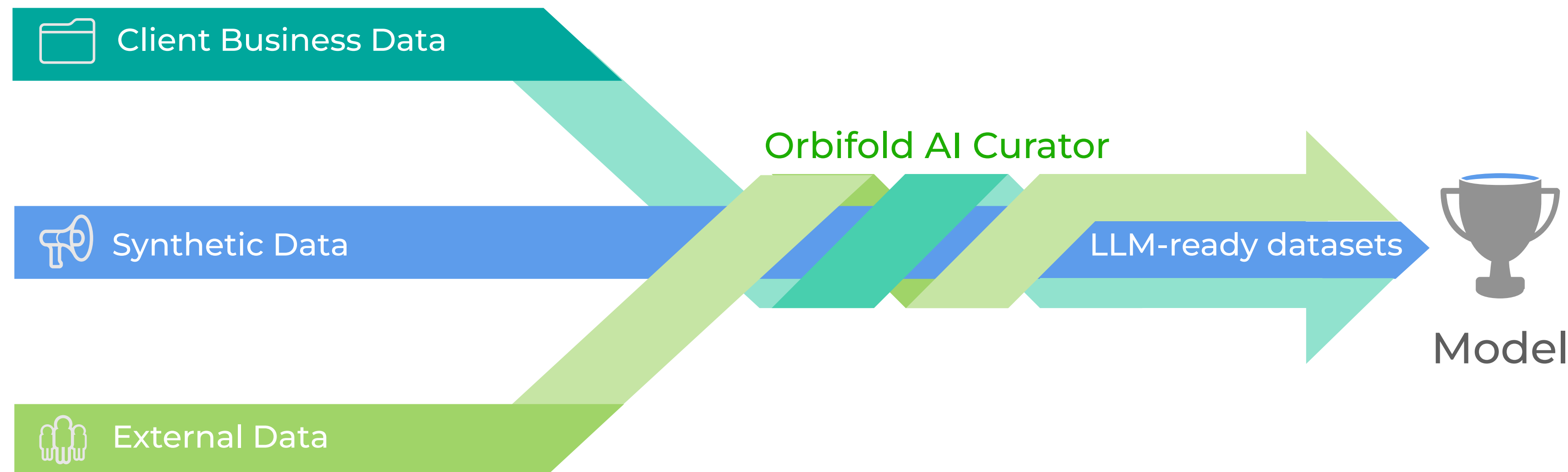
Large enough to transform the model into a domain expert

All-inclusive

Extensive enough to address the majority of use cases

Granular

Detailed, capturing the subtleties where true insights reside

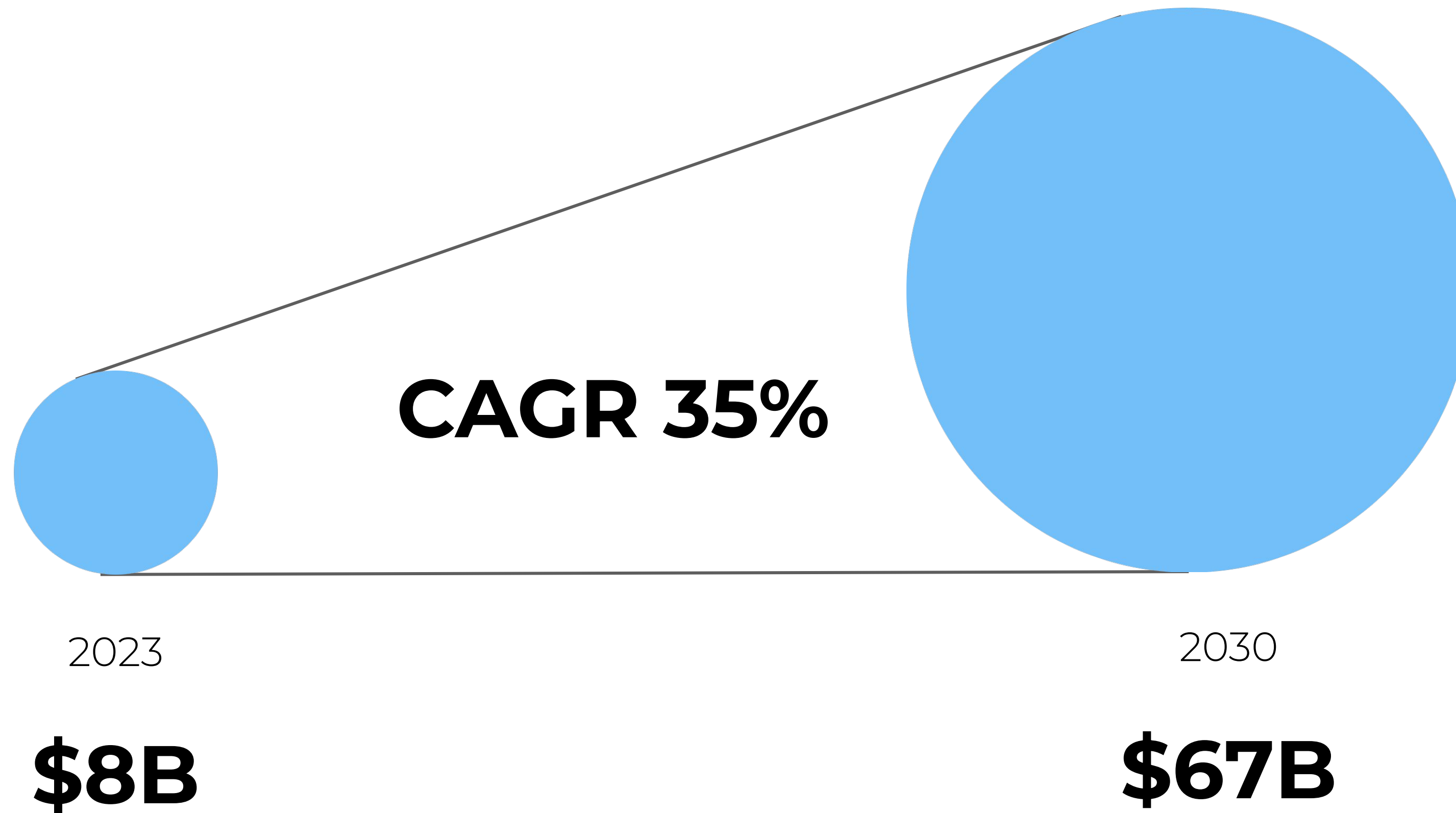


Sample | Prune | Structify | Augment | Desensitize | Evaluate

Why Now

Data curation is booming

■ Data curation



Other Players

	Orbifold AI	Scale AI	Datology AI
Funding	N/A	\$1.6B	\$57.6M
Valuation	N/A	\$13.8B	\$200M
Focus	Efficient Data	Data Annotation/RLHF	Data Sampling
Differentiator	Quality over cost	Low labor cost	Automation

Our business partners *



Our clients *



10+ incoming clients

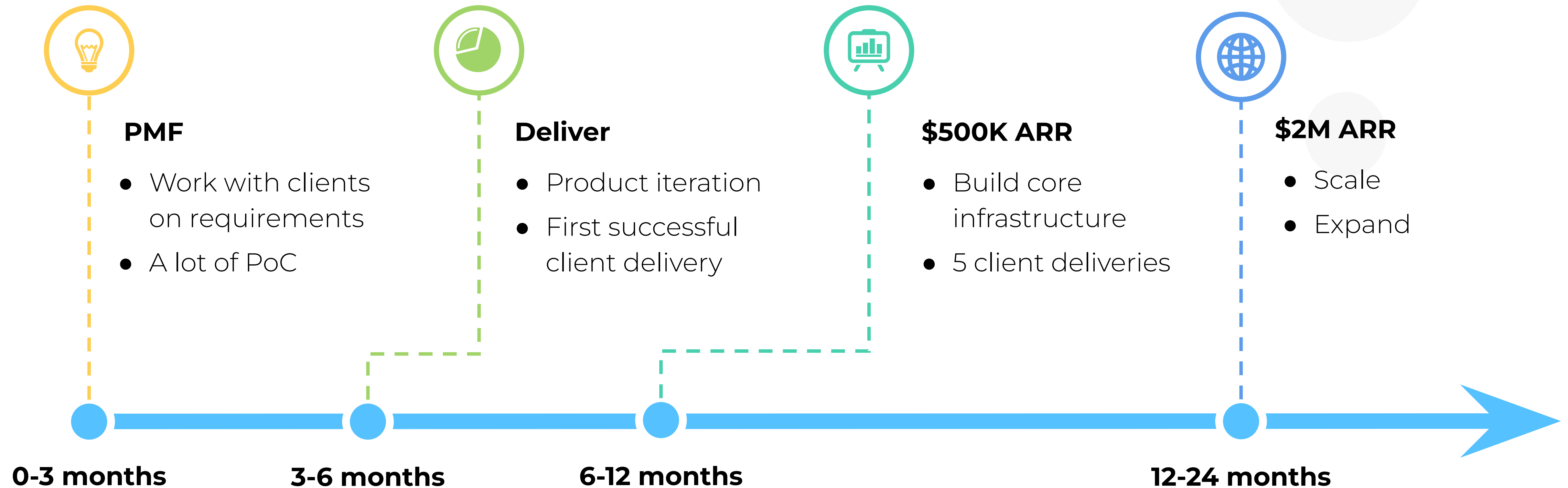
* Please feel free to contact us for more info



Our business model

Usage-based SaaS

A two year plan for sustained future growth



Orbifold by the numbers

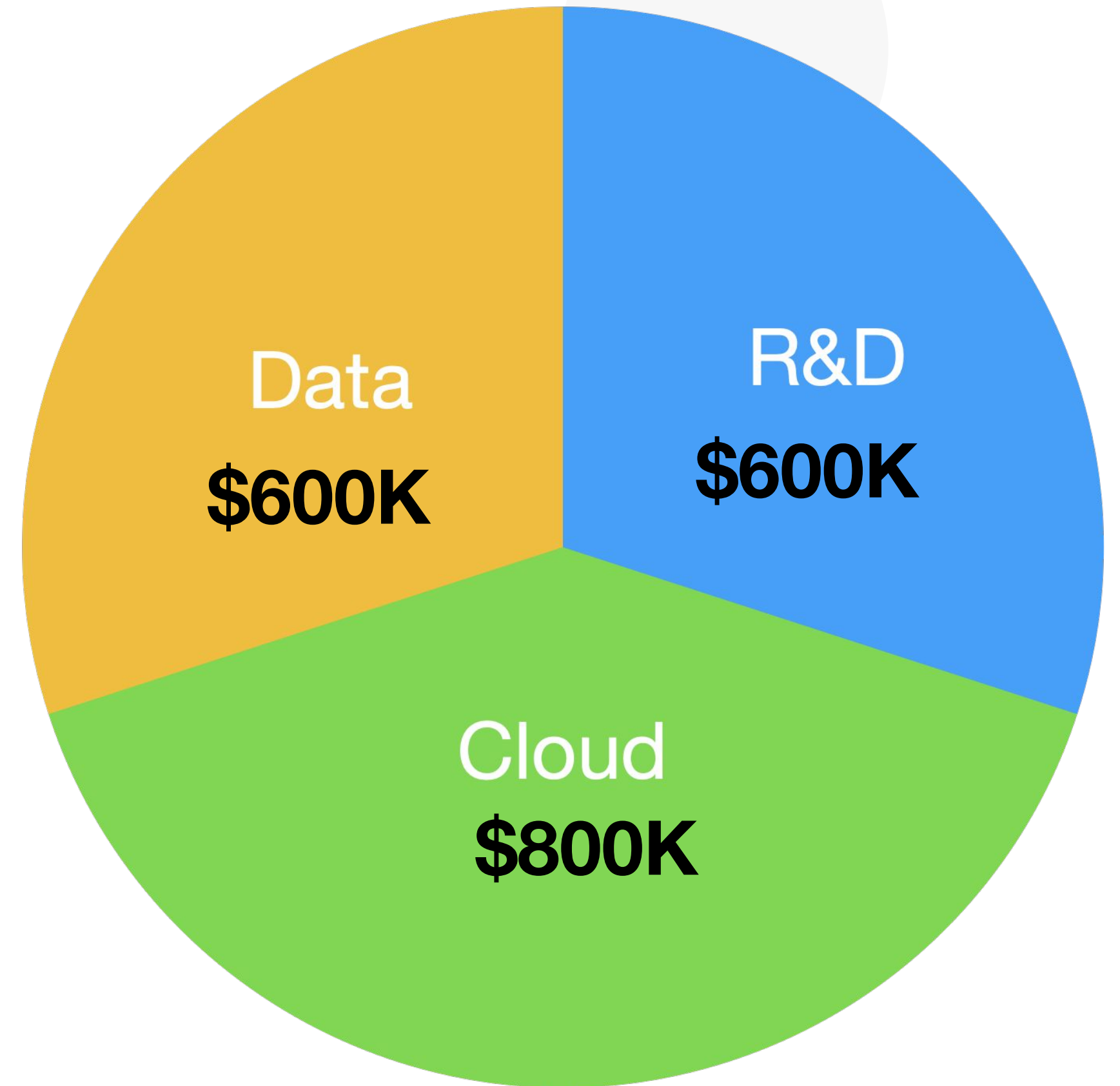
Raising \$3M

7+ clients

1.5 year runway

Our plan

- R&D for 2 founders, 3 engineers
- Cloud storage & compute
- Data procurement for external data



Yearly Expenses

Appendix: Product Demo

