



Snorkel Flow

Programmatic Data Development Program

Thank you for downloading this Snorkel.AI resource. Carahsoft is the distributor for Snorkel.AI's AI and ML solutions available via NASA SEWP V, ITES-SW2, NASPO ValuePoint, and many more contract vehicles.

To learn how to take the next step toward acquiring Snorkel.AI's solutions, please check out the following resources and information:



For Scale.AI overview:
carah.io/SnorkelAI



For upcoming events:
carah.io/SnorkelAI_Events



For additional resources:
carah.io/SnorkelAI_resources



For additional Artificial Intelligence:
carah.io/ai-solutions



To set up a meeting:
SnorkelAI@carahsoft.com

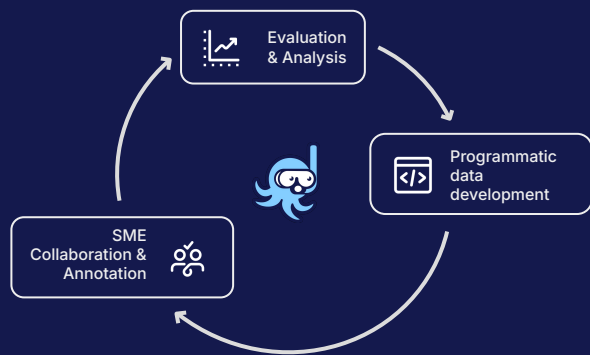


To purchase, check out the contract vehicles available for procurement:
carah.io/procurement

For more information, contact Carahsoft or our reseller partners:
888-606-2770 | SnorkelAI@carahsoft.com

Snorkel Flow

Programmatic Data Development Platform



Specialized, enterprise-ready AI is all about the right data. Label and curate data programmatically for faster, higher-quality, and more adaptable AI.

Snorkel Flow turns AI data development into *software development*, resulting in data + models that are:

- **Faster:** 10-100x faster than manual approaches
- **More accurate:** Rapid iteration to correct error modes and improve quality
- **More adaptable and transparent:** Edits and audits as easy as modifying or reviewing code

Rapidly and iteratively develop high quality datasets for:

LLM Evaluation

Create golden evaluation datasets and tune specialized auto-evaluators for **specialized, fine-grained** evaluation.

LLM Tuning & Development

Label and curate datasets for LLM fine-tuning, alignment, distillation, and in-context learning.

RAG System Tuning & Development

Label and curate datasets for tuning custom embeddings, retrieval, and re-ranker models, as well as chunking/tagging.

Used by leading enterprises, including **7/10 top US banks**



CHUBB®



Snorkel Data-as-a-Service

Fully managed AI dataset development

Uniquely high-quality, highly specialized data for LLM evaluation and development—using our unique platform technology + a highly curated network of subject matter experts

Snorkel's unique technology leads to data that is:

- **Higher quality:** Programmatic and manual quality controls for higher quality, more efficiently
- **Highly specialized:** More efficient use of higher-quality subject matter experts- e.g. PhDs, professors, expert developers, and more
- **Cost effective:** Market-leading quality without market-leading cost

Spotlight example: Snorkel works with a leading OSS LLM developer to build uniquely challenging, specialized, and high-quality datasets, beating out all other major vendors:

80%+

FRONTIER LLM
FAILURE RATE

100+

PHD-LEVEL
SUB-TOPICS

98%+

ACCURACY AT FIRST
QA PASS