AI Acceleration

Same Answers, Faster

The output of language models can vary based on the prompt.

This can have a huge impact on downstream task performance.

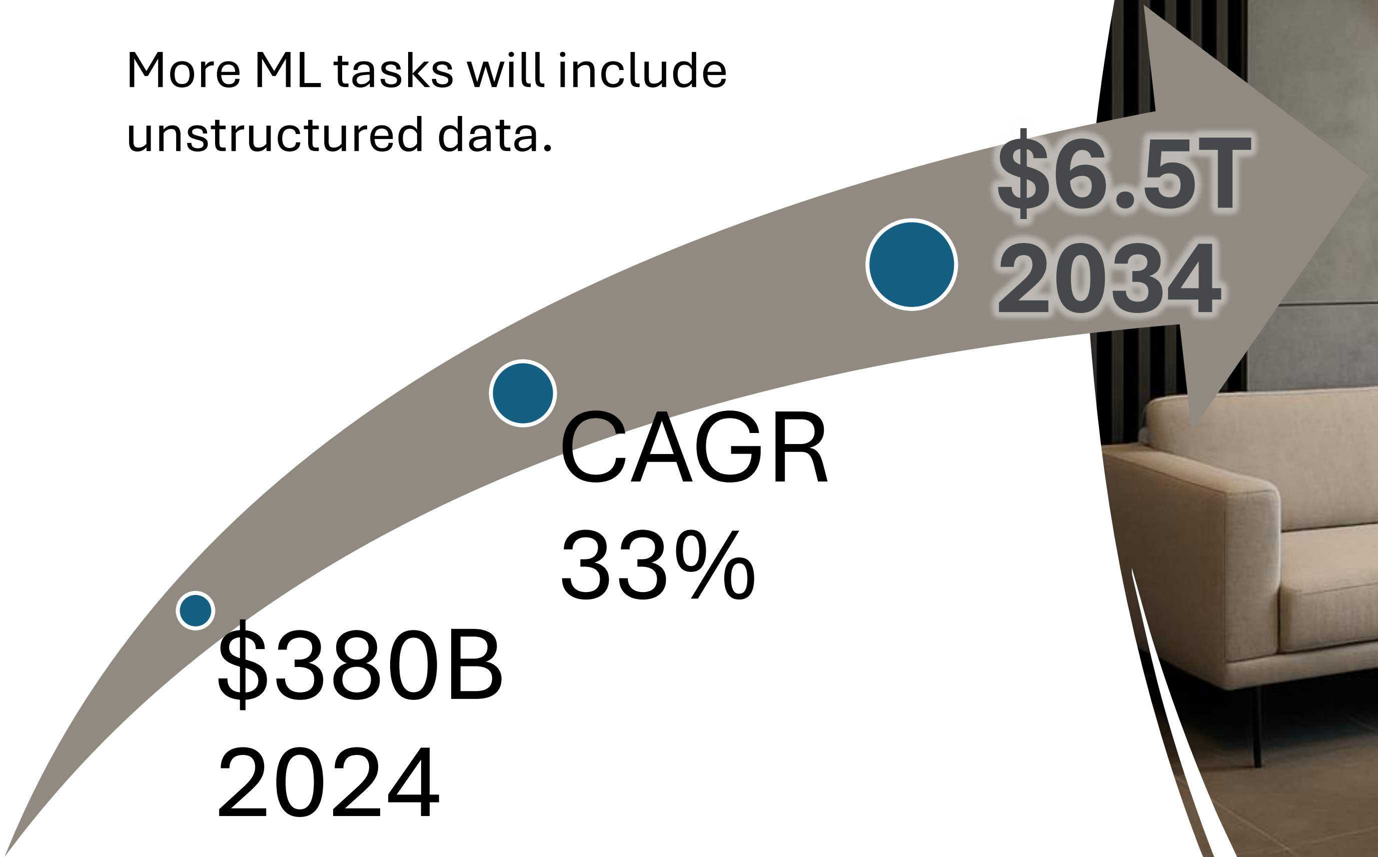
Tasks such as:

- Classification
- Information Extraction
- Intent and Sentiment Detection

In addition to accuracy, *Better Prompting* has implications for governance, security and tone.

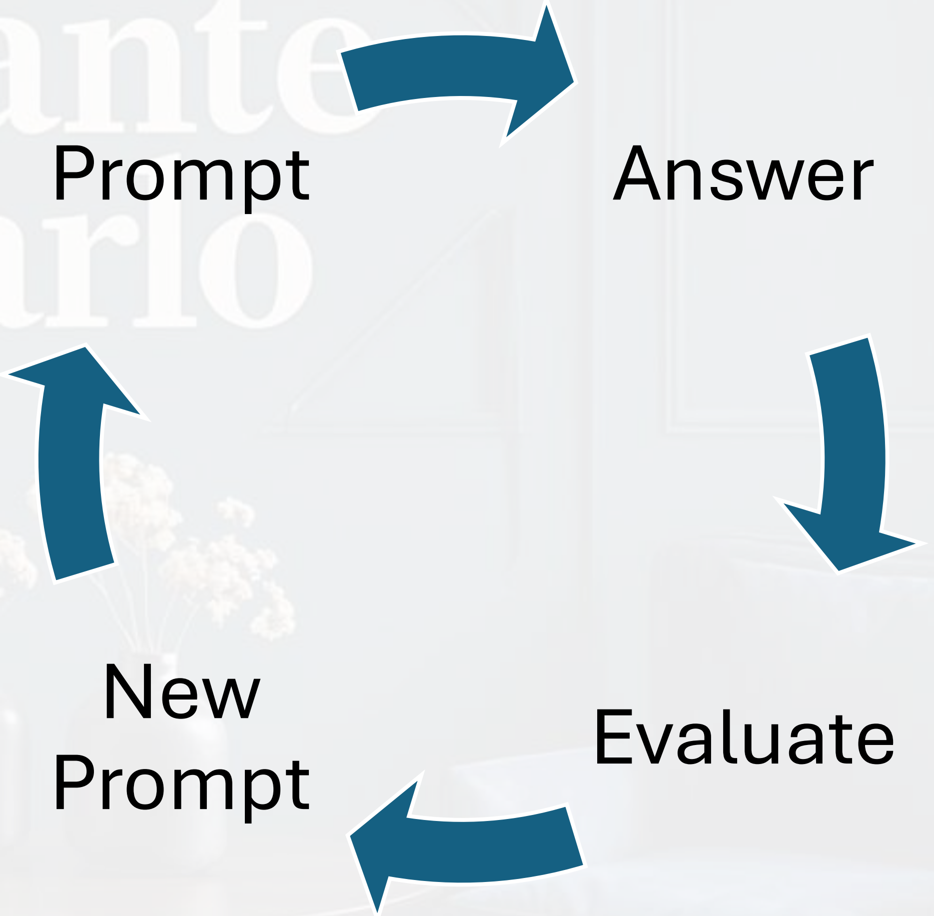
The Prompt Economy

More ML tasks will include unstructured data.



Prompt Optimization

- Quante Carlo accelerates a key part of the Automated Prompt Optimization Process: 300% - 1000% faster
- Same answer, faster
- In a market test, QC was 5x faster overall for the same performance as DSPy



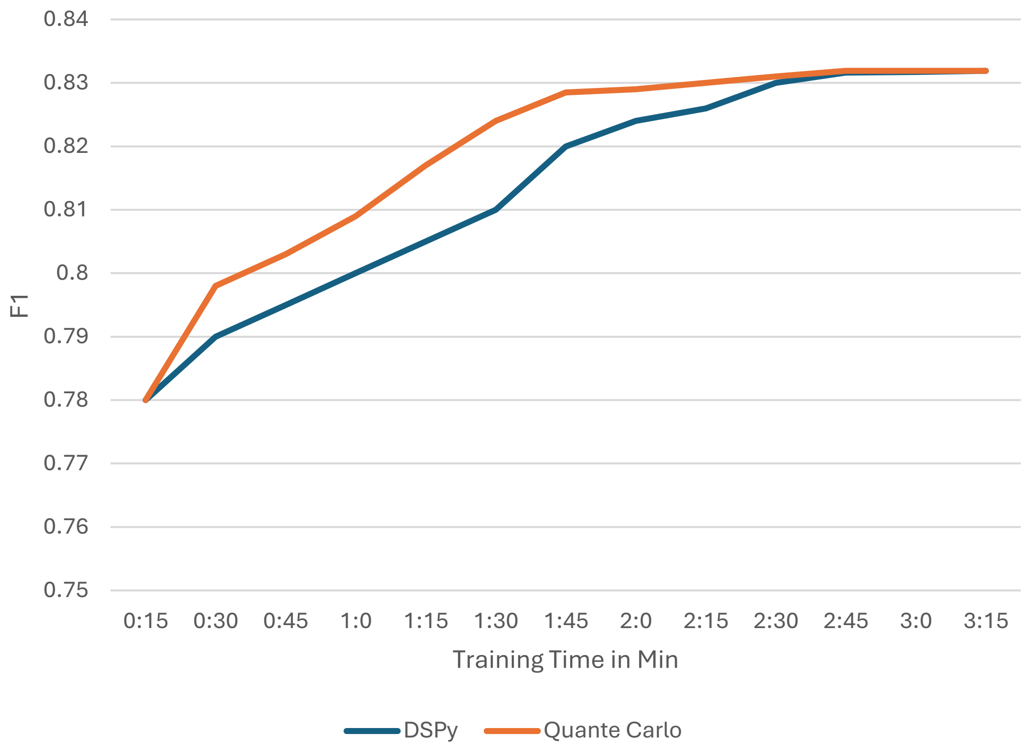
QC outperforms the current state of the art.

Improving the time it takes to make the next guess also allows the next guess to be better. The process converges to the right answer in fewer steps.

In production

- Liability Determination
- Subrogation Document Validation
- Total Loss (Quick Tow)
- Claims Prioritization

Images, Audio, Code, Tools (Excel)



Competitive Landscape

Because of QC's unique technology, all of these competitors are potential customers.

Large Companies with Prompt Tools Offerings	Revenue derived from AI
AWS	$20B - $30B
Microsoft	$13B
Anthropic	$850M - $2.2B
NVIDIA	$130B
Oracle	$5B - $10B
Salesforce Trailhead	$900M
Cohere	$85M
Dataiku	$300M

Newer companies focused on prompt workflow management	Revenue	Raised
Kore.ai	$209.8M	73.5M
TELUS Digital AI	$2,658M	
Arize AI	$13.6M	$131M
Geniusee	$25.4M	
Narrative IO		$14M – $19M
MrGlasco.ai	< $1M	
Prompt Layer		$4.8M
LangChain		$35M
Toolhouse		$9.6M
AiFA Labs	$13M	
Lyzer	< $1M	
Firebase	$2.7M	
Aipromptstudo		< $1M
MrGlasco.ai	< $1M	
Airia		$2.3M
airprompt.ai	$0	$0
Portkey		$3M

Defensible

220-year-old open problem

- Most solutions are compromises or attempts to work around the problem.
- New problems are being framed faster than they are being solved, so, not all open problems get attention

Algorithms are expected to be open source.

- Due to IP climate in the US, usually everything is either open source or can easily be reverse engineered given the necessary resources (talent, infrastructure)

Meta does not have this.

- This is how I found out about this use case.
- QC is the only company with batch Bayesian Optimization on CUDA.

History

QC was discovered while solving a problem in high-performance simulation for financial risk management.
Solving a 220-year-old, more fundamental, open problem in the field of probability and statistics, I looked for other use cases

- Power Grid
- Structural Engineering
- Logistics
- Networking
- Cybersecurity
- AI/ML
- Recommendations
- Materials Science
- Robotics, Gaming

Plan

R&D < 2020

Dec 2024: 1st Prompt Engineering

Customer

Dec ‘25: +5 Customers

2026:

Key Partnership

Seed Round

- Prompt Engineering Platform
- 3rd party module integrated with other platforms

2028+: HP Tuning; AI/ML Platform Integration

Opportunity = $250K

Mark Shipman

- 27 Years of AI/ML/DS Experience Including Ops
- Carnegie Mellon (Math, CS), Columbia (Engineering)
- 1 Founder, 100% Equity, $0 Expenses