HOW TO COOK GOOD AI PRODUCTS WITH WHAT YOU ALREADY HAVE IN YOUR DATA WAREHOUSE

Julia Neagu | CEO, Quotient AI
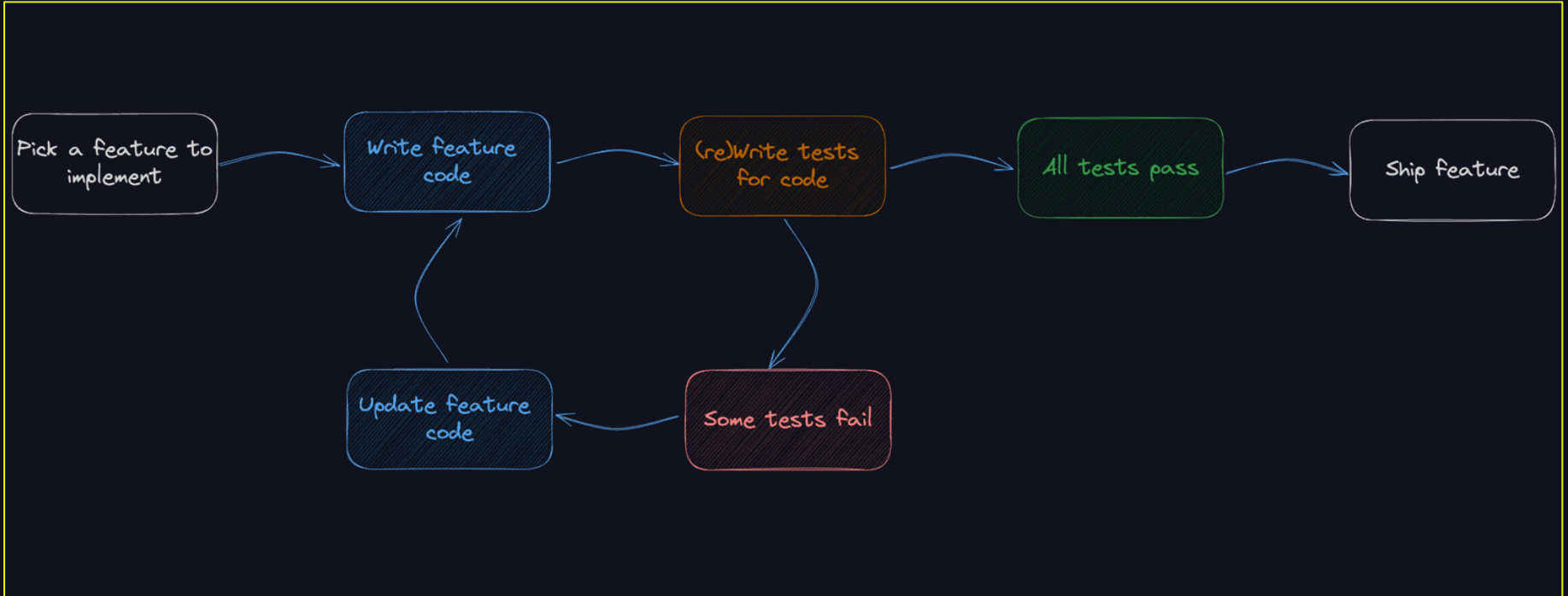
QuotientAI

"*90% of enterprises are not confident going to production with gen AI solutions*"
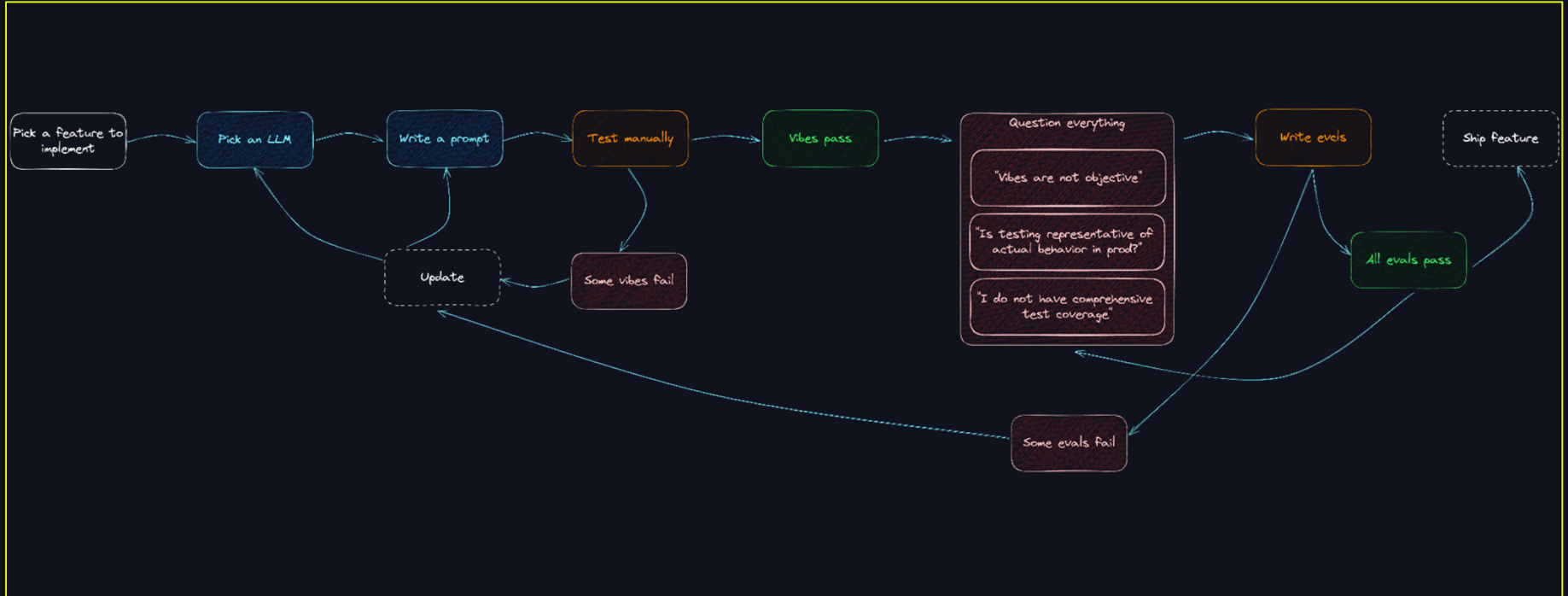
why?

HOW IT STARTED

Writing, testing, and shipping code with deterministic behavior is ~~easy~~ *easier*...

HOW IT'S GOING

... but having realistic, reproducible, and compressive testing for LLMs is *not*

THE ~~3~~ ~~4~~ 5 RULES OF ~~TESTS~~ EVALS

Evaluations that lead to consistent product ships must be:
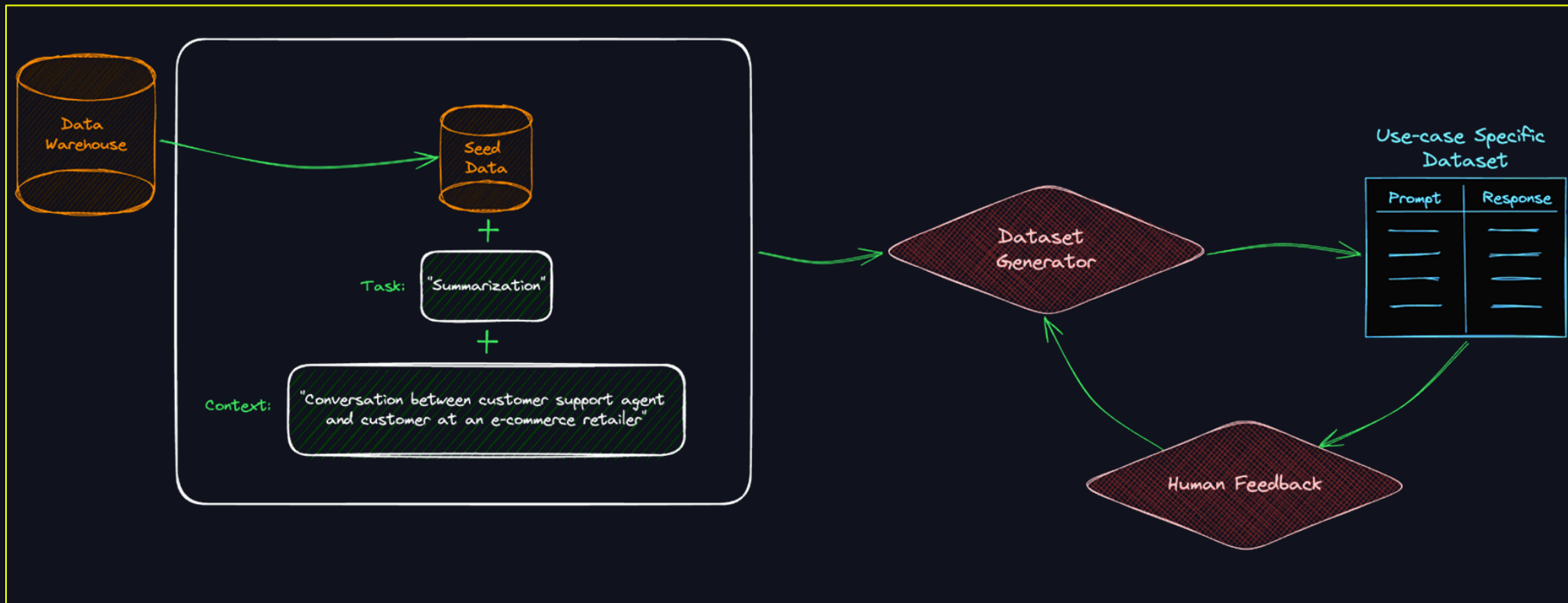
1. **Realistic.**
   Reflecting actual production scenarios accurately.

2. **Aligned.**
   Correlated with human judgement.

3. **Comprehensive.**
   Encompassing a wide range of production scenarios.

4. **Reproducible.**
   Producing the same results under unchanged conditions.

5. **Secret.**
   Not part of the training data.

how?

DATA AI SUMMIT

automatic evaluations

w/ human-in-the-loop feedback

starting from real data

DATA AI SUMMIT

WAREHOUSE-TO-REFERENCE-DATA GENERATION

Existing data can be the starting point of use-case-specific reference data for evaluation and fine-tuning.

DATA AI SUMMIT

THE 5 RULES OF EVALS

Evaluations that lead to consistent product ships must to be:

1. **Realistic.**
   Reflecting actual production scenarios accurately.

2. **Aligned.**
   Correlated with human judgement.

3. **Comprehensive.**
   Encompassing a wide range of production scenarios.

4. **Reproducible.**
   Producing the same results under unchanged conditions.

5. **Secret.**
   Not part of the training data.

DEMO

www.quotientai.co

find us at the Expo!

booth #e3