

Making Sense of Your Experience

The Three Gulfs Framework

Franck Albinet

franckalbinet@gmail.com

Independent Data Science & AI Consultant

October 8, 2025

What Did You Just Experience?

Debrief: Your First Encounter

Let's Hear From You

Quick round of sharing:

- What worked well in your paper extraction attempt?
- What was frustrating or confusing?
- Did the AI give you what you expected?

Let's take a moment to think

Common Patterns

A Recurring Theme Across Users

When working with AI, a recurring theme emerges:

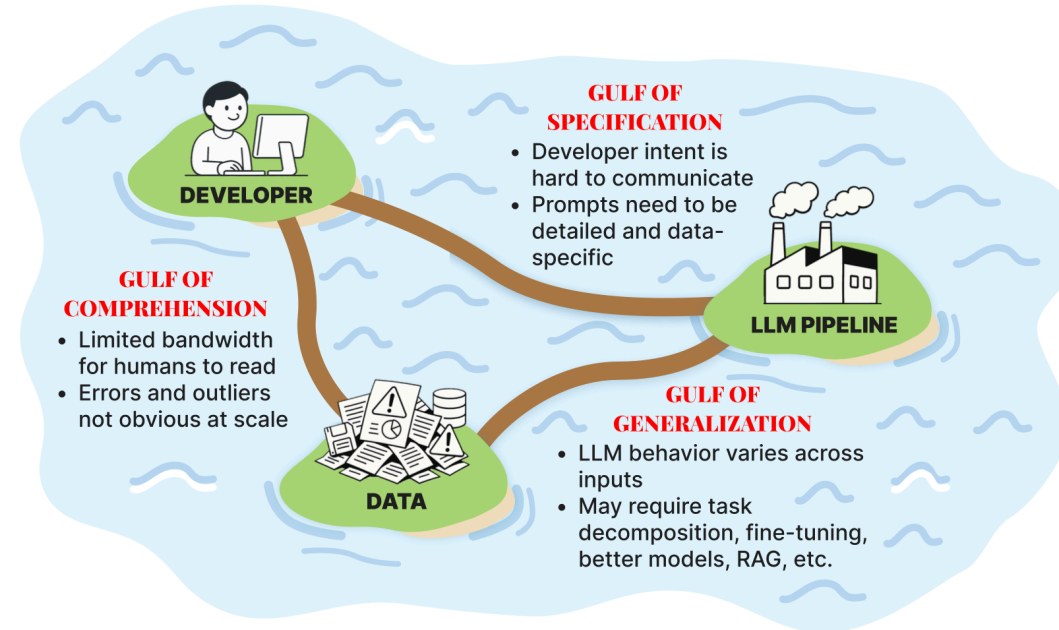
- **“It gave me something, but not quite what I wanted”**
- **“The output format kept changing”**
- **“I wasn’t sure how to describe what I needed”**
- **“It worked for one input but not another”**

Sound familiar? There’s a reason for this...

The Three Gulfs Framework

Understanding AI Pipeline Challenges

A Framework to Make Sense of It All



Three Gulfs Model - showing Developer, LLM Pipeline, and Data with the gulfs between them

This framework helps explain what you just experienced and guides our journey through this course

Gulf 1: Comprehension

You ↔ Your Data

The challenge: Understanding your input data (scientific papers) and how your AI pipeline behaves on that data at scale.

What you just experienced:

- You couldn't easily check if the AI extracted everything important
- Hard to know if it would work the same way on your next 50 papers
- Difficult to spot patterns in what it missed or got wrong

Critical insight: We need systematic ways to understand our data and AI performance at scale

Gulf 2: Specification

You ↔ AI Pipeline

The challenge: The gap between what you want the AI to do and what you actually communicate in your prompts.

What you just experienced:

- “Extract key information” - but what exactly is “key”?
- Should it be bullets or paragraphs?
- How detailed should methodology be?
- What format do you actually want?

Key point: Seemingly clear instructions often contain hidden ambiguities

Gulf 3: Generalization

Your Data ↔ AI Pipeline

The challenge: Even with perfect prompts, AI may behave inconsistently across different inputs.

What you might experience next:

- Your prompt works great for experimental papers, fails on theoretical ones
- Different paper structures confuse the AI
- Older vs. newer publications get different treatment

Important: Each application requires bridging the Three Gulfs anew - there are no universal solutions

Why This Framework Matters

Guiding Our Course Journey

Understanding the Three Gulfs helps us:

1. **Expect iteration** - Perfect prompts don't exist from the start
2. **Plan for evaluation** - We need systematic ways to assess performance
3. **Design for diversity** - Our approach must work across different paper types
4. **Stay methodical** - Each gulf requires specific strategies to bridge

This week's structure: We'll address each gulf systematically through hands-on practice

Connecting to Your Experience

Which Gulf Did You Feel Most?

Think about your hands-on session:

- **Gulf of Comprehension:** Couldn't easily verify the output?
- **Gulf of Specification:** Struggled to describe what you wanted?
- **Gulf of Generalization:** Worried it won't work on other papers?

Most likely: You experienced all three! That's completely normal.

The Week Ahead

How We'll Bridge Each Gulf

Days 2-3: Addressing the Gulf of Specification - Systematic prompt engineering - Iterative refinement techniques

Day 3: Addressing the Gulf of Comprehension - Evaluation methodology - Understanding performance at scale

Days 4-5: Addressing the Gulf of Generalization - Real-world applications - Handling diverse paper types and use cases

Questions & Reflection

Before We Move On

Final Thoughts

- Does this framework help make sense of your experience?
- Which gulf feels most challenging for your work?
- What questions do you have?

Next: Collaborative discussion on expectations, concerns, and ethics
(Session 1E)