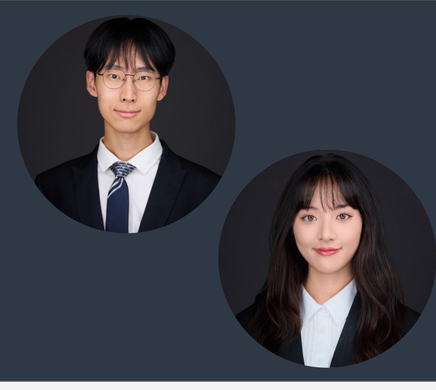
COMCAST

RAGsolve

A Universal AI Engine for Customer Issue Resolution

Capstone Project Sponsors: Benjamin Land, Ian Tongs
Faculty Advisor: Swati Gupta
MBAn students: Jieyi Wang (jieyi@mit.edu), Yuki Yu (yuki620@mit.edu)



OVERVIEW

CONTEXT

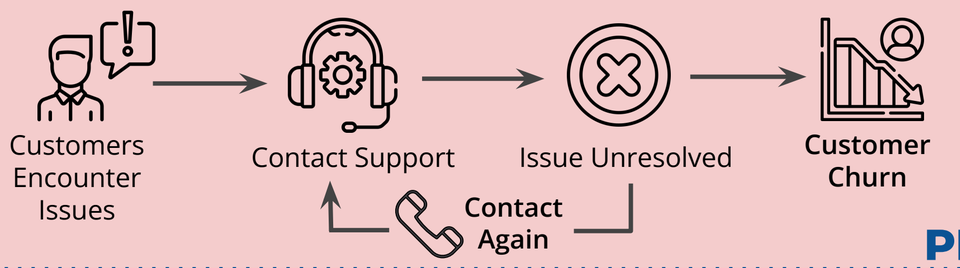
At Comcast, customer support can cause major costs — especially when issues go unresolved.

$100M
Annualized cost of handling customer issue calls

16.7%
Share of total churn attributable to unresolved issues

PROBLEM

Not resolving issues means more calls and lost customers.



DATA

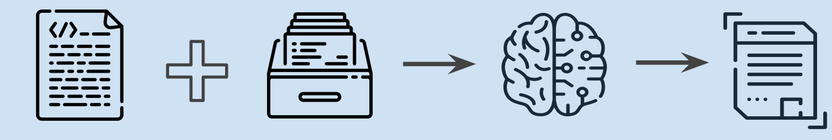
So by leveraging historical data on how we solved issues:

Transcripts
Digitized Transcripts
6 months, ~1.5 TB

Customer activity
Whether they churn, callback, etc

GOAL

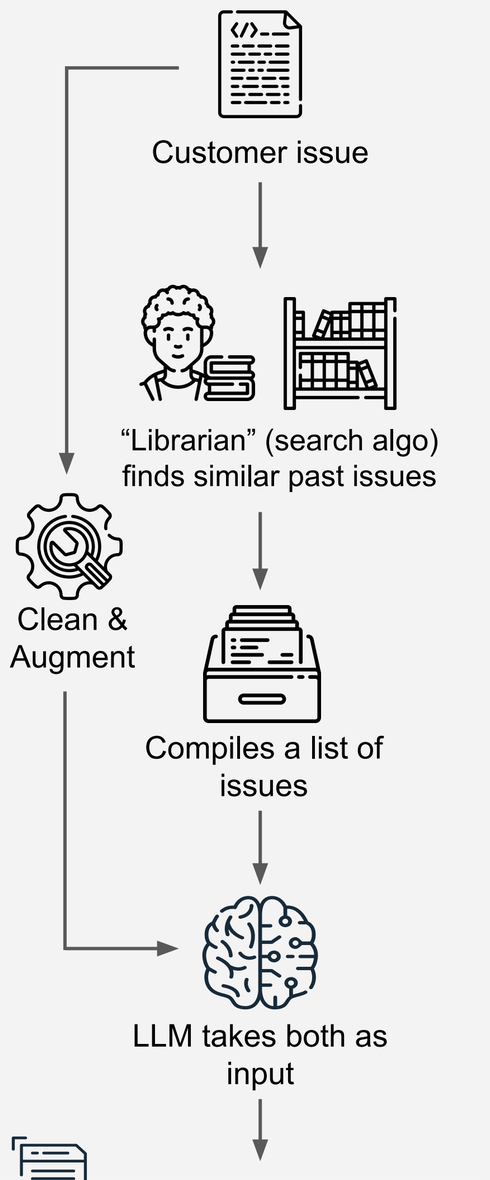
We build an AI engine (LLM + RAG) to generate better solutions.



APPROACH

LLMs are powerful "learners" — they generate best when given relevant, high-quality information.
So, we build them a "library" of past cases and train a smart "librarian" to retrieve best ones at runtime.

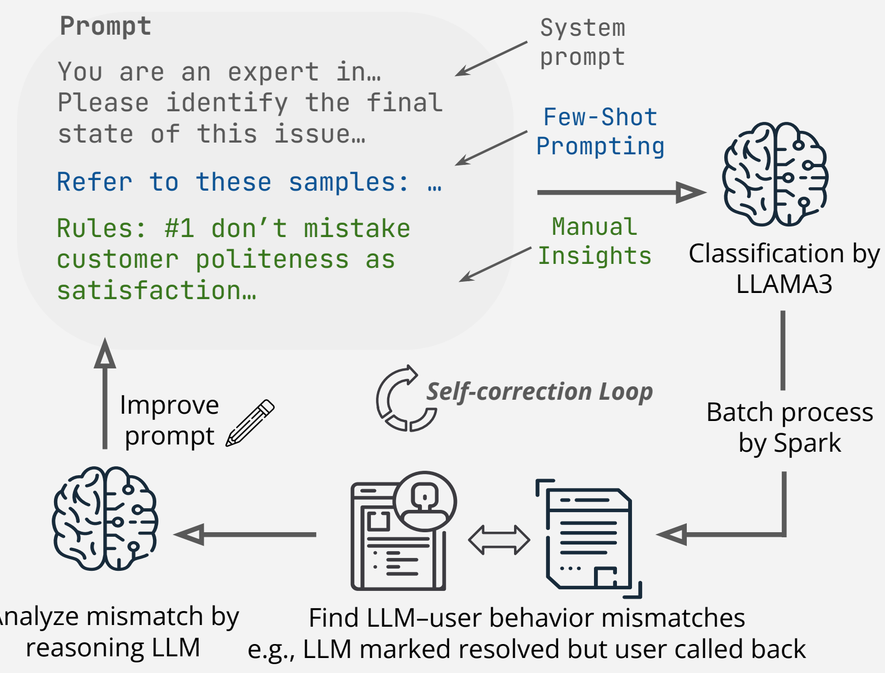
Pipeline ①



Build A Library ②

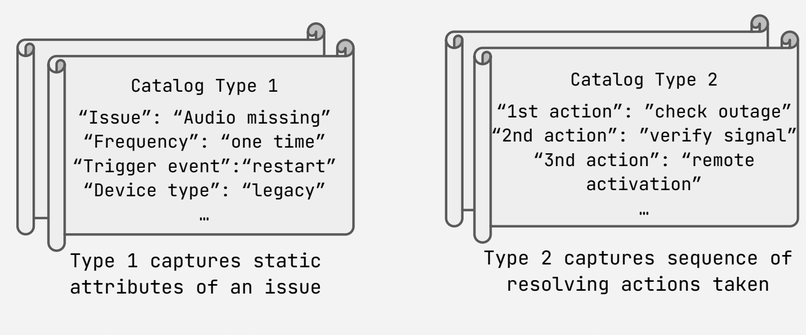
Curate the Shelves

Use LLM to distinguish between resolved and unresolved issues; retain mostly resolved ones to ensure quality.

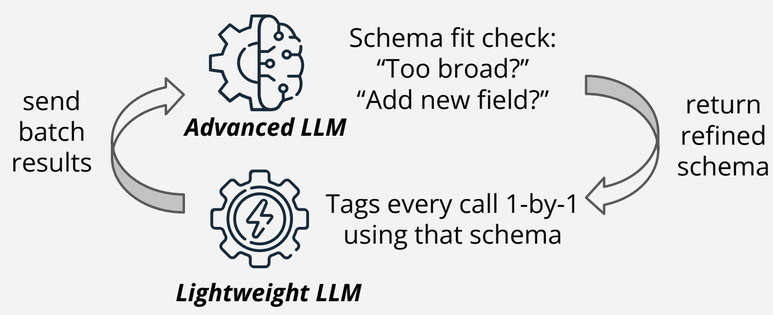


Design the Catalogs

Extract structured attributes from raw text. This structured catalog serves as a map for locating similar cases — which supports more precise retrieval later.



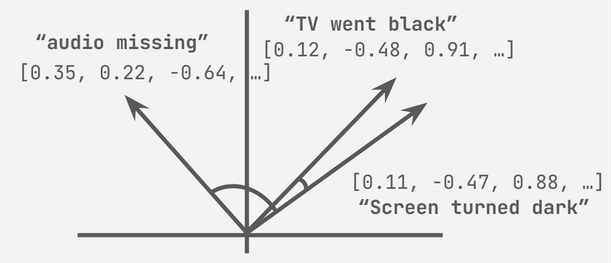
The schema of the catalog is generated in an iterative way:



Train A Smart Librarian ③

Utilize Vector Retrieval

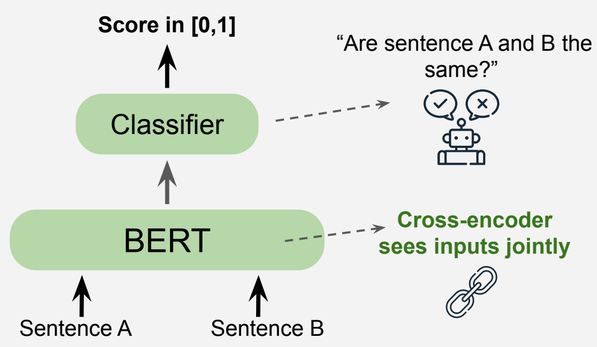
Embed cases in vector space for a fast, first-pass retrieval based on meaning.



Issues with similar meanings are closer in vector space

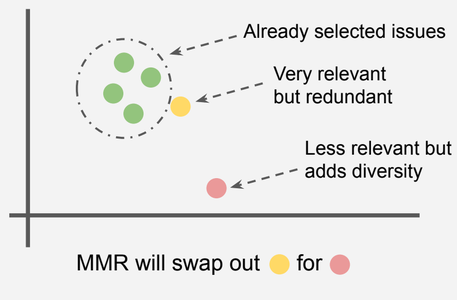
Rerank with Precision

Use cross-encoder and structured catalog matching together to surface the most relevant cases.



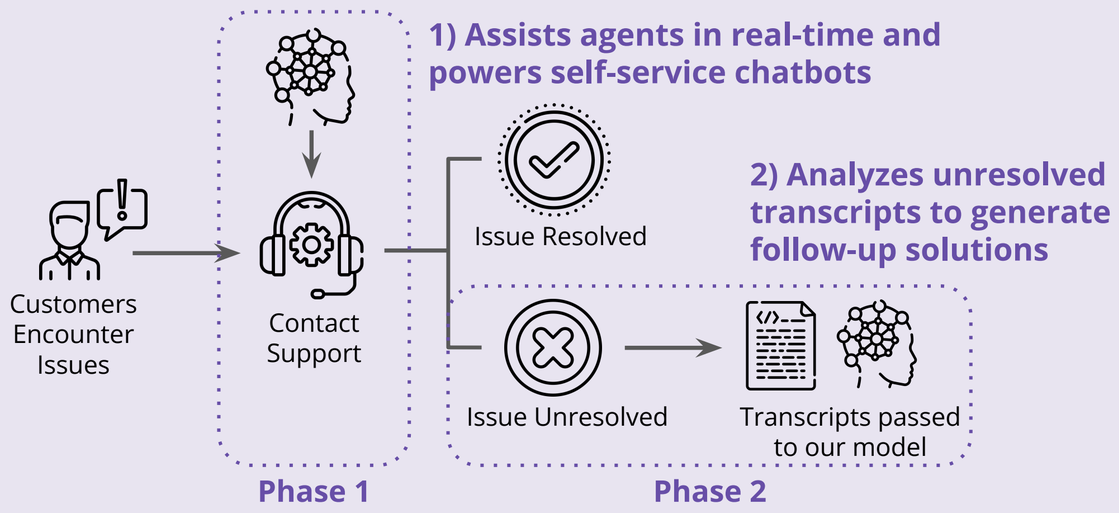
Diversify the Results

Add few unresolved cases for contrast and apply MMR (Maximal Marginal Relevance) to reduce redundancy.



USE CASES

Our AI engine can apply at two critical phases in the support flow:



RESULTS & IMPACT

By powering more accurate resolutions, our AI engine **drives up retention and cuts costs**:

+10pp
Increase in accuracy / resolution rate

+$30M
Annualized savings from less calls and churns

We also created **ecosystem values** for the data science team:

The catalog doubles as a knowledge graph for broader analysis.

A reusable framework that ties resolutions to financial impact.