

LEARNER WORKSHEET - Problem Decomposition (Week 3)

Project: Reddit Reply Assistant

STEP D1 - DEFINE DONE

What does the final output look like?

A natural, non-promotional Reddit reply that contextually engages with an existing thread related to AI Voice or Voice Datasets (my working company's offering), written in a conversational tone appropriate to the subreddit.

- Each output will be between 50-100 words
- Format: Text
- Shouldn't use salesy language
- No links unless already present in thread
- Must directly reference the thread's context

Output Contract / Schema:

```
{
  "thread_url": "",
  "subreddit_name": "",
  "topic": "",
  "text_size": "",
  "date_posted": "",
  "sentiment_score": "",
  "output_text": ""
}
```

STEP D2 - DUMP THE STEPS

List ALL actions that come to mind.

Do not worry about order or quality.

Messy Step List:

- Find relevant subreddits discussing AI Voice / Voice Datasets
- Identify suitable Reddit threads
- Scrape the thread contents, comments, thread URL
- Analyzing the content sentiment, intent and context
- Chunk and embed an existing Knowledge base in vectorDB about our services
- Define tone and persona
- Draft a contextual reply
- Validate tone, length, and non-promotional nature
- Store the text in a Google sheet

STEP D3 - DISTILL TO MVS (≤6 NODES)

Cluster, merge, remove noise, and name each node with a verb-first title.

Node	Verb-First Name	Why This Node Exists
N1	Collect Reddit Thread	Fetch thread + comments as raw data
N2	Analyze & Process Thread Context	Clean text + identify intent & sentiment
N3	Retrieve Relevant Knowledge	Pull only relevant service knowledge
N4	Generate Contextual Reply	Produce natural Reddit-style response
N5	Validate & Self-Check	Ensure tone, length, non-promo
N6	Export & Log Final Reply	Return structured output and store it

(If fewer than 6, even better.)

STEP D4 - DESIGN NODE TYPES

Choose one per node: **Agent**, **Workflow**, or **Human**.

Node	Type	Reason
N1	Workflow	Scraping / fetching
N2	Agent	Requires understanding context & sentiment
N3	Workflow	Vector retrieval and filtering
N4	Agent	Conversational, non-promotional generation
N5	Agent	Quality control & rule enforcement
N6	Workflow	Formatting output and logging to Google Sheet

STEP D5 - DECIDE ACCEPTANCE CHECKS

Define pass/fail rules for each node.

Node	Acceptance Checks
N1	Thread text + ≥ 5 comments fetched
N2	Intent identified + sentiment score present
N3	≥ 1 relevant knowledge chunk retrieved
N4	50–100 words, conversational, no CTA
N5	No sales language, matches subreddit tone
N6	JSON valid and Google Sheet row written successfully

STEP D6 - DELIVER THE TASK GRAPH

Fill in each field of the node card. This becomes your agent blueprint.

Node	Name	Type	Inputs	Acceptance Check	Artifact	Next
N1	Collect Reddit Thread	Workflow	thread_url	Thread + ≥5 comments	raw_thread	N2
N2	Process Thread Context	Agent	raw_thread	Intent + sentiment detected	context.json	N3
N3	Retrieve Relevant Knowledge	Workflow	context.json	≥1 relevant chunk	knowledge_chunks	N4
N4	Generate Contextual Reply	Agent	context + knowledge	50–100 words, no CTA	draft_reply	N5
N5	Validate & Self-Check	Agent	draft_reply	Non-promo, tone valid	validated_reply	N6
N6	Export & Log Final Reply	Workflow	validated_reply	JSON valid + Sheet updated	final_output.json	END

EXTRA SECTION - EDGE MAPPING

Draw arrows between nodes showing sequencing, dependencies, and loops.

N1 → N2 → N3 → N4 → N5 → N6

(if N5 fails → return to N3 once)

OPTIONAL SAMPLE WORKFLOW FOR PRACTICE

Choose one if you don't want to use your own:

- Competitor research
 - Pre-call brief generator
 - Blog → Carousel
 - Podcast → Insights → Tweet thread
 - Lead qualification
 - Social media research → content creation
-

SUBMISSION CHECKLIST

- ✓ Output contract defined
- ✓ Messy step-list created
- ✓ ≤ 6 distilled nodes
- ✓ Node types selected
- ✓ Acceptance checks written
- ✓ Task graph table completed
- ✓ Edges mapped