

- The same project might appear in different category types - include mentions across all categories

**\*\*SEARCH METHODOLOGY:\*\***

1. **\*\*Identify key entities\*\*** in bullet content: names, places, amounts, equipment, topics, processes
2. **\*\*Search ALL batches\*\*** for any mention of these entities in ANY category
3. **\*\*Group by content similarity\*\***, not category similarity
4. **\*\*Include 8-12 mentions\*\*** regardless of how they're categorized

Focus on complete topic clusters that enable summarization of:

- **\*\*Entity or Relationship Histories\*\***: interactions, developments, or changes related to a specific person, organization, group, or other identifiable entity across the entire plan (complete relationship or entity arc)
- **\*\*End-to-End Processes\*\***: steps, stages, or phases of a specific process, workflow, or methodology from initiation to conclusion across the entire plan (no missing steps)
- **\*\*Resource or Asset Lifecycles\*\***: mentions of acquisition, allocation, usage, modification, and outcomes for a specific resource, asset, or material across the entire plan
- **\*\*Decision and Strategy Journeys\*\***: details of decision-making, planning, and strategy development from problem identification through implementation across the entire plan
- **\*\*Problem/Challenge Resolution Narratives\*\***: instances of identifying, analyzing, addressing, and resolving a particular issue or challenge across the entire plan
- **\*\*Timeline-Driven Developments\*\***: events and updates showing chronological evolution of a specific project, initiative, or topic across the entire plan
- **\*\*Knowledge or Skill Development Sequences\*\***: progress updates, milestones, and learning activities related to acquiring or improving a specific skill or knowledge area across the entire plan
- **\*\*Discussion and Agreement Processes\*\***: discussions, debates, negotiations, and agreements related to a specific matter across the entire plan

Prioritize bullet point groups that:

- Contain rich, interconnected information suitable for synthesis
- Enable questions like "Can you summarize my interactions with X?" or "Summarize the [X] process"
- Include both factual/quantitative details and qualitative/narrative elements for well-rounded summaries
- Have varying complexity levels (simple single-topic vs. complex multi-faceted stories)
- Allow testing of information condensation across multiple conversation sessions
- Include both factual details and narrative elements for comprehensive summarization
- Create opportunities to test synthesis of scattered information into coherent narratives

Return your analysis in this exact JSON format where each object contains 8-12 related bullet points:

```
[{"capability": "summarization", "batch_numbers": "1, 1, 2, 3, 4, 5", "bullet_numbers": "5, 22, 18, 19, 23, 31",
  "bullet_points": ". **Character & Relationship:Close Friend:** Jeremiah, 37, met at festival, recommended consultant. | . **Conflict & Resolution:Relationship Boundaries:** Jeremiah requested exclusive studio access; agreed to 3 hours only. | . **Character & Relationship:Close Friend:** Jeremiah helped install MLV, bringing snacks from bakery. | . **Character & Relationship:Close Friend:** Jeremiah invited me to music club to test prototype. | . **Character & Relationship:Close Friend:** Jeremiah brought dinner during late work session. | . **Goals & Progress:Milestone Celebration:** Hosted listening party with Jeremiah, Jon, and Tonya ."}]
```

Important formatting notes:

- The "batch\_numbers" and "bullet\_numbers" correspond to each other positionally
- "1, 1, 2, 3, 4, 5" and "5, 22, 18, 19, 23, 31" means: Batch 1 Bullet 5, Batch 1 Bullet 22, Batch 2 Bullet 18, Batch 3 Bullet 19, Batch 4 Bullet 23, Batch 5 Bullet 31
- Each bullet point separated by " | "
- Use comma-separated values for batch\_number and bullet\_number
- Vary group sizes based on topic complexity and richness
- Focus on groups that enable comprehensive summarization questions
- Include both simple single-topic and complex multi-topic groups
- **\*\*NO LIMIT on number of bullet points\*\*** - include as many as needed for complete coverage

**CRITICAL NOTES:**

- **\*\*ANALYZE BULLET CONTENT, NOT CATEGORY NAMES\*\***: Search for mentions of entities in the actual text
- **\*\*IGNORE CATEGORY LABELS\*\***: The same entity mentioned in different category types should all be grouped together
- **\*\*COMPREHENSIVE ENTITY SEARCH\*\***: For each entity/topic/process, scan ALL batches and ALL categories for any mention
- Include 8-12 mentions regardless of bullet point category if they reference the same entity/topic in the content

Select 7-9 groups of bullet points with COMPLETE mention coverage that would enable the most sophisticated and comprehensive summarization questions across different complexity levels.

NOTE: Only output the list without any explanation before or after the list.

## Listing 8: Candidate selection summarization prompt

This is a plan that contains detailed bullet points about a topic. This plan is used to generate realistic chat conversations between a user and an AI assistant, which are then used to evaluate the long-term memory capabilities of LLMs.

Your task is to analyze this plan and select bullet points that would be most effective for testing instruction following abilities when incorporated into chat conversations.

Analyze this plan and identify bullet points that contain user instructions ideal for testing whether the LLM remembers and follows user-given instructions.

**\*\* INPUT DATA**

- **\*\*PLAN\*\***: <plan>

Focus on bullet points with:

- **\*\*User Instruction\*\*** category/label
- **\*\*Explicit instruction statements\*\***: "Always", "Never", "When I ask about X, do Y"
- **\*\*Behavioral directives\*\***: How the AI should respond or behave