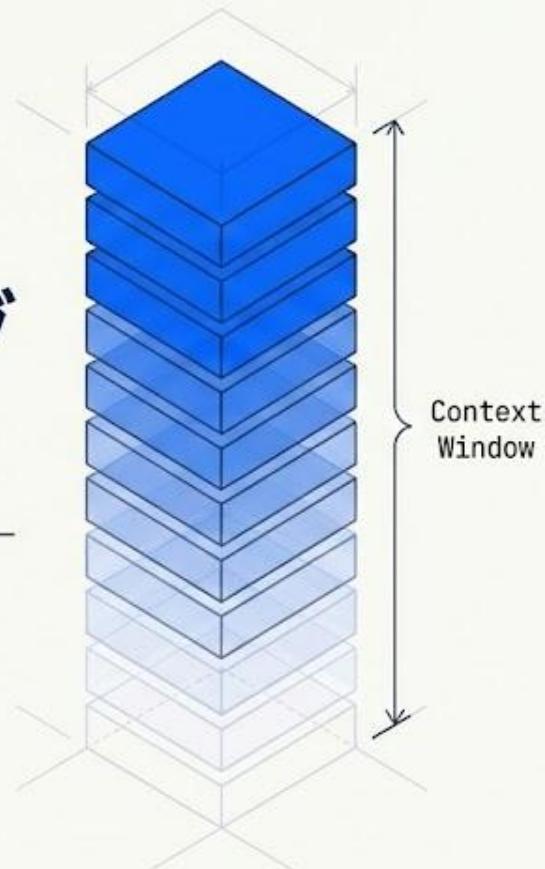


コンテキストエンジニアリング

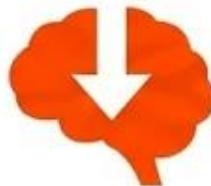
記憶の寿命と管理術

～なぜAIは長話で「精度」が落ちるのか？
コンテキストウィンドウとスレッド分割～



開発現場で起きる「AIの劣化」現象

これらはバグではありません。仕様による必然的な結果です。



IQの低下

チャットが長引くと、初期の重要な指示（「英語禁止」など）が無視され、回答の質が下がる。



フォーマット崩壊

JSONやCSV形式での出力が、途中から自然言語混じりの雑なものに変わり、パースエラーを起こす。



空にしたはずのごみ箱

「リセットした」「話題を変えた」つもりでも、壁打ちで溜まった前提是完全には消えない。見えないゴミが、次の答えに混ざり込む。



修正の泥沼化

「修正して」と指示を重ねても、一度染み付いた思考のクセや悪い前提が完全には抜けず、スレッド内での教育・修正が効かなくなる。

【定義】そもそも「コンテキスト」とは何か？

コンテキストとは、単なる
「会話履歴」ではありません。

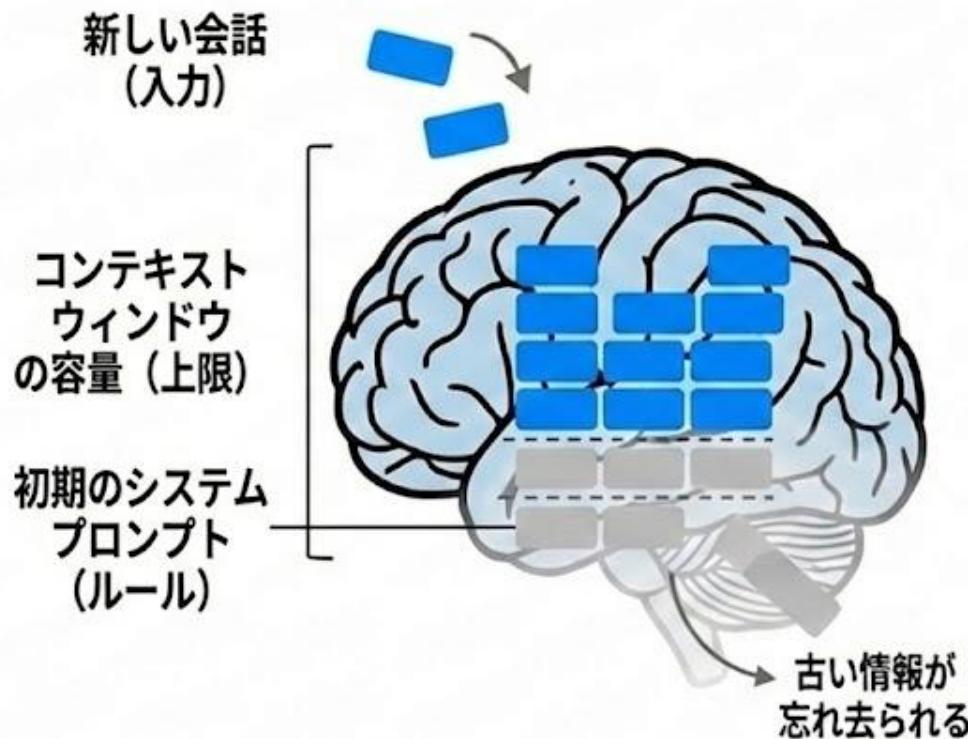
AIが次のトークンを生成する
際に参照する、その時点で
与えられているすべての入力
情報の集合です。



誤解と正解

✗	誤解：AIは人間のように「前の会話」を脳に記憶している。
✓	正解：AIは毎回、「現在の入力範囲（コンテキストウィンドウ）」を参照し、その中から確率的に次のトークンを生成している。

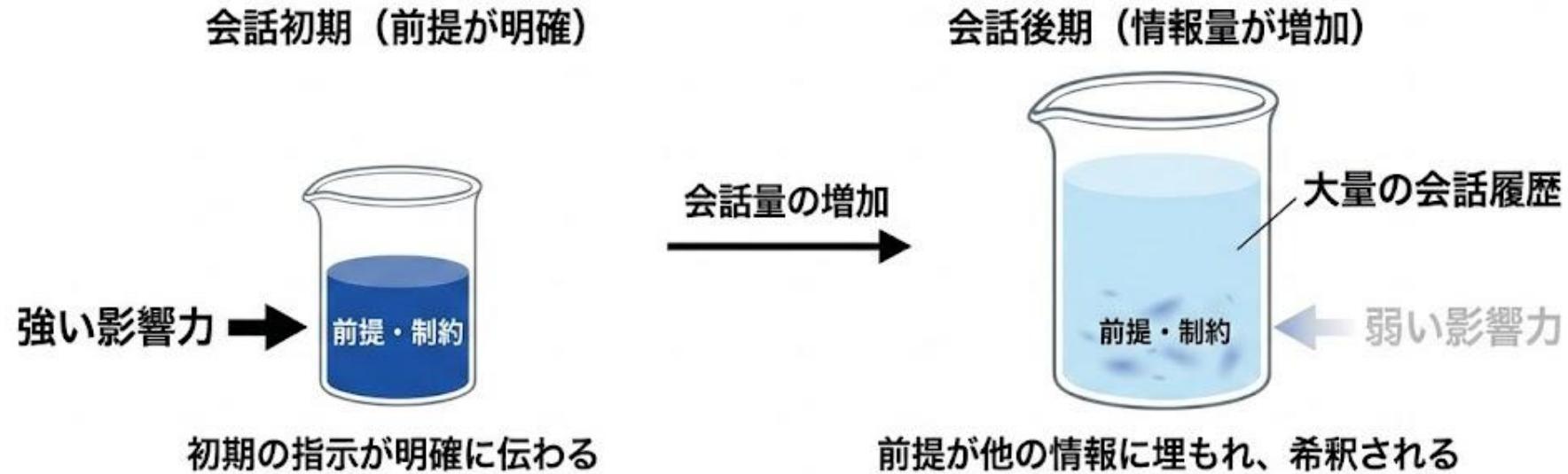
【原理】「コンテキストウィンドウ」と記憶の限界



なぜ「約束」を忘れるのか？

- ✓ LLMには一度に処理できる情報量（トークン数）の上限がある。
- ✓ 新しい会話、情報を入れつづけるとパンクして古い情報を忘れてしまう
- ✓ 結果として、最初に定義した役割や出力ルールが参照されなくなる

【課題】会話量の増加で起きる「Context Dilution（前提の希釈）」



トークン上限に達していなくても、精度は劣化します。チャットが長くなり入力情報が増える について、過去に共有した前提や制約が相対的に影響力を失い、前の内容が参照されにくくなるContext Dilutionが起きるためです。

【対処法】新規スレッドによる再初期化

Bad: 複雑化したコンテキスト (Contaminated Context)



- ・ノイズの蓄積
- ・前提の希釈・残留
- ・指示の優先度が不明確

精度が不安定 / 修正困難

Good: 新規スレッド (Clean Context)

- ・ノイズがない
- ・初期指示が明確
- ・注意配分が安定

再現性の高い出力



ノイズや誤った前提が蓄積した状態では、後から修正を重ねても期待した挙動に戻らないことがあります。そのような場合は、新しくチャットを作成してコンテキストをきれいにしてから使用することが重要です。

でも、毎回リセットするのは大変



ここまで見てきたのは、会話が進むにつれてコンテキストがどのように劣化し、なぜ新しくチャットを作るのが有効なのかという「仕組み」の話です。

ただ、

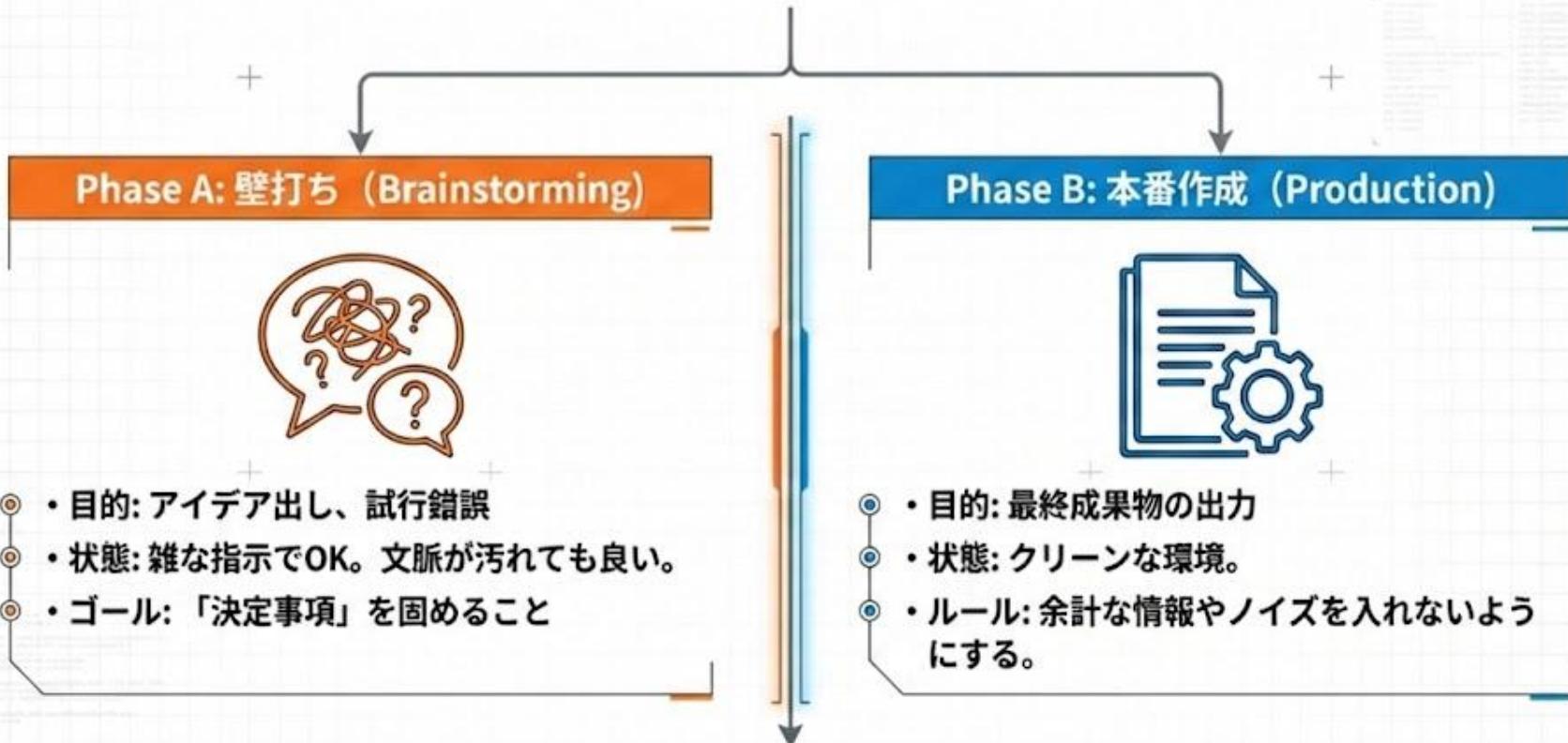
毎回、新しいチャットを作って前提を書き直すのは、正直しんどい。

次からは、

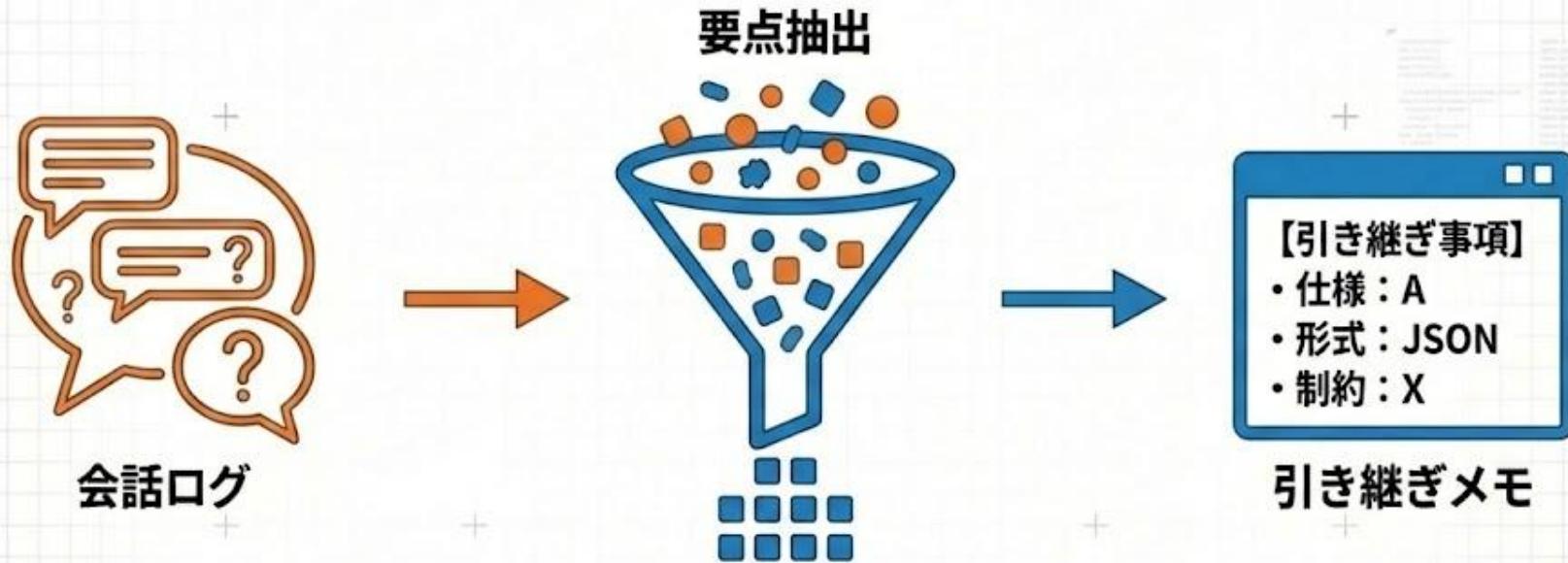
新しいチャットを前提としつつ、コンテキストをうまく活用するための実践的な運用方法を考えてみました。

【実践①】運用フロー：「壁打ち」と「本番」を分ける

1つのスレッドですべてを完結させようとしない



【実践②】要点の抜き出しと引き継ぎ



「壁打ち」スレッドで良い結果が出したら、そのチャットを使い続けるのではなく、重要な決定事項だけを抜き出します。抜き出した要点は、次のチャットに引き継ぐためのメモ（引き継ぎ事項）として保存し、新しいチャットの前提として使用します。

【実践③】前提を明示して再スタートする (Grounding)

新しいチャットを開き、実践②で整理した「決定事項」を冒頭にまとめて渡します。

このように、前提を明示した状態で作業を再開することを
グラウンディング (Grounding) と呼びます。

Prompt Input Box

--- Role ---

あなたは社内コミュニケーションの専門家です。分かりやすい説明資料を作成してください。

--- GROUNDING CONTEXT ---

以下の【前提（引き継ぎ事項）】に基づいてください。

【前提（引き継ぎ事項）】

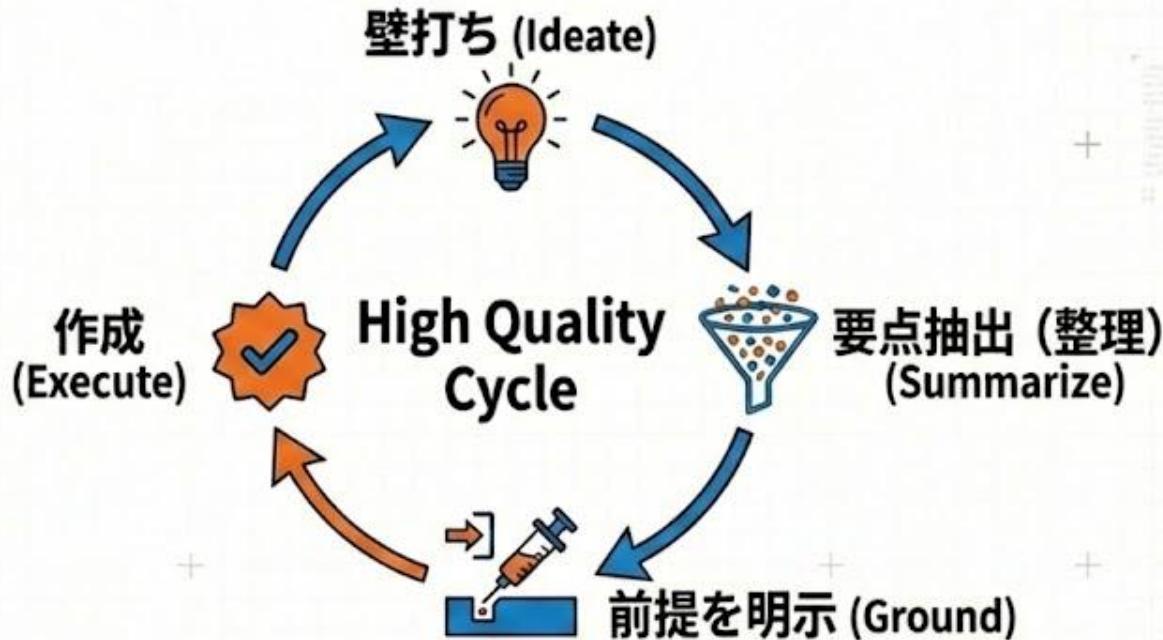
- ・対象：非エンジニア
- ・目的：社内向け説明
- ・トーン：分かりやすく、簡潔
- ・文字数：800字程度

前提を最初に
まとめて渡す

--- TASK ---

この前提に基づいて、文章を作成してください。

【総集編】コンテキスト管理のサイクル



「記憶管理は完璧になりました。しかし、AIはクリアな頭脳で『もっともらしい嘘』をつきます。」

次回、【第3回】AIの「嘘」と「確率」の正体（ハルシネーション対策）

