

# プロジェクト概要： 大規模テキスト情報の構造化と視覚化

Team: Scaler  
Singularity Society Bootcamp #3  
May 2025 ~ March 2026

# ソフトウェアの概要

- 大量の情報を、情報の粒度(詳細度)に応じて構造化  
-> テキスト中心で、後に動画、画像、音声などにも拡張
- 構造化した情報を、インタラクティブな2D/3D空間に視覚的に表現
- 従来の線形的なテキスト情報が持つ「情報のオーバーロード」や「全体像の把握の困難さ」といった課題を解決
- AIによる要約・分析機能を統合し、膨大な情報から迅速に必要な知識・記憶を抽出

# 抽象的な理念・世界観

## 1. 拡張: 情報の“広さ”への追求

- どのように個人が本質的に知りたい情報を発見させるのか？
- “人間は、見たものを欲求する。欲求の対象を意図的には探さない。”<sup>1</sup>

## 2. 収縮: 情報の“深さ”への追求

- どのように情報の深度や詳細度を深めていくのか？
- 追求すべき情報への深掘り、不必要な情報の削減

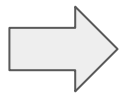
## 3. 回帰: 思考過程の可視化

- どのような思考プロセスを経て、AIとの会話・質問を行ったのか？
- どのような思考過程で、どのような解決を見出せたのか？



# 課題 & ソリューション

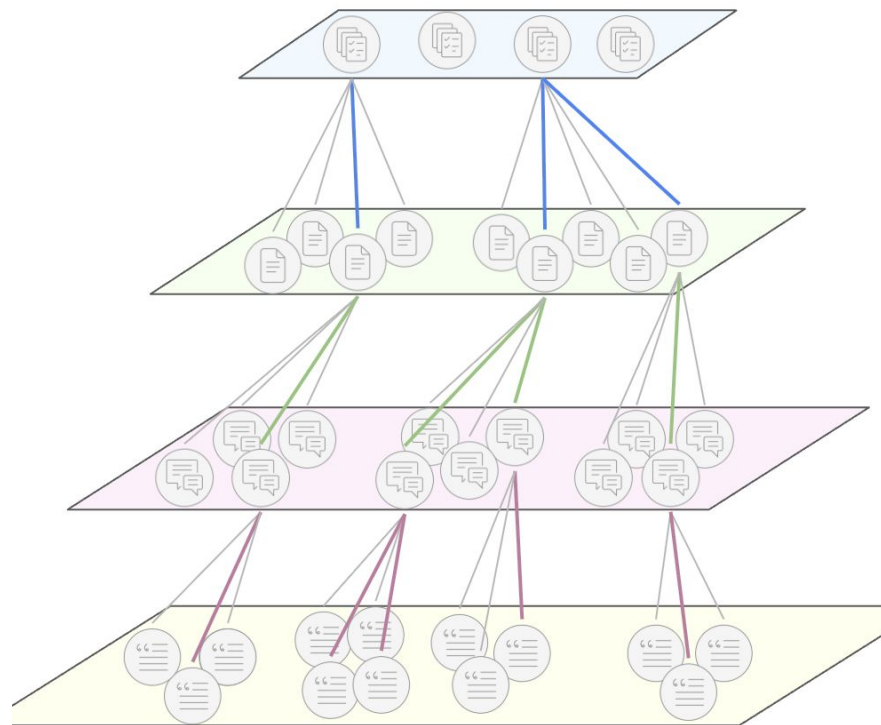
- **線形的な情報** : 章/節での階層化(目次)でも、情報間の関連性や全体像の直感的な把握が困難  
=> コンテキスト間の関連性と全体像をグラフで表現、ミドルレイヤー (サマリ・クラスタ)の導入
- **情報過多** : あらゆる粒度の情報が同じ空間・平面に混在、不必要な情報が露呈・認知  
=> ベクトル空間を情報の詳細度別の多層構造を実現、情報の“広さ・深さ”のキレを強調
- **困難な記憶の整理** : 一時的な暗記に頼ったことによる非一貫的な情報管理  
=> 拡張性のある階層構造、“流れ”で記憶をたどることでデータ(事実)を引き出す
- **思考過程の保管** : 一般的なAIチャット履歴では、自分の思考過程の整理が不可能  
=> Git形式でのチャット履歴の管理、各ノードにはその時点の状態を記録



グラフと共に行う段階的な思考、読解、理解  
探求欲を刺激、要点の掘り下げ、事実を提示

# 手法・アプローチ

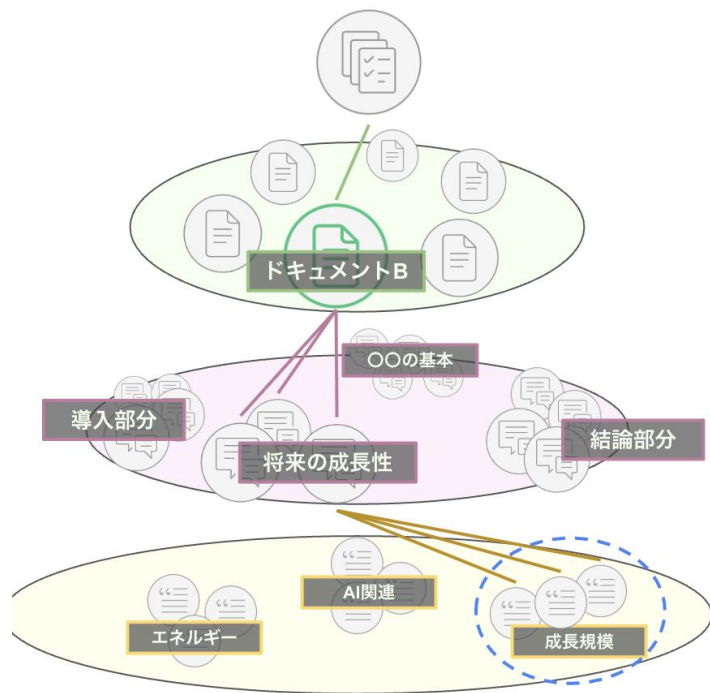
- ベクトル化されたテキストデータを、情報の粒度に応じて異なる空間に階層的に配置
- ユーザーは、まず全体を俯瞰し、そこから興味のある部分へと段階的にドリルダウン
- これにより、効率的な知識の探索と深い理解を支援する、認知科学的なアプローチを提供します。



# アーキテクチャ

- **『マイクロサービス』:**  
ドキュメント、AI&チャット、グラフキャンバス、ベクトル検索などの独立したサービス
- **外部サービス:**  
機械学習(Python), SQL(Postgre), NoSQL(DynamoDB), ベクトルDB(Qdrant), LLM(Gemini API, OpenAI, Claude), クラウド(AWS), 認証(Clerk),
- **ネットワーク:**
  - Public API: RESTful + API Gateway (Kong or AWS)
  - フロント<=>バック: RESTful + WebSockets
  - サービス間: gRPC
- **プログラミング言語:**
  - フロント: NextJS
  - バック: Go(メイン)、Python(機械学習・データ分析系)

# 機能(1): ユーザーの視覚と階層的空間の連携



ドキュメント内の文章と構造化ダイアグラム上での俯瞰情報とのマッチング実現

## ドキュメントB

導入：

\_\_\_\_\_

本文：

1. ○○とは？

\_\_\_\_\_

2. ○○の詳細

\_\_\_\_\_

3. ○○の将来

\_\_\_\_\_

まとめ：

\_\_\_\_\_

カーソル

- ☐ ドキュメントA
- ☒ ドキュメントB
- ☐ ドキュメントC
- ☐ ドキュメントD
- ☐ ドキュメントE
- ☐ ドキュメントF



ユーザーの視覚