S.A.G.E.

Smart Assistant for Guided Enablement

GW Business

TEAM 3 – Section 80

Meet The Team

Mason Levitt

Ambuja Sharma

Pramodh Kaushik Mandem

Himavarsha Yerrapothu

Usama Mohammad

S.A.G.E.
Smart Assistant for Guided Enablement

Agenda

1. Executive Summary
2. Business Need
3. Solution Overview
4. Data & Technical Process
5. Live Demo
6. Value Add for Businesses
7. Next Steps
8. Appendix

AI-enabled project onboarding transforms lost time into billable value and faster client readiness.

Overview

- **Situation**
  - Onboarding is a critical but often overlooked driver of project quality, early momentum, and client trust.
  - Consulting teams rely on rapid project onboarding for client delivery.
  - Key knowledge dispersed across docs, drives, emails, and internal tools.
- **Complication**
  - Consultants spend weeks coming up-to speed on new/existing projects.
  - Manager dependence and information dispersion slows productivity and consistency.

Our Impact

- **Key Question**
  - How can AI cut onboarding time and standardize knowledge transfer?
- **Solution**
  - An expert and interactable knowledge source trained specifically on project data across formats and sources.
  - Delivers short answers, context narratives, and visual outputs, calibrated across multiple data repositories.
  - Operated using a B2B SaaS model

S.A.G.E.
Smart Assistant for Guided Enablement

Inefficient project onboarding delays mobilization, consumes leadership bandwidth, and erodes institutional knowledge across teams.

Operational loss from redundant ramp-up[1]

- Teams spend **~127 hours per year** just recovering focus after interruptions, rather than diving straight into project value.[4]
- Constant switching between tasks/tools (context switching) can cost up to **40% of productivity**, delaying ramp-up.[5]

Managerial loss via bandwidth drain and oversight loss

- Managers spend **10–20% of weekly time** re-explaining project context instead of leading delivery.[2]
- Early attrition from poor onboarding and team connectedness can cost firms up **to 200% of an employee's annual salary**.[6]

Financial impact of inefficient onboarding[3]

Cost Drivers	Annual Impact
Ramp-Up Inefficiency	**$99K**
Consultant Productivity Loss	**$312K**
Managerial Time Loss	**$180K**
Total Estimated Loss	**$591K**

All estimates are based on assumptions of a firm with 50 consultants ($65/hr), 10 managers ($90/hr), and teams of 5 consultants + 1 manager completing 6 projects per year. ***We recognize these are low-end estimates,*** *but we want to be inclusive of usage from all firms. More details of cost estimation in appendix.*

1 Boston Consulting Group (2023); 2 Harvard Business Review (2025); 3 Society for Human Resource Management (SHRM) (2023); 4 Dropbox (2023); 5 Drag App (2025); 6 SHRM (2025)

S.A.G.E. Smart Assistant for Guided Enablement

By simulating synthetic data for an education consulting project for GW, we can replicate onboarding onto to a real consulting engagement.

Our Approach

Our approach required us to have rich and in-depth data regarding:

1) Our Principal Client: Kestrel Education

2) Our Secondary Client: GWU (Kestrel's Project)

The performance of the onboarding assistant is directly tied to how it can sift through complex and vast data sources to find the most relevant and effective answer to the user's query.

Data Produced

To train and test our agent, we produced the following data:

- Foundational Documents
- Cadence Documents
- Client Deliverables
- Governance and Controls
- Quantitative Datasets

Desired Outputs

Situation	Expected Output
Basic Administrative Facts	Short-form text and basic information on retrieval
Context-Driven Narratives	A detailed visualization with enriched graphics
High-level Analytics	A combination of text and basic visuals if necessary to supplement
Trend-based visualizations	A detailed visualization with enriched graphics

Outputs focused on project-specific details produced from synthetic and externally sourced data sources.

Synthetic Data

Foundational Documents (Project Charter, SoW, Stakeholder Map)

Weekly Cadence (Standups, 1-1s, Memos)

Client Deliverables (Deliverable summaries, decks)

Governance (decision log, change logs)

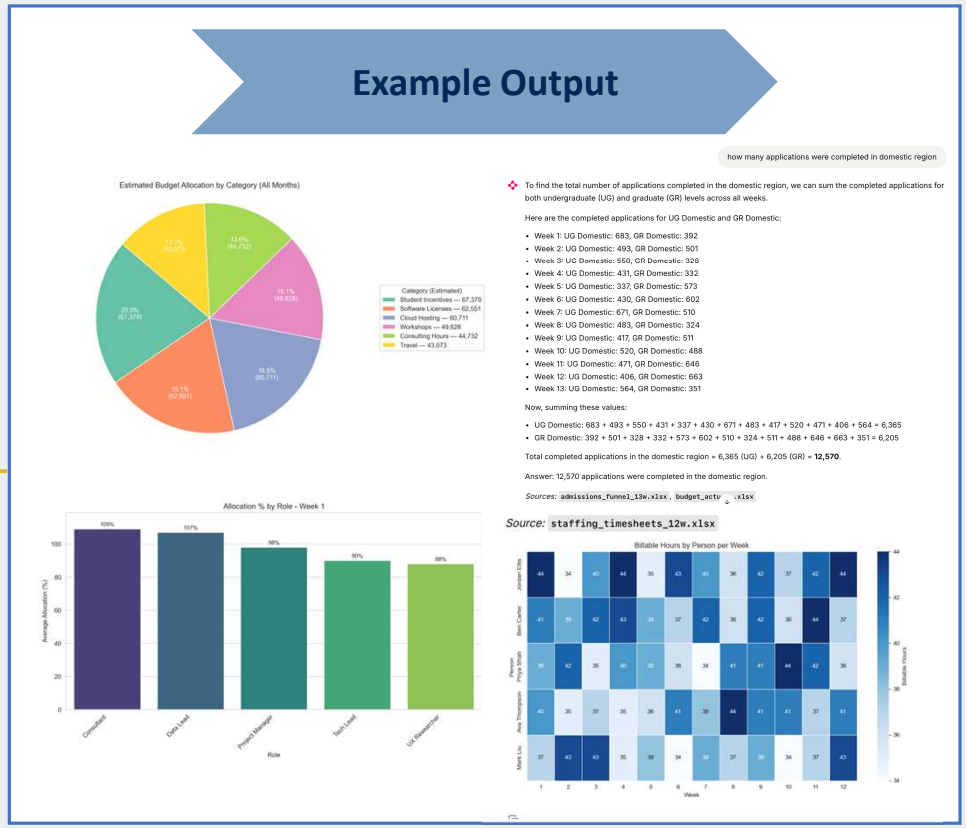
Quantitative Datasets (Admissions funnel, staffing timesheets)

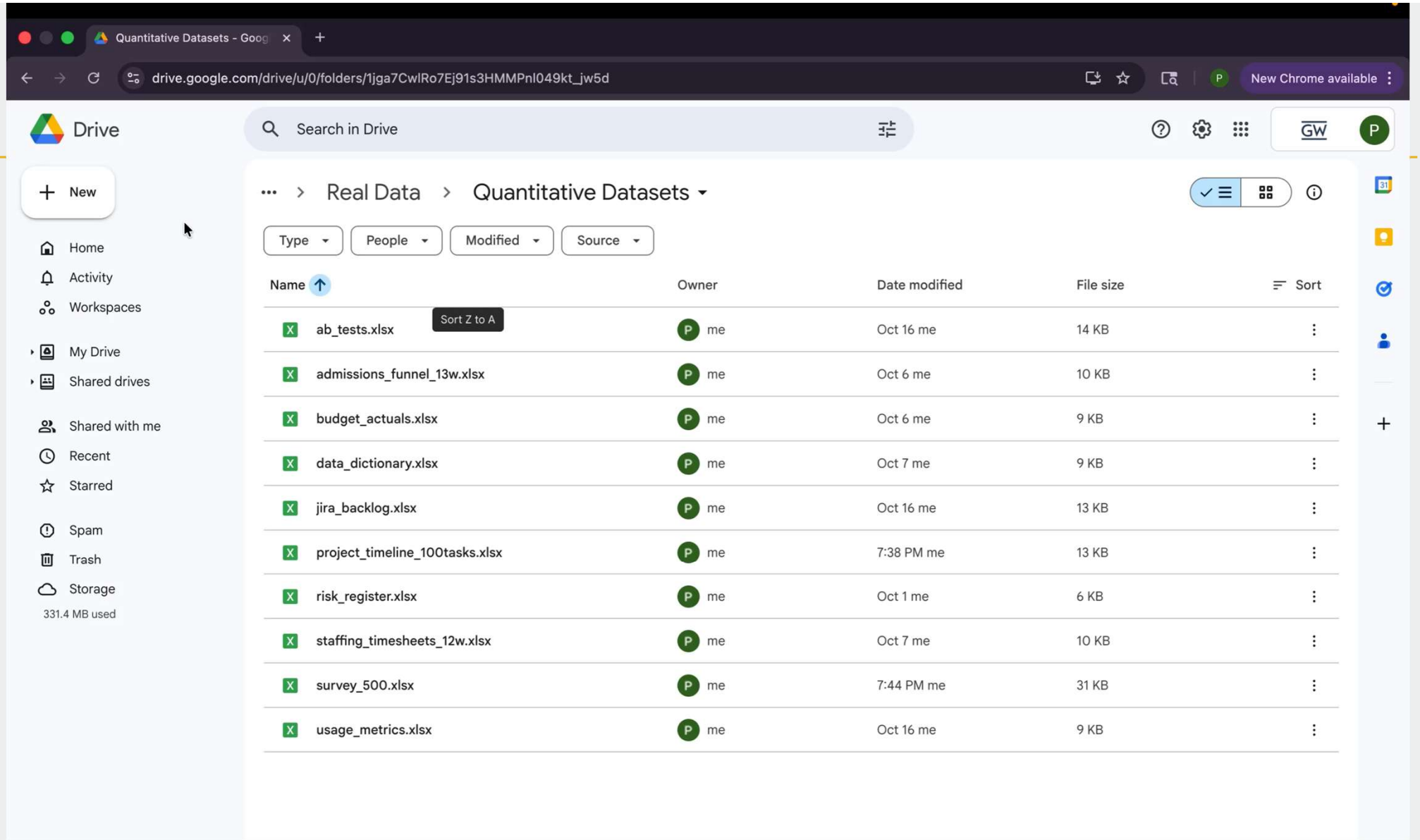
Miscellaneous (Consultant Guidebook)

Sourced Data

1. US Undergraduate Enrollment Statistics
2. US Undergraduate Test Scores
3. US Undergraduate Financial Aid and Tuition Statistics

Example Output




Quantitative Datasets - Goog
drive.google.com/drive/u/0/folders/1jga7CwlRo7Ej91s3HMMPnl049kt_jw5d
New Chrome available
Drive
Search in Drive
GW
New
Home
Activity
Workspaces
My Drive
Shared drives
Shared with me
Recent
Starred
Spam
Trash
Storage
331.4 MB used
Real Data
Quantitative Datasets
Type
People
Modified
Source
Name
Sort Z to A
Owner
Date modified
File size
Sort
ab_tests.xlsx me Oct 16 me 14 KB
admissions_funnel_13w.xlsx me Oct 6 me 10 KB
budget_actuals.xlsx me Oct 6 me 9 KB
data_dictionary.xlsx me Oct 7 me 9 KB
jira_backlog.xlsx me Oct 16 me 13 KB
project_timeline_100tasks.xlsx me 7:38 PM me 13 KB
risk_register.xlsx me Oct 1 me 6 KB
staffing_timesheets_12w.xlsx me Oct 7 me 10 KB
survey_500.xlsx me 7:44 PM me 31 KB
usage_metrics.xlsx me Oct 16 me 9 KB

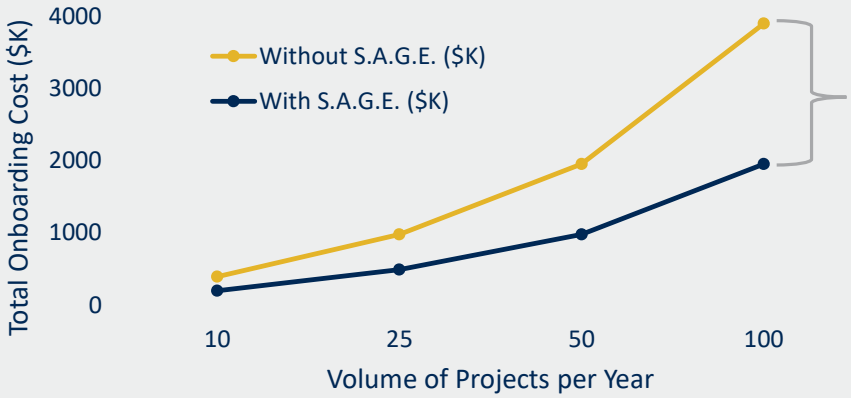
Live Demonstration

AI-driven project onboarding converts ramp-up time into realized value through decreased labor hours and cost reductions.

Estimated Time and Value Savings per Project

Role	Baseline (hrs)	After Bot (hrs)	Hours Saved	Hourly Rate	Savings
Consultant (× 4)	60 hrs	20 hrs ea.	160 hrs	$65/hr	$10,400
Manager (× 1)	40 hrs	10 hrs	30 hrs	$90/hr	$2,700
Partner (× 1)	10 hrs	5 hrs	5 hrs	$150/hr	$750
Total per Project			**195 hrs**		**≈ $13.9 K**

Savings Extrapolated Across a Firms Portfolio



At 100 projects per year, firms would **save $1.95M in labor costs and 19,500 hours** of recovered delivery capacity with S.A.G.E.

Drivers of Time Savings

- **Consultant Efficiency Gains:** Automated search and context retrieval cut onboarding time by about **65%**, consistent with onboarding productivity data.[1]
- **Manager Bandwidth Recovered:** Managers spend **15–20%** of their week re-explaining context; automation reduces this by **75%**.[2]
- **Partner Oversight Reduced:** AI-generated summaries shorten senior review time by **50%**, aligning with mobilization efficiency benchmarks.[3]

1 StrongDM (2025); 2 Harvard Business Review (2025); 3 Saviom (2025); 4 Enboarder (2024)

S.A.G.E. Smart Assistant for Guided Enablement

Next steps include defining security, integrating with enterprise technology tools, and establishing B2B market positioning.

Product Readiness – Defining Security

Define baseline security model (SSO, RBAC, audit logging, data retention). Outline approach for tenant isolation and document-level access control.

Map future requirements to common standards (SOC 2–style controls, GDPR alignment).

Platform Integration – SharePoint & Enterprise Stack

Design a SharePoint connector to index sites, libraries, and permissions-aware content. Prototype ingestion of SharePoint docs alongside existing drives to keep one unified corpus.

Plan extensible integration layer for additional systems (e.g., Teams, Jira, CRM).

Go-to-Market – B2B SaaS

Position S.A.G.E. as a multi-tenant B2B SaaS onboarding layer for project-based teams. Start with targeted pilots for consulting services firms, priced per seat or per project.

Build case studies around time-to-ramp, manager bandwidth saved, and ROI to support broader rollout.

Thank you!

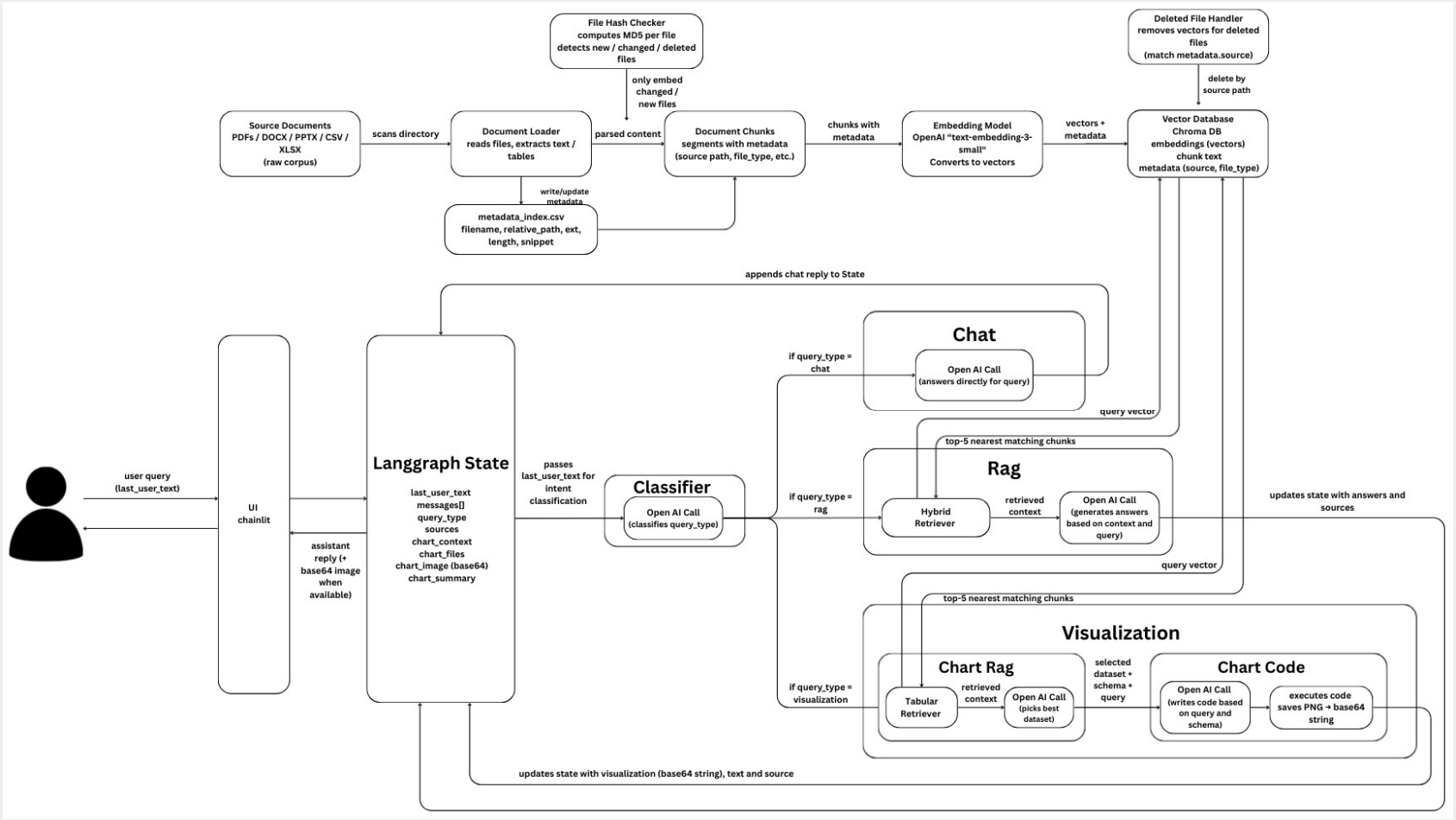
Questions?

GW Business

S.A.G.E.
Smart Assistant for Guided Enablement

Appendix

Technical Architecture Diagram



Financial Impact Analysis of Inefficient Onboarding

Financial impact of inefficient onboarding[3]

Cost Driver	Assumption	Source of Assumption	Annual Impact
Ramp-Up Inefficiency	127 hours per team[4] ÷ 5 consultants/team x $65/hr/consultant x 10 teams/company x 6 projects/year/team	• Avg. 127 hrs/yr lost to focus recovery after interruptions[4] • Based on typical 5-person teams completing 6 projects/year	$99K
Consultant Productivity Loss	40% productivity loss[5] x 6 onboarding weeks/year x 40 hours/week x 50 consultants x $65/hour/consultant	• Context switching reduces productivity up to 40%[5] • Assumption of 1 onboarding week per project (6 projects per year)	$312K
Managerial Time Loss	10% productivity drag[2] x 2000 hours (1 FTE) x $90/hr/managers x 10 managers	• Managers spend 10–20% of time re-explaining project context[2]	$180K
Total Estimated Loss			**$591K**

All estimates designed off assumptions of a firm with 50 consultants ($65/hr), 10 managers ($90/hr), and teams of 5 consultants + 1 manager completing 6 projects per year.

1 Boston Consulting Group (2023); 2 Harvard Business Review (2025); 3 Society for Human Resource Management (SHRM) (2023); 4 Dropbox (2023); 5 Drag App (2025); 6 SHRM (2025)

S.A.G.E. Smart Assistant for Guided Enablement

Implementation and Security

Tech Implementation

- MD5 hash checker detects new/changed/deleted files and updates vectors automatically.
- Document loader ingests PDFs, DOCX, PPTX, CSV, XLSX → segmented into metadata-rich chunks.
- Embeddings generated using OpenAI *text-embedding-3-small* and stored in ChromaDB.
- LangGraph state machine classifies queries (Chat, RAG, Visualization) and manages conversation flow.

Privacy/Security

- File Integrity – MD5 hash checker and deleted file handler keep only current, authorized files in the system.
- Metadata Control – Each document chunk carries source/path metadata, enabling precise tracking and controlled access.
- Secure Execution - Visualization code runs in a sandboxed environment to prevent unsafe or malicious actions.
- Encryption - Data stored in ChromaDB is encrypted, and retrieval ensures only relevant, authorized chunks are accessed.

File Transfer

- Secure uploads via encrypted channels (HTTPS/TLS).
- Supports multiple formats: PDF, Word, PowerPoint, Excel, CSV.
- Auto-indexing embeds uploaded files into vector database for immediate retrieval.
- Deleted file handler removes outdated vectors to maintain accuracy.

Technical Deepdive

System Architecture

LangGraph State
- ✓ Classifier routes queries to 3 specialized nodes
- ✓ Chat, RAG, and Visualization pipelines
- ✓ Stateful conversation management

Chainlit Frontend
- ✓ Async message handling
- ✓ Real-time chart rendering
- ✓ Session-based state persistence

Vector Database
- ✓ ChromaDB with 501indexed documents
- ✓ Hybrid retrieval (semantic + keyword)
- ✓ Auto-sync with incremental updates

Hybrid Retrieval System

Dual-Mode Retrieval
- ✓ Semantic: OpenAI embeddings (text-embedding-3-small)
- ✓ Keyword: BM25 algorithm for exact matching
- ✓ Weighted ensemble (50/50 or 40/60 for tabular)

Specialized Retrievers
- ✓ General RAG: Balanced hybrid (5 docs, MMR)
- ✓ Tabular: 70% keyword weight for CSV/XLSX precision
- ✓ 150 tabular documents optimized for charts

Search Strategies
- ✓ MMR for diversity (avoids duplicates)
- ✓ Similarity for precision (exact matches)
- ✓ Query expansion with domain synonyms

AI-Powered Visualization

Dynamic Code Generation
- ✓ GPT-5 generates matplotlib/seaborn code
- ✓ Schema-aware: analyzes column types & data
- ✓ Automatic chart type selection

Execution Pipeline
- ✓ Retry logic with error-aware regeneration
- ✓ Base64 encoding for web display

Smart Dataset Selection
- ✓ Tabular-aware retriever prioritizes spreadsheets
- ✓ Context-rich prompting with data previews

Competitor Analysis

Competitor	Description	Opportunity for S.A.G.E.
Guru[1]	• Centralizes team knowledge into curated cards, wikis, and searchable collections. • AI answers workflow and process questions using structured content.	• Requires teams to manually build cards instead of autonomous response generation. • Weak integration with SharePoint and Google Drive repositories.
Notion AI[2]	• Workspace combining docs, wikis, tasks, and internal hubs with built-in AI tools. • Strong for team collaboration and working with uploaded content.	• Limited ingestion of legacy files; teams must rebuild content inside Notion or carefully adjust file types to use full scope of AI toolkit. • AI features struggle with large enterprise repositories.
Confluence[3]	• Documentation platform for meeting notes, requirements, and structured project pages. • Integrates with Jira to support engineering-focused workstreams.	• AI can only summarize pages stored in Confluence; ineffective on any external drives or slide decks. • Requires extensive manual page creation and upkeep on workflows.
Loopio[4]	• AI-supported content library built for automating RFP responses and proposal creation from a pool of pre-existing documents. • Simplifies answer reuse and content assembly for sales organizations.	• Not built for onboarding but instead focused solely on RFP workflows instead of overall project context. • Cannot parse client deliverables or extract historical knowledge.
Dock.US[5]	• Provides shared client hubs for onboarding, project plans, and deliverable exchanges. • Offers robust templates, checklists, and CRM tools for users.	• Designed for client handoffs, not internal onboarding services. • Limited ability to search, interpret, or unify internal documents stored across multiple systems or repositories.

1 Guru (2025); 2 Notion (2025); 3 Atlassian (2025); 4 Loopio (2025); 5 Dock (2025)

S.A.G.E. Brand Package

S.A.G.E.
Smart Assistant for Guided Enablement

S.A.G.E
Smart Assistant for Guided Enablement

Role / Usage	Color	Color Name	HEX Code	Description
Primary Brand Color		Navy Blue	#002855	Professional, trusted base tone for text and icons
Secondary Accent Color		Buff Gold	#E4B429	Highlight color used for underline and accents
Neutral / Supporting Color		Fog Gray	#E5E6E7	Background or neutral tone
Text / Body Copy Color		Cool Gray	#A7A8AA	Subtle supporting text color
Optional Light Background		White	#FFFFFF	Clean background for contrast and clarity

S.A.G.E. palette is inspired by the GW Office of Communications and Marketing[1] palette and logos were designed with support of ChatGPT.

1 George Washington University