



The ultimate AI business intelligence platform

<https://good-bi.vercel.app>

<https://github.com/Vshnv2001/good-bi>

| <b>Names</b>          | <b>Matriculation Number</b> |
|-----------------------|-----------------------------|
| Lok Jian Ming         | A0236537Y                   |
| Ong Pei Cong Lester   | A0245751A                   |
| Shirshajit Sen Gupta  | A0244147H                   |
| Vaishnav Muralidharan | A0235268Y                   |

# Table of Contents

|   |           |
|---|-----------|
| <b>Prologue: Known Issues.....</b>  | <b>4</b>  |
| <b>Milestone 0: Problem Description.....</b>  | <b>4</b>  |
| <b>Milestone 1: Competitors.....</b>  | <b>5</b>  |
| ClickUp.....  | 5         |
| Teradata.....   | 6         |
| Power BI.....   | 7         |
| Our Product.....  | 8         |
| GoodBI vs ClickUp: Addressing a Need for Dedicated BI Automation.....                               | 8         |
| GoodBI vs TerraData: Simplifying Complex Data Analysis for Small to Medium Enterprises.....         | 9         |
| GoodBI vs Power BI: Combining Simplicity with Customization.....                                    | 10        |
| Market Need Tackled by GoodBI.....  | 11        |
| <b>Milestone 2: Application Description.....</b>  | <b>12</b> |
| Overview.....   | 12        |
| User Stories.....   | 12        |
| <b>Milestone 3: Tech Moat.....</b>  | <b>13</b> |
| <b>Milestone 4: Target Users and User Acquisition Plan.....</b>                                     | <b>14</b> |
| <b>Milestone 5: Features.....</b>   | <b>15</b> |
| Importing data.....   | 15        |
| Dashboard.....  | 15        |
| AI-Powered Insights.....  | 16        |
| Future Features and Expansions.....   | 16        |
| 1. Collaboration on shared dashboards.....  | 16        |
| 2. Predictive Analytics.....  | 16        |
| 3. AI Powered Recommendations.....  | 16        |
| 4. Integration with ERP, CRM and Financial Systems.....   | 17        |
| <b>Milestone 6: Monetisation Strategy.....</b>  | <b>18</b> |
| Pricing strategy.....   | 18        |
| 1. Basic Plan – \$0/month (Freemium Tier).....  | 18        |
| 2. Pro Plan – \$10/user/month (Comparable to ClickUp's mid-tier).....                               | 18        |
| 3. Enterprise Plan – \$49/user/month (Competitively priced against Teradata's scalable plans).....  | 19        |
| 4. Custom/Unlimited Plan – Custom Pricing (Teradata-like custom pricing for high-volume needs)..... | 19        |
| Production Costs.....   | 20        |
| Perceived Value.....  | 20        |
| Competition.....  | 20        |
| Target Market Segmentation.....   | 21        |
| <b>Milestone 7: Why LLMs?.....</b>  | <b>22</b> |
| <b>Milestone 8: Prompts Used.....</b>   | <b>24</b> |
| SQL Generation.....   | 24        |
| Base Prompt.....  | 24        |

|  |           |
|--|-----------|
| Few-Shot Examples.....                           | 24        |
| Guardrails.....                                  | 26        |
| <b>SQL Validation.....</b>                       | <b>27</b> |
| Base Prompt.....                                 | 27        |
| Rules/Issues to Check.....                       | 27        |
| Output Format.....                               | 28        |
| <b>Milestone 9: Choice of LLM.....</b>           | <b>29</b> |
| <b>Milestone 10: Product Name and Logo.....</b>  | <b>34</b> |
| Product Name.....                                | 34        |
| Product Logo.....                                | 36        |
| Brand Identity.....                              | 38        |
| <b>Milestone 11: Choice of Technologies.....</b> | <b>39</b> |
| User Interface.....                              | 39        |
| Database.....                                    | 42        |
| Authentication.....                              | 43        |
| Engineering Methodology.....                     | 44        |
| Hosting.....                                     | 46        |
| Front-end hosting.....                           | 46        |
| Backend-hosting.....                             | 47        |
| <b>Milestone 12: Common Workflows.....</b>       | <b>48</b> |
| Workflow 1: Data integration.....                | 48        |
| Workflow 2: Switching between projects.....      | 49        |
| Workflow 3: Insight creation.....                | 50        |
| <b>Milestone 13: UI Considerations.....</b>      | <b>51</b> |
| Data Privacy Transparency.....                   | 51        |
| Minimal Form for Creating Insights.....          | 53        |
| Preview Page.....                                | 54        |
| Feedback Mechanism.....                          | 55        |
| Use of Simple and Practical Language.....        | 56        |
| <b>Milestone 14: Landing page.....</b>           | <b>57</b> |
| <b>Milestone 15: Analytics.....</b>              | <b>58</b> |
| <b>Milestone 16: Marketing.....</b>              | <b>59</b> |
| <b>Milestone 17: Embeddings.....</b>             | <b>61</b> |
| <b>Milestone 18: Social Networks.....</b>        | <b>61</b> |

## Prologue: Known Issues

Our application does not work in Safari and incognito tabs in Google Chrome due to the use of third-party cookies in our implementation of authentication. For more information, see this [GitHub issue](#). While there exists a workaround to use header-based sessions instead of cookies, using header-based sessions might make our application vulnerable to XSS attacks.

## Milestone 0: Problem Description

In today's data-driven world, companies across various industries increasingly rely on business intelligence (BI) to gain a competitive edge. Business intelligence involves the collection, analysis, and visualisation of data to inform strategic decisions and drive performance. Specialised talents with expertise in BI play a crucial role in this process, as they can transform complex datasets into actionable visual insights such as key performance indicators (KPIs), making it easier for decision-makers to interpret and act upon the information. However, while larger organisations often have the resources to invest in these tools, smaller companies face significant challenges. Financial constraints can make it difficult for smaller companies to afford the necessary technology, making them miss out on the opportunity to leverage on their data effectively. This disparity creates a valuable market opportunity to cater specifically to the needs of these smaller businesses.

According to Statista, the BI software industry worldwide is projected to reach US\$27.79 billion in revenue in 2024, based on an estimated growth rate of 5.52%<sup>1</sup>. In 2023, there were an estimated 358 million small-medium enterprises (SMEs) across the globe<sup>2</sup>. Assuming an average revenue of US\$500 per customer per year, the total addressable market for SMEs stands at US\$179 billion a year.

---

<sup>1</sup>According to [Business Intelligence Software - Global | Statista](#)

<sup>2</sup>According to [SMEs worldwide by region 2023 | Statista](#)

# Milestone 1: Competitors

Our 3 closest competitors are ClickUp, Teradata and Power BI.

## ClickUp

<https://clickup.com/features/ai/kpi-generator>

ClickUp does not aim to enhance BI or perform ingestion, it instead generates a template for the possible KPIs the task may need (vanilla GPT). ClickUp aims to be a “Super App” for product managers.

### Pros

- **AI-Powered Offerings:** ClickUp Brain offers robust AI features that can automatically generate tasks, project summaries, and action items, saving teams a lot of manual work.
- **Flexible Pricing:** ClickUp offers a forever-free plan with essential features, which is great for small teams or individuals.
- **Seamless Integrations:** ClickUp Brain integrates with over 1,000 apps like Slack, Zoom, and Google Drive.

### Cons

- **Feature Bloat/Performance:** ClickUp has too many features which can make it difficult to work with for some users, resulting in performance issues.
- **Learning Curve:** Despite its flexibility, some users find the platform overwhelming due to the sheer number of features and customisation options.
- **Indirect Competition:** Broad feature space; the tool competes more with Notion than with us. It does not provide data ingestion capabilities.
- **User Dissatisfaction:** Numerous complaints online on poor user experience with unconventional billing practices, feature flaws etc.

### Sources

- <https://workwellremote.com/in-depth-clickup-review-2024-pros-and-cons/>
- <https://jeremynoronha.com/clickup-ai-review/>
- [https://www.reddit.com/r/clickup/comments/1fdd8g6/collaboration\\_on\\_clickup\\_docs\\_i\\_why\\_does\\_clickup/](https://www.reddit.com/r/clickup/comments/1fdd8g6/collaboration_on_clickup_docs_i_why_does_clickup/)

## Teradata

<https://www.teradata.com/platform/clearscape-analytics>

Teradata is a database analytics software with 3 main services: business analytics, cloud products, and consulting. While the company provides business analytics services, it does not have any LLM integrations, and does not do KPI generation.

### Pros

- **High Performance and Scalability:** Teradata is known for its parallel processing capabilities, which allows it to handle massive datasets efficiently.
- **Advanced Analytics:** The platform includes built-in machine learning algorithms and supports AI and MLOps for data analysis.
- **Flexibility:** Teradata supports various deployment options, including on-premises, cloud, and hybrid setups, offering flexibility for different business needs.

### Cons

- **Cost:** Compared to other RDBMS solutions, the software and professionals needed for maintenance are costly.
- **No LLMs:** While Teradata provides analytics services for ML professionals, it does not use LLM-based tools, nor does it explicitly target KPI generation.
- **Learning Curve:** The system is rather complex, and has a learning curve for analytics functions.

### Sources

- <https://www.trustradius.com/products/teradata-vantage/reviews?qs=pros-and-cons>
- <https://www.capterra.com/p/133231/Teradata-Database/reviews/>

## Power BI

Power BI is a popular data analytics and interactive visualisation tool from Microsoft.

### Pros

- **Excel-like:** Anyone familiar with Excel can learn Power BI, making it perfect for business users.
- **Affordable:** Power BI has a free version with basic functionality and a relatively low-cost pro version with cloud services.
- **Community Support:** There is a large user community and great documentation, guides, and tutorials, making it easy to start learning the tool.

### Cons

- **Sluggish with Large Datasets:** Power BI struggles to handle large datasets, leading to slower performance. Additionally, the desktop version has limitations in data capacity.
- **Steeper Learning Curve for Advanced Features:** While basic usage is straightforward, advanced functionalities like Data Analysis Expressions (DAX) and Power Query might require more time and expertise to master.
- **No Interpretability:** Power BI requires the user to interpret results on their own, thus requiring business expertise.

### Sources

- <https://acuvate.com/blog/microsoft-power-bi-pros-and-cons/>
- <https://www.altexsoft.com/blog/power-bi-pros-cons/>

## Our Product

GoodBI's value proposition lies in addressing a significant gap in the BI market: providing an affordable, easy-to-use solution that automates key aspects of data analysis for non-technical users, while still being robust enough to handle meaningful data queries. Let's break down the competitive landscape in more detail and explain how GoodBI meets a specific market need.

### GoodBI vs ClickUp: Addressing a Need for Dedicated BI Automation

ClickUp is primarily a project management and productivity platform, not a dedicated BI tool. It offers task tracking, time management, and collaboration features, with some data visualisation capabilities via integrations (e.g., with third-party tools like Tableau or Power BI).

### Where GoodBI Stands Out

GoodBI is a dedicated BI platform that allows users to directly interact with their data through SQL-based queries, with no need for integrations with external BI tools. ClickUp lacks native automated data explanation and visualisation suggestions. In contrast, GoodBI provides an intuitive experience by explaining the query outputs in layman's terms and automatically recommending appropriate graphs. This reduces the time and expertise needed to gain insights from data.

- **Market Need:** Teams often find that platforms like ClickUp don't provide deep data analysis and still require third-party BI tools to derive actionable insights. GoodBI fills this gap by being an all-in-one solution for analysis, without the complexity and cost of external add-ons.
- **Affordability:** ClickUp, while more affordable than enterprise-grade tools, still doesn't offer BI-specific pricing tiers. GoodBI, designed with a focus on delivering BI at a lower cost, appeals to small and medium-sized teams that don't need the broader scope of project management features, but do need accessible BI capabilities.

## GoodBI vs TerraData: Simplifying Complex Data Analysis for Small to Medium Enterprises

TerraData is a highly sophisticated platform aimed at enterprises with large, complex data environments. It offers advanced features for data warehousing, query optimization, and large-scale analytics, but it comes at a high cost and with significant complexity.

### Where GoodBI Stands Out

GoodBI democratises data analysis for businesses that cannot afford or do not need the full power of TerraData. While TerraData is perfect for handling big data and performing complex analytical queries, it is often over-engineered for SMEs or teams with smaller datasets.

- **Ease of Use:** TerraData requires significant technical expertise, often needing data engineers or data scientists to operate effectively. In contrast, GoodBI's user-friendly interface and automated explanations allow users with limited technical background to run SQL queries, understand the results, and generate meaningful visualisations without needing a dedicated analytics team.
- **Market Need:** Many small and medium businesses and even startup teams face a gap where they can't afford to hire specialised data scientists but still need meaningful insights from their data. GoodBI meets this need by offering an affordable and easy-to-operate BI solution with the automation necessary to make data analysis accessible to a wider audience.
- **Affordability:** TerraData is often cost-prohibitive for smaller teams. GoodBI, by focusing on a more lightweight yet effective tool, captures the market segment that wants insights without the hefty costs and complexity associated with enterprise tools like TerraData.

## GoodBI vs Power BI: Combining Simplicity with Customization

Power BI is one of the most widely-used business intelligence tools and is known for its deep integration with the Microsoft ecosystem. It offers powerful data visualisation, report building, and integration options but requires a moderate to high level of technical skill to fully leverage its features.

### Where GoodBI Stands Out

- **Ease of Use:** Power BI, while highly capable, presents a steeper learning curve for non-technical users. GoodBI bridges this gap by offering an intuitive user experience, where the system handles much of the complexity behind the scenes (e.g., query explanation, graph suggestions).
- **Automated Graph Suggestions:** Power BI provides excellent flexibility in creating reports and dashboards, but it doesn't automatically suggest the best visualisation for the data. GoodBI's ability to **automatically suggest and generate graphs** based on user queries is a differentiator that caters to users who are not sure what the best way to represent their data is.
- **SQL-First Approach:** Power BI allows for complex data manipulation through its native tools (Power Query, DAX) but requires a separate learning process. GoodBI's SQL-first approach is more familiar to many users and allows direct query-based interaction, while still remaining simple through the explanation of results.
- **Market Need:** While Power BI serves mid-to-large businesses well, it often requires in-house expertise or investment in training. GoodBI taps into the need for simplicity and affordability, targeting users who want to perform essential BI tasks without going through a steep learning curve or integrating with a broader software ecosystem like Microsoft.

## Market Need Tackled by GoodBI

GoodBI identifies a gap in the market for simple, automated, and cost-effective BI solutions that serve a variety of businesses, particularly:

- **Small to medium enterprises** that cannot afford expensive, complex enterprise tools like TerraData but need powerful analytics.
- **Teams with limited technical expertise** that require a user-friendly interface to query, interpret, and visualise data without relying on in-house data scientists.
- **Non-technical users** who benefit from automated query explanations and visualisation suggestions, reducing the barrier to understanding and acting on data insights.
- **Companies on a budget** that need an affordable alternative to tools like Power BI and TerraData without sacrificing key functionality like SQL querying, graph generation, and insights.

GoodBI effectively carves out its space by focusing on **automated simplicity, affordability, and accessibility** in the business intelligence market. While competitors offer more comprehensive, complex, or project-oriented solutions, GoodBI shines as a focused BI tool that bridges the gap between ease of use and powerful analysis, tailored to the needs of SMBs and teams looking for a lightweight but effective solution.

# Milestone 2: Application Description

## Overview

We are introducing an exciting solution, GoodBI (pronounced /gʊd'baɪ/), a game-changing GenAI-powered enterprise resource planning (ERP) application designed to transform how businesses generate KPIs and actionable insights from their raw data.

GoodBI has several objectives. It is built to break down barriers, making advanced BI accessible to companies of all sizes. By harnessing the power of AI, GoodBI aims to simplify the traditionally complex BI processes, eliminating the need for specialised skills or expensive software. With faster, automated insights generation at a fraction of the cost, GoodBI empowers businesses to make smarter, data-driven decisions with ease.

## User Stories

1. As a small business owner with limited resources, I want to upload my raw data easily, so that I can quickly generate insights without technical expertise.
2. As a data analyst, I want to quickly generate KPIs without barriers faced from tools, so that I can focus on more strategic analysis and deliver greater value to my organisation.
3. As a data analyst, I want to customise the generated insights, so that they align with specific business requirements and goals.
4. As a medium-sized business owner, I want to handle large datasets efficiently, so I can analyse performance metrics at scale.

## Milestone 3: Tech Moat

Our product's unique competitive advantage stems from the integration of cutting-edge large language models (LLMs) that empower executives to effortlessly manage their business intelligence needs without the technical expertise typically required. Traditionally, businesses rely on BI developers or teams with deep knowledge of BI tools to set up and manage their dashboards and KPIs. We've completely redefined this process.

By simply uploading CSV files or other data sources, our AI assistant takes over, generating key performance indicator (KPI) suggestions and offering an intuitive, user-friendly interface. Executives can now create customised dashboards on the fly, based on a simple form describing what they need, without any steep learning curve or dependency on technical staff. This simplicity offers an enormous time-saving benefit, allowing decision-makers to focus on interpreting results and driving strategy rather than navigating complex software.

This ease of use creates a powerful inertia for clients to remain loyal to our product. Even if competitors manage to replicate our features, the value we offer extends beyond functionality – it's in the seamless experience and instant results. Switching to another platform would not only be time-consuming but also result in a significant loss of productivity and a poor return on investment, as executives would need to rebuild dashboards and KPIs from scratch.

Our moat is further reinforced by the intelligent use of AI to handle sophisticated BI tasks, removing technical barriers and enabling executives to stay agile in today's fast-paced business environment. The ability to get immediate insights without external help positions our platform as an essential tool that enhances both efficiency and decision-making. As a result, clients are more likely to stay with our solution for the long term, as the combination of convenience, AI-driven insights, and time savings outweighs any potential benefits of switching to a competitor.

In addition to that, porting over data and dashboards from our web portal is a cumbersome and time-consuming process, increasing the reluctance of users to switch over to other platforms. Hence, the nature of our solution gives us an inherent moat.

## Milestone 4: Target Users and User Acquisition Plan

Our primary target users are subject matter experts and executives who have specific questions regarding their company data and want to visualise data in a certain way but face a bottleneck from a BI developer who acts as a middleman between them and their dashboard. We aim to reduce their communication overhead by allowing them direct and instant access to KPIs, with the ability to modify the KPIs on the fly.

Our secondary users are BI developers themselves who have domain knowledge but are restricted by the slowness of having to use complicated BI tools to get the same repetitive things done. BI developers normally get specifications as to what KPIs to build, and building tiles on a dashboard is often a repetitive process. GoodBI can help speed up this process to save their bandwidth for other tasks.

However, acquiring target users will be tricky. Given that we are a startup with no credibility yet, it will be difficult for us to secure any clients without being squeezed on price negotiations. Furthermore, executives will not want to let investors know that they are investing huge money into a new startup with no clientele to solve their business and technical needs.

As a result, we will adopt the strategy of **whale hunting**, where we identify a large or well-known player and pitch the app to them. We either sell the application to them at a heavy bargain or basically for free (convert them to beta testers), and put their positive feedback and reviews on our website.

This gives us not only credibility but also access to their network. We build relationships with top-level executives who provide us access to other executives and managers who could potentially use our product.

Furthermore, given that we adopt a **top-down SaaS model**, pitching to executives or the ‘top dogs’ within a company will enable them to effect company-wide change if they like our product. Executives are always on the lookout to diversify, so the data field is not exactly a zero-sum game. Hence, in a way, acquiring customers doesn’t necessarily reduce our competitors’ profit margins.

## Milestone 5: Features

Our MVP consists of the bare minimum set of features our client would need to generate at least a small set of dashboards with insights. Hence, we decided to implement an end-to-end ingestion pipeline that leveraged multiple LLMs to generate insights and KPI charts based on the user's natural language query.

We aimed for a high accuracy of output (KPI charts) that matched closely with user queries, as this can impact their perception of the usefulness and trustworthiness of Generative AI in solving their business needs.

### Importing data

Users can manually upload their CSV data sources to our web app for storage in our database for analytics to be generated. Since popular CRM tools like Salesforce and HubSpot offer the ability to export existing data in the CSV format, supporting manual uploads for this file type simplifies the data ingestion process for users with existing BI solutions. This allows us to highlight the efficiency and ease-of-use of our data ingestion process to them. This also sets the stage for future automated imports once CRM integrations are implemented.

### Dashboard

Users are presented with a clear dashboard view of AI-generated KPIs and insights. These insights can be grouped into projects, and users can seamlessly switch between different projects directly on the dashboard without leaving the page. Users are able to customise their dashboard experience by dragging and resizing the different insights on the dashboard and adding any additional insights as needed. This implementation aligns with the purpose of our app to offer our target user the ability to effortlessly engage and interact with their generated KPIs and business insights.

## **AI-Powered Insights**

Users can generate KPIs and business insights from their previously uploaded raw data by providing simple, natural language descriptions of the insights they wish to visualise. These commands are passed to a large language model (LLM) for interpretation, which generates a query string based on the names and descriptors of the data columns. The query is then executed on the user's raw data, producing the relevant visualisations for the user. The inclusion of this feature allows us to demonstrate how complex and repetitive tasks in business insights generation can be simplified into a one-click solution through the help of LLMs. By eliminating the need for manual query writing and visualisation creation, this enhances user efficiency and the efficiency of business insights generation.

## **Future Features and Expansions**

### **1. Collaboration on shared dashboards**

Users can collaborate on shared dashboards to review, discuss, and agree on KPIs and business metrics in real time. Each team member can suggest modifications, add new metrics, or highlight important trends, ensuring that the insights reflect a comprehensive view of the business. This collaborative approach not only improves decision-making but also ensures that the resulting enhancements are aligned with the collective goals of the organisation.

### **2. Predictive Analytics**

Predictive analytics will be provided to users based on the data gathered via their generated KPIs and business insights. These foresights generated by analysing historical data trends and patterns allows organisations to make informed decisions, anticipate potential challenges, and strategically plan for upcoming opportunities. This feature enhances decision-making by offering data-driven projections, helping users stay proactive and better prepared for future business scenarios. It also deepens the value of the platform by moving beyond reactive insights to forward-looking strategies.

### **3. AI Powered Recommendations**

AI-generated recommendations on improving business metrics are provided to users based on the insights generated. These recommendations leverage the power of large language models (LLMs) to analyse the data holistically, offering suggestions that are both specific and highly impactful. This showcases the ability of our app to offer insights with enhanced values that could guide users toward actionable improvements.

#### **4. Integration with ERP, CRM and Financial Systems**

The seamless integration of our app with users' existing ERP, CRM, and financial systems will allow them to take strategic actions based on the AI-powered insights we provide. This integration facilitates a more cohesive and efficient workflow, allowing insights to be immediately actionable within their current operational setup. By connecting with these systems, we enhance the utility of our insights while promoting continuous user engagement with our platform. This streamlined integration also simplifies the adoption of our app, embedding it into the user's daily business processes and ensuring a smoother transition into our AI-powered decision-making through our product.

# Milestone 6: Monetisation Strategy

## Pricing strategy

### 1. Basic Plan – \$0/month (Freemium Tier)

**Target:** Small businesses and startups testing AI-powered analytics.

#### Features

- AI-generated KPI dashboard (limited to 3 KPIs)
- Integration with basic data sources (Excel, CSV)
- Up to 500MB of data processing per month
- Standard email support

**Value Proposition:** Provide small teams with a no-cost solution to test the AI's capabilities, driving adoption through simplicity and free access to essential features. This creates a clear upgrade path as businesses grow.

### 2. Pro Plan – \$10/user/month (Comparable to ClickUp's mid-tier)

**Target:** Small and medium-sized businesses needing deeper insights and broader integrations.

#### Features

- AI-generated KPIs (up to 10 KPIs)
- Integration with CRM tools (Salesforce, HubSpot) and cloud storage (Google Drive, Dropbox)
- 5GB of data processing per user per month
- Basic predictive analytics (AI-driven forecasts)
- Data visualisation and export capabilities (PDF, Excel)
- Shared dashboards with collaboration tools
- Standard email and chat support

**Value Proposition:** Offer more powerful insights with multiple integrations and predictive analytics for growing businesses, making it an affordable and practical choice for mid-market companies.

### **3. Enterprise Plan – \$49/user/month (Competitively priced against Teradata's scalable plans)**

**Target:** Larger companies handling complex data from multiple sources. We have these features that will be built on demand for our enterprise clients.

#### **Features**

- Unlimited AI-generated KPIs with real-time insights
- Integration with ERP, CRM, and financial systems (SAP, Oracle, QuickBooks)
- 50GB of data processing per user per month
- Advanced predictive analytics and AI-based recommendations
- Industry-specific KPI templates (finance, healthcare, retail)
- Custom dashboards with API access for external tools
- Role-based access and permission controls
- 24/7 priority support, onboarding, and training

**Value Proposition:** Address the needs of large enterprises that require complex data integration, security, and industry-specific insights, but at a fraction of the cost of traditional enterprise BI solutions like Teradata.

### **4. Custom/Unlimited Plan – Custom Pricing (Teradata-like custom pricing for high-volume needs)**

**Target:** Enterprises with extensive, tailored AI and data requirements.

#### **Features**

- Fully customisable AI-driven KPI models
- Unlimited data processing
- Custom machine learning model integration
- Advanced security features (encryption, multi-factor authentication)
- Dedicated data science consulting
- On-premise or hybrid cloud deployment
- SLA-backed 24/7 premium support with a dedicated account manager

**Value Proposition:** Tailor-made for large enterprises that need customised solutions, real-time analytics, and extensive data storage. This plan offers flexibility for companies with specific regulatory or operational requirements, similar to Teradata's model.

## Production Costs

- **AI Infrastructure & Development:** Our platform uses advanced AI and machine learning models to process large amounts of data. Developing and maintaining these models, along with the required cloud infrastructure (for data storage, compute power, etc.), directly contributes to production costs.
- **Data Processing:** Offering competitive plans means the platform has to manage varying levels of data ingestion, storage, and computation. While smaller businesses can get by with limited data processing, larger businesses generate much more, increasing the cost. Tiered and custom pricing accounts ensure that costs are covered for high data users.

## Perceived Value

- **AI-Powered Insights:** Our platform is an AI-driven solution that simplifies and accelerates KPI generation compared to traditional BI tools. AI reduces the complexity and manual work associated with BI, providing **high perceived value** to users looking for efficiency gains.
- **Democratising BI:** Making advanced analytics accessible without specialised skills enhances the perceived value for SMEs, which typically lack the budget for costly analytics teams.
- **Freemium Model:** Offering a free tier helps to drive adoption by lowering the barrier to entry, showing potential users the value before they commit to paid plans.

## Competition

- **ClickUp:** ClickUp's competitive pricing for collaborative project management influenced the **lower pricing tiers** in our model (e.g., \$10/user/month for the Pro plan). This keeps the product affordable for SMEs while still offering advanced AI capabilities.
- **Teradata:** Teradata's pricing model focuses on enterprise-scale operations, often with custom solutions. Drawing inspiration from this, the **Enterprise and Custom Plans** are designed for large-scale businesses that need more customisation and scalability. Custom pricing ensures flexibility, allowing the platform to compete with Teradata on high-end features without compromising profitability.

## Target Market Segmentation

- **SMEs vs. Large Enterprises:** SMEs generally have tighter budgets but are willing to pay for value-adding tools. Thus, the **Pro and Enterprise tiers** were priced competitively to provide essential features at a reasonable cost.
- **Scalability for Enterprises:** Large enterprises typically require vast amounts of data processing, real-time insights, and advanced security, which leads to higher pricing for the **Enterprise and Custom tiers**. This aligns with the expectations and budgets of larger businesses.

## Milestone 7: Why LLMs?

In the traditional business intelligence (BI) process, language serves as the primary means for executives to communicate their analytical requirements to BI developers. Executives, who typically possess deep domain knowledge but lack technical expertise in BI tools or SQL, must rely heavily on the ability of BI developers to interpret their instructions correctly. The BI developer, often without the executive's domain context, is responsible for translating these high-level business needs into data queries and visualisations. However, this process introduces several challenges. Miscommunications can occur, leading to misunderstandings of the executive's intent. Furthermore, even when the instructions are clear, the developer may not always know which visualisations will be most appropriate for the specific business context. As a result, the entire BI design process becomes dependent on the effectiveness of communication and knowledge transfer between the executive and the developer.

This is where large language models (LLMs) provide a transformative solution by acting as a natural language layer of abstraction between the executive and the technical complexities of BI tools. LLMs bridge the gap, allowing executives to bypass the need for a BI developer altogether. Through the use of natural language inputs, executives can describe their requirements, goals, and desired outcomes in the same way they would communicate with a human developer. The LLM interprets these instructions in real time, generating appropriate KPI suggestions, data queries, and visualisations instantly – without the need for the executive to possess any technical knowledge of BI tools or SQL.

The power of LLMs lies in their ability to understand complex domain-specific language, context, and nuances within the executive's instructions. By accurately interpreting and translating these inputs into actionable visualisations, LLMs ensure that the final output aligns perfectly with the executive's business needs. This eliminates the risk of miscommunication and reduces the time traditionally spent on back-and-forth exchanges between executives and BI developers.

Moreover, LLMs enable on-the-fly dashboard creation, where executives can adjust parameters, ask follow-up questions, and refine their data visualisations through simple, conversational inputs. This level of dynamic interaction provides an unparalleled level of flexibility and agility, empowering business leaders to extract insights from their data in real-time without technical intervention.

In essence, LLMs are central to the BI process because they democratise data access and analysis, making sophisticated BI capabilities accessible to those without technical backgrounds. They remove the dependency on BI developers, streamline the workflow, and allow executives to leverage their domain expertise directly. The result is faster, more accurate visualisations that support better decision-making, all while reducing friction and delays that commonly plague traditional BI setups.

Thus, the integration of LLMs into our application offers a seamless and efficient solution, enabling non-technical users to harness the full potential of their data with minimal effort and maximum impact. This paradigm shift not only saves time but also enhances the overall effectiveness of business intelligence, making LLMs indispensable to the future of BI technology.

## Milestone 8: Prompts Used

One of the more elaborate prompts that we engineered was for the SQL Generation and Validation. This was a 2 stage process incorporating Persona and Output Automator Patterns, followed by multiple instructions setting up guardrails.

### SQL Generation

#### Base Prompt

Uses the Persona Pattern.

You are an AI assistant that generates SQL queries based on user questions, database schema, and unique nouns found in the relevant tables. Generate a valid SQL query to answer the user's question. If there is not enough information to write a SQL query, respond with "NOT\_ENOUGH\_INFO". A schema name will be provided as well, which should be used to identify the relevant tables and columns.

#### Few-Shot Examples

The base prompt is then followed by few-shot examples, each example elaborating on a different user query type.

Here are some examples:

1. What is the top selling product?

schema\_name: 59320165-aa75-409c-9c92-5a5ae59ea28

table\_name: sales

Answer:

```
SELECT product_name, SUM(quantity) as total_quantity
FROM "59320165-aa75-409c-9c92-5a5ae59ea28"."sales"
WHERE product_name IS NOT NULL AND quantity IS NOT NULL AND
product_name != '' AND quantity != '' AND product_name != 'N/A'
AND quantity != 'N/A'
GROUP BY product_name ORDER BY total_quantity DESC LIMIT 1
```

2. What is the total revenue for each product?

schema\_name: 2415r3qd-31edasc34-31rf3q-9464yh9

table\_name: sales

Answer:

```
SELECT product_name, SUM(quantity * price) as total_revenue
FROM "2415r3qd-31edasc34-31rf3q-9464yh9"."sales"
WHERE product_name IS NOT NULL AND quantity IS NOT NULL AND price
IS NOT NULL AND product_name != '' AND quantity != '' AND price !=
'' AND product_name != 'N/A' AND quantity != 'N/A' AND price !=
'N/A'
GROUP BY product_name ORDER BY total_revenue DESC
```

3. What is the market share of each product?

```
schema_name: g34tfv42t-24erfvty35-898756yg
table_name: sales
```

Answer:

```
SELECT product_name, SUM(quantity) * 100.0 / (SELECT SUM(quantity)
FROM "g34tfv42t-24erfvty35-898756yg"."sales") as market_share
FROM "g34tfv42t-24erfvty35-898756yg"."sales"
WHERE product_name IS NOT NULL AND quantity IS NOT NULL AND
product_name != '' AND quantity != '' AND product_name != 'N/A'
AND quantity != 'N/A'
```

```
GROUP BY product_name ORDER BY market_share DESC
```

4. Give me a query of the tracks with the higher number of streams
with the released year of 2023 from the Spotify dataset.

```
schema_name: 32r32g533-14rf35532r-t6u56hrt
table_name: spotify
```

Answer:

```
SELECT track_name, streams FROM
"32r32g533-14rf35532r-t6u56hrt"."Spotify"
WHERE released_year = '2023' AND track_name IS NOT NULL AND
streams IS NOT NULL AND track_name != '' AND track_name != 'N/A'
ORDER BY streams DESC LIMIT 25
```

5. Plot the distribution of income over time.

```
schema_name: 32r32g533-14rf35532r-t6u56hrt,
table_name: users
```

Answer:

```
SELECT income, COUNT(*) as count FROM
"32r32g533-14rf35532r-t6u56hrt"."users"
WHERE income IS NOT NULL AND income != '' AND income != 'N/A'
GROUP BY income
```

THE RESULTS SHOULD ONLY BE IN THE FOLLOWING FORMAT, SO MAKE SURE TO ONLY GIVE TWO OR THREE COLUMNS:

`[[x, y]] or [[label, x, y]]`

## Guardrails

We then follow it up with guidelines on what to do and not do.

For questions like "plot a distribution of the fares for men and women", count the frequency of each fare and plot it. The x axis should be the fare and the y axis should be the count of people who paid that fare.

SKIP ALL ROWS WHERE ANY COLUMN IS NULL or "N/A" or "". Ensure that the conditions in the WHERE clause are valid with respect to the column types.

For example, for string columns, ensure that the comparison is done with a string value. For numeric columns, ensure that the comparison is done with a numeric value. Just give the query string. Do not format it. Make sure to use the correct spellings of nouns as provided in the unique nouns list.

Limit the number of results to `{max_results}`.

# SQL Validation

## Base Prompt

You are an AI assistant that validates and fixes SQL queries. Your task is to:

Check if the SQL query is valid. If there are any issues, fix them and provide the corrected SQL query.

## Rules/Issues to Check

Check for the following issues:

- Single quotes not double quotes around empty strings, double quotes ONLY make delimited identifiers, and "" isn't a meaningful identifier.
- For numeric types, do not check for empty strings or "N/A".
- Ensure all table and column names are correctly spelled and exist in the schema.
- Ensure that the conditions in the WHERE clause are valid with respect to the column types. For example, for string columns, ensure that the comparison is done with a string value. For numeric columns, ensure that the comparison is done with a numeric value.
- Ensure that the correct table, schema, and column names are used. The correct format for table name is "schema\_name.table\_name".
- The schema name and table name should not be modified, if it is dash-separated, it should remain dash-separated.
- If the table name is incorrect, provide the correct table name.
- For every column, if the column name is incorrect, provide the correct column name.
- Check for any other issues that may cause the query to fail.

If no issues are found, return the original query.

## Output Format

We enforce JSON outputs for all of our other prompts. This ensures consistency in outputs and reduces variability.

Respond in JSON format with the following structure. Only respond with the JSON:

```
{ {  
    "valid": boolean,  
    "issues": string or null,  
    "corrected_query": string  
} }
```

For example:

```
1. { {  
    "valid": true,  
    "issues": null,  
    "corrected_query": "None"  
} }  
  
2. { {  
    "valid": false,  
    "issues": "Column user(s) does not exist",  
    "corrected_query": "SELECT * FROM users WHERE age > 25"  
} }  
""" ,  
    ),  
    (  
        "human",  
        """==User question: {question}  
  
==Schema name: {schema_name}  
  
==SQL query: {sql_query}  
  
==Table name: {table_name}  
  
==Column details:  
"""
```

```
{column_details}
```

Validate the SQL query and provide the corrected query.

In the end we again remind the LLM of its core task to reduce forgetfulness in models. We also asked the LLM to explain the reasoning behind selecting the graph type it determines is most suitable.

You are a data analyst expert in visualizing data.

Given the question and SQL query, identify the graph type that best represents the data output.

Give a short explanation of why the graph type is appropriate.

Allowed output graph types are: ["none",  
"scatter", "bar", "horizontal\_bar", "line", "pie", "scatter"]

Sample JSON format:

```
{  
  "graph_type": string,  
  "reason": string  
}
```

## Milestone 9: Choice of LLM

Our primary consideration in selecting a suitable LLM is its availability via an API. Given our limited computational resources and the challenges associated with self-hosting, using an API is a more practical and cost-effective solution at the introductory stage of our product. This approach allows us to focus on developing the features of our app rather than managing LLM infrastructure. Additionally, strong reasoning and data analysis capabilities are crucial for us, as a key feature of our app is to identify the relevant data needed for the automated generation of business insights and KPIs.

With these considerations in mind, here are the 4 models we have shortlisted for further evaluation:

- GPT-4o by OpenAI
- Claude 3.5 Sonnet by Anthropic
- Llama 3.1 (405B) Instruct Turbo by Meta
- Gemini 1.5 Pro by Google

Benchmarks:

[LiveBench](#)

| Model                              | Reasoning Average | Data Analysis Average | Combined Average |
|------------------------------------|-------------------|-----------------------|------------------|
| chatgpt-4o-2024-08-06              | 52.00             | 54.43                 | 53.22            |
| claude-3-5-sonnet-20240620         | 58.67             | 56.74                 | 57.70            |
| meta-llama-3.1-405b-instruct-turbo | 53.33             | 53.51                 | 53.42            |
| gemini-1.5-pro-exp-0827            | 49.33             | 50.83                 | 50.08            |

[Vellum AI](#)

| Model               | Reasoning Average (GPQA) |
|---------------------|--------------------------|
| GPT-4o              | 53.60                    |
| Claude 3.5 Sonnet   | 59.40                    |
| Meta Llama 3.1 405b | 51.10                    |
| Gemini 1.5 Pro      | 46.20                    |

As our app includes the generation of SQL queries from text, we identified a few comparisons that help us to make our choices for the LLM

1. [Waii](#)

- a. Schema Understanding: GPT-4o demonstrates a slight edge in selecting relevant tables and summarising large schemas (up to 1251 tables), while Claude 3.5 Sonnet excels at providing detailed and easy-to-understand column descriptions.
- b. Query Complexity: Both models perform effectively across simple and complex queries, though GPT-4o has a slight accuracy advantage in handling more complex tasks.
- c. Performance and Efficiency: GPT-4o is notably faster, generating complex queries 42.4% quicker than Claude 3.5 Sonnet, and is more efficient, using 34.3% fewer tokens.
- d. Interactive Queries: GPT-4o outperforms Sonnet in handling follow-up questions and maintaining context across interactions.

2. Prem (BIRD dataset)

a. Accuracy (comparison by difficulty level)

| Models                | Simple | Moderate | Hard | Overall |
|-----------------------|--------|----------|------|---------|
| llama-3.1-40<br>5B    | 60     | 35       | 42   | 45.66   |
| gpt-4o                | 56     | 35       | 41   | 44      |
| claude-3.5-s<br>onnet | 42     | 27       | 31   | 33.33   |

b. Valid Efficiency Score (VES)

| Models                | Simple  | Moderate | Hard    | Overall |
|-----------------------|---------|----------|---------|---------|
| llama-3.1-40<br>5B    | 95.3519 | 38.899   | 41.733  | 58.66   |
| gpt-4o                | 54.9621 | 37.9472  | 41.9917 | 44.96   |
| claude-3.5-s<br>onnet | 44.4475 | 24.9191  | 31.3255 | 33.56   |

Overall Performance Ranking

| Model               | Total Combined Ranking     |
|---------------------|----------------------------|
| GPT-4o              | 9                          |
| Claude 3.5 Sonnet   | 8                          |
| Meta Llama 3.1 405b | 7                          |
| Gemini 1.5 Pro      | 8 (excluding Prem ranking) |

Cost

| Model                               | Input Price (USD per 1M tokens) | Output Price (USD per 1M tokens) |
|-------------------------------------|---------------------------------|----------------------------------|
| GPT-4o                              | 5.00                            | 15.00                            |
| Claude 3.5 Sonnet via Anthropic API | 3.00                            | 15.00                            |
| Claude 3.5 Sonnet via AWS Bedrock   | 3.00                            | 15.00                            |
| Meta Llama 3.1 405b via             | 5.32                            | 16.00                            |

|                |      |       |
|----------------|------|-------|
| AWS Bedrock    |      |       |
| Gemini 1.5 Pro | 3.50 | 10.50 |

Given the overall satisfactory performance of GPT-4o, which consistently ranked second in most benchmarks that we have listed above and third in one, we found it to be a robust choice for our project. While it is tempting to choose the model which performed best on paper, we decided to start with GPT-4o due to our team's familiarity with ChatGPT models, especially considering the tight timeline we face. In an evaluation conducted by Waii, GPT-4o has also been noted to outperform Claude 3.5 Sonnet, indicating its superior capabilities in Text to SQL tasks. With these factors in mind, we selected GPT-4o as our large language model (LLM) for this project, and we were able to achieve results with the accuracy required to meet our objectives. As we experimented with both models via Langchain, we also found that the responses from GPT-4o were more accurate than Claude 3.5 Sonnet, especially while working with complex queries. Responses from GPT-4o were also noticeably faster than the ones from Claude 3.5 Sonnet.

Other factors behind choosing GPT-4 included the following:

1. **Ease of Integration:** GPT-4 was designed with integration in mind, offering robust APIs that can be effortlessly embedded into existing systems. This feature allowed us to incorporate GPT-4 into our tool without the need for extensive customization or re-engineering. Its flexible architecture ensures that it can work smoothly with our data processing pipelines and UI, which significantly accelerates the development timeline and ensures a reliable user experience.
2. **Comprehensive Documentation:** A critical factor in our choice was the high-quality documentation that supports GPT-4. The extensive and clear documentation ensures that we can quickly navigate, troubleshoot, and implement complex features. This reduces the development cycle, allowing us to maintain a fast pace of innovation and enhance the tool without worrying about technical roadblocks or ambiguous guidelines.
3. **LangChain Ecosystem:** GPT-4 is well-supported by LangChain, an ecosystem that enables sophisticated workflows involving language models. This ecosystem allows us to create complex chains of tasks, automate data interactions, and enhance the overall functionality of our application. LangChain simplifies advanced operations, such as integrating external databases, improving data processing efficiency, and connecting multiple LLM-powered components within the tool. By leveraging

LangChain's capabilities, we can scale the number agents we use and offer deeper functionality with minimal overhead.

4. **Plug-and-Play Capabilities:** Another key reason for choosing GPT-4 is its plug-and-play nature. GPT-4's out-of-the-box capabilities are well-suited for handling diverse tasks like understanding natural language inputs, generating KPI suggestions, and creating visualizations. Its adaptability means that we can introduce new features quickly, respond to user feedback effectively, and integrate it into various use cases without requiring heavy modifications or deep customizations. This plug-and-play functionality ensures that GPT-4 can be fine-tuned to meet our evolving needs without slowing down our development cycles.

By choosing GPT-4, we've equipped our BI platform with a powerful, flexible, and scalable language model that not only fits our current technical needs but also positions us to continue enhancing the tool with future advancements in AI technology.

#### Model parameters

- **Model:** GPT-4o
- **Parameters:** More than 200 billion parameters, surpassing GPT-4-175B (exact count not disclosed)
- **Temperature:** 0 (deterministic, no randomness)
- **Output:** Fixed, repeatable, most probable response

# Milestone 10: Product Name and Logo

## Product Name

Our product is named GoodBI (pronounced /gʊd'baɪ/). GoodBI stands for “Good Business Intelligence” and it embodies our commitment to delivering high quality insights that empower businesses to thrive. In today’s data-driven world, having access to accurate and actionable information is crucial for making informed business decisions. Our name is also a word play on “goodbye”, with a similar pronunciation. This highlights our vision to simplify business intelligence, making it much easier and cheaper to generate KPIs, even without the help of BI developers.

We mainly considered names that were a combination of other words, since it would help in creating a shorter name that is more memorable. Some of the other names we considered included:

- Genlytics: This name is a combination of generative and analytics, but we decided against this name as there were multiple applications and businesses with the same name<sup>3</sup>.
- GenSight: This name is a combination of generative and insights. Similarly, we rejected this name as there were multiple companies with the same name<sup>4</sup>.
- FireBI: We felt that this name as it sounded too similar to Firebase, an app development platform by Google, and users might get confused.
- ChatBI: We felt that this did not capture our identity entirely, since this name could be mistaken for a chat application.

---

<sup>3</sup> See <https://www.linkedin.com/company/sbbs-pvt-ltd>, <https://github.com/zfahssaan/genlytics>, and <https://www.genlytics.ai/>

<sup>4</sup> See <https://www.gensight-biologics.com/> and <https://gensight.com/>

Eventually, we picked GoodBI. Beyond the meaning of the name, GoodBI was also a suitable name for several reasons:

- GoodBI is short and memorable, and sounds similar to the word goodbye.
- GoodBI is easy to reproduce by spelling.
- The domain name, although the domain goodbi.com has been registered since 2011, it is not in use currently and is available for sale.

| goodbi.com         |   | Updated 1 second ago |
|--------------------|---|----------------------|
| Domain Information |   |                      |
| Domain:            | goodbi.com                                  |                      |
| Registrar:         | Alibaba Cloud Computing (Beijing) Co., Ltd. |                      |
| Registered On:     | 2011-05-04                                  |                      |
| Expires On:        | 2029-05-04                                  |                      |
| Updated On:        | 2019-10-30                                  |                      |
| Status:            | ok  |                      |
| Name Servers:      | ns3.afternic.com<br>ns4.afternic.com        |                      |

- While the social media handle “goodbi” is taken on both Instagram, TikTok and X (formely Twitter), we can opt for a longer handle “goodbusinessintelligence” which is available on all the platforms. While this might take longer to type, search indexing on most social media platforms are done on both handles and names, so we will still be searchable on the social media platforms.
- Low similarity with other brands. There is a horror film with the name [GoodBi](#), but since it is not a software application, it is not problematic. The only other competing product of concern is Good.BI. However, it does not appear very established despite being registered in 2019.

| good.bi            |                                      | Updated 1 second ago |
|--------------------|--------------------------------------|----------------------|
| Domain Information |                                      |                      |
| Domain:            | good.bi                              |                      |
| Registrar:         | Key-Systems                          |                      |
| Registered On:     | 2019-11-07                           |                      |
| Expires On:        | 2024-11-07                           |                      |
| Updated On:        | 2023-11-06                           |                      |
| Status:            | ok                                   |                      |
| Name Servers:      | ns1.netangels.ru<br>ns2.netangels.ru |                      |

- Web searches for the word Good BI do not get corrected or filtered on search engines we tested on (Google, Bing, DuckDuckGo).

## Product Logo

With the name in mind, we designed our product logo with the wordmark. In designing our logo, we had several considerations.

Firstly, the logo has to be visible at different sizes. Since we are building a web application, the logo will serve as a favicon, which is typically 16x16 pixels when displayed to users. It is crucial to ensure that it maintains clarity even at a reduced size. This means that the logo should feature simple shapes and limited details.

Secondly, the logo should be versatile in color usage. It should work well on different backgrounds, and should also work well when displayed in a monochrome format. This gives us the ability to adapt to various themes such as light and dark backgrounds on our website and on our marketing materials. This ensures a consistent brand experience across the application. Choosing a small number of colors to use for the logo would be ideal.

Thirdly, the logo should be meaningful and relevant. The logo should bring across the essence of our software and provide some idea about what it does. Given the serious nature of business, our wordmark also required a more formal look.

With these considerations in mind, we brainstormed several possible concepts for the logo. The concepts are as follows:

- A hand wave, since our product name is a word play on goodbye.
- A chart formed by the fingers, since we display data visualisations.
- A thumbs up, as our product name has the word “good”.
- An eye, a bulb or a brain, to capture the idea of “business intelligence”.
- A lightning bolt to signify speed.

We did several design iterations for the logo based on the concepts.

#### First idea



GoodBI

This was the first idea for our logo, based on the chart concept. We did not choose this logo as it did not satisfy the visibility and versatility criterias.

#### Second idea



This was the second idea for our logo, based on the chart and hand wave concepts. While it did satisfy the visibility and versatility criterias, it did not feel recognisable enough.

#### Third idea



This was the third idea for our logo, based on the hand wave and lightning bolt concepts. Similarly, we did not pick this as it did not feel recognisable enough, and the lightning bolt in the palm looked imbalanced.

We decided on the following logo



This logo comprises a bulb and a hand wave. The bulb captures the idea of “business intelligence” and the hand wave represents our word play on “goodbye”. For the wordmark, we chose a geometric sans-serif font to ensure it looked professional and readable at various resolutions. At the same time, this wordmark features subtle curves to bring across the user-friendly of our platform. As for the brand color, we chose a medium shade of red, reflecting the urgency and importance of the insights our software provides. This choice also underscores the powerful capabilities of our solution in driving data-driven decision making.

## Brand Identity

We also picked a secondary color complementary to the red to be used on our website or marketing materials.

A graphic showing the primary and secondary colors, and the logo font.

Primary  
**BitterSweet**  
**#FF6052**

Secondary  
**Majorelle Blue**  
**#695AF2**

Logo font  
**Figtree**  
**Extra Bold**

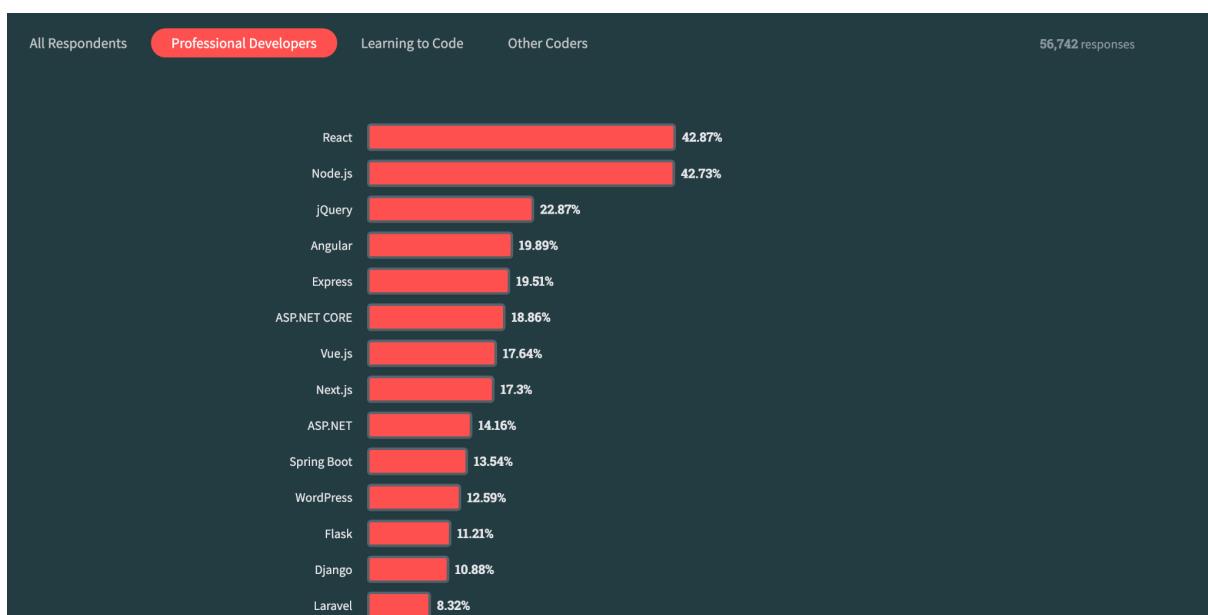
Below the color palette, there are seven small "GoodBI" icons in a row.

# Milestone 11: Choice of Technologies

## User Interface

Our front-end tech stack is built on Next.js. We evaluated several major front-end frameworks such as React, Svelte, Vue and Angular. However, given the short timeframe to complete the project, we decided to opt for a React-based framework due to our team's familiarity with React. Since most of us had prior experience with React, we were able to hit the ground running much more quickly with React. The component-based architecture of React is also useful to help us to break down the interface into smaller, reusable components which makes the application more maintainable as it grows to include more business logic.

Another benefit of React is that it is a mature and well-established framework. React has widespread adoption including by industry leaders such as Apple, Netflix, Airbnb and many others. According to the Stack Overflow Developer Survey 2023, React is the most common web technologies used by professional developers, much more than competing frameworks like Angular and Vue.



The large developer community provides us with access to numerous open-source libraries built upon React that can streamline our workflow and accelerate development. This also meant that it would be much easier to get help for any issues we may face during development. In the future, if we were to continue developing our product, it would also be easier to hire developers who are familiar with React, rather than other frameworks due to the larger community.

In addition, React is an open-source project actively maintained and supported by Meta (formerly Facebook), one of the largest technology companies in the world. This backing ensures that React stays up to date with the latest developments. At the time of writing, React 19 is currently in development, and we believe React will be supported for many more iterations to come.

We also opted to use Next.js over React on Vite or Remix. Next.js has several additional features on top of React that we felt might be useful:

1. Built-in routing: While packages like [React Router](#) exist, having it baked directly into Next.js reduces the amount of setup and code to write to manage routes manually.
2. Optimisation features: Next.js has built in image optimisation using sharp or squoosh, which improves the website's performance and user experience, and reduces the need for us to do it manually. This also helps to improve the search engine optimization of our site.
3. Developer documentation: Next.js has very good developer documentation, which helps us to implement features quickly.

Next.js is also an open-source project in active development by Vercel, a cloud platform as a service provider. While some developers have raised concerns with Next.js and Vercel, such as [vendor lock-in](#) by strongly integrating some features such as edge functions with Vercel's hosting environment, we felt that this was not an issue given the active community support and documentation on deploying on other platforms (though more setup may be required when deploying on other platforms). Moreover, Next.js, being built on React, also provides a flexible foundation to transition to other frameworks like Remix or Astro if the need arises. Many of these alternative frameworks provide documentation to assist in such transitions, like this [one from Astro](#), helping to mitigate the risk of vendor lock-in.

For styling, we chose TailwindCSS together with shadcn. We considered other alternatives such as styled UI frameworks like Bootstrap, Chakra UI and MUI among others. However, We felt that TailwindCSS was more suitable as we had built a Figma mock-up with custom designed components, and using styled UI frameworks like Bootstrap may incur additional setup time by overriding their default styles.

We also considered other unstyled UI libraries such as Bulma, but we picked Tailwind over due to our familiarity with TailwindCSS's utility-first classes approach over Bulma's semantic classes approach. The familiarity also helps other members of the team more quickly understand and modify styles as required.

We also briefly thought of going with regular CSS, however, we were concerned with readability and setup time. Using pure CSS requires more setup time if we were to use postCSS to include plugins such as autoprefixer. Writing TailwindCSS code is also much faster than writing CSS code, since it reduces the need to write media queries and encapsulates CSS into a class.

As an added benefit, TailwindCSS also integrates well with Next.js and other popular frameworks, making it extremely versatile, even if we decide to move to another framework in the future. TailwindCSS also has a simple configuration file to expand the classes and styles to suit our design easily. If ever necessary, porting out of TailwindCSS is not an issue. Since it is largely an encapsulation over pure CSS, moving from TailwindCSS can be done by converting the utility classes into a CSS file, though this might be a time consuming process as the application expands.

One common issue brought up is the HTML bloat that possibly comes with TailwindCSS, which may impact readability.

```
<button
  class='inline-flex items-center justify-center whitespace nowrap rounded-md
          text-sm font-medium ring-offset-background transition-colors
          focus-visible:outline-none focus-visible:ring-2 focus-visible:ring-ring
          focus-visible:ring-offset-2 disabled:pointer-events-none disabled:opacity-0'
>
</button>
```

This was also the reason why we decided to include another layer of encapsulation using shadcn. Shadcn is a collection of reusable components based on Radix Primitives and TailwindCSS, which can be copied into our project. Since we had custom styling in mind, we only considered unstyled libraries such as HeadlessUI, Radix Primitives, react-aria and shadcn. Such libraries also help us to handle accessibility issues.

HeadlessUI was an appealing option to us as it was built by the same team behind TailwindCSS, however, the selection of components from HeadlessUI is very limited (there are only 16 components provided at the time of writing), and would likely require another library to complement it, hence we decided against it.

React-aria was another option that we looked into for its flexibility, and that it offers more components than HeadlessUI. However, we were put off by the slight learning curve in comparison to Radix Primitives and shadcn, which is a concern for us given the tight timeline of the project.

Finally, shadcn was picked over Radix Primitives as it offers slightly more components, and included some sensible default styling that was easy to modify out of the box with TailwindCSS. Shadcn was also preferred since we could choose what components to include into our application, rather than installing all components at one go. Our main worry with using Shadcn would be the future development of Radix Primitives, however, we felt that it was the most suitable for our product at this point of time. In addition, even if components change or stop receiving support, the modular design of Radix Primitives also allows us to port to other components one by one.

## Database

For our backend, we chose a PostgreSQL database server deployed on Aiven, a free service to deploy relational databases. Postgres, a relational database, was chosen over NoSQL as they are easier and most commonly used by large enterprises who wish to glean insights and trends into their data, thus mandating that their data be structured in a tabular form.

Aiven was chosen over cloud services like AWS RDS and GCP Postgres Managed database due to its ease of deployment and troubleshooting errors.

To test locally, we built our own PostgreSQL image that we ran. Given that tables are created dynamically (upon the user uploading their data), we are not using an ORM but rather embedding raw SQL Queries in the endpoints, creating schemas for each user and adding tables to accommodate their datasets.

## Authentication

Our authentication system is made using [SuperTokens](#) as it provides a managed service to manage users and creates a useful abstraction over Session Authentication. We also considered implementing OAuth but did not choose it because of:

1. Complexity and Overhead: OAuth is primarily designed for third-party app authorization and adds unnecessary complexity when you only need authentication for your app's users. The multi-step OAuth flow is more suited to applications where resources from other services are accessed, which isn't needed for internal user sessions in a Business Intelligence app.
2. User Experience: OAuth often requires redirects and external permission grants, which can disrupt the user experience. In contrast, session-based authentication with SuperTokens keeps the login flow simple and entirely within the app.
3. Security Considerations: OAuth is prone to vulnerabilities like token leakage if not implemented correctly, especially when dealing with implicit flows. SuperTokens, on the other hand, handle session management securely without exposing the risks associated with managing refresh tokens manually, as seen in some OAuth implementations.
4. First-Party Application Focus: Since GoodBI is focused on internal BI users, managing our own user's sessions with SuperTokens makes more sense than OAuth, which is designed more for third-party authorization rather than first-party authentication.

By choosing SuperTokens, we ensure a streamlined, secure, and less complex authentication flow suited to our app's specific needs.

Furthermore, its libraries ensured that our backend code quality was much cleaner, as we did not have to implement authentication endpoints.

## LLMs

For GoodBI, our GenAI Business Intelligence app, we initially explored several approaches for using LLM technology to handle user queries effectively.

1. **Initial Choice:** Our first approach was to format the tabular data, write it to a README, and pass it along with the user's query to the assistant. However, this approach had downsides. The client data was often large, causing GPT to hallucinate more frequently when handling vast amounts of information. Additionally, it could easily exceed the token limit, and the role of the LLM wasn't very specialized. We also noticed that GPT's SQL queries were prone to errors.
2. **Second Approach:** We then considered using a network of OpenAI assistants, each performing a specific task. While this worked well, we realized that it would be difficult to switch models later on if we wanted to replace the LLM agent with a different one.
3. **Our Final Choice:** We decided to build a knowledge graph of LLM agents, each managed through LangChain wrappers. This approach gave us a flexible interface to direct the model on what to do while allowing us to specialize the prompts. By specializing each agent's task, we significantly reduced the likelihood of hallucinations.

## Engineering Methodology

1. We initialized all LLMs in the knowledge graph with the heaviest but most accurate model (GPT-4), ensuring that we didn't overdo prompt engineering early on. This also gave us the flexibility to scale down to smaller models over time.
2. We created a clear pipeline that defined each step involved, with every step assigned to a specific LLM.
3. We provided response templates to streamline the forward-feeding of information between agents, making the process more efficient and reducing the chances of errors.

## Pipeline

- **User uploads table:** The user uploads a table (e.g., CSV) to the platform.
- **MetadataAgent:** The table is ingested by the MetadataAgent, which generates short descriptions for each column header to provide context for downstream agents.
- **User makes a query:** The user submits a query related to the dataset.
- **PruningAgent:**
  - The query, along with column headers, descriptions (metadata), and the first 3 rows of the table, is passed to the PruningAgent.
  - The PruningAgent analyzes the query to identify key columns and extract important noun phrases relevant to the query.
- **SQLAgent:**
  - The parsed data from the PruningAgent is sent to the SQLAgent.
  - The SQLAgent first generates the SQL query needed to retrieve the correct data.
  - It then validates the query to ensure accuracy and correctness.
- **Query execution:** The validated SQL query is executed, retrieving the necessary data.
- **InterpreterAgent:**
  - The query results are passed to the InterpreterAgent, which summarizes the data in a human-readable format for the user.
- **GraphTypeDeterminerAgent:**
  - The query results are also fed into the GraphTypeDeterminerAgent.
  - This agent determines the most appropriate graph type for visualizing the data based on the user query and data characteristics.
- **DataFormatterAgent:**
  - The selected graph type, along with the SQL query results, are sent to the DataFormatterAgent.
  - The DataFormatterAgent formats the data into a presentable structure suitable for graph generation.
- **KPISuggesterAgent (optional step):**
  - The KPISuggesterAgent takes the metadata of the table and outputs potential KPIs that are relevant to the dataset.

This pipeline ensures efficient handling of user queries and optimal presentation of data, from raw ingestion to summarized results and graphical visualizations.

# Hosting

## Front-end hosting

For the front-end hosting, we ultimately decided on Vercel, primarily because of the seamless integration with Next.js, which is the framework we chose to build GoodBI on. Vercel provides an optimised environment for Next.js and enables other features such as image optimisation and API routes for performance and development efficiency.

We did consider other platforms such as Cloudflare Pages, Netlify, and GitHub pages, but several factors led us to favour Vercel over these platforms. We only considered platforms with a free starting cost to reduce our business expenses in the MVP stage.

Netlify is a platform that is similar to Vercel. Netlify offers many similar features, such as one-click deployment and branch deployments. It also supports Next.js. However, we felt that it lacks the seamless integration provided by Vercel, and has a slightly lower bandwidth limit in comparison to Vercel.

Cloudflare Pages offers good integration with Cloudflare's CDN and security features and also supports Next.js. However, we found the setup process to be a bit more complicated in comparison to Vercel. Furthermore, in the future, If we were to use Next.js's server-side rendering features, the setup process on Cloudflare is much more challenging.

GitHub Pages is the simplest out of the four platforms, and has good integration with GitHub where our code is hosted on. However, it lacks many of the advanced functionality required for more dynamic websites like GoodBI, and we felt it was more suitable for static websites.

At the same time, we acknowledge that Vercel has its downsides as well. As our project scales, Vercel's pricing plans can become very expensive costing \$20/user/month minimally on their Pro plan. Moreover, due to the zero-configuration approach, Vercel offers very limited control over infrastructure. This means there is little flexibility in configuring servers or optimising performance settings. Despite these downsides, given our current scale, we felt that Vercel matched our existing needs the best out of all the platforms. As we scale, we might consider moving to platforms such as Microsoft Azure or Amazon Web Services.

## **Backend API and Database Hosting**

For the backend, we deployed a FastAPI application using Render because it offers a seamless deployment experience with automatic scaling and easy configuration. Render provides built-in support for FastAPI, which allows us to focus on building the application without worrying too much about infrastructure management. Additionally, its pricing model is transparent (free) and cost-effective, making it a great fit for startups and smaller projects. The automatic SSL, custom domains, and automatic deployment from our github repository also made it an appealing choice for us to manage everything in one place.

We considered hosting our app on fly.io. However, it could be deployed only from the command line, making it unsuitable for automatic deployment from our github repositories whenever we pushed to main. Render also had better documentation and user support.

As mentioned before, we used Aiven for hosting our Postgres Server for our users' tables and managed our user authentication info through a Managed Supertokens Core that populated a PostgresDB every time a new user signed up to our app.

# Milestone 12: Common Workflows

## Workflow 1: Data ingestion

This workflow pertains to situations where users wish to add more data to the platform. There are two scenarios: first, when users want to add data to a new dataset; and second, when users want to add data to an existing dataset.

Both scenarios are performed using a similar workflow.

1. Enter the 'Datasets' page.
2. Click on 'Add data' on the page, and it will bring you to the add data page.
3. Enter a dataset name in the combobox
  - a. To create a new dataset, enter a name for the dataset, and select create. The names for datasets are unique, i.e. a new dataset cannot have the same name as an existing dataset.
  - b. To add to an existing dataset, search for the name of an existing dataset and select it.
4. Enter a description for the dataset if required.
5. Select a CSV file to upload as the data.
6. Click 'Add data' to add the data, a toast will show if it is successful.

An alternative workflow would be to separate the two scenarios into two sub-workflows.

### Creating a new dataset

1. Enter the 'Datasets' page.
2. Click on 'Create dataset' on the page.
3. Enter a new dataset name.
4. Enter a description if required.
5. Select a CSV file as the input data.
6. Click 'Create dataset' to create the dataset.

### Add to an existing dataset

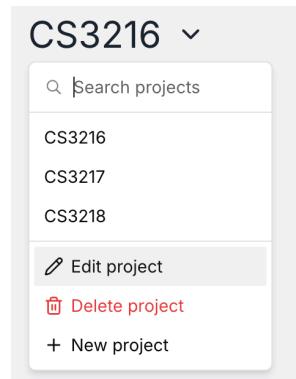
1. Enter the 'Datasets' page.
2. Find an existing dataset, either using search or manually, and click on the button on the top right and select 'Add data' and a modal will open.
3. Select a CSV file as the input data.
4. Click 'Add data' to add the data to the existing dataset.

We felt that the first workflow where both scenarios were merged met the usability aspect of user experience much more strongly. The merged workflow was easier to learn (learnability) and easy to remember how to use (memorability), since there is only one single flow for adding data. The first workflow is also more effective and efficient to use compared to having two separate sub-workflows. We also included toast components to adhere to Nielsen's Heuristic of "Visibility of system status", so users know that the data has been successfully added.

## Workflow 2: Switching between projects

This workflow is for when users want to switch to another project on their account to view different insights.

1. Enter the 'Dashboard' page.
2. Click on the current project name on the top left.
3. Search and select the desired project from the dropdown.



This workflow is important as there are multiple projects that a business owner might be involved in, and they might want to separate the insights from different projects. Initially, we considered having a separate project page which links to a group of insights.

A screenshot of the GoodBI dashboard. At the top, there is a navigation bar with the GoodBI logo, 'Projects', 'Dataset', and user profile icons. Below the navigation is a search bar and a 'Create project' button. The main area is titled 'Project' and lists two projects: 'Project 1' (uploaded on 10 Jan 2024) and 'Project 2' (uploaded on 15 Jan 2024). Both projects have a small ellipsis icon next to them.

However, we decided against this flow because adding another screen would complicate the user experience. Introducing another step to view their insights might lead to confusion, especially for first-time users. A more streamlined process is essential for enhancing usability by increasing effectiveness, enabling users to gain access to insights in as few clicks as possible.

## **Workflow 3: Insight creation**

This workflow is for creating new insights on the user's dashboard. This is the core functionality of our application, and is where the generative AI is used.

1. Enter the 'Dashboard' page.
2. Click 'Add new insight' on the page.
3. Enter the inputs required.
4. Select 'Create' to go to the preview page.
5. Select 'Confirm' on the preview page to add the insight to the dashboard.

We also explored the possibility of consolidating this workflow into a single modal, eliminating the separate preview page. While having a single modal makes it faster for users to generate insights, given the AI-driven nature of this feature, we decided against this flow. A preview page brings significant benefit for an application like ours (which is discussed [below](#)).

Additionally, a dedicated page for insight creation provides a foundation for future feature expansions. This separate space would allow us to introduce new features, such as a visual node-based insights creator, without disrupting the existing user flow. Users can seamlessly learn and adapt to these new functionalities within the familiar context of the insight creation page.

# Milestone 13: UI Considerations

## Data Privacy Transparency

In artificial intelligence applications, users are often concerned with how their data is being used and processed. This is even more important in our domain as we deal with sensitive business data. In conformance with [OECD AI Principle 1.3 Transparency and Explanability](#), and [Google's guideline on “Be transparent about privacy and data settings”](#), we have included several sections in our user interface to clarify how data is being processed.

We have included a “Frequently Asked Questions” section on our landing page for information on how we handle their data.

### FAQ

## Frequently Asked Questions

- What is GoodBI? ▾
- How does GoodBI work? ▾
- What type of data is supported? ▾
- Is my business data safe? ▾
- How much does it cost? ▾

For users seeking more detailed information about our data processing, we have included a Terms of Service and Privacy Policy page that outlines the specifics of our data guidelines.



Login

Sign Up

## Privacy Policy

Last updated 21 September 2024

### About

At GoodBI, (“GoodBI”, “we”, “us”, or “our”), we are committed to protecting your privacy. This Privacy Policy outlines how we collect, use, share, and protect your information when you use our services (“Services”). This includes interactions such as:



Login

Sign Up

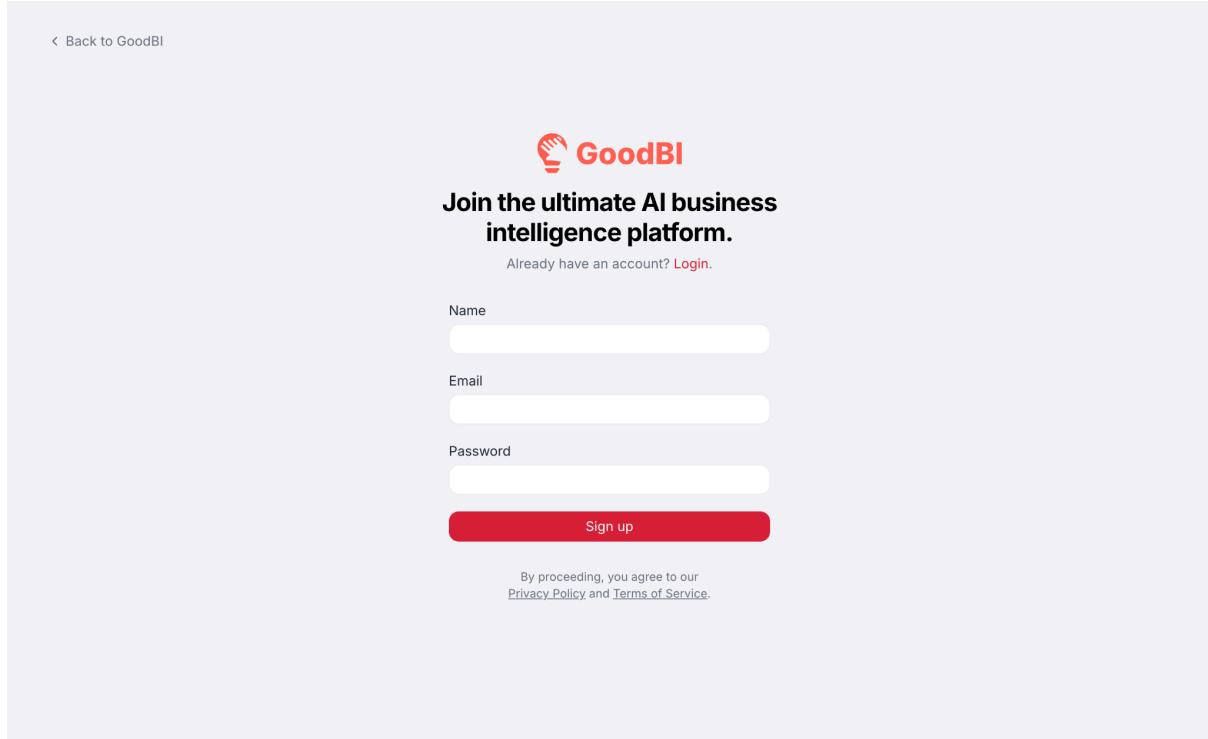
## Terms of Service

Last updated 21 September 2024

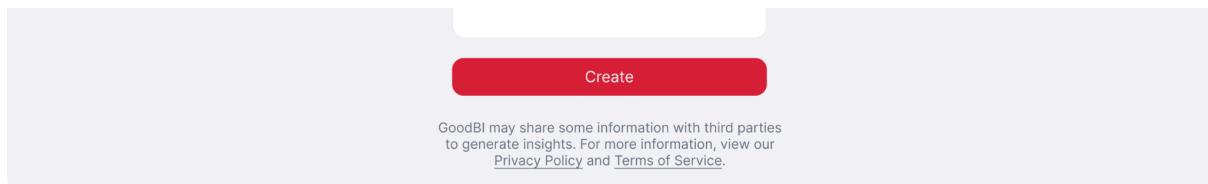
### About

Welcome to GoodBI (“GoodBI”, “we”, “us”, or “our”). By accessing or using GoodBI, you agree to comply with and be bound by these Terms of Service. If you do not agree to these terms, please do not use GoodBI.

Our sign up and login pages also provide direct links to the Privacy Policy and Terms of Service, placed below the “Sign up” and “Login” buttons where users will clearly see it. We considered placing this information right at the bottom of the screen, but we felt that users might miss it, thus, we placed it near the buttons where users are most likely to press on.



We have also included a data warning below the page to create insights, since that is when some user data is passed into AI in our web application.



## Minimal Form for Creating Insights

Following the guideline “[Automate more when the risk is low](#)”, we have decided to simplify our form for creating insights and reduce the number of required input fields. This is meant to enhance user experience and efficiency on our platform.

GoodBI will automatically generate recommended insights, which will be displayed on a preview page ([discussed below](#)). Users will still have the flexibility to control additional parameters on the preview page if necessary, before adding the insights to their dashboards. We find that this approach has the best balance between automation and user control.

In our first iteration, we planned for the input form to include all possible fields. This was because we wanted to give users more control, and also to provide more information in our text prompt to AI to ensure higher accuracy and correctness.

### Create a new insight

Dataset

Type of chart

Date range

Y range

Title

KPI description

GoodBI may share some information with third parties to generate insights. For more information, view our [Privacy Policy](#) and [Terms of Service](#).

However, as we worked on the generative AI integration, we found that our implementation was able to generate fairly accurate suggestions even without several of the input fields. As such, in our second iteration, we decided to move this flexibility over to the preview page, and reduce the number of input fields required to enhance efficiency.

Create a new insight

Title

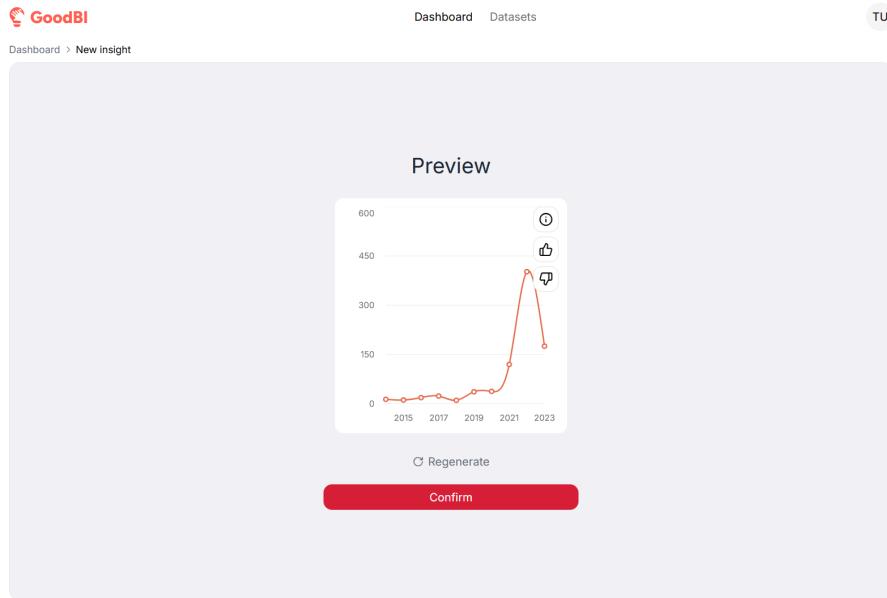
KPI description

**Create**

GoodBI may share some information with third parties to generate insights. For more information, view our [Privacy Policy](#) and [Terms of Service](#).

## Preview Page

To adhere to Google's guideline "[Let users supervise automation](#)", we have implemented a preview page to display the generated insights before they are added to the user's dashboard.

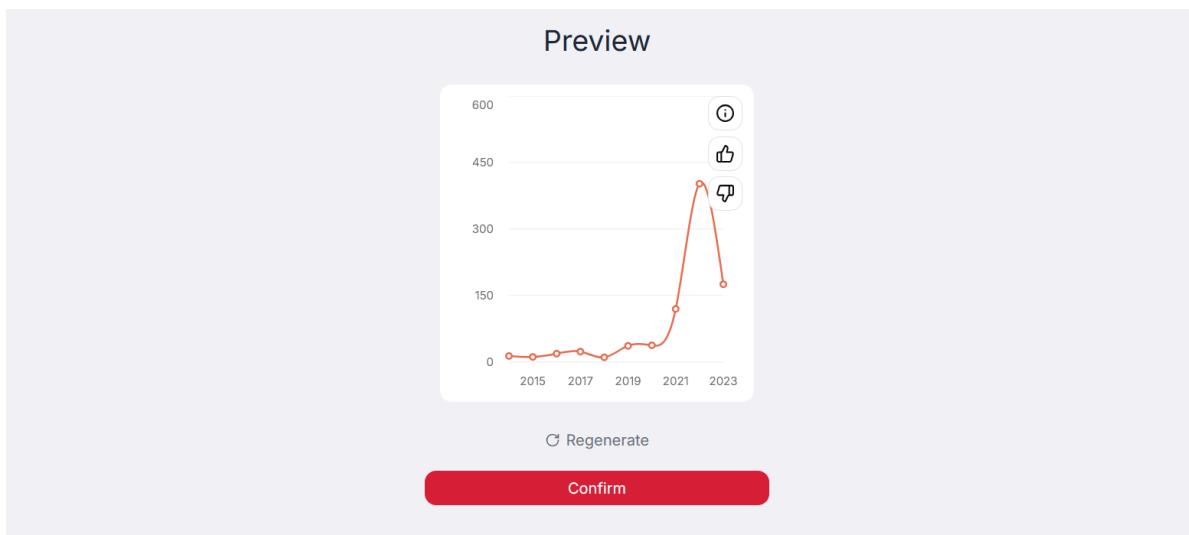


The preview page presents the generated insights in a clear format, allowing users to examine the output and assess its correctness and accuracy. By offering this layer of oversight, users can catch and correct any potential errors produced by the application, ensuring that only accurate insights are added.

Furthermore, this complies with Google's guideline on "[Give control back to the user](#)". The 'Regenerate' button provides a way for users to correct the system by prompting the system with extra inputs like chart type.

## Feedback Mechanism

To further enhance user supervision of automation and align with Google's guidelines of "[Let users give feedback](#)", we have integrated a feedback mechanism directly into the preview page, allowing users to provide immediate feedback, both positive and negative on the insights generated. This feedback feature empowers users to flag any inaccuracies, suggest improvements, or affirm correct insights, ensuring a continuous feedback loop that holds the system accountable for errors as outlined in Google's "[Be accountable for errors](#)" guideline.



When users provide their feedback on the preview page, they are promptly informed that their input has been received and will be reviewed by our team. For negative feedback, we also prompt users to contact us about the issue, so we can respond to them as quickly as possible. This approach creates a transparent process that reinforces accountability and responsiveness. We not only improve the quality of insights over time, but we also build user trust by demonstrating that their feedback actively shapes the performance and accuracy of our platform.

## Use of Simple and Practical Language

In line with Google's guidelines "[Set the right expectations](#)" and "[Explain the benefits, not the technology](#)", we made a conscious effort to minimise technical jargon across our website's user interface, including key areas such as the landing page.

This helps to set the right expectations by communicating to users what our platform does and how it can add value without overwhelming users with confusing and complex terms that they may not understand. By focusing on quantified, practical benefits, we guide users to understand how our platform directly addresses their needs.

For example, we initially wrote these descriptions under the two of the three core features in the landing page.

- AI-powered: Instantly generate key insights from your data, using the latest large-language models from OpenAI.
- Data visualisations: Get real-time data visualisations from the click of a button.

However, upon reviewing these descriptions, we felt that they did not set the right expectations, and used more technical jargon than necessary and some users may not understand the language.

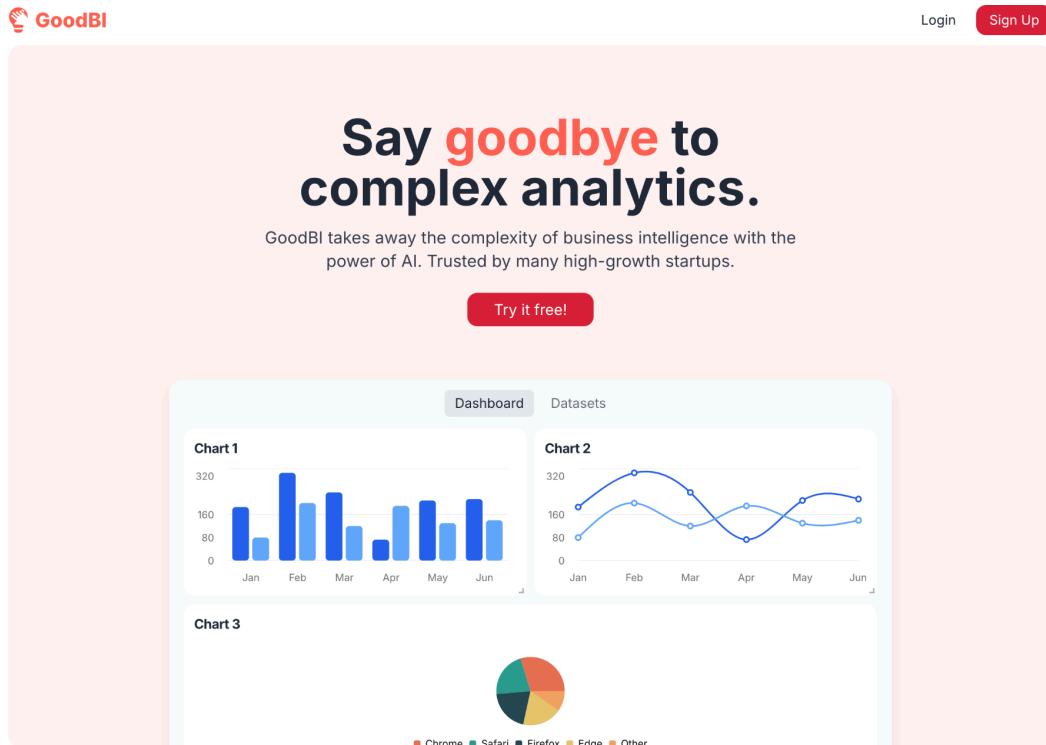
Hence, we reworded these points into the following:

- AI-powered: GoodBI learns from your data, and generates actionable key insights from your data in a few minutes to empower your business goals.
- Data visualisations: Unlock dynamic, real-time data visualisations at the click of a button! Transform your data into bar, line, or pie charts effortlessly.

This improves user comprehension and builds confidence in GoodBI, making potential users to grasp our core value proposition and feel more connected to the product.

## Milestone 14: Landing page

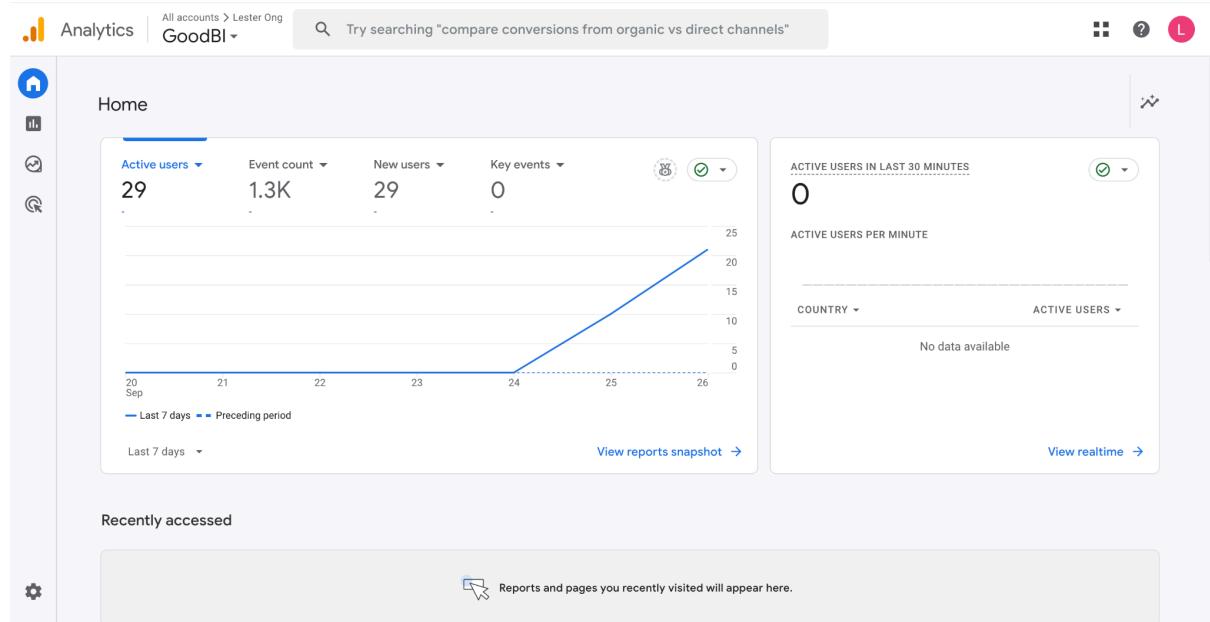
The landing page can be found at <https://good-bi.vercel.app>.



The landing page comprises several sections: hero, product (what our product can do), features of our product including integration and security, pricing and a frequently asked questions section.

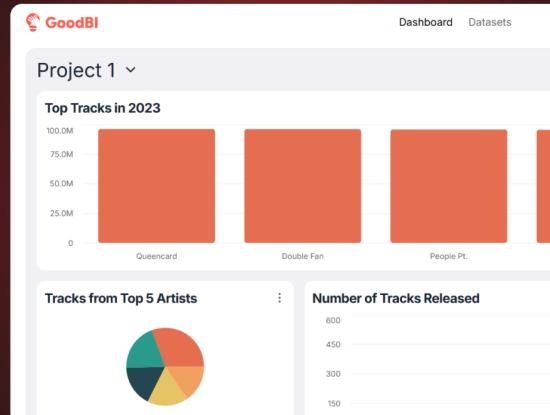
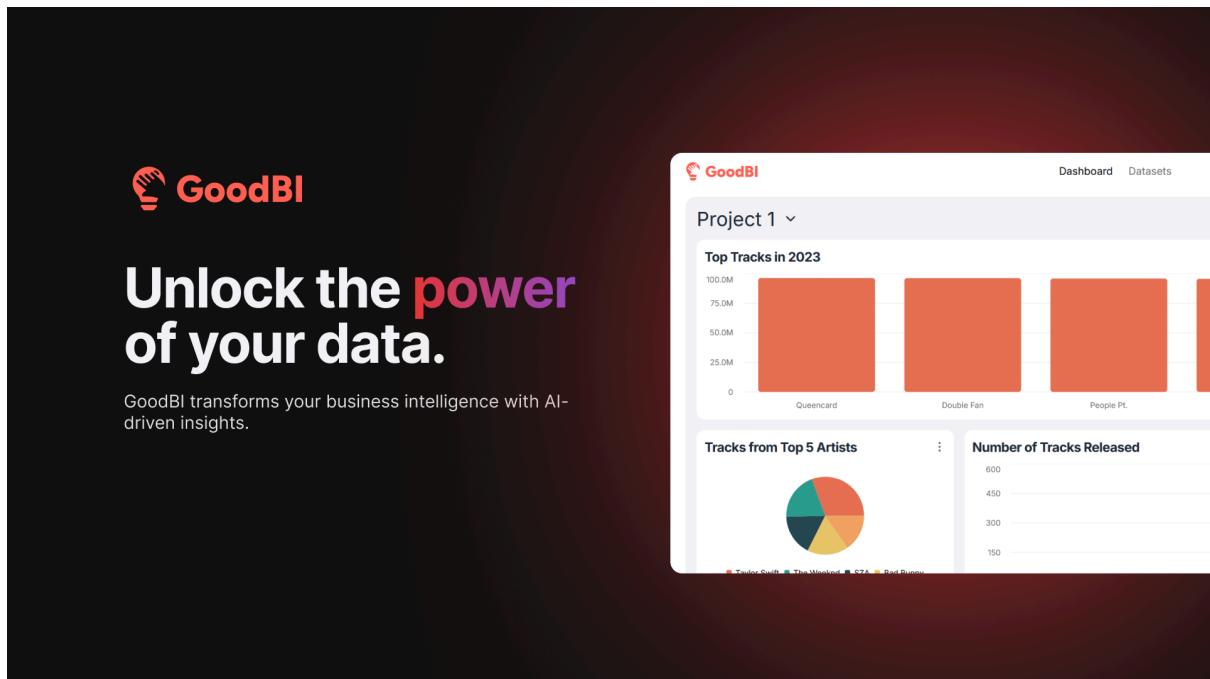
## Milestone 15: Analytics

We have embedded Google Analytics into our website. At the time of writing, it appears we have gained significant traction, with a total of 29 active users on our website over the last week.



## Milestone 16: Marketing

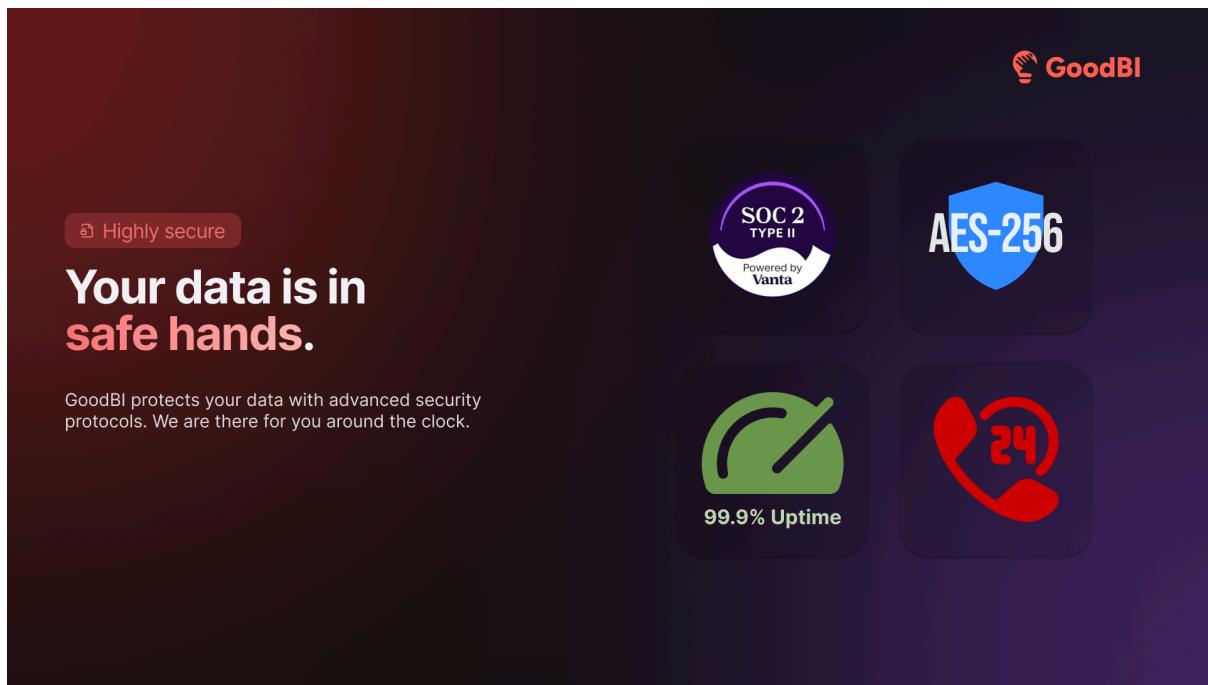
We created 5 screens for our marketing, highlighting the key features of GoodBI.



This image shows a preview screen for GoodBI. On the left, there is a pie chart with four segments. Below the pie chart, there is a modal window titled "Create a new insight". The modal contains fields for "Title" and "KPI description", both of which are currently empty. At the bottom of the modal is a red "Create" button. In the bottom right corner of the main preview area, there is a purple button labeled "Powered by AI". To the right of the preview area, there is a call-to-action text: "Generate insights in with a single click." Below this text, a smaller note reads: "GoodBI empowers you to make your insights in minutes, ensuring you stay ahead of your competition."

The GoodBI mobile application interface features a dark background with a red header containing the GoodBI logo. Below the header, a red button labeled "Data-driven insights" is visible. The main content area displays a line chart titled "Number of Tracks Released" with data points for January through May. The chart shows values of approximately 222, 320, 222, 11, and 222 respectively. To the right of the chart is a summary bar with three colored segments: green, orange, and red.

The GoodBI dashboard interface features a dark background with a red header containing the GoodBI logo. The main content area displays two charts: a bar chart on the left showing categories like "ouble Fan", "People Pt.", "Everything", and "fukumean" with corresponding orange bars; and a line chart on the right titled "Number of Tracks Released" showing a sharp increase from 2014 to 2021, peaking around 450 tracks released in 2021. A search bar and an "Add new insight" button are located at the top of the dashboard. To the right of the charts, a blue button labeled "Customisable" is shown. Below the charts, a text block states: "Put insights where they belong. GoodBI's dashboard is fully customisable. Put your insights wherever you want them to be."

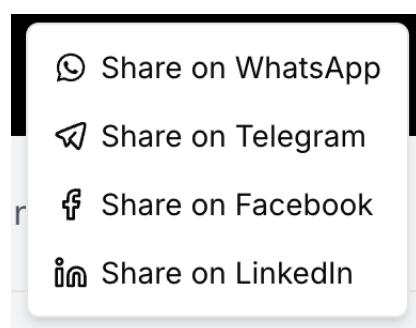


## Milestone 17: Embeddings

We did not use embeddings in our application.

## Milestone 18: Social Networks

As a platform primarily targeted towards businesses, integrating social plugins can help enhance user engagement and expand networking opportunities to attract more users. To cater to the needs of our business audience, we have selected the following social platforms for integration: LinkedIn, Facebook, WhatsApp and Telegram. The share links are found in the footer when logged out, and under the profile dropdown when logged in.



LinkedIn is the leading business social network, boasting over 900 million users globally<sup>5</sup>, and is projected to grow further. This makes it a powerful tool to connect with industry professionals and other potential clients for GoodBI. By integrating LinkedIn, we allow our existing users to share insights directly from our platform, showcasing our capabilities and driving organic awareness among industry peers.

According to Hootsuite, 91% of the organisations in the world have a presence on Facebook<sup>6</sup>. This widespread adoption underscores the platform as a significant tool to create increased brand exposure of our platform to other businesses which are not onboard with us yet.

The integration of WhatsApp and Telegram sharing allows us to connect with niche and smaller businesses that prioritise direct, real-time communication within their industry circles.. Smaller businesses often operate in tight-knit communities where word-of-mouth recommendations and direct communication are crucial. With this integration, we allow these businesses to easily introduce our platform to their peers, driving awareness and adoption in sectors that may otherwise be harder to reach through traditional platforms such as LinkedIn or Facebook.

While we prioritised the integration of popular social platforms to enhance the visibility and engagement of GoodBI, we made a deliberate choice not to include X (formerly Twitter). This decision was influenced by the negative press surrounding the social network, including concerns over hate speech, platform instability, controversial policy changes and an overall decline in user trust.

These issues have raised questions about the platform's reliability and suitability for business-focused connections leading us to focus on the aforementioned social networks that align more closely with our values of user privacy and safety. By avoiding integration with X, we aim to maintain the credibility of our platform and provide a trusted environment to connect with other businesses.

---

<sup>5</sup> <https://www.statista.com/forecasts/1147197/linkedin-users-in-the-world>

<sup>6</sup> [https://hootsuite.widen.net/s/mgqjjznhxs/hootsuitesocialtrends2024\\_report\\_en](https://hootsuite.widen.net/s/mgqjjznhxs/hootsuitesocialtrends2024_report_en)