**תמונה שמכילה עיצוב

התיאור נוצר באופן אוטומטי**

**User1st Ltd**

**A**pplication **R**equirements **D**ocument

Data-Driven Marketing Intelligence Platform

MarketinGuru

**Ltd**

**Evyatar Kopans**

**Yuval Goldstein**

**Peleg Ben-Barak**

**Ron Franco**

November 26, 2024

Version 1.1

**1. Introduction**

**1.1 Vision**

*"Without data, you're just another person with an opinion." – William E. Deming.*

The concept of an Ideal Customer Profile (ICP) has been a cornerstone of marketing for decades. Basically, it's the answer to the question *"Who will buy my product/ services?".* Back then, before the emergence of data-driven methods and super computers, traditional approaches often relied on intuition and broad generalizations, leaving too much to chance. In today’s competitive world however, where sales are as essential to companies as air is to life, data-driven methods have revolutionized marketing, transforming it into a precise, measurable, and highly impactful discipline. These advances allow businesses not only to understand their customers better but also to craft strategies that resonate deeply and deliver tangible results.

Our project, *MarketinGuru*, harnesses this revolution. It is more than just a platform; it is our gateway to intelligent, data-driven marketing. By combining advanced algorithms and tools with comprehensive data analysis, *MarketinGuru* will enable us to redefine our company’s approach to marketing, provide an unprecedented ability to identify and adapt to our Ideal Customer Profile with precision, ensuring that every marketing effort is purposeful and effective.

Because at the end of the day, every company is a marketing company. Whether you sell software, shoes, or services, success depends on how well you understand your customers and meet their needs. With *MarketinGuru*, we’re not just keeping up with the competition—we’re setting the standard. This platform empowers us to connect meaningfully with our customers, adapt to market changes, and drive sustained growth. It’s not just about marketing; it’s about transforming how we think, operate, and succeed in the modern world.

**תמונה שמכילה טקסט, צילום מסך, עיגול, גרפיקה

התיאור נוצר באופן אוטומטיתמונה שמכילה טקסט, צילום מסך, עיגול, גרפיקה

התיאור נוצר באופן אוטומטיתמונה שמכילה טקסט, צילום מסך, עיגול, גרפיקה

התיאור נוצר באופן אוטומטי**

**1.2 The problem domain**:

In contemporary marketing teams within companies, there is often difficulty in targeting the right clients for a product and determining which actions will attract or repel the ideal customer. While a trial-and-error approach works for some companies on a certain scale, effective targeting requires analytical tools that support informed marketing and development decisions. These decisions, in turn, drive company growth and increase revenue.

There are three main challenges in marketing intelligence analysis:

1. **Information gathering:**

The first challenge is identifying reliable sources for marketing intelligence, distinguishing important data from redundant information, and categorizing it effectively—tasks that are not easy to accomplish. Making informed marketing decisions requires massive amounts of data, but obtaining and organizing this data is complex. Common sources include social media, market trend analyses, and analyses of active user behavior, among others. Finally, organizing this data in a concise and analyzable manner presents another challenge, as it's impossible to consider all collected data in every marketing decision.

1. **Make correct analysis of the data:**

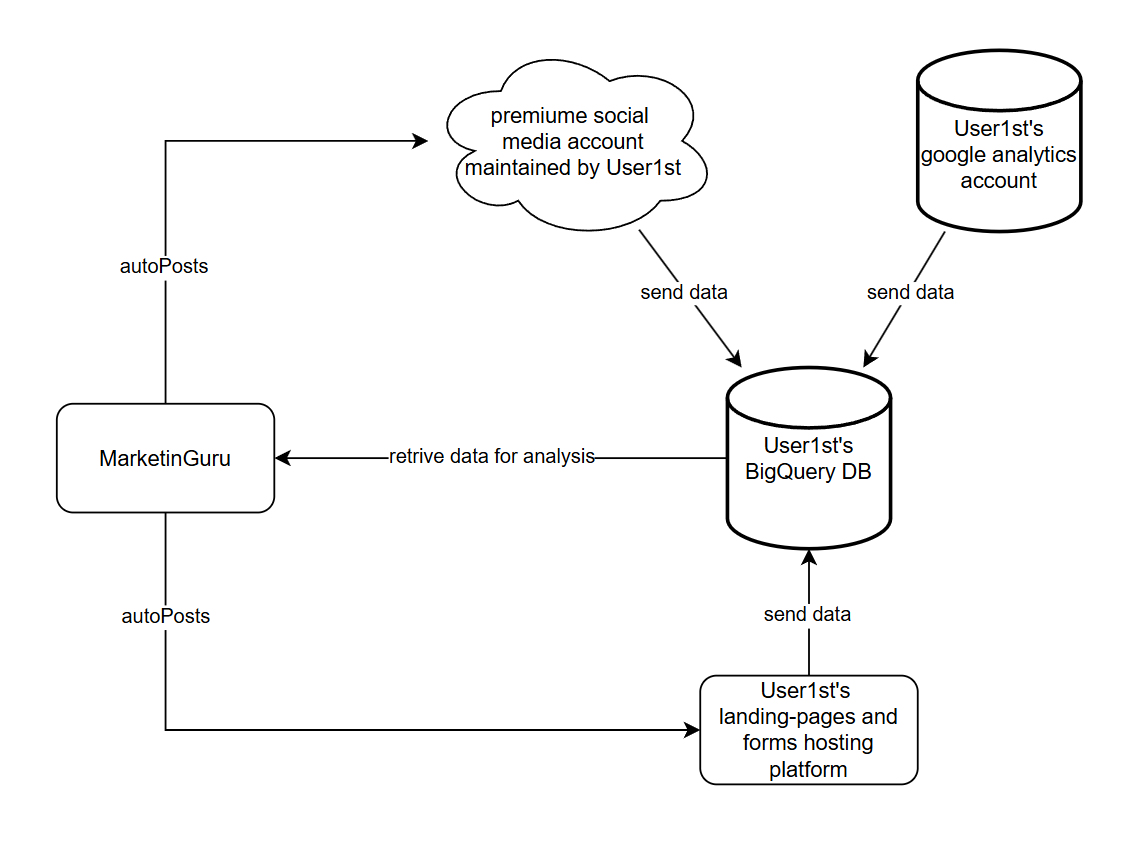
Once data is collected, using it wisely to make predictions is essential but challenging. Assessing whether these predictions have been “optimal” (i.e., the most productive) can sometimes be an unanswerable question. Additionally, deriving the correct conclusions from the data is not a trivial task.

1. **Presenting the information in a coherent and understandable way**

Converting large datasets into a human-readable format can be challenging and highly subjective; what may be clear to one person could be incomprehensible to another.

The goal of our project is to provide marketing professionals with a framework that automates parts of the analysis process. This includes auto-collection of relevant information from various sources, automated analysis of this data, and presenting the analysis results in a readable and understandable way. Of course, some of the challenges mentioned above will shift from problems faced by the marketing professional to problems we, as the developers of the product, will need to solve.

**1.2 Context:** What should be added to this diagram?



* **BigQuary**: is a Google product to which User1st subscribes. It is a data warehouse that provides fast querying capabilities and a platform for data analytics. User1st sends a large amount of relevant information from various components (described in the diagram) to this database. Our system will integrate with BigQuery to retrieve the data necessary for our operations.
* **Google analytics**: User1st also subscribes to Google Analytics, another Google product. This service provides insights into traffic analysis, audience demographics, acquisition tracking, behavior analysis, and more. Information from Google Analytics is sent to BigQuery, which our system then retrieves.
* **Landing pages and forms**: User1st maintains landing pages and forms for user acquisition and registration. Information from these pages is also sent to BigQuery for later retrieval. Our system will allow new landing pages to be posted with customized design and text at the click of a button.
* **Social media accounts**: User1st manages several social media accounts. Data gathered from these accounts—such as campaign success, liked posts, and demographics of primary visitors—is also sent to BigQuery for later use. Our system will communicate with these accounts, enabling posts to be made across multiple social media platforms simultaneously to ease the workload for the marketing team and maintain consistency across accounts.

**1.3 Goal:**

The main goal of our project is to provide a marketing analysis framework for the User1st marketing team that is personally tailored to their needs. Success will be achieved if our platform identifies the ideal customer profile (ICP) for User1st and simplifies the day-to-day work of marketing personnel.

**The tool will offer several key functionalities, including:**

* **Automatic Data Collection**: Gather data from multiple sources, such as social media campaigns, frequently searched keywords on search engines, and demographics of active and potential users.
* **Automated Data Analysis**: Analyze the collected data to identify the ICP and detect trends in the dynamic market.
* **Customizable Data Presentation**: Display data in customizable charts and diagrams for the convenience of the marketing team.
* **Landing Page Auto-Posting**: Provide an interface to automatically create and post customizable landing pages to increase User1st’s digital footprint, thereby enhancing data collection.
* **Automated Social Media Posting**: Enable automatic posting on multiple social media platforms simultaneously, further augmenting the data collected.

The main challenges we will address include integrating with User1st's existing data collection methods (such as Google Analytics, premium social media accounts, User1st landing pages, and forms) and enhancing these sources. We will also focus on filtering essential information from irrelevant data, developing a suitable dashboard for the marketing team, and integrating with User1st’s social accounts and landing page hosting to enable automation tools.

**Once the core system is implemented and major obstacles are addressed, we envision expanding the system as follows:**

* **Social Media Crawlers**: Develop crawlers to scan social media platforms, identifying potential influencers for User1st campaigns and potential clients.
* **AI-Powered Market Prediction**: Utilize state-of-the-art AI techniques to forecast market shifts and evolving ICPs.
* **ROI Analysis**: Analyze campaign return on investment (ROI) to detect when a campaign has reached the point of diminishing returns.
* **User Action Analysis**: Examine actions taken by users on the User1st site, identifying behaviors that may deter purchases.
* **Cookie-Based Behavior Analysis**: Use cookies to analyze user behavior and facilitate targeted advertising.

**1.4 Stakeholders:**

* **User1st’s Marketing Team (Informed Customers):**

The marketing team will be the primary user of our system. Requirements are primarily elicited from them, and they provide feedback on whether a feature is redundant, important, or highly useful.

* **User1st’s R&D Team (Technology Experts):**

User1st's research and development team will help us determine which customer-suggested requirements are feasible given the company’s resources, which would be too costly to maintain, and what can be achieved within the project's timeframe. The R&D team will also assist us in seamlessly integrating our system with the company’s existing infrastructure.

* **User1st’s Legal/ Compliance Department (Law Experts):**

The legal department will help us ensure that proposed features and requirements are lawful and will provide guidance on adjustments if any elements are non-compliant.

Anyone else?

**1.5 Software context:**

**Input:**

As mentioned, the primary input for our system is user data collected from various sources. The following list provides a guideline for the types of data our platform will gather:

* From google analytics:
* **Age, Gender, Location**: can help with understanding the ICP
* **Device and Browser Type**: devices type (mobile, desktop) and browsers (Chrome, Safari, etc.) that users prefer helps in optimizing website design and functionality.
* **Referral Sources**: where users come from (direct, search engines, social media, other websites). can help in understanding which channels drive the most traffic.
* **Page Views and Session Duration**: helping assess content effectiveness.
* **Bounce Rate**: Indicates if users leave immediately after visiting a page, which can reveal issues with relevance and user experience.
* **Conversion Rate**: Measures the percentage of users who complete specific goals, such as signing up, purchasing, or subscribing.
* **Click Path Analysis**: Tracks the sequence of pages visited by users, helping identify common paths or drop-off points.
* **Frequency and Recency**: Shows how often and how recently users return, helping gauge brand loyalty.
* **Interest Categories**: Using Google’s affinity data, marketers can understand users' broader interests (e.g., travel, fitness), which helps in targeted ad placements and content strategy.
* From social media:
* **Demographics**: understanding the potential user age, gender, location, job titles, income, education level, behaviours, interests, and values. This helps in refining the ICP.
* **Likes, Comments, Shares, and Reactions**: helps in understanding the ICP and the effectiveness of a campaign
* **Click-Through Rate**: The percentage of people who click on your ad or post out of the total number who saw it. Can gauge campaign effectiveness
* **Ad Placement Insights**: how ads performed across different placements (e.g., feed, stories, sidebar), helping optimize campaign decisions.
* **Behavioral Data**: Insights into how users interact with your content (e.g., time spent, clicks, engagement with specific parts of your post or video) help optimize future campaigns.
* From landing pages and forms (ask what data can be collected?):

Another form of input is the auto-posting tool which is pretty straight forward:

* Auto-posting on social media: takes text of the post, an optional image, and a list of social networks to which we want to post.
* Auto-posting landing pages: takes the customizable element of a landing page (text, color, logo, images and so on) and generate a new landing page with those properties.

**Processing:**

The primary processing function of the software involves making predictions based on collected data through machine learning and statistical analysis. Examples of potential conclusions include identifying the ICP, determining which campaigns are successful, discovering keywords that attract the target audience, and identifying locations where campaigns succeed or fail.

The AI models to be used are yet to be determined.

**Output:**

There are four main output channels:

* **Dashboard Display**: Collected data is presented in the software UI (dashboard) in a customizable way, allowing each user to select specific charts and query the system accordingly.
* **Marketing Suggestions**: Based on collected data, the system provides a set of predefined recommendations regarding the target audience, campaign success, and more.
* **Social Media Posts**: Posts generated by the auto-posting mechanism.
* **Landing Pages**: Generated by the auto-posting landing page mechanism.
* Add Interface
* Add Events
* Talk about server responsibilities?

**TABLE OF CONTENT**

1. **Usage Scenarios** 
   1. User Profiles — The Actors
   2. Use-cases
   3. Special usage considerations

**The Actors – who are the actors?**

* **Marketing Staff**  
  Marketing staff are primary users of the MarketingGuru platform, benefiting from its analytical and targeting capabilities to enhance strategic decisions. They use the platform for:
  + Optimized user analysis, enabling tailored marketing strategies.
  + Identification of target user candidates through data-driven profiling.
  + Enhanced product development and market positioning insights.
  + Improved search engine optimization strategies.
* **MarketingGuru System Server**  
  The MarketingGuru server operates as an automated marketing assistant, processing and analyzing marketing data in real-time to support staff with insights throughout the marketing lifecycle.
* **BigQuery**  
  While not a traditional "actor," BigQuery serves as a pivotal backend service, enabling high-performance data analysis for MarketingGuru. Through BigQuery, the platform can efficiently handle large volumes of data for in-depth analytics.
* **External Data Retrieval Services**  
  MarketingGuru integrates with various external sources, including Meta, Google, and LinkedIn, to collect data on user behavior, demographics, and popular search terms. This integration allows MarketingGuru to maintain an up-to-date understanding of the market and refine its targeting algorithms.
* **Customers (lead engagement)**

**Use Cases** – with regard to s.e principles, does the marked categories correctly describe the necessary use-case

* **Campaign Management via API**
  + Post-campaign data retrieval for individual social networks.
  + Simultaneous posting and analysis across multiple social platforms.
* **Data Analysis in BigQuery**
  + Using BigQuery for running specific queries to uncover insights. Sample queries could focus on engagement metrics, conversion rates, and demographic analyses.
* **Data Visualization in Looker**
  + Displaying analysis results with customizable visualizations for reporting and strategy refinement.
* **Data Aggregation from Multiple Sources**
  + Gathering user interaction data from social networks, Google Analytics, and landing pages, and uploading it to BigQuery for consolidated analysis.
* **Landing Page Optimization** / Generation
  + Implementation and tracking of UTM parameters on landing pages.
  + Creation and deployment of landing pages using a no-code solution, followed by server integration for lead generation tracking.
* **Additional Functionalities** (NTH)
  + Automate personalized communication by sending forms via email to potential leads, facilitating engagement through direct and customized outreach.

**Special Usage Considerations – anything else?**

* **Compliance and Privacy**
  + Ensure adherence to data privacy regulations (such as GDPR and CCPA) when retrieving and handling user data.
* **Data Quality and Accuracy**
  + Regular checks on data quality and updates to ensure accuracy in target profiling.
* **Scalability and Performance**
  + As data volume grows, maintain system performance and scalability by optimizing BigQuery queries and Looker visualizations.

**2. Functional Requirements**

|  |  |  |
| --- | --- | --- |
| Title | Requirement | Explanation/notes |
| Social media posting | The system should allow the marketing team to post a campaign to Meta, LinkedIn and Google within one click. | A campaign is defined as a paid advertising action which User1rst pays for to Meta, Google or LinkedIn. The system should allow the user a selection between these alternatives.   - What if we have different features across the different platform? - How is the bidding process done – what if each platform has a different bidding algorithm? How should we choose between different features?  - How do we fund and pay for the campaign? |
| Data collection | The system should collect User’s Geographic location, gender, age, education, proffession data from Google Analytics, advertising campaigns (LinkedIn, Google, Meta?) The system should collect the same (?) data from manually filled forms and landing pages. | How do we compare between different data points/vectors? Each platform may have its own statistics -For Google Analytics User1st has an implementation – should we use it or build a new one? Should we build something similar for Meta and LinkedIn. |
| Forms | The system will analyse x, y, z from the forms the potential clients fill out. | Should we use backpropogation on data from Google Analytics? Do we have feature vectos set up? Is our goal to find the output for each feature? Is there an algorithm in mind? |
| Data visualisation | The system will display x, y, z from BigQuery |  |
| Data visualisation | The system should display x, y, z on the ICP dashboard | Working with Looker |

Nice to haves – or Must?

|  |  |
| --- | --- |
| AI Model | The system will calculate a lead score for potential customers. |
| Data collection | The system will contain a stateful key-value storage for potential clients (key- UserID, value – cookies) |
|  |  |
|  |  |
|  |  |

**Non-Functional Requirements**

Performance (Speed, Capacity, Throughput, etc.) – are there any specific Performance attributes we must abide to?

Quantify the performance requirements for your software system.

Campaign posting should take no longer than 1 minute after a click.  
Should not use an exponential time algorithm.  
May be irrelevant.

Reliability & Stability – same question for all non-functional requirements

Is your system required to withstand certain hardware, software, network failures? Is your system required to support data recovery, error-correction, etc? describe and quantify the factors that effect the reliability & stability of the software system.

Once published – the system needs to be available at least 95% of

SAFETY AND SECURITY

The website should have https connection.

On which computer/server should the system run?  
What security measures do we need to take?

Where is the data stored?

What can we do with the users’ data?

How do we handle and secure users’ data?

PORTABILITY

The system should run on Google Chrome 131 or newer.

USABILITY

The application is to be web based.

AVAILABILITY

The system should be available at least 12 hours a day, between 8am and 8pm Israel Standard Time? (working time)

**5. Risk Assessment & Plan for the Proof of Concept**

**Overview**

To assess the risks in our system’s design, we must break down each functionality and analyze it individually. This process will deepen our understanding of the design and allow us to explore a broader range of implementation options for our current high-level system architecture.

**תמונה שמכילה טקסט, תרשים, קו, צילום מסך

התיאור נוצר באופן אוטומטי**

In this chapter, we’ll focus on the seven main bottlenecks in our design. The goal here is to create a proof-of-concept (PoC) prototype that tackles each bottleneck separately, testing how well our components work together—things like campaign posting across platforms, pulling in data from LinkedIn and Google Analytics, and visualizing it in the ICP Dashboard. The PoC will let us experiment with limited functionality to see if we can solve specific technical challenges, especially around data management, visualization, and security.

Building this PoC will give us a clearer sense of what’s needed, help us understand the limitations and strengths of tools like BigQuery, and allow us to make smarter design choices upfront. We’ll outline test cases for each bottleneck, covering happy, sad, and bad paths, to show how the PoC will reduce project risks. Ultimately, this step will help us get a real feel for what the final system might look like and build confidence in the design before diving into full development.

|  |  |  |
| --- | --- | --- |
| **Risk Category** | **Risk Description** | **Mitigation Strategy – w/ Ariel** |
| **Integration for Cross-Platform Posting** | Difficulty in implementing simultaneous posting across LinkedIn, Meta, and Google. | Test each platform individually in the PoC to resolve compatibility issues before full integration.  **Test using** platform-specific API calls and logging responses. |
| **Data Retrieval Challenges** | Difficulty in retrieving data from LinkedIn, Meta, Google Analytics, and other sources. | Establish connections to each source in the PoC to confirm data access and permissions.  **Test using** API response verification for each source and error handling for permission issues. |
| **Data Presentation and Visualization** | Difficulty in presenting data clearly for marketing teams, especially using Looker. | Create basic Looker visualizations in the PoC; gather feedback to refine dashboard functionality.  **Test using** sample data visualizations and feedback from end-users on usability and clarity. |
| **Data Privacy and Security** | Risk of handling personal user information securely. | Set up basic encryption and access controls in the PoC; test role-based permissions, third-party authentication mechanism (googleAuth, Facebook Login..)  **Test using** simulated access attempts by different roles and check for unauthorized data access attempts. |
| **Complexity of Working with BigQuery** | Learning curve and complexity in setting up BigQuery. | Run a small dataset through BigQuery in the PoC; familiarize team with its features.  **Test using** simple queries on sample data to assess performance and data handling efficiency. |
| **Challenges in Identifying ICP** | Difficulty in accurately identifying the Ideal Customer Profile (ICP). | Test ICP detection on sample data; refine methodology iteratively.  **Test using** sample customer data, analyzing ICP algorithm accuracy and refining based on results. |
| |  | | --- | | **Implementing a NoCode Solution for Landing Pages** |  |  | | --- | |  | | |  | | --- | | Challenges in building a flexible, user-friendly no-code tool for landing page creation. |  |  | | --- | |  | | Develop a basic no-code prototype in the PoC to test functionality and gather user feedback on usability.  **Test using** sample landing page builds and collect feedback from non-technical users. |



**Plan for the Proof of Concept – please review the current plan (not finished, ETA's not accurate right now) and comment any ideas for testing/implementation/ "functionality already exists"**

Our PoC will validate the core components of "MarketingGuru" in a manageable, phased approach. Each phase targets specific high-risk areas, so we can address potential blockers early. Given our team of four developers, I’ve outlined the necessary resources and time for each step.

**Risk 1: Integration for Cross-Platform Posting**

**Objective:** Verify that we can post campaigns simultaneously across LinkedIn, Meta, and Google.

1. **Implement Individual Platform Posting**
   * **Steps:**
     + Start with a LinkedIn API integration.
     + Implement basic campaign posting functions (e.g., text posts, image uploads).
     + Repeat the process for Meta and Google.
   * **Testing:**
     + Use API response verification and logging to confirm successful posts.
     + Track potential issues with rate limits and incompatible formats.
   * **ETA:**
   * **Developers Required:** 2
2. **Combine Platforms for Simultaneous Posting**
   * **Steps:**
     + Develop a single function that can trigger campaign posting across all platforms.
     + Ensure API calls are synchronized and handle errors for each platform independently to avoid interruptions.
   * **Testing:**
     + Test simultaneous posting with small sample data.
     + Monitor for rate limit issues and API error handling.
   * **ETA:**
   * **Developers Required:** 2

**Risk 2: Data Retrieval Challenges**

**Objective:** Confirm our system can retrieve data from all sources (LinkedIn, Meta, Google Analytics, user forms) and consolidate it for use.

1. **Establish Data Retrieval for Each Source**
   * **Steps:**
     + Set up API connections for LinkedIn, Meta, Google Analytics, and user forms.
     + Build basic functions for retrieving data (e.g., campaign engagement metrics, user-submitted forms).
   * **Testing:**
     + Test each data retrieval function individually, verifying data integrity.
     + Confirm authentication and permissions for each source.
   * **ETA**
   * **Developers Required:**
2. **Data Consolidation and Error Handling**
   * **Steps:**
     + Create a centralized function that consolidates data from all sources.
     + Develop error handling to manage data inconsistencies or missing fields.
   * **Testing:**
     + Test with mixed data from all sources, checking that data is accurately combined.
     + Simulate permission issues to verify error handling.
   * **ETA:** 1 week
   * **Developers Required:** 1

**Risk 3: Data Presentation and Visualization**

**Objective:** Use Looker to present consolidated data in a clear, actionable format.

1. **Set Up Looker Integration**
   * **Steps:**
     + Connect Looker to the PoC database containing consolidated data.
     + Build sample visualizations (e.g., engagement rates, campaign performance).
   * **Testing:**
     + Test visualizations using real data.
     + Collect feedback from the marketing team on the clarity and usability of visualizations.
   * **ETA:** 0.5-1 week
   * **Developers Required:** 1-2
2. **Refine Dashboard Based on Feedback**
   * **Steps:**
     + Adjust visualizations and add new metrics based on feedback.
   * **Testing:**
     + Validate that updates improve clarity and functionality.
   * **ETA:** 1 week
   * **Developers Required:** 1

**Risk 4: Data Privacy and Security**

**Objective:** Ensure data storage and access controls meet security standards, focusing on user privacy.

1. **Implement Basic Security Framework**
   * **Steps:**
     + Set up encryption for data storage.
     + Implement basic role-based access controls (e.g., admin vs. regular user).
     + Use third-party tool for authentication (GoogleAuth, Facebook Login… )
   * **Testing:**
     + Run tests with different roles to verify access permissions.
     + Confirm encryption functionality and check for unencrypted data leakage.
   * **ETA:** 1 weeks
   * **Developers Required:** 2
2. **Review and Stress-Test Security Protocols**
   * **Steps:**
     + Conduct internal security audits.
     + Simulate unauthorized access attempts to test system robustness.
   * **ETA:** 1 week
   * **Developers Required:** 2

**Risk 5: Complexity of Working with BigQuery**

**Objective:** Confirm BigQuery’s suitability for large-scale data processing, specifically for querying and analysis.

1. **Initial BigQuery Setup**
   * **Steps:**
     + Configure BigQuery with a sample dataset.
     + Develop queries to analyze key metrics (e.g., campaign reach, user engagement).
   * **Testing:**
     + Test queries for speed and accuracy.
     + Monitor performance with increasing dataset size.
   * **ETA:** 1 week
   * **Developers Required:** 1
2. **Optimize Query Performance**
   * **Steps:**
     + Experiment with different indexing and query optimization techniques.
   * **Testing:**
     + Benchmark query performance and refine based on test results.
   * **ETA:** 1 week
   * **Developers Required:** 1

**Risk 6: Implementing a NoCode Solution for Landing Pages**

**Objective:** Develop a simple no-code landing page generator to validate ease of use for non-technical users.

1. **Develop Basic No-Code Landing Page Creator**
   * **Steps:**
     + Create a drag-and-drop interface for landing page creation.
     + Include basic components like text, images, and call-to-action buttons.
   * **Testing:**
     + Conduct usability tests with marketing staff.
     + Collect feedback on functionality and ease of use.
   * **ETA:** 2 weeks
   * **Developers Required:** 2
2. **Refine Based on Feedback**
   * **Steps:**
     + Make improvements to the user interface and add any necessary features identified in feedback.
   * **Testing:**
     + Test updated features and confirm improvements in user experience.
   * **ETA:** 1 week
   * **Developers Required:** 2

**Risk 7: Challenges in Identifying ICP**

**Objective:** Test and refine algorithms for identifying the Ideal Customer Profile (ICP) from collected data.

1. **Preliminary ICP Detection Algorithms**
   * **Steps:**
     + Develop initial algorithms to identify customer profiles based on key metrics (e.g., engagement, demographics).
   * **Testing:**
     + Run algorithms on sample data to evaluate accuracy.
     + Analyze results for potential adjustments.
   * **ETA:** 2 weeks
   * **Developers Required:** 2
2. **Refine ICP Algorithms**
   * **Steps:**
     + Refine algorithms based on initial results, focusing on precision and relevance.
   * **Testing:**
     + Rerun tests and validate improvements in identifying meaningful customer profiles.
   * **ETA:** 1 week
   * **Developers Required:** 2