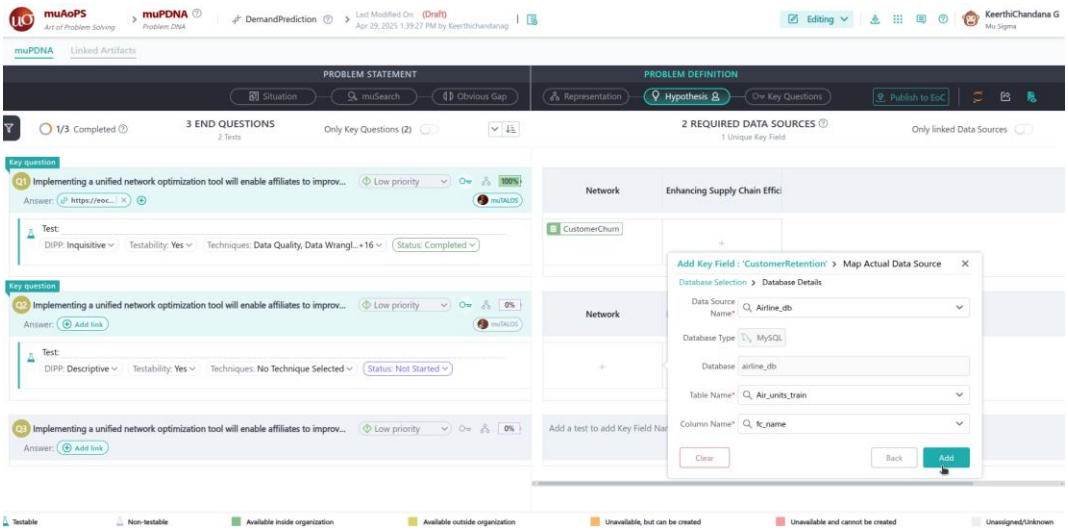
- Username
- Password

- Use Test Connection to verify the setup

If Selecting an Existing Database:



(Fig 115: Adding Existing Database)

- Choose the existing database from the dropdown
- Select the Table Name and Column Name required for mapping
- Complete the linking by confirming these selections

Once this step is complete, the data source is successfully connected to Key Question.

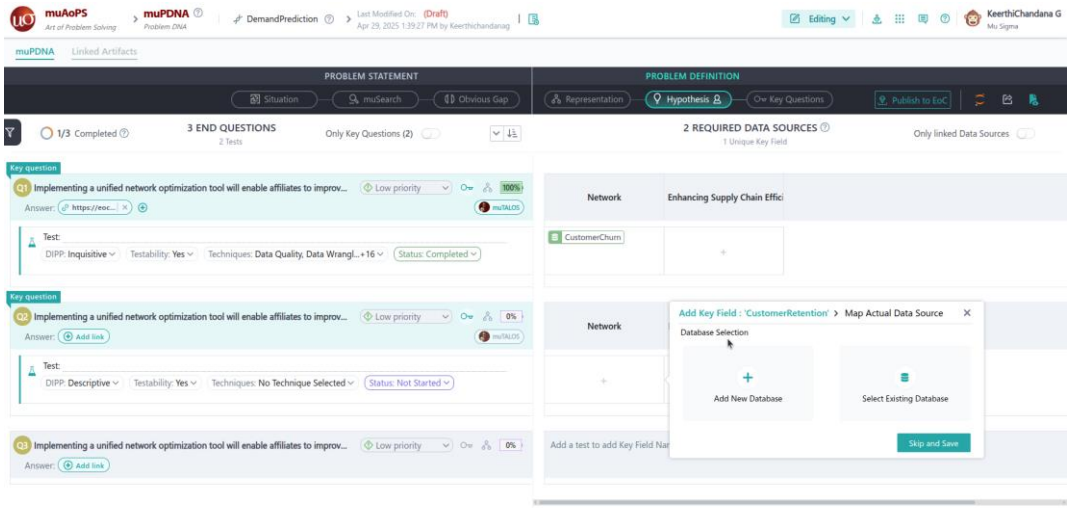
6.muTalos Button Activation and Next Steps

With the test setup and data connection completed:

- The muTalos button remains active next to the Key Question
- Clicking this button now launches the muTalos interface, where the system takes over to automate the reasoning, code generation, and execution process

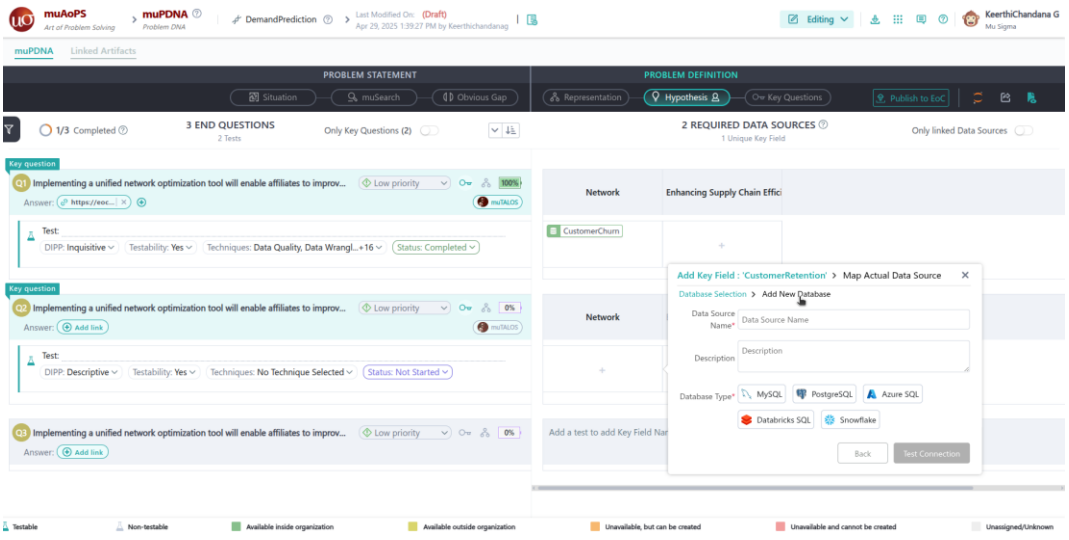
From this point forward, the AoPS to Code process moves into the muTalos environment, where problem configuration, reasoner setup, and automated testing are executed seamlessly.

Clicking this muTalos button initiates the transition from the hypothesis-building phase to the automated reasoning and code execution phase within muTalos.



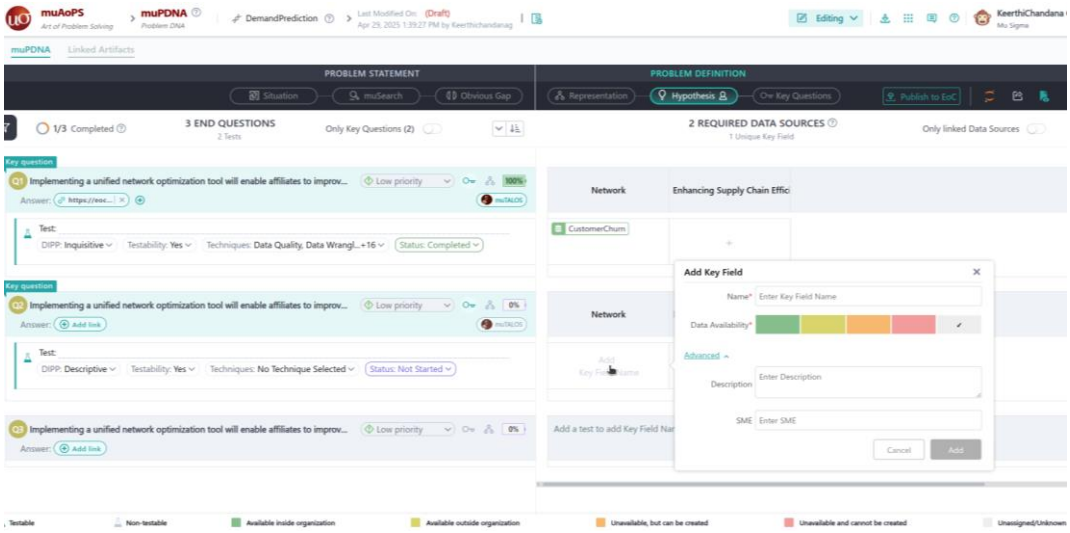
(Fig 113: Add Data Source)

If Adding a New Database:



(Fig 114: Adding new Database)

- Enter the Data Source Name and Description
- Choose the Database Type (MySQL, PostgreSQL, Azure SQL, Databricks SQL, Snowflake)
- Fill in the connection details:
  - Host
  - Port
  - Database Name



(Fig 112: Add Key Field)

- Enter the Key Field Name relevant to the test (for example, Customer Churn)
- Define the Data Availability by selecting one of the available statuses:
  - Available inside the organization
  - Available outside the organization
  - Unavailable but can be created
  - Unavailable and cannot be created
  - Unassigned/Unknown
- Provide a brief Description of the data field
- Mention the SME (Subject Matter Expert) responsible for this field

5.Connecting to a Database for the Key Field

After defining the key field:

Proceed to map the actual data source by selecting the method of database connection.

The system offers two options:

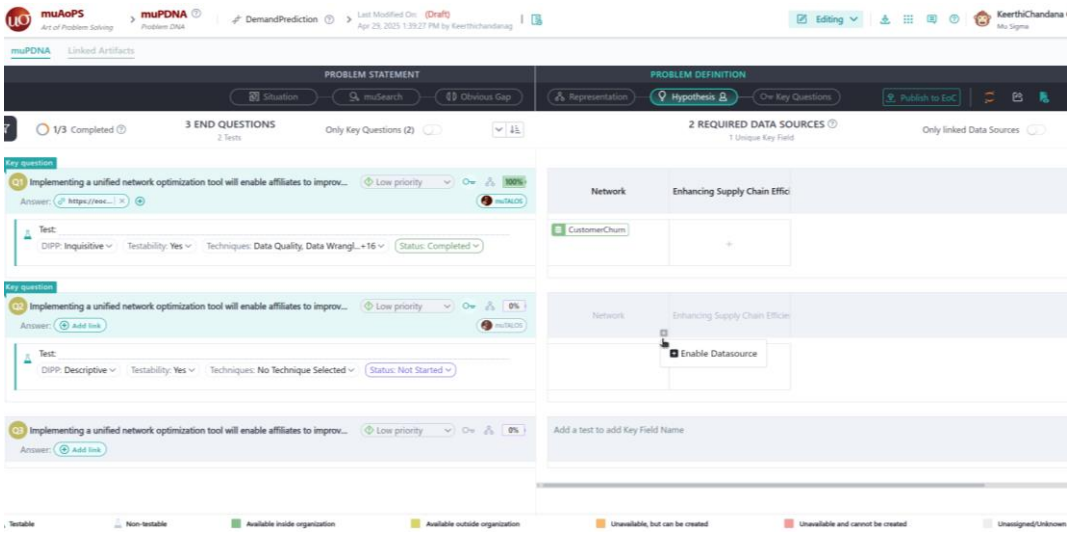
- Add New Database (for first-time connection setup)
- Select Existing Database (if the database is already available)

- Define Testability by selecting either Yes or No
  This indicates whether the hypothesis can be tested through available data and analytical methods
- Choose the appropriate Techniques from the dropdown list
  Techniques may include Data Quality, Data Wrangling, Statistical Analysis, Predictive Modeling, Segmentation, and several others

4.Enabling Data Source for the Test

After configuring the test:

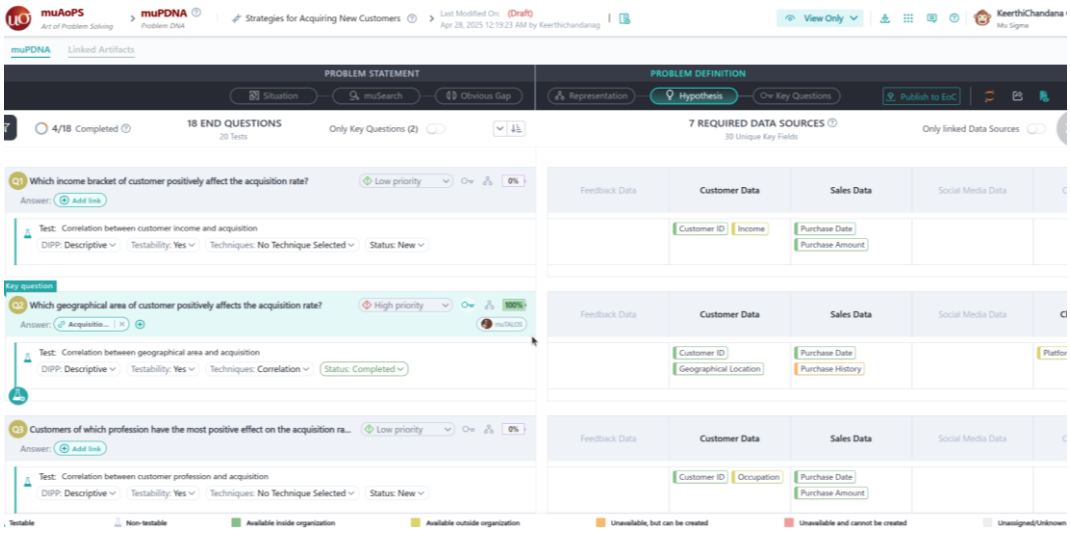
- Navigate to the Required Data Sources section located next to the test configuration
- Click on the box to enable the data source for the question



(Fig 111: Enable Data Sources)

After enabling the Data source, you need to add a Key Field. To do this, click on the + icon to add it.

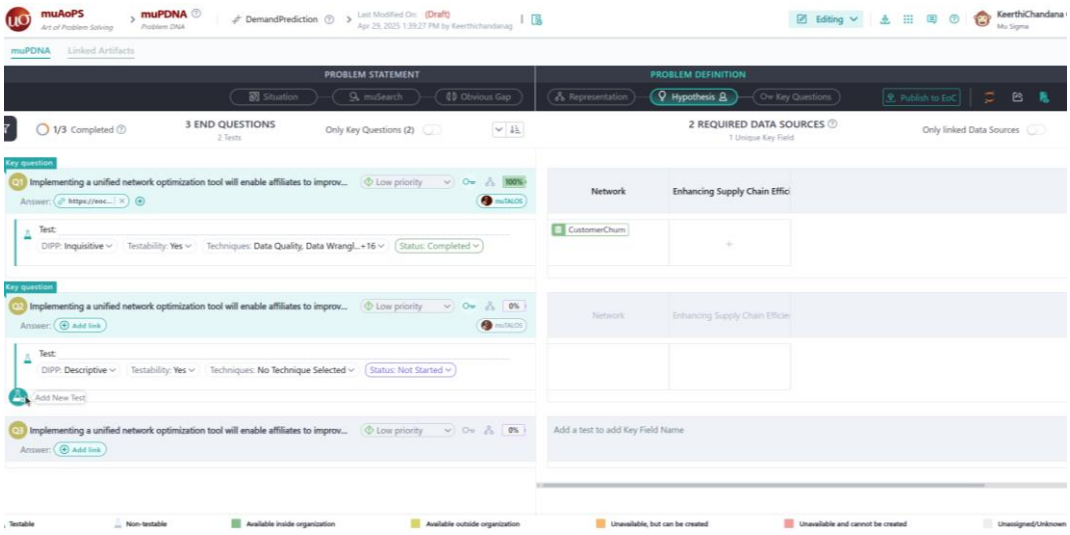
A pop-up titled “Add Key Field” will appear. In this window:



(Fig 109: Hypothesis)

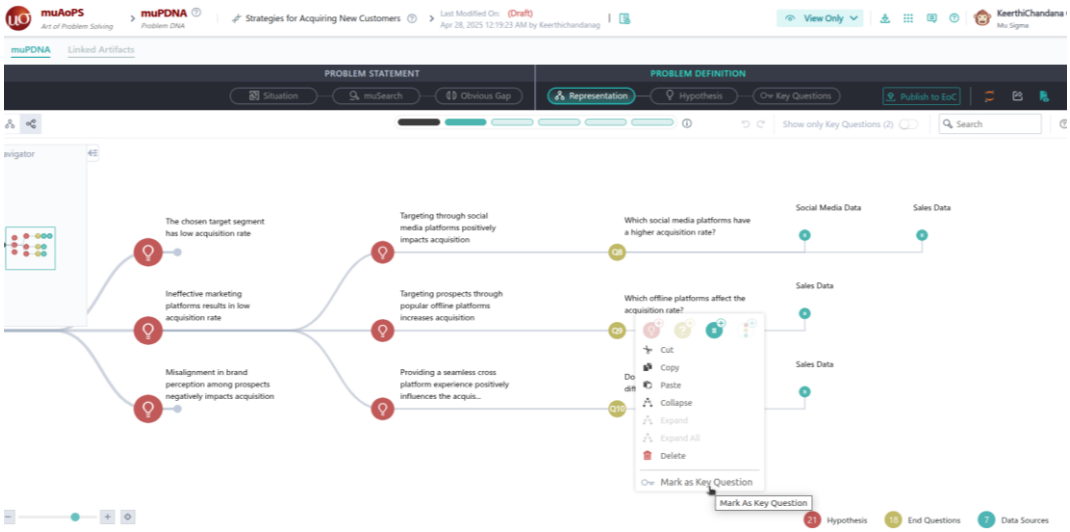
3.Adding Tests for Key Questions in the Hypothesis Section

Once the Key Question is visible in the Hypothesis section with the muTalos button enabled, the next step is to configure Tests for that question.



(Fig 110: Adding Test)

- Click on Add Test below the Key Question
- Select the DIPP category, choosing one of the following:



(Fig 108: Mark as Key Question)

In the Representation screen:

- Go to the list of questions under the problem breakdown.
- Find the question that is important for analysis and validation.
- Right-click on that question to open the menu.
- Select “Mark as Key Question” from the options.

2.Key Question Flow to Hypothesis Section

Once you do this, the question is marked as a Key Question, meaning it becomes critical for testing, finding data, and validating hypotheses.

After marking a question as a Key Question in the Representation screen:

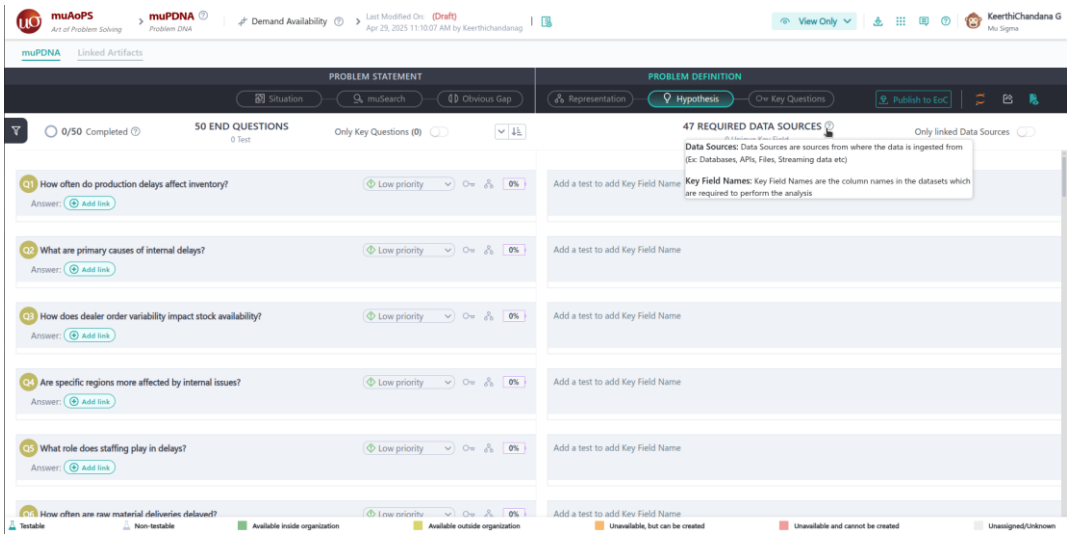
- This Key Question will automatically appear in the Hypothesis section.
- In the Hypothesis section, each Key Question is displayed along with options to configure tests, link data sources, and set up validation logic.

Important Note:

When a question is marked as a Key Question, the system immediately activates the muTalos button beside that question in the Hypothesis screen. This button is essential for launching the automated code execution process after test setup.

The presence of the muTalos button confirms that the question is ready for testing and aligned with the AOPS to Code workflow.

- Option to Add Tests and Key Fields: Users can add test cases or key field names to validate their hypotheses
- Only Linked DataSources: Displays only the datasources that are linked to the test



(Fig 107: Data Sources)

C. List of Hypothesis Questions

Each question is displayed with:

- Priority Indicator: Marks whether the question is of low, medium, or high importance
- Answer Link Field: Provides an option to attach answers
- Progress Indicator: Shows the percentage of work completed for each question
- Assigned User: Displays the person responsible for answering the question
- muTalos: The muTalos button is used for the AoPS to Code process, transforming problem representations into automated tests.
  Below is the detailed process for AoPS to code:

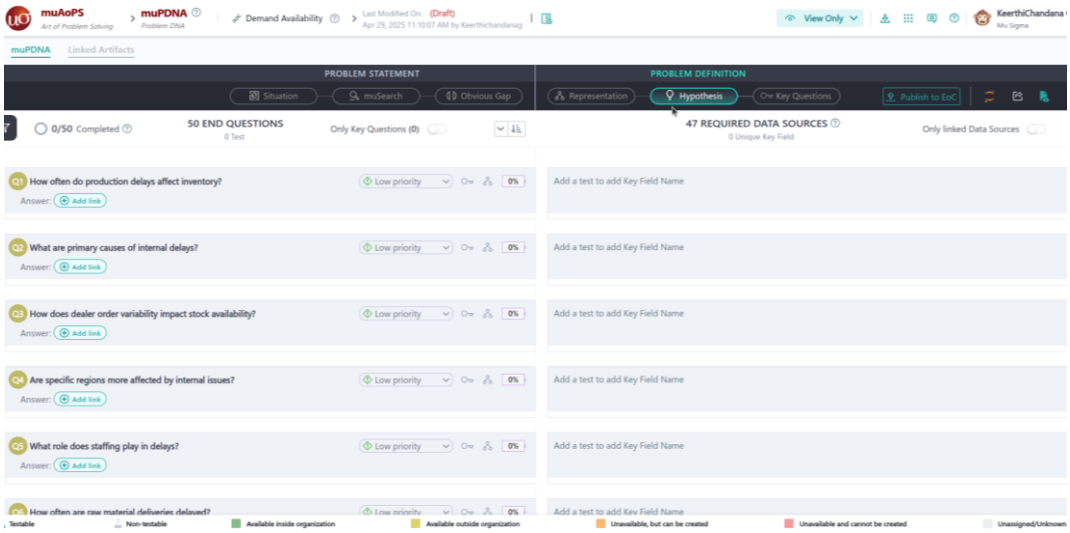
AoPS to Code:

1. Starting from the Representation Screen

The process starts in the Problem Definition phase, under the Representation screen of the muPDNA framework. Here, you focus on breaking the main problem into smaller, measurable sub-questions.

- Available outside organization: Filters hypotheses requiring external data sources
- Unavailable but can be created: Data that does not currently exist but can be generated through surveys, experiments, or other methods
- Unavailable and cannot be created: Data that is impossible to obtain or generate
- Unassigned/Unknown: Data availability is not yet determined

Hypothesis Tab



(Fig 106: Hypothesis)

The main section of the Hypothesis tab displays a list of key questions that help validate different hypotheses.

A. Header Section

- Total End Questions Count: Displays the total number of hypothesis-related questions
- Completion Progress: A tracker showing how many questions have been completed
- Only Key Questions Toggle: Allows users to filter and view only the most important questions

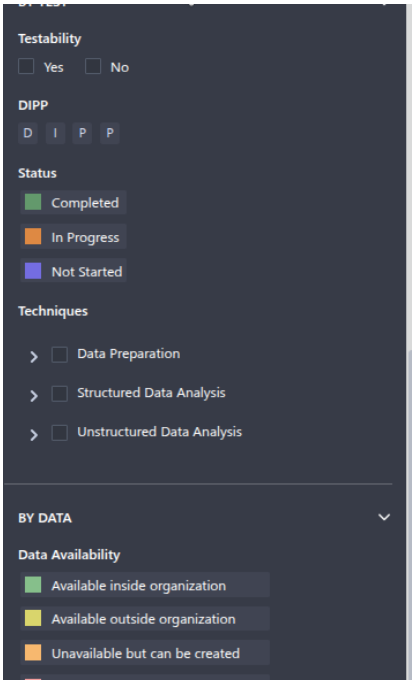
B. Required Data Sources

- Total Data Sources Count: Indicates how many data sources are required for hypothesis validation

3.By Test

This filter is useful for tracking which hypotheses have been tested and which ones require validation.

- Testability (Yes/No): Filters whether a hypothesis can be tested with available data
- DIPP (Descriptive, Inquisitive, Predictive, Prescriptive): Filters tests based on their purpose
- Status (Completed, In Progress, Not Started): Helps track the stage of each hypothesis test
- Techniques Used: Filters based on the type of data analysis used:
  - Data Preparation: Cleaning and organizing data before analysis
  - Structured Data Analysis: Analyzing data in structured formats like tables
  - Unstructured Data Analysis: Processing text, images, or other unstructured data



(Fig 105: By Test, By Data)

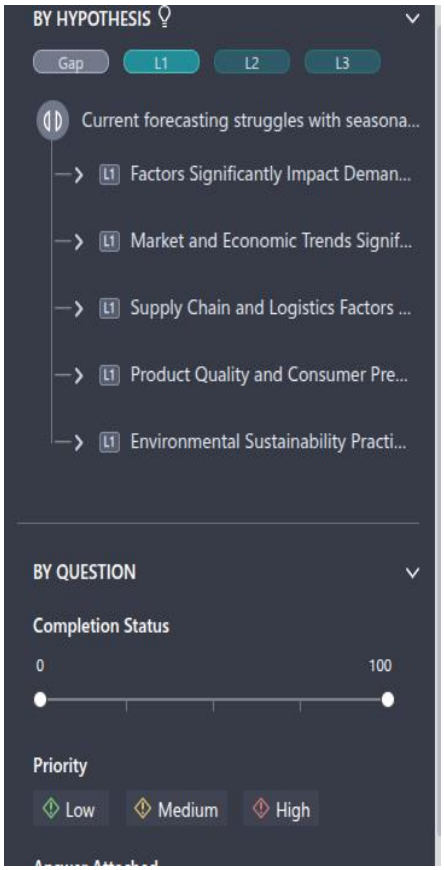
4.By Data

This filter helps users check the availability of required data for hypothesis validation.

- Available inside organization: Filters hypotheses based on data that are available internally

- **L1, L2, L3 Levels**: Represent different depths of hypotheses, with L1 being broad categories and L3 being more specific

Example: Filtering for "Supply Chain and Logistics Factors" will show only the hypotheses related to that aspect of the problem



(Fig 104: By Hypothesis, By Questions)

2.By Questions

This filter helps in focusing on specific key questions related to the problem. Key questions help in understanding the root cause of an issue.

- Completion Status Slider: Allows users to filter questions based on their progress (0-100%)
- Priority Levels: Questions can be categorized as Low, Medium, or High priority
- Answer Attached: Users can filter based on whether an answer has already been attached

Step 1: Understanding the Problem Statement

- Start with a well-defined problem statement that highlights the key challenge
- Example: Forecasting struggles with seasonal and regional demand shifts, causing stockouts and excess inventory

Step 2: Breaking Down into Hypotheses

- Identify key factors influencing the problem
- Each hypothesis represents a potential cause of the issue

Step 3: Adding Key Questions

- Key questions help in validating or rejecting the hypotheses
- These questions should be specific and linked to data sources

Step 4: Assigning Data Sources

- Identify relevant data sources to support the key questions
- Data sources should provide measurable insights

Step 5: Marking Key Questions

- Highlight the most important key questions
- This ensures focus on the most critical aspects of the problem

Step 6: Managing the Structure

- Use options like cut, copy, paste, collapse, and expand to refine the representation
- Ensure a logical flow from the problem statement to hypotheses, key questions, and data sources

3.9 Hypothesis

The Hypothesis tab in muPDNA helps in structuring and validating problem statements by organizing key questions and tracking progress.

Left Panel – Filters

Filters help users narrow down and focus on specific hypotheses, questions, tests, and data sources

1.By Hypothesis

This filter allows users to organize and filter hypotheses based on different levels (L1, L2, L3) and categories. Hypotheses represent assumptions or potential causes of a problem, and filtering by hypothesis helps focus on a particular set of related assumptions

Why is it Important?

- It helps identify the key challenge
- It shows what needs to be fixed
- It helps businesses focus on solutions

Example of an Obvious Gap

If a company struggles with demand forecasting, an obvious gap could be:

"The company cannot accurately predict demand changes, leading to stock shortages or excess stock."

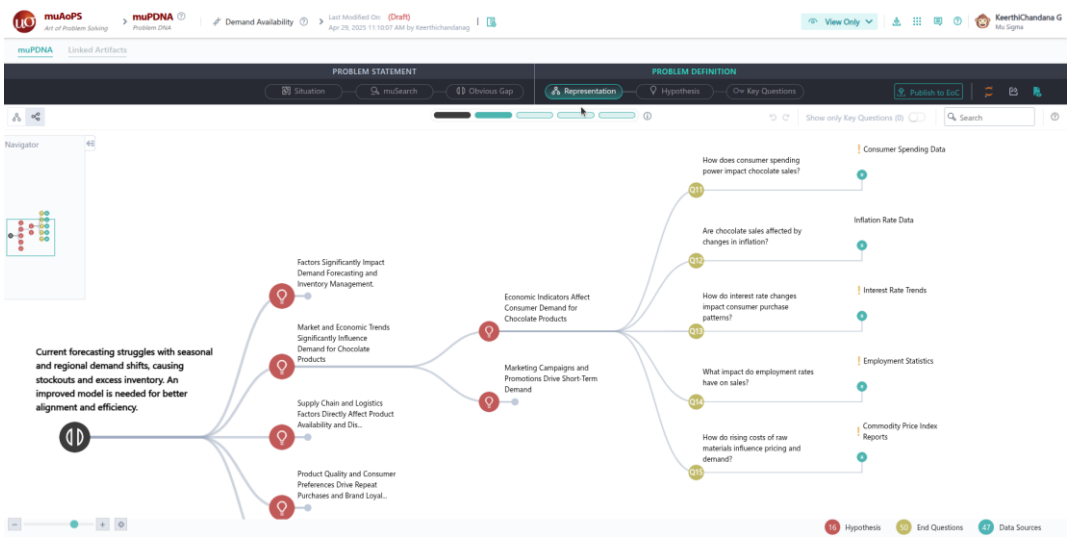
How to Find an Obvious Gap?

- Look at the Current State: What problem exists now?
- Think about the Future State: What is the goal?
- Find the main issue stopping progress

3.8 Representation

The Representation tab helps in structuring the problem by breaking it down into:

- Hypotheses: Possible reasons behind the problem
- Key Questions: Critical questions that need to be answered
- Data Sources: Information required to validate the key questions



(Fig 103: Representation)

complex searches on Google Scholar but also actively finds similar search terms, expanding the scope of exploration and delivering more comprehensive insights. In short, it's not just a responder; it's a dynamic navigator enhancing your search experience

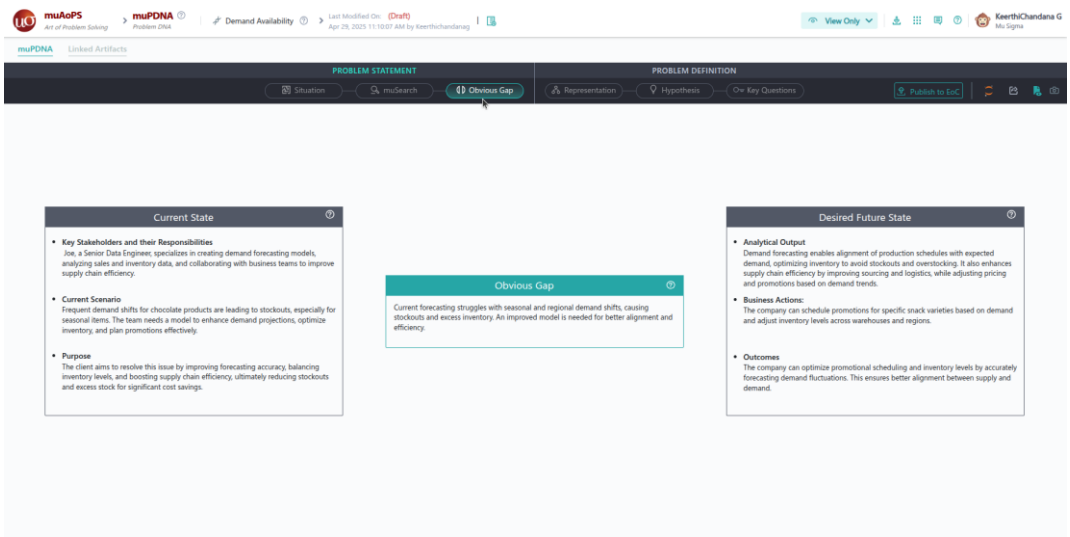
- **Generating Search Term Details:** The final step involves the synthesis of the obtained information. The AI4NS systematically organizes and presents the details related to the search term. The output is a detailed and well-structured overview.

In essence, AI4NS amalgamates the power of semantic search, advanced embeddings, and a sophisticated language model to streamline the research process. By harnessing the capabilities of Google Scholar, it not only fetches information but transforms raw data into meaningful insights, fostering more efficient and informed research for muPDNA.

3.7 Obvious gap

What is an Obvious Gap?

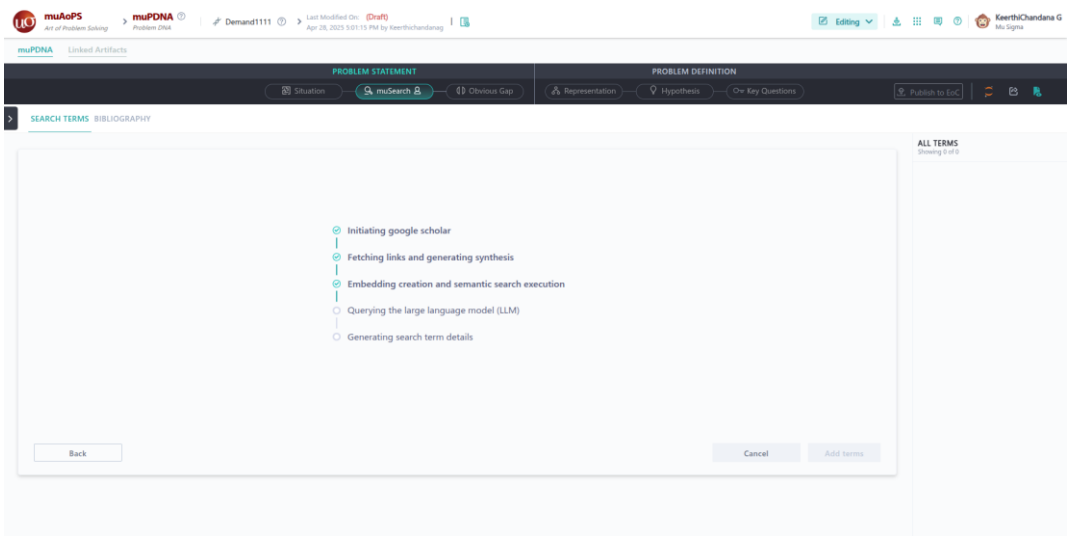
An Obvious Gap is the main problem stopping a stakeholder from moving from the Current State (where they are now) to the Desired Future State (where they want to be).



(Fig 102: Obvious Gap)

- All the generated search terms would be added to the muSearch. Users can see all the generated terms with AI4NS symbol at the right side of the muSearch tab

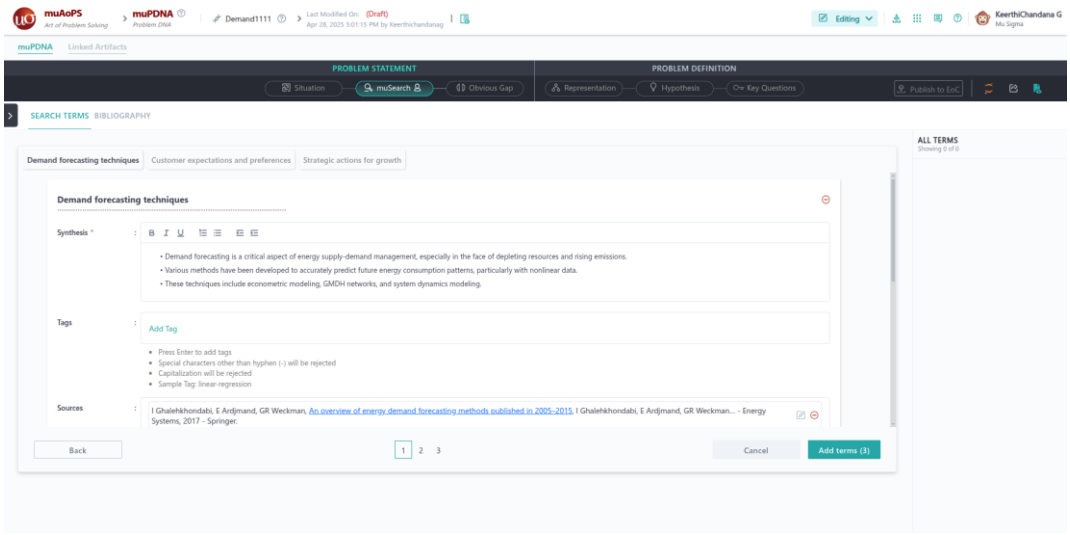
3. Working of AI4NS



(Fig 101: Working of AI4NS)

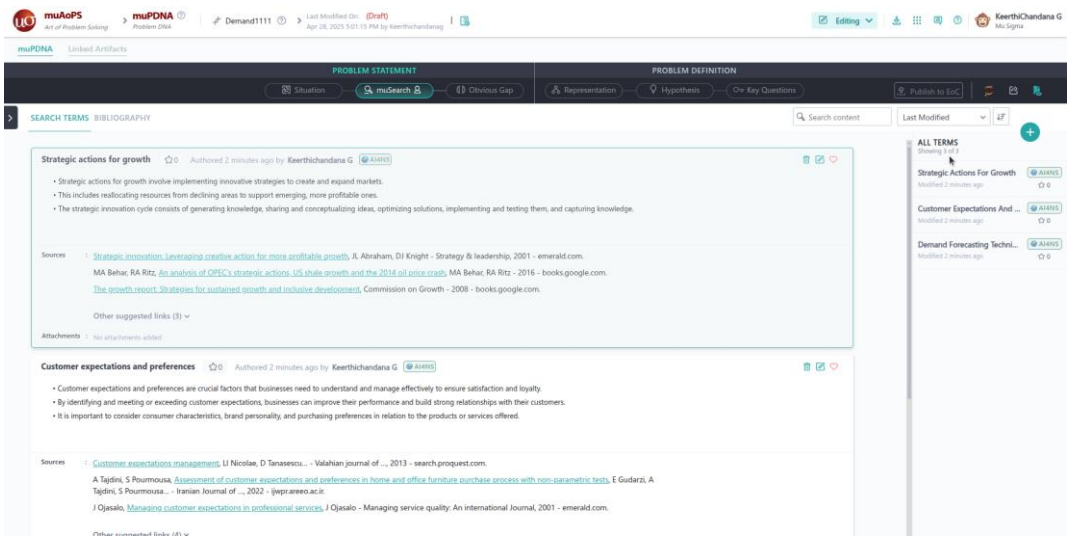
- **Initiating Google Scholar**:The first step involves AI4NS accessing Google Scholar, a repository of scholarly articles, papers, and publications. This serves as the primary source for gathering relevant information
- **Fetching Links and Synthesis**:Upon initiation, the AI diligently fetches links related to the given search term. Simultaneously, it generates a synthesis, providing a succinct overview of the selected search term. This synthesis acts as a quick reference point, offering an initial understanding of the topic selected
- **Embedding Creation and Semantic Search Execution**:One of the key strengths of this AI lies in its ability to create embeddings for the search terms. Embeddings, in this context, represent a mathematical representation of the semantic meaning of the words. These embeddings facilitate a more nuanced and context-aware search, ensuring that the retrieved information is not only relevant but also semantically connected to the user's query
- **Querying the Large Language Model (LLM):** The LLM is the brain behind our system, trained to understand and respond to natural language queries. It not only executes

- Now the AI4NS will generate details for the selected terms. Users can also cancel by clicking on cancel
- It may take some time to generate details while following all the backend steps



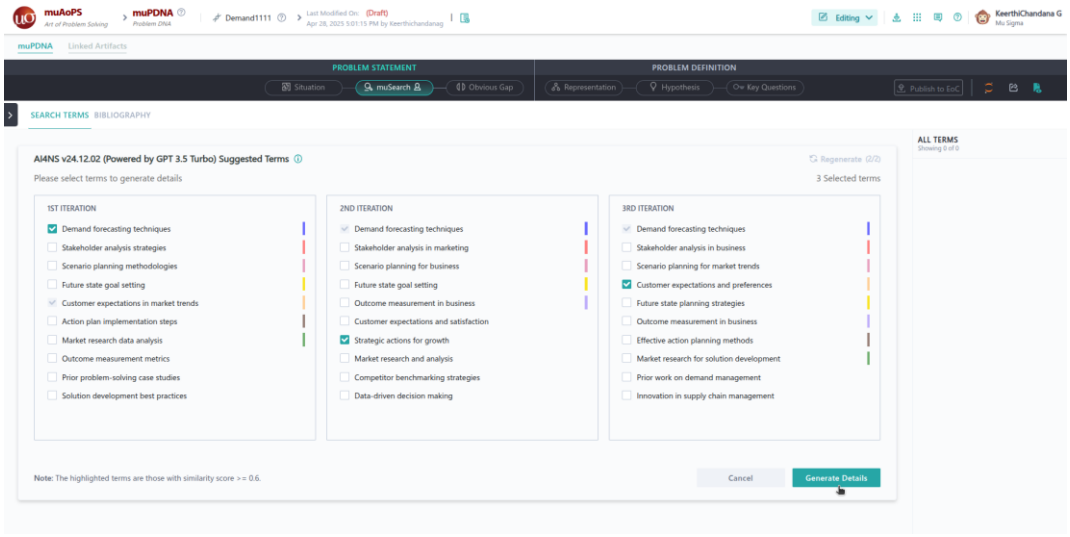
(Fig 99: muSearch Tab for generated details for selected search terms)

- Now users can review and edit the generated details for all the search terms
- Click on ‘add terms ‘after reviewing and editing details



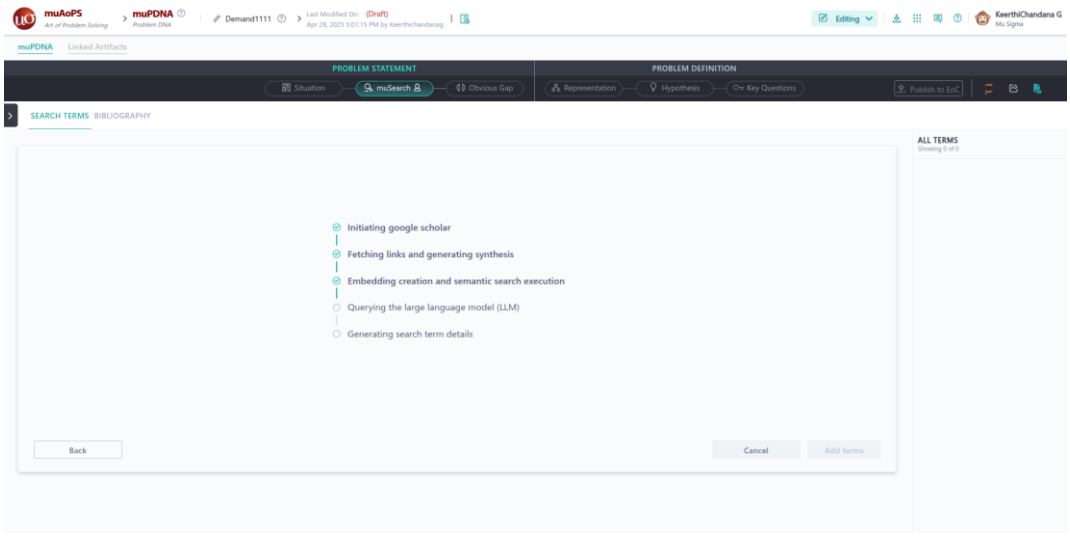
(Fig 100: muSearch Tab showing all search terms updated by AI4NS)

- Now it will generate search terms over third iteration with creativity parameter 0.5 to 0.75

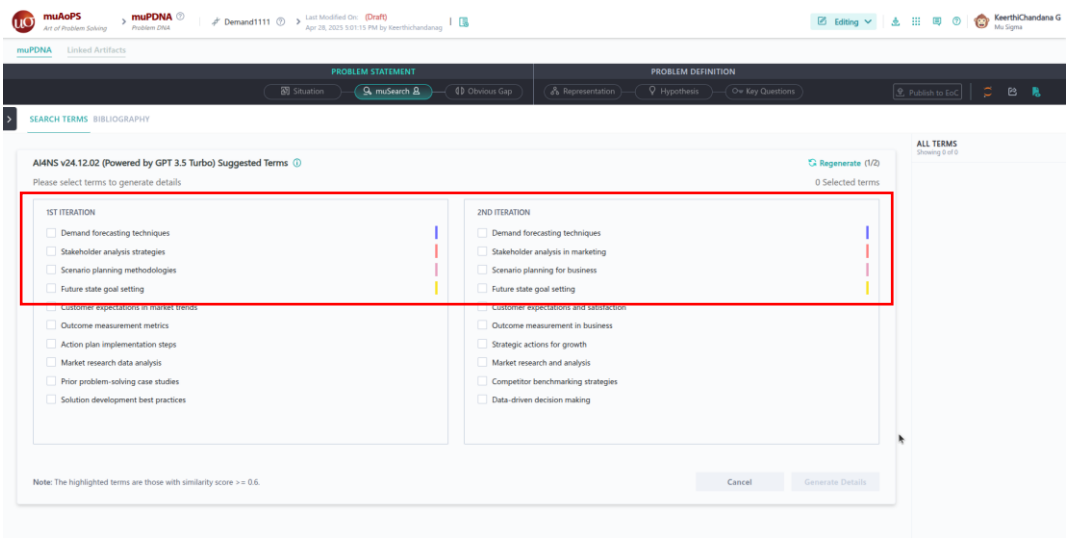


(Fig 97: muSearch Tab for selecting search terms)

- In this iteration, also terms with similarity score greater than or equal to 0.6 will be highlighted for users
- Users can select the search terms that meet their expectations and then click on generating details

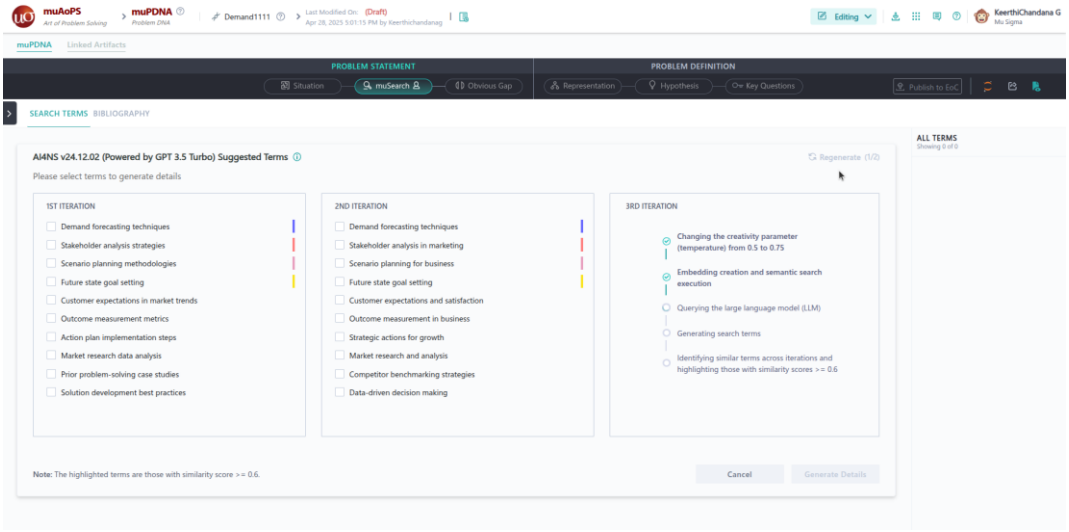


(Fig 98: muSearch Tab for generating details for selected search terms)



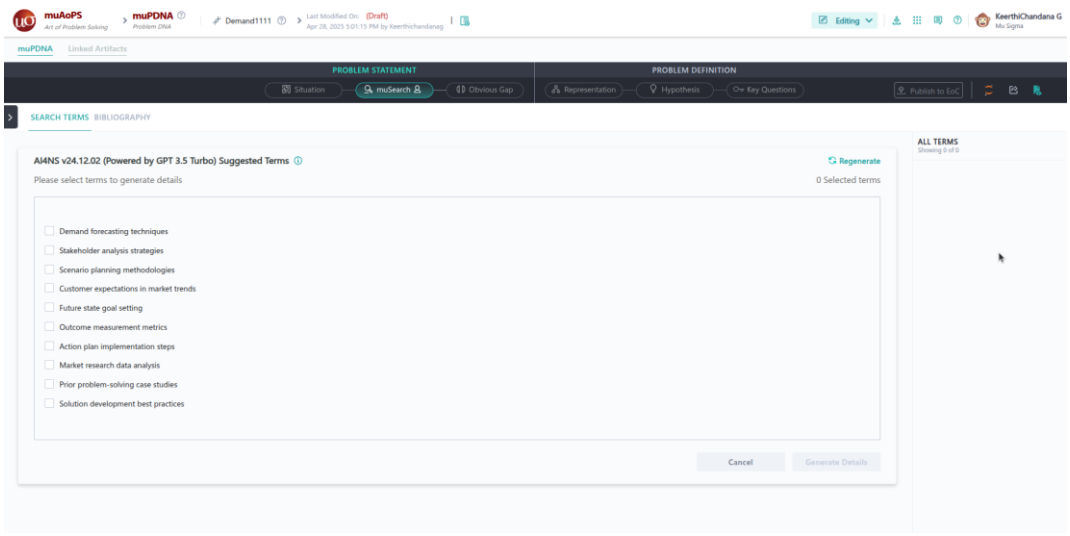
(Fig 95: muSearch Tab for generated search terms over second iteration)

- Terms with similarity score greater than or equal to 0.6 will be highlighted for users aiding them in choosing the most relevant terms
- If the search results still not meet user's expectations again, then can click on "regenerate" to iterate for the third time



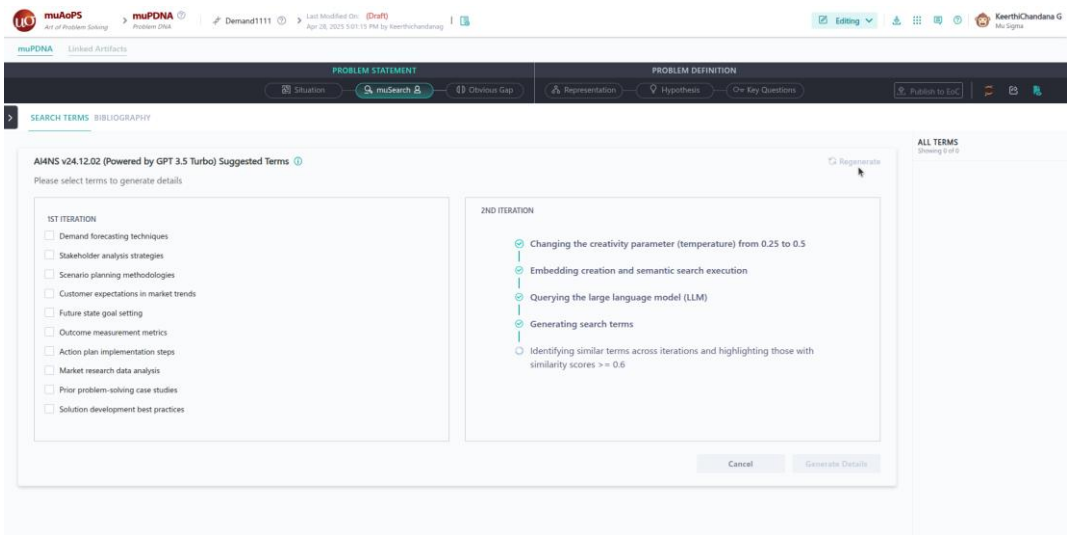
(Fig 96: muSearch Tab for generating search term over third iteration

- Cancel:Stop generating the search terms at any time by clicking on cancel



(Fig 93: muSearch Tab for initially generated search terms)

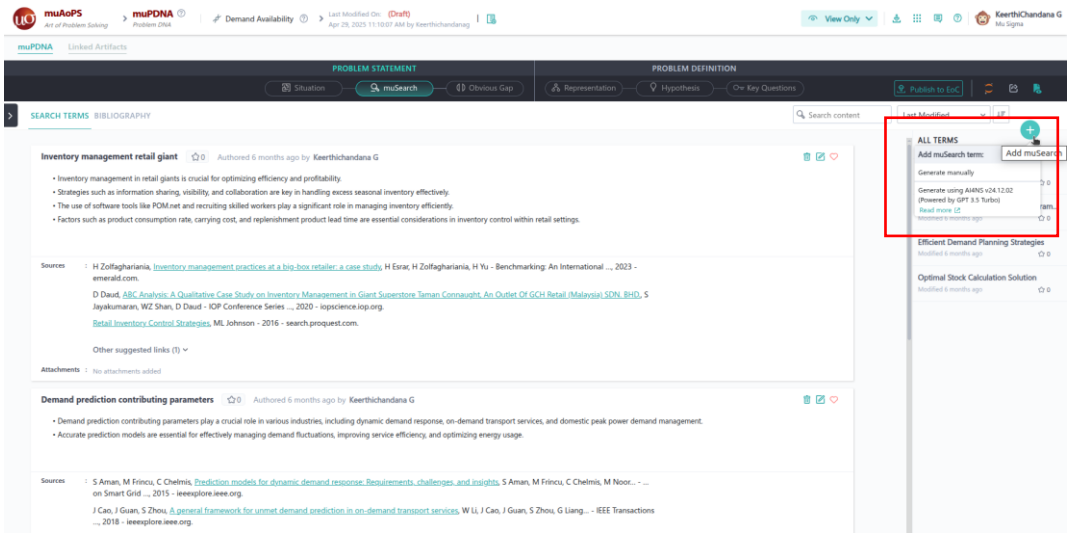
- Regenerate: If the initial search results do not meet user's expectations, they can click on "regenerate" to iterate again, ensuring that the searches precisely align with their specific needs and preferences



(Fig 94: muSearch Tab for generating search term over second iteration)

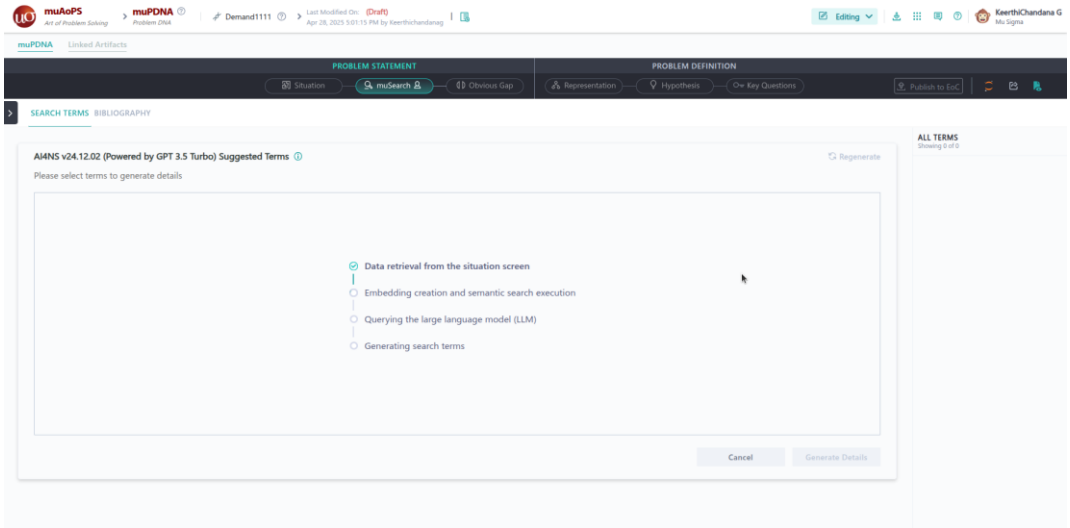
- Now it will generate search terms over second iteration with creativity parameter 0.25 to 0.5

Appending muSearch to an existing muPDNA:



(Fig 91: muSearch Tab for adding to existing muPDNA)

- Generate a muSearch term through AI4NS by clicking on the plus (+) icon at the top right corner of the muSearch page
- Then click on Generate using AI4NS (Artificial Intelligence for Natural Stupidity). AI4NS will start regenerating search terms



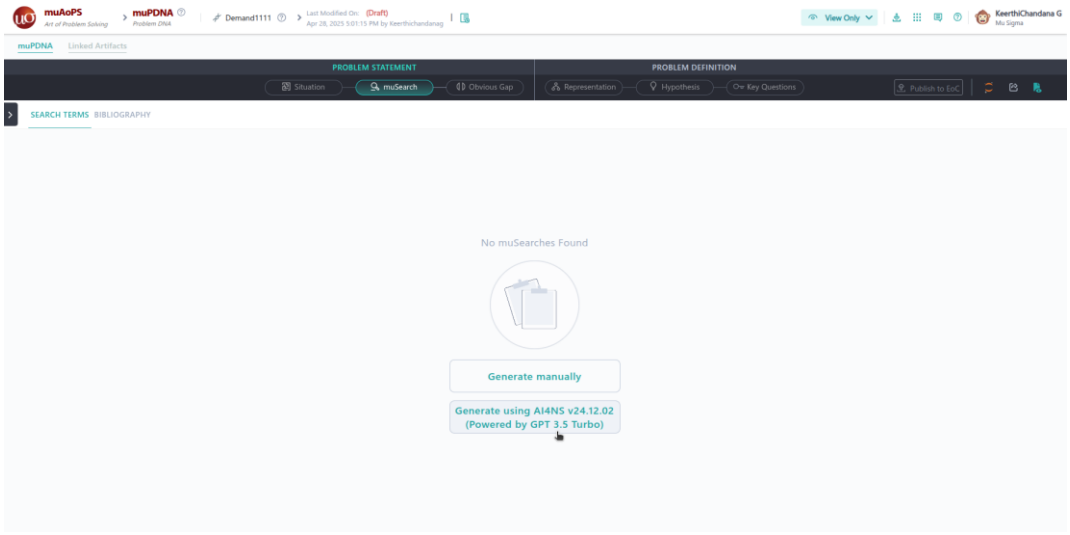
(Fig 92: muSearch Tab for generating search terms initially)

Users can manually input a search term when they already have a specific topic in mind. This is useful for well-defined research areas where users already know what they want to explore

2.Creating muSearch through AI4NS

Enable muSearch for both new and existing muPDNAs effortlessly.

Adding muSearch to a new muPDNA



(Fig 90: muSearch Tab for adding to new muPDNA)

To create a muSearch term using AI4NS, select the “Generate using AI4NS” option on the muSearch page. AI4NS will then automatically begin generating search terms for the new muPDNA.

Main Features:

- List of Sources: Shows books, research papers, and articles related to user topic
- Source Details: Each reference includes the author, title, year, and a link to the original content
- Editing Options:
    - Edit: Modify a reference
    - Delete: Remove an entry
    - Favorite: Save important references
- Other Suggested Links: Get additional related sources for deeper research

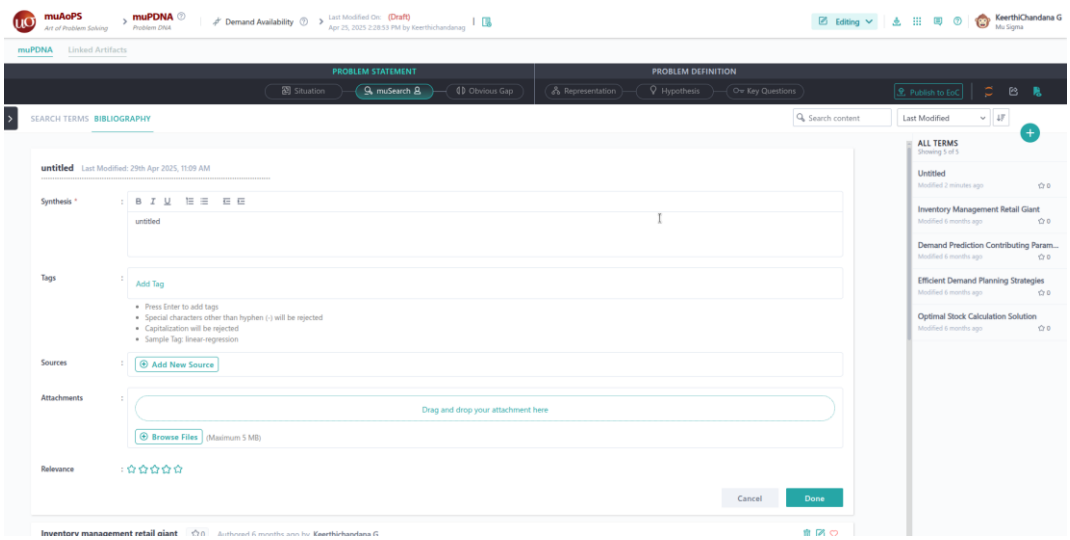
3.Right-Side Filter Options

The right-side panel in muSearch provides filtering and categorization options for search terms and bibliography entries. Users can:

- View All Search Terms: A list of previously created search terms is displayed
- Sort by Last Modified: Organize search terms based on recent updates
- Add muSearch Term: A "+" button in the top-right corner allows users to create a new search term using two methods: Generating Search Terms: Manual vs AI4NS

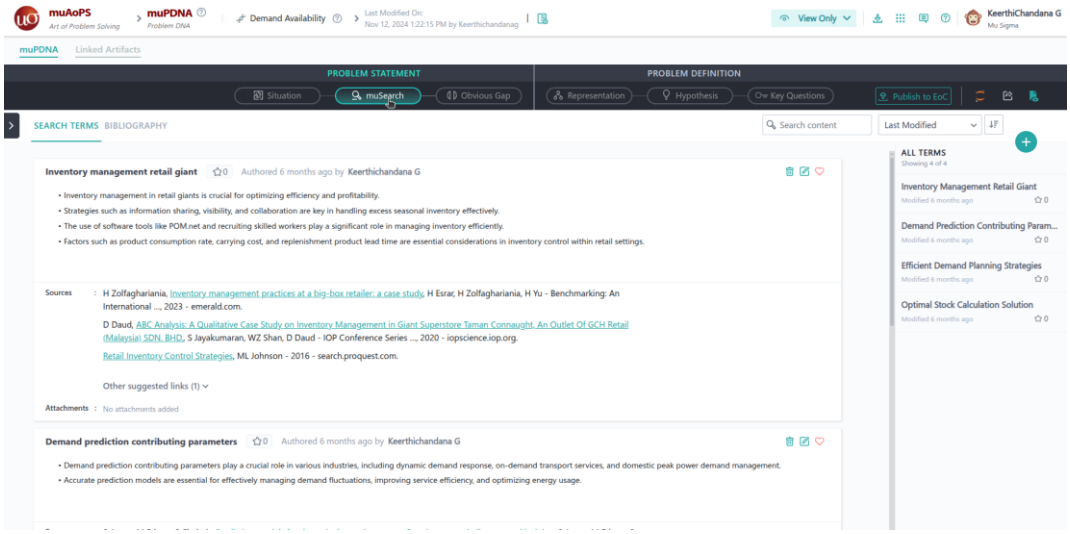
Add muSearch Term:

1.Generate Manually



(Fig 89: Generating muSearch through manually)

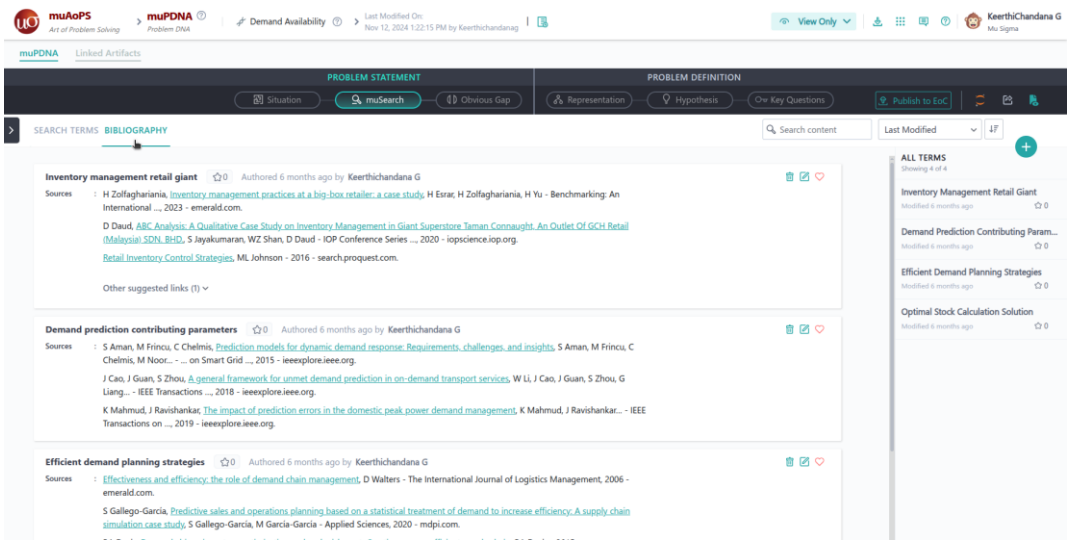
- Actions (Edit, Delete, Favorite): Users can edit, delete, or mark a search term as a favorite using the available icons



(Fig 87: muSearch)

2.Bibliography

The Bibliography section in muSearch helps users to store and organize research references for easy access.



(Fig 88: Bibliography)

3.6 muSearch

muSearch is an advanced research and bibliography management tool designed to assist users in efficiently gathering, organizing, and referencing information related to a business problem. It combines primary and secondary research methods, utilizing internal company resources such as the EOC, the vast internet, and AI-driven techniques like AI4NS (Artificial Intelligence for Natural Stupidity) to generate search terms.

The process of muSearch involves carefully searching for and collecting important ideas, terms, and knowledge that are crucial for solving the targeted business problem. It fosters teamwork and structured exploration to ensure comprehensive understanding and coverage of the problem space.

muSearch is structured into two main sections:

- **Search Terms**: Collect and organize relevant keywords and phrases, either manually or through AI4NS.
- **Bibliography**: Manage and generate bibliographic references for the gathered resources.

There are two ways to add search terms in muSearch:

1. **Generate Manually**: Users independently brainstorm and input relevant search terms.
2. **Generate using AI4NS**: Users leverage AI assistance to automatically generate effective and diverse search terms.

1. Search Terms

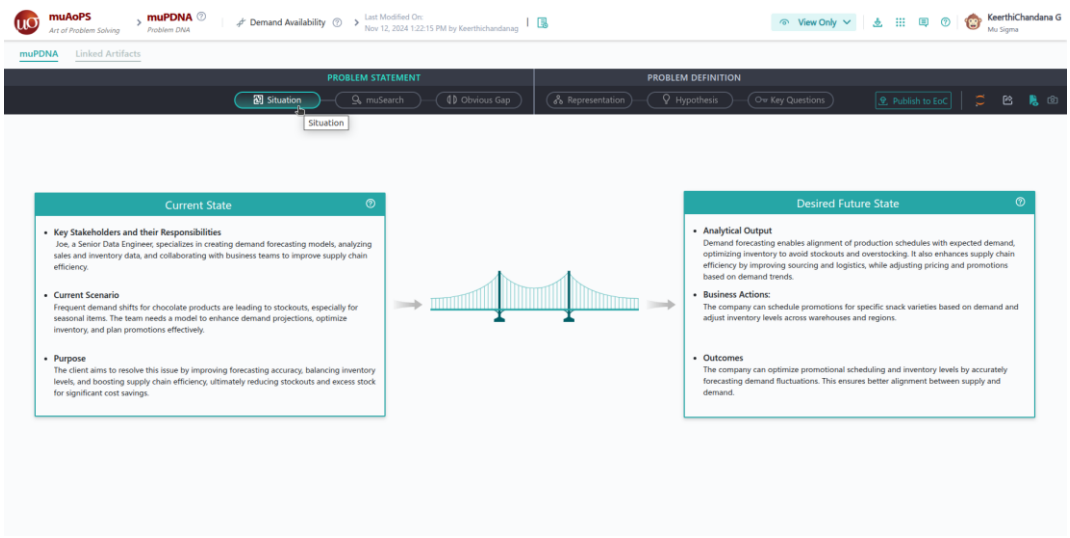
The Search Terms section in muSearch allows users to store, categorize and organize important research topics with related references. Each search term entry includes:

- Title: A short and descriptive name for the topic
- Author & Timestamp: The name of the user who created the entry and when it was last modified
- Key Research Insights: A summary of key points related to the topic
- Sources: A list of reference links, including journal articles, case studies, and relevant documents
- Additional Suggested Links: Extra reading materials suggested by AI, if available
- Attachments: Users can add attachments to support their research

Purpose

- Defines why the problem needs to be solved
- Highlights business goals and objectives
- Shows how solving the problem will add value to the organization

By structuring the Situation effectively, users can ensure that they are addressing the right problem in a well-defined manner before moving to deeper analysis and solution-building.



(Fig 86: Situation)

2.Desired Future State

The Desired Future State represents the ideal outcome after addressing the problem. It ensures that the solution aligns with business goals and is measurable.

Analytical Output:

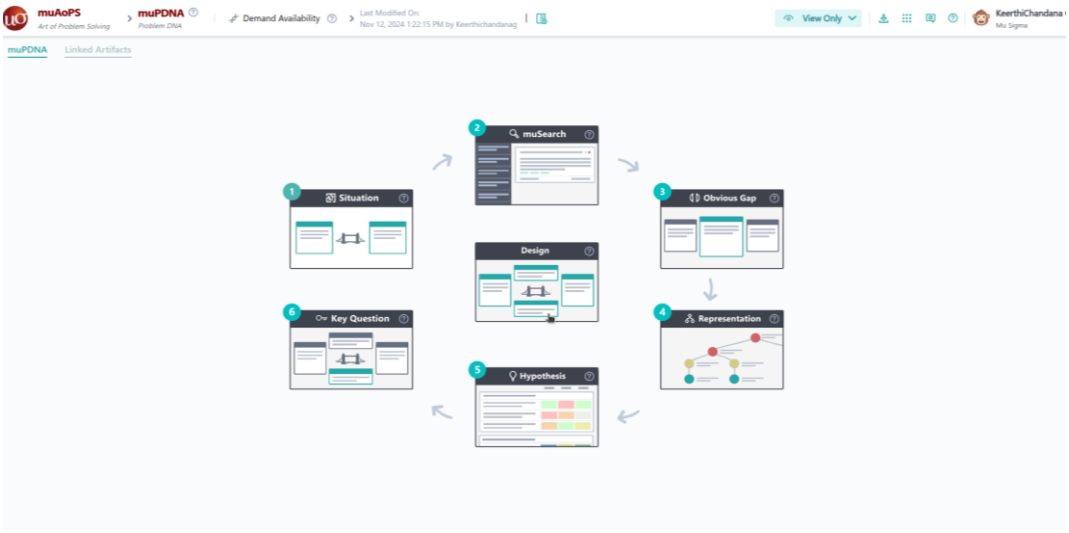
- Defines the specific output expected after solving the problem
- Includes measurable deliverables such as reports, dashboards, insights, or process improvements

Utilization of Output:

- Explains how businesses will use the outputs to make better decisions
- Ensures that the solution is actionable and relevant to the stakeholders

Required Input:

- Identifies the necessary data, resources, and tools needed to achieve the desired state



(Fig 85: Steps in muPDNA)

3.5 Situation

The Situation is the foundation of problem-solving in muPDNA. It helps in establishing a clear understanding of current challenges and setting the right direction for finding solutions. By defining the Current State and the Desired Future State, all stakeholders can align their understanding before diving into problem-solving techniques.

Key Components of a Situation

1.Current State

The Current State helps in understanding the existing problem, the context in which it occurs, and why it is significant. It consists of:

Key Stakeholders and Their Responsibilities

- Identify people or teams involved in the problem
- Defines their roles and responsibilities in addressing the issue
- Help in understanding who will be responsible for implementing solutions

Current Scenario

- Describes the problem the business is facing
- Identifies the trigger that caused the issue
- Explaining the impact of the problem on business operations

Updating muPDNA

- Users can modify a muPDNA by switching from View-Only Mode to Editing Mode using the dropdown on the top right
- Changes can be saved by clicking on the 'Save' option in the app header
- Version control helps in tracking modifications, improvements, and key decisions

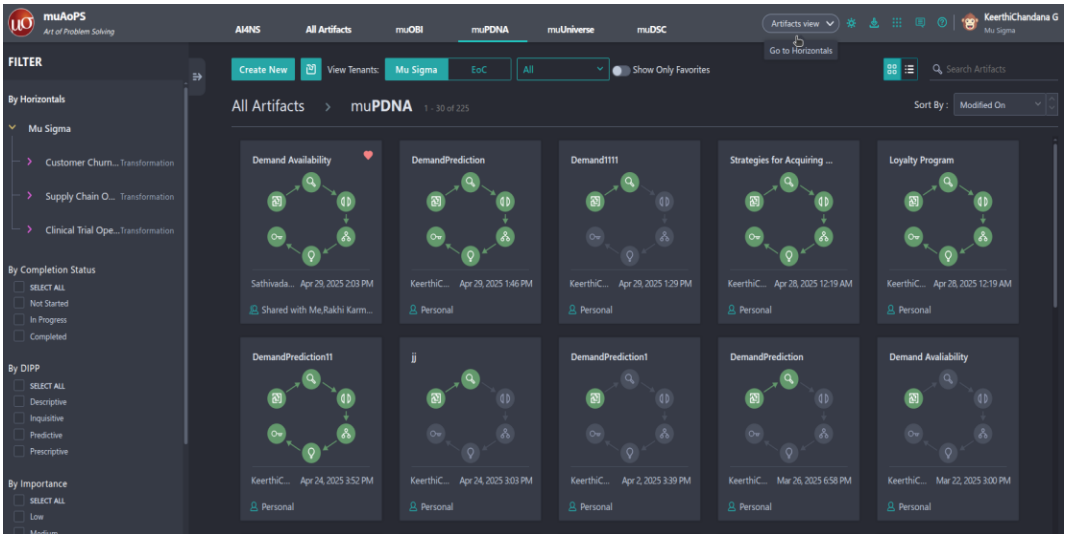
Deleting muPDNA

- Users can manage or remove muPDNAs if necessary, ensuring only relevant ones are maintained
- By effectively using these CRUD operations, users can create, manage and refine their muPDNA for better problem-solving and decision-making

3.4 Explanation of Terminology

- **Situation:** Describe the Current State and Desired Future State
- **muSearch:** Search and understand the context of the problem
- **Obvious Gap:** State the Obvious Gap
- **Representation:** Represent the problem through Hypotheses, End Questions, and Data Sources
- **Hypothesis:** State the tests/data requirements for End Questions and Data Sources
- **Key Questions:** Refine the Gap and define the Key Question based on the current context of the problem
- **Design:** Provides a visual representation of the structured problem, outlining the transition from the Current State to the Desired Future State

Note: *A tile will become active only after content is entered in the preceding tile. Users can access a tile only if it is active, meaning it contains content.*



(Fig 84: muPDNA Artifacts Tab)

3.3 CRUD Operations of muPDNA

CRUD operations (Create, Read, Update, Delete) in muPDNA allow users to efficiently manage Problem DNAs. Here's how users can perform these operations:

Creating a muPDNA

- To create a muPDNA (Problem DNA) in muAoPS, users can follow the standard artifact creation process:
- To create a muPDNA, start by navigating to the 2. Common Utilities section. From there, access 2.2 Toolbar and refer to 4. Creating a New Artifact to initiate the process creating a new muPDNA
- If the user has an existing muPDNA they want to use, 2. Common Utilities section. From there, access 2.2 Toolbar and refer to section 5, 'Importing Artifact,' for easy import

Reading muPDNA

- Users can access different muPDNAs they have created or imported
- Multiple versions of the same muPDNA can be saved and accessed through the dropdown menu, ensuring version control

Three Principles Behind a Good muPDNA Project

1. The process should be iterative: Representation and Hypotheses should drive the best possible design, requiring refinement over time.
2. Top-down structuring leads to better design: Rather than a scattered approach, slowing down thinking deeply about the problem leads to better solutions.
3. The right question should be challenging yet answerable: The problem should be framed in a way that makes solutions insightful and actionable.

3.2 Navigation to muPDNA

To access muPDNA within muAoPS, follow these steps:

- Login to muAoPS: Open muAoPS and enter the Credentials
- Go to the Homepage: After logging in, the homepage will be displayed

(Fig 83: Navigation to muPDNA)

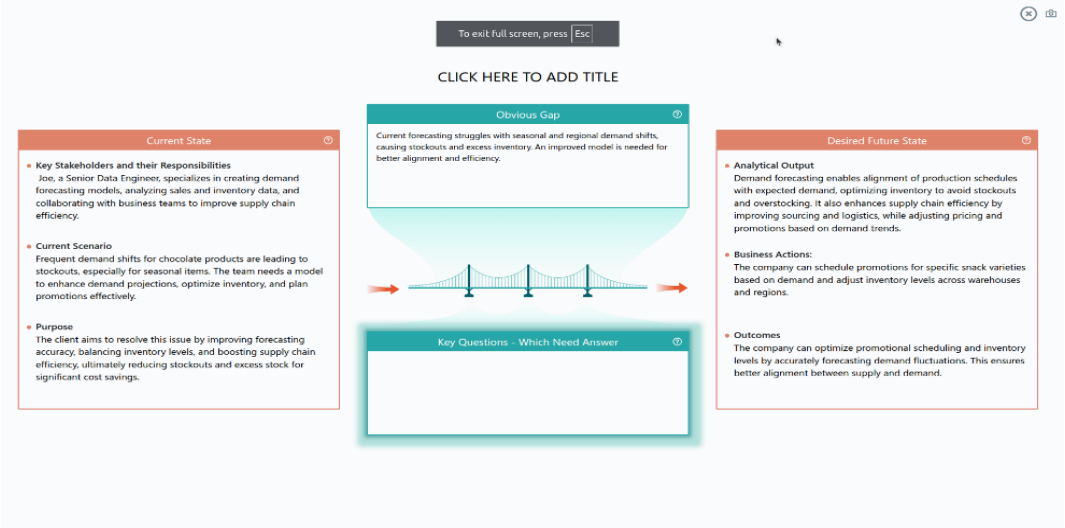
- Click on the muPDNA Icon: click the muPDNA icon to access the muPDNA section This will take the user to the 'muPDNA' tab where all the artifacts available to the user are listed

3.1.2 Why do we need muPDNA

The need for muPDNA arises from the fact that many businesses struggle with defining problems clearly, leading to ineffective solutions. Without a structured approach, companies may jump to conclusions, misinterpret data, or focus on irrelevant issues.

Benefits of muPDNA:

- **Ensures Clarity in Problem Definition:** Helps teams properly define and structure business problems before jumping into solutions
- **Promotes Structured Thinking:** Encourages a top-down approach, ensuring a logical flow from problem identification to solution
- **Avoid Premature Solutions:** Focuses on understanding the root cause rather than providing quick fixes
- **Enhances Data-Driven Decision Making:** Ensures that hypotheses are tested using appropriate data, leading to reliable and efficient solutions
- **Iterative and Adaptive:** The methodology allows for continuous refinement, making it flexible and scalable for different business scenarios



(Fig 82: muPDNA)

3.muPDNA

3.1 Introduction to muPDNA

muPDNA (Mu Sigma's Problem DNA) is a structured problem-solving framework designed to enable organizations to define and deconstruct complex business challenges. It ensures a systematic approach to problem definition, representation, and hypothesis formation, leading to more effective decision-making.

Much like DNA serves as the fundamental blueprint of life, muPDNA acts as the foundational structure for solving business problems with precision and clarity. It emphasizes structured thinking, hypothesis-driven analysis, and iterative problem refinement, making sure that businesses arrive at data-backed, well-designed solutions rather than jumping to premature conclusions.

3.1.1 muPDNA what and how?

muPDNA provides a step-by-step methodology to define business problems clearly and systematically. It structures problems in a way that promotes the best possible design by focusing on asking the right questions.

It consists of three key strands:

1.Design

- Understand the Current State and Desired Future State
- Identify gaps in business processes based on extensive research
- Define a key contextual question that is challenging yet solvable

2.Representation

- Identify key hypotheses that represent the problem.
- Use a top-down approach to break the problem into structured components
- Avoid premature solutions and focus on understanding the problem first

3.Hypothesis

- Formulate hypotheses to test possible solutions to the problem
- Design structured tests to answer key business questions
- Identify and analyze the data required to execute these tests

By iterating through these steps, muPDNA ensures a well-framed problem statement that leads to the most effective solution.