I. Title Slide (Slide 1)

- **Text:** Designing and Development of a Large Language Model for Scholarly Articles at Foster Wheeler Library, Batangas State University - Alangilan
    - [Your Name(s)]
    - [Department/Faculty]
    - Batangas State University
    - [Date]

II. Introduction/Background (Slides 2-4)

- **Slide 2: The Evolving Role of Libraries:**
    - **Text:** Libraries are in a period of dynamic transformation, moving beyond their traditional role as physical repositories of books and journals. The digital revolution has fundamentally altered how individuals access and process information. While libraries continue to be vital centers of learning, the sheer volume and accessibility of digital content, particularly scholarly articles, present new challenges. Users often struggle to efficiently navigate this information overload and find resources relevant to their specific research needs. This project recognizes this evolving role of libraries and aims to contribute to their ongoing adaptation by designing and developing an AI-powered solution to enhance access to scholarly articles.
- **Slide 3: Problem Statement/Research Gap:**
    - **Text: At Batangas State University - Alangilan's Foster Wheeler Library, students and researchers encounter several key obstacles that hinder their research progress. Firstly, access to advanced research tools and digital resources is often limited, making it difficult for students to efficiently explore and analyze the vast body of scholarly literature. Secondly, the process of synthesizing information from numerous academic articles is time-consuming and can be cognitively demanding.** Students frequently struggle to identify the core arguments, key findings, and connections between different studies. **Thirdly, summarizing research papers, a crucial step in the research process, is a laborious undertaking, requiring users to read and digest entire documents to extract the most pertinent information. Finally, the library's current digital infrastructure may not fully support the demands of modern research, potentially lacking the collaborative tools and seamless access needed for effective scholarship.** This project aims to address these challenges by designing and developing an AI-powered solution specifically tailored to the needs of the Foster Wheeler Library and its users.
- **Slide 4: The Promise of AI and LLMs:**
    - **Text: Artificial intelligence, and more specifically, Large Language Models (LLMs), offer a promising avenue for addressing the challenges outlined. LLMs are sophisticated AI models trained on massive datasets of text and code. These models possess the remarkable ability to understand and generate human-like text,**

making them exceptionally well-suited for a wide range of natural language processing (NLP) tasks. Integrating LLMs into library resources and research workflows has the potential to revolutionize how students and researchers interact with scholarly literature. By leveraging the power of LLMs, we can design systems that dramatically improve information access, enhance research efficiency, and even create personalized learning experiences tailored to individual student needs. This project focuses on the design and development of an LLM-based system explicitly intended to address the information access and research support needs of the Batangas State University - Alangilan library community.

III. Research Questions and Objectives (Slides 5-6)

- **Slide 5: Research Questions:**

The university library is one of the most crucial places in a university that becomes a go-to resource for many students seeking various types of information. Most students struggle to locate and access the research papers required for their studies. They face difficulties with tasks like scanning extensive research papers and identifying pertinent information quickly. With the emergence of advancements in artificial intelligence, especially the integration of the Large Language Model, there is significant potential to help resolve these challenges. Recognizing these challenges, this study aims to develop and implement an AI Large Language model for research paper writing. The following research questions will guide the study:

1. How effective can the proposed lLarge Language Model achieve high precision in extracting and summarizing the research papers especially when dealing with large volumes of research papers?
2. Can the accuracy and functionality of the proposed Large Language Model be assessed in comparison to the currently available GPT model?
3. To what extent can the implementation of the Large Language Model into the library's digital resources enhance students' research paper writing capabilities in accessing and utilizing academic research?

- **Slide 6: Research Objectives:**

IV. Literature Review (Slides 7-8)

The main objective of this study is to develop and implement an AI Large Language Model for scholarly articles using advanced technologies such as machine learning algorithms, natural language processing, and software development. Specially , It aims to:

1. To gather data set consist of:

a. Academic papers from programs from BSCpE, BSIT and BSCS at Batangas State University-Alangilan Campus from 2020 to 2024.

2. To identify and use the most suitable LLM model architecture to train, validate, and test the dataset in terms of:

a. Increasing data sizes and complexity tasks.

b. Performance for the precision and speed in text extraction and summarization task.

c. Ability to fine-tune for better contextual understanding and output relevance.

3. To develop a website that provides the following:

a. User- friendly interface for students to easy use with features like answering questions, giving personalized feedback and offering help in real-time

b. Summarize of research papers, highlighting important points for user's question

- **Slide 7: Related Work:**
    - **Text: The use of AI in education has been extensively researched, with numerous studies demonstrating its potential to personalize learning experiences and improve student outcomes.** For example, [cite a relevant study, providing a brief summary of its findings]. Furthermore, **the application of Large Language Models in library and information science contexts has shown significant promise in areas such as information retrieval, question answering, and text summarization.** [Cite a relevant study on LLMs in libraries and briefly describe its key contributions]. **This project builds upon this existing body of research by focusing specifically on the design and development of an LLM-based system tailored to the unique needs of academic research within a university library setting. It aims to address the specific gap of providing readily accessible and intelligent assistance for students navigating the complexities of scholarly literature.**
- **Slide 8: Theoretical Framework (If Applicable):**
    - **Text: This project is grounded in the principles of Information Retrieval (IR) models. IR models provide a theoretical framework for understanding how information is organized, indexed, and retrieved in response to user queries. In this context, the LLM is conceptualized as an advanced and intelligent IR system. It goes beyond traditional keyword-based search by leveraging its natural language processing capabilities to understand the semantic meaning of user queries and retrieve relevant information from a large corpus of scholarly articles.** [Specifically mention the IR model (e.g., Vector Space Model, BM25) and explain how it's relevant to your LLM design, for example: "The Vector Space Model is particularly relevant as it allows for representing both user queries and scholarly articles as vectors in a high-dimensional space, enabling the LLM to calculate semantic similarity and retrieve the most relevant articles based on contextual understanding."]. This theoretical foundation provides a robust framework for designing and evaluating the effectiveness of the LLM as a research assistant.

V. Methodology (Slides 9-11)

- **Slide 9: Research Design:**

- **Text: This project adopts a design and development research methodology, which is appropriate for creating and evaluating a functional system. The Agile software development methodology was selected for its iterative and incremental approach, providing the flexibility needed to adapt to evolving requirements and user feedback throughout the development lifecycle. The Agile process emphasizes close collaboration with stakeholders, frequent feedback loops, and continuous testing, all of which are critical for ensuring that the final system effectively addresses the needs of the users.**

- **Slide 10: Data Collection and Treatment:**
  - **Text: The primary data source for this project consists of scholarly articles from the BSCpE, BSIT, and BSCS programs at Batangas State University - Alangilan, published between 2020 and 2024.** A mixed approach was employed for data collection. **Web scraping techniques** were used to gather articles that were already available in digital format from online repositories and academic databases. [Mention specific web scraping tools/libraries used, e.g., "Python's Beautiful Soup and Scrapy libraries were utilized for this purpose."]. For printed copies of theses and dissertations, **Optical Character Recognition (OCR) technology** was employed to convert the physical documents into digital text. [Mention specific OCR software/libraries used, e.g., "Tesseract OCR was used for this conversion."]. **Once the data was collected, it underwent a rigorous cleaning and preprocessing phase.** This involved removing irrelevant information (e.g., headers, footers, tables of contents), correcting errors introduced by OCR, and tokenizing the text. **Tokenization** involved breaking down the text into individual words or sub-word units, which are the fundamental units of input for the LLM. [Mention specific NLP libraries used for preprocessing, e.g., "**spaCy and NLTK were used for text cleaning and tokenization."].**

- **Slide 11: System Architecture and Design:**
  - **Text:** The proposed system architecture is designed to be modular and scalable, comprising several key components working in concert. At **the core of the system is the GPT-3 Large Language Model, accessed via the OpenAI API. GPT-3** was chosen for its state-of-the-art natural language processing capabilities, its ability to generate human-quality text, and its proven performance in various NLP tasks, including text summarization and question answering. **The OpenAI API acts as the bridge between our system and the powerful GPT-3 model, allowing us to send requests and receive responses.** Three distinct design proposals were initially considered, each offering a different balance of complexity, functionality, and resource requirements. **Design 1 represented a baseline implementation using a less computationally intensive LLM and a basic web interface. Design 2 enhanced the LLM capabilities and incorporated a more interactive user interface with real-time feedback mechanisms. Design 3, the ultimately selected design, leverages the full power of GPT-3, integrates AI fairness tools to mitigate bias, and employs a robust web framework for enhanced performance, scalability, and security. Design 3 was chosen due to its superior performance in NLP tasks, its enhanced security features, its long-term sustainability potential, and its adherence to ethical**

considerations in AI development. [Include a clear and well-labeled diagram of the system architecture, showing the interactions between the user, the web application, the OpenAI API, the GPT-3 model, the database of scholarly articles, and any other relevant components.

VI. Implementation and Development (Slides 12-14)

- **Slide 12: Software and Technology:**
    - **Text:** This project leverages a range of software and technologies carefully selected for their specific strengths and suitability for the design and development goals. **Python** served as the primary programming language due to its versatility, extensive libraries for data science and machine learning, and strong community support. **Pandas,** a powerful data manipulation library, was instrumental in cleaning, transforming, and analyzing the large dataset of scholarly articles. **spaCy,** a library specializing in advanced natural language processing, was employed for tasks such as tokenization, part-of-speech tagging, named entity recognition, and dependency parsing, enabling a deeper understanding of the text's structure and meaning**. Hugging Face Transformers** provided access to pre-trained LLMs, including GPT-3, and offered tools for fine-tuning these models on our specific dataset. **The OpenAI API** served as the interface to the GPT-3 model, allowing our system to utilize its advanced language processing capabilities. **PyTorch,** a widely used deep learning framework, was chosen for model training and optimization, offering flexibility and efficient computation, particularly on GPUs. For the **web application, Django**, a high-level Python web framework, provided a robust and scalable backend for handling user requests, managing data, and securing the application. **Vue.js, a progressive JavaScript framework, was used to develop the frontend, creating a dynamic, responsive, and user-friendly interface. Vuetify, a UI component library for Vue.js, enhanced the visual appeal and usability of the interface, providing pre-built components and styling. Finally, AI Fairness 360, a comprehensive toolkit for examining and mitigating bias in AI systems, was integrated into the development process to ensure ethical compliance and fairness in the LLM's outputs, promoting responsible AI development.**
- **Slide 13: Software Design and Procedure:**
    - **Text:** The software development process adhered to the Agile methodology, which emphasizes iterative development, continuous feedback, and close collaboration with stakeholders. This approach was chosen for its flexibility and adaptability, allowing the project to evolve based on user feedback and changing requirements. The development process was divided into several key phases:
        - **Requirements Identification:** This initial phase involved gathering requirements from potential users, including students, faculty, and library staff, through surveys, interviews, and focus group discussions. The goal was to thoroughly understand their needs, pain points, and expectations for the system.
        - **Planning:** Based on the gathered requirements, detailed plans were developed for each iteration (or sprint) of the development cycle. These plans outlined the

specific features to be implemented, the tasks involved, and the timelines for completion.

- **Design:** The system architecture, database schema, user interface design, and API specifications were defined in this phase. Careful consideration was given to user experience, system performance, security, and scalability.
- **Implementation and Integration:** This core phase involved the actual coding and development of the system. The LLM was fine-tuned on the prepared dataset, and the web application was built to connect to the LLM via the OpenAI API. Integration with the library's existing systems was also considered.
- **Testing:** Rigorous testing was conducted throughout the development process, including unit testing, integration testing, and user acceptance testing. The aim was to identify and fix bugs, ensure the system's functionality, and validate that it met the user requirements.
- **Release:** Once the system was deemed stable and functional, it was released to a select group of users for initial feedback and testing in a real-world environment.
- **Evaluation and Monitoring:** Post-release, the system's performance was continuously monitored, and user feedback was actively collected. This information was used to identify areas for improvement and inform future development iterations. [Include the flowchart of the AI GPT Model System here, explaining the steps involved in processing a user query, from initial input and preprocessing, through query understanding and model invocation, to answer generation, post-processing, response delivery, feedback collection, continuous learning, bias evaluation, and model updates.]

- **Slide 14: UI Design:**
  - **Text:** The user interface (UI) was designed with a strong emphasis on user-friendliness, accessibility, and intuitiveness. The goal was to create an interface that would be easy for students and researchers to learn and use, regardless of their technical expertise. Key design principles included:
    - **Simplicity:** The interface was kept clean and uncluttered, avoiding unnecessary complexity.
    - **Intuitive Navigation:** Clear and logical navigation was implemented to allow users to easily find the information and features they need.
    - **Accessibility:** The UI was designed to be accessible to users with disabilities, adhering to accessibility guidelines and best practices.
    - **Responsiveness:** The interface was designed to be responsive across different devices, including desktops, laptops, tablets, and smartphones. Key features of the designed UI include:
    - **Question Answering:** A prominent text input area allows users to type in their research questions in natural language.

        - **Summarization:** A feature is provided to allow users to easily request summaries of research papers.
        - **Personalized Feedback:** The system provides personalized feedback and recommendations based on user interactions and preferences. This might include suggesting related articles, relevant keywords, or alternative search strategies.
        - **Real-time Assistance:** Integrated help and support features offer real-time guidance to users as they interact with the system. [Include high-quality screenshots of the designed user interface, highlighting the key features and design elements. Explain the rationale behind the design choices and how they contribute to a positive user experience.]

VII. Results and Evaluation (Slides 15-17)

- **Slide 15: Expected Output and Features:**
    - **Text:** The designed AI Large Language Model (LLM) is expected to provide several key features that will significantly enhance the research experience for students and faculty at Batangas State University - Alangilan. These features are designed to directly address the challenges identified in the problem statement and to support users at each stage of their research process. The core features include:
        - **Intelligent Text Summarization:** The LLM will be capable of condensing lengthy research papers into concise and informative summaries, extracting the most important points, key findings, and methodological approaches. This will allow users to quickly grasp the essence of a paper without having to read the entire document, saving them valuable time and effort.
        - **Context-Aware Question Answering:** Users will be able to ask questions related to the content of research papers in natural language, and the LLM will provide relevant and accurate answers based on its understanding of the text's context. This interactive question-answering feature will facilitate deeper comprehension of complex research topics.
        - **Personalized Feedback and Recommendations:** Based on user interactions and search queries, the LLM will provide personalized feedback and recommendations. This could include suggesting related research papers, identifying relevant keywords, or offering alternative search strategies. This personalized guidance will help users explore their research topic more effectively.
        - **Real-time Research Assistance:** The system will provide real-time assistance and support to users as they navigate the research process. This could include explanations of complex concepts, guidance on using the system's features, and troubleshooting assistance. This feature will ensure that users can effectively utilize the LLM and overcome any challenges they may encounter.
- **Slide 16: Design Constraints and Trade-offs:**

- **Text:** The design and development of the LLM system were subject to several constraints and trade-offs, which were carefully considered during the design process. These constraints included:
    - **Technical Constraints:** The computational resources available, the complexity of training and fine-tuning large language models, and the limitations of current AI technology posed certain technical challenges. For instance, the availability of powerful GPUs for training was a key consideration.
    - **Sustainability Constraints:** The long-term sustainability of the system, including the costs of maintenance, updates, and infrastructure, was a critical factor. The design aimed to minimize these costs while ensuring the system's continued functionality and effectiveness.
    - **Security Constraints:** Protecting the security of the system and the privacy of user data was a paramount concern. Robust security measures were incorporated into the design to prevent unauthorized access, data breaches, and other security vulnerabilities.
    - * **Time Constraints:** The project had to be completed within a defined timeframe. This required careful prioritization of features and efficient use of development resources. Trade-offs had to be made between the desire for a highly comprehensive system and the practical limitations of the project timeline.
    - * **Ethical Constraints:** Ethical considerations related to AI development, such as data privacy, bias mitigation, transparency, and accountability, were carefully addressed. The design incorporated mechanisms to minimize bias in the LLM's outputs and ensure responsible use of AI technology.

The selection of Design 3 represented a strategic trade-off. While it required a longer development timeline compared to Designs 1 and 2, it offered significant advantages in terms of performance, security, scalability, and ethical compliance. The decision was made to prioritize these long-term benefits over a faster, but potentially less robust and less ethically sound, implementation. Design 3's use of GPT-3, its integration of AI fairness tools, and its robust web framework were deemed essential for creating a truly effective and responsible research assistant. The longer development time was considered a worthwhile investment in the long-term value and sustainability of the system for the university library.

- **Slide 17: Evaluation Metrics (Proposed):**
    - **Text:** Since the system is currently in the design and development phase, a full-scale evaluation with real user data is not yet possible. However, we have defined a set of proposed evaluation metrics that will be used to assess the system's performance and effectiveness once it is fully developed and deployed. These metrics include:
        - **Accuracy:** The accuracy of the LLM's question-answering capabilities will be measured by comparing its responses to a gold standard set of answers created

by human experts. This will assess the LLM's ability to provide correct and relevant information.

- **Precision and Recall:** Precision will measure the proportion of the LLM's retrieved information that is actually relevant to the user's query. Recall will measure the proportion of all relevant information that is successfully retrieved by the LLM. These metrics will provide insights into the LLM's ability to retrieve both accurate and comprehensive information.
- **F1-Score:** The F1-score, which is the harmonic mean of precision and recall, will be used as a综合 metric to evaluate the overall performance of the LLM in information retrieval.
- **User Satisfaction:** User satisfaction will be assessed through surveys and feedback forms collected from students and researchers who use the system. These surveys will gather data on the system's usability, effectiveness, and overall user experience.
- **Bias Detection and Mitigation:** AI Fairness 360 tools will be used to analyze the LLM's outputs for potential biases related to gender, race, or other sensitive attributes. These analyses will help to identify and mitigate any biases in the system, ensuring fairness and ethical compliance.

These proposed evaluation metrics will provide a comprehensive assessment of the LLM's performance and its impact on student research. The data collected from these evaluations will be used to further refine the LLM and improve its effectiveness over time.