Growing up in Taiwan and studying Electrical Engineering, I was immersed in the world of semiconductors. The precision of giants like TSMC is legendary, but it relies on countless hours of repetitive, enormous human labor, from monitoring fabrication data to identifying microscopic defects. My perspective was transformed during my Machine Learning course, where I realized these very challenges were prime applications for AI. My personal conviction to eliminate repetitive work found its focus: to build intelligent systems capable of automating complex industrial workflows. This ambition raised an essential question that has since guided my academic and professional pursuits: how can these powerful techniques be transformed into **efficient, reliable, real-world applications with tangible impact?** To find the answer, I immersed myself in both academic research and industry, focusing specifically on **ML and Large Language Model applications.** My career goal is to become an **AI R&D Engineer** or an **AI Scientist** at a leading tech company, where I can build intelligent systems that automate complex workflows and streamline routine tasks for people. However, to transition from building novel applications to architecting robust and scalable systems, I recognize the need for a more rigorous foundation in software engineering and system design. Therefore, I am convinced that the Master of Science in Electrical and Computer Engineering in (MSECE) program at Carnegie Mellon University is the essential next step to bridge this gap and achieve my full potential.

My foundational **research experience** progressed from applying traditional ML models to enhance network reliability at WMNLab to using advanced GNNs for security prediction at Academia Sinica, consolidating my belief in ML's power to solve impactful, real-world problems. As LLMs began to proliferate, I saw immense opportunities for optimization and pivoted my focus to this domain in the SPMLLab and RLLab, where **equipped me with LLM architecture concepts and LLM efficiency for real-world applications**. My works centered on a core question: how can we make LLMs more efficient and effective for assisting people in creative tasks? My first project was a multi-agent discussion framework that successfully enhanced LLM creativity, published at COLM 2024. However, I soon identified the practical limitations of this approach, which were high latency and computational cost. This led me to a more refined question: how can we achieve these creative gains efficiently? My subsequent research tackled this directly by exploring representation engineering, achieving a similar creative boost within a single model while drastically reducing computational costs by 95%. This journey from one solution to a more refined one taught me a critical lesson: the best approach is often a trade-off between performance, cost, and scalability under various circumstances. This realization has ignited my desire to master the fundamental principles of AI systems to build truly efficient and reliable AI, a pursuit I aim to deepen through CMU's 18-662 Principles and Engineering Applications of AI.

After establishing a foundation in academic research, I sought to understand how AI concepts translate into viable, real-world products. I started from an internship at Compal Electronics to an opportunity to join the AI startup Genibuilder as a **founding AI R&D Engineer**. Getting this role was a pivotal moment in my life. It served as the industry's validation of my specialized skills, affirming that my knowledge and hard work could generate significant, tangible value. More importantly, this experience crystallized my ambition: if this was the impact I could make with my current expertise, I was determined to discover what I could achieve by pursuing true mastery. This realization is what drives me to seek a Master's degree to deepen my expertise and push the boundaries of what is possible in the field of artificial intelligence. During my year there, my primary project

was building an AI agent to replace aesthetic clinic consultants. This was a multifaceted challenge that encompassed ensuring factual accuracy, maintaining legal compliance, and enabling sales functions and appointment scheduling via robust API and system design. Driven by my principle of automating inefficient workflows, I also tackled a critical internal bottleneck: the time-consuming manual evaluation of our agent. To solve this, I designed an AI system that authentically simulated client-customer dialogues, reducing the team's manual review time by over 80%. These projects confirmed my ambition to build elegant and impactful AI applications that solve real-world problems.

My year of industry experience at a fast-paced startup was a pivotal journey. I had the opportunity to manage end-to-end systems, from AI system design and full-stack architecture to service deployment. While this breadth of responsibility was an incredible learning opportunity, it also highlighted the **gap between my current skills and the sophisticated design** and engineering required to build robust, commercial-grade AI products. This realization has solidified my desire to deepen my expertise through the Master's program at CMU. I am particularly drawn to your curriculum like 18-763: Systems & Toolchains for AI Engineers. Furthermore, my time in the industry has made me recognize the importance of being in an environment that fosters deep technical innovation. While my work in Taiwan has been valuable, the focus is often on applying existing AI models through API calls rather than exploring the core principles behind them. To achieve my goals, I need to be in an environment that encourages a deeper dive into the **core** of these AI techniques. I am confident that the rigorous coursework, research focus, and community of ambitious peers in the Master's program at CMU provides the exact environment I need to grow and achieve my aspirations.

I believe the Master's program in Electrical and Computer Engineering at CMU is the ideal fit for my academic and career goals. I aim to take courses focused on real-world system design and large-scale data processing, which will strengthen my ability to architect the impactful applications I am passionate about. I am particularly drawn to **Professor Sarah Cen's** research on designing and governing AI in diverse domains, which perfectly aligns with my journey of building applications to solve distinct problems in fields like legal tech, medical consulting, and sports analytics. Her recent work, *AI Supply Chains: An Emerging Ecosystem of AI Actors, Products, and Services*, discussed and analyzed the current workflow of AI, which is absolutely insightful. I am excited to be able to contribute to her lab and expand on this same vision. I am equally interested in **Professor Beidi Chen's** work on foundation model understanding and reasoning enhancement. This directly connects with my recent research, where I sought to improve LLM creativity but with method inside the model's activations. I believe her guidance would be invaluable in deepening my understanding of model internals and advancing my focus on LLM interpretability. Finally, after my year as an AI engineer, I deeply appreciate the importance of efficient resource allocation, a challenge faced by any company without a dedicated GPU cluster. For this reason, I am interested in **Professor Greg Ganger's** research on systems and cloud computing, which I am confident will provide the foundational knowledge I need to design scalable and cost-effective AI solutions.

The same **consistency** that drives my passion for AI reflected in my decade-long commitment to basketball. Serving as a point guard on school basketball team from elementary school to college has honed my skills that directly translate to my technical pursuits. Point guard is analogous to that of an engineer: I was tasked with analyzing complex, real-time situations on the court, breaking down problems, and leading my teammates to execute the optimal strategy-acting as the "coach on the court." This

experience was essential in developing my abilities in **rapid analytical problem-solving and leadership** under pressure. I am confident that these skills will be invaluable as I tackle complex research questions and collaborate within the diverse community of Master's program.

In conclusion, my journey has been guided by a core philosophy: a deep trust in the iterative process of identifying a problem, building a solution, and systematically refining it. I am confident that the training at Carnegie Mellon University will help me reach my full potential and support me in achieving my career goals. My enthusiasm and strong foundational knowledge also make me well-suited for this program, and I can contribute to the Master's program in Electrical and Computer Engineering.