DAME

Authors: Jack Corrao, Jack McGarrity, Matthew Sims, Xander Steele, & Garrett Young

December 16th, 2024

I. Introduction

Our project aims to design and build a robot inspired by TARS from *Interstellar*, combining the advanced functionality with a simple, humanlike design. Our TARS robot aka DAME will incorporate the agility mechanics found in the movie, sarcastic voice-communication, and artificial intelligence. By the end, we plan on having a companion-like robot that is able to converse, maneuver, and interact with its environment semi-autonomously.

II. Problem Statement and Proposed Solution

In today's society, there is a growing demand for interactive and intelligent companions that extend beyond virtual assistants like Siri or Alexa. While these AI systems have made significant strides in voice recognition and response, they lack a physical presence that can engage users on a more personal and tangible level. Existing robotic companions in the market are often limited by high costs, simplistic interactions, or lack of adaptability, leaving a gap for a more accessible and emotionally engaging solution. This underscores the need for a small, versatile robot that can serve as a personable AI companion, offering both functional assistance and meaningful interactions.

Our project aims to address this need by designing and creating a compact robot inspired by TARS from the movie *Interstellar*. This robot will function as an AI companion, capable of understanding and responding to user inputs through speech and visuals, thereby providing a more immersive interactive experience. By integrating advanced robotics with AI technologies, we intend to develop a companion that is not only capable of assisting with everyday tasks but also of forming a personal connection with the user. This endeavor will contribute to the field of personal robotics by making sophisticated AI companionship more accessible and by pushing the boundaries of human-robot interaction.

To develop a compact AI companion robot inspired by TARS from *Interstellar*, the proposed solution integrates mobility, artificial intelligence, voice recognition, display capabilities, motor control, speech synthesis, and remote control functionality. The robot aims to interact with users through speech, visual displays, and physical movements, providing a multifaceted user experience.

III. System Requirements

1. **Mechanical Body**:
    a. Requirement: The robot body must resemble the design of TARS from the movie *Interstellar*, but scaled down to be cost-effective and compact enough to fit on tables and other small surfaces. The design must be modular to ensure ease of

mobility and practical use. The robot must be lightweight to allow lifting and repositioning by a single user. Additionally, the body must be durable enough to protect the inner electrical components and withstand bumping into objects and tipping over.

2. **Mobility**
    a. Requirement: The robot must be capable of moving and navigating its environment in a manner inspired by TARS from *Interstellar*. (TARS-like Mobility - see Walk Modulation 2). It should be able to walk forward and turn without falling over. Mobility should not damage joints or the robot body.
3. **Remote Control**
    a. Requirement: Users must be able to remotely/wirelessly operate the robot with a controller, ensuring that the robot responds promptly and accurately to inputs. The user should be able to control the robot remotely at a distance of at least 20 ft.
4. **Artificial Intelligence**
    a. Requirement: The robot must integrate advanced AI capabilities enabling the robot to process complex user inputs, engage in natural conversations, and learn from interactions in real-time. AI must support adaptable and context-aware behavior.
5. **Voice Recognition**
    a. Requirement: The robot must accurately implement speech-to-text conversion to process and interpret speech and respond to wake/sleep commands.
6. **Speech Synthesis**
    a. Requirement: The robot must be able to communicate verbally using text-to-speech technology.
7. **Connectivity and Networking**
    a. Requirement: The robot must be able to connect to the internet to enable AI, voice recognition, and speech synthesis capabilities.
8. **Display Screen**
    a. Requirement: The robot must provide visual feedback and user interfaces via an integrated screen. The robot should display text as it speaks in real time.
9. **Overall User Interface and Usability**
    a. Requirement: Robot must be reasonably easy to set up and operate, with a straightforward user interface. It must reliably interface with the user in real time, with minimal processing delays or errors. The robot should support basic functions such as volume control, sleep/wake modes, power on/off, and reset.
10. **Power**
    a. Requirement: The robot must operate wirelessly using a rechargeable battery system that provides at least an hour of continuous operation per charge. The user must be able to charge the robot easily by plugging it into a standard wall outlet. The power system must be capable of supporting all electrical components

reliably. The batteries must not overheat, or be easily overcharged or overdischarged.

11. Safety

a. Requirement: We will ensure proper safety protocols are followed during testing to prevent electrical fires, overheating, and electrical shocks. High voltage components will be insulated and inaccessible to the user. Systems should automatically shut down if voltage, current, or temperature of components exceed safe values and preventative circuitry should be in place to attempt to mitigate these issues.

IV. System Block Diagram

4.1 Overall System

The DAME project is going to use a combination of mechanical and electrical systems. All of the components and parts will be stored inside a frame that will be designed and fabricated by our team using a 3-D printer. The frame will need to be big enough to house the rest of the components, while having structural integrity, and being light enough to allow for easy motion. All of the software and other components will be controlled by a central microcontroller. The controller needs to have two cores for processing multiple signals at once, wifi and bluetooth connectivity, easily programmable to work with multiple motors, compatible with GPT API's, and be compact enough to fit comfortably within the frame. As of now, the ESP32-WROOM-32 microcontroller fits all of these criteria. The microcontroller will need to connect to multiple systems outside of frame, so the antenna will need to be placed in a location not blocked by the frame, most likely out of the top of the frame. This microcontroller will be central to our project relaying all information for DAME. DAME will be able to move with a remote controller, interact with humans, interact with , connect to power, and be able to mimic to some degree TARS from the movie *Interstellar*. Below is the overall block diagram for all of the systems.

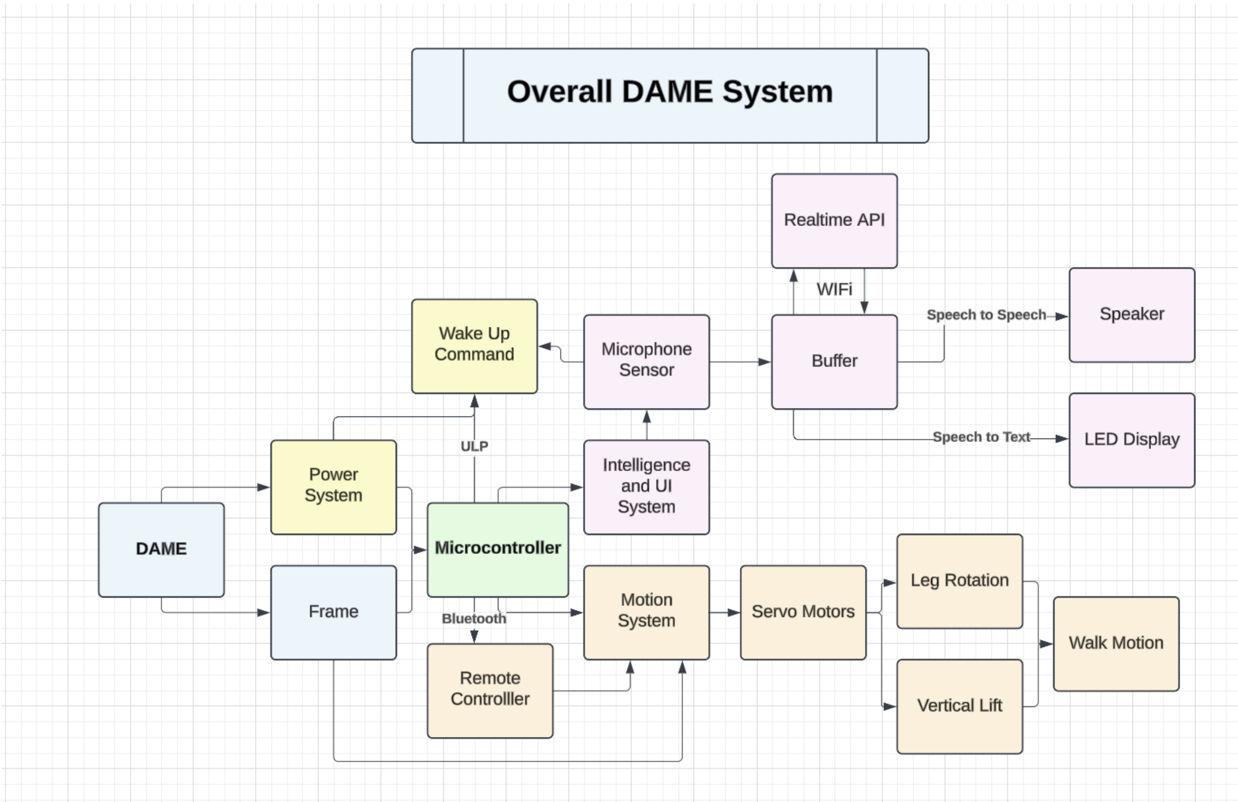

Figure 1. shows the block diagram for the overall system for DAME.

4.2 Motion

A vital aspect to imitate the TARS robot is the ability to move. For this, DAME will use a system of servo motors, controlled by the user with a remote controller. The motor system will be connected to the frame and will have two jobs in order to imitate the movement of TARS. They will need to be able to lift the central body for ease of movement, and then be able to rotate the body and legs in a way that will pull or push DAME forward, backwards, or turn it. They will also have to be light and compact enough to allow for smooth movement and ease of integration into the frame. As of now, we will likely be using the SG90 micro-servo controller, as they are compact and should have the capabilities to move DAME adequately. Should we need more power to move the system, we will move to motors with higher power capabilities. These motors will be controlled by a remote controller allowing for the user to control the movement. This remote controller will be connected to the ESP32 with its bluetooth capabilities.

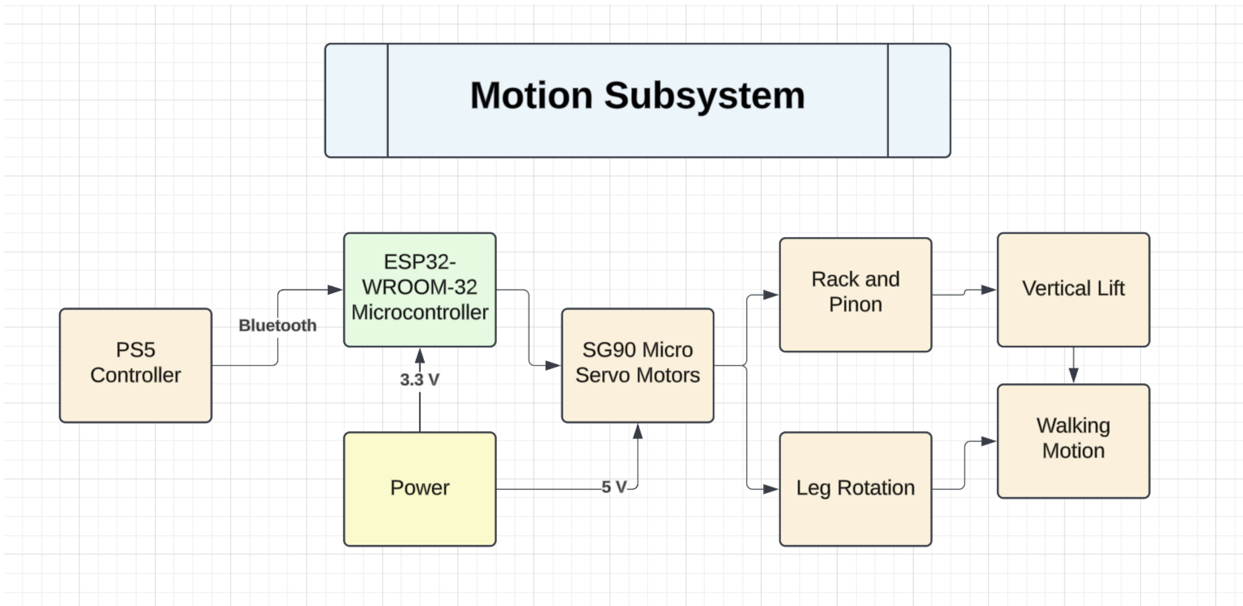

Figure 2. Shows the block diagram for the motion subsystem.

4.3 Intelligence and UI

DAME will need to be able to interact with humans, namely listening, comprehending, and communicating responses. The "listening" will be accomplished using a microphone sensor. Ideally it will be able to operate under low power which will be necessary, as we will implement a wake word to power the system on. This microphone will link to openAI's Realtime API. The Realtime API has both speech to speech capabilities, as well as speech to text capabilities, both of which are requirements for the communication aspect. The information collected by the microphone will be processed by the Realtime API and a response will be generated by the GPT API. This response will be communicated to the user in two ways. Text will be generated on a 5" LED screen with the initial message and the response, keeping a record of the entire conversation. The LED screen will be connected to the microcontroller and will be an output for the text. The other way DAME will communicate to the user with a speaker connected to the microcontroller. The speaker will utilize Realtime's speech to speech capabilities and will ideally be able to respond in real time,as if it's having a conversation with the user. It's important that both of these communication capabilities are able to be done in real time, which is why the Realtime API was chosen for all of this processing and relaying of information.

https://platform.openai.com/docs/guides/realtime?text-generation-quickstart-example=stream - Realtime API

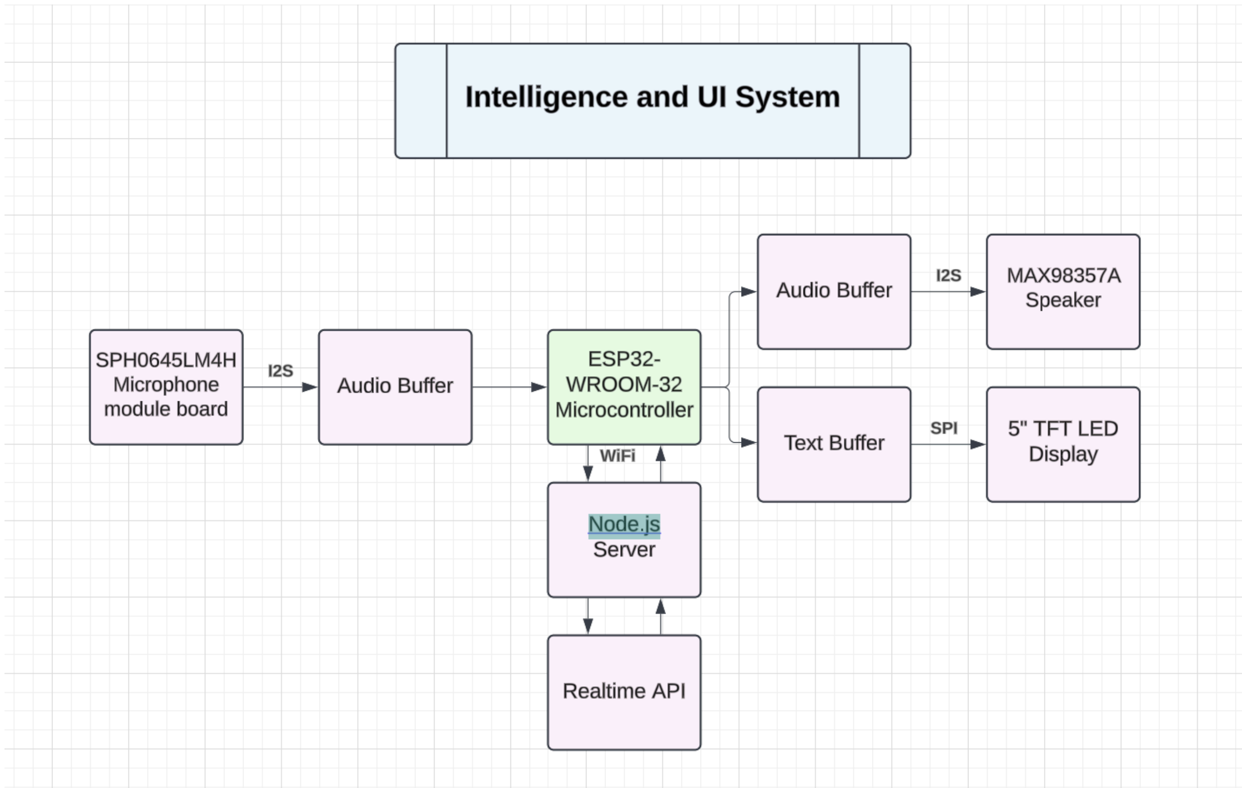

Figure 3. Shows the block diagram for the intelligence and User Interface subsystem.

4.4 Power

The power system must be compact enough to fit into the frame, light enough to not impede motion, and ample enough to provide enough power for all of the other systems and components with a reasonable battery life. The LiPo batteries are good options for DAME's power system, as they should be reasonably sheltered within the frame, have a small light design, and provide enough power to power the system's within DAME. Next to the batteries, we'll need to place voltage regulators which will step down the voltage to the 3.3 Volts that the microcontrollers, microphone, speaker, and other systems use. One other issue is that we want DAME to always be running an ultra low power mode, even if it's not necessarily on. This will allow for the microphone's to pick up on any wake commands to be processed by ESP Skainet, and allow for DAME to be powered on with a voice command.

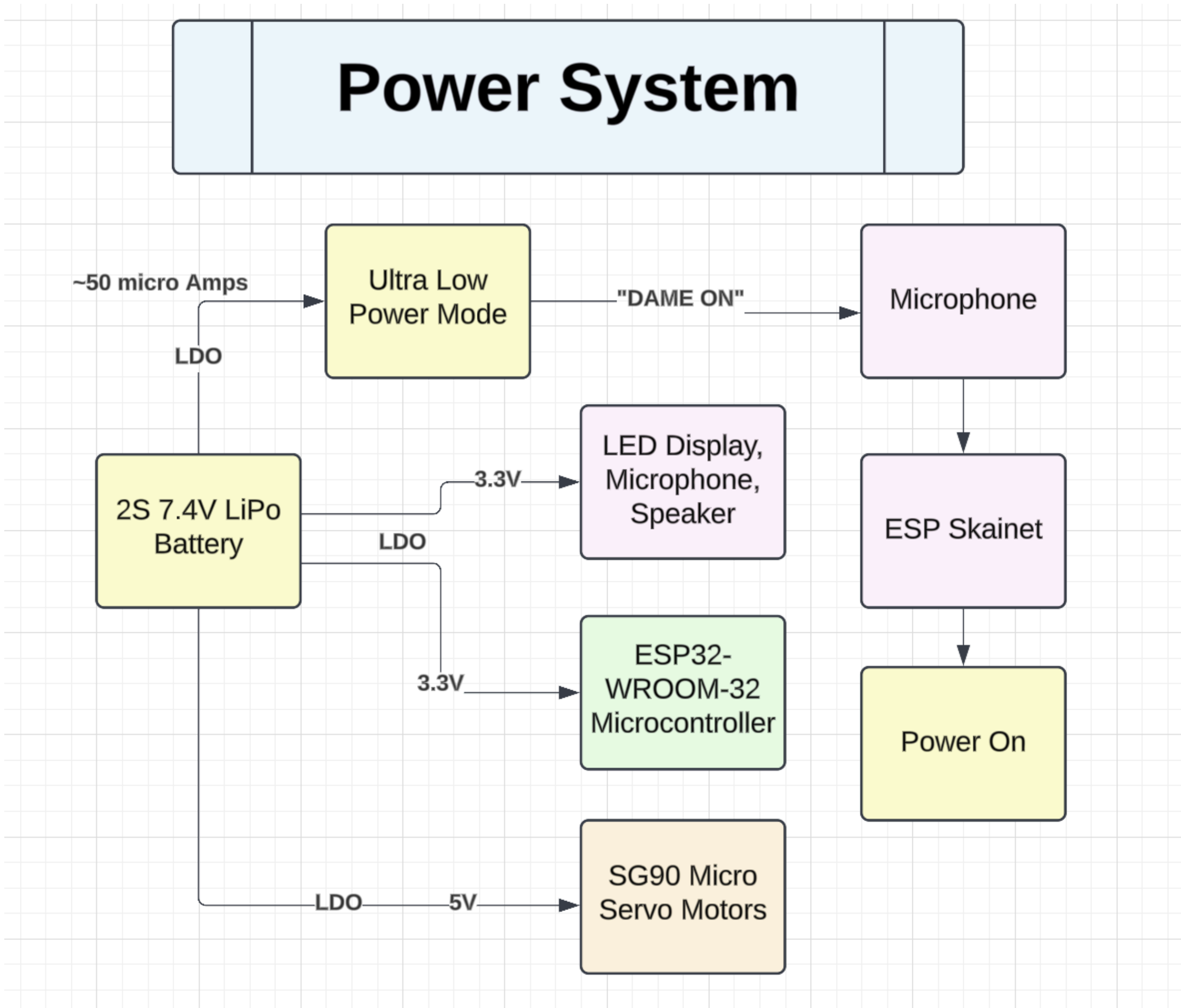

Figure 4. Shows the block diagram for the power system

4.5 Future Enhancements

There will likely be many enhancements that we would like to add to DAME that we won't be able to because of the time frame. Ideally, in the future, DAME would be able to move on its own, using computer vision and collision sensors, giving it the appearance of being even more lifelike. The computer vision would also allow for higher level conversations talking about the environment around DAME. We plan to add the ability to connect to Spotify, allowing for the control of music and vibes of the area in which it resides. Smoothing out the movement will be something we'll constantly strive for, and making it larger, closer to the actual size of TARS would be a goal as well.

V. High Level Design Decisions

A. Mechanical Subsystem

a. Frame:
   i. Design will consist of three rectangular prisms of equal height (around 12 inches) and thickness. These prisms will be mounted together by servo motor axes to resemble *Interstellar's* TARS. The middle prism will have twice the width of the other two prisms. While *Interstellar's* TARS consists of four conjoined prisms of equal width, our motion requirements only require three prisms. The wider middle prisms also provides additional spacing to house our motors, battery, display, and PCBs.
   ii. Prisms will be designed using CAD software and 3D printed in plastic. This satisfies our lightweight requirement while ensuring a more streamlined iterative design process.
   iii. The frame will feature an opening in the torso to allow an antenna to poke through for strong WiFi and bluetooth connectivity.

b. Motion
   i. The "walk" will be accomplished by two mechanisms: leg rotation and torso lift. The middle prism or "torso" will lift linearly before the legs rotate it, swinging the torso and clearing the ground. The torso will then lower itself using the lift mechanism and the legs will rotate back to the robot's neutral position. After this, one "step" will be completed. While this description may not perfectly satisfy our desired motion output, we believe the lift and rotation mechanisms will allow us to accomplish the general walk and may be optimized with creative tweaks and reinforcements after testing.
   ii. Design will include four servo motors. Two motors will actuate the leg rotations (one motor per leg) and two motors will actuate the "lift" mechanism (one motor per leg). We will integrate a rack and pinion mechanism to convert the rotational motion to a linear lift motion in each leg-torso joint. The rack and pinion offers precise linear motion, simple construction, scalability, and durability, suiting our needs well.
   iii. Motors should be light, small, and capable of generating the necessary torque requirements. We are looking at the SG90 or 31311S as potential options; the SG90 is a smaller less powerful motor than the 31311S. We plan to begin designing around the SG90 as the smaller package is

desirable for fitting the mechanisms in our small scale bot. If power is lacking, we will look to the 31311S or other more powerful but space-permitting variant.

iv. To control the movement of the robot, we will implement remote controller functionality over bluetooth.

B. Intelligence and UI Subsystem

a. UI:

i. The primary conversational interface consists of a microphone and speaker. These are to be mounted in the torso of the robot against openings to ensure the audio can be inputted and outputted.

1. Microphone: For our mic, we choose Adafruit's SPH0645LM4H module board. This is a MEMS microphone with an I2S interface popular for small robots and embedded systems applications. It is low power and small with good noise rejection, making it an effective choice for our system.
2. Speaker: For our speaker interface, we choose Adafruit's MAX98357A I2S 3W Class D Amplifier which we will pair with a small 8 ohm, 3W speaker. The I2S interface ensures straightforward comms with the MCU. This setup is similarly a popular choice for small robots and embedded systems: it is small and should be able to provide good sound quality suited to our requirements.

ii. The secondary interface is a display mounted on the face of the torso which displays the text responses of the robot.

1. As HDMI or other parallel interface displays are too intensive for a simple MCU, we are looking at TFT LCD displays with accompanying modules. This would allow simplified comms via SPI and offload a lot of the intensive computing to the separate module. The 5" 800x480 TFT Display without Touchscreen from Adafruit paired with the RA8875 Driver Board would offer SPI control of a nice display, although this is a higher cost option (~$80 total).

b. Intelligence:

i. The design will center around a Node.js server that will act as a "middleman" to handle API processing and sending inputs and outputs to the MCU. This approach simplifies network interactions, offloads processing to improve overall system performance, improves security, and enhances scalability. It opens up more processing resources for the MCU to manage all the various tasks without sacrificing output quality.

ii. Example workflow:
    1. MCU to Node.js: The MCU sends Websocket message to Node.js containing speech mp3.
    2. Node.js to OpenAI: The Node.js server handles authentication, formatting of the API request, error handling, and processes the received API data.
    3. Node.js to MCU: After processing data (extracting speech mp3 and text transcript), Node.js sends it back to the MCU in a simplified formatted response.

iii. We will use OpenAI's Realtime API as it offers speech-to-speech input-output which is optimal for our setup. This allows us to offload all of the STT and TTS to OpenAI, streamlining our processes. Furthermore, OpenAI's intelligence models are near the top of the industry and their customization options will allow us to cater responses to best fit the sarcastic tone we are looking for.

iv. We will use Espressif's Skainet WakeNet tool to handle the "wake" feature required by our robot. This is a straightforward tool that will allow us to turn-on listening via a specific voice command. We will offer additional functionality to control speech input using a button on the remote controller.

C. Power Subsystem

a. We need bus voltages of 5V for motors and 3.3V for everything else—MCU, speaker, microphone, display. To suit our power requirements, we plan to use a 2S LiPo battery pack (7.4 V) and LDOs to supply 5V and 3.3V. This selection is rechargeable and not excessively heavy, which is important for our requirements.

b. We plan to mount the battery inside the bottom of the torso to lower the center of mass. Having weight at the bottom of the torso is essential for stability and balance, especially for the DAME's walk to not land him on his face.

VI. Known Unknowns

The *Interstellar* TARS robot is a marvel of science fiction. It simultaneously possesses futuristic, yet humanistic characteristics in both its mobility and user-interface. With that being said, there are going to be certain anticipated challenges that will make it difficult to fully construct the TARS-robot within the time period of a single semester. The three main known unknown obstacles that we can expect for our project are motor joints for mobility, balancing our robot's weight in both the static, upright position and during motion, and developing a conversational voice user-interface through GPT.

TARS consists of 4 equidimensional rectangular prisms in neighboring columns. The outer two are used as appendages that extend outward, latch onto the ground, then rotate backwards in order to bring the center two prisms forward. The main issue with this design is that any single-axis motor rotation would cause the outside two arms to catch on the floor it is traversing, and topple the robot backwards. Therefore, we plan on having an additional axis of rotation, and will also include 3 rectangular prisms rather than 4 prisms. The secondary axis would allow us to rotate/slide the outer two prisms upwards, overcome the floor-plane, then rotate forward to carry out the rotating appendage motion. In human anatomy, this would be the equivalent of shrugging the shoulders upwards so they raise an inch above their current lowest point, then extending them forward to walk. Three rectangular prisms instead of four will allow for easier housing of the components within the larger central rectangular prism. There will be a number of anticipated obstacles for this known unknown such as having the shoulder joint at the correct height so that the torque doesn't cause the appendage to snap off, having the motion of these two axes work together to provide a fluid movement, and having strong enough motors to lift the chassis forward.

One of the hardest, yet initially overlooked obstacles we expect is balance. When the robot is in transit, it will essentially act as a tripod, with the two arms outstretched and the main chassis behind. This will displace the weight and offer TARS stability. However, when the robot is standing on all fours (each rectangular prism column is flush), the center of gravity must be low enough to ensure that the robot doesn't go face forward or fall backwards. Therefore, while the UI and motor joints may be on the top half of the chassis, the bulk of our electrical components should be in the bottom half. We also plan on including an accelerometer to measure the balance and potentially have motorized weight distribution to counteract any imbalance.

Finally, the third main known unknown we can anticipate is creating a voice-based UI that can respond to questions in a conversational manner rather than a typical CHAT-GPT response. While the MCU→Node.js→OpenAI API setup sounds effective for meeting our needs, we won't know how natural the conversation will be when tested. For example, there might be large delays during processing that make the conversation less smooth. We will need a strong WiFi connection and optimized buffers and processing algorithms to smoothly and efficiently handle all steps of computation. This area of design will require much research, experimentation, and testing.

VII. Major Component Costs

1) Mechanical Systems
    a) SG90 Servo motor (~ $2 per motor)
    b) Additional mechanism materials (gears, rack, axels, etc) (~$25)

        c) Frame Materials (3D-Printed)
2) Microcontroller and Processing Unit
    a) Dual-Core ESP32 WROOM 32 (WiFi/Bluetooth) (~ $10)
3) Wireless Communications
    a) RC (PS or XBox controller) (~ Common)
4) Sensors
    a) Camera for Computer Vision (~ $20) *
    b) MEMS Microphone ($7)
5) Interface
    a) 5” TFT Display ($27.50)
    b) RA8875 Display Driver ($40)
    c) 8 ohm 3W speaker ($5)
    d) Adafruit’s MAX98357A I2S 3W Class D Amplifier ($6)
6) Power
    a) 2S 7.4 V LiPo battery ($11)
    b) HTRC LiPo Battery charger ($10)
7) Software
    a) PlatformIO, libraries pen-source)
    b) Control Algorithms
    c) Node.js server
    d) Speech and Command Processing using Speech APIs (~ $30) *
8) Fabrication
    a) 3D-Printing (~$40)
    b) Custom PCB (~$40)

VIII. Conclusions

In conclusion, our project focuses on designing and building a robot inspired by TARS from *Interstellar*, blending advanced functionality with a straightforward, humanlike design. Our TARS robot, named DAME, will feature agility mechanics reminiscent of those in the movie, along with voice communication and artificial intelligence capabilities. Ultimately, our goal is to create a semi-autonomous companion robot capable of conversing, maneuvering, and interacting with its environment effectively. The world today has a need for a small and personable AI companion and our project aims to address this exact need. We plan to integrate advanced robotics with AI technologies, creating a companion that can assist with everyday tasks while forming a personal connection with the user. We hope to advance the field of personal robotics by making sophisticated AI companionship more accessible, expanding the possibility of human-robot interaction.

Our DAME robot is a sophisticated machine, but dividing it into three main subsystems—motion, intelligence and UI, and power—will enable us to successfully develop our AI companion. These subsystems, when integrated, will enable mobility,motor control, artificial intelligence, voice recognition, display capabilities, speech synthesis, and remote control functionality. Despite breaking down our design into these attainable subsystems, we expect there to be several challenges along the way. The main obstacles we expect to face are developing motor joints for mobility, balancing our robot through varied use, and developing a friendly voice user-interface through GPT. We believe being able to adapt to difficulties as they occur along with our approach to each of these challenges will ensure success despite the roadblocks accompanying each.

Through a careful sectioned approach to our DAME robot we propose an attainable personal AI companion robot. Dividing our project into three main subsystems, while proactively addressing the challenges that are likely to arise, will help us maintain a focused and effective development plan. Additionally if time allows, we plan to incorporate computer vision, collision sensors, and/or connectability to Spotify.