Democratic People’s Republic of Algeria
Ministry of Higher Education and Scientific Research
University of 20 August 1955 – Skikda

Faculty of Sciences
Department of Computer Science

Project for obtaining a master's degree + patent certificate under Ministerial Decision 1275

Smart Sign Bridge: Smart Glasses for Real-Time Speech-to-Sign Language Translation Empowering Deaf Communication

---

Prepared by:
NOUAR Wissam
ZIREK Anfal

Supervised by:
Dr.Cheikh Mohamed

Specialization: **M**aster **2 Artificial Intelligence**

Defended Before the Jury:

President:

Examiner:

BMC Expert:

Commercial Partner Representative:

July 2025

Acknowledgements

In the preamble to this thesis, we thank God who helps us and gives us the patience and courage during these long years of study. We also wish to thank our very dear parents who sacrifice every day for us in order to see us succeed in our lives.

*We would also like to warmly thank our working group, [**ZIREK Anfal, NOUAR Wissam**], for the exemplary collaboration, mutual support, and dedication throughout this study. Our individual contributions were essential in enriching our discussions and deepening our analysis.*

*We express our deep gratitude to our research director Dr. **Cheikh Mohamed** for his wise advice, availability, and constant encouragement throughout this project.*

Our thanks are also dedicated to the jury members for kindly agreeing to review our work.

Finally, we extend our most sincere thanks to all our families, loved ones, and friends for their unconditional support, understanding, and constant encouragement.
***Thank you**.*

Dedication

*We dedicate this final year project to**:***

Our parents who supported and encouraged us throughout our years of study.

Our brothers and sisters who shared all the emotions with us during the completion of this work.

Our loved ones, our friends, to all those who have given us love, perseverance.

Our professors who have ensured to guide us throughout our journey.

All the colleagues from the class.

Thank you

Abstract

Deaf and hard-of-hearing societies live in a deaf world that is dominated by hearing. Their lives take the shape of a continuous motion of barriers. With no chance to communicate openly with speech, they do what they can have as a result of sheer determination. They illuminate their lives with the hope of technological advancement. Our innovation, Smart Sign Bridge, a groundbreaking AI-driven smart glasses, promises to transform deaf and hard-of-hearing lives. The glasses utilize state-of-the-art artificial intelligence in the form of speech recognition and natural language processing to provide an intuitive, real-time interface that translates voice into emotive 3D sign language avatars hovering in front the eyes.

Smart Sign Bridge makes communication easy in diverse contexts offices, classrooms, hospitals, and parties accepting and interpreting words delivered from diverse sources such as lectures, dialogue, and public address. Furthermore, it is adaptable with local sign languages, i.e., Algerian Sign Language, to support cultural diversity and inclusivity. By empowering independent, hands-free communication, Smart Sign Bridge maximizes independence and social interaction. Finally, our ultimate aspiration is to advance Smart Sign Bridge even more so that it can make a real difference in the life of the deaf and hard-of-hearing population to result in a society where their voice, as they convey it through signs, is being listened to and valued.

Keywords:

Artificial Intelligence (AI), Speech Recognition, Natural Language Processing (NLP), Real-time Translation, 3D Avatars, Human-Computer Interaction, Assistive Technology.

ملخص

تعيش مجتمعات الصم وضعاف السمع في عالم أصم تهيمن عليه السمع. تتخذ حياتهم شكل حركة مستمرة من الحواجز. بدون فرصة للتواصل المفتوح بالكلام، يفعلون ما يستطيعون فعله نتيجة للعزيمة الصادقة. يضيئون حياتهم بأمل التقدم التكنولوجي. ابتكارنا، Smart Sign Bridge، نظارات ذكية مدفوعة بالذكاء الاصطناعي، تعد بتحويل حياة الصم وضعاف السمع. تستخدم النظارات الذكاء الاصطناعي المتقدم في شكل التعرف على الكلام ومعالجة اللغة الطبيعية لتوفير واجهة بديهية في الوقت الفعلي تترجم الصوت إلى رموز لغة الإشارة ثلاثية الأبعاد العاطفية التي تحوم أمام العيون.

تجعل Smart Sign Bridgeالتواصل سهلاً في سياقات متنوعة مثل المكاتب، والفصول الدراسية، والمستشفيات، والحفلات، حيث تقبل وتفسر الكلمات المنطوقة من مصادر متنوعة مثل المحاضرات، والحوار، والخطابات العامة. علاوة على ذلك، فإنه قابل للتكيف مع لغات الإشارة المحلية، أي لغة الإشارة الجزائرية، لدعم التنوع الثقافي والشمولية. من خلال تمكين التواصل المستقل بدون استخدام اليدين، تعزز Smart Sign Bridgeالاستقلالية والتفاعل الاجتماعي إلى أقصى حد. أخيرًا، طموحنا النهائي هو تطوير Smart Sign Bridgeبشكل أكبر بحيث يمكنه إحداث فرق حقيقي في حياة السكان الصم وضعاف السمع، مما يؤدي إلى مجتمع تُستمع فيه أصواتهم، كما يعبرون عنها بالإشارات، وتُقدَّر قيمتها.

الكلمات المفتاحية:

الذكاء الاصطناعي(AI) ، التعرف على الكلام، معالجة اللغة الطبيعية(NLP) ، الترجمة الفورية، الصور الرمزية ثلاثية الأبعاد، التفاعل بين الإنسان والحاسوب، التكنولوجيا المساعدة.

Résumé

Les sociétés sourdes et malentendantes vivent dans un monde sourd dominé par l'audition. Leurs vies prennent la forme d'un mouvement continu de barrières. Sans possibilité de communiquer ouvertement par la parole, ils font ce qu'ils peuvent grâce à une détermination sans faille. Ils illuminent leur vie avec l'espoir des avancées technologiques. Notre innovation, Smart Sign Bridge, des lunettes intelligentes révolutionnaires alimentées par l'IA, promet de transformer la vie des sourds et malentendants. Les lunettes utilisent une intelligence artificielle de pointe sous forme de reconnaissance vocale et de traitement du langage naturel pour fournir une interface intuitive en temps réel qui traduit la voix en avatars de langue des signes 3D émouvants flottant devant les yeux.

Smart Sign Bridge facilite la communication dans divers contextes bureaux, salles de classe, hôpitaux et événements, en acceptant et en interprétant les mots provenant de diverses sources telles que les conférences, les dialogues et les discours publics. De plus, il est adaptable aux langues des signes locales, c'est-à-dire la langue des signes algérienne, pour soutenir la diversité culturelle et l'inclusivité. En permettant une communication indépendante et sans les mains, Smart Sign Bridge maximise l'indépendance et l'interaction sociale. Enfin, notre aspiration ultime est de faire progresser Smart Sign Bridge encore davantage afin qu'il puisse faire une réelle différence dans la vie des personnes sourdes et malentendantes, pour aboutir à une société où leur voix, telle qu'elle est exprimée par des signes, est écoutée et valorisée.

Mots-clés :

Intelligence Artificielle (IA), Reconnaissance Vocale, Traitement du Langage Naturel (TLN), Traduction en Temps Réel, Avatars 3D, Interaction Humain-Ordinateur, Technologie d'Aide.

Table of contents

Chapter 03: System design and implementation

List of figures

List of tables

General introduction

General introduction

Picture a world in which all the laughter that is shared at a family reunion, every classroom lesson taught, and every emergency communication in a hospital is audible to all whether or not they can hear. For the deaf and hard-of-hearing worldwide, this is not their world. These vibrant members of society, full of talent, vigor, and varying perceptions, face daily challenges far greater than the absence of sound. The pain of being excluded from interactions, the frustration of laboring an environment where co-workers are not acquainted with sign language and the fear of losing valuable health information due to a lack of interpreters instill a profound sense of isolation.

Research indicates that deaf individuals experience much higher rates of loneliness, depression, and anxiety because of these barriers. In most of the globe, especially in low-resource communities, access to assistive devices like hearing aids or cochlear implants is limited, and when possible, these devices often are not compatible with the population who utilize sign language as their core tool for communication. Human interpreters, although valuable, are scarce and costly, and text-based software can be clumsy, disrupting the natural flow of human communication. These issues emphasize a pressing need for innovative solutions that not only bridge communication gaps but also honor the cultural and linguistic identity of deaf communities.

This project presents "Smart Sign Bridge," an innovative idea to transform the lives of deaf and hard-of-hearing individuals through AI-powered smart glasses. Smart Sign Bridge isn't merely a technological advancement, it's a passionate commitment to accessibility, designed to translate spoken language into expressive sign language movements in the moment, cast discreetly into the user's line of vision. By providing users with the freedom to join conversations naturally at school, in the workplace, or any other places, Smart Sign Bridge aims to revive the elegance of communication and the dignity of independence.

The dissertation consists of four chapters, and each of them weaves the threads of science, technology, and social benefit together.

Chapter 01 evaluates the human auditory system, the heterogenous etiologies of deafness, the profound social and emotional challenges faced by deaf individuals and underlines the limitations of assistive technologies that are currently available.

Chapter 02 explores the critical contribution of artificial intelligence, especifically machine learning, speech recognition, and natural language processing, to the facilitation of real-time communication platforms.

Chapter 03 presents the innovative design and technical realization of the Smart Sign Bridge prototype, highlighting its synergy of AI, wearable technology, and user-focused features to provide inclusive, culturally appropriate communication.

Chapter 04 sketches out business vision with affordability, scale, and worldwide access to ensure that Smart Sign Bridge is able to serve multilingual communities, including deaf people who use regional sign languages like Algerian Sign Language.

Inspired by values of solidarity and creativity, Smart Sign Bridge seeks to break down borders, to create belonging, and to make a world where deaf people can connect, learn, and thrive in confidence and ease.

Chapter 01: Human Auditory System and Assistive Technologies for Deaf Individuals

1 Introduction :

Hearing is one of the foremost basic faculties for human communication, empowering people to connecte, learn, and interface with the world around them. For those who involvement hearing misfortune or are born hard of hearing, this fundamental channel of communication is disturbed, driving to significant challenges in existence. Essentially, muteness whether innate or acquired further compounds these troubles, making obstructions to expression and social integration. Understanding the causes, types, and consequences of hearing loss is fundamental to creating viable arrangements that can bridge these communication crevices. This chapter provides a comprehensive review of the state of the art in assistive technologies for deaf individuals.

Figure 1 : Deaf person

2 Human auditory system:

The auditory system processes how we hear and understand sounds within the environment. Peripheral and central structures comprise this organ system. The outer, middle, and inner ear are the peripheral auditory structures. The ear is a very efficient transducer (i.e., a device that changes energy from one form to another), changing sound pressure in the air into a neural-electrical signal that is translated by the brain as speech, music, noise, etc. The external ear, middle ear, inner ear, brainstem, and brain each have a specific role in this transformation process [1].

3 The importance of human auditory system:

The human auditory system has a significant impact in numerous fields.

3.1 Part in communication:

- **Discourse recognition:** The sound-related framework permits us to listen and get it talked dialect, which is the essential mode of communication for most individuals.
- **Social interaction:** Hearing empowers us to lock in in discussions, recognize feelings through tone of voice, and construct connections.
- **Dialect improvement:** For children, the sound-related framework is fundamental for learning to talk and get it dialect. Hearing misfortune in early childhood can essentially delay dialect securing.

3.2 Natural mindfulness:

- **Security:** The sound-related framework makes a difference us distinguish potential threats, such as drawing closer vehicles, cautions, or caution yells.
- **Spatial introduction:** Hearing permits us to find the source of sounds, which is basic for exploring our environment.
- **Natural sounds:** Regular sounds, like winged creatures chirping or music, improve our encounters and contribute to our quality of life.

3.3 Cognitive and enthusiastic advancement:

- **Learning:** Hearing is fundamental for instruction, as most instructing strategies depend on talked instruction. Hearing misfortune can prevent scholarly execution.
- **Memory and consideration:** The sound-related framework makes a difference us prepare and keep in mind data, such as addresses or discussions.
- **Enthusiastic Well-Being:** Sounds like music or an adored one's voice can bring out feelings and give consolation.

3.4 Proficient and Social Integration:

- **Work environment communication:** Compelling hearing is significant for taking part in gatherings, taking after enlightening, and collaborating with colleagues.
- **Social interest:** Hearing empowers people to lock in in bunch exercises, go to occasions, and keep up social associations [2].

4 Ear constitution:

The ear is the organ of hearing and balance, that includes a few fundamental structures for hearing. Here is an overview of the constitution of the ear based on available information. The human ear may be an advanced organ composed of numerous key parts that work together to capture, process, and transmit sound-related data to the brain. [2].

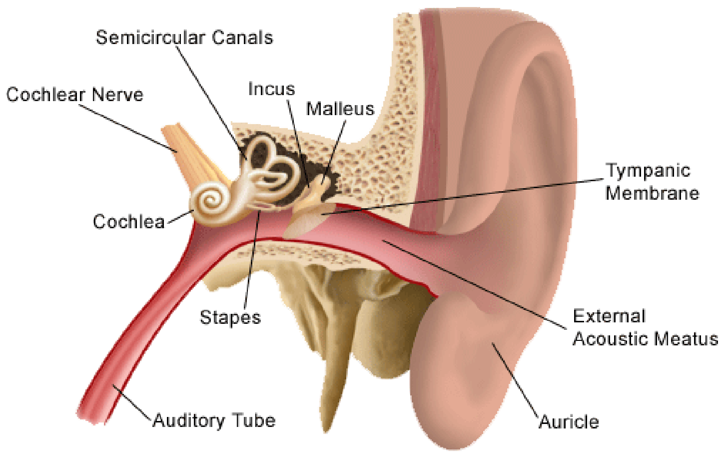

Figure 2 : Anatomy and Physiology of the ear

The parts of the ear include:

External ear (outer ear), consisting of:

- **Pinna (auricle):** This is the outside part of the ear.
- **External auditory canal (tube) :** This is the tube that connects the outer ear to the inside or middle ear.
- **Tympanic membrane (eardrum):** The tympanic membrane divides the external ear from the middle ear.

Middle ear (tympanic cavity), consisting of :

- **Ossicles:** Three small bones that are connected and transmit the sound waves to the inner ear. The bones are called: Malleus, Incus, Stapes.
- **Eustachian tube:** A canal that links the middle ear with the back of the nose. The eustachian tube helps to equalize the pressure in the middle ear. Equalized pressure is needed for the

proper transfer of sound waves. The eustachian tube is lined with mucous, just like the inside of the nose and throat.

Inner ear, consisting of:

- **Cochlea:** This contains the nerves for hearing.
- **Vestibule:** This contains receptors for balance.
- **Semicircular canals:** This contains receptors for balance [3].

5 Deafness :

5.1 Hearing loss:

A person who is not able to hear as well as someone with normal hearing hearing thresholds of 20 dB or better in both ears is said to have hearing loss. Hearing loss may be mild, moderate, moderately severe, severe or profound. It can affect one ear or both ears and leads to difficulty in hearing conversational speech or loud sounds [4].

Hearing loss exists when there is diminished sensitivity to the sounds normally heard. The term hearing impairment is usually reserved for people who have relative insensitivity to sound in the speech frequencies. The severity of a hearing loss is categorized according to the increase in volume above the usual level necessary before the listener can detect it [5].

People who are hard of hearing usually communicate through spoken language and can benefit from hearing aids, cochlear implants, and other assistive devices as well as captioning.

Deaf people mostly have profound hearing loss, which implies very little or no hearing. They can benefit from cochlear implants. Some of them use sign language for communication [4].

Figure 3: Hearing loss

5.2 causes of hearing loss and deafness:

Although these factors can be encountered at different periods across the life span, individuals are most susceptible to their effects during critical periods in life. Here are the main causes:

- **Genetic:** Hearing loss is a sensory disorder that is often found, about one in a thousand births. Hearing loss can be caused by genetic factors, environmental factors, and the interaction of these two factors. Genetic factors play about 50 to 75 percent as a cause of hearing loss.
- Hearing loss related to genetic factors (congenital hearing loss) can be found in two forms: syndromic disorders and non-syndromic disorders. About 70 percent of hereditary hearing loss is in the form of non-syndromic disorders.
- GJB2 gene mutation is the main cause of congenital deafness because due to the modification of codons that encode the connexin 26 protein, the stability of mRNA will be affected. Worldwide, about 50 percent of the GJB2 gene mutation causes non-syndromic recessive congenital deafness [6].

Figure 4: Gene mutation.

- **Neurological disorders:** Neurological disorders such as multiple sclerosis and strokes can have an effect on hearing as well. Multiple sclerosis (MS) is an autoimmune disease where the immune system attacks the myelin sheath, a covering that protects the nerves. Once the myelin sheaths are destroyed there is no possible way at present to repair them. Without the myelin to protect the nerves, nerves become damaged, creating disorientation for the patient. This is a painful process and may end in the debilitation of the affected person until they are paralyzed and have one or more senses gone. One of those may be hearing. If the auditory nerve becomes damaged, then the affected person will become completely deaf in one or both ears. There is no cure for MS. Depending on what nerves are damaged from a stroke, one of the side effects can be deafness [7].

- **Childhood and adolescence:** The auditory system is susceptible during childhood and adolescence to conditions that might lead to hearing loss, including chronic ear infections, fluid buildup in the ear and serious infections like meningitis. Chronic suppurative otitis media is characterized by long-term ear discharge and infections that can damage middle ear structures, whereas chronic nonsuppurative otitis media (often called glue ear) leads to buildup of fluid that dampens sound and can cause delayed speech development. Meningitis and other systemic infections can damage the auditory nerve or inner ear, risk permanent hearing impairment if not addressed promptly [7].
- **Adulthood and older age:** In adults, chronic diseases such as diabetes or cardiovascular disorders, for example, are also associated with hearing loss. Such conditions can restrict or damage blood flow to the inner ear or impair auditory nerves, gradually reducing hearing ability over time. Otosclerosis (abnormal bone growth around the middle ear bones) typically develops in adulthood. This gradual disorder causes the stapes bone to become immobilized, resulting in conductive hearing loss that can progress without surgical or medical treatment.

  age-related sensorineural hearing loss, or presbycusis. This progressive, bilateral form of hearing loss is caused by degeneration of the inner ear or auditory nerve, so sounds have to be louder to register, and it becomes difficult to hear high-frequency sounds and understand speech in noisy situations.

  Sudden sensorineural hearing loss (SSNHL) is one of the most researched causes of rapid hearing loss, caused by a medical emergency that typically occurs in one ear over the course of hours or days. Related to viral infections, vascular problems, or autoimmune conditions, it needs urgent treatment to ensure a better chance of recovery [7].

6 Diffrent forms of deafness:

6.1 Types of deafness:

Functionally, the human ear can be divided into two major divisions, the conductive division, associated with the areas responsible for air conduction (the outer ear and the middle ear) and the sensorineural division associated with the inner ear. Accordingly, the three main types of hearing loss are classified as conducive, sensorineural, and mixed hearing losses.

- **Conductive hearing loss:** CHL is a type of hearing loss characterized by having better hearing thresholds for bone-conducted signals compared with air-conducted signals. CHL is usually associated with dysfunction located in the outer and/or middle ear while having a normal inner ear function. In CHL, the audiogram typically shows normal bone conduction (0–25 dB) and abnormal air conduction threshold levels (higher than 25 dB). According to the American Speech-Language-Hearing Association, a difference greater than 10 dB is considered a significant air–bone gap and requires the use of masking to eliminate a response from the ear not being tested, hence obtaining true thresholds from the test ear [8].

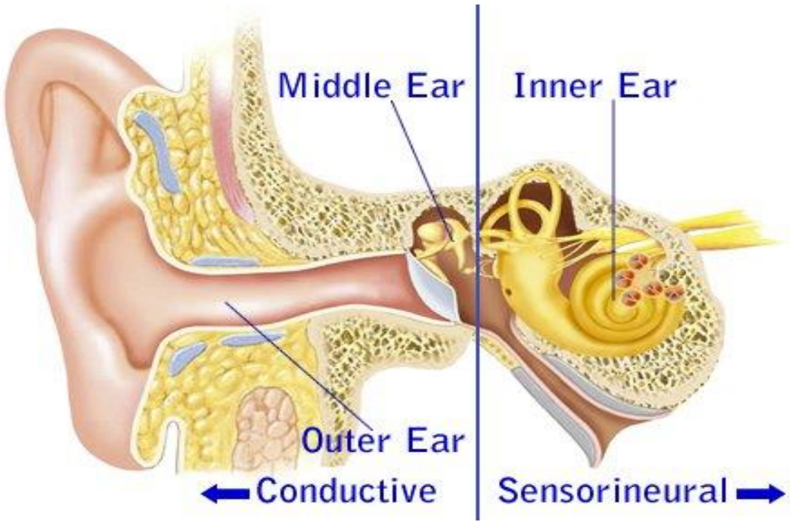

Figure 5: Types of hearing loss

- **Sensorineural hearing loss:** SNHL is a hearing loss that occurs as a result of damage in the cochlea or beyond, that is, either along the 8th cranial nerve or in the brain. SNHL can cause complete loss of hearing, despite the outer ear and middle ear being normal. Individuals with SNHL demonstrate similar air and bone conduction thresholds. The sensory component is usually due to the damage to the organ of Corti or to an inability of hair cells to stimulate the auditory nerve [8].
- **Mixed hearing loss:** Mixed hearing loss is a type of hearing loss that has a combination of conductive and sensorineural damage in the same ear. Cases where both an air–bone gap greater than 10 dB and an elevated bone conduction threshold are observed suggest a mixed hearing loss. While the conductive component may be treated, the sensorineural component is more of a challenge [8].

6.2 **Degree of hearing loss:**

Degree of hearing loss refers to the severity of the loss. Severity of hearing loss is based on thresholds at individual frequencies. Once the type and degree of loss are established, an

appropriate intervention may be assigned. This may include hearing aids, aural rehabilitation, cochlear implants, medical intervention, or surgery.
The table below shows one of the more commonly used classification systems. The numbers are representative of the patient's hearing loss range in decibels (dB HL) [9].

.

Degree of hearing loss	**Hearing loss range (dB HL)**
Normal	-10 to 15
Slight	16 to 25
Mild	26 to 40
Moderate	41 to 55
Moderately severe	56 to 70
Severe	71 to 90
Profound	90 +

Tableau 1: Degree of hearing loss

7 Challenges faced by deaf individuals :

Deaf individuals encounter a variety of challenges in their daily lives:

- **Communication Barriers:** It can be difficult for deaf people to express themselves and fully participate in conversations because spoken language is the primary mode of interaction for most people. While sign language is an essential tool for many deaf people, not everyone in the hearing community understands or uses it, which creates a barrier in everyday interactions, from casual conversations to professional settings. Another common communication method, lip reading, is only partially accurate and highly dependent on the speaker's mouth being visible, which frequently results in misunderstandings. Written communication is also helpful, but it can be slow and impractical in situations that are fast-paced or face-to-face [10].

- **Social isolation:** Deaf persons frequently experience social isolation as a result of their inability to interact with hearing people. When spoken language predominates at social events, group discussions, and community meetings, they could feel left out. Because it becomes difficult to establish and sustain relationships with hearing peers, this exclusion can lead to feelings of loneliness and melancholy. Since this seclusion can eventually affect their general and mental health, it is critical to establish more welcoming social settings.
- **Educational Challenges:** Deaf children frequently face major challenges in educational settings. Adequate accommodations, such captioning services, sign language interpreters, or teachers with deaf education training, are lacking in many schools. Deaf kids might find it difficult to stay up with their peers without these services, which could result in poorer academic achievement and fewer possibilities for postsecondary education. These difficulties are further compounded by the dearth of accessible educational resources, such as captionless films or textbooks created for deaf students. Their personal growth and career prospects may be negatively impacted for some time by this educational disparity [10].
- **Access to healthcare:** For those who are deaf, accessing healthcare services can be very difficult. Many medical professionals lack the tools or training necessary to interact with deaf patients in an effective manner, such as visual aids or sign language interpreters. Miscommunication, incorrect diagnosis, or insufficient therapy may result from this. Important health information, like emergency warnings or medical instructions, may also be difficult for deaf people to obtain because it is frequently provided in spoken or written forms that may not be completely accessible.
- **Employment discrimination:** Discrimination and obstacles are commonplace for Deaf people in the job. Many companies don't know what deaf workers can do or don't want to make the required accommodations, including visual aids or sign language interpreters. Because of this, it may be challenging for deaf people to get employment, grow in their careers, or even carry out simple responsibilities at work. Misunderstandings or poor communication with hearing coworkers can also result in irritation and feelings of exclusion, which further restricts their ability to advance professionally [11].

8 Assistive technologies for deaf and hard-of-hearing individuals:

The integration of assistive technologies into daily communication processes has become an essential component in promoting the inclusion of individuals with disabilities. For the deaf and hard-of-hearing individuals, these technologies provide critical means of bridging communication gaps with hearing community. This section critically analyzes existing solutions, identifying limitations that boost the need for more sophisticated, inclusive alternatives.

8.1 Auditory centric assistive technologies:

8.1.1 Hearing aids:

Functionality of hearing aids: Hearing aids are personally worn electrical devices that process and amplify the incoming sounds according to an individual's unique hearing configuration. Before consulting a hearing aid specialist, ear conditions should be treated medically and, if possible, surgically. Hearing aids are designed to improve speech understanding of the patient with hearing impairment.

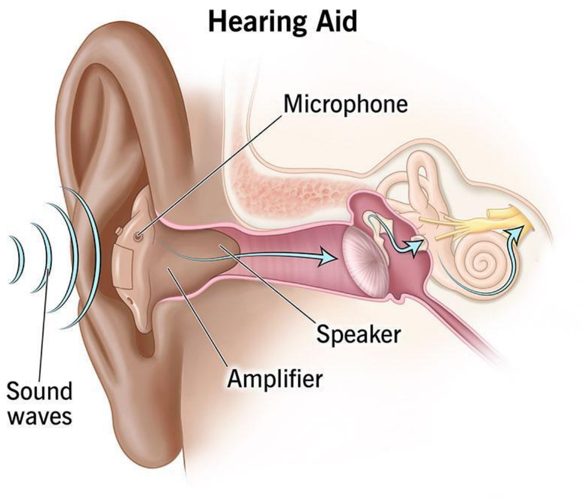

Figure 6: Hearing Aid

They contain different components that work together to amplify sound. The three basic hearing aid parts include a microphone, an amplifier and a speaker (reciever). Via the microphone, sounds enter. The sound waves are then transformed into electrical impulses by the microphone before being sent to the amplifier. Lastly, the amplifier uses a tiny speaker to deliver the sounds to the ear [12].

Real-world limitations of hearing aids: Despite major advances in digital signal processing, hearing aids continue to underperform in acoustically complex environments. A recent study reported that 62% of hearing aid users identify amplified background sounds as a key obstacle, impairing their ability to selectively attend to conversations and track speech effectively [13].

While some devices feature AI-based denoising and directional microphones, most commercial hearing aids still lack fully adaptive, real-time AI-driven directional filtering that can dynamically extract speech from noise in complicated auditory environments [14].

Furthermore, high costs and limited insurance coverage continue to restrict access to these devices for many individuals globally, contributing to widespread underutilization [15].
Critically, user preference surveys and cultural studies show that many Deaf people with profound hearing loss find hearing aids inadequate for their communication needs, particularly those raised in sign language culture who prefer visual communication modes over auditory amplification [16].

8.1.2 Cochlear implants:

Mechanism of cochlear implants: A cochlear implant is an electronic system designed to produce useful hearing sensations to a person with severe to profound nerve deafness by electrically stimulating nerves inside the inner ear [17].

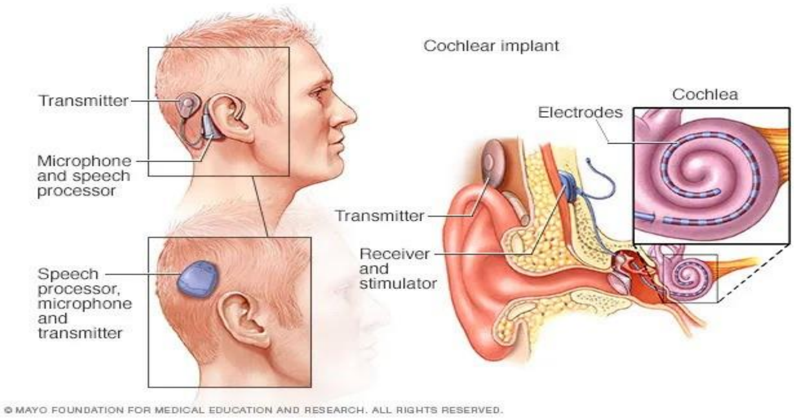

Figure 7: Cochlear implant

The cochlear implant system is made up of an internal part (implanted under the skin surgically) and an external part (worn on the skin surface of the head). The external part (which typically looks like a hearing aid) is often referred to as a 'speech processor'. The microphone(s) on the speech processor picks up sound and converts it into an electrical signal. This electrical signal is passed to the antenna and sent to the receiver on the internal part of the implant, under the child's scalp. The antenna is held near to the internal receiver using magnets.

The internal part converts the signal into a tiny electrical current that is passed down an electrode that has been put in the cochlea. This stimulates the auditory nerve. In most people with hearing loss, the hearing nerve works normally and stimulating it electrically can stimulate the brain and reproduce hearing. [18]

Accessibility and sound quality challenges: Despite being an essential tool for many people with profound hearing loss, cochlear implants are constrained by unmet needs for accessibility and sound quality, as well as high surgical risks and financial constraints. While new AI features are on the horizon, a substantial portion of users still experience unnatural sound perception [19].

About 10% of cases result in surgical problems, such as flap necrosis and device extrusion, which call for careful pre- and post-operative care.

In low- and middle-income countries, where only 5-10% of people who could benefit receive the intervention, the expensive expense of the device, surgery, and lifelong rehabilitation constitutes a significant obstacle [20], this leads to a significant unmet medical demand for reasonably priced hearing aids [21]. Concerns about cultural identity and communication autonomy also continue to exist among Deaf communities, which affects how cochlear implants are perceived and accepted as a solution.

8.1.3 Assistive listening devices (ALD):

Functionality of ALDs: Assistive listening devices (ALDs) are specialized instruments designed to enhance auditory accessibility and amplify sound for people with hearing loss, particularly in situations where cochlear implants or hearing aids are insufficient. These devices enhance the signal-to-noise ratio, making speech clearer in noisy environments or poor acoustics. They encompass technologies like frequency modulated (FM), infrared, and hearing loop systems for large venues, as well as portable amplifiers and remote microphones for personal use. [23] [24]

Infrastructure and portability limitations: Despite their intended role in improving auditory access, ALDs remain limited by infrastructure dependency, lack of portability, and user control [25]. ALDs are mostly effective when they are installed automatically in public areas 78% of deaf people who travel to foreign countries report that they cannot use them in airfields, hotels, or meeting rooms [26]. ALDs also lack intelligent features, such as noise-source differentiation or user-adaptive filtering, which may otherwise enable deaf users to communicate in noisy environments [27].

8.2 Human-centered sign language communication:

8.2.1 Sign language interpreting services:

Sign Language Interpreting Services (SLIS) are professional communication facilitation services that bridge the interaction between individuals who use signed languages (such as American Sign Language, Arabic SL) and those who use spoken languages. As defined by the National Registry of Interpreters for the Deaf, interpreting involves the complex process of conveying meaning, intent, and cultural context between these distinct linguistic modalities.

Sign language interpreters do much more than translate words, they build bridges between worlds. In classrooms, hospitals, courtrooms, and offices, these professionals serve as vital links, ensuring Deaf individuals can fully participate in conversations that shape their lives, education, and well-being. Their work demands extraordinary skill. Beyond being perfectly bilingual, interpreters must be culturally fluent, understanding the unspoken norms and values of both Deaf and hearing communities. [28]

Figure 8: Tactile interpreting

However, there are different types of interpreters based on the needs of the client, such as but not restricted to:

- **Oral interpreting:** used for those who do not sign or not proficient in ASL. An oral interpreter is someone trained in how to silently mouth speech and use gestures for non-signing deaf consumers. The use of facial expressions and gestures are added to enhance understanding for those who lip read.
- **Tactile interpreting:** a common means of communication used by people with DeafBlindness. Tactile is based using SL with the touch “hand over hand” using their hands loosely on top of the DeafBlind individual’s hands to receive the information.
- **Protactile interpreting:** a relatively new method of tactile communication that allows an individual who is DeafBlind to have more reliable information about the visual elements of their environment (people smiling/nodding). Protactile allows people to establish tactile communication channels in different ways.

8.2.2 Scalability challenges and gaps in human interpretation:

The demand for human sign language interpreters far exceeds availability, especially in rural and emergency settings where access is critically limited. According to the Journal of Deaf Studies and Deaf Education, this lack causes lengthy wait periods for necessary services, with emergency medical interpretation frequently taking longer than 48 hours. Manufactured human interpreting services cannot scale to meet real-time, ubiquitous communication needs of Deaf users. Furthermore, no existing automated system provides reliable low-latency (200 ms) sign language interpretation at high cost, an essential opportunity for AI and computer vision technologies to further their progress in gesture recognition and real-time translation models [29].

8.3 Smart wearables for real-time speech translation:

Smart devices for real-time translation, when designed for deaf and hard-of-hearing individuals, are a specialized category of assistive technologies that facilitate communication and enhance situational awareness by combining a variety of modern technologies, including speech recognition, image recognition, artificial intelligence, and, in certain cases, augmented reality(AR).

The core objective of these smart systems is to significantly enhance the lives of deaf individuals by bridging communication gaps between deaf individuals and the predominantly hearing world and improving accessibility, especially in dynamic real-time settings where delays or misunderstandings can lead to social exclusion or even physical danger. Traditional communication tools, such as lip reading, written text, and sign language interpreters, often fall short in providing instant, autonomous, and inclusive interaction, particularly in public or unpredictable environments. Smart real-time translation devices aim to bridge this gap by providing non-intrusive, wearable solutions that are context-aware and responsive [30].

8.3.1 User-Centered Barriers in Smart Glasses:

The application of smart glasses for real-time speech-to-text translation for deaf and hard-of-hearing users presents some significant challenges. Among the greatest challenges is the utilization of textual output, which may be recalcitrant to the communication mode of numerous users, particularly those for whom sign language is the primary communication method. Studies have indicated that a large proportion of deaf adult's view text-based AR systems as not being synchronized with their communication needs, thereby resulting in reduced usage and increased abandonment[32].

Environmental conditions also affect the functioning of these kinds of devices. Speech systems may function poorly under noisy or dynamic environments and lead to transcription errors of more than 30%, hence undermining the validity of the communication aid [33]. Moreover, the cognitive effort of reading scrolling or time-lagged text on small screens can cause fatigue for the user, rendering it less feasible for widespread use.Technical limitations also hinder mass acceptance. Most smart glasses require persistent internet connectivity and synchronization with smartphones to function effectively, restricting their use in areas where connectivity is poor. Battery capacity constraints also pose problems, as power requirements of real-time processing can cause the need to recharge the device multiple times, disrupting user experience [34].

Expense is a concern; the high-end cost of pricey smart glasses, generally more than $1,100, plus probable subscription costs, places them beyond the reach of most consumers in the absence of sponsorships.

These complex issues point to the need for more affordable, accessible, and simpler-to-use wearable devices that are sensitive to the different needs of the deaf and hard-of-hearing population.

8.4 AI-powered wearables for inclusive communication:

Major limitations in existing assistive technologies clearly demonstrate the necessity for a holistic, context-aware, and real-time solution. A next-generation wearable must focus on not just real-time speech recognition, but also direct speech-to-sign language translation, depending on the language and cultural needs of Deaf individuals. Relative to current smart glasses that use captioning, a more effective solution would be ASL output, offline capability, and solid performance in noisy or unpredictable environments. The device would also need to break through accessibility barriers, such as cost and literacy demands, that still undermine the effectiveness of mainstream solutions [35]. By matching technological possibility with the actual needs and everyday lives of deaf and hard-of-hearing users, AI-enabled smart glasses have the potential to turn assistive communication into an authentically user-driven experience.

9 Conclusion:

Deaf people confront numerous hurdles, including communication barriers, social isolation, educational obstacles, workplace discrimination, and limited access to key services. These difficulties stem from a lack of awareness, inadequate accommodations, and societal structures that frequently prioritize hearing-centric modes of interaction. Addressing these issues requires not only technological innovations, such as assistive devices and communication tools, but also a broader cultural shift toward inclusivity and accessibility. By understanding and addressing these challenges, we can create a fairer society that supports deaf individuals to fully engage in all parts of life. This chapter emphasizes the critical need for solutions especially for those who rely on sign language, paving the way for the development of technologies such as smart glasses to bridge the communication gap.

Chapter 02: Speech Recognition and NLP Technologies

1 Introduction:

In a world where technological advancements are rapidly evolving, artificial intelligence plays a crucial role in enhancing many aspects of daily life. This chapter focuses on the impact of AI on smart glasses designed for speech-to-sign language translation. We will examine the technological foundations of these glasses, highlighting the importance of machine learning and pattern recognition. Additionally, we will explore practical applications of these technologies, such as speech recognition, natural language processing, and their contribution to improving communication accessibility for the deaf and hard-of-hearing community.

2 What is artificial intelligence:

Artificial Intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think and act like humans. It involves the development of algorithms and computer programs that can perform tasks that typically require human intelligence such as visual perception, speech recognition, decision-making, and language translation [36]

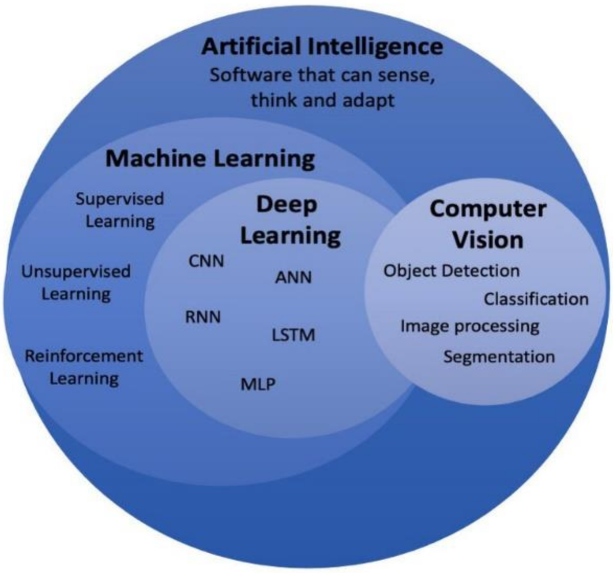

Figure 9: Artificial intelligence and its subfields.

2.1 Machine learning:

Machine learning is the science that enables machines to act without being explicitly programmed, by learning from data and past experiences through mathematical models. Anything that can be stored digitally can serve as data for machine learning (images, signals, text).

The process begins with providing high quality data, which is then used to train our machines (computers) by creating machine learning models based on the data and other methods. The algorithms we use are determined by the type of data we have and the task we are attempting to automate.

Machine learning algorithms are utilized in a wide range of applications, including medicine, email filtering, and computer vision, where developing traditional algorithms to do the required tasks is difficult or impossible [37]

2.2 Types of ML:

- **Supervised learning:** Supervised learning is a type of machine learning where a model is trained on labeled data meaning each input is paired with the correct output. the model learns by comparing its predictions with the actual answers provided in the training data. Over time, it adjusts itself to minimize errors and improve accuracy. The goal of supervised learning is to make accurate predictions when given new, unseen data. For example, if a model is trained to recognize handwritten digits, it will use what it learned to correctly identify new numbers it hasn't seen before.

  Supervised learning can be applied in various forms, including supervised learning classification and supervised learning regression, making it a crucial technique in the field of artificial intelligence and supervised data mining.

  A fundamental concept in supervised machine learning is learning a class from examples. This involves providing the model with examples where the correct label is known, such as learning to classify images of cats and dogs by being shown labeled examples of both. The model then learns the distinguishing features of each class and applies this knowledge to classify new images [38].

- **Unsupervised learning:** Unsupervised learning is a branch of machine learning that deals with unlabeled data. Unlike supervised learning, where the data is labeled with a specific category or outcome, unsupervised learning algorithms are tasked with finding patterns and relationships within the data without any prior knowledge of the data's meaning. Unsupervised machine learning algorithms find hidden patterns and data without any human intervention, i.e., we don't give output to our model. The training model has only input parameter values and discovers the groups or patterns on its own.

  There are two main types of unsupervised learning:

- **Clustering:** Clustering algorithms group similar data points together based on their characteristics. The goal is to identify groups, or clusters, of data points that are similar to each other, while being distinct from other groups.
- **Dimension Dimensionality:** reduction algorithms reduce the number of input variables in a dataset while preserving as much of the original information as possible. This is useful for reducing the complexity of a dataset and making it easier to visualize and analyze [38].
- **Reinforcement machine learning:** In reinforcement learning an agent learns to interact with an environment by performing actions and receiving rewards or penalties based on its actions. The goal of reinforcement learning is to learn a policy, which is a mapping from states to actions, that maximizes the expected cumulative reward over time.

  There are two main types of reinforcement learning:
- **Model-based reinforcement learning:** The agent learns a model of the environment, including the transition probabilities between states and the rewards associated with each state-action pair. The agent then uses this model to plan its actions in order to maximize its expected reward.
- **Model-free reinforcement learning:** The agent learns a policy directly from experience without explicitly building a model of the environment. The agent interacts with the environment and updates its policy based on the rewards it receives [38].

2.3 Artificial neural network:

2.3.1 Definition of ANN:

ANNs are computational models inspired by the human brain, designed to process and analyze complex data. They consist of interconnected artificial neurons that work together to make decisions and recognize patterns.

ANNs contains 3 layers of nodes namely the input layer, single or multiple hidden layers, and output layers. The input layer uses a matrix of the rows and columns which contains values of every data point and is used in the hidden layer for further processing. ANN contains two types of functions namely, Summation Function and Activation Function [39].

2.3.2 Operating principal of an ANN :

The neural network is organized by layers, which can be divided into the input layer, hidden layer and output layer, according to their functions. The number of neurons in each layer should be determined by the input, output and accuracy requirements. The output of a neuron is composed of two parts: one is the linear combination of the values of the previous layer, and the other is the nonlinear transformation of the activation function. The weight (w) and bias (b) parameters in the linear model are continuously updated according to the output error when the network is established [40].

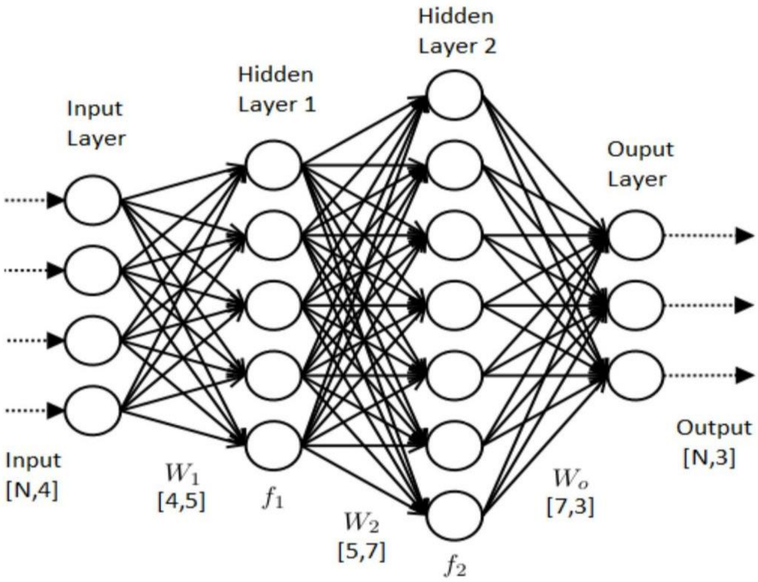

Figure 10: Architecture of ANNs

One update round is referred to as an iteration, and the iterations will not stop until the ending rule is satisfied. In the process of nonlinear transformation, different activation functions can be selected to adapt to the different data structures.

2.4 Deep learning:

2.4.1 Definition:

Deep Learning is a branch of Artificial Intelligence (AI) that enables machines to learn from large amounts of data.

Deep Learning is transforming the way machines understand, learn, and interact with complex data. Deep learning mimics neural networks of the human brain, it enables computers to autonomously uncover patterns and make informed decisions from vast amounts of unstructured data.

It uses neural networks with many layers to automatically find patterns and make predictions. very useful for tasks like image recognition, language translation, and speech processing.

Deep learning models learn directly from data, without the need for manual feature extraction [41].

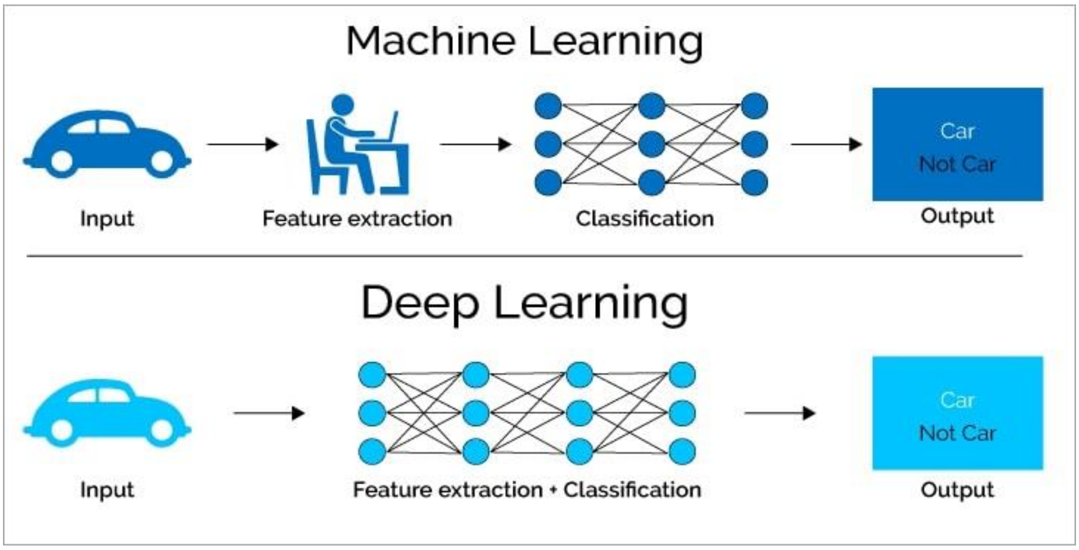

Figure 11: Machine learning VS Deep learning

2.4.2 Application of deep learning:

- **Language translation:** Deep learning models can translate text from one language to another, making it possible to communicate with people from different linguistic backgrounds
- **Object detection and recognition:** Deep learning models are used to identify and locate objects within images and videos, making it possible for machines to perform tasks such as self-driving cars, surveillance, and robotics.
- **Image classification:** Deep learning models can be used to classify images into categories such as animals, plants, and buildings. This is used in applications such as medical imaging, quality control, and image retrieval.
- **Image segmentation:** Deep learning models can be used for image segmentation into different regions, making it possible to identify specific features within images.
- **Automatic Text Generation:** Deep learning model can learn the corpus of text and new text like summaries, essays can be automatically generated using these trained models.
- **Sentiments analysis:** Deep learning models can analyze the sentiment of a piece of text, making it possible to determine whether the text is positive, negative, or neutral.
- **Speech recognition:** Deep learning models can recognize and transcribe spoken words, making it possible to perform tasks such as speech-to-text conversion, voice search, and voice-controlled devices.
- **Medical Diagnostics:** Detecting diseases from X-rays, MRIs, and other medical scans.
- **Recommendation Systems:** Personalizing suggestions for movies, music, and shopping.

- **Autonomous Vehicles:** Enabling self-driving cars to recognize objects and make driving decisions.
- **Fraud Detection:** Identifying unusual patterns in financial transactions and preventing fraud.
- **Gaming:** Enhancing AI in games and creating realistic environments in virtual reality.
- **Predictive Analytics:** Forecasting customer behavior, stock prices, and weather trends.
- **Generative Models:** Creating realistic images, deepfake videos, and AI-generated art.
- **Robotics:** Automating industrial tasks and powering intelligent drones.
- **Customer Support:** Enhancing chatbots for instant and intelligent customer interactions.

2.5 Speech recognition:

2.5.1 Definition:

Speech recognition refers to the ability of machines or computers to identify and interpret spoken language, converting it into a textual or command format. It is a technology that enables humans to interact with devices or systems using voice commands, allowing for hands-free control and efficient communication. This technology has evolved significantly over the years, moving from basic voice recognition to advanced systems capable of understanding continuous speech and handling various accents and languages.

Speech recognition systems are designed to process audio signals, analyze them, and generate textual transcriptions or execute specific commands based on the input. These systems rely on probabilistic models, such as acoustic models and language models, to create robust representations of human speech and language structures [42].

2.5.2 Fundamental concepts of speech recognition:

The fundamental concepts of speech recognition involve converting speech signals into text or commands through identification and understanding processes. This technology typically consists of several modules, including front-end processing, recognition, and post-processing.

- **Speech signal and preprocessing:**The speech signal is a complex audio signal that includes the spoken words, pauses, and background noise. Preprocessing is a critical step in speech recognition, as it reduces the amount of processing required in later stages. This involves blocking the speech samples into frames and extracting unique patterns from each frame. Preprocessing helps in enhancing the quality of the speech signal and preparing it for further [43].
- **Acoustic model:** The acoustic model is a probabilistic model that represents the relationship between speech sounds and their acoustic features. It is trained on a large dataset of speech

samples and their corresponding transcriptions. The acoustic model is responsible for identifying the sequence of sounds (phonemes) in the speech signal [44].

- **Language model:** The language model is another probabilistic model that predicts the likelihood of a sequence of words occurring in a language. It is trained on large text corpora and is used to constrain the possible interpretations of the speech signal, ensuring that the recognized text is grammatically correct and contextually appropriate [44].
- **Lexical model:** The lexical model, also known as the pronunciation model, contains information about the pronunciation of words in a language. It maps words to their phonemic representations, which are then used by the acoustic model to match the speech signal [44].
- **Decoder:** The decoder is the component responsible for combining the outputs of the acoustic and language models to generate the final transcription. It searches for the most likely sequence of words that match the speech signal, considering both acoustic and linguistic constraints [44].

2.5.3 Architecture of speech recognition systems:

The architecture of speech recognition systems typically consists of several key components, each performing a specific function in the recognition process. These components work together to convert the speech signal into a textual representation or execute commands based on the input.

- **Speech signal acquisition:** The first step in the architecture is the acquisition of the speech signal. This is typically done using a microphone, which captures the audio input from the user.
- **Preprocessing:** The acquired speech signal is then preprocessed to enhance its quality and prepare it for further analysis. This step may include Noise reduction, normalization, and segmentation of audio into frames and feature extraction
  Techniques like Fourier Transform or Mel-Frequency Cepstral Coefficients (MFCCs) extract spectral features that represent speech characteristics.
- **Feature extraction:** Feature extraction is the process of converting the speech signal into a set of acoustic features that can be analyzed by the acoustic model. Common features include Mel-Frequency Cepstral Coefficients (MFCCs) and spectrograms.

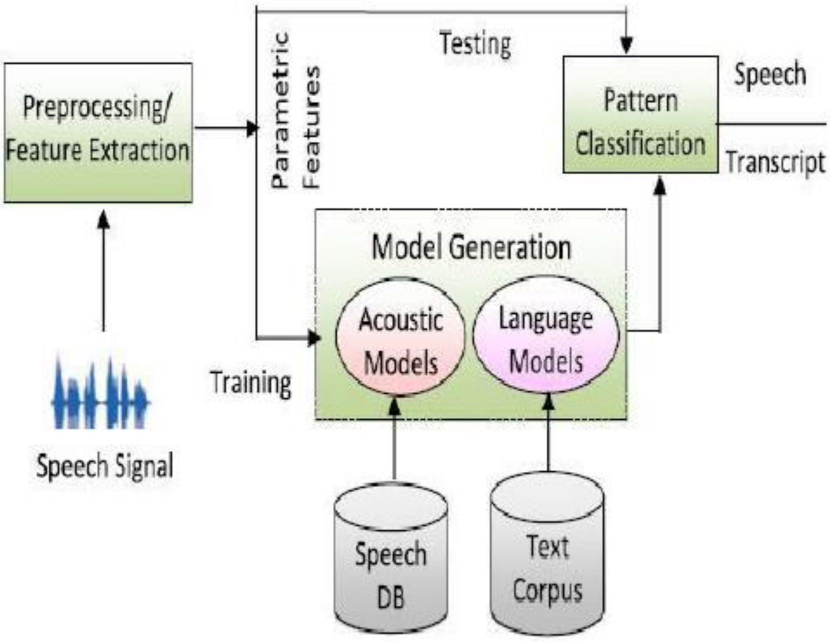

Figure 12: Architecture of speech recognition systems

- **Acoustic model:** The acoustic model processes the extracted features to identify the sequence of phonemes in the speech signal. This model is typically implemented using techniques such as Hidden Markov Models (HMMs) with Gaussian Mixture Models (GMMs) or like Deep Neural Networks (DNNs), transformers like Whisper.
- **Language model:** The language model processes the sequence of phonemes generated by the acoustic model to predict the most likely sequence of words. This model is typically implemented using n-gram models or recurrent neural networks (RNNs), Transformer-based models (e.g., BERT for contextual understanding).
- **Decoder:** The decoder combines the outputs of the acoustic and language models to generate the final transcription. It searches for the most likely sequence of words that match the speech signal, considering both acoustic and linguistic constraints.
- **Postprocessing:** The final step in the architecture is postprocessing, which may include spell-checking, grammar-checking, and formatting the output to ensure it is accurate and readable [45].

2.5.4 Applications and use cases of speech recognition:

Speech recognition technology has a wide range of applications across various industries, from consumer electronics to healthcare and education. Some of the most common applications and use cases include:

- **Voice assistants:** Voice assistants, such as Siri, Alexa, and Google Assistant, use speech recognition to understand and execute voice commands. These assistants can perform tasks such as setting reminders, playing music, and answering questions.
- **Speech-to-text systems:** Speech-to-text systems convert spoken words into written text, allowing users to dictate documents, emails, and messages. This is particularly useful for individuals with disabilities or those who prefer hands-free typing.
- **Voice-controlled devices:** Speech recognition is used in voice-controlled devices such as smart speakers, televisions, and home automation systems. These devices allow users to control their surroundings using voice commands.
- **Transcription services:** Speech recognition is used in transcription services to convert audio and video recordings into written text. This is commonly used in industries such as media, education, and law enforcement.
- **Language learning:** Speech recognition technology is used in language learning applications to help users improve their pronunciation and speaking skills. These applications can provide real-time feedback on pronunciation and intonation.
- **Medical applications:** In the healthcare industry, speech recognition is used for medical transcription, voice-controlled medical devices, and patient-doctor communication. It allows healthcare professionals to dictate notes and commands without the need for manual input.
- **Customer service:** Speech recognition is used in customer service applications such as interactive voice response (IVR) systems. These systems allow customers to interact with automated agents using voice commands, reducing the need for human intervention.

2.6 Natural language processing:

2.6.1 Definition:

Natural language processing is a subfield of Artificial Intelligence (AI), which assists computers and machines to understand human language. It enables computers to interpret and manipulate human language in order to automate the tasks between machines and humans.

It evolved from computational linguistics, which uses computer science to understand the principles of language, but rather than developing theoretical frameworks, NLP is an engineering discipline that seeks to build technology to accomplish useful tasks. NLP can be divided into two overlapping subfields: natural language understanding (NLU), which focuses on semantic analysis or determining the intended meaning of text, and natural language generation (NLG), which focuses on text generation by a machine [46].

2.6.2 How NLP works:

NLP models work by finding relationships between the constituent parts of language for example, the letters, words, and sentences found in a text dataset. NLP architectures use various methods for data preprocessing, feature extraction, and modeling. Some of these processes are

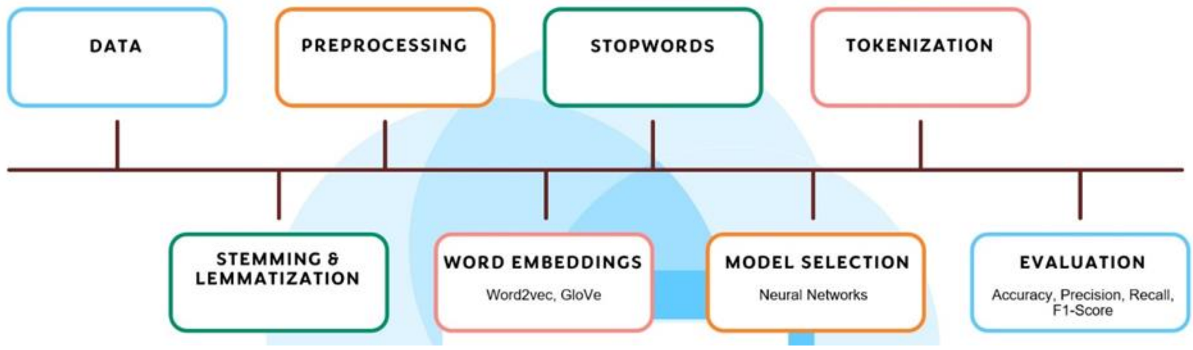

Figure 13: Building NLP model.

- **Data preprocessing:** It is frequently necessary to preprocess text before a model processes it for a particular job in order to enhance model performance or to convert words and characters into a format that the model can comprehend. Data preparation is given priority in the emerging field of data-centric AI.

  There are several methods that can be applied to this data preprocessing:

- **Tokenization :** Machines find it challenging to comprehend the semantics and context of a sentence or paragraph. Tokenization is the process of breaking down words or sentences into semantic units known as 'tokens'. This enables us to deal with shorter chunks of text, which can be easily understood even when read independently from the context of the text. Fig. 6 explains the process of tokenization [47].

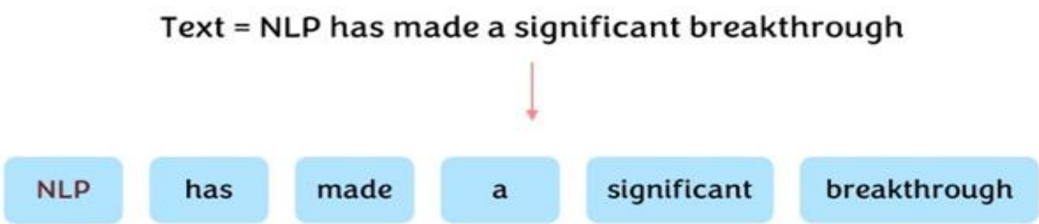

Figure 14: The process of tokenization.

Stop Words Removal: Stop Words Removal is a text pre-processing technique that eliminates words with zero significance so that the model can focus on meaningful words. Eliminating such words does not change the meaning or the context of a sentence. While they must be eliminated to avoid junk space in the database and consequently enhance the computation time.

Stop words include - the, a, I, an, be, you, had, has, were, was, will, does, is - and many more. Fig. illustrates the elimination of stop words [47].

Figure 15: Stop words removal.

Stemming & Lemmatization: It is the process of analyzing variation of words to reveal their origin or root word. This pre-processing technique helps in understanding the context of a sentence.Stemming reduces the words to their root form. It eliminates duplicate words with identical relevance to ensure tasks are performed more accurately. It works by removing the last letters of the word. Fig. 8 illustrates the working of stemming by removing the last letters of a word.Lemmatization also reduces the words to their original form, but they perform analysis in a group of words to find its root word, which is also known as 'Lemma'. Lemma is a word that represents a group of words ('recover' as in Fig. 8). This group of ords is called 'lexeme' (recover, recovering, recovers, recovered as in Fig. 8) [47].

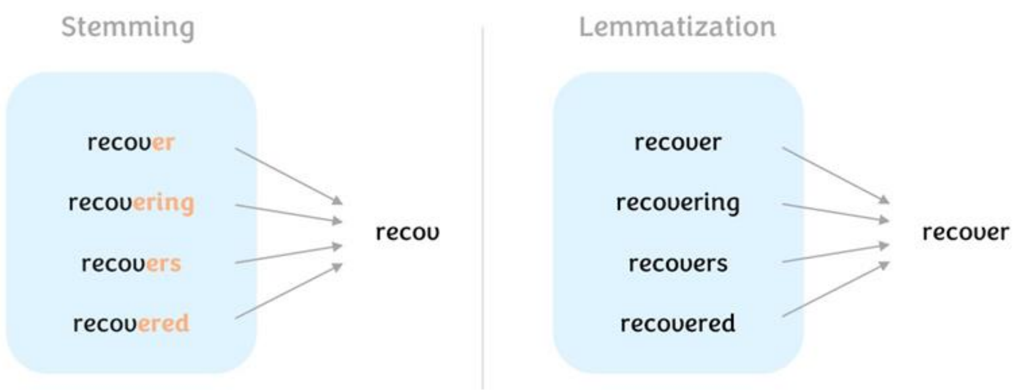

Figure 16: Stemming and lemmatization.

- **Word Embeddings:** Word Embeddings is an NLP technique where words are converted into numbers or vectors. To convert text into numerical format, significant techniques are required which are called Vectorization or Word Embeddings in NLP. There are various word embedding methodologies, which include TF-IDF (Term Frequency Inverse Document Frequency), BOW (Bag-of-Words), Count Vectorization, N-grams Vectorization, Word2Vec and GloVe.

- **Word2Vec :**The process of ‘word embeddings’ involves turning each word into a numerical representation of the word (a vector).

  Each word is converted into a single vector, which is then trained in a manner resembling a neural network.

  Based on a word's usage in the text, Word2Vec can infer a word's meaning with a high degree of accuracy.

- **GloVe:** GloVe represents a global vector, which is an unsupervised learning technique that generates word vector representations.

  The advantage of GloVe is that it integrates global statistics to build word vectors, while Word2Vec depends on local statistics (local context knowledge about words) for generation of word vectors.

  It is based on word-context matrix factorization algorithms. In GloVe, a co-occurrence matrix is used to determine the semantic relationship between words [47].

- **NLP Model Selection and Architecture Design:** Model Selection is the process of selecting an appropriate machine learning model or deep learning model for an NLP application. The dataset will be divided into training and testing sections, based on which the model will be trained and then tested.
- **Training NLP Models :** NLP pre-trained models are useful for NLP tasks like translating text, predicting missing parts of a sentence or even generating new sentences. NLP pre-trained models can be used in many NLP applications such as chatbots and NLP API, etc.
- **Performance Evaluation Metrics :** To evaluate and interpret natural language data, such as text or speech, NLP requires an integration of machine learning techniques. The most common evaluation metrics include accuracy, precision, recall and f1-score [47].

2.6.3 Natural language processing techniques:

Most of the NLP tasks discussed above can be modeled by a dozen or so general techniques. It’s helpful to think of these techniques in two categories: Traditional machine learning methods and deep learning methods.

Traditional Machine learning NLP techniques:

- **Logistic regression:** is a supervised classification algorithm that aims to predict the probability that an event will occur based on some input. In NLP, logistic regression models can be applied to solve problems such as sentiment analysis, spam detection, and toxicity classification.
- **Naive bayes:** is a supervised classification algorithm that finds the conditional probability distribution P(label | text).

  And predicts based on which joint distribution has the highest probability. The naive assumption in the Naive Bayes model is that the individual words are independent.

  In NLP, such statistical methods can be applied to solve problems such as spam detection or finding bugs in software code.
- **Decision trees:** are a class of supervised classification models that split the dataset based on different features to maximize information gain in those splits.
- **Latent Dirichlet Allocation (LDA):** is used for topic modeling. LDA tries to view a document as a collection of topics and a topic as a collection of words. LDA is a statistical approach. The intuition behind it is that we can describe any topic using only a small set of words from the corpus [46].

  Deep learning NLP Techniques:
- **Convolutional Neural Network (CNN):** The idea of using a CNN to classify text was first presented in the paper "Convolutional Neural Network for Sentence Classifcation" by Yoon Kim. The central intuition is to see a document as an image. However, instead of pixels, the input is sentences or documents represented as a matrix of words [46].

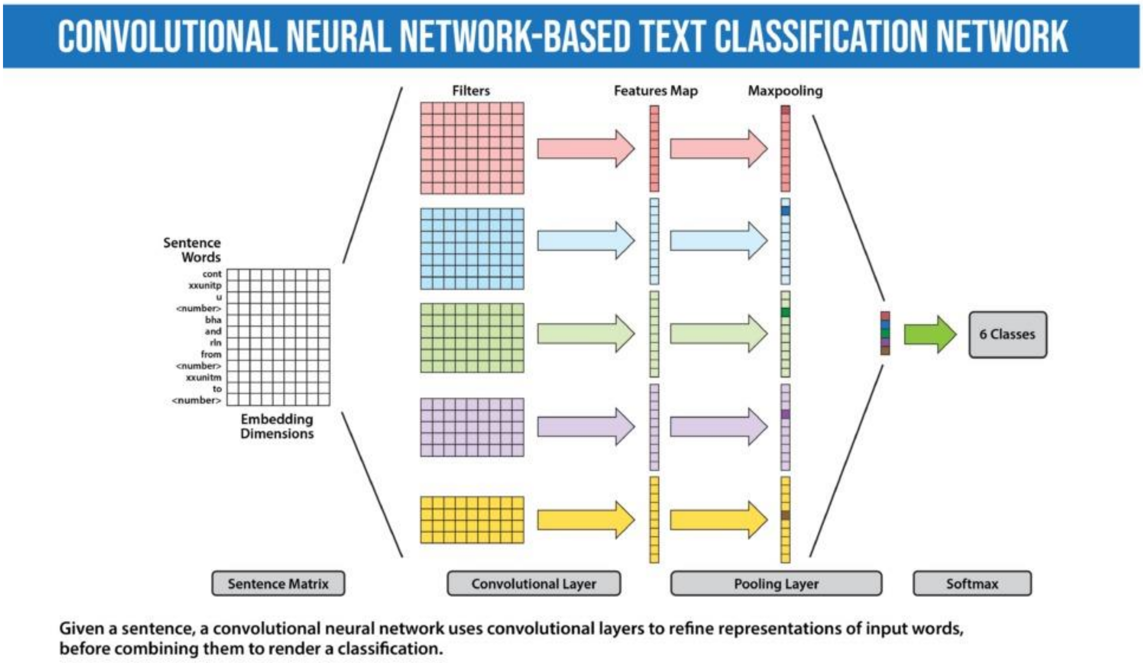

Figure 17: Convolutional neural networks for sentence classification.

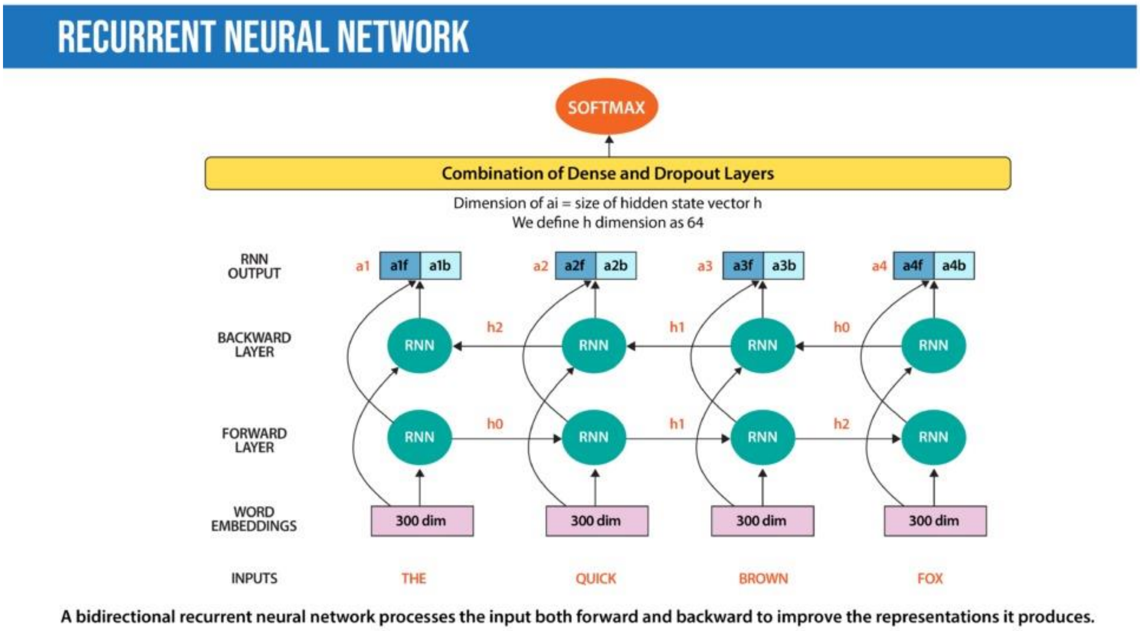

Figure 18: Recurrent neural network.

- **Recurrent Neural Network (RNN):** Many techniques for text classification that use deep learning process words in close proximity using n-grams or a window (CNNs). They can see "New York" as a single instance. However, they can't capture the context provided by a particular text sequence. They don't learn the sequential structure of the data, where every word is dependent on the previous word or a word in the previous sentence [46].

RNNs remember previous information using hidden states and connect it to the current task. The architectures known as Gated Recurrent Unit (GRU) and long short-term memory (LSTM) are types of RNNs designed to remember information for an extended period. Moreover, the bidirectional LSTM/GRU keeps contextual information in both directions, which is helpful in text classification. RNNs have also been used to generate mathematical proofs and translate human thoughts into words.

- **Autoencoders:** are deep learning encoder-decoders that approximate a mapping from X to X, i.e., input=output. They first compress the input features into a lower-dimensional representation (sometimes called a latent code, latent vector, or latent representation) and learn to reconstruct the input. The representation vector can be used as input to a separate model, so this technique can be used for dimensionality reduction. Among specialists in many other fields, geneticists have applied autoencoders to spot mutations associated with diseases in amino acid sequences [46].

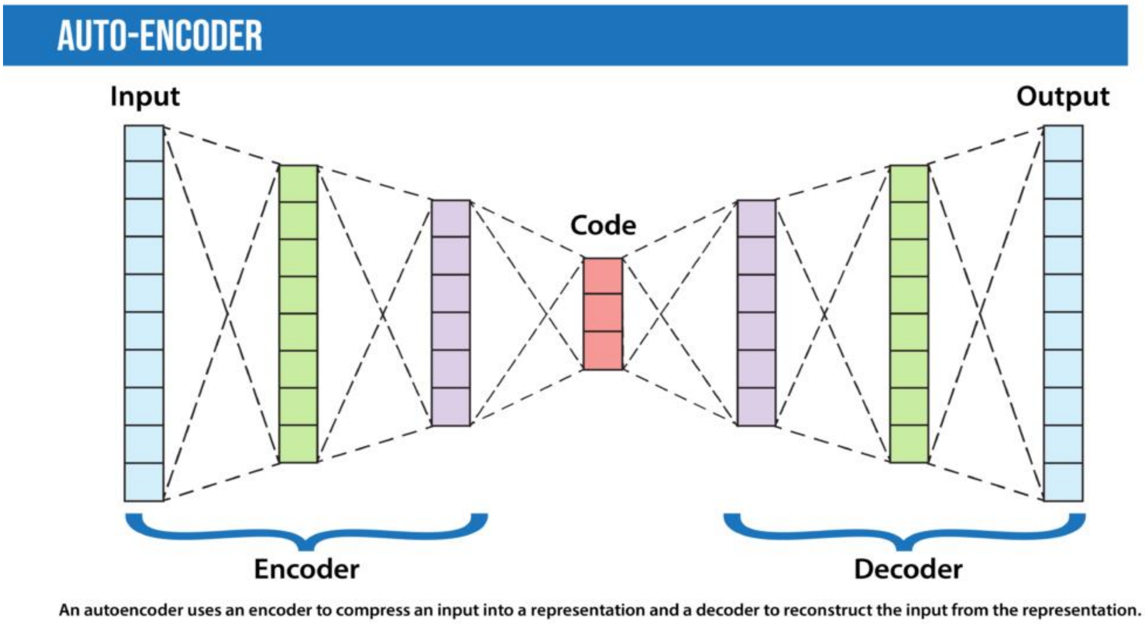

Figure 19: Auto-Encoder.

- **Transformers:** The transformer, a model architecture first described in the 2017 paper "*Attention Is All You Need*" (Vaswani, Shazeer, Parmar, et al.), forgoes recurrence and instead relies entirely on a self-attention mechanism to draw global dependencies between input and output. Since this mechanism processes all words at once (instead of one at a time) that decreases training speed and inference cost compared to RNNs, especially since it is parallelizable. The transformer architecture has revolutionized NLP in recent years, leading to models including Bloom, jurassic_X, and turning-NLG It has also been successfully applied to a variety of different vision taskes, including making 3D images [46].

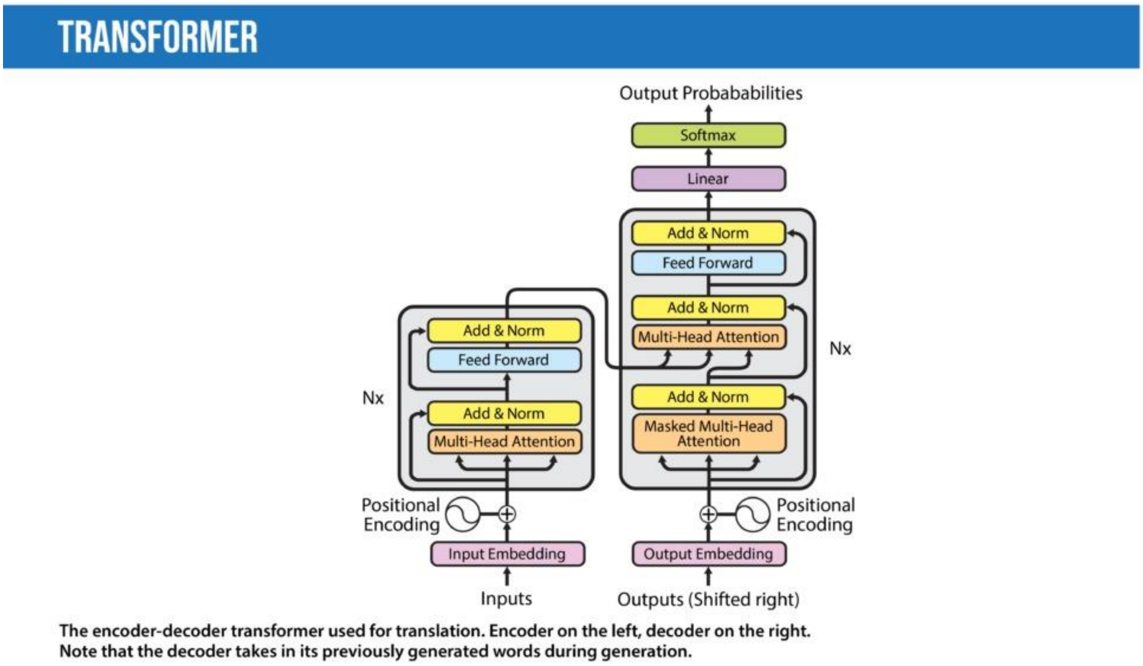

Figure 20: Transformer.

2.6.4 Applications and use cases of NLP:

NLP is a fairly generic term that covers a very wide range of applications. Here are the most popular applications:

- **Sentiment analysis**: Sentiment analysis is the process of classifying the emotional intent of text. Generally, the input to a sentiment classification model is a piece of text, and the output is the probability that the sentiment expressed is positive, negative, or neutral. Typically, this probability is based on either hand-generated features, word n-grams, TF-IDF features, or using deep learning models to capture sequential long- and short-term dependencies. Sentiment analysis is used to classify customer reviews on various online platforms as well as for niche applications like identifying signs of mental illness in online comments.
- **Machine translation:** The development of machine translation algorithms has truly revolutionized the way texts are translated today. Applications, such as Google Translate, are able to translate entire texts without any human intervention.

  Because natural language is inherently ambiguous and variable, these applications do not rely on word-for-word replacement, but require true text analysis and modeling, known as Statistical Machine Translation.

- **Text generation:** more formally known as natural language generation (NLG), produces text that's similar to human-written text. Such models can be fine-tuned to produce text in different genres and formats including tweets, blogs, and even computer code. Text generation has been performed using Markov processes, LSTMs, BERT, GPT-2, and other approaches. It is particularly useful for autocomplete and chatbots.
- **Autocomplete:** predicts what word comes next and autocomplete systems of varying complexity are used in chat applications like WhatsApp. Google uses autocomplete to predict search queries. One of the most famous models for autocomplete is GPT-2, which has been used to write articles, and much more.
- **Chatbots:** NLP methods are at the heart of how today's chatbots work. While these systems are not completely perfect, they can now easily handle standard tasks such as informing customers about products or services, answering their questions, etc. They are used across multiple channels, including the Internet, applications and messaging platforms. The opening of the Facebook Messenger platform to chatbots in 2016 contributed to their development.
- **Marketing:** Marketers also use NLP to find people who are likely to make a purchase.

  They rely on the behavior of Internet users on websites, social networks and search engine queries. This type of analysis allows Google to generate a significant profit by offering the right advertisement to the right people. Each time a visitor clicks on an ad, the advertiser pays up to 50 dollars!

  More generally, NLP methods can be used to build a rich and comprehensive picture of a company's existing market, customers, issues, competition, and growth potential for new products and services.

2.6.5 Opportunities and challenges and of NLP:

The integration of NLP technologies can enhance efficiency, improve decision-making, and facilitate better human-computer interactions. However, issues such as algorithmic bias, data privacy, and the need for contextual understanding remain prevalent. Below are key aspects of the opportunities and challenges associated with NLP.

- **Opportunities in NLP:** Automation in Recruitment: NLP streamlines processes like resume parsing and candidate matching, significantly reducing time-to-hire.

  Healthcare Applications: NLP aids in diagnosing conditions, extracting information from clinical notes, and identifying biomarkers in oral cancer research.

Generative AI Development: Advances in large language models (LLMs) enhance NLP capabilities, leading to improved applications in education and healthcare.

- **Challenges in NLP:** Algorithmic Bias: NLP systems can perpetuate biases present in training data, leading to unfair outcomes in applications like recruitment and content moderation.
- **Data Privacy Concerns:** The handling of sensitive information raises ethical issues, particularly in healthcare settings.

  Contextual Understanding: NLP struggles with nuances in language, such as cultural differences and contextual meanings, which can hinder its effectiveness.

3 Conclusion:

This chapter has explored how artificial intelligence is transforming communication through smart glasses for speech-to-sign language translation. By leveraging machine learning and computer vision, these devices enable real-time visual communication, offering new possibilities for accessibility. While significant progress has been made, challenges remain in improving accuracy and expanding accessibility. Nevertheless, these technological advances represent an important step toward a more inclusive future where communication barriers can be overcome. The continued development of such assistive technologies holds promise for empowering the deaf and hard-of-hearing community through innovative AI solutions.

Chapter 03: System design and implementation.

1 Introduction:

The development of innovative technological solutions is crucial to improving communication accessibility for deaf and hard-of-hearing individuals. Our smart glasses project introduces an intelligent wearable system that bridges communication gaps through real-time speech-to-sign language translation. This chapter presents the conceptual framework, technical implementation, designed to empower users through cutting-edge AI and ergonomic hardware integration.

2 General presentation of the project:

Deaf individuals face significant challenges in communicating with hearing individuals who do not understand sign language, often leading to social isolation and limited access to information in real-time conversations. Traditional assistive technologies, such as text-based communication aids, are often slow, cumbersome, or fail to convey the nuances of sign language, leaving a gap in effective solutions. This underscores the need for innovative technologies that enable seamless, real-time communication to enhance the quality of life and social integration of deaf individuals.

On the other hand, existing wearable translation devices, while promising, often suffer from limitations such as high costs, complex interfaces, or insufficient accuracy in translating spoken language into expressive sign language gestures. These shortcomings highlight the demand for a user-friendly, efficient, and accessible solution tailored to the unique needs of the deaf community.

Our contribution lies in the design and development of smart glasses that translate spoken language into sign language in real time, addressing the following key requirements:

- **Real-Time Speech-to-Sign Translation:** Convert spoken language into accurate and expressive sign language gestures displayed through an integrated AMOLED screen with prism projection for optimal visibility with minimal latency for natural communication and fluid conversations.
- **Intuitive Display:** Present sign language through an animated avatar projected via a prism, ensuring clear and unobtrusive viewing.
- **User-Friendly Operation:** Enable simple activation through a single button press, making the system accessible to users of all technical abilities.

- **High Accuracy and Expressiveness:** Ensure precise transcription of speech and detailed representation of sign language, including hand gestures, body movements, and facial expressions.
- **Seamless Integration:** The lightweight glasses and prism-based display resemble ordinary eyewear, ensuring comfort and discretion in daily use.
- **Affordability and Scalability:** Develop a cost-effective solution that can be enhanced with new features through software and hardware updates and adapted to various sign languages. The system aims for global accessibility while maintaining strict user privacy standards.
- **Enhanced Social Inclusion:** Facilitate natural interactions between deaf and hearing individuals, reducing communication barriers.

3 Functioning principle:

Our project represents a groundbreaking advancement in assistive communication technology for deaf individuals. The system leverages wearable hardware and artificial intelligence to translate spoken language into expressive sign language gestures in real time, enabling seamless interaction between deaf and hearing individuals. The functioning principle is as follows:

The process begins when the user presses a button on the glasses, activating a wireless earphone's microphone to capture spoken input, which is sent via Bluetooth to an Android mobile application. The application employs an advanced AI model to transcribe the audio into text, providing a solid foundation for translation., This text is then transmitted to a server via an API call, where another AI model processes it to generate dynamic sign language there are transformed into a sequence of 3D avatar animations. These animations are subsequently sent back to the mobile application. These animations wirelessly transmit them via Wi-Fi to an ESP32 modulo integrated with the glasses. The ESP32 equipped with AMOLED screen, project the gestures of the avatar via a prism into the user's field of view for clear, intuitive viewing, enabling seamless and expressive communication.

4 Practical realization:

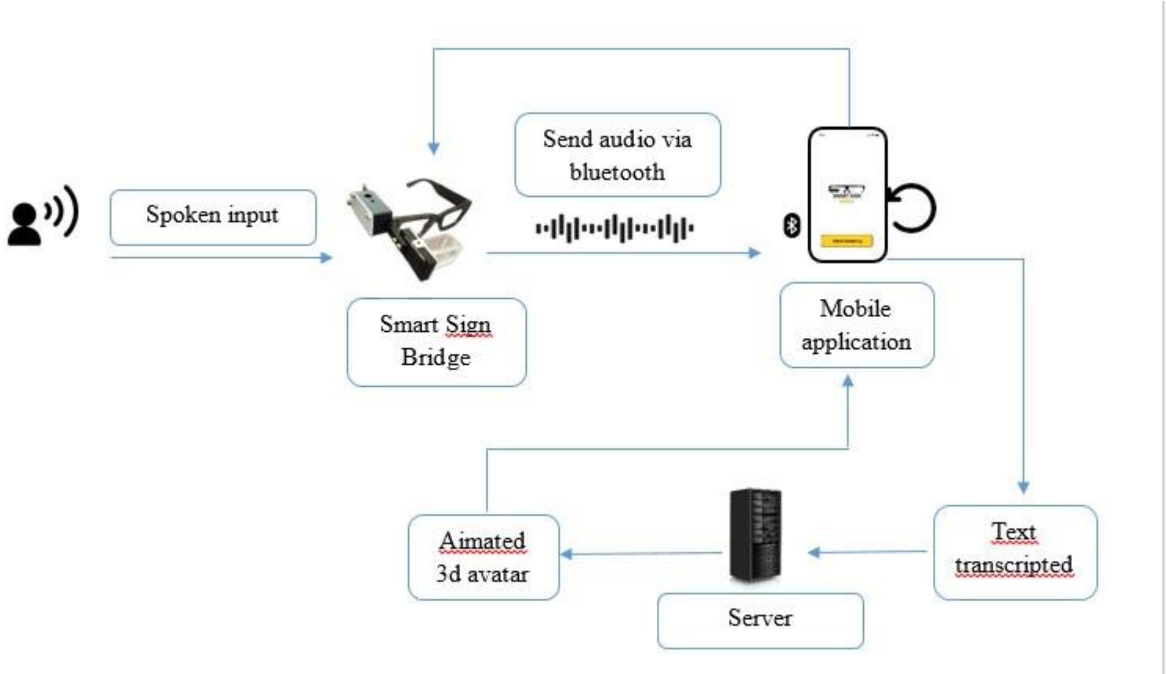

Figure 21: Final implemented circuit.

5 System architecture:

5.1 Flowchart:

The following flowchart explains the various features built into our smart glasses and how they work together to give users total support.

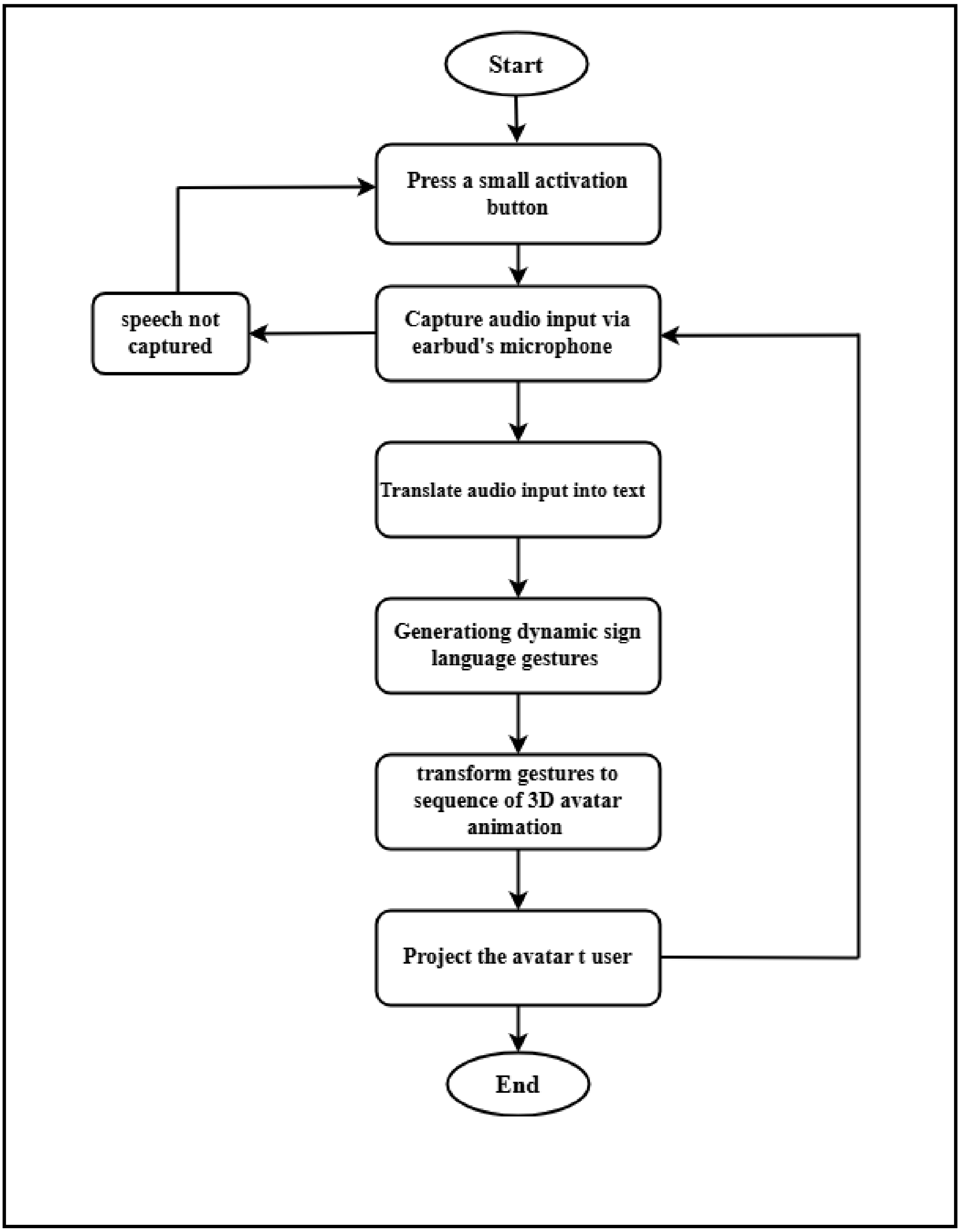

Figure 22: Functioning flowchart.

5.2 Artificial intelligence system:

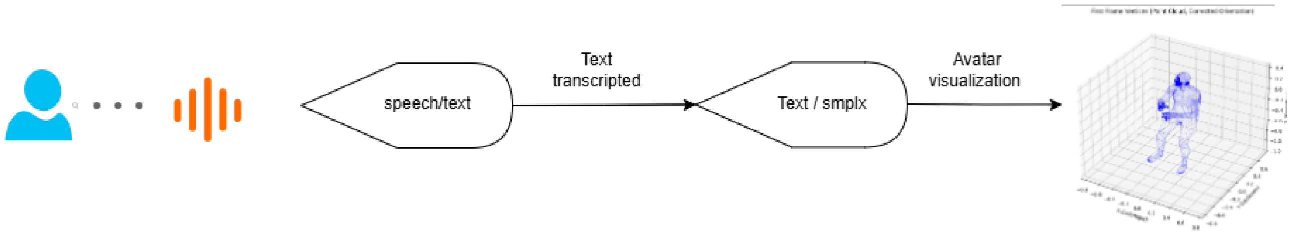

Figure 23: AI system

The AI model embedded within the smart glasses is a two-pipeline setup for real-time speech-to-sign language interpretation and 3D avatar animation, enabling smooth and accessible communication for deaf and hard-of-hearing users who rely on sign language. In the first stage, the speech-to-text (STT) phase, which utilizes the pre-trained model to translate audio input captured using a microphone into accurate transcriptions of text. It employs the model's robust speech recognition to handle a variety of accents and noise in the background to provide a sound textual foundation. The output is fed into the second stage of the system, which offers context awareness and motion naturalness of the sign gestures. In the second stage, the system employs the usage of a text encoder to analyze the transcribed text to yield semantic context and SL specific patterns of language and translate them to a sequence of gestures. There is then a motion generation module that translates these instructions into smplx parameters with an added context-aware mechanism to facilitate coherence between gestures and enhance naturalness while transitioning between movements. The architecture progresses to the 3D avatar rendering module that employs the SMPL-X body model to convert the generated parameters to 3D mesh data, portraying the movement of the avatar as a sequence of frames (point clouds) prior to exporting the animation in the form of a video. This modular and scalable process seamlessly combines speech recognition, context-dependent text processing, motion synthesis, and 3D rendering to provide a robust basis for the speech-to-3D SL avatar system.

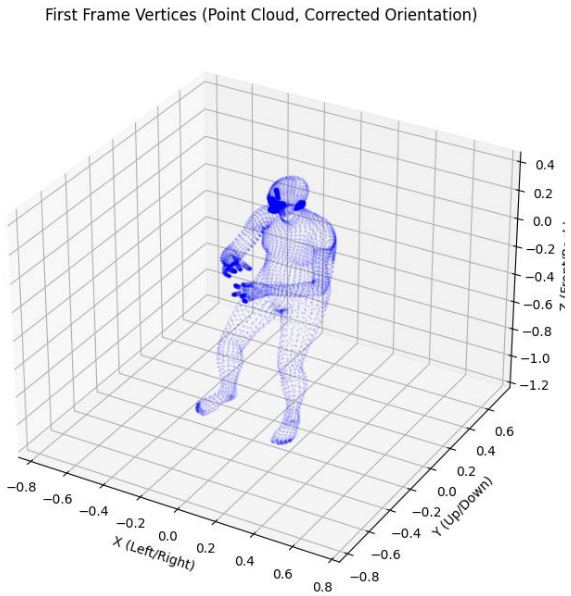

Figure 24: 3D Avatar pose.

In the current prototype, the AI system outputs sign language gestures based on American sign language (SL), this choice was made primarily due to the availability of large-scale, annotated datasets that are crucial for training and validating the model, especially for motion generation using 3D avatars. And the publicly accessible resources, making it a practical starting point for developing and evaluating the technical feasibility of the system. The use of ASL in the prototype is considered a contextual and methodological compromise, which will be addressed in the localization phase of the project.

In addition, the system design is also modular and extensible, and it allows future upgrades in language models, gesture precision, and avatar reality, thereby a scalable and sustainable solution to assistive technology.

5.3 Dataset:

To train the second phase of the model (text-to-ASL) of the smart glasses system, we utilized the ASL subset of the SignAvatars dataset a large-scale, high-quality dataset generated for sign language animation with virtual avatars. The ASL subset offers thousands of annotated American sign language video sequences with aligned corresponding SMPL-X pose parameters that capture detailed full-body movement, hand articulation, and facial expressions. These pose annotations directly map natural language to 3D avatar animation, enabling realistic sign language synthesis. The dataset supports both isolated signs and continuous signing and can be applied in conversational contexts rather than just vocabulary-level translation.

Data	Video	Frame	Duration (hours)	co-articulated	Pose Annotation	Signer
RWTH-Phoenix-2014T (Camgoz et al., 2018)	8.25K	0.94M	11	C	-	9
DGS Corpus (Hanke et al., 2020)	-	-	50	C	2D keypoints	327
BSL Corpus (Schembri et al., 2013)	-	-	125	C	-	249
MS-ASL (Joze & Koller, 2018)	25K	-	25	I	-	222
WL-ASL (Li et al., 2020)	21K	1.39M	14	I	2D keypoints	119
How2Sign (Duarte et al., 2021)	34K	5.7M	79	C	2D keypoints, depth*	11
CSL-Daily (Huang et al., 2018)	21K	-	23	C	2D keypoints, depth	10
SIGNUM (Von Agris et al., 2008)	33K	-	55	C	-	25
AUTSL (Sincan & Keles, 2020)	38K	-	21	I	depth	43
Forte et al. (2023)	0.05K	4K	-	I	body mesh vertices*	-
SignAvatars (Ours)	70K	8.34M	117	Both	SMPL-X, MANO, 2D&3D keypoints	153

Tableau 2: Dataset's modality.

Each sample in the dataset pairs transcribed ASL English sentences with temporally aligned SMPL-X poses, enabling the training of the model for dynamic sign gesture translation from natural language. As a preprocessing step, pose sequences were normalized, time-aligned, and split into fixed-length windows for efficient and stable training. Text annotations were tokenized and input to the encoder, and the decoder produced the corresponding motion parameters. Together, this enabled the model to learn expressive, context-dependent gesture synthesis. The quality, range, and linguistic diversity of the subset of ASL in the SignAvatars dataset were instrumental in generating smooth, accurate, and comprehensible sign language animations for the 3D avatar, hence rendering the assistive system more effective.

6 Evolutive perspectives:

As a prototype, Smart Sign Glasses could demonstrate the feasibility of real-time speech-to-ASL translation using AI and 3D avatar animation. However, to address more effectively the diverse needs of the global deaf and hard-of-hearing community, several upgrades are planned in the future that both augment the software functionality as well as the hardware setup of the system.

- **Software enhancements:** From a software perspective, the current system only enables English spoken input and provides outputs in American Sign Language (ASL). Future enhancement of the solution will see the inclusion of several spoken languages, including Arabic and French, to make it accessible to a larger population of users. We also plan to integrate regional sign languages, including Algerian Sign Language (LSA) and other forms of dialects. This linguistic inclusivity is especially important in multicultural societies, where dialect and language are closely connected to identity and access.

Additionally, the smart glasses and the mobile app will be enhanced to have personalization features through which users will be able to customize the 3D avatar with respect to style, appearance, and cultural context.

- **Hardware enhancements:** On the hardware side, the current prototype is made up of a LilyGO T-Display S3 microcontroller that has a 1.91-inch AMOLED display which runs off a lithium battery, all contained in a simple, functional enclosure. While this setup was adequate for proof of concept in showing the underlying functionality, future iterations will aim to provide a more advanced, wearable solution. We plan to explore the use of high-resolution micro displays, such as used in professional augmented reality (AR) glasses, to facilitate better visibility, comfort, and ease of use in various lighting conditions. The housing will be further made ergonomic and aesthetically pleasing to increase user comfort on extended use. With greater capacity and longer life batteries, extended daily use will be possible.

  In the end, these innovations showcase our commitment to advancing our smart glasses from a working prototype to an inclusive, scalable, and user-centered assistive technology. Through technical innovation combined with thoughtful design, we aspire to empower users to have more natural, accessible, and dignified communication experiences regardless of language, location, or physical ability.

7 Implementation:

To produce our Smart sign Bridge, a careful selection of materials and software tools was made to meet functional requirements and performance standards. This section describes the essential physical components as well as the software development environments chosen for this innovative project.

7.1 Hardware components:

To respond to the functions assigned to our smart glasses, we began by defining the various work tools that we will require.

7.1.1 Glasses frame:

The glasses frame constitutes the main structure of the smart glasses. holding all the components in a lightweight and ergonomic shape, chosen to be comfortable for wear for a prolonged period and to position the display in the peripheral field of view of the user.

Figure 25: The glasses frame.

7.1.2 Display:

The LILYGO T-Display-S3 AMOLED integrates a 1.91-inch RM67162 IPS AMOLED display with the ESP32-S3R8 microcontroller, providing a high-end platform for the rendering of real-time sign language translation in our wearable smart glasses solution.

Figure 26: LILYGO T-Display-S3 AMOLED.

- **Board specifications:**

MCU	ESP32-S3R8 Dual-core LX7 microprocessor
Wireless Connectivity	2.4 GHz Wi-Fi & Bluetooth5 (LE)
Development	Arduino, PlatformIO-IDE, Micropython
Flash	16MB
PSRAM	8MB
Bat Voltage Detection	IO04
Onboard functions	Boot (IO00) + Reset + IO21 Button

Tableau 3: Board specifications [48].

- **1.91 inch RM67162 IPS AMOLED :**

Resolution	240 X RGB X 536(H)
Interface	QSPI
Active Area	19.8*44.22mm
Driver IC	RM67162
Viewing Angle	IPS Full View Angle

Tableau 4: Screen specifications [48].

7.1.3 Microphone:

The microphone is the specific component responsible for capturing the spoken audio. It’s a critical part of the system, as it initiates the speech input process.

7.1.4 Earphones :

The wireless earphone serves as the housing or integrated device that contains the microphone and facilitates Bluetooth connectivity to the smartphone. Including the earphone emphasizes the practical, user-friendly aspect of the hardware, as it’s the wearable component the user interacts with.

Figure 27: Wireless earphones.

7.1.5 Small activation button:

The button triggers the wireless earphone’s microphone to begin recording, making it an essential part of the hardware-user interaction.

7.1.6 Battery:

Lithium-ion battery with mAh capacity powers the smart glasses, providing an optimal balance between operating time and mobility. The battery is selected to power *-* hours of continuous use, representative of daily usage for real-time translation. The small battery and lightweight minimize the overall weight of the system, significant for user comfort in a wearable device.

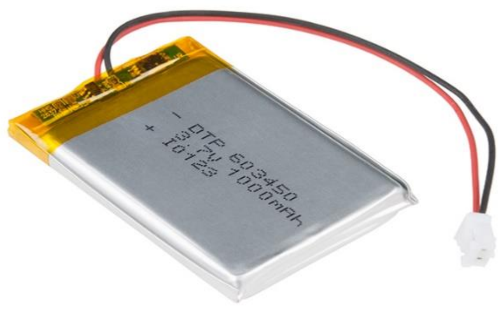

Figure 28: Lithium-ion battery.

.

7.1.7 Cables:

Cables allow the LILYGO T-Display-S3 AMOLED, microphone, battery, and button within the smart glasses to be linked. Selected based on their flexibility and reduced size, the cables enable reliable electrical connections within the ergonomic constraints of the frames of the glasses. They facilitate audio and power transfer to the ESP32-S3 and maintain system operation when in use.

7.1.8 3D case design :

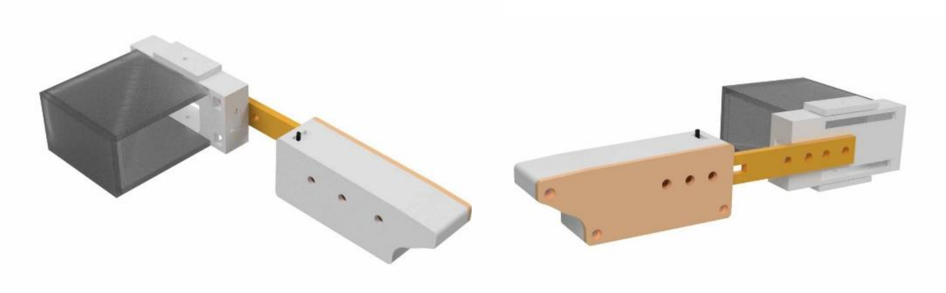

Figure 29: case design.

The 3D printed case is fabricated using plastic, encases the display, battery, and associated circuitry the case is designed using AutoDesk to fit snugly within the glasses frame, with openings for the microphone, button, and USB-C port. The 3D-printed case offers component protection with a compact and lightweight profile.

7.1.9 The prism:

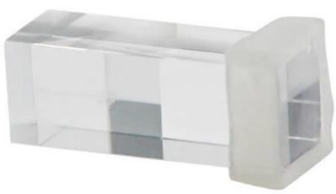

Figure 30: The prism.

The prism used in this project, constructed in Plexiglass and a mirror, is utilized as an optical component to direct the output of the 1.91-inch RM67162 IPS AMOLED display on the LILYGO T-Display-S3 to the user's eyes, enabling an augmented reality-like experience for viewing the sign gestures translation.

The Plexiglass, a lightweight and transparent acrylic plastic, is shaped to form a small prism that reflects the light from the display, with the incorporated mirror offering optimal reflection efficiency to project clear ASL sign gestures without obstructing the environmental perception of the user.

The prism works towards minimizing visual distortion as well as ergonomic integration, positioning the display output in line with the user's peripheral vision for real-time translation visibility.

7.2 Software tools and libraries:

7.2.1 Python:

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together.

Figure 31: Python – logo.

The primary programming language for the development of the STT, SLP, and 3D avatar generation pipeline was Python. Owing to its vast ecosystem of libraries and frameworks, readability, and flexibility, it was ideal for implementing complex machine learning models as well as multimedia processing. Python's support for both CPU and GPU calculations enabled efficient implementation of the deep learning components, while its wide usage in the research community facilitated simple integration with other tools and datasets [49].

7.2.2 Dart :

Dart is a client-optimized language for developing fast apps on any platform. Its goal is to offer the most productive programming language for multi-platform development, paired with a flexible execution runtime platform for app frameworks.

Dart also forms the foundation of flutter. Dart provides the language and runtimes that power Flutter apps, but Dart also supports many core developer tasks like formatting, analyzing, and testing code [50].

.

Figure 32: Dart-logo

Dart was selected as the programming language for developing the Flutter application of this project, offering a reliable platform for developing cross-platform apps that could run successfully on mobile, web, and desktop platforms. Its object-oriented nature and compatibility with Flutter's reactive framework helped to develop a responsive user interface that could communicate with the pre-trained sign language production (SLP) model. This choice ensured the eventual application was deployable across different devices, which was in accordance with the vision for accessibility and broad use of the project.

7.2.3 PyTorch :

PyTorch is an open-source machine learning framework that accelerates the path from research prototyping to production deployment. Built to offer maximum flexibility and speed, PyTorch supports dynamic computation graphs, enabling researchers and developers to iterate quickly and intuitively. Its Pythonic design and deep integration with native Python tools make it an accessible and powerful platform for building and training deep learning models at scale.[]

PyTorch was the chosen deep learning framework for developing and executing the SLP model. Its auto-differentiation capability and dynamic computation graph provided a degree of freedom in model design and optimization, while its GPU acceleration option significantly enhanced the inference speed of the neural networks [51].

Figure 33: PyTorch – logo.

7.2.4 Transformers (Hugging Face):

Transformers is a library of pretrained natural language processing, computer vision, audio, and multimodal models for inference and training. Use Transformers to train models on your data, build inference applications, and generate text with large language models.

The Hugging Face Transformers library was employed in leveraging pre-trained models and tokenizers, specifically openai/whisper-tiny for speech-to-text and bert-base-uncased for text encoding in the SLP model. The library allowed fast deployment of cutting-edge natural language processing and speech recognition models without deep training from the ground up. Its pre-trained weights and transfer learning capabilities facilitated the project to utilize community developments and attain high accuracy in text as well as speech processing under offline conditions [52].

7.2.5 NumPy :

NumPy is the fundamental package for scientific computing in Python. It is a Python library that provides a multidimensional array object, various derived objects (such as masked arrays and matrices), and an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more.[53]

Figure 34: NumPy-logo

NumPy was utilized for numerical computations, i.e., vertex data manipulation and frame stacking during the 3D rendering process. Its mathematical functions and vectorized array operations enabled convenient manipulation of large datasets, e.g., 3D coordinates of avatar meshes across multiple frames. This library served as the computational backbone for manipulating and managing the spatial data generated by the model.

7.2.6 Pandas :

Figure 35: Pandas – logo.

Pandas aims to be the fundamental high-level building block for doing practical, real world data analysis in Python. Additionally, it has the broader goal of becoming the most powerful and flexible open-source data analysis / manipulation tool available in any language.

Pandas was utilized for data inspection tasks, such as verifying the .pkl files from the dataset at the development stage. It proved to clarify and cross-checking the input data structures and contributing to the resilience of the data preprocessing steps [54].

7.2.7 Flutter Framework:

Figure 36: Flutter – logo.

Flutter is an open-source framework for building beautiful, natively compiled, multi-platform applications from a single codebase [55].

Flutter, Google's open-source UI toolkit, was utilized to build the cross-platform application that integrated the pre-trained SLP model. Its ability to compile to native code for mobile, web, and desktop was complemented by its reactive paradigm, enabling the creation of an easily accessible interface for users to interact with the avatar animation. Usage of this tool was instrumental in transforming the backend model into a functional application.

7.2.8 Dart SDK:

Dart SDK provided the development environment and compiler to write and run the Flutter application. It included native support for the Dart programming language in the form of facilities for running, debugging, and optimizing code to allow the app to be written and shipped effectively on target platforms according to the deployment intentions of the project.

7.3 Developement environnements :

7.3.1 kaggle notebook :

Figure 37: Kaggle-logo.

Kaggle Notebooks is a cloud computing platform provided by Kaggle to write, execute, and share code primarily data science and machine learning endeavors. The notebooks allow users to perform reproducible and collaborative analysis, with open access to Kaggle datasets as well as GPU resources for free model training. Kaggle Notebooks are similar to Jupyter Notebooks but are deeply integrated with Kaggle's environment, enabling simple experimentation and sharing among the Kaggle community [56].

Used to test and prototype the AI pipeline. Its provision of free ‘GPU T4 x2’ resources accelerated the training stage and inference stage, and its collaborative features allowed for iterative model building and testing. This platform was instrumental in the real-world deployment of the project in the short timeline.

7.3.2 Visual studio code :

Figure 38: Vs code-logo

Microsoft Visual Studio is an integrated development environment (IDE) developed by Microsoft that facilitates the entire software development life cycle. Visual Studio gives rise to computer programs like websites, web applications, web services, mobile applications using Microsoft software development platforms like Windows API, Windows Forms, and Windows Presentation Foundation (WPF). Visual Studio has a comprehensive set of tools with a code editor that includes IntelliSense (code completion), code refactoring, an inbuilt debugger (machine-level as well as source-level), code profilers, designers for GUI, web, class, and database schema. Visual Studio is also pluggable using plug-ins such as source control systems including Git and Subversion, and supports many programming languages such as C, C++, C#, Visual Basic .NET, F#, JavaScript, etc.. [58].

7.3.3 Arduino IDE :

Figure 39: Arduino IDE-logo

Arduino is an open-source electronics development platform that includes both hardware and software. On the hardware side, it is based on development boards equipped with microcontrollers and multiple input/output pins to connect to various electronic components. On the software side, Arduino provides a user-friendly integrated development environment (IDE) that allows programming these boards using a simplified C/C++-based programming language [57].

7.3.4 WiFi:

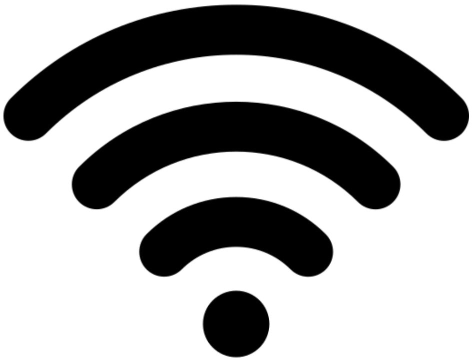

Figure 40: Wi-Fi logo.

WiFi is a wireless technology that allows electronic devices to connect to the internet and communicate with each other without a physical cable. This uses radio waves to transmit the data between a WiFi router and compatible devices like smartphones, computers, and smart home gadgets. These WiFi networks are common in homes, offices, and public spaces providing convenient internet access and local connectivity. This technology has become an essential part of modern digital life enabling wireless internet browsing, file sharing, and device communication in various settings [59].

7.3.5 Bluetooth:

Bluetooth short-range wireless technology enables two devices to connect directly without requiring supporting network infrastructure such as a wireless router or access point. Today, Bluetooth technology is most commonly used by people around the world to connect devices such as wireless headphones, keyboards, mice, and speakers to both PCs and mobile devices.Bluetooth technology operates on radio frequencies in the 2.4 GHz range [60].

Figure 41: Bluetooth logo.

7.3.6 Autodesk :

Autodesk, founded in 1982 by John Walker, is a leading American software company that develops and sells computer-aided design (CAD), 3D modeling, engineering, and entertainment software. While Autodesk is best known for its flagship product, AutoCAD, it also provides a wide range of industry-specific tools for architecture, engineering, construction, manufacturing, media, and entertainment [61].

Figure 42: AutoDesk- logo.

7.4 The prototype:

Figure 43: The prototype.

7.5 Application user interfaces:

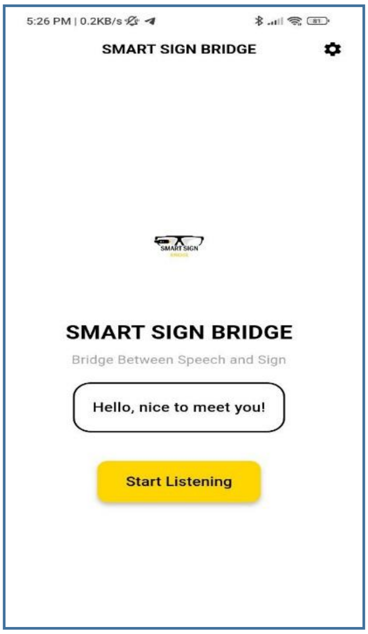

Figure 44: The application’s home interface

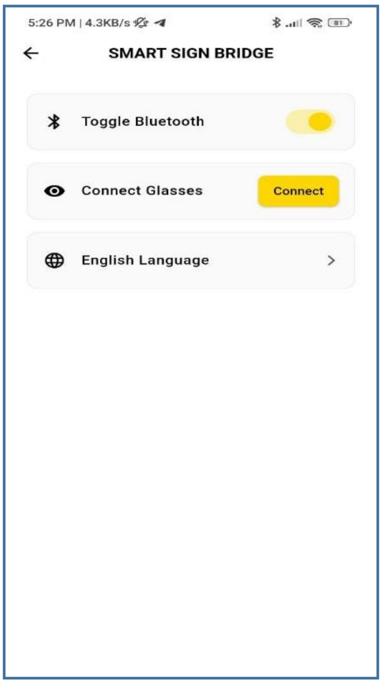

Figure 45: Application settings (Bluetooth, connect glasses)

Figure 46: Speech to text translation interface

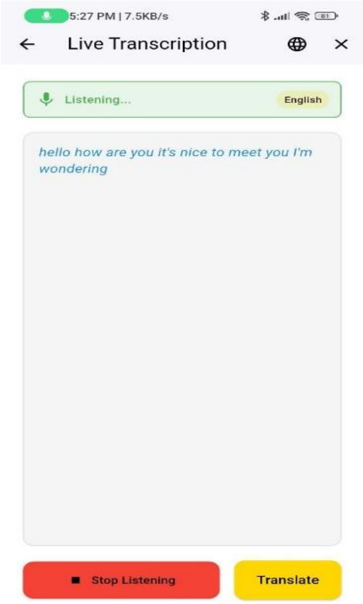

Figure 47: Example of speech recording

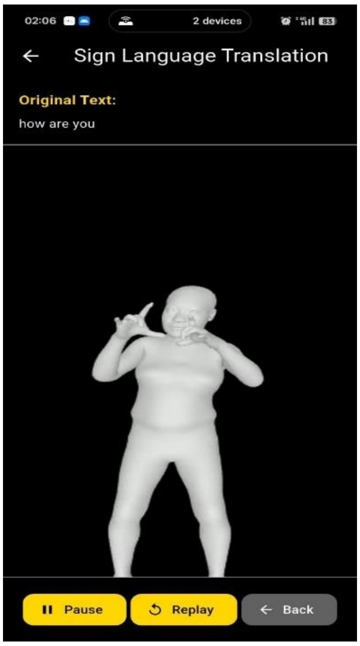

Figure 48: Animated 3D avatar

8 Conclusion:

This chapter has detailed the design and implementation of the "Smart Sign Bridge," an innovative project aimed at enhancing communication accessibility for deaf and hard-of-hearing individuals. By integrating cutting-edge technologies such as speech recognition, AI-powered sign language production, and holographic display systems, we have developed a functional prototype capable of real-time speech-to-sign-language translation. Key milestones included selecting optimized hardware components, implementing efficient AI models for local processing, and designing an intuitive mobile application interface. While the current prototype successfully demonstrates the concept's feasibility, future enhancements will focus on expanding language support, improving hardware comfort and battery life, and refining the AI models for greater accuracy and responsiveness.

In summary, the smart sign glasses represent a meaningful advancement in assistive technology, empowering users with greater independence and social inclusion. This work not only provides an immediate solution but also lays the foundation for future developments to broaden accessibility and adapt to evolving user needs.

Chapter 04:

Presentation of the commercial side

1 Axis 01: Project presentation

1.1 The idea of the project

The idea came from a growing awareness of the gaps in services available for deaf and hard-of-hearing individuals. Due to the lack of tailored solutions addressing their specific communication needs and the underrepresentation of this group in existing technologies, we engaged in extensive discussions with our supervisor about potential solutions. These conversations, following a suggestion from our supervisor, led to the development of an innovative concept: smart glasses that integrate artificial intelligence and real-time translation to enhance communication accessibility, making them more advanced compared to current offerings.

We found that conventional solutions like employing human interpreters or depending on mobile applications are either not private, not scalable, or simply not natural. Most of these solutions force users to look down at a screen, hold a device, which breaks the natural flow of interaction. These constraints necessarily illustrated to us that there must be a better, more immersive solution. This brought us to the idea of "Smart Sign Bridge" a pair of AI-powered smart glasses designed exclusively to facilitate real-time speech-to-sign language translation. The purpose is to bridge the communications gap by translating spoken words into 3D animated sign language gestures and projecting them directly into the user's field of vision through an AMOLED display and prism projection system integrated into the glasses. Our innovation goes beyond translation in a superficial sense. We aim to create an intelligent, hands-free, and portable sign language communication assistant for multiple sign languages (e.g., Algerian Sign Language) and the best-in-class user experience through a companion smartphone app. The smartphone app captures audio through earbuds, processes speech through cloud-based AI systems, and transmits the translated gestures to the glasses for real-time visual display. The system also encompasses settings customization, periodic updates, and avatar personalization. Other than its technological innovation, Smart Sign Bridge is also affordable, allowing the greatest numbers of people to use it especially where there is low availability of assistive devices. By doing so, we would like to increase inclusion and independence among users and allow them to communicate more freely and confidently in society. Project development is in the capable hands of two master's students who specialize in Artificial Intelligence. The one works on backend development, including the AI model, speech processing. The other works on mobile app development and UI/UX. The two worked together on the design and construction of the prototype of the smart glasses to achieve an ideal integration of hardware and software components. The project will be done within a specially equipped workshop with the right tools and technology for embedded development, experimentation with AI, and hardware testing. The workshop will also serve as the headquarters of our startup, where we will do testing, verification, and fine-tuning of the device in collaboration with assistive technology experts and members of the deaf community. Through this careful and practical route, we believe that Smart Sign Bridge not only will solve an underlying communication problem, but will also be a key to social innovation, inclusivity, and global evolution of assistive technologies.

1.2 Value propositions

Smart Sign Bridge project bridges an important gap in communication and accessibility of deaf and hard of hearing individuals through a wearable, AI-powered solution for instant speech-to-sign translation. Beyond the technical prototype, the project carries several levels of worth in individual, societal, professional, and technological arenas. The value propositions are established below:

1.2.1 Value to individuals:

- **Enhance autonomy:** Smart Sign Bridge enables deaf individuals to automatically follow spoken conversations in real time on their own, without relying on others or interpreters.
- **Usability in the real world:** Wearability of the device means it is best suited for real-word usage in common, real-life situations, hospitals, classrooms, or public buildings where instant convenient communication is required.

1.2.2 Value to Society:

- **Promotion of inclusion and equality:** By breaking communication barriers, Smart Sign Bridge facilitates social integration, especially for groups that typically get marginalized due to hearing disability.
- **Support for minority languages:** Integration of local sign languages (e.g., Algerian Sign Language) promotes cultural and linguistic diversity in the digital environment.
- **Awareness and visibility:** The project raises awareness of issues facing deaf communities and demonstrates how AI can be implemented for the greater good.

1.2.3 Value to the Professional and Working World:

- **Equal Access to Services and Employment:** Using Smart sign bridge tools, deaf individuals can participate more deeply in job interviews, meetings, or training sessions without the need for real-time human interpretation.

- **Improved Institutional Accessibility:** The public institution (e.g., university, administration, hospital) can leverage the solution to better serve citizens and employees with hearing disability.

1.2.4 Technological and Scientific Value:

- **Integrated Innovation:** Instead of building new technology from scratch, the project integrates speech recognition, AI natural language processing, gesture synthesis, and embedded AR into a fully functional, working system, something few such bundles can even offer on the market, much less in low-resource situations.
- **Validation by Prototype:** It demonstrates that integration is technologically viable with low-power, low-cost hardware, thus making it deployable in emerging economies.
- **Scalability:** The architecture is modular and such future extensions (e.g., reverse translation, offline, and support for more languages) are feasible without major reengineering.

1.2.5 Economic value:

- **Cost effective solution:** Using low-cost components lowers the barrier to universal use by schools, clinics, and families.
- **Startup Potential:** The initiative offers a good foundation for an AI startup in social tech and health tech with potential funding via accessibility-targeted grants, innovation programs, or incubators.
- **Job creation and ecosystem support:** Manufacturing, localization, support, and training of such a product would create specialized job opportunities in healthcare, education, and AI.

1.3 The working team

The working team is composed of two master's students in artificial intelligence Zirek Anfal and Nouar Wissam, under the academic supervision of Dr. Mohamed Cheikh doctor in artificial intelligence and computer security, The two students collaborated throughout all phases of the project, from design to research, system design, prototype development, and documentation. Together, they came up with the smart glasses prototype Smart Sign Bridge that can translate oral language into sign language using a 3D avatar, with the goal of making communication more accessible to deaf and hard-of-hearing individuals.

1.4 The objectives of the project:

The primary objective of our project is to design and develop innovative smart glasses to provide advanced communication assistance for individuals who employ sign language as their most frequent mode of interaction. These smart glasses will be equipped with an advanced AMOLED screen and prism technology. That will project holographic 3D avatar animations that represents sign language gestures in real time. This project named "**Smart Sign Bridge**" reflecting its vision to empower and connect individuals through accessible sign language technology. Furthermore, we aim to integrate complete multilingual features to support a diverse range of sign languages enhancing its global reach and inclusivity. To ensure widespread accessibility, the smart glasses offered at a low price, making this transformative technology available to a broad audience in need of communication support.

1.5 Project implementation plan:

			M 1	M2	M 3	M 4	M 5	M 6	M 7
1		Preliminary studies and planning	✓	✓					
2		AI pipeline and mobile application developement		✓	✓				
3		Infrastructure and equipment demand			✓	✓	✓		
4		Equipment installation, system integration and work environment preparation				✓	✓		
5		Test, evaluation and Optimization						✓	
6		Finalization, marketing and product launch							✓

Tableau 5: Calendar

2 Axis 02: Innovative aspects

2.1 Nature of innovations

The innovations introduced in Smart Sign Bridge are technological as well as use-case driven, with a strong emphasis on integration and accessibility:

- **System level integration:** The project's innovation is not in the creation of new algorithms but in innovative integration of pre-existing parts (speech recognition, NLP, gesture animation, and AR display) into a single logical, cohesive, and functional prototype designed for accessibility.
- **User Centric design in wearable format:** Unlike the overwhelming majority of translation systems that rely on computers or phones, Smart Sign Bridge attends to user-friendly convenience, portability, and privacy seriously by incorporating display of the translation directly into the user's line of sight.
- **Real time translation loop:** The system mitigates communication latency and lag through a slim real-time processing pipeline from audio input, cloud translation, to visual output a key characteristic for the sustainability of sign language interpretation in informal discussions.
- **Scalability:** The system architecture is module-based in order to accommodate future breakthroughs such as bidirectional translation (sign-to-speech), offline processing, and integration with education or public service platforms.
- **Access through low cost materials:** The project uses low-cost and low-energy materials, which enhances the adoption of the innovation in low-resource environments and enhances its potential for large-scale deployment and societal impact.

2.2 Areas of innovation

Smart Sign Bridge innovates at the intersection of several technological and societal fields:

- **Assistive technology:** Smart Sign Bridge provides added value to assistive technology through the introduction of a real-time speech-to-sign translation system for deaf and hard of hearing individuals, with a viable and portable implementation.
- **Artificial intelligence and language processing:** The project leverages AI techniques such as speech recognition and natural language processing (NLP) to decode spoken language into structured sign language output in real-time.

- **Embedded and wearable systems:** By integrating the display system into low-power microcontroller-based systems of smart glasses (e.g., ESP32), the project offers an embedded and wearable solution that is light and portable as opposed to screen-based solutions, something that is a novelty in sign language translation tools.
- **Inclusive digital communication:** The inclusion of regional sign languages (e.g., Algerian Sign Language) is a breakthrough in the localization of accessibility mechanisms such that solutions are customized to linguistic and cultural environments often neglected in global implementations.

3 Axe 03: Description of the invention

3.1 Title of the invention:

« **Smart Sign Bridge** » Advanced Smart Glasses for assisting deaf and hard-of-hearing individuals.

3.2 Summary of the Invention:

The patent for the " Smart Sign Bridge for Real-Time for Speech-to-Sign Language Translation" presents an innovative solution designed to enhance communication for deaf and hard-of-hearing individuals who rely on sign language, by integrating advanced wearable technology and artificial intelligence. These smart glasses incorporate cutting-edge features that translate spoken language into expressive sign language gestures in real time, enabling users to interact seamlessly with their environment and hearing individuals. Utilizing sophisticated AI, the glasses provide a dynamic and accessible communication tool, fostering social inclusion and independence.

The patent highlights the limitations of prior art, noting the inadequacy of existing assistive devices that fail to offer real-time sign language support or cater specifically to the needs of deaf users. Unlike previous solutions, the proposed smart glasses are tailored to address these gaps, offering a practical and affordable option.
The primary objective of the patent is to deliver an effective tool that boosts the quality of life for deaf individuals, enhancing their confidence and enabling full participation in daily activities. The smart glasses represent a significant advancement in assistive communication technology, providing a versatile solution to overcome communication barriers.
In summary, the patent describes a promising innovation with the potential to positively impact the lives of millions of deaf individuals globally. By combining the latest technological advancements, the "Smart Sign Bridge" offer a groundbreaking approach to improving autonomy and inclusion for deaf and hard-of-hearing people in society.

3.3 Technical field to which the invention belongs:

In the field of assistive communication technology, various devices have been developed to support individuals with hearing impairments, blending technology and design for unique experiences. However, these devices often fall short, being expensive, complex, or not specifically designed for the diverse needs of deaf users. This is where our smart glasses come into play, integrating recent advancements in artificial intelligence to enhance the daily lives of deaf and hard-of-hearing individuals. They enable real-time translation of spoken language into sign language, assisting users in engaging with their surroundings. The glasses support interaction in various settings, recognizing spoken input and converting it into visual sign language representations. Additionally, they facilitate social connections by providing an accessible communication bridge. In essence, this innovation aims to improve the quality of life for those facing hearing-related challenges, offering a practical and effective tool to navigate and connect with the world around them.

3.4 Objective of the invention:

The objective of our invention is to provide deaf and hard-of-hearing individuals with an innovative tool that enhances their quality of life. These smart glasses integrate the latest technological advancements with a thoughtful design, aiming to increase user independence and autonomy by offering easier access to communication and interaction with their environment. The goal is to create a product that promotes social inclusion, boosts self-confidence, and enables these individuals to participate fully in modern society.

3.5 Essence of the invention:

The "Real-Time Smart Glasses for Speech-to-Sign Language Translation" are a pioneering assistive device aimed at improving the quality of life for deaf and hard-of-hearing individuals through advanced technology. Equipped with a T-Display S3 AMOLED screen, a prism, a button for activation, a wireless earphone with an integrated microphone, and a lithium battery, the glasses capture and process spoken language in real time. Supported by a smartphone application and a server-based system, the glasses utilize AI to translate spoken input into sign language gestures, which are displayed as an animated avatar. This capability allows users to engage with their environment and others effectively. The device also enhances social interaction by providing a reliable communication tool. With its user-friendly design and innovative features, the smart glasses offer a new level of independence, making them a valuable asset for deaf and hard-of-hearing users worldwide.

4 Axe 04: The Claims

Smart Glasses for Real-Time Speech-to-Sign Language Translation.

- This invention stands out for its ability to provide innovative communication assistance to deaf and hard-of-hearing individuals. It integrates artificial intelligence and advanced technologies to enhance their quality of life and offer seamless access to communication.
- The system is characterized by its capacity to translate speech into sign language in real time, instantly converting spoken words into expressive gestures via an animated avatar, thus providing valuable support for social interaction.
- This invention includes a contextual translation feature, enabling users to understand and engage in conversations across various indoor and outdoor environments. This capability facilitates effective communication navigation and social integration.
- The technology offers an intuitive interface with a button for activation, allowing simple and practical use to ensure a smooth user experience.
- The advanced artificial intelligence algorithms utilized analyze audio data and generate precise gestures, ensuring fast and effective assistance.
- The economic model of this invention is designed to make this technology affordable for deaf and hard-of-hearing individuals, promoting greater social inclusion and an improved quality of life.
- People with hearing impairments who rely on sign language

5 Axis 05: List of appendices:

5.1 Prototype:

Figure 49: The external face of the first prototype of Smart Sign Bridge

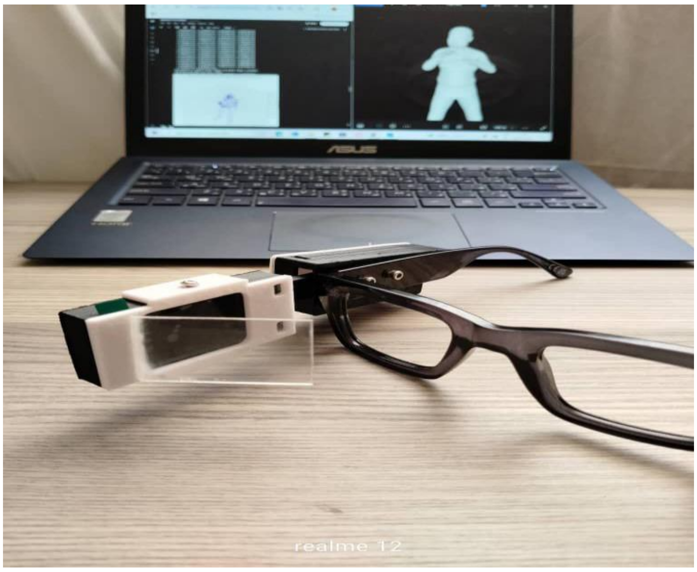

Figure 49: The inner side face of the first prototype of Smart Sign Bridge

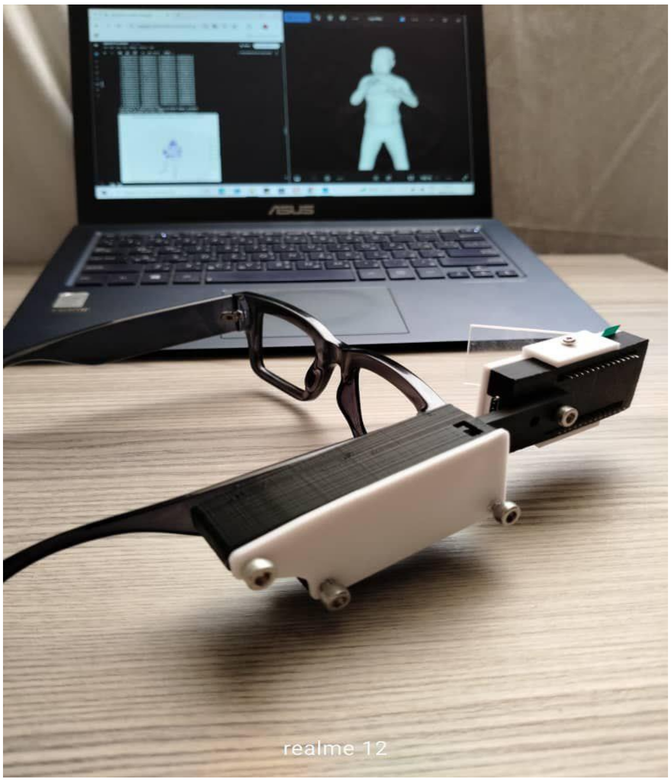

Figure 51: The Side view of the first prototype of Smart Sign Bridge

5.2 Business model canvas (BMC)

Business Model Canvas

Key partners

- Technology providers and AI developers
- Partnerships with deaf advocacy organizations
- Collaborations with educational institutions
- Suppliers of wearable hardware components
- Research institutions

Key Activities

- Product development and testing
- AI model training and updates
- Manufacturing and assembly
- User support and training
- Marketing and distribution

Key Resourcs

- Research and development team
- AI and software expertise
- Partnerships with Electronic Equipment Manufacturers
- Server infrastructure for AI processing

Value Propositions

- Real-time speech-to-sign language translation
- Enhanced communication accessibility
- User-friendly wearable design
- intuitive user interface
- Inclusive and expressive avatar display
- Affordable and scalable solution

Customer Relationship

- Personalized support for deaf users
- Online community to share experiences and advice
- Ongoing technical assistance
- Feedback integration for improvements
- Accessible customer service

Channels

- Online sales through websites and e-commerce platforms
- Partnerships with distributors and retailers
- Social media and digital marketing
- Collaborations with NGOs and accessibility events

Customer Segments

B2C:

•Deaf individuals

•People with hearing impairments who rely on sign language.

B2B:

•Educational and training institutions that work with people with special needs

•Healthcare providers

Government and NGOs

	• Development and maintenance facilities			

Cost structure	Revenue Stream
• Costs of materials for hardware production • Research and development expenses • Marketing and promotion costs • Operational costs (staff, servers) • Development of technology and training to develop and expand the functionality of smart glasses	• Sales of smart glasses • subscriptions for software updates and premium usage • Contracts with institutions and NGOs (Non-governmental organizations) • Maintenance and support services • Collaborations for custom solutions

Tableau 6: BMC representation of « Smart Sign Bridge ».

5.3 Costs:

Years	Task List	Expenses Hardware	Quantity	Smartphone Application Expenses	Total
1st year	• Smart glasses development • Mobile application development • Maintenance	**17,000.00 DA**	**100**	**70,000.00 DA**	**1,770,000.00 DA**
2nd year	• Smart glasses development • Mobile application development • Maintenance	**17,000.00 DA**	**300**	**210,000.00 DA**	**5,310,000.00 DA**
3rd year	• Smart glasses development • Mobile application development • Maintenance	**17,000.00 DA**	**500**	**350,000.00 DA**	**8,850,000.00 DA**
4th year	• Smart glasses development • Mobile application development • Maintenance	**17,000.00 DA**	**700**	**490,000.00 DA**	**12,390,000.00DA**

Tableau 7: Cost table « Smart Sign Bridge ».

5.4 Profits:

Years	Selling price of glasses	Number of glasses	Total
1st year	**50,000.00 DA**	**100**	**5,000,000.00 DA**
2nd year	**50,000.00 DA**	**300**	**15,000,000.00 DA**
3rd year	**50,000.00 DA**	**500**	**25,000,000.00 DA**
4th year	**50,000.00 DA**	**700**	**35,000,000.00 DA**

Tableau 8: Profits table « Smart Sign Bridge ».

General conclusion

General conclusion

Living with reliance on sign language in a society that prioritizes verbal communication is extraordinarily challenging. For the majority of the deaf and hard-of-hearing individuals around the world, being exists in terms of living as a quest to ensure commonality between their visual language and society, which prefers to turn a blind eye to it. Irrespective of whether their life began with hearing gradually worsening or with a world that was never loud due to conditions existing at birth, genetic causes like GJB2 mutations, or infection such as meningitis, the result is a severe loss of autonomy. The inability to easily engage with the daily conversations, is one that creates isolation and psychological strain, with research observing high levels of depression and anxiety within deaf populations. Professionals too might not be able to create solutions that fully honor the language and cultural identity of the sign language users, so many must live in a world not created for them. But the absence of perfect solutions does not exclude hope; new, affordable technology offers a promise, with implications that extend beyond technical shortcuts.

This makes us understand deaf and hard-of-hearing people who communicate through sign language must enter settings that rarely accommodate them, remodelling their attitude and way of life to gain access to the world in general. Assistive technology, such as our groundbreaking Smart Sign Bridge, can significantly enable such adoption. Designed as a final-year project by two enthusiastic master's students, Smart Sign Bridge uses artificial intelligence to translate spoken words into emotive 3D sign language avatars in real time, delivered through a lightweight, user-friendly interface.

Due to time constraints, we could not realize the project's full potential within our academic endeavor. Still, we are determined to continue working on Smart Sign Bridge, to explore options such as offline translation, the ability to work with multiple sign languages such as Algerian Sign Language, and personalized avatars to better suit deaf communities globally. Our aim is to make Smart Sign Bridge more helpful in a way that allows sign language people with independence, self-respect, and a sense of belonging to society in spite of the fact that it excludes them, creating a world that includes them where their voices, as uttered in signs, do matter and are worth appreciating.

References

[1] Diana C. Peterson et al., “Neuroanatomy, Auditory Pathway,” in StatPearls [Internet], NCBI Bookshelf (updated October 24, 2023).

[2]F. E. Musiek and J. A. Baran, *The Auditory System: Anatomy, Physiology, and Clinical Correlates*, 2nd ed. San Diego, CA: Plural Publishing, 2020.

[3] Stanford Children's Health, "Anatomy and Physiology of the Ear," Stanford Medicine Children's Health, [Online]. Available: https://www.stanfordchildrens.org/en/topic/default?id=anatomy-and-physiology-of-the-ear-90-P02025.

[4]World Health Organization, "Deafness and hearing loss," WHO, Mar. 20, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss.

[5] National Institute on Deafness and Other Communication Disorders (NIDCD), "Hearing Loss and Older Adults," [Online]. Available: https://www.nidcd.nih.gov/health/hearing-loss-older-adults.

[6] G. I. Prabowo, “Study of Gene Mutation in Congenital Deaf Patients in Surabaya,” Universitas Airlangga Official Website, Aug. 19, 2019. [Online]. Available: https://unair.ac.id/study-of-gene-mutation-in-congenital-deaf-patients-in-surabaya/.

[7]Hearing Health Resource, *Types of Hearing Loss*. [Online]. Available: [PDF document].

[8]W. B. Alshuaib, J. M. Al-Kandari, and S. M. Hasan, ‘Classification of Hearing Loss’, Update On Hearing Loss. InTech, Dec. 02, 2015. doi: 10.5772/61835.

[9] Harvard Medical School Center for Hereditary Deafness, *Common Causes of Hearing Loss: For Parents & Families*, 2004. [PDF].

[10]National Deaf Center on Postsecondary Outcomes, “Deaf People and Educational Attainment,” 2020. [Online]. Available: https://www.nationaldeafcenter.org.

[11]E. Punch, D. Hyde, and G. Power, “Career and workplace experiences of Australian university graduates who are deaf or hard of hearing,” *Journal of Deaf Studies and Deaf Education*, vol. 11, no. 4, pp. 479–492, 2006.

[12]C. C. M. Professional, “Hearing aids,” Cleveland Clinic, Mar. 19, 2025. https://my.clevelandclinic.org/health/treatments/24756-hearing-aids.

[13]https://hearingup.com/blog/best-hearing-aid-features-2025-dr-cliff.

[14] https://www.hearingtracker.com/hearing-aids/best-hearing-aids-background-noise.

[15] https://www.hearingaid.org.uk/hearing-aid-technology/hearing-aid-technology-2025.

[16] https://www.hearingaid.org.uk/hearing-aid-technology/hearingaid-technology-2025.
[17]Great Ormond Street Hospital, “Cochlear implants,” GOSH Hospital Site. https://www.gosh.nhs.uk/conditions-and-treatments/procedures-and-treatments/cochlearimplants/
[18]Center for Devices and Radiological Health, “What is a Cochlear Implant?,” U.S. Food And Drug Administration, Feb. 04, 2018. https://www.fda.gov/medical-devices/cochlearimplants/what-cochlear-implant
[19]https://www.biotechdaily.com.au/wp/wp-content/uploads/2025/02/COH-Feb-212025.pdf.
[20]https://www.businesswire.com/news/home/20250319879319/en/Cochlear-Implants-Market-Report-2025-A-Global-and-Regional-Analysis-with-Focus-on-Patient-Type-and-End-User-2024-2034-Featuring-Profiles-of-Leading-Players-CochlearLimited-Sonova-Holding-More---ResearchAndMarkets.com)
[21](https://www.businesswire.com/news/home/20250319879319/en/Cochlear-Implants-Market-Report-2025-A-Global-and-Regional-Analysis-with-Focus-on-Patient-Type-and-End-User-2024-2034-Featuring-Profiles-of-Leading-Players-Cochlear-Limited-Sonova-Holding-More---ResearchAndMarkets.com)
[23] E. Ostrowski, “Assistive listening devices for hearing loss,” Healthy Hearing, Feb. 03, 2025. https://www.healthyhearing.com/help/assistive-listening-devices (accessed Apr. 11, 2025).
[24]American Academy of Audiology, “Assistive Listening and Alerting Devices - American Academy of Audiology,” American Academy of Audiology, Feb. 08, 2022. https://www.audiology.org/consumers-and-patients/managing-hearing-loss/assistivelistening-and-alerting-devices/ (accessed Apr. 11, 2025).
[25]https://www.health.gov.au/sites/default/files/2025-04/hearing-servicesprogram-minimum-specifications-for-subsidised-devices-report.pdf.
[26](https://rnid.org.uk/wp-content/uploads/2025/02/Final-Auracast-Report110225.pdf.
[27]https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4285000/.
[28] State of Connecticut and Department of Aging and Disability Services, “Sign language interpreting services.” [Online]. Available: https://portal.ct.gov/-/media/dcf/multicultural_affairs/2023/ads-asl-interpreters-fact-sheet2022.pdf?rev=d19db61f883242c980f1a7a87fd64c79&hash=B3B764502F3CD4AB51EC140EA005F7AA)
.[29](https://www.linkedin.com/pulse/how-ai-addressing-shortage-sign-languageinterpreters-wb3ec)

.[30] (M. Busari, “Smart Wearable Devices for Deaf People : Integrating Haptic Feedback with Image Recognition,” Feb. 2025.)
[32](https://dl.acm.org/doi/pdf/10.1145/3132525.3134781)
[33](https://www.sigaccess.org/newsletter/2023-01/Emma_McDonnell.html).
[34](https://abilitynet.org.uk/news-blogs/smart-glasses-resurgence-technical-andpractical-barriers?utm_source=chatgpt.com)
[35](https://www.researchgate.net/publication/388732891_Exploring_Deaf_And_Hard_of_Hearing_Peoples'_Perspectives_On_Tasks_In_Augmented_Reality_Interacting_With_3D_Objects_And_Instructional_Comprehension).
[36]Kartik, “Artificial Intelligence Tutorial,” GeeksforGeeks, 25 Jun. 2025. [Online]. Available: https://www.geeksforgeeks.org/artificial-intelligence/artificial-intelligence/.
[37] GeeksforGeeks, “Machine Learning (ML) – Machine Learning Tutorial,” GeeksforGeeks, Apr. 10, 2025. [Online]. Available: https://www.geeksforgeeks.org/ml-machine-learning/.
[38] T. O. Ayodele, “Types of machine learning algorithms,” in *New Advances in Machine Learning*, vol. 3, no. 19–48, pp. 5–1, 2010.
[39]DataCamp, “What are Neural Networks?,” DataCamp Blog, Dec. 2023. [Online]. Available: https://www.datacamp.com/blog/what-are-neural-networks.
[40] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2009.
[41] J. Holdsworth and M. Scapicchio, “What Is Deep Learning?,” IBM, Jun. 17, 2024. [Online]. Available: https://www.ibm.com/think/topics/deep-learning. [Accessed: Jul. 2, 2025].
[42]TechTarget, “What is Speech Recognition?,” TechTarget Search Customer Experience, 2024.[Online].Available:https://www.techtarget.com/searchcustomerexperience/definition/speech-recognition.
[43] J. Hui, “Speech Recognition — Acoustic, Lexicon & Language Model,” Medium, Sep. 15, 2019. [Online]. Available: https://jonathan-hui.medium.com/speech-recognition-acoustic-lexicon-language-model-aacac0462639.
[44] Y. Dong and L. Deng, “An overview of HMM-Based large vocabulary continuous speech recognition,” *Symmetry*, vol. 11, no. 8, p. 1018, 2019. [Online]. Available: https://www.mdpi.com/2073-8994/11/8/1018.
[45]S. Karpagavalli and E. Chandra, “A Review on Automatic Speech Recognition Architecture and Approaches,” *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9, no. 4, pp. 393–404, 2016. [Online]. Available:

https://www.researchgate.net/publication/302915903_A_Review_on_Automatic_Speech_Recognition_Architecture_and_Approaches.
[46]DeepLearning.AI, “Natural Language Processing Resources,” DeepLearning.AI. [Online]. Available: https://www.deeplearning.ai/resources/natural-language-processing/
[47] A. Upadhyay, “NLP 101 – NLP Architecture Explained,” *E2E Networks Blog*, Jul. 14, 2023. [Online]. Available: https://www.e2enetworks.com/blog/nlp-101-nlp-architecture-explained/.
[48] LILYGO®, “T-Display S3 AMOLED,” LILYGO®. https://lilygo.cc/products/t-display-s3amoled?srsltid=AfmBOor_sRBvzUp_oBqEjcpaCcXPH8H0a3EE058e4AAVPO6ghfGkx2VM .
[49] “What is Python?,” Python.org. https://www.python.org/doc/essays/blurb/ .
[50] “Dart overview,” Dart. https://dart.dev/overview#language .
[51] “PyTorch – PyTorch.” https://pytorch.org/projects/pytorch/ .
[52] https://huggingface.co/.
[53] https://numpy.org/doc/stable/user/whatisnumpy.html.
[54] https://pandas.pydata.org/about/.
[55]https://flutter.dev/#:~:text=Flutter%20is%20an%20open%20source,applications%20from%20a%20single%20codebase.
[56] https://www.kaggle.com/docs/notebooks.
[57] https://docs.arduino.cc/software/ide-v1/tutorials/arduino-ide-v1-basics.
[58] https://learn.microsoft.com/en-us/visualstudio/get-started/visual-studio-ide?view=vs-2022].
[59] Cisco, “What Is Wi-Fi? – Definition and Types,” Cisco Products & Technologies, published three months ago. [Online]. Available: https://www.cisco.com/c/en/us/products/wireless/what-is-wifi.html.
[60] Intel, “What Is Bluetooth® Technology?,” Intel Wireless Products and Technologies. [Online]. Available: https://www.intel.com/content/www/us/en/products/docs/wireless/what-is-bluetooth.html.
[61]“History of Autodesk, Inc.,” *FundingUniverse Company Histories*. [Online]. Available: https://www.fundinguniverse.com/company-histories/autodesk-inc-history/.