

Software Requirements Specification for Software Engineering: subtitle describing software

Team 13, Speech Buddies

Mazen Youssef

Rawan Mahdi

Luna Aljammal

Kelvin Yu

October 10, 2025

Contents

1	Purpose of the Project	vi
1.1	User Business	vi
1.2	Goals of the Project	vii
2	Stakeholders	vii
2.1	Client	vii
2.2	Customer & Hands-On Users of the Project	viii
2.3	Other Stakeholders	viii
2.4	Personas	viii
2.5	Priorities Assigned to Users	ix
2.6	User Participation	ix
2.7	Maintenance Users and Service Technicians	ix
3	Mandated Constraints	x
3.1	Solution Constraints	x
3.2	Implementation Environment of the Current System	x
3.3	Partner or Collaborative Applications	x
3.4	Off-the-Shelf Software	x
3.5	Anticipated Workplace Environment	x
3.6	Schedule Constraints	x
3.7	Budget Constraints	x
3.8	Enterprise Constraints	xi
4	Naming Conventions and Terminology	xi
4.1	Glossary of All Terms, Including Acronyms, Used by Stakeholders involved in the Project	xi
4.2	Technical Terminology	xi
4.3	Medical Terminology	xi
5	Relevant Facts And Assumptions	xii
5.1	Relevant Facts	xii
5.2	Business Rules	xii
5.3	Assumptions	xii
6	The Scope of the Work	xii
6.1	The Current Situation	xii
6.2	The Context of the Work	xii

6.3	Work Partitioning	xii
6.4	Specifying a Business Use Case (BUC)	xii
7	Business Data Model and Data Dictionary	xii
7.1	Business Data Model	xii
7.2	Data Dictionary	xiii
8	The Scope of the Product	xiii
8.1	Product Boundary	xiii
8.2	Product Use Case Table	xiii
8.3	Individual Product Use Cases (PUC's)	xiii
9	Functional Requirements	xiii
9.1	Functional Requirements	xiii
10	Look and Feel Requirements	xiii
10.1	Appearance Requirements	xiii
10.2	Style Requirements	xiv
11	Usability and Humanity Requirements	xiv
11.1	Ease of Use Requirements	xiv
11.2	Personalization and Internationalization Requirements	xiv
11.3	Learning Requirements	xiv
11.4	Understandability and Politeness Requirements	xiv
11.5	Accessibility Requirements	xiv
12	Performance Requirements	xv
12.1	Speed and Latency Requirements	xv
12.2	Safety-Critical Requirements	xv
12.3	Precision or Accuracy Requirements	xv
12.4	Robustness or Fault-Tolerance Requirements	xv
12.5	Capacity Requirements	xv
12.6	Scalability or Extensibility Requirements	xvi
12.7	Longevity Requirements	xvi
13	Operational and Environmental Requirements	xvi
13.1	Expected Physical Environment	xvi
13.2	Wider Environment Requirements	xvi
13.3	Requirements for Interfacing with Adjacent Systems	xvi

13.4	Productization Requirements	xvii
13.5	Release Requirements	xvii
14	Maintainability and Support Requirements	xvii
14.1	Maintenance Requirements	xvii
14.2	Supportability Requirements	xvii
14.3	Adaptability Requirements	xviii
15	Security Requirements	xviii
15.1	Access Requirements	xviii
15.2	Integrity Requirements	xviii
15.3	Privacy Requirements	xviii
15.4	Audit Requirements	xix
15.5	Immunity Requirements	xix
16	Cultural Requirements	xix
16.1	Cultural Requirements	xix
17	Compliance Requirements	xix
17.1	Legal Requirements	xix
17.2	Standards Compliance Requirements	xix
18	Open Issues	xx
19	Off-the-Shelf Solutions	xx
19.1	Ready-Made Products	xx
19.2	Reusable Components	xx
19.3	Products That Can Be Copied	xx
20	New Problems	xx
20.1	Effects on the Current Environment	xx
20.2	Effects on the Installed Systems	xx
20.3	Potential User Problems	xxi
20.4	Limitations in the Anticipated Implementation Environment That May Inhibit the New Product	xxi
20.5	Follow-Up Problems	xxi

21 Tasks	xxi
21.1 Project Planning	xxi
21.2 Planning of the Development Phases	xxi
22 Migration to the New Product	xxii
22.1 Requirements for Migration to the New Product	xxii
22.2 Data That Has to be Modified or Translated for the New System	xxii
23 Costs	xxii
24 User Documentation and Training	xxiii
24.1 User Documentation Requirements	xxiii
24.2 Training Requirements	xxiii
25 Waiting Room	xxiii
26 Ideas for Solution	xxiii
26.1 Browser-Based Extensions	xxiii
26.2 Personalized ASR Fine-Tuning	xxiii
26.3 LLM Command Mapping	xxiii
26.4 Noise Filtering	xxiv
26.5 Open-Source Integrations	xxiv
26.6 Accessibility Enhancements	xxiv
26.7 Data Encryption	xxiv

Revision History

Date	Version	Notes
------	---------	-------

1 Purpose of the Project

1.1 User Business

Individuals with speech impairments face significant barriers when interacting with digital devices. VoiceBridge addresses this gap by providing an accurate, inclusive, and accessible speech-to-control system that enables users to communicate with their devices using their speech, regardless of clarity of articulation. Building on familiar technology, such as personal computers and mobile devices, may be one of the most cost-effective and easily adoptable approaches for improving an individual with disability's autonomy and access to the world around them. The rise of ASR technology and Artificial Intelligence (AI) integrations in the industry provides a novel landscape of opportunities to improve accessibility interfaces. VoiceBridge exploits bleeding-edge technology for a practical and impactful application.

1.2 Goals of the Project

ID	Goal	Description
G1	Accurate Speech Transcription	Reliably convert impaired or slurred speech into text.
G2	Command Mapping	Translate recognized speech into actionable browser commands.
G3	User Independence	Enable users with speech impairments to browse autonomously.
G4	Lightweight & Accessible Design	Keep the system simple, fast, and cost-effective.
G5	Cross-Browser Compatibility	Support major browsers (e.g., Chrome, Edge, Firefox).
G6	Robust Error Handling	Detect and recover gracefully from common failures.
G7	Data Privacy & Security	Protect user data and ensure secure local processing.
G8	Customizable Interface	Allow users to adjust sensitivity, shortcuts, and feedback modes.
G9	Scalable Architecture	Design the system for future integration beyond browsers.

Table 1: Project Goals for VoiceBridge

2 Stakeholders

2.1 Client

The primary client for the VoiceBridge project is the organization or individual funding or commissioning the system. The client is primarily concerned with achieving the following goals: **G1**, **G5**, and **G9**, ensuring accurate speech transcription, cross-browser compatibility, and scalable architecture.

2.2 Customer & Hands-On Users of the Project

The primary customers are also the users of the project, them being individuals with speech impairments, who seek independence and autonomy through technology. Their needs directly relate to: **G1**, **G2**, **G3**, and **G8**.

2.3 Other Stakeholders

Secondary stakeholders include experts in linguistics, speech processing, and healthcare domains:

- Speech researchers and linguistics specialists, including the project supervisor, Dr. Christian Brodbeck, who provide insight toward **G1** and **G6**.
- Healthcare professionals and speech therapists who advise on usability and accessibility, contributing to **G3** and **G8**.
- Accessibility advocates and organizations interested in promoting the application, aligned with **G4** and **G5**.
- Software developers who implement and maintain the system, supporting **G9** and **G6**.

Tertiary stakeholders include caregivers and professionals who interact with end users, supporting **G3** and **G7**.

2.4 Personas

Potential end users of VoiceBridge include:

- **Amira**, a 45-year-old with Parkinson's disease, uses the system to log into Gmail and send emails.
- **David**, a stroke survivor, uses the system to browse the web and make purchases.

2.5 Priorities Assigned to Users

The highest priorities are assigned to end users with speech impairments, as their experience with the system defines its success. Secondary priorities include caregivers and technical experts who support the end users in using and maintaining the system.

2.6 User Participation

Individuals matching the target user profiles will be recruited for prototype testing and personalization development. Their participation directly supports refining goals **G1**, **G2**, **G3**, and **G8**.

2.7 Maintenance Users and Service Technicians

Maintenance Users:

- **Role:** End-users or caregivers performing basic troubleshooting and initiating support requests.
- **Responsibilities:**
 - Reporting errors or unexpected system behavior (**G6**).
 - Installing application updates (**G9**).
 - Managing user-specific configurations (**G8**).

Service Technicians:

- **Role:** Trained technical staff with deeper access to system logs and back-end services.
- **Responsibilities:**
 - Investigating reported issues (**G6**).
 - Ensuring transcription accuracy (**G1**).
 - Deploying updates and patches (**G9**).

- Ensuring compatibility with operating systems and accessibility frameworks ([G5](#)).
- Performing preventive maintenance, including performance monitoring and optimization ([G1](#), [G6](#)).

3 Mandated Constraints

3.1 Solution Constraints

Insert your content here.

3.2 Implementation Environment of the Current System

Insert your content here.

3.3 Partner or Collaborative Applications

Insert your content here.

3.4 Off-the-Shelf Software

Insert your content here.

3.5 Anticipated Workplace Environment

Insert your content here.

3.6 Schedule Constraints

Insert your content here.

3.7 Budget Constraints

Insert your content here.

3.8 Enterprise Constraints

Insert your content here.

4 Naming Conventions and Terminology

4.1 Glossary of All Terms, Including Acronyms, Used by Stakeholders involved in the Project

Insert your content here.

4.2 Technical Terminology

ASR - Automatic Speech Recognition

TTS - Text To Speech

STT - Speech To Text

4.3 Medical Terminology

Aphasia - A condition that robs you of the ability to communicate. It can affect your ability to speak, write and understand language, both verbal and written. Aphasia usually occurs suddenly after a stroke or a head injury. But it can also come on gradually, as in the case of a brain tumor or a progressive neurological disease.

ALS - Amyotrophic Lateral Sclerosis

Dysarthria - A motor speech disorder that makes it hard to speak. It is caused by damage to the nervous system, which can affect the muscles used for speaking. People with dysarthria may have slurred or slow speech, and they may have difficulty controlling the pitch, volume, and rhythm of their speech.

5 Relevant Facts And Assumptions

5.1 Relevant Facts

Insert your content here.

5.2 Business Rules

Insert your content here.

5.3 Assumptions

Insert your content here.

6 The Scope of the Work

6.1 The Current Situation

Insert your content here.

6.2 The Context of the Work

Insert your content here.

6.3 Work Partitioning

Insert your content here.

6.4 Specifying a Business Use Case (BUC)

Insert your content here.

7 Business Data Model and Data Dictionary

7.1 Business Data Model

Insert your content here.

7.2 Data Dictionary

Insert your content here.

8 The Scope of the Product

8.1 Product Boundary

Insert your content here.

8.2 Product Use Case Table

Insert your content here.

8.3 Individual Product Use Cases (PUC's)

Insert your content here.

9 Functional Requirements

9.1 Functional Requirements

Insert your content here.

10 Look and Feel Requirements

10.1 Appearance Requirements

The interface shall have a clean and minimal design to reduce cognitive load. Key elements should be visually distinct, with consistent spacing, and color usage to support quick recognition of actions.

Since it's a browser integration, the user interface should minimally interfere with the visibility of the content on the page. The interface should only capture the user's attention as functionally needed (i.e., listening to user prompts, confirming user intent), but should otherwise blend in with the browser interface.

10.2 Style Requirements

The system shall maintain a professional and neutral visual style suitable for general workplace use. Colors, icons, and fonts should prioritize clarity over branding at this stage. Future iterations may incorporate custom styling or theming.

11 Usability and Humanity Requirements

11.1 Ease of Use Requirements

The system should minimize user effort by providing a simple, intuitive interface. Key actions should be accessible within 3-4 interactions, with clear feedback after each action.

11.2 Personalization and Internationalization Requirements

The system should support basic personalization (e.g., remembering user preferences) and allow easy adaptation for different languages or regions at a later stage. For the PoC, English support is sufficient.

11.3 Learning Requirements

The system should be learnable within 10 minutes without prior training or documentation. Users should be able to complete core tasks on their first attempt through the interfaces navigation tutorial upon first time launch. Additional documentation should be supplemental but not necessary.

11.4 Understandability and Politeness Requirements

The system should use clear, direct, and neutral language in responses. Error messages or clarifications should remain polite and informative.

11.5 Accessibility Requirements

The interface should be navigable using standard assistive tools and offer clear text contrast and legible font sizes. Full accessibility compliance is not

required at the PoC stage but should be feasible for future iterations.

12 Performance Requirements

12.1 Speed and Latency Requirements

Under normal operating conditions, latency requirements can be broken down into:

- Speech interpretation: 5 s after end of utterance

- Command generation & execution: 10-15 seconds after end of speech interpretation

12.2 Safety-Critical Requirements

The system shall enforce guardrails to prevent unsafe or unintended actions, requiring validation and user confirmation for potentially disruptive operations, and provide warnings or fail-safes for errors.

12.3 Precision or Accuracy Requirements

ASR accuracy shall be at least 70% in stationary noise conditions. The system shall achieve at least 80% command recognition precision under stationary noise conditions for the PoC.

12.4 Robustness or Fault-Tolerance Requirements

The system shall remain stable under fluctuating network conditions and noisy input. Fallback mechanisms (e.g., retry logic, error messaging) shall ensure graceful issue handling.

12.5 Capacity Requirements

The system shall support at least 20 concurrent users without service degradation if the product is commercialized.

12.6 Scalability or Extensibility Requirements

The system architecture shall allow horizontal scaling to handle increased traffic and modular extensions for new interaction capabilities.

12.7 Longevity Requirements

The system shall be designed to operate reliably over a minimum of 5 years, with maintainable and updatable components to support long-term product evolution.

13 Operational and Environmental Requirements

13.1 Expected Physical Environment

The product shall be operable in a variety of typical office, home, or institutional environments where users perform their daily tasks. The system shall be robust to stationary background noise such as air conditioning, computer fans, and ambient hum. It is not required to reliably handle non-stationary noise, including multiple people speaking or sudden loud interruptions. No modifications to the host operating system, browser, or network configuration shall be required.

13.2 Wider Environment Requirements

The primary interface shall be web-based, accessible via standard web browsers, to maximize user accessibility and support flexible use cases.

13.3 Requirements for Interfacing with Adjacent Systems

1. The system shall integrate with existing browser-based platforms and may interface with external language interpreter modules.
2. Open-source components (e.g., browser interaction agents and libraries) may be incorporated, ensuring compatibility and maintainability.

13.4 Productization Requirements

1. The product shall be deployable for multiple users within an organization, supporting secure user accounts and personalized ASR profiles.
2. The design shall allow packaging and distribution without requiring technical setup by end users. The product shall have straightforward installation or access via pre-configured web access and automatic model initialization.
3. Productization shall include logging suitable for monitoring performance and usage in a hosted environment.
4. The system shall include mechanisms for updates to features with minimal disruption to users.

13.5 Release Requirements

The product shall follow a defined release cycle, providing minor updates quarterly and major updates semi-annually. Each release must maintain backward compatibility with user data, personalization settings, and existing features. Release planning will account for maintenance effort, compute resources, and compliance obligations.

14 Maintainability and Support Requirements

14.1 Maintenance Requirements

Code must be modular, documented, and testable to support scalability, debugging, and future updates. Maintenance must be possible by developers who were not the original authors.

14.2 Supportability Requirements

The product shall provide an accessible Help Page and a Frequently Asked Questions (FAQ) section with clear instructions. It shall be displayed with high-contrast visuals, simple language, and auditory and visual aids.

14.3 Adaptability Requirements

The system must run on common workplace platforms via a web browser, including Windows 10 or later, macOS 12 Monterey or later, and Linux distributions such as Ubuntu 20.04 LTS or later, supporting modern web browsers.

15 Security Requirements

15.1 Access Requirements

Only authorized end users shall be able to access personalized ASR features, voice command execution, and saved transcripts. Access shall be role-based:

- **Primary users (end users):** can access their own data.
- **Secondary users (supervisors/SMEs):** may have read-only access to assist with support or troubleshooting.
- **Tertiary users (caregivers):** may have limited access to assist the end user, can view basic usage history and transcripts but cannot modify settings.

Access to external services (e.g., LLM APIs) shall be rate-limited to ensure system stability, control costs, and prevent abuse.

15.2 Integrity Requirements

The system must provide real-time confirmation and validation of commands before execution to ensure they match user intent.

15.3 Privacy Requirements

The system must gather explicit user consent for storing voice and personal data. All user data used for model improvement must be anonymized. Database and personalized ASR models must maintain integrity through secure, versioned backups. Unauthorized changes or corruption of user data must be prevented.

15.4 Audit Requirements

The system shall maintain secure logs of major actions and commands, including logging events, access to profile and personal data, and command execution failures. Logs shall be protected and retained in a secure database.

15.5 Immunity Requirements

The system must be resilient to accidental misuse. It shall handle noisy input robustly, avoid executing unintended commands, and operate safely within rate limits to prevent resource overload.

16 Cultural Requirements

16.1 Cultural Requirements

The system shall maintain a culturally neutral and respectful tone when prompting users, avoiding slang, bias, and discriminatory language. It must include ethical guardrails to prevent the generation of harmful content or execution of potentially dangerous commands.

It shall support inclusive and accessible design to serve users across diverse cultural backgrounds.

17 Compliance Requirements

17.1 Legal Requirements

The project shall comply with the Personal Information Protection and Electronic Documents Act (PIPEDA) regarding the collection, storage, and handling of personal information.

17.2 Standards Compliance Requirements

The application must comply with the Web Content Accessibility Guidelines (WCAG) 2.0, Level AA guidelines to ensure usability by individuals with disabilities.

18 Open Issues

The primary open issue is maintaining high ASR accuracy across the diverse and severe spectrum of dysarthric speech. This requires extensive training and validation data, as well as careful model tuning.

19 Off-the-Shelf Solutions

19.1 Ready-Made Products

The project will evaluate existing, specialized speech recognition models and applications (e.g., Whisper model and Project Euphonia) as performance baselines.

19.2 Reusable Components

Potential reused components could include existing Text-to-Speech (TTS) modules for feedback and LLMs to map user commands into structured actions.

19.3 Products That Can Be Copied

Open-source browser automation agents (e.g., The AI browser agent) may be integrated for command execution via voice input.

20 New Problems

20.1 Effects on the Current Environment

The product shall operate without modifying the user's OS, browser, or network configuration. This separation prevents unintended impact from incorrect commands.

20.2 Effects on the Installed Systems

The product shall not bypass firewalls, alter security settings, access banned sites, install untrusted content, or perform any malware execution.

20.3 Potential User Problems

Due to the non-deterministic nature of dysarthric speech patterns, a user's exact speech type may not be fully captured by the model, leading to higher training overhead for personalization before the system becomes reliably usable for them. Users may experience misinterpretations requiring retries, which can cause frustration.

20.4 Limitations in the Anticipated Implementation Environment That May Inhibit the New Product

Variability in dysarthric speech may necessitate frequent retraining of ASR models.

Real-time performance varies between devices, depending on processing power and low-latency operation.

20.5 Follow-Up Problems

The local operation limits the ability to perform remote problem diagnosis if any issues arise.

21 Tasks

21.1 Project Planning

Breakdown of major tasks:

Our main tasks are centered on preparing the system by training and tuning the dysarthric ASR model, integrating the core command interpreter, designing the accessible interface, and conducting rigorous testing.

21.2 Planning of the Development Phases

A detailed schedule of development phases, milestones, and dependencies is outlined in the PoC and Development Plan.

22 Migration to the New Product

22.1 Requirements for Migration to the New Product

The system shall support new users with no prior ASR and interpreter experience through an onboarding process.

The system shall support transitioning users migrating from an existing system by allowing for easy uploading of recordings of their speech training data.

22.2 Data That Has to be Modified or Translated for the New System

All previous transcripts, audio files, and personalization data shall be tied to secure user accounts.

Users shall be able to access their data from any supported device after logging in and authentication.

23 Costs

The primary costs for this product are related to machine learning computation and hosting. Development and training will initially leverage Compute Canada credits provided by the supervising research team, along with Colab Pro (CAD 13.99/month).

No external hosting costs are anticipated for the initial release; however, optional hosting may be required for commercial distribution and deployments. If hosting is included, approximate costs for cloud-based deployment on Google Cloud Platform (GCP) include: training and inference on GPUs (\$50 CAD/month for pilot use), storage (\$3 CAD/month), and minimal networking (\$2–5 CAD/month). Costs scale with the number of users and training frequency. All estimates are approximate and intended for planning purposes.

24 User Documentation and Training

24.1 User Documentation Requirements

The system shall provide clear, accessible documentation, including an instructional guide and FAQ. (As noted in Section 14b, supporting requirements already cover this in detail.)

24.2 Training Requirements

Training shall not require formal instruction; the documentation and interface shall support self-directed onboarding.

25 Waiting Room

The system shall display launch, loading, or processing pages during transitions to provide users with clear feedback and reduce confusion during waiting periods.

26 Ideas for Solution

26.1 Browser-Based Extensions

The product is primarily web-based. Future considerations could include desktop or mobile apps to broaden the scope and accessibility (see Section 13 for environment and interface requirements).

26.2 Personalized ASR Fine-Tuning

Consider incremental model tuning paired with real-time streaming of speech input. This could be an approach to adapt to individual speech patterns while minimizing the training time overhead, related to the adaptability requirements discussed in earlier sections.

26.3 LLM Command Mapping

User intent parsing via context-aware LLMs could improve natural language command interpretation, as mentioned in Section 13.

26.4 Noise Filtering

Beyond stationary noise, adaptive filtering techniques could be adapted to enhance recognition in busier environments. Section 13 already specifies stationary noise handling.

26.5 Open-Source Integrations

Potential use of browser automation or voice control frameworks can be investigated, as mentioned in Section 16 for off-the-shelf solutions.

26.6 Accessibility Enhancements

Additional visual, auditory, or haptic cues for feedback and help features could enhance usability. Section 14 highlights basic supportability requirements that these enhancements would build on.

26.7 Data Encryption

Anonymized logging and privacy approaches can balance model improvement with privacy in mind. Section 15 discusses privacy requirements in detail.

Appendix — Reflection

The purpose of reflection questions is to give you a chance to assess your own learning and that of your group as a whole, and to find ways to improve in the future. Reflection is an important part of the learning process. Reflection is also an essential component of a successful software development process.

Reflections are most interesting and useful when they're honest, even if the stories they tell are imperfect. You will be marked based on your depth of thought and analysis, and not based on the content of the reflections themselves. Thus, for full marks we encourage you to answer openly and honestly and to avoid simply writing "what you think the evaluator wants to hear."

Please answer the following questions. Some questions can be answered on the team level, but where appropriate, each team member should write their own response:

1. What went well while writing this deliverable?
2. What pain points did you experience during this deliverable, and how did you resolve them?
3. How many of your requirements were inspired by speaking to your client(s) or their proxies (e.g. your peers, stakeholders, potential users)?
4. Which of the courses you have taken, or are currently taking, will help your team to be successful with your capstone project.
5. What knowledge and skills will the team collectively need to acquire to successfully complete this capstone project? Examples of possible knowledge to acquire include domain specific knowledge from the domain of your application, or software engineering knowledge, mechatronics knowledge or computer science knowledge. Skills may be related to technology, or writing, or presentation, or team management, etc. You should look to identify at least one item for each team member.
6. For each of the knowledge areas and skills identified in the previous question, what are at least two approaches to acquiring the knowledge or mastering the skill? Of the identified approaches, which will each team member pursue, and why did they make this choice?