Software Requirements Specification for Bridging Gaps: AI for Diagram Accessibility: subtitle describing software

Team 22, Reading4All
Nawaal Fatima
Dhruv Sardana
Fiza Sehar
Moly Mikhail
Casey Francine Bulaclac

October 9, 2025

Contents

Revision History

Date	Version	Notes
Date 1	1.0	Notes
Date 2	1.1	Notes

1 Purpose of the Project

1.1 User Business

Insert your content here.

1.2 Goals of the Project

Insert your content here.

2 Stakeholders

2.1 Client

Insert your content here.

2.2 Customer

Insert your content here.

2.3 Other Stakeholders

Insert your content here.

2.4 Hands-On Users of the Project

Insert your content here.

2.5 Personas

Insert your content here.

2.6 Priorities Assigned to Users

Insert your content here.

2.7 User Participation

Insert your content here.

2.8 Maintenance Users and Service Technicians

Insert your content here.

3 Mandated Constraints

3.1 Solution Constraints

MD-SL 1. *The solution design must comply with at least the Level AA of the Web Content Accessibility Guidelines (WCAG) 2.1 standards*

Rationale: This ensures that the solution ensures inclusivity for users with visual, auditory, or cognitive impairments
Fit Criterion: The solution must past all tests using WCAG automated testing tools and manual tests.
Priority: High

MD-SL 1. *The solution must be implemented as a website*

Rationale: A website will allow for automated testing against the WCAG standards which ensures accesibility for users and allow users to upload images/figures to generate alternative text
Fit Criterion: The website must be functional and allow users to generate alternative text by uploading images and figures
Priority: High

MD-SL 1. *The solution must support common image formats (e.g. JPEG, PNG, etc.)*

Rationale: The website will enable users to upload images or figures to generate the alternative text, therefore the solution must be able to handle the different types of image formats
Fit Criterion: The product must successfully process at least one image of each required format including JPEG and PNG images and figures
Priority: High

3.2 Implementation Environment of the Current System

MD-IE 1. *The product must be able to run on standard laptop environments, including operatins systems (OS) such as macOS, Windows, and Linux*

Rationale: This ensures that the product is compatible with major operating systems to allow the product to be accessible to users, regardless of their laptop environment
Fit Criterion: The product must successfully install and operate on the latest releases of macOS, Windows, and Linux, verified through installation and functionality testing on each OS
Priority: High

3.3 Partner or Collaborative Applications

MD-PA 1. *The product must not conflict with other accessibility tools (e.g. screen readers, screen magnifiers, dictation software)*

Rationale: This is to ensure that the product does not limit or intefere with other accessibility tools that meets the users' needs
Fit Criterion: The product must operate simultaneously with at least one other accessibility tool, verified through interoperability testing
Priority: High

3.4 Off-the-Shelf Software

There are a number of existing AI generated alternative text off-the-shelf software in the market today. The following highlights a few of these tools, including their functions, benefits, and limitations:

1. **Azure AI Vision Image Analysis**: This service by Microsoft can extract a wide variety of visual features from images. Image Analysis offers image captioning models that generate one-sentence descriptions of an image's visual content. Limitations of this product is that it only generates one simple sentence, and that the image captions are only available in English.

2. **ALTTEXT.AI**: This service allows users to upload images and generate alternative text. The website supports over 100 languages and many modern image formats. A significant limitation of this project is that it doesn't guarantee compliance with WCAG which limits accessibility.

3. **accessiBe**: This service is an accessibility platform built for developers and engineers that plugs into their SDLC to detect and remediate WCAG issues at code level. The tools offers AI alt-text descriptions for images and allows users to review and edit the alt text. A limitiation of this tool is that it uses overlays that sit on top of a website to fix issues at run-time. This is an issue because overlays can conflict with assistive technologies and miss context-specific WCAG requirements creating a false sense of real accessibility compliance.

3.5 Anticipated Workplace Environment

The anticipated workplace environment for this product is academic settings such as universities, where students may require alternative text to interpret images and figures within their coursework and study materials.

3.6 Schedule Constraints

MD-SC 1. *The final product must be completed and tested by the end of the academic term (April 2026)*

Rationale: This is to ensure that the final product is functional and meets all requirements at the end of the academic year
Fit Criterion: All deliverables are submitted, and the final product is tested and operable by April 2026
Priority: High

3.7 Budget Constraints

MD-BC 1. *The project budget must include compensation for user testers, set at maximum $150 per participant for two rounds of usability testing.*

Rationale: This is to ensure that user testers are compensated

for their meaningful feedback, and that our testing aligns with ethical practices
Fit Criterion: There must be record of participants being compensated between the range of $100 and $150 for two rounds of testing.
Priority: High

3.8 Enterprise Constraints

MD-EC 1. *The product must comply with the Accessibility for Ontarians with Disabilities Act (AODA)*

Rationale: This ensures that the product meets the legal requirements in Ontario and guarantees that the product is accessible to users with diverse needs
Fit Criterion: AODA requires compliance with WCAG standards, which ensures that the product meets AODA regulations. Compliance must be verified through both automated WCAG testing tools and manual accessibility testing.
Priority: High

4 Naming Conventions and Terminology

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

Insert your content here.

5 Relevant Facts And Assumptions

5.1 Relevant Facts

- This project is being developed for a Software Engineering Capstone course with a fixed timeline
- The solution is targeted primarily for laptop and/or desktop environments, but can later be extended for mobile platforms use

5.2 Business Rules

The business rules established among the team are as follows:

- **Adherence to Project Schedule**: All deliverables and milestones must be completed according to the established project schedule. Any anticipated delays must be communicated in advance.
- **Pull Request Requirement**: All pull requests made by a team member must be reviewed by three other members before being merged into the `main` branch. The reviewers must provide approval or feedback within 24 hours of the pull request.
- **Team Communication Standard**: All team members must communicate respectfully and professionally during all discussions, meetings, and written communication.
- **Testing Requirements**: All code contributions must include appropriate unit, intergration, and functionality tests to ensure correctness and reliability. Accessibility testing must also be performed for all product features.

5.3 Assumptions

The following assumptions are made when using the product:

- Users will be operating on modern browsers such as Chrome, Safari, and Firefox
- Users will have access to stable internet connection when installing and using the product
- Users will have basic knowledge of installing and enabling web extensions

6 The Scope of the Work

6.1 The Current Situation

Currently, alternative text generation tools are able to provide sufficient descriptions for simple images and figures. However, for more complex visuals

such as engineering diagrams, the generated alt text is often misleading, incomplete, or inefficient at conveying the intended meaning.
Accurate alternative text is particularly essential for individuals with visual or cognitive impairments, as it enables fair access to academic content. Without reliable descriptions, students may experience barriers to learning and miss critical information conveyed in diagrams and figures.
The current limitations of existing generated alternative text tools are as follows:

- **Inaccurate Alternative Text**: Generated alt text may emphasize unimportant details and overlook key elements, resulting in misleading or confusing interpretations.
- **Oversimplification of Complex Figures**: Current tools frequently oversimplify technical or academic diagrams, failing to capture essential details required for learning.
- **High Manual Effort**: In many cases, subject matter experts must manually create alt text, which is time-intensive and not scalable across large volumes of academic content.

6.2 The Context of the Work

The product will be in the form of a website that integrates into existing accessibility workflows by providing accurate descriptions from images that can be read aloud by screen readers. The product will complement existing screen readers by ensuring accurate generated alternative text from uploaded images and figures of academic work are available. Figure1 shows how the product will integrate with existing screen readers.

6.3 Work Partitioning

Table 1 shows the work partitioning for completing the project. It includes major events, their inputs and outputs, and the summary of the event.

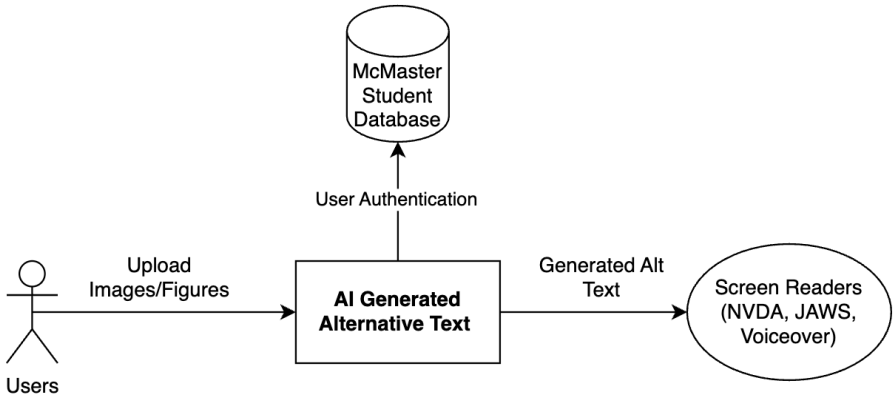

Figure 1: Work Context Diagram

Table 1: Work Partition for the System

Event Name	**Input**	**Output**	**Summary**
Login	Username, Password		User logs in using their McMaster account
Upload Images/Figures	PNG/JPEG files	Uploaded File Reference	User uploads their files to generate alternative text
OCR Text Extraction	Uploaded Images/Figures	Detected Text	System reads the text embedded in the uploaded files
Generate Alternative Text	Uploaded Images/Figures, Extracted OCR, Model Parameters	Generated Alt Text, Quality Metric from ML Model	System analyzes the image and extracted OCR data to generate accurate alternative text
View History	User Login, Stored Uploads and Generated Alt Text	List of Previously Generated Alt Text	The system retrieves and displays a user's history of uploaded images along with their associated generated alt text

6.4 Specifying a Business Use Case (BUC)

The project has one primary business use case, which aims to achieve the goal of providing users with visual and cognitive impairments an efficient and accessible way to generate accurate alternative text for academic images and figures.

Preconditions:

- The user has access to the website
- The user has files containing diagrams or images requiring alternative text

Scenario:

1. The user logs into the system using their McMaster student credentials
2. The user uploads one or more files (PNG, JPEG) containing diagrams
3. The AI model analyzes the uploaded file(s), performs OCR to extract any visible text, and generates alternative text describing each image/figure accurately
4. Screen readers use the generated alternative text to read aloud and convey the uploaded image
5. The generated alternative text can be copied to clipboard or dowloaded as a `.txt` file by the user if needed
6. The user can view previously uploaded files and generated alternative text for future reference

Postcondition:

- The user obtains accurate and accessible alternative text that complies with AODA and WCAG 2.1 standards.

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

The diagram below shows the components within the system and how they connect. The components that this project will aim on building include a user interface, an alternative text generation ML model, a session manager. Furthermore, these components will utilize or communicate with a screen reader software, McMaster Authentication system and external AI/ML Frameworks.

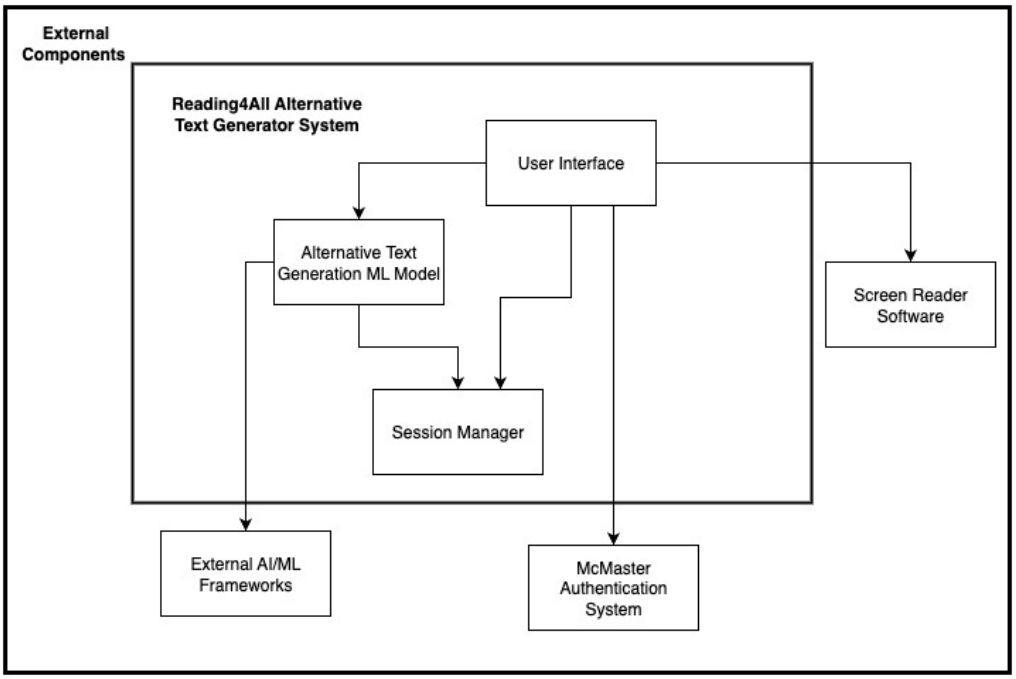

Figure 2: Product Boundary Diagram

8.2 Product Use Case Table

The table below discusses the different use cases.

Table 2: Labels and Their Usage

PUC#	PUC Name	Actor(s)	Input & Output(s)	Requirement
PUC 1	Login Using McMaster Credentials	McMaster Student and Faculty , McMaster Authentication System	User Credentials (input), Authentication Results (output)	FR 7
PUC 2	Upload Image	McMaster Student and Faculty	JPEG or PNG (input), Image upload status (output)	FR 1, UHR-EUR 3
PUC 3	Generate Alternative Text	McMaster Student and Faculty, Alternative Text Generation Model	Uploaded image (input), Generated alternative text	FR 2, FR 3
PUC 4	Copy or Download Text	McMaster Student and Faculty	User decision to copy or download (input), text copied to clipboard or downloaded	UHR-PIR 1
PUC 5	View History of Inputted Images and their Alternative Text	McMaster Student and Faculty	User request to view history (input), Display of previously inputted images and their generated text within a session	FR 5

8.3 Individual Product Use Cases (PUC's)

PUC 1: Login Using McMaster Credentials
Trigger: User selects "Login" and is directed to McMaster sign in page.
Preconditions:

- The user is registered person in McMaster system and has valid credentials.

Actors: McMaster Student or Faculty, McMaster Authentication System.
Outcome: McMaster validates user's credentials are validated and they are

given access to system.
Input: McMaster System username and password.
Output: User enters system or an error message is displayed.

PUC 2: Upload Image
Trigger: User selects "Upload Image" and chooses a file
Preconditions:

- User successfully logged into the system.

Actors: McMaster Student or Faculty
Outcome: The selected image is uploaded and stored for later text generation.
Input: Image file (JPEG or PNG)
Output: A confirmation message is displayed if the image was successfully uploaded, or an error message otherwise is displayed.

PUC 3: Generate Alternative Text
Trigger: User selects *Generate Alternative Text* for an uploaded image.
Preconditions:

- A valid image has been successfully uploaded to the system.

Actors: McMaster Student or Faculty.
Outcome: The system generates a descriptive alternative text for the uploaded image.
Input: User selection to generate alternative text
Output: Generated alternative text is displayed to the user.

PUC 4: Copy or Download Generated Alternative Text
Trigger: User selects *Copy* to clipboard or *Download* .txt after generating alternative text.
Preconditions:

- System has successfully generated alternative text.
- User is satisfied with generated alternative text and has made any desired changes.

Actors: McMaster Student or Faculty.
Outcome: The user receives the alternative text through their preferred

method.
Input: User decision to copy or download.
Output: Text is copied to clipboard or downloaded as .txt file on the users device.

PUC 5: View History of Uploaded Images and Generated Alternative Text
Trigger: User selects *View History* option.
Preconditions:

- User is logged in with an active session.
- User has previously uploaded at least one image and generated text within the session.

Actors: McMaster Student or Faculty.
Outcome: The user views a list of their images and the corresponding generated alternative text within the current session.
Input: User request to view session history.
Output: Display of uploaded images and their corresponding generated alternative text.

9 Functional Requirements

9.1 Functional Requirements

FR 1. *The system must accept technical diagrams in the format of JPEG and PNG.*

Rationale: The system must process JPEG/PNG images in order to output alternative text.
Fit Criterion: The system successfully takes as accepts JPEG/PNG images and provides feedback to users when an invalid file type is inputted.
Priority: High

FR 2. *The system shall generate alternative text of uploaded gyms.*

Rationale: The main purpose of the system is to make scientific diagrams more accessible by generating better alternative-text.
Fit Criterion: For a set of test diagrams, the alternative text generated must meet the pre-determined criteria.
Priority: High

FR 3. *The system shall output alternative text in a format readable by screen readers.*

Rationale: Students with disabilities utilize screen readers to access digital content; therefore, the alternative text must be displayed in away that enables screen readers to read it correctly. Furthermore, if the alternative text output format is not compatible with screen readers, then students cannot benefit from the application output.
Fit Criterion: The alternative text output must be readable by at least three commonly used screen readers.
Priority: High

FR 4. *The system shall allow users to edit the outputted alternative texts.*

Rationale: Providing users with an option to edit the outputted text, enables them to adjust the output to better meet their needs if needed.
Fit Criterion: Users can add or delete text in any part of the outputted alternative text and save their changes.
Priority: High

FR 5. *The system shall store and display all inputted images and their generated alternative texts within a session.*

Rationale: Storing previously inputted images and their generated alternative texts, allows users to easily review or reuse them without re-uploading.
Fit Criterion: Users can see view all previously inputted images with their generated alternative texts during the same session.
Priority: Medium

FR 6. *The system must accept keyboard input for navigation.*

Rationale: Many users, including those with disabilities, use keyboard inputs to navigate through applications, the system must support this as a way to navigate.

Fit Criterion: Users can navigate to all the main functions and areas of the system using their keyboard.
Priority: High

FR 7. *The system must validate users during login to confirm they are McMaster University students.*

Rationale: User verification will ensure that only McMaster University students have access to the system, ensuring that the system is used by the intended users.
Fit Criterion: Users can only gain access to the system features after their McMaster University credentials are successfully validated.
Priority: High

10 Look and Feel Requirements

10.1 Appearance Requirements

LFR-AR 1. *The system must allow all text on the interface to be resized up to 200 %, without any loss of functionality or content.*

Rationale: Allowing text resizing will enable users with low vision to more easily utilize the system. This also ensures the system meets WCAG Success Criterion 1.4.4 Resize Text. User verification will ensure that only McMaster University students have access to the system, ensuring that the system is used by the intended users.
Fit Criterion: All text, excluding any captions and images of text can be enlarged to 200 % on a standard browser zoom (ex. Google Chrome) without any overlapping, hidden content, or broken features.
Priority: High

LFR-AR 2. *The system must not use color as the only method to provide information, indicate actions or prompt user input.*

Rationale: Users with color vision deficiencies or other visual impairments may not detect color differences accurately. This also ensures the system meets WCAG Success Criterion 1.4.1 Use of Color.

Fit Criterion: Any use of color communicates information to the user or requests information from he user must be appear with text.
Priority: High

LFR-AR 3. *The system must ensure sufficient contrasts of text and images of text.*

Rationale: Sufficient color contrast is important as it enables users with low vision or color vision deficiencies to easily read any system text. This also ensures the system meets WCAG Success Criterion 1.4.3 Contrast (Minimum)
Fit Criterion: All text and images of text in the system interfaces has a contrast ratio of at least 4.5:1
Priority: High

LFR-AR 4. *The system must provide alternative text for all non-text content..*

Rationale: Users with visual impairment often use screen readers to navigate through software systems; therefore, it is essential that all images have sufficient alternative text, so that the purpose of the images can understood. This also ensures the system meets WCAG Success Criterion 1.1.1 Non-text Content .
Fit Criterion: All images and non-text element are joined with descriptive alternative text that communicate their meaning.
Priority: High

10.2 Style Requirements

LFR-SR 1. *The system interface must follow a simple and modern design style.*

Rationale: A simple interface will improve the systems usability as it better highlights the system's features, while also ensuring the system is visually appealing.
Fit Criterion: The system uses a clean layout with a maximum of three colors, consistent font styles and sizes, as well as only has key design elements that support usability.
Priority: High

LFR-SR 2. *The system interface must use McMaster University branding while maintaining accessibility standards and a modern style.*

Rationale: As the system is targeted towards McMaster University students, using the schools branding will build trust with users and ensure the system aligns with McMaster's identify. However, using McMaster branding must not interfere with usability and accessibility criteria..
Fit Criterion: The system interface includes McMaster University' official logo and meets the WCAG 2.1 contrast and non-text content success criteria.
Priority: High

11 Usability and Humanity Requirements

11.1 Ease of Use Requirements

UHR-EUR 1. *The system interface must allow users to efficiently use the system features.*

Rationale: It is important the users can quickly access and use the system features, as they may be generating multiple alternative text outputs in a single session.
Fit Criterion: Users can upload images to the system and generate alternative text in 5 steps or fewer.
Priority: High

UHR-EUR 2. *The system interface must be easy for users to remember how to use after not using it for some time.*

Rationale: Users should be able to quickly recall how to use the system without needing to relearn the features. An intuitive design will make it easier for returning users to find and use key features.
Fit Criterion: Users who have not used the system in a month, can successfully login, upload an image and generate alternative text within 5 minutes, without needing any assistance.
Priority: Medium

UHR-EUR 3. *The system interface must provide users with clear and immediate feedback for all actions.*

Rationale: Providing the users with feedback ensures they understand of the outcome of their actions and whether they are using the system correctly. This reduces confusion and makes users more confident while using the system.
Fit Criterion: The system provides textual feedback within 1 second after a user interaction, such as uploading an image.
Priority: High

UHR-EUR 4. *The system interface must provide clear instructions, prevent common errors and allow users to easily correct them.*

Rationale: Providing easy to follow instructions will help ensure that users can easily use the system features and prevent errors. Additionally, if a user makes a mistake, they should easily be able to revert it.
Fit Criterion: In user testing, at least 80% of users can complete tasks without errors. When a user error occurs, the system explains the issue and how to recover within 2 seconds.
Priority: High

11.2 Personalization and Internationalization Requirements

UHR-PIR 1. *The system interface must allow users to choose how generated alternative text is stored or copied.*

Rationale: Providing users with the option to either copy generated text to their clipboard or download it as file, helps tailor the output to the users specific needs.
Fit Criterion: After generating the alternative text users can choose to "Copy to Clipboard" or "Download as .txt" from the interface and system successfully completes the chosen option.
Priority: High

11.3 Learning Requirements

UHR-LR 1. *The system must be easy for low-vision users to learn and operate with screen readers.*

Rationale: The system should be intuitive for users with low vision to use without prior training. Additionally, the system being highly compatible with screen readers, allows users to more easily navigate and use the system.
Fit Criterion: In user testing, at least 90% of first time users with low vision using a screen reader can upload an image and generate alternative text within 5 minutes without assistance.
Priority: High

11.4 Understandability and Politeness Requirements

UHR-LR 1. *The system must be only display essential information and hide all technical details.*

Rationale: The system should only communicate the information needed to use the system. Displaying any technical details may cause the user to be confused and make the system less usable.
Fit Criterion: In user testing, users do not encounter any technical terms, code outputs or information that is not relevant to them.
Priority: High

11.5 Accessibility Requirements

UHR-AR 1. *The system must meet the WCAG 2.1 Level AA accessibility standards.*

Rationale: The Accessibility for Ontarians with Disabilities Act (AODA) requires organizations to meet WCAG 2.0 Level AA for websites. Therefore, meeting WCAG 2.1 Level AA ensures the system meets AODA standards and is accessible for users with disabilities.
Fit Criterion: The system will be evaluated using an accessibility testing tool such as Pope Tech and Wave Web Aim to

ensure WCAG 2.1 criteria is met.
Priority: High

12 Performance Requirements

12.1 Speed and Latency Requirements

PR-SL 1. *The tool shall generate alt-text for uploaded images within a reasonable time frame.*
Rationale: Ensures users, including those using assistive technologies, do not experience delays that hinder accessibility.
Fit Criterion: The system shall return generated alt-text within **3 seconds** for images $\leq$2 MB and within **8 seconds** for images $\leq$10 MB under normal load conditions.
Priority: High

PR-SL 2. *The web interface shall load and render accessibility components efficiently.*
Rationale: Improves user experience and responsiveness for screen-reader users and keyboard navigation.
Fit Criterion: All interactive elements shall respond within **300 ms** of user input under typical conditions.
Priority: Medium

12.2 Safety-Critical Requirements

PR-SCR 1. *The tool shall ensure that no personally identifiable data from uploaded images is stored or shared without consent.*
Rationale: Protects user privacy and adheres to ethical AI standards.
Fit Criterion: Uploaded images are deleted from temporary storage within **60 seconds** of processing unless explicitly saved by the user.
Priority: High

PR-SCR 2. *The tool shall not produce alt-text containing offensive, biased, or harmful language.*
Rationale: Ensures ethical AI output and inclusivity.
Fit Criterion: 0 % of generated outputs shall contain content flagged

by moderation filters as offensive or biased.
Priority: High

PR-SCR 3. *The interface shall adhere to WCAG 2.1 Level A accessibility guidelines to prevent stress or strain on users' eyes and ensure comfortable interaction.*
Rationale: Provides a visually safe, inclusive experience for all users, including those with visual or cognitive impairments.
Fit Criterion: Verified through front-end accessibility testing that confirms conformance with WCAG 2.1 Level A success criteria.
Priority: High

12.3 Precision or Accuracy Requirements

PR-PAR 1. *The generated alt-text shall adequately describe the image content with minimal omissions or irrelevant details.*
Rationale: Ensures the description fulfills its accessibility purpose.
Fit Criterion: At least **85 %** of outputs rated "Sufficient" or better on the sufficiency scale by testers.
Priority: High

PR-PAR 2. *The alt-text shall maintain appropriate length and readability.*
Rationale: Prevents overly short or verbose outputs that reduce usability.
Fit Criterion: ≥ 90 % of outputs rated "Proper Length" on the user-testing scale.
Priority: Medium

PR-PAR 3. *The overall accessibility and usability of the alt-text shall be acceptable to testers.*
Rationale: Evaluates real-world effectiveness of generated descriptions.
Fit Criterion: Median user rating ≥ 3 ("Mostly Accessible/Usable") on the 0–3 or 0–4 scales; no outputs below 2.
Priority: Medium

12.4 Robustness or Fault-Tolerance Requirements

PR-RFT 1. *The system shall gracefully handle unsupported or corrupted image inputs.*
Rationale: Prevents crashes and maintains system stability.
Fit Criterion: Invalid files trigger a clear error message within 2 seconds without interrupting service.
Priority: High

PR-RFT 2. *The backend shall recover automatically from isolated process failures.*
Rationale: Ensures continued operation without developer intervention.
Fit Criterion: System recovers within 5 seconds after fault detection.
Priority: High

12.5 Capacity Requirements

PR-CR 1. *The system shall support limited concurrent usage suitable for a proof-of-concept deployment.*
Rationale: Demonstrates feasibility and reliability for initial testing without production-level scaling.
Fit Criterion: Supports at least **5 simultaneous requests** with response times $\leq$ 10 seconds.
Priority: Medium

PR-CR 2. *Storage shall accommodate pilot testing datasets.*
Rationale: Ensures smooth prototype validation without capacity issues.
Fit Criterion: The system can temporarily store metadata for up to **500 images per day** without data loss.
Priority: Low

12.6 Scalability or Extensibility Requirements

PR-SER 1. *The architecture shall allow integration of improved ML models or multilingual capabilities in future phases.*
Rationale: Enables progressive enhancement and future accessibility expansion.
Fit Criterion: New models or language modules can be incorporated

without restructuring existing components.
Priority: Medium

12.7 Longevity Requirements

PR-LR 1. *The codebase shall be maintainable and adaptable to updates in WCAG guidelines, Python libraries, and ML frameworks.*
Rationale: Ensures long-term usability and compliance even after the pilot phase.
Fit Criterion: Minor updates or migrations require ≤ 2 person-days per quarter.
Priority: Medium

PR-LR 2. *The prototype shall maintain compatibility with at least the next two Python releases.*
Rationale: Ensures sustainability of the pilot for educational and testing purposes.
Fit Criterion: Verified through annual testing on supported Python versions.
Priority: Low

13 Operational and Environmental Requirements

13.1 Expected Physical Environment

Insert your content here.

13.2 Wider Environment Requirements

Insert your content here.

13.3 Requirements for Interfacing with Adjacent Systems

Insert your content here.

13.4 Productization Requirements

Insert your content here.

13.5 Release Requirements

Insert your content here.

14 Maintainability and Support Requirements

14.1 Maintenance Requirements

Insert your content here.

14.2 Supportability Requirements

Insert your content here.

14.3 Adaptability Requirements

Insert your content here.

15 Security Requirements

15.1 Access Requirements

SR-AR 1. *The system shall restrict access exclusively to McMaster University users through institutional Single Sign-On (SSO) authentication.*
Rationale: Restricting access ensures only authorized users within McMaster can use the system during the pilot phase, reducing the risk of unauthorized use or data exposure.
Fit Criterion: All users must log in using verified McMaster SSO credentials before accessing the platform. Unauthenticated requests are automatically rejected.
Priority: High

SR-AR 2. *All actions performed by users shall be tied to their authenticated session.*

Rationale: Linking actions to a user's authenticated identity enables traceability and controlled access to system features.
Fit Criterion: Each upload or alt-text generation event is associated with a unique McMaster user ID through SSO session tracking.
Priority: Medium

15.2 Integrity Requirements

SR-IR 1. *All communication between the frontend, backend, and machine learning services shall use encrypted HTTPS (TLS 1.2 or higher).*
Rationale: Encryption prevents interception and tampering of sensitive data such as authentication tokens or image files.
Fit Criterion: All HTTP requests must be redirected to HTTPS; unencrypted requests are rejected by the web server.
Priority: High

SR-IR 2. *Uploaded images shall remain unmodified during processing and analysis.*
Rationale: Preserving file integrity ensures consistent and accurate generation of alt-text.
Fit Criterion: File hash comparison verifies that image files remain identical throughout the upload and analysis process.
Priority: High

15.3 Privacy Requirements

SR-PR 1. *Uploaded images shall be deleted immediately after processing is complete.*
Rationale: Protects user privacy and ensures compliance with institutional data governance policies.
Fit Criterion: Uploaded files are stored temporarily in memory or on a secure local directory and deleted within 60 seconds after alt-text generation.
Priority: High

SR-PR 2. *Generated alt-text shall not contain personally identifiable information (PII) or sensitive content.*
Rationale: Prevents disclosure of private information and ensures responsible AI usage.

Fit Criterion: The model output is passed through a content moderation filter that rejects or flags any alt-text containing PII or inappropriate language.
Priority: Medium

15.4 Audit Requirements

SR-AU 1. *System usage logs shall record authentication events, uploads, and generation activities for accountability and debugging.*
Rationale: Audit logs enable traceability, assist in debugging, and ensure ethical research practices.
Fit Criterion: Logs record timestamps, user IDs, and non-sensitive metadata while excluding image or generated text content.
Priority: Medium

SR-AU 2. *Access to audit logs shall be restricted to authorized project administrators.*
Rationale: Limits access to potentially sensitive operational data and protects user confidentiality.
Fit Criterion: Logs are stored in a restricted-access directory with read permissions granted only to administrators.
Priority: Medium

15.5 Immunity Requirements

SR-IM 1. *The system shall validate and sanitize all uploaded files to prevent malicious or unsupported file types.*
Rationale: Protects against injection attacks, corrupted uploads, or execution of non-image files.
Fit Criterion: Only files with valid image types (.png, .jpg, .jpeg, .svg, .webp) are accepted; unsupported or script files are automatically rejected.
Priority: High

SR-IM 2. *The system shall block access from networks or domains outside McMaster University's infrastructure.*
Rationale: Restricting network access minimizes exposure to external threats during the proof-of-concept phase.

Fit Criterion: Requests must originate from verified McMaster SSO tokens or IP ranges associated with university networks.
Priority: High

16 Cultural Requirements

16.1 Cultural Requirements

Insert your content here.

17 Compliance Requirements

17.1 Legal Requirements

Insert your content here.

17.2 Standards Compliance Requirements

Insert your content here.

18 Open Issues

This section outlines unresolved questions and decisions that may impact the overall success of the system. The following items require additional research, testing, or discussion to ensure the project's successful completion.

- The ML/AI model architecture the team will use to generate alternative text will need research and testing to ensure optimal accuracy and correctness.
- The optimal length of the generated alternative text requires further research to determine how many characters provide an accurate description without causing confusion or distracting from the main idea of the diagram.

19 Off-the-Shelf Solutions

19.1 Ready-Made Products

Insert your content here.

19.2 Reusable Components

Insert your content here.

19.3 Products That Can Be Copied

Insert your content here.

20 New Problems

20.1 Effects on the Current Environment

Insert your content here.

20.2 Effects on the Installed Systems

Insert your content here.

20.3 Potential User Problems

Insert your content here.

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

Insert your content here.

20.5 Follow-Up Problems

Insert your content here.

21 Tasks

21.1 Project Planning

Insert your content here.

21.2 Planning of the Development Phases

Insert your content here.

22 Migration to the New Product

22.1 Requirements for Migration to the New Product

Insert your content here.

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

Insert your content here.

23 Costs

The total cost of developing this project is primarily based on the effort contributed by the student development team and faculty supervisors. As the project utilizes open-source technologies (e.g., Python, TensorFlow, Flask, and React) and university-hosted infrastructure, no direct monetary expenditure is incurred. The project is scheduled to be completed within the academic term (MVP ready by April 2026), and resource allocation is focused on efficient time management and workload balancing rather than financial cost.

23.1 Metrics for Estimation

To estimate overall development cost in terms of time and effort, the following key metrics have been considered:

- Number of image input/output workflows supported by the tool.

- Number of core functional requirements (e.g., image upload, alt-text generation, user authentication, evaluation metrics).
- Number of non-functional requirements (e.g., accessibility compliance, latency, privacy, and scalability).
- Number of deliverables and milestones within the development timeline.
- Team size and individual role distribution (frontend, backend, model integration, documentation).

23.2 Estimation Approach

Each deliverable has been estimated based on the effort required to implement, test, and document it within the given academic term. The estimates are derived from prior experience with similar web-based machine learning projects and adjusted for the learning and research effort required to integrate accessibility and WCAG compliance features. Time allocation also accounts for model fine-tuning, front-end accessibility testing, and usability evaluation with pilot users.

23.3 Cost Breakdown

• **Development Effort:** Based on a team of four student developers and one faculty supervisor, the estimated time allocation per team member is as follows:

- Initial research, planning, and requirement analysis: **40 hours per team member.**
- Model integration and backend implementation: **120 hours per team member.**
- Frontend development and accessibility compliance: **80 hours per team member.**
- Testing, debugging, and refinement: **60 hours per team member.**
- Documentation and presentation preparation: **30 hours per team member.**

Total estimated effort: 330 hours per team member.
• **Tools and Software:** All software components used in the project (Python, Flask, React, TensorFlow, Machine Learning Models, and GitHub) are open-source or free for academic use. Therefore, there are no direct licensing or software procurement costs.
• **Testing Environment:** Testing will be conducted using McMaster-hosted or open-source platforms for model inference and user testing. The team might decide to conduct user testing and award attendees 10 dollar Amazon.com gift cards as a reward for participation. No additional hardware purchases are required beyond existing student laptops and cloud compute credits provided for academic purposes.

23.4 Estimated Cost

The total development effort is estimated at approximately **1,320 person-hours** across all team members (4 × 330 hours). Assuming an average academic hourly equivalent of **$20 per hour** for estimation purposes, the notional cost of development would be approximately **$26,400 CAD**.

However, as this project is conducted as part of a capstone academic course and leverages free university and open-source resources, the **actual monetary cost is $0 CAD**. The effective cost of the project lies entirely in time, research, and human effort required to meet the performance, accessibility, and security requirements outlined in previous sections.

24 User Documentation and Training

24.1 User Documentation Requirements

1. User Manual

    - **Purpose**: The user manual will serve as a user guide and provide detailed information and instructions on the final product and how to use it effectively
    - **Target Audience**: Academic students, instructors, and other professionals
    - **Content**: Website navigation and instructions, usage examples, product purposes, and best practices

24.2 Training Requirements

Users of the final product will require minimal to no training as we aim to ensure that the tool is as accessible, simple, and intuitive as possible. For any additional guidance, a user manual will be created along with any relevant tutorials on how to use the features within the website.

25 Waiting Room

This section lists potential ideas and features that are out of scope for the current project, however, may be valuable for future updates.

- Support for multilingual alternative text generation (e.g., French and Spanish)
- A browser extension that automatically generates alternative text on websites or learning platforms (e.g. *D2L*) using our model
- Compatibility with mobile platforms to extend accessibility across users' preferred devices

26 Ideas for Solution

This section discusses potential ways to acheive some of the functionality discussed throughout this report, including image upload and processing, alternative text generation and session history. These ideas have been documented so they can be referenced later during development.

Image Upload and Processing This functionality can be achieved through a front-end interface, where users are prompted to upload an image using an upload button or by dragging their file into the drop box. Furthermore, to minimize errors, this will only allow JPEG and PNG image files. Once the image has been uploaded, it will be displayed to the user with the image file name, so users can confirm the correct file was chosen. If the upload fails, the system will display an error message explaining the issue. Furthermore, this can be achieved using the HTML5 File API, which supports reading and processing file data, specifically obtained through input or drag and drop.

Alternative Text Generation This functionality can be implemented using a vision-language model (VLM), which combines natural language models with computer vision. The model can learn from both images and text to solve various problems. The model can be trained using sample scientific images, paired with examples of descriptive alternative text, allow it to generate accurate and high quality descriptions for new images.

Session History After the user is satisfied with the generated alternative text, the system will store the image and its final description in the browser's session storage as a JSON record. This allows the data to be stored temporarily and can easily be displayed to the user when they request their history.

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?