iNTUtion
Handbook

IEEE
NTU Student Branch

Sponsored by:

Jane Street®

5th - 8th FEBRUARY 2026

Table of contents

TIMELINE

02

5TH FEBRUARY, 6PM

OPENING CEREMONY

5TH FEBRUARY, 6.30PM

PROBLEM STATEMENT REVEAL AND HACKING BEGINS!

5TH FEBRUARY, 7.30PM

DINNER

6TH FEBRUARY, 6.30 PM

HACKING ENDS

6TH FEBRUARY, 7PM

ROUND 1 JUDGING BEGINS

TIMELINE

7TH FEBRUARY, 11AM
RELEASE OF FINALISTS

8TH FEBRUARY, 12.30PM
FINALE OPENING CEREMONY AND LUNCH!

8TH FEBRUARY, 1.30PM-4.30PM
PRESENTATION FOR FINAL ROUND BEGINS..

8TH FEBRUARY, 5.30PM
ANNOUNCEMENT OF WINNERS AND CLOSING CEREMONY!

WHAT IS iNTUtion?

INTUtion 2026 is an in-person hackathon focused on building innovative AI-driven solutions to real-world problem statements. Organized by the IEEE NTU Student Chapter, the event brings together passionate student teams to design, prototype, and present impactful ideas. INTUtion aims to help participants sharpen their technical, problem-solving, and pitching skills while gaining exposure to industry-relevant challenges and judging standards.

PROBLEM STATEMENT

Despite rapid digital innovation, many digital interfaces remain inaccessible due to **rigid, one-size-fits-all designs** that fail to adapt to diverse user abilities. **Poor compatibility** with assistive technologies, high cognitive load from **complex interfaces**, and a **lack of multimodal alternatives** prevent users with visual, auditory, motor, or cognitive impairments from using digital tools independently and effectively. These issues are compounded by limited personalization options and inconsistent accessibility support across platforms and updates, turning everyday digital interactions into significant barriers rather than enablers.

Submissions are expected to go beyond concepts or UI mockups and instead provide a functional, end-to-end system, involving multimodal interaction (e.g. text, speech, vision, or UI adaptation), that can be executed, evaluated, and demonstrated live.

PROBLEM STATEMENT

The goal is to let teams clearly define the interface accessibility problem they are addressing, justify the use of AI, and deliver a solution that is usable, testable, and presentable.

Examples:

1. **Adaptive UI mode:** Automatically adjusts layout, contrast, text size, and touch targets to fit different user needs, addressing rigid one-size-fits-all interface designs. (Example Tech stack: React / Next.js, Tailwind CSS, CSS media queries, JavaScript accessibility APIs, WCAG-compliant ARIA attributes, device sensor APIs (where applicable))
2. **Multimodal content converter:** Presents the same information in text, audio, and visual formats, reducing exclusion caused by single-modality content.
3. **Accessibility health checker:** Detects common accessibility issues in real time and suggests fixes, improving compatibility with assistive technologies

*These examples are to support brainstorming and to provide a better overview of the problem statement. They are not concrete ideas that you must choose from.

Judging Criteria

Criterion	Description
Impact (25%)	• Is the specific accessibility barrier (visual, auditory, motor, cognitive) clearly identified and grounded in a real user scenario? • Does the solution directly address a failure of rigid, one-size-fits-all interfaces described in the problem statement? • Does the system measurably reduce user burden (e.g. cognitive load, interaction steps, error rate, time-to-completion)? • Are improvements demonstrated relative to a baseline interface (before vs after)? • Does the solution enable users to complete tasks independently without external assistance? • Does it reduce reliance on manual accessibility settings or static configurations?
Real-time performance & latency (25%)	• How quickly does the system process and respond to multimodal inputs (speech, text, vision, interaction signals)? • Are interactions fast enough to avoid breaking usability for users with impairments? • Are latency–accuracy trade-offs explicitly justified for accessibility (e.g. faster feedback over perfect transcription)? • Does the system degrade gracefully under slower networks or weaker devices? • Does the system remain stable during continuous interaction, not just single requests? • Can it handle concurrent users or repeated interactions without failure?
Design (25%)	• Is this a fully functional pipeline, not disconnected components? • Are AI models, backend services, and UI layers cleanly integrated? • Does the interface adapt dynamically to user needs rather than relying on fixed accessibility settings? • Are adaptations understandable, predictable, and reversible for the user? • Is the UI intuitive and usable for the target impairment group? • Are design choices (contrast, layout, interaction mode) justified based on accessibility needs? • Does the system work reliably during a live demo without manual intervention? • Is the solution clearly presented, reproducible, and technically sound?
Innovation & Creativity (25%)	• Does the solution move beyond standard accessibility features (screen readers, captions, dark mode)? • Is AI used in a creative or unexpected way to solve the accessibility problem? • Is there a clear “wow factor” rooted in accessibility impact, not just technical complexity? • Does the solution meaningfully stand out from common accessibility tooling? • Does the system show clear paths for extension to other impairments, interfaces, or platforms? • Can the approach evolve with more data, users, or modalities?

Judging Criteria

Criterion	9–10 (Excellent)	6–8 (Good)	3–5 (Fair)	0–2 (Poor)
Impact (25%)	Clearly defined accessibility barrier; strong reduction in user effort or cognitive load; enables independent use; broadly applicable across users or platforms	Clear problem definition; visible accessibility improvement; supports partial independence; limited generalization	Problem loosely defined; minor or indirect accessibility benefit; narrow scope	No clear accessibility problem; purely cosmetic changes; no real benefit
Real-time Performance & Latency (25%)	Responsive multimodal interaction; low or well-justified latency; stable under continuous use; suitable for live deployment	Generally responsive; occasional delays; usable in most scenarios	Noticeable lag; inconsistent responsiveness; degrades usability	Non-interactive or batch-based; latency breaks usability; unstable
Design (25%)	Fully functional end-to-end system; well-integrated components; adaptive behavior aligned to user needs; reliable live demo	Functional system; reasonable integration; some adaptive elements; demo mostly works	Partial implementation; weak integration; limited adaptivity; fragile demo	UI mockup or static prototype; disconnected parts; demo fails
Innovation & Creativity (25%)	AI meaningfully enables personalization, adaptation, or multimodality; non-obvious approach; strong differentiation; clear extension potent	AI adds value; moderate novelty; some differentiation	Conventional approach; AI largely replaceable by rules or heuristics	No clear AI role; gimmicky or unnecessary use of AI

SUBMISSION FORMAT

Teams **must** submit a valid github link that **must** include clear execution instructions and repository structure in a readme.md file. If the project is not executable by judges due to any reason like api keys, private tokens, etc.. teams **must** provide a valid working deployed web-app link.

Teams **must** also submit a supporting pitch document. This can either be a video presentation (2 minutes maximum) or a slide deck PDF (12 slides maximum). Highlight the accessibility problem, describe the solution and its features based on the judging criteria.

SUBMISSION PORTAL

Submit the above on the google forms link provided below. Submit the video as a shared google drive link or an unlisted youtube video link. For text documents, submit a pdf.

https://shorturl.at/n0mCB

SUBMISSION

ROUND 2

Top 10 teams selected for Round 2 can refine and polish their end products for the final presentation.

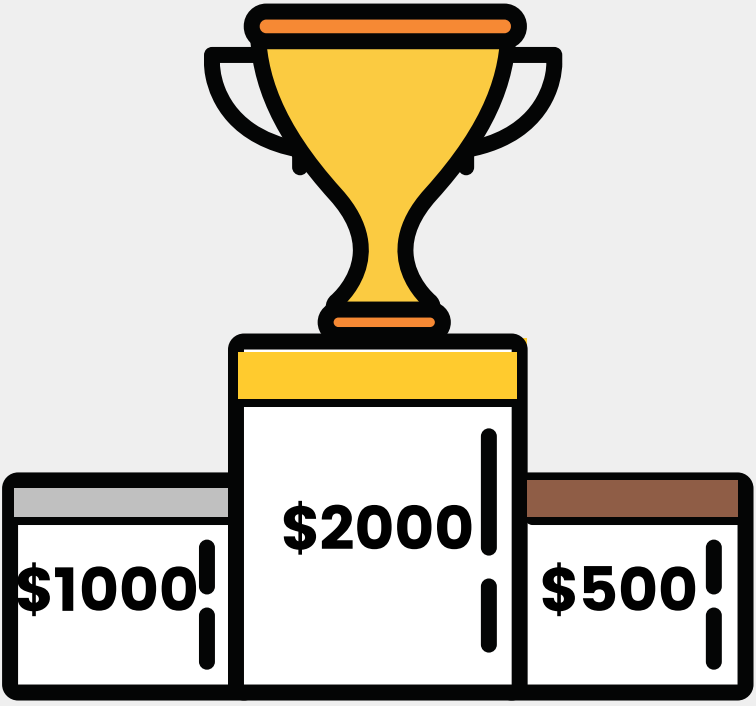
There is no submission for the final round, only a presentation.

PRESENTATION GUIDELINES

The presentation must be 8-10 minutes long, including a live demo showcasing your developed idea. No restrictions on length of slide deck.

After the presentation, there will be 5 minutes for a Q&A session where judges will ask technical and non-technical questions regarding your idea.

PRIZES



$500 for best freshman

Good Luck!

USEFUL

RESOURCES

MultiModal

Multimodal learning integrates data from different modalities, such as text, images, audio, and video, enabling advanced AI models to understand and process information across various contexts. Below is an overview of commonly used datasets and models in multimodal AI.

Note that use of MultiModal is neither mandatory nor will make a difference in judging criterion as long as the solution is creatively and innovatively solving the problem statement.

DATASETS

LAION-5B
A large-scale dataset containing 5 billion image-text pairs sourced from the web, used for training models like Stable Diffusion for tasks such as image generation and cross-modal retrieval.

COCO (Common Objects in Context)
A dataset of 300,000 labeled images with captions, commonly used for image captioning, object detection, and scene understanding.

Visual Genome
Provides 108,000 images annotated with objects, attributes, relationships, and region-specific captions, focusing on scene graph generation and semantic understanding.

VQA (Visual Question Answering)
A dataset of images paired with natural language questions and answers, aimed at evaluating models on visual-text reasoning tasks.

AudioSet
Contains over 2 million 10-second audio clips labeled with sound events, supporting applications like audio-visual synchronization and sound recognition.

MultiModal

CLIP (Contrastive Language-Image Pretraining)

- Associates images and text by learning shared embeddings, enabling zero-shot classification, content-based retrieval, and captioning.

DALL·E

- Generates high-quality, diverse images from textual descriptions using multimodal embeddings, ideal for creative content generation.

BLIP (Bootstrapped Language-Image Pretraining)

- Combines image and language embeddings for tasks like captioning and retrieval, offering efficient multimodal solutions.

VisualBERT

- Integrates image region embeddings with text using a transformer architecture for tasks like question answering and reasoning.

VideoCLIP

- Extends CLIP to video by adding temporal information, enabling video understanding, action recognition, and text-to-video retrieval.

Useful Resources

01 OpenCLIP

An open-source implementation of OpenAI's CLIP, this project has trained several models on various data sources and compute budgets, ranging from small-scale experiments to larger runs, including models trained on datasets such as LAION-400M, LAION-2B, and DataComp-1B.

02 Multimodal AI

An article that explores the fundamentals of multimodal AI, its applications, and the integration of different data types for comprehensive understanding:

https://www.datastax.com/guides/multimodal-ai?utm_source=chatgpt.com

Contact Us

All communication will be via Email or Telegram.
Make sure that you have joined the telegram chat below:

Email: IEEENTU-Branch@e.ntu.edu.sg

Tele: https://t.me/+NIkz-sRyntE1MzM1 or QR below: