SAM NGOPI

ALWAYS ON AI FOR DIGITAL LITERACY INCLUSIVITY

Empowering Communities with Inclusive Digital Literacy Through AI

MUHAMMAD FAQIH

MOCH. ALFAN MIFTACHUL HUDA

RIDHO AULIA RAHMAN

BREAKING BARRIERS: AI ASSISTANT FOR UNIVERSAL DIGITAL ACCESS

Visual impairment affects over 253 million people globally, including 36 million who are completely blind, stemming from various conditions like cataracts and diabetic retinopathy (Ganesan et al., 2022; Wang et al., 2023). Traditional aids such as canes and guide dogs provide limited support, while AI-driven assistive technologies have emerged as transformative solutions. These include deep learning models like CNNs for real-time object recognition and LSTM networks for descriptive feedback, enabling more effective navigation and communication for visually impaired persons (Ashiq et al., 2022; Kumar & Jain, 2022). Integrating these advancements, this project introduces a mobile application that acts as a screen summarizer, navigator, and voice assistant to enhance smartphone interaction for visually impaired users.

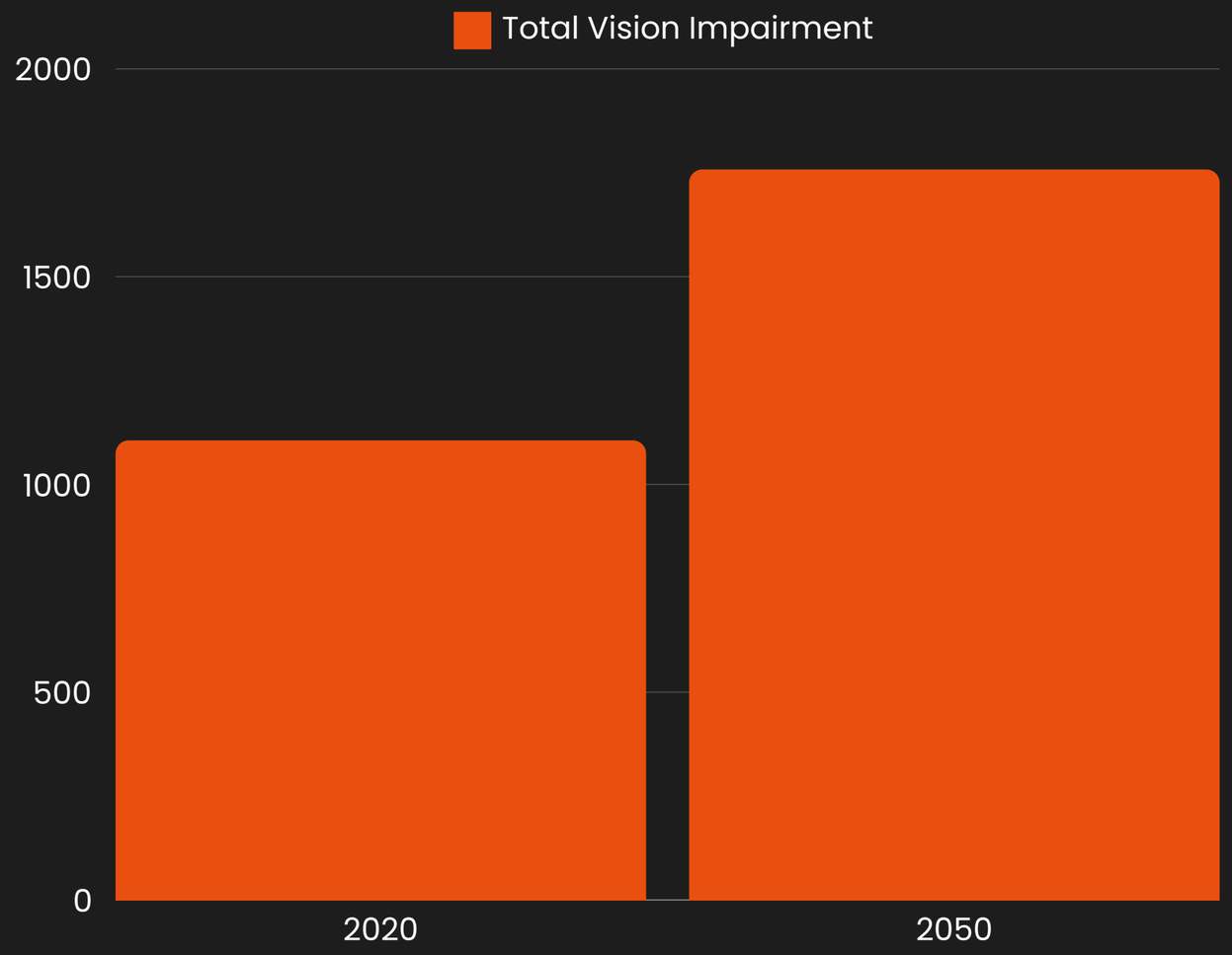
Image Generated by AI

According to the WHO (World Health Organization)
important impacts and implications of visual impairment

Early vision impairment causes developmental delays and lower educational achievements in children.

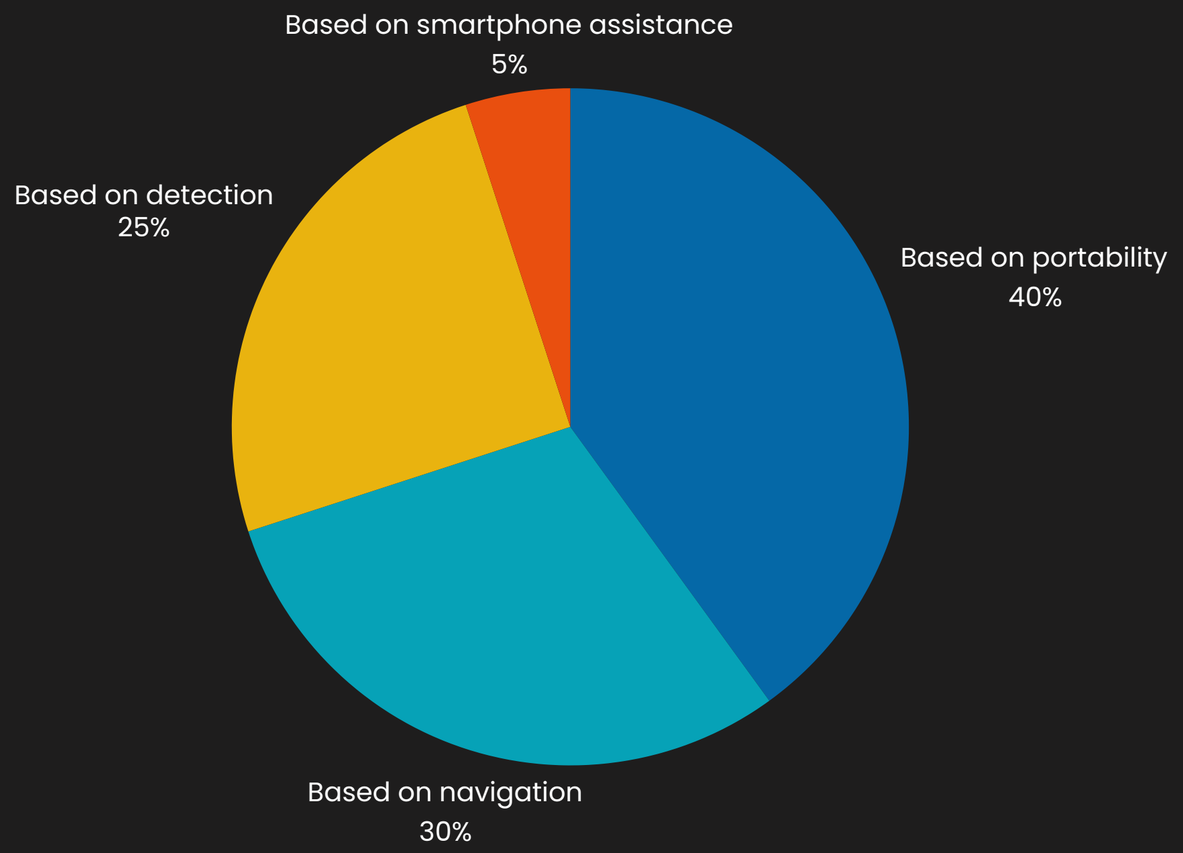
Among adults, vision impairment reduces quality of life, increases unemployment, depression, and anxiety risks.

Vision impairment results in a global productivity loss of $411 billion, despite the cost to meet unmet vision needs being only $25 billion.



As noted in the IAPB IAPB vision atlas The number of people with vision loss is growing Projections show that vision loss will increase by 55%, or 600 million people over the next 30 years.

IAPB: International Agency for the Prevention of Blindness



Assistive devices for visually impaired People

Based on the study from the article Towards Assisting Visually Impaired Individuals: A Review on Current Status and Future Prospects (Mashiata et al., 2022), the smartphone assistance category only accounts for 5% of available solutions.

This low percentage indicates that smartphone based solutions remain underexplored for the visually impaired.

INTRODUCING

AOI

ALWAYS-ON INTELLIGENCE

AI Assistant for Universal Digital Access

WHAT & WHY AOI?

AOI's core intelligence is powered by a robust, context-aware AI engine that operates directly on your device. It continuously monitors your screen and environment, building a deep understanding of your unique habits and preferences. This allows AOI to provide assistance that is not only helpful, but also feels natural and anticipatory.

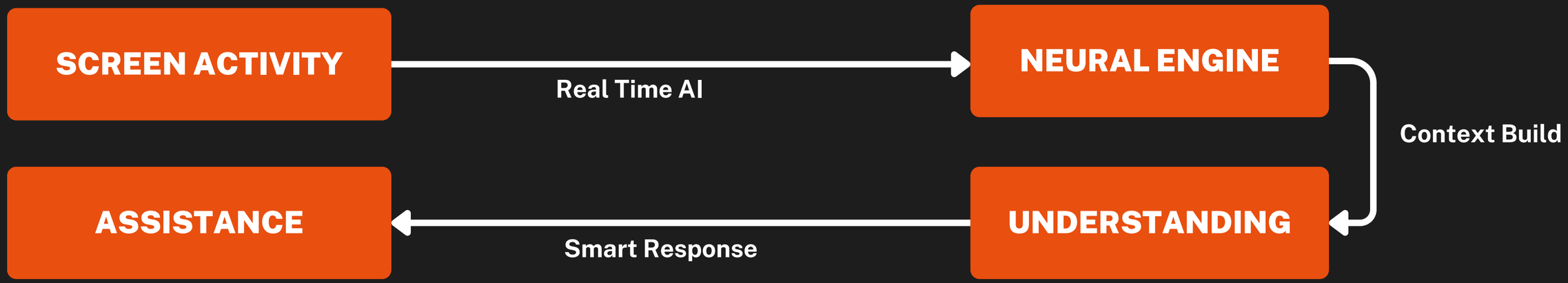
CONTROL ALL OF YOUR DEVICE WITH JUST YOUR VOICE

Imagine you're trying to fill out an online form, but the small text and intricate layout make it challenging. With AOI, you can simply say, "Hey AOI, what's on this page?" AOI will instantly analyze the screen and provide a detailed, narrated description of the form fields, instructions, and even suggestions on how to efficiently complete the task.

PRODUCT INNOVATION

THE VISION : BEYOND AN ASSISTANT

Unlike traditional assistive technologies that require constant user input, AOI is designed to be an always-present, proactive companion that seamlessly integrates into your daily digital life. Its sophisticated AI and context-aware capabilities set it apart, empowering you to regain control and independence in the digital world.



TECHNOLOGY FOR EVERYONE
WHO WILL USE AOI?

10 million people with disabilities in Indonesia

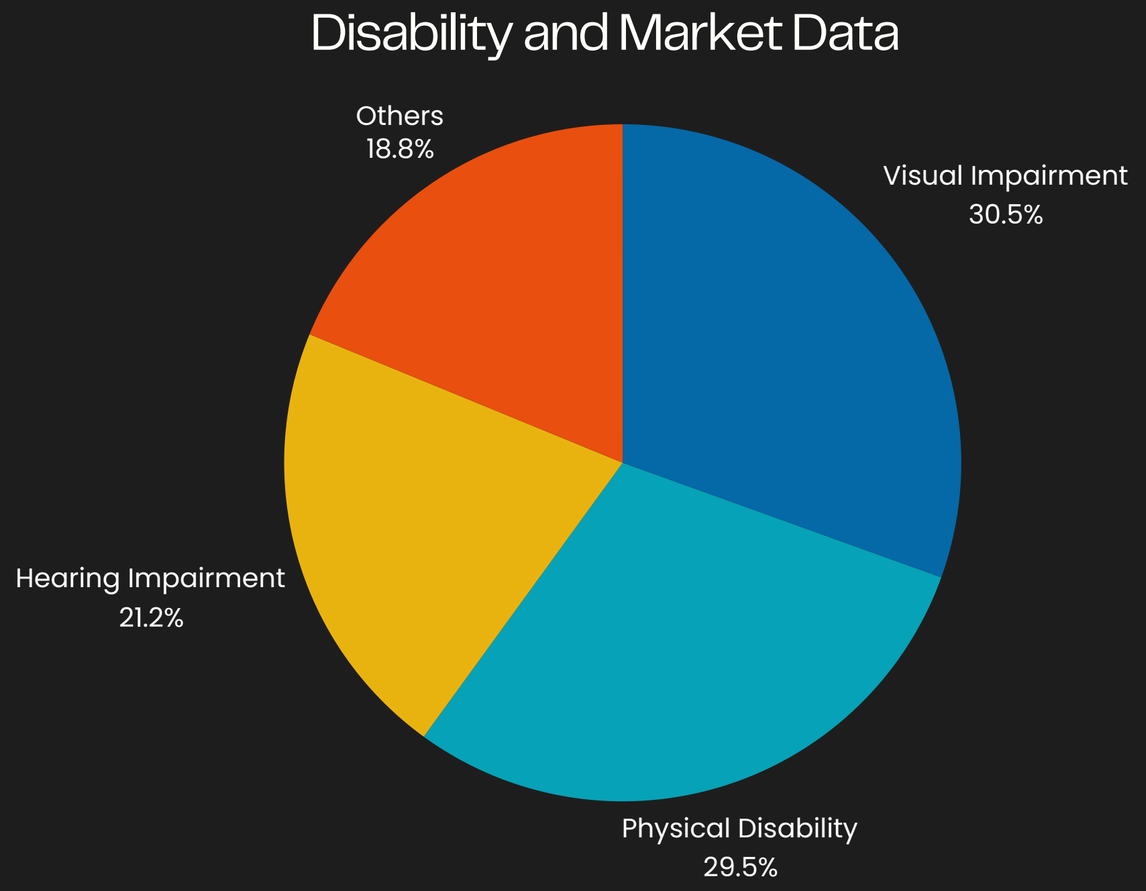
- **3 million have difficulty seeing**
- **2.9 million have difficulty moving**
- **2.1 million have difficulty hearing**
- **1.9 million others (various challenges)**

Elderly People (Parents & Grandparents):

- **Now: 29.3 million elderly people**
- **By 2035: Growing to 48.2 million**
- **Problem: 4 out of 10 elderly struggle with phones/gadgets**

Business Opportunity:

- **Current market value: $460 million**
- **Growth: Increases 12.3% every year**
- **Almost 90% of Indonesians have smartphones**



Source : Indonesian Ministry of Social Affairs Annual Report 2023

UNDERSTANDING USER NEEDS: MAKING TECHNOLOGY WORK FOR EVERYONE

Key User Insights 2024

(Based on National Digital Survey 2023–2024)

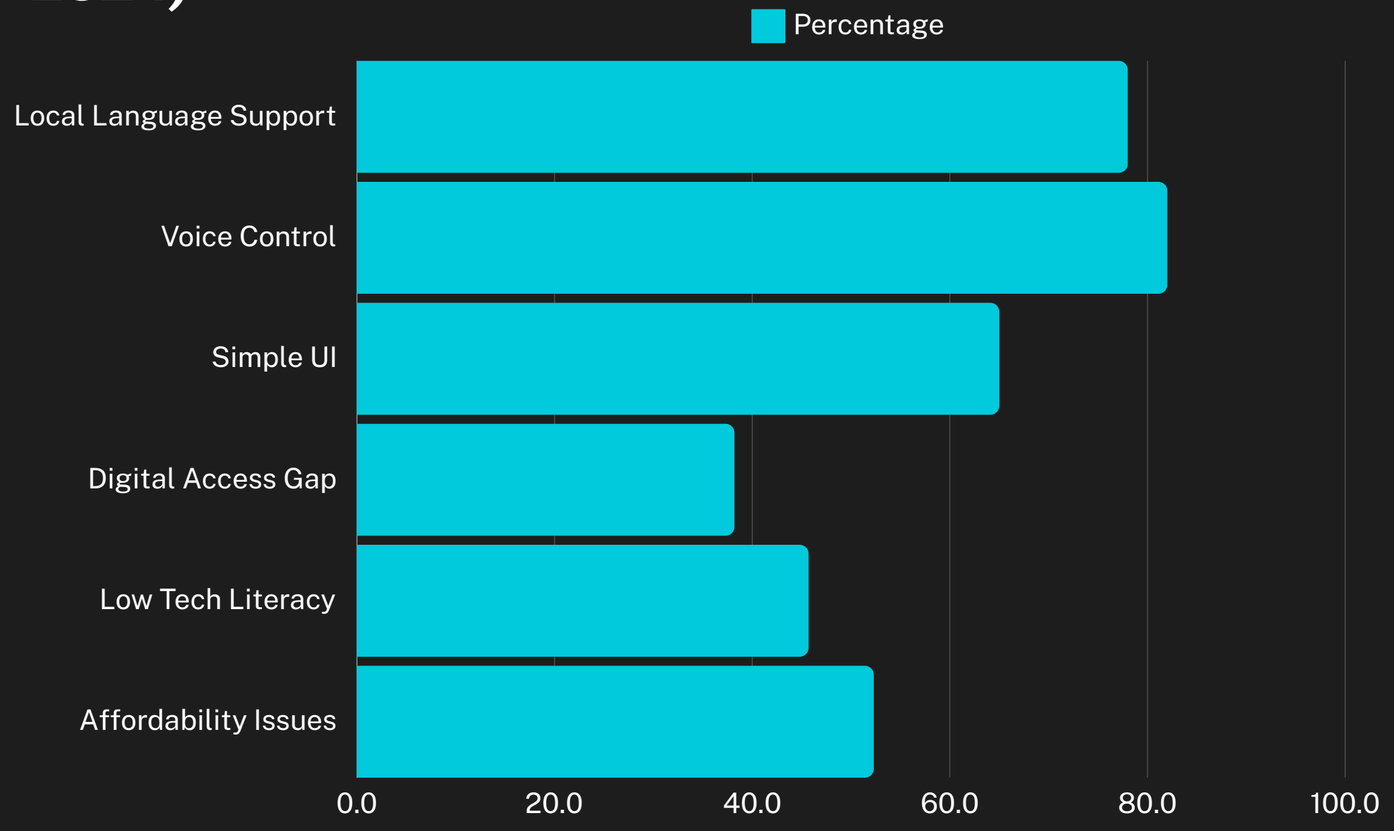
What People Want:

- 78% - "We need apps in our local language"
- 82% - "Voice commands are easier than typing"
- 65% - "Make it simple to use"

Main Challenges:

- 38% - No access to technology
- 46% - Don't know how to use it
- 52% - Too expensive to buy

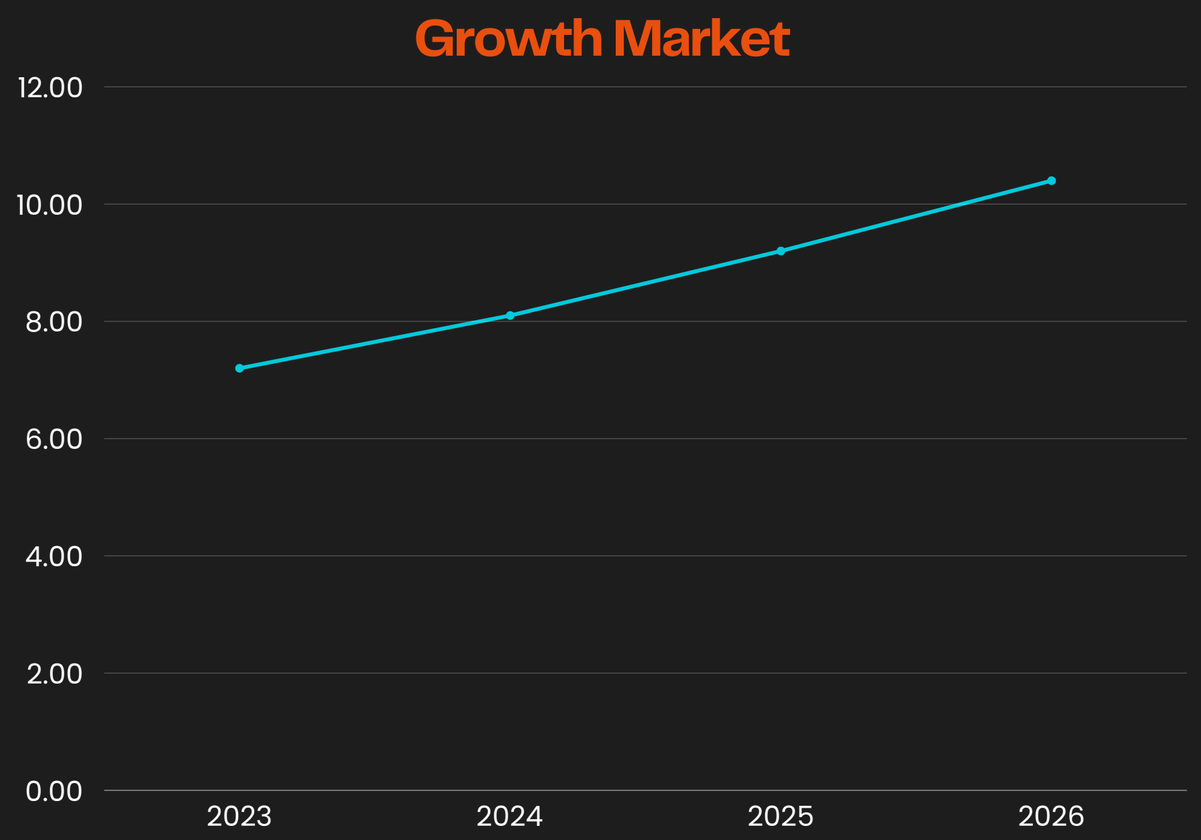
Market Growth: 2024: $520M → 2026: $670M (13.5% annual growth)



WHY AOI IS DIFFERENT

- **Proactive AI-Powered Actions**
  - Predicts user context & needs
  - Reduces manual interaction steps
  - More intuitive than reactive competitors
- **Personalized Learning**
  - Adapts to user preferences automatically
  - Custom navigation based on usage patterns
  - Machine learning optimization
- **Enhanced Voice Control**
  - Natural conversation flow
  - Continuous command sequences
  - Reduced need for command repetition
- **Integrated Visual AI**
  - Combines object recognition with device control
  - Seamless OCR integration
  - All-in-one accessibility solution

AOI revolutionizes accessibility by combining AI-powered proactive assistance, personalized learning, advanced voice control, and integrated visual recognition - creating a comprehensive, intuitive, and adaptable solution that truly understands and anticipates user needs.

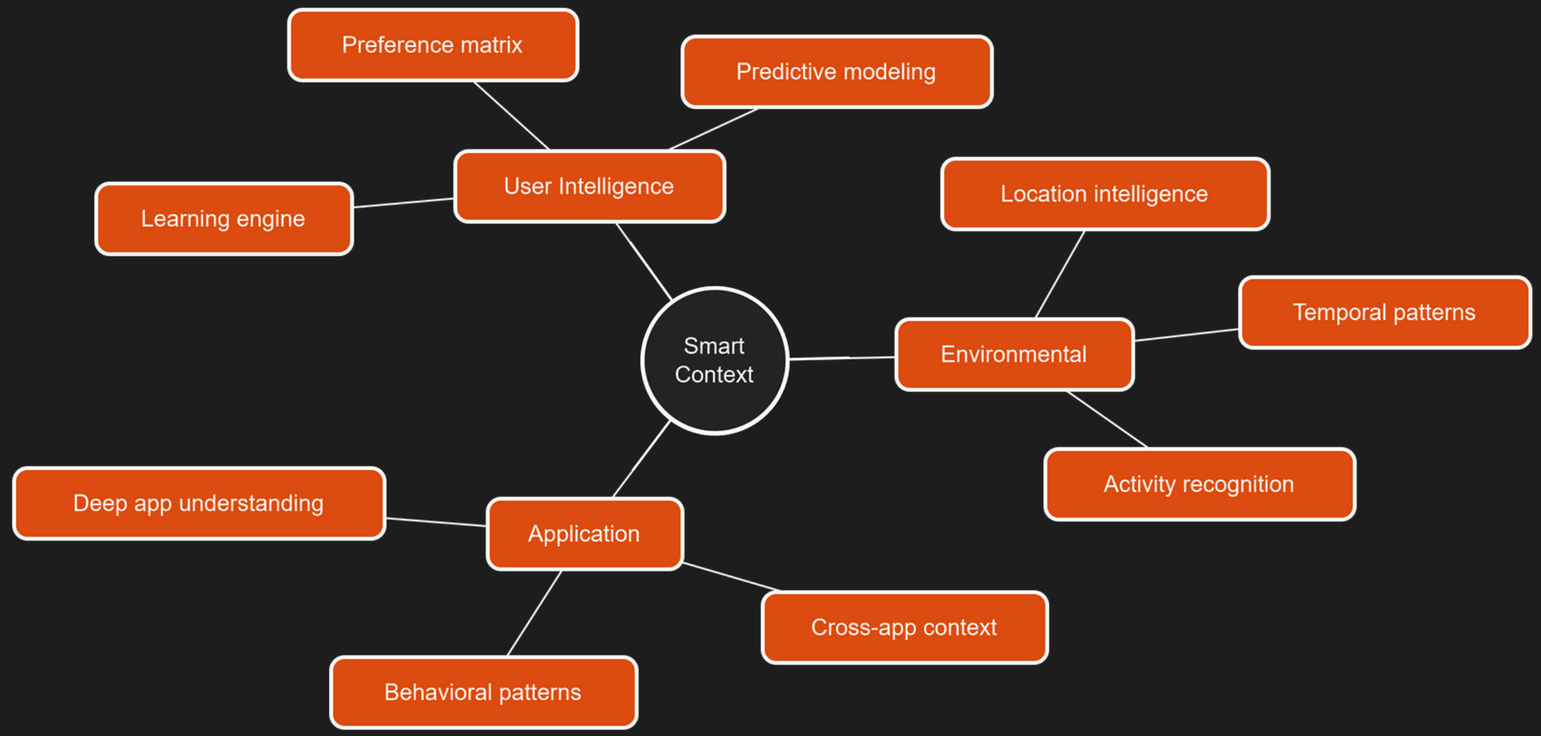


A trend line indicates consistent and positive growth over the period, showing a steady increase year over year.

Source : Indonesian Ministry of Social Affairs Annual Report 2023

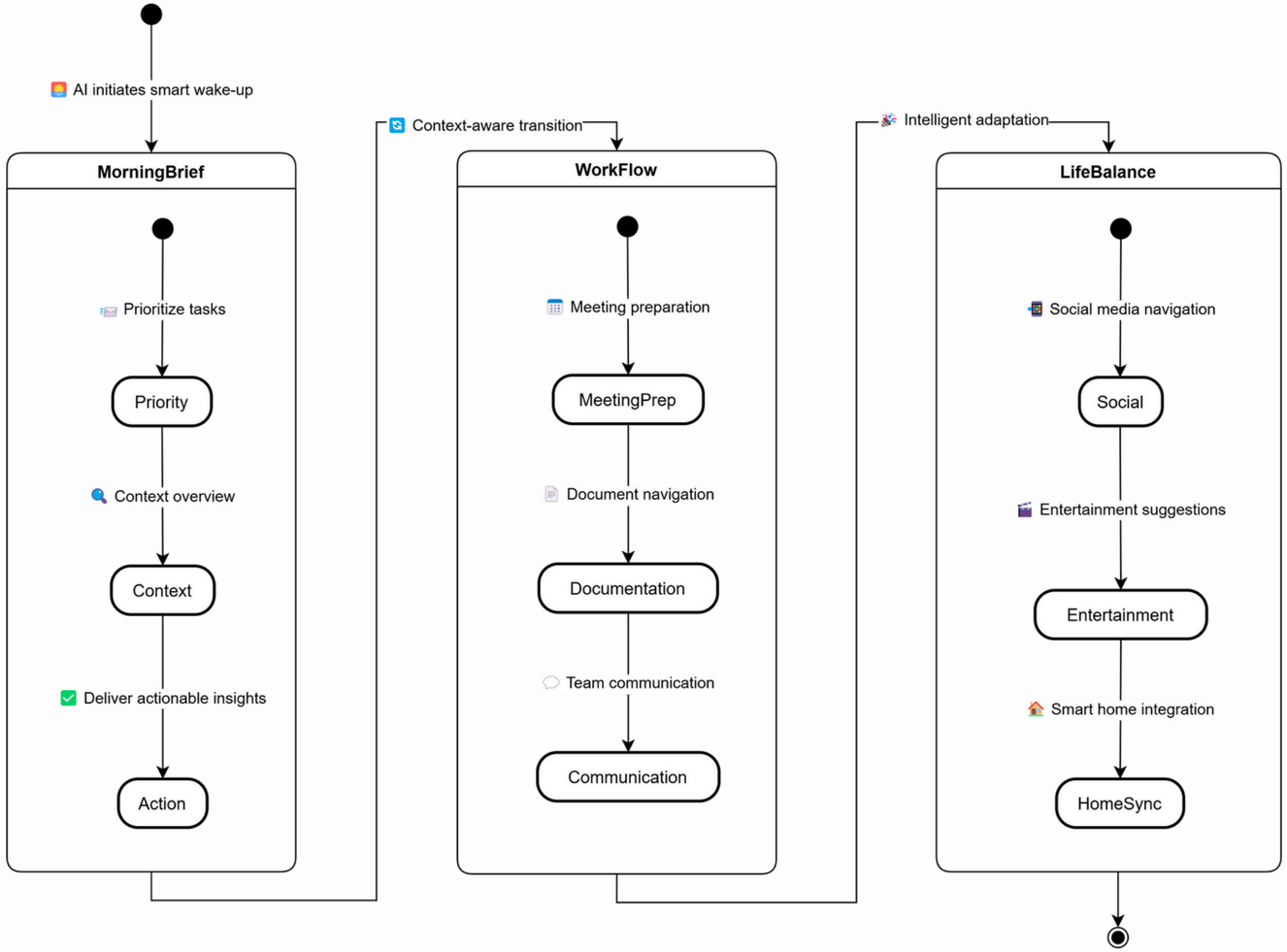
PRODUCT INNOVATION

CONTEXTUAL INTELLIGENCE ENGINE



Our multi-layered context understanding system creates a comprehensive awareness of user needs and environmental factors

A DAY OF EMPOWERED DIGITAL EXPERIENCE



MEET SARAH: BREAKING DIGITAL BARRIERS

Sarah, 34, is a marketing executive who experiences low vision. Like 1.3 billion people globally living with visual impairments, she faces daily challenges in digital interaction. But her story represents more than just accessibility - it's about transforming how everyone interacts with technology.

TRANSFORMATIVE JOURNEY

MORNING EMPOWERMENT

Sarah starts her day with intelligent prioritization. The system understands her morning routine, providing: Contextual email summaries reducing information overload, Smart calendar organization with proactive meeting preparation, Personalized news briefing tailored to her interests and reading preferences

LIFE ENHANCEMENT

Beyond work, the system adapts to her personal life: Intuitive social media navigation, Smart entertainment recommendations, Automated home device synchronization

PROFESSIONAL EXCELLENCE

In her work environment, the solution becomes her intelligent companion: Real-time document navigation, Contextual meeting assistance improving productivity, Seamless cross-application workflow reducing task switching time.

This isn't just an accessibility solution - it's a glimpse into the future where technology truly understands and adapts to human needs. Sarah's journey demonstrates how intelligent assistance can transform digital interaction for everyone, from those with specific needs to power users seeking enhanced productivity.

"IMAGINE A WORLD WHERE TECHNOLOGY UNDERSTANDS YOU SO WELL, YOU FORGET IT'S EVEN THERE."

TECHNICAL EXELLENCE : SYSTEM ARCHITECTURE

Frontend (Flutter)

- Flutter App
- Accessibility API
- Gesture Detector
- Voice Input Module

Middleware Layer

- Accessibility Data Processor
- UI Action Executor
- Speech-to-Text Handler

Backend Layer

Screen Reader Module

- NLP Model (BERT/T5)
- Context Extraction
- Text-to-Speech

Screen Clicker Module

- Speech Recognition
- NLU Model
- Action Processing

AI Multimodal Module

- Image Processor
- Multimodal Processor
- Recommendation Engine

FRONTEND LAYER

The frontend or "Flutter" layer of the system represents the user-facing components and interfaces. This layer includes:

Flutter App

The main application that users interact with, built using the Flutter framework.

Speech-to-Text Handler

Converts user speech to text via speech recognition, then passes it to the Voice Control System for processing. This allows hands-free interaction with AOI.

Accessibility API

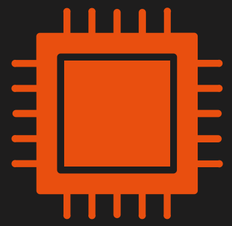
An interface that allows the application to provide accessible features for users with disabilities.

Voice Input Module

A component that enables voice-based input and interaction within the application.

MIDDLEWARE LAYER

The middleware layer in AOI functions to ensure that each core component can interact seamlessly. It involves receiving, processing, and transmitting data between various modules within the system. This layer functions to enable smooth integration among AI, OCR, and voice control features in AOI, ensuring real-time and context-appropriate responses.



Accessibility Data Processor

Manages data from accessibility modules (OCR, screen readers) and processes it for compatibility with other systems. This allows AOI to map and describe the screen layout for users.

UI Action Executor

Enables AOI to perform actions on the UI based on user commands, like opening apps or pressing buttons, allowing touch-free device control.

Speech-to-Text Handler

Converts user speech to text via speech recognition, then passes it to the Voice Control System for processing. This allows hands-free interaction with AOI.

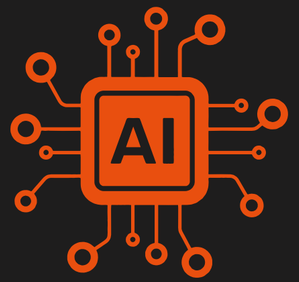
BACKEND LAYER

Screen Reader Module

This module leverages natural language processing (NLP) models like BERT and T5 to extract contextual meaning from the on-screen content. The Context Extraction component then synthesizes this information into a comprehensive understanding, which is ultimately converted into natural-sounding audio output using a high-quality text-to-speech (TTS) engine.

Screen Clicker Module

The Screen Clicker Module handles voice-based interactions, employing speech recognition models like DeepSpeech or Wav2Vec to transcribe user commands. The Natural Language Understanding (NLU) component, powered by GPT-2 or LLaMA, then interprets the intent behind these voice inputs and translates them into UI actions to be executed by the frontend.



AI Multimodal Module

The AI Multimodal Module is the heart of AOI contextual awareness and intelligent assistance. It integrates visual processing (OpenCV) with multimodal AI models like CLIP and Visual Transformers to achieve a deep understanding of the user's environment, activities, and preferences. This insight powers the Recommendation Engine, which can suggest relevant actions and provide proactive assistance tailored to the user's specific needs.

TensorFlow

CONCLUSION

AOI represents a transformative leap in the realm of digital accessibility, bridging the gap between technology and the diverse needs of all users. By fusing advanced AI capabilities, intuitive multimodal interactions, and a deep understanding of user context, AOI redefines the relationship between humans and digital devices. This innovative system empowers individuals with visual, auditory, or motor impairments to regain control over their digital lives, unlocking new levels of independence, productivity, and social integration.

AOI's cutting-edge architecture, grounded in principles of inclusive design and privacy-preserving technology, lays the foundation for a future where digital accessibility is not an afterthought, but a fundamental right. As we continue to push the boundaries of what's possible, AOI stands as a beacon of hope, ushering in an era where technology truly celebrates the diversity of human experiences and empowers everyone to thrive in the digital landscape.

REFERENCE

Ashiq, F., Asif, M., Ahmad, M. B., Zafar, S., Masood, K., Mahmood, T., ... & Lee, I. H. (2022). CNN-Based Object Recognition and Tracking System to Assist Visually Impaired People. IEEE Access, 10, 14819-14831. https://doi.org/10.1109/ACCESS.2022.3148036

Ganesan, J., Azar, A. T., Alsenan, S., Kamal, N. A., Qureshi, B., & Hassanien, A. E. (2022). Deep Learning Reader for Visually Impaired. Electronics, 11(3335). https://doi.org/10.3390/electronics11203335

Kumar, N., & Jain, A. (2022). A deep learning-based model to assist blind people in their navigation. Journal of Information Technology Education: Innovations in Practice, 21, 95-114. https://doi.org/10.28945/5006

Wang, J., Wang, S., & Zhang, Y. (2023). Artificial intelligence for visually impaired. Displays, 77, 102391. https://doi.org/10.1016/j.displa.2023.102391

World Health Organization. (n.d.). *Blindness and visual impairment*. World Health Organization. Retrieved November 7, 2024, from https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

International Agency for the Prevention of Blindness. (n.d.). Projected change. Vision Atlas. Retrieved November 7, 2024, from https://www.iapb.org/learn/vision-atlas/magnitude-and-projections/projected-change/

Mashiata, M., Ali, T., Das, P., Tasneem, Z., Rahman Badal, M. F., Sarker, S. K., Hasan, M. M., Abhi, S. H., Islam, M. R., Ali, M. F., Ahamed, M. H., Islam, M. M., & Das, S. K. (2022). Towards assisting visually impaired individuals: A review on current status and future prospects. Biosensors and Bioelectronics: X, 12, 100265. https://doi.org/10.1016/j.biosx.2022.100265