SLIIT
Discover Your Future

BSc (Hons) in Information Technology Specializing Data Science

Research Project - IT4010 Ethical Analysis Report

Group ID: 24-25j-281

Project Title: Signify – AI-Enhanced E-Learning Platform for Hearing-Impaired Children

	Identification and Analysis
Data Privacy Concerns	**Overview:** Signify processes various types of sensitive data from children aged 4 to 12 who have hearing impairments—a highly vulnerable group. The platform collects biometric video data (hand gestures), audio extracted from videos, textual inputs (from chatbots and quizzes), and also derives a GIF dataset representing sign language gestures. Managing this data ethically is critical to protect the privacy and rights of the children. **Detailed Considerations:** • **Biometric Data from Video and GIFs:** The platform captures live video streams of children performing American Sign Language (ASL) gestures through mobile device cameras during interactive AR sessions. To enable real-time feedback, this video data must be processed efficiently. Additionally, a GIF dataset was created by converting educational videos into short animated clips representing signs. This dataset is used for real-time audio-to-text-to-sign conversions in Component 2 and for chatbot text-to-sign translations in Component 3. While the GIF dataset abstracts gesture representation, it is based on videos that potentially contain sensitive biometric information. Ensuring that such data cannot be reverse-engineered or traced back to individuals is vital. • **Audio and Textual Data:** Audio is extracted from educational videos and converted into text using speech recognition. Both audio and transcribed text are then analyzed for sentiment and emotion detection, adding another layer of sensitive emotional profiling. User queries in the chatbot and quiz interaction data also contain personal educational information. • **Consent and Transparency:** Ethical operation requires obtaining clear, informed

	consent from parents or guardians. This consent must explain exactly what types of data are collected, how they are processed (including the creation and use of the GIF dataset), how long data is stored, and how it is used—for example, to provide real-time feedback, personalized course recommendations, and emotional support. • **Data Minimization:** The system must only collect data essential for delivering the educational experience. For instance, real-time video data for gesture recognition should be processed locally or transiently without long-term storage. Textual and quiz data stored for performance tracking must avoid storing direct personal identifiers. • **Anonymization/Pseudonymization:** Stored datasets, especially performance logs and chat interactions, should be pseudonymized. This prevents direct association with specific children and reduces privacy risks. • **Compliance with Child Data Protection Laws:** Given the users are minors, Signify must comply with GDPR-K, COPPA, or similar regulations that require additional protections such as parental controls, data access rights, and breach notification protocols. • **Potential Risks:** Inadequate privacy protection risks exposing children's biometric or emotional data, potentially leading to stigmatization, unauthorized profiling, or psychological harm.
Data Security	**Overview:** Securing the collected data from unauthorized access or breaches is fundamental, especially considering the sensitivity of video, audio, emotional, and educational performance data.

	Detailed Considerations: • **Encryption:** All data—video streams during AR sessions, audio clips, text transcripts, GIF datasets, chatbot interactions, and quiz results—must be encrypted both when transmitted over networks and when stored. This ensures confidentiality even if data is intercepted or improperly accessed. • **Role-Based Access Control (RBAC):** Access to sensitive data is limited to authorized personnel only—such as developers working on gesture recognition, educators reviewing student progress, or system administrators monitoring platform health. This is implemented through RBAC and authentication measures. • **Secure Cloud Infrastructure:** Signify uses Firebase and FastAPI backend services hosted on secure cloud platforms compliant with industry standards (ISO 27001, SOC 2). These platforms provide built-in safeguards against unauthorized access, malware, and data leakage. • **Authentication and Anti-Spoofing:** While the system currently relies on traditional authentication, future plans include biometric anti-spoofing for face login. This reduces risks of unauthorized access by mimicking authorized users. • **Continuous Security Monitoring:** Real-time monitoring systems and routine security audits detect vulnerabilities or breaches early. Incident response plans are in place to mitigate damage. • **Backup and Recovery:** Secure data backups ensure data availability and integrity, protecting against accidental loss or ransomware attacks.

	• **Risks if Neglected:** Data breaches could expose children's biometric or emotional profiles, leading to serious privacy violations and loss of trust.
Bias and Fairness	**Overview:** Machine learning models trained on limited or skewed data may encode biases that negatively affect accuracy and fairness, risking exclusion or harm to subgroups of users. **Detailed Considerations:** • **Dataset Diversity:** o **Gesture Recognition:** The ASL datasets (including the GIF dataset) contain a broad representation of hand shapes, skin tones, and environmental conditions, aiming to reduce racial or physical bias that could misclassify gestures of certain children. o **Speech Recognition and Sentiment Analysis:** Models are trained with varied speech samples representing diverse accents and emotional expressions common among children with hearing impairments to prevent linguistic or cultural bias. o **Chatbot and Recommendations**: The Q&A dataset includes questions across multiple grades and subjects, with collaborative filtering ensuring personalized recommendations that do not favor particular demographic groups unfairly. • **Continuous Evaluation:** Performance metrics like confusion matrices and fairness metrics are regularly analyzed across different demographic slices to detect bias. For example, if the gesture recognition model performs worse on darker skin tones, targeted data augmentation and retraining occur. • **Fairness Algorithms:** When bias is detected, fairness-aware training methods

	and post-processing adjustments are applied to correct skewed model behaviors. • **Human Review:** Educators monitor AI outputs, ensuring that biased predictions or unfair recommendations are identified and mitigated before impacting learners. • **Risks if Unaddressed:** Biased models could systematically disadvantage some children, reducing learning efficacy and violating principles of educational equity.
Transparency and Accountability	**Overview:** To foster trust and responsible use, Signify must be transparent about its AI systems and accountable for their outcomes. **Detailed Considerations:** • **Documentation:** The project maintains detailed documentation of datasets (including the GIF dataset derivation), model architectures (CNN for gestures, Whisper for speech recognition, T5 for summarization, Random Forest and ARIMA for prediction), training procedures, and known limitations. This documentation is made accessible to educators and guardians. • **Explainability:** AI outputs—such as real-time gesture correctness feedback, sign language translations, chatbot responses, and personalized learning recommendations—are designed to be interpretable. For example, students and educators can see why a gesture was marked incorrect or how quiz scores influence course suggestions. • **Human-in-the-Loop:** The system provides mechanisms for human oversight—

	educators or guardians can override or question AI decisions, ensuring critical learning feedback is not blindly automated. • **Feedback Mechanisms:** Users can report errors or ethical concerns, with the development team regularly incorporating this feedback into model retraining and system updates. • **Ethical Governance:** Periodic internal reviews and potential external audits ensure continued compliance with ethical AI principles and regulatory frameworks. • **Risks of Non-Transparency:** Black-box AI systems may cause mistrust, misinterpretation of results, and reduced user engagement.
Impact on Stakeholders	**Overview:** Signify impacts various stakeholders, each with unique concerns and benefits. **Detailed Considerations:** • **Hearing-Impaired Children:** The platform offers immersive, adaptive learning tailored to their needs, improving access and engagement. Emotional analysis supports content adaptation. However, inaccurate AI feedback or misunderstood emotions could cause frustration or decreased motivation. • **Parents and Guardians:** They play a critical role in consent and supervision. Transparency about data use and AI capabilities empowers informed decision-making and oversight. • **Educators:** AI augments teaching effectiveness but requires training to interpret AI outputs critically. Educators serve as ethical moderators, ensuring AI supports rather than replaces human judgment.

	• **Developers and Researchers:** Ongoing responsibility to uphold privacy, fairness, and transparency as the platform evolves, responding to emerging ethical challenges. • **Educational Institutions and Society:** Signify promotes inclusivity but must address technology access inequities to avoid widening educational disparities.