## The P3 Situational Learning Sandbox

#### **Concept:** An interactive, AI-driven learning and role-playing simulation platform that guides learners through a structured process of **Prepare, Practice, and Perform**, enabling them to develop and apply real-world skills confidently. This platform leverages LLMs to generate dynamic scenarios and provide personalised feedback, making it ideal for soft skills, critical thinking, and practical application of knowledge, especially in vocational training or professional development across Southeast Asia.

**How it Addresses the Challenge:**

The "Situational Learning Sandbox" directly addresses the core challenges of **personalisation, accessibility, language diversity, and engagement** by:

* **Personalisation through Dynamic Scenarios:** The LLM dynamically generates and adapts role-play scenarios based on the learner's chosen topic, skill level, and past performance. It acts as multiple characters within the simulation, adjusting its responses to the learner's actions and providing immediate, personalized feedback on their communication, decision-making, and problem-solving.
* **Accessibility (Varied Infrastructure):** The platform prioritizes efficiency and user preference. The core role-play interaction can be conducted via **text-based input/output or voice-based input/output**. This flexibility caters to learners with diverse preferences, varying digital literacy levels, and different bandwidth capabilities. Simplified, optional graphics can enhance engagement without heavy data requirements.
* **Language Diversity:** Learners can choose their preferred Southeast Asian language for the simulation, with the LLM (like SEA-LION) responding accordingly and incorporating culturally relevant contexts into the scenarios. This is excellent for practicing professional communication in diverse linguistic settings.
* **Engagement:** Role-playing is inherently engaging. The dynamic nature of LLM-driven scenarios keeps learners invested, and the "game-like" feel reduces the intimidation often associated with traditional learning.
* **Equitable Access:** This platform can bridge skill gaps for learners who lack access to internships or real-world practice, making them more competitive. One LLM can facilitate countless simultaneous simulations, offering high-quality practice at scale.

**The Prepare-Practice-Perform Framework in Action:**

The "Situational Learning Sandbox" is built around a robust three-phase learning cycle:

### **1: Prepare – Learn with Instant AI Coaching**

This initial learning phase equips learners with essential knowledge and strategy before they enter simulation practice. It’s designed to be **immediate, relevant, and contextual**, with real-time AI support throughout.

#### **Skill Modules by Profession and Industry**

Learners access concise, LLM-generated learning modules tailored to specific **soft skills and professional tracks**, such as:

* **Industries**: Retail, Call Centers, Logistics, Sales, Hospitality, and more
* **Skills**: Interviewing, Sales Conversations, Customer Service, Negotiation, etc.

Modules are available in **text**, **simplified visuals**, and **audio** formats to cater to varying learning preferences and infrastructure conditions across Southeast Asia.

#### **Certifiable Learning Journeys (Gamified)**

To motivate progression and show proof of skill development, learners can pursue **Certifiable Tracks**, e.g.: Complete 10 Negotiation Simulations to earn your ‘Negotiation Ready’ Badge”

These journeys can be **co-developed or endorsed by employers and vocational schools**, creating industry-relevant skill recognition and opening pathways to job readiness.

#### **Instant AI Coaching & Contextual Learning**

As learners move through modules, the LLM functions as an always-available personal coach:

* Answering questions in real time
* Offering examples (e.g., “Give me a STAR answer for a customer service role”)
* Providing immediate feedback on learner queries and drafted responses

The system embeds **regional and cultural context** into all explanations, ensuring relevance for local norms, behaviors, and communication expectations.

#### **Scenario Briefings & Strategic Planning**

Before entering any role-play simulation, learners receive detailed **scenario briefings** generated by the LLM, outlining:

* The context of the simulation
* Characters involved
* Key objectives and challenges to anticipate

This sets up the learner with a clear **mental model** and strategic framework.

#### **Knowledge Bank & Pre-Simulation Prompting**

Learners can tap into a searchable **Knowledge Bank** of best practices, regional norms, example phrases, and suggested responses.

They can also draft and test “scripts” or interaction plans with the AI for feedback before live role-play: “How does this sound if I’m negotiating salary in Malaysia?”  
 “What would be a better way to open this pitch in Thai culture?”

#### **“What Good Looks Like” (WGLL) Benchmarks**

After completing the preparation module, the LLM can generate and walk learners through **exemplar scenarios**—ideal responses and communication approaches: “Here’s what a strong conflict-resolution conversation looks like for a retail team lead in Singapore.”

These WGLL examples create **mental anchors for excellence**, giving learners clarity on success criteria.

### **2: Practice – Use AI Role-Play to Rehearse**

This is where learning becomes immersive and interactive. Learners enter simulated real-world scenarios powered by AI, allowing them to **practice, make mistakes, and improve in a safe, judgment-free environment**.

#### **Dynamic Role-Play with Adaptive Characters**

Learners engage in **real-time simulations** with the AI acting as various personas—such as a hiring manager, prospective client, upset customer, or senior stakeholder. The AI adjusts the scenario in real time based on the learner’s choices, tone, and strategy. Multimodal interaction is supported:

* Text-based for low-bandwidth users
* Voice-based (speech-to-text + text-to-speech) for more natural conversations

Scenarios are generated dynamically, reflecting the learner’s selected industry, language, skill level, and past performance.

#### **Cultural Intelligence Engine (CIE)**

Every simulation is layered with cultural sensitivity and regional relevance through the **Cultural Intelligence Engine**, which:

* Adapts responses and tone based on cultural etiquette and local norms (e.g., hierarchy in Thai workplaces, indirectness in Malaysian negotiation)
* Adjusts difficulty and phrasing to align with regional communication styles
* Embeds culturally relevant cues (e.g., dress codes, local names, context-specific greetings)

#### **Real-Time, Personalized Feedback**

The AI delivers **instant, constructive feedback** as the learner progresses through the role-play, focused on:

* Verbal communication (clarity, tone, empathy, assertiveness)
* Strategic approach (problem-solving, objection handling, persuasion)
* Cultural alignment (appropriateness of response, tone, and timing based on local norms)

The feedback is **actionable and specific**, helping learners understand both what they did well and what to improve.

#### **"What If" Scenario Exploration**

Learners can pause or rewind the simulation at key decision points to explore alternate paths:

“What if I had responded more assertively here?”  
 “How would the customer react if I had offered a refund instead of a discount?”

This feature builds **critical thinking and decision-making agility**.

#### **Performance Metrics**

The AI tracks and analyzes **key performance indicators**, such as:

* Communication clarity
* Confidence and tone
* Empathy and active listening
* Success in meeting simulation objectives

Each session ends with a **performance summary report** that includes:

* Quantitative scores
* Qualitative observations
* Progress against certifiable journey milestones

### **3: Perform – Apply Skills Confidently in Real-World Situations**

The ultimate goal is the successful transfer of learned skills to actual situations.

* **Actionable Insights:** Post-simulation, the LLM provides a comprehensive debriefing, summarizing key learnings, identifying recurring patterns in the learner's performance, and suggesting specific strategies for real-world application.
* **Personalised Drills:** Based on performance data, the LLM generates short, targeted drills or mini-scenarios focusing on areas where the learner needs more practice.
* **Self-Reflection Prompts:** The platform prompts learners to reflect on their simulated experiences and how they relate to real-world interactions they might encounter, fostering deeper understanding and skill integration.
* **Confidence Building:** Through repeated, safe practice and clear feedback, learners build confidence in their abilities, preparing them to apply their honed skills effectively in their personal and professional lives.
* **Community Integration (Optional):** Learners could share anonymized successful (or challenging) "perform" scenarios (with AI guidance) in a community forum for peer learning and inspiration.

The "Situational Learning Sandbox" is an ambitious project that combines AI, real-time interaction, and localised content. Here's a comprehensive tech stack breakdown, categorized for clarity:

**I. Core AI & Language Processing**

* **Large Language Models (LLMs):**
  + **SEA-LION (Primary Choice):** Given the explicit mention and focus on Southeast Asian languages and cultural relevance, SEA-LION (or similar localized LLMs like SeaLLMs, Sailor) would be the ideal core. This provides the foundation for dynamic scenario generation, personalized feedback, and culturally intelligent responses.
  + **Fallback/Supplemental LLMs:** For broader capabilities or in case SEA-LION needs augmentation, consider integrating with general-purpose LLMs like:
    - OpenAI GPT-4o, GPT-3.5 (for robust language understanding and generation, possibly fine-tuned for specific tasks).
    - Google's Gemini Pro or other specialized Google Cloud AI services.
    - Anthropic's Claude.
    - Self-hosted open-source LLMs (e.g., Llama, Mistral) if privacy, cost, or extreme customization are paramount, but this adds significant infrastructure overhead.
* **Speech-to-Text (STT) API:**
  + **Cloud Providers:** Google Cloud Speech-to-Text, Microsoft Azure Speech Service, Amazon Transcribe, OpenAI Whisper API. These offer high accuracy and support for many languages.
  + **Region-Specific Models:** Prioritize providers or models with strong performance in various Southeast Asian accents and dialects.
* **Text-to-Speech (TTS) API:**
  + **Cloud Providers:** Google Cloud Text-to-Speech, Microsoft Azure Speech Service, Amazon Polly, ElevenLabs, OpenAI TTS API.
  + **Natural Sounding Voices:** Select APIs that offer natural-sounding voices and, ideally, localized voices for Southeast Asian languages.
* **Natural Language Processing (NLP) Libraries/Frameworks:**
  + **SpaCy, NLTK, Hugging Face Transformers:** For pre-processing text, sentiment analysis (for feedback on tone), entity recognition, and general language understanding tasks. These can complement the LLM for specific analytical needs.
* **Cultural Intelligence Engine (CIE) Implementation:** This would likely be a combination of:
  + **Dedicated Datasets/Knowledge Graphs:** Curated data on cultural norms, etiquette, communication styles, and regional specificities in SEA.
  + **LLM Fine-tuning/Prompt Engineering:** The core LLMs would be extensively fine-tuned or prompted to incorporate and apply this cultural intelligence dynamically within scenarios.
  + **Rule-based Systems:** For very specific, non-negotiable cultural rules, a rule-based system might be layered on top of the LLM output to ensure strict adherence.

**II. Web Application Development**

* **Frontend (Client-side):**
  + **JavaScript Frameworks:**
    - **React.js:** Highly popular, component-based, excellent for building complex and interactive UIs. Strong community and rich ecosystem.
    - **Vue.js:** Easier learning curve than React, progressive framework, good for incremental adoption.
    - **Angular:** Robust, opinionated framework for large-scale enterprise applications.
  + **State Management:** Redux (for React), Vuex (for Vue), NgRx (for Angular) – to manage complex application state, especially for real-time interactions and user progress.
  + **UI Component Libraries:** Material-UI, Ant Design, Chakra UI – for consistent and performant UI elements, especially helpful for simplified graphics that are not data-heavy.
  + **WebRTC/WebSocket Library:** For voice-based interactions and real-time feedback.
    - **Socket.IO:** A popular choice for real-time, bi-directional communication.
    - **WebRTC frameworks/libraries:** For direct peer-to-peer audio/video if more complex communication features are envisioned (e.g., actual human mentors joining a session).
* **Backend (Server-side):**
  + **Programming Languages & Frameworks:**
    - **Python (with FastAPI/Django/Flask):** Excellent for AI/ML integration, extensive libraries, and strong community. FastAPI is good for high-performance APIs, Django for rapid development with ORM, Flask for lightweight services.
    - **Node.js (with Express.js/NestJS):** Ideal for real-time applications due to its asynchronous nature. Good for a unified JavaScript stack (frontend and backend).
    - **Go (with Gin/Echo):** For high-performance, concurrent services, especially if handling a massive number of simultaneous simulations.
  + **API Gateway:** NGINX or AWS API Gateway, Google Cloud API Gateway, Azure API Management for managing, securing, and scaling API requests to the LLMs and other services.
  + **Authentication & Authorization:**
    - **OAuth 2.0 / OpenID Connect:** For secure user authentication (e.g., integration with Google, Microsoft, or custom identity providers).
    - **JWT (JSON Web Tokens):** For stateless authentication between the client and backend.
    - **User Management System:** Custom or off-the-shelf solutions (e.g., Firebase Authentication, Auth0) for managing user accounts and roles.

**III. Data Storage & Management**

* **User Data & Progress (Relational Database):**
  + **PostgreSQL:** Robust, open-source, good for structured data like user profiles, learning paths, module completion, and aggregated performance metrics.
  + **MySQL:** Another popular open-source relational database.
* **Dynamic Scenario Data & LLM Interactions (NoSQL/Document Database):**
  + **MongoDB:** Flexible document-based database, ideal for storing the dynamically generated scenario details, specific LLM prompts, responses, and detailed session logs (which might have varying structures).
  + **Cassandra / ScyllaDB:** For highly scalable, distributed storage of large volumes of session data if immense scale is anticipated.
* **Knowledge Bank / Content Storage (Blob Storage/Search Index):**
  + **Cloud Object Storage:** Amazon S3, Google Cloud Storage, Azure Blob Storage for storing learning modules (text, audio, simplified visuals), certifiable track content, and WGLL examples.
  + **Elasticsearch / Algolia:** For building a searchable knowledge bank with fast full-text search capabilities.
* **Real-time Analytics & Caching:**
  + **Redis:** In-memory data store for caching frequently accessed data (e.g., user preferences, popular modules) and for real-time data processing (e.g., tracking active simulations, leaderboard updates).

**IV. Infrastructure & Deployment**

* **Cloud Provider:**
  + **AWS, Google Cloud Platform (GCP), Microsoft Azure:** All offer comprehensive suites of services, including compute (VMs, containers), AI/ML services (for hosting/accessing LLMs, STT/TTS), databases, storage, and networking. GCP is particularly strong in AI/ML services.
* **Containerization:**
  + **Docker:** For packaging applications and their dependencies into portable containers, ensuring consistent environments across development, testing, and production.
* **Orchestration:**
  + **Kubernetes (K8s):** For deploying, scaling, and managing containerized applications. Essential for handling fluctuating user loads and LLM inference demands. Managed Kubernetes services (EKS, GKE, AKS) simplify operations.
* **Serverless Computing (for specific components):**
  + **AWS Lambda, Google Cloud Functions, Azure Functions:** Can be used for specific, event-driven tasks like processing post-simulation reports, badge awarding, or background data processing to reduce operational overhead.
* **Content Delivery Network (CDN):**
  + **Cloudflare, Amazon CloudFront, Google Cloud CDN:** To cache static assets (images, audio, video) and deliver them quickly to learners across Southeast Asia, improving accessibility and performance, especially in regions with varying bandwidth.
* **Monitoring & Logging:**
  + **Prometheus, Grafana:** For infrastructure monitoring and visualization.
  + **ELK Stack (Elasticsearch, Logstash, Kibana) / Splunk / Cloud-native logging (CloudWatch Logs, Stackdriver Logging, Azure Monitor):** For centralized logging and analysis of application events and errors.

**V. Development & Operations (DevOps)**

* **Version Control:** Git, GitHub/GitLab/Bitbucket.
* **CI/CD (Continuous Integration/Continuous Delivery):** Jenkins, GitLab CI/CD, GitHub Actions, AWS CodePipeline, Google Cloud Build, Azure DevOps. For automating build, test, and deployment processes.
* **Infrastructure as Code (IaC):** Terraform, AWS CloudFormation, Azure Resource Manager, Google Cloud Deployment Manager for defining and provisioning infrastructure.

**Considerations for Southeast Asia:**

* **Low Bandwidth Optimization:**
  + **Prioritize text-based interactions.**
  + **Simplified graphics, optimized images/audio formats.**
  + **Efficient data serialization (e.g., Protobuf, MessagePack) over JSON for smaller payloads.**
  + **CDN for static assets.**
* **Multi-language Support:** Ensure all frameworks and libraries support Unicode and character sets for SEA languages. The chosen LLM is crucial here.
* **Regional Cloud Presence:** Utilize cloud regions closest to target user bases in SEA to minimize latency.
* **Cost Optimization:**
  + Leverage serverless options where appropriate.
  + Optimize LLM inference costs (batching requests, model distillation if possible).
  + Choose cost-effective database solutions.

This tech stack provides a robust foundation for building the "Situational Learning Sandbox," emphasizing scalability, personalization, and accessibility for the Southeast Asian context.

## **Layout & Navigation Flow (Conceptual)**

### **1. Home Dashboard (Landing Page)**

* **Top Section:** Personalized greeting ("Hello, [Learner Name]!"), current progress snapshot (e.g., "50% through 'Customer Service Pro' Journey").
* **"Continue Learning" Card:** Prominent card to pick up exactly where the learner left off (e.g., "Continue 'Negotiation Basics' Module" or "Re-do 'Upset Client' Simulation").
* **"Recommended for You" Section:** AI-curated suggestions for new modules or practice scenarios based on past performance and stated interests. Presented as scrollable cards.
* **"Your Journeys" Section:** Overview of active Certifiable Learning Journeys with progress bars and next milestone.
* **Quick Access Buttons:** Icons for "Start New Practice," "Explore Modules," "View Feedback."
* **Bottom Navigation Bar:** Home, Prepare, Practice, Perform, Profile.

### **2. Prepare Phase - Module View**

* **Module Title & Overview:** Clear title, brief description, and learning objectives.
* **Content Sections:** Organized into digestible, scrollable sections (e.g., "Introduction," "Key Concepts," "Examples," "Best Practices", “Methodology”).
  + Content can be text, simplified images, or embedded audio players.
* **“What good looks like”:** After completing the preparation module, the LLM can generate and walk learners through example scenarios, ideal responses and communication approaches.
* **Instant AI Coach Widget:** A persistent chat icon or small chat window at the bottom/side. Tapping it opens the AI coaching interface for real-time questions or drafting.
* **Knowledge Bank Search Bar:** Integrated search within the module for quick access to best practices.
* **"Next Section" / "Mark Complete" Buttons:** Clear navigation through the module.

### **3. Practice Phase - Simulation Interface**

* **Scenario Selection:** Dropdown options or the different Categories
* **Scenario Briefing (Pre-Simulation):** A dedicated screen before starting, outlining context, characters, and objectives. Includes a "Start Simulation" button.
* **Main Simulation Area (Chat-like):**
  + **AI Character Display:** Name/persona of the AI (e.g., "Hiring Manager," "Upset Customer") clearly displayed.
  + **Conversation Log:** Scrollable chat history. AI responses are distinct (e.g., different color bubble, avatar).
  + **User Input Field:** Large, clear text input box.
  + **Real-time Feedback Cues:** As the conversation progresses, subtle indicators (e.g., a small "i" icon next to an AI response or user input) appear. Tapping it reveals a concise feedback pop-up (e.g., "Consider more empathy here," "Good use of active listening!").
  + **"Try Again" Button:** Accessible during the simulation to "rewind" to a previous point and try a different response.
  + **"End Simulation" Button:** Clear option to conclude the session.
* **Session Assessment Scores ( Quantitative ):** the ratings of the answers as follows ( each out of 5 ):
  + Relevance -
  + Structured ( STAR ) -
  + Specific -
  + Honest -
  + Confident - but not arrogant -
  + Aligned with the role -
  + Outcome Orientated -
* **Session Feedback: ( Qualitative ):** Descriptive feedback and suggestions for improvements based on the above 7 categories and how the simulation went for the user.
* **Session History:** Session history replay with Options to download the script in voice or text transcript.

### **4. Perform Phase - Performance Summary**

* **Overall Score/Rating:** A prominent, easy-to-understand score or qualitative rating (e.g., "Competent," "Needs Practice").
* **Key Performance Indicators (KPIs):** Visual representation (e.g., a radar chart or simple bar chart) of scores across different metrics (Communication Clarity, Empathy, Problem-Solving, Cultural Alignment).
* **Qualitative Observations:** AI-generated summary of strengths and areas for improvement, using clear, actionable language.
* **Actionable Insights:** Specific, personalized recommendations for real-world application (e.g., "Next time, try using the 'STAR' method for behavioral questions").
* **Personalised Drills:** Links to mini-scenarios or modules specifically designed to address identified weaknesses.
* **Self-Reflection Prompts:** Open-ended questions to encourage deeper learning and connection to real-world experiences.
  + AI's Role in Reflection
    - Summarisation: The AI could summarise the learner's reflection, helping them consolidate their thoughts.
    - Gentle Probing: If a reflection is very brief, the AI might offer a follow-up prompt: "That's a good point. Could you elaborate on why that specific action would have changed the outcome?"
    - Connection to Resources: Based on the reflection, the AI might suggest relevant Knowledge Bank articles or specific parts of a Skill Module for review.
* **"Share Progress" (Optional):** Button to share anonymized performance with community (if enabled).
* **Performance Badge:**

### **Success Criteria for the Situational Learning Sandbox PoC**

The PoC's primary goal is to **demonstrate the technical feasibility and core value proposition** of the Situational Learning Sandbox. We aim to show that our AI-driven simulation can deliver personalised, engaging, and culturally relevant learning experiences in a controlled environment.

**1. Technical Feasibility & Core Functionality (Must-Haves):**

* **LLM Integration & Dynamic Scenario Generation:**
  + **Criterion:** The LLM successfully generates a variety of plausible role-play scenarios based on pre-defined topics (e.g., customer service, sales negotiation).
  + **Measurement:** Qualitative assessment of scenario realism and diversity by subject matter experts and pilot users.
* **Adaptive AI Character Response:**
  + **Criterion:** The AI character (LLM) intelligently adapts its responses to the learner's input (text-based) in real-time, maintaining context and progressing the simulation logically.
  + **Measurement:** AI response relevance score (e.g., 80% of AI responses are contextually appropriate and move the scenario forward) as assessed by human evaluators.
* **Basic Personalised Feedback:**
  + **Criterion:** The system provides immediate, basic feedback on learner performance within the simulation (e.g., "Good attempt at empathy," "Consider a more direct approach").
  + **Measurement:** Feedback relevance and clarity as rated by pilot users (e.g., average satisfaction score of 4 out of 5 for feedback).
* **Text-Based Multimodal Operation:**
  + **Criterion:** The core role-play interaction (Prepare and Practice phases) functions reliably via text input/output.
  + **Measurement:** System uptime during pilot (e.g., 99% availability) and absence of major text processing errors.

**2. User Experience & Engagement (Key Indicators):**

* **Ease of Use:**
  + **Criterion:** Pilot users find the platform intuitive and easy to navigate for the core "Prepare" and "Practice" functions.
  + **Measurement:** User onboarding completion rate (e.g., 90% of pilot users complete initial setup) and average task completion time for a simulated scenario.
* **Engagement & Stickiness:**
  + **Criterion:** Pilot users actively engage with the simulations and express a desire to continue using the platform.
  + **Measurement:** Average session duration (e.g., 15+ minutes per session) and qualitative feedback on "game-like" feel and enjoyment.
* **Perceived Learning Value:**
  + **Criterion:** Users perceive that the platform helps them practice and improve relevant skills.
  + **Measurement:** Post-simulation survey results (e.g., X% of users agree/strongly agree that the platform helped them improve communication/problem-solving).

**3. Cultural Relevance (Initial Validation):**

* **Basic Cultural Context Integration:**
  + **Criterion:** The AI demonstrates an ability to incorporate basic, pre-defined cultural nuances into scenarios and feedback, relevant to the selected SEA language/context.
  + **Measurement:** Qualitative assessment by native speakers/cultural experts on appropriateness of language and tone in scenarios.

**4. Scalability Potential (Early Indicators):**

* **System Performance:**
  + **Criterion:** The PoC can handle a limited number of concurrent users (e.g., 10-20 simultaneous simulations) without significant degradation in performance (response time).
  + **Measurement:** Average AI response time (e.g., < 3 seconds) under load conditions.
* **Data Capture for Learning Analytics:**
  + **Criterion:** The system successfully captures raw interaction data (user inputs, AI responses, timestamps) for future analysis.
  + **Measurement:** Verification of data logs containing relevant interaction data.

**5. Stakeholder Buy-in & Feedback:**

* **Stakeholder Endorsement:**
  + **Criterion:** Key internal stakeholders (product, tech, potential sales leads) and external pilot partners express confidence in the PoC's potential.
  + **Measurement:** Formal sign-off or positive feedback reports from pilot partners and internal review sessions.