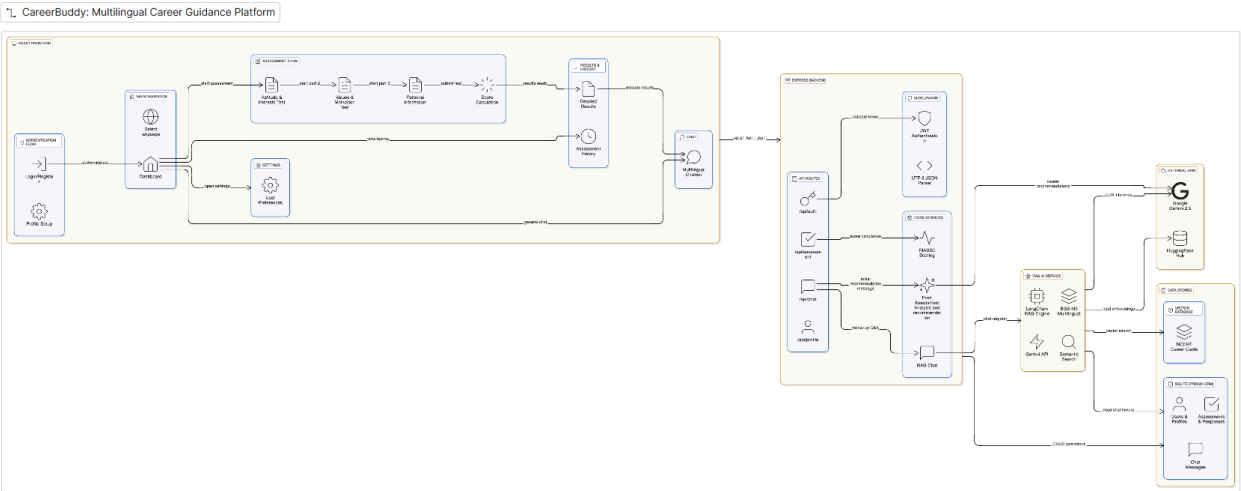
HCI Task 3 Report

Neerja Kasture - 22110165 Tapananshu Manoj Gandhi - 22110270
Ashmit Choker - 22110040 Dewansh Singh Chandel - 22110072

---

1. System Design & Architecture

The application employs a three-tier client-server architecture comprising a presentation layer (React-TypeScript), application logic layer (Express.js REST API), and data persistence layer (SQLite with Prisma ORM). The frontend implements component-based architecture with reusable UI modules (Radix UI, Tailwind CSS) following the separation of concerns principle. Backend modularity is achieved through dedicated route handlers (authentication, assessments, chat, profile) secured via JWT middleware. External AI integration utilizes the service layer pattern, encapsulating Google Gemini API calls within abstracted service classes to enable provider-agnostic implementation. The data model enforces relational integrity through Prisma ORM, defining entities for users, assessments, test responses, and chat messages with cascade deletion policies. Assessment workflow management implements strict state isolation via unique identifiers, preventing data cross-contamination between evaluation sessions.



System architecture diagram [link here]

---

2. Compromises in Prototyping

Horizontal Compromises

Horizontal prototypes provide a **broad overview** of the system but with **limited depth** in some areas.

Only Core Assessments Implemented

Although the final product might include broader psychometric modules (e.g., personality, cognitive ability), the prototype focused only on the three essential assessments—RIASEC, values, and preferences with a more limited question set.

Multilingual Chatbot Support Provided

While six languages are supported, conversational nuance, fine-tuned translations, and code-switching were not deeply optimized—just enough to show feasibility. Additionally, more languages can be integrated into this system easily but we kept it to this just for demonstration purposes.

Limited knowledge base of careers

The knowledge base of careers was limited to those mentioned in NCERT career cards. Of course, the overall number of career choices is a lot more

Vertical Compromises

Vertical prototyping focuses on **deep functionality** in a few selected areas while leaving others shallow.

AI Service Reliability: Limited retry mechanisms and request queuing for AI API calls, accepting occasional failure scenarios rather than implementing complex error recovery systems that increase vertical complexity.

Connectivity Requirements: System lacks offline progressive web app functionality, reducing vertical complexity in state synchronization at the cost of continuous connectivity requirements.

Deep Implementation of Database Constraints

Strict isolation of assessments, chat messages, and recommendations was implemented thoroughly to demonstrate reliability—while other database features (e.g., user analytics, logging) were only minimally implemented.

---

3. Conceptual Design

Design Metaphor

The system employs a **guided journey metaphor** where students progress through structured evaluation stages toward personalized career insights, mirroring traditional career counseling but augmented with AI intelligence. The conversational AI acts as a virtual career counselor providing on-demand guidance.

Design Space

Five functional domains comprise the system:

User Management: Authentication and profile administration with demographic data collection (age, grade, language preferences).

Assessment Orchestration: Multi-test workflow implementing RIASEC vocational theory, values alignment mapping, and preference evaluation with progress persistence enabling session resumption.

Recommendation Generation: Hybrid engine combining rule-based RIASEC code matching with generative AI synthesis (Google Gemini), producing confidence-scored recommendations with explanatory reasoning and actionable next steps.

Conversational Guidance: Dual-mode chatbot architecture—general career exploration (session-independent) and results-specific discussion (assessment-bound)—maintaining separate contexts to ensure information isolation.

Historical Analytics: Assessment tracking and visualization enabling longitudinal career development tracking.

Design Decisions

Input Modality: Likert-scale questionnaires selected over open-ended responses to enable quantitative psychometric analysis while reducing cognitive load.

Recommendation Format: Structured presentation chosen to facilitate comparison and actionable planning.

Interaction Pattern: Conversational chatbot integrated alongside static recommendations to accommodate diverse learning styles and follow-up inquiries.

---

4. Prototype Description and Functionalities

The web application prototype provides an end-to-end career guidance platform designed to offer personalized, accessible, and reliable recommendations for students. At its core, the system delivers a structured multi-stage assessment experience that evaluates vocational interests, personal values, and individual preferences. Users can easily register, log in, and complete assessments at their own pace, with automatic progress saving to support flexible participation.

Once assessments are submitted, the platform uses Google Gemini AI to generate personalized career recommendations, offering clear reasoning, confidence indications, and practical next steps. A multilingual, Retrieval-Augmented Generation (RAG) chatbot further enhances the experience by enabling students to explore careers through natural conversation in six major Indian languages. To ensure high-quality, contextual information, the system uses processed and cleaned data derived from NCERT Career Cards. The chatbot is equipped with guardrails to ensure safe, age-appropriate, and reliable guidance throughout the interaction

Students can review their assessment history, revisit detailed results, and access related chat conversations through a unified dashboard. Profile settings allow basic information management, while built-in safeguards ensure that each assessment and its corresponding recommendations or chat interactions remain securely linked. Overall, the prototype demonstrates a streamlined, student-friendly approach to AI-supported career exploration within a structured and research-grounded framework.

Complete Functionality List

Assessment Workflow Management

- Start new assessment with automatic assignment of sequential assessment numbers
- Detection and resumption of existing in-progress assessments
- Three-test assessment flow: Aptitude, Values, Personal preferences
- Automatic creation of test response records for each assessment stage
- Progress tracking with current question index for each test
- Real-time progress saving during test completion
- Assessment status management
- User ownership verification for all assessment operations

RIASEC Career Matching System

- RIASEC score calculation from questionnaire responses
- Top 3 RIASEC code identification based on highest scores
- Retrieval of matching career titles from predefined career lists, this list was taken from NCERT career cards as it is a good starting point for a comprehensive list
- Career database matching using RIASEC codes, these were taken from ONet Online Occupation Search Database which has RIASEC code for every career.

- Integration of RIASEC results into recommendation generation

AI-Powered Recommendation Generation

- Google Gemini AI integration for recommendation synthesis
- Processing of complete assessment data (all three tests)
- Generation of career recommendations using RIASEC identified careers, modified by values and personal preferences
- Detailed reasoning explanation for each recommendation
- Summary generation of overall assessment results
- Fallback response generation when AI service fails based purely on deterministic RIASEC score
- Multi-language support in recommendation generation

RAG Chatbot System

- Retrieval-Augmented Generation for accurate career information using NCERT career cards as source
- Runs on Gemini API, also supports GROQ as a backup
- Context-aware conversation handling with message history
- Session-based chat management with unique session IDs
- Multi-language support, having multilingual embeddings in the vector database itself
- Conversation history retrieval with chronological ordering
- Assessment-specific context loading for relevant responses
- Parallel RAGbot initialization during reccommendation generation by recommendation bot for optimization
- Contextual message history (last 10 messages currently due to api request size constraints)
- Error handling with fallback responses for chat failures
-

Assessment History & Analytics

- Retrieval of all completed assessments for user
- Display of assessment completion timestamps
- Test count display for each assessment
- Chat availability indicator for each assessment
- Assessment deletion with cascade removal of test responses and chat messages

Home Dashboard

- Easy access buttons to take assessments, view previous assessment history or chat with career chatbot.
- Easy access button for settings page
- Intuitive and friendly design
- In-progress assessment detection with progress verification
- Display of saved test responses for in-progress assessments
- Completion statistics calculation (total assessments, completed count)

Authentication & User Management

- User registration with username, password, name, grade, age, email, and phone number.
- JWT-based secure login authentication with token generation.
- Password hashing using bcrypt for security
- Token-based middleware authentication for protected routes
- User profile retrieval with selective field exposure (excludes password)
- Profile update functionality (name, grade, age, email, phone)
- Password change with current password verification
- Complete account deletion with cascade removal of all related data
- Setup completion tracking to determine if user has completed onboarding

API Endpoints & Routing

- `/api/assessment/start` - POST: Create new or resume assessment
- `/api/assessment/test/:testType` - GET: Retrieve test questions and progress
- `/api/assessment/save` - POST: Save test progress
- `/api/assessment/submit` - POST: Submit completed test
- `/api/assessment/complete` - POST: Process assessment and generate recommendations
- `/api/assessment/history` - GET: Retrieve assessment history
- `/api/assessment/:id` - DELETE: Delete assessment
- `/api/assessment/:id/responses` - GET: Get detailed test responses
- `/api/chat/start` - POST: Initialize chat session
- `/api/chat/message` - POST: Send message and receive AI response
- `/api/chat/history` - GET: Retrieve chat history
- `/api/chat/history` - DELETE: Delete chat messages
- `/api/` - GET: Home dashboard data
- `/api/profile` - GET: Retrieve user profile
- `/api/profile/update` - PATCH: Update user profile
- `/api/profile/delete` - DELETE: Delete user account

Security & Authorization

- JWT middleware for route protection
- User ownership verification for all resource access
- Password verification before allowing password changes
- Token validation on every protected endpoint
- Unauthorized access prevention (403 responses)
- Resource not found handling (404 responses)

Error Handling & Validation

- Comprehensive error logging throughout the application
- Input validation for required fields
- Test type validation against allowed values
- Assessment ID validation (integer parsing, NaN checking)
- JSON parsing error handling with fallbacks
- 400 Bad Request responses for invalid inputs
- 500 Internal Server Error responses with fallback data
- Environment variable validation (API keys)

Multi-Language Support

- Six supported languages: English, Hindi, Telugu, Tamil, Bengali, Gujarati
- Language preference storage in user profile
- Default language fallback to English

5. User Evaluation and Results

Evaluation Method

We conducted a summative usability evaluation using the System Usability Scale (SUS). Participants interacted with the system in a natural, real-world setting and subsequently completed an online SUS questionnaire delivered via Google Forms. The collected data, consisting of standardized SUS responses, provided quantitative usability metrics including mean score, median score, and score variability, enabling a robust assessment of overall system usability.

The evaluation was carried out through this google form : [link]

The **System Usability Scale (SUS)** is a standardized 10-item Likert scale questionnaire designed to measure perceived system usability. The SUS utilizes alternating positive and negative statements rated on a 5-point scale (1=Strongly Disagree, 5=Strongly Agree) to minimize response bias.

SUS Questionnaire Items:

1. I can very well imagine using the system regularly
2. I find the system unnecessarily complex
3. I find the system easy to use
4. I think that I would need technical support to use the system
5. I find that the various functions of the system are well integrated
6. I think there are too many inconsistencies in the system
7. I can imagine that most people learn to master the system quickly
8. I find the operation very cumbersome
9. I felt very confident using the system
10. I had to learn a lot of things before I could work with the system

Participants

Participants were recruited following informed consent procedures, completing demographic information (age, grade, language, current field etc.) from among student community of IIT Gandhinagar. We performed testing on non target user due to lack of availability of target user i.e 10th grade students.

Statistical Results

SUS scores were calculated using standard protocol: for odd items (1,3,5,7,9), scale position minus 1; for even items (2,4,6,8,10), 5 minus scale position; sum multiplied by 2.5.

Aggregate Measures:

- **Mean SUS Score**: 83.18
- **Standard Deviation**: 11.296
- **Median Score**: 85.0

The mean score of 83.18 significantly exceeds the industry benchmark of 68 (average usability) and surpasses the 80.3 threshold for excellent usability, placing the system in the top 10-15% of evaluated products. This is further supported by a **median score of 85**, suggesting that most users provided very strong usability ratings. The standard deviation of **11.30** indicates moderate variability, implying that while the majority of participants rated the system highly, a small number reported comparatively lower—but still acceptable—scores. Overall, the SUS outcomes

confirm that users found the system easy to use, efficient, and supportive of confident interaction, aligning with the positive trends observed in the item-level analysis.

Item-Level Performance:

- **Functional Integration (Q5)**: Mean=4.36, SD= 0.68
- **Ease of Use (Q3)**: Mean=4.54, SD=0.9
- **Regular Usage Intention (Q1)**: Mean=3.72, SD=1
- **System Complexity (Q2)**: Mean=1.45, SD=0.68 (low perceived complexity)
- **User Confidence (Q9)**: Mean=4.54, SD= 0.82
- **Learning Requirements (Q10)**: Mean=1.90, SD=1.04

Interpretation

- The system is perceived as **easy to use**, **simple**, and **well-integrated**, with users reporting **high confidence**.
- These are hallmarks of strong usability.
- The only middling area is **regular use intention**, which likely depends on perceived usefulness, not just usability. In our case, this is because testing was done on college students instead of grade 10 students due to lack of availability of target demographic.
- A minority of users may need more support or clearer onboarding, as seen from variability in the learning-requirements item.

---

6. Ethics and Risk Assessment

Participant Protection

All evaluation participants provided **informed consent** before beginning evaluation. Participants were informed of evaluation purposes, their rights to withdraw at any time, data usage policies, and anonymity guarantees. The informed consent form acted as a contract between participants and researchers, ensuring transparency in the evaluation process.

Data Privacy and Security

Participants were not asked to provide any personally identifiable information (PII), and all responses were collected anonymously. This reduced risks related to data misuse, identity

exposure, or personal profiling. Data was stored securely on Google Forms with restricted access limited to the team.

User data including assessment responses, demographic information, and chat histories are stored with appropriate security measures. JWT-based authentication ensures secure access control with token-based session management. Database constraints maintain data integrity and prevent unauthorized cross-user access.

Risk Mitigation

Usability Risks: Evaluation identified minimal consistency issues (Q6: mean=1.81) and low cumbersome operation scores (Q8: mean=1.36), confirming acceptable risk levels for user interaction.

Bias Considerations: The SUS methodology's alternating positive/negative statements minimizes response bias, while demographic data collection enables analysis of potential usage disparities across user groups.

Minimizing User Fatigue and Discomfort
The questionnaire was intentionally kept short (10 SUS items + minimal demographic or follow-up questions), reducing cognitive load, survey fatigue, or frustration. Participants were informed that the activity would take only a few minutes, which helped manage expectations and prevent discomfort.

Remote, Low-Risk Environment
Because the evaluation took place through an online form, participants could complete it at their own pace and in a comfortable environment, minimizing any risk of situational stress, pressure, or observation.

7. Conclusion

The high-fidelity prototype successfully demonstrates a comprehensive AI-powered career guidance platform with strong usability performance (Mean SUS: 83.18, SD: 11.296). The evaluation validates core design decisions including the guided journey metaphor, three-tier architecture, RIASEC-based assessment framework, and conversational AI integration. Strategic compromises in horizontal scalability (SQLite database, JWT authentication without distributed session management) and vertical complexity (AI service reliability, chatbot sophistication, offline functionality) enabled rapid prototyping while maintaining functional completeness for demonstration purposes.

Item-level analysis confirms exceptional functional integration (Q5: mean=4.36), strong ease of use (Q3: mean=4.54), and positive regular usage intention (Q1: mean=3.72), with low perceived complexity (Q2: mean=1.45) and minimal inconsistencies (Q6: mean=1.81). Minor variability in user confidence (Q9: SD=0.82) and learning requirements (Q10: SD=1.04) indicates

opportunities for enhanced onboarding support, particularly for users with diverse technical backgrounds.

The working demonstration encompasses complete user workflows from authentication and profile setup through multi-stage assessment completion to AI-generated career recommendations and conversational exploration across six languages. The prototype successfully meets HCI design principles including user-centered design, iterative evaluation, accessibility considerations through multi-language support, and ethical evaluation practices with informed consent. Future iterations should address identified horizontal scaling limitations for production deployment while maintaining the validated user experience that positions the system in the top 10-15% of evaluated digital products.

References:

1. System Usability Scale methodology sourced from Wikipedia - System Usability (https://de.wikipedia.org/wiki/System_Usability_Scale)
2. RIASEC codes for careers sourced from O*NET OnLine
3. Career context for RAG sourced from NCERT
4. Response sheet for usability testing: https://docs.google.com/spreadsheets/d/1F5_wyr1wdNZY_MzDDeKy2SoOYYVaG9EPR2QJIn9SZ2c/edit?usp=sharing