Smart-Sheet: AI-Powered Excel Assessment System

Comprehensive Design Documentation

Executive Overview

Smart-Sheet represents a paradigm shift in Excel skills assessment, moving from static, predetermined question sets to a dynamic, AI-driven evaluation system. The platform employs sophisticated AI agents to create personalized learning trajectories, generate contextually relevant business scenarios, and provide real-time adaptive questioning based on candidate performance.

Key Innovation Points

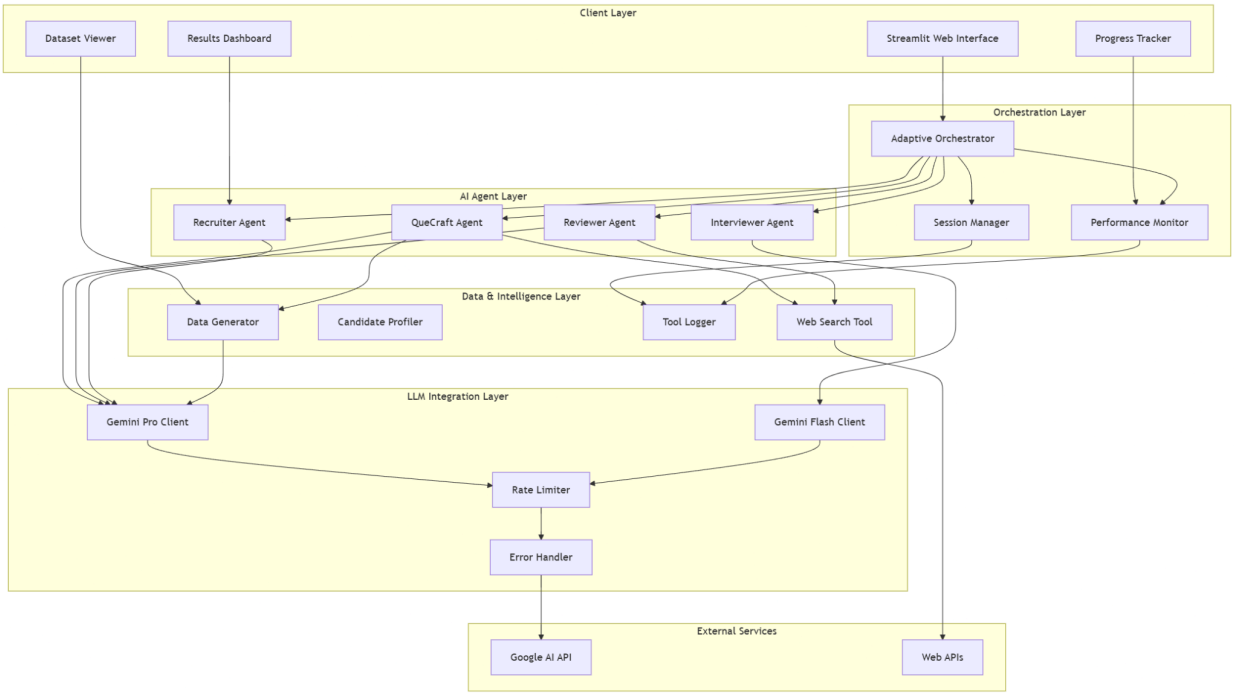
- **Response-Driven Architecture**: Questions generated in real-time based on candidate performance
- **Contextual Data Integration**: AI-generated datasets that perfectly match assessment scenarios
- **Multi-Agent AI System**: Specialized agents handling different aspects of the assessment process
- **Adaptive Learning Trajectories**: Dynamic difficulty adjustment and skill focus based on demonstrated capabilities

---

System Architecture

Core Components Overview

The Smart-Sheet system employs a sophisticated multi-agent architecture with clear separation of concerns and specialized AI components working in concert to deliver adaptive Excel assessment.



1. Adaptive Orchestrator

The **Adaptive Orchestrator** (`InterviewOrchestrator`) serves as the central intelligence hub, coordinating all assessment activities and maintaining the adaptive state machine.

Core Responsibilities:

- **Interview State Management**: Tracks progression through initialization, active assessment, and completion phases
- **Candidate Profile Evolution**: Maintains dynamic profile including strengths, weaknesses, performance trends, and preferred difficulty levels
- **Adaptive Decision Making**: Determines when to continue, adjust difficulty, or conclude assessment based on performance stability and skill coverage
- **Question Generation Coordination**: Manages batch generation of questions (max 2 at a time) based on real-time performance analysis

Key Data Structures:

```
candidate_profile = {
    "strengths": List[str],
    "weaknesses": List[str],
    "skill_trajectory": List[Dict],
```

```
    "preferred_difficulty": str,
    "areas_needing_focus": List[str]
}

trajectory_decisions = [{
    "timestamp": str,
    "reasoning": str,
    "trajectory": str,
    "questions_generated": int
}]
```

State Machine Flow:

1. **Initialization**: Profile creation, initial question generation
2. **Assessment Loop**: Response evaluation → Profile update → Decision point → Question generation
3. **Completion**: Final assessment, report generation

2. AI Agent Ecosystem

QueCraft Agent

Purpose: Adaptive question generation with contextual business scenario creation

Capabilities:

- **Case-Based Question Design**: Creates realistic business scenarios requiring Excel solutions
- **Contextual Data Integration**: Coordinates with Data Generator for scenario-specific datasets
- **Adaptive Reasoning**: Analyzes candidate profile to generate targeted questions
- **Progressive Complexity**: Adjusts question difficulty based on demonstrated capability

LLM Integration: Uses Gemini Pro for complex reasoning and scenario generation

Reviewer Agent

Purpose: Response evaluation and skill assessment

Capabilities:

- **Methodology Assessment**: Evaluates Excel approach rather than syntax perfection
- **Multi-Dimensional Scoring**: Assesses technical accuracy, problem-solving logic, and communication clarity
- **Follow-Up Detection**: Identifies when clarification questions are needed
- **Performance Tracking**: Maintains skill-specific performance history

Evaluation Framework:

- Excel Methodology (40%)
- Problem-Solving Logic (30%)
- Practical Application (20%)
- Communication Quality (10%)

Recruiter Agent

Purpose: Final decision making and comprehensive assessment synthesis

Capabilities:

- **Holistic Performance Analysis**: Synthesizes all assessment data for final recommendation
- **Risk Assessment**: Identifies potential concerns and hiring risks
- **Competency Mapping**: Maps performance to specific role requirements
- **Development Recommendations**: Provides targeted improvement suggestions

Interviewer Agent

Purpose: Conversational interface and candidate experience management

Capabilities:

- **Natural Conversation Flow**: Maintains engaging, supportive interview atmosphere
- **Question Presentation**: Formats technical scenarios in conversational manner
- **Encouragement and Guidance**: Provides hints and clarifications when needed
- **Session Management**: Handles interview transitions and conclusions

LLM Integration: Uses Gemini Flash for responsive, conversational interactions

3. Data Generator

The **MockDataGenerator** represents a significant innovation in assessment technology, using AI to create contextually perfect datasets for each question scenario.

Core Innovation:

Traditional assessment systems use static, generic datasets. Smart-Sheet generates datasets that perfectly match the specific Excel challenge being presented, including realistic business context and data quality issues.

Generation Process:

1. **Context Analysis**: LLM analyzes question requirements and business context
2. **Structure Design**: Creates appropriate columns, data types, and relationships
3. **Data Population**: Generates realistic values with intentional quality issues
4. **Format Conversion**: Outputs in HTML table format for web display

Supported Business Contexts:

- Marketing Campaign Analysis (ROAS calculations, performance metrics)
- Sales Performance Tracking (revenue analysis, trend identification)
- Financial Reporting (budget analysis, expense categorization)
- HR Analytics (employee performance, compensation analysis)
- Inventory Management (stock levels, reorder calculations)

Design Choices and Rationale

Multi-Agent Architecture Decision

Rationale: Different aspects of the assessment process require distinct types of AI reasoning:

- **Complex Reasoning Tasks** (Question generation, evaluation) → Gemini Pro
- **Conversational Tasks** (Interface management, encouragement) → Gemini Flash
- **Specialized Functions** (Data generation, web search) → Dedicated tools with AI integration

Benefits:

- **Separation of Concerns**: Each agent has a clear, focused responsibility

- **Scalability**: Individual agents can be optimized and scaled independently
- **Maintainability**: Changes to one assessment aspect don't affect others
- **Extensibility**: New agents can be added for additional capabilities

Dual LLM Strategy

Gemini Pro for Complex Reasoning:

- Question generation requiring deep business context understanding
- Performance evaluation with multi-dimensional analysis
- Adaptive trajectory decision making
- Contextual data structure design

Gemini Flash for Conversational AI:

- Real-time candidate interaction
- Question presentation and formatting
- Encouragement and hint provision
- Session flow management

Trade-offs Addressed:

- **Performance vs. Capability**: Flash provides faster responses for interactive elements while Pro handles complex reasoning
- **Cost Optimization**: Strategic use of more expensive Pro model only when sophisticated reasoning is required
- **User Experience**: Immediate conversational responses maintain engagement while complex tasks process in background

Real-Time Question Generation

Design Decision: Generate questions in small batches (max 2) based on real-time performance rather than pre-generating full question sets.

Benefits:

- **True Adaptivity**: Questions directly respond to demonstrated skills and weaknesses
- **Personalization**: Each candidate receives unique question sequence
- **Efficiency**: No wasted questions on already-demonstrated capabilities
- **Depth**: Can probe deeper into specific skill areas as needed

Technical Challenges Addressed:

- **Response Time**: Batch generation minimizes API calls while maintaining responsiveness
- **Quality Control**: Each generated question includes validation and fallback mechanisms
- **State Management**: Careful tracking of generated vs. presented vs. answered questions

Contextual Data Creation

Innovation: AI-generated datasets that perfectly match assessment scenarios rather than generic sample data.

Implementation Strategy:

- **LLM Context Analysis**: Understanding specific Excel challenge requirements
- **Business Scenario Matching**: Creating realistic data that supports the intended Excel exercise
- **Quality Issue Injection**: Including missing values, formatting inconsistencies, and other real-world data problems
- **Scale Optimization**: Right-sized datasets (30-100 rows) for assessment context

---

AI Model Integration

LLM Client Architecture

The system employs a sophisticated LLM client management system designed for reliability, performance, and cost optimization.

Rate Limiting Strategy

```
class RateLimiter:
    def __init__(self, max_requests: int, time_window: int = 60):
        self.max_requests = max_requests  # Pro: 15/min, Flash: 60/min
        self.time_window = time_window
        self.requests = []
```

Implementation Benefits:

- **Cost Control**: Prevents API cost spikes from excessive usage
- **Reliability**: Avoids rate limit violations that could disrupt assessments

- **Performance Optimization**: Batches requests efficiently within limits

Error Handling Framework

Multi-Level Error Recovery:

1. **Network Level**: Automatic retry with exponential backoff
2. **API Level**: Graceful degradation when services are unavailable
3. **Content Level**: Fallback to simpler prompts when complex requests fail
4. **User Experience Level**: Transparent error handling maintaining assessment flow

Model Selection Logic

Gemini Pro Use Cases:

- Adaptive question generation requiring complex business scenario understanding
- Multi-dimensional response evaluation with nuanced scoring
- Final hiring decision synthesis with comprehensive analysis
- Contextual dataset structure design

Gemini Flash Use Cases:

- Real-time conversational responses during assessment
- Question presentation and formatting
- Candidate encouragement and hint provision
- Session flow and transition management

Integration with MockDataGenerator

The Data Generator represents a breakthrough in assessment technology, using AI to understand question context and generate perfectly matched datasets.

LLM-Driven Data Design Process:

1. **Question Analysis Phase**:

```
prompt = f"""
Analyze this Excel interview question and design a dataset structure:
QUESTION: {question_context}

Design a dataset that:
```

1. Directly supports the Excel challenge described
2. Contains exact columns mentioned or implied
3. Includes realistic business data relevant to scenario
4. Has data quality issues making the challenge meaningful
"""

2. **Structure Generation**: LLM designs column schema with data types, business context, and quality issues
3. **Data Population**: Faker library generates realistic values following LLM-designed patterns
4. **Quality Issue Injection**: Intentional missing values, formatting inconsistencies, edge cases

Advanced Capabilities:

Business Context Awareness:

- Marketing scenarios → Campaign data with spend, revenue, ROAS calculations
- Financial analysis → Budget data with categories, variances, calculations
- HR analytics → Employee data with performance metrics, compensation analysis

Excel Challenge Alignment:

- VLOOKUP questions → Separate main and lookup tables with common keys
- Pivot table scenarios → Transactional data suitable for summarization
- Formula challenges → Data with edge cases requiring error handling

---

Assessment Process Design

Initialization Phase

Candidate Profile Creation:

```
candidate_profile = {
    "role_context": str,           # Business context for scenarios
    "initial_difficulty": str,     # Starting assessment level
    "focus_areas": List[str],      # Optional skill preferences
    "preferred_difficulty": str,   # Dynamically adjusted
    "strengths": List[str],        # Discovered during assessment
    "areas_needing_focus": List[str]  # Skills requiring development
```

}

Initial Question Generation:

- **Contextual Analysis**: Consider role, difficulty preference, focus areas
- **Batch Creation**: Generate 2 questions with complementary skill coverage
- **Dataset Integration**: Create contextual data for questions requiring it
- **Validation**: Ensure questions are appropriate for initial assessment

Adaptive Assessment Loop

Core Loop Structure:

1. **Question Presentation** → 2. **Response Collection** → 3. **Evaluation** → 4. **Profile Update** → 5. **Decision Point** → 6. **Question Generation** (if continuing)

Performance Tracking Mechanisms:

Real-Time Metrics:

- Skill-specific performance scores
- Response quality trends
- Time-to-competency in different areas
- Error pattern recognition

Adaptive Triggers:

- **Difficulty Adjustment**: 3+ consecutive high scores → increase difficulty
- **Skill Pivoting**: Demonstrated mastery → explore different skill areas
- **Remediation Focus**: Consistent struggles → reinforce fundamentals
- **Depth Exploration**: Strong foundation → advanced technique assessment

Decision Point Logic:

Continue Assessment Conditions:

- Skill coverage incomplete (< 5 different Excel areas tested)
- Performance trajectory still evolving
- Candidate demonstrating learning/improvement
- Maximum question limit not reached

Completion Triggers:

- Performance stabilization (score variance < 15 points over 4 questions)
- Adequate skill coverage achieved
- Clear competency level established
- Maximum assessment time/questions reached

Personalized Feedback Generation

Multi-Dimensional Analysis:

- **Skill Competency Mapping**: Excel proficiency across different functional areas
- **Problem-Solving Assessment**: Analytical approach and methodology quality
- **Communication Evaluation**: Ability to explain Excel approaches clearly
- **Growth Trajectory Analysis**: Learning patterns and improvement indicators

Adaptive Report Features:

- **Personalized Recommendations**: Targeted skill development based on performance gaps
- **Role-Specific Insights**: Assessment results contextualized for specific business roles
- **Learning Path Suggestions**: Structured approach to addressing identified weaknesses
- **Strength Amplification**: Strategies for leveraging demonstrated Excel capabilities

---

Technical Infrastructure

Core Dependencies and Technology Stack

Backend Framework:

- **Streamlit**: Web application framework providing rapid prototyping and deployment capabilities
- **Python 3.8+**: Core runtime environment with comprehensive library ecosystem
- **Google Generative AI SDK**: Direct integration with Gemini Pro and Flash models

Data Processing:

- **Pandas**: DataFrame operations for dataset generation and manipulation

- **Faker**: Realistic test data generation with business context awareness
- **JSON5**: JSON parsing for LLM response handling

AI and ML Libraries:

- **google-generativeai**: Official Google AI integration library
- **requests**: HTTP client for web search and external API integration
- **python-dotenv**: Environment variable management for secure API key handling

Scalability Architecture

Concurrent Assessment Handling:

Session Isolation:

- Each assessment maintains independent state with unique session IDs
- Candidate profiles stored in memory with minimal cross-session dependencies
- Tool logging provides per-session performance tracking

Resource Management:

- **Rate Limiting**: Per-client API throttling prevents resource exhaustion
- **Memory Optimization**: Efficient data structures minimizing memory footprint
- **Garbage Collection**: Automatic cleanup of completed assessment sessions

Performance Monitoring:

Tool Performance Tracking:

```
class ToolLogger:
    def log_tool_call(self, tool_name, agent_name, execution_time, success):
        # Comprehensive logging of all AI agent interactions
        # Performance metrics for optimization
        # Error pattern analysis for reliability improvements
```

System Health Monitoring:

- API response time tracking
- Success rate monitoring across different agents
- Resource utilization metrics
- Error frequency and pattern analysis

Deployment Strategies

Local Development Environment:

```
# Environment Setup
git clone https://github.com/adityalakhani/Smart-Sheet
pip install -r requirements.txt
export GOOGLE_API_KEY=your_api_key
streamlit run app.py
```

Production Deployment Considerations:

Scalability Requirements:

- **Horizontal Scaling**: Multiple Streamlit instances behind load balancer
- **Database Integration**: PostgreSQL or MongoDB for persistent candidate data
- **Caching Layer**: Redis for session state management and response caching
- **Container Orchestration**: Docker containers with Kubernetes for auto-scaling

Security Implementation:

- **API Key Management**: Secure key rotation and environment-based configuration
- **Data Encryption**: TLS/SSL for all communications, encrypted data storage
- **Access Control**: Authentication and authorization for administrative functions
- **Audit Logging**: Comprehensive activity tracking for compliance

Monitoring and Observability:

- **Application Performance Monitoring**: New Relic or Datadog integration
- **Error Tracking**: Sentry for real-time error monitoring and alerting
- **Business Metrics**: Custom dashboards for assessment completion rates, user satisfaction
- **Infrastructure Monitoring**: Server resource utilization and availability

---

Future Enhancements

Planned Features

1. Gemini Live API Integration for Voice-Enabled Assessments

Vision: Transform text-based assessment into natural conversation using Gemini Live API for real-time voice interaction.

Implementation Strategy:

- **Google Meet Integration**: Seamless video conferencing with AI interviewer
- **Real-Time Speech Processing**: Natural language understanding of candidate explanations
- **Conversational Assessment Flow**: Dynamic question asking based on verbal responses
- **Multi-Modal Analysis**: Combined analysis of verbal explanations and screen sharing

Technical Architecture:

```
class VoiceAssessmentOrchestrator:
    def __init__(self):
        self.gemini_live_client = GeminiLiveAPI()
        self.meet_integration = GoogleMeetAPI()
        self.speech_processor = SpeechToTextProcessor()

    async def conduct_voice_assessment(self, candidate_id):
        # Real-time voice interaction with adaptive questioning
        # Natural conversation flow with AI interviewer
        # Dynamic assessment based on verbal explanations
```

Benefits:

- **Natural Interaction**: More realistic interview experience
- **Nuanced Assessment**: Voice tone and communication style analysis
- **Accessibility**: Better for candidates with typing difficulties
- **Engagement**: More interactive and engaging assessment experience

2. In-Screen Excel Proctoring System

Purpose: Real-time monitoring of Excel usage during practical assessments to prevent malpractice and provide better evaluation metrics.

Core Components:

Screen Monitoring Module:

- **Excel Window Tracking**: Monitor Excel application usage and activity
- **Formula Auditing**: Track formula creation and modification patterns
- **Function Usage Analysis**: Analyze which Excel functions are being used
- **Time-on-Task Metrics**: Detailed timing analysis of different assessment phases

Behavioral Analysis Engine:

- **Natural Usage Patterns**: Establish baselines for typical Excel navigation
- **Anomaly Detection**: Identify suspicious patterns like copy-paste from external sources
- **Skill Verification**: Confirm that claimed approaches match actual Excel usage
- **Performance Validation**: Cross-reference verbal explanations with actual actions

Implementation Framework:

```
class ExcelProctoringSystem:
    def __init__(self):
        self.screen_monitor = ScreenCaptureAPI()
        self.excel_analyzer = ExcelActivityAnalyzer()
        self.behavior_engine = BehaviorAnalysisEngine()

    def monitor_assessment(self, session_id):
        # Real-time Excel activity monitoring
        # Behavioral pattern analysis
        # Malpractice detection and flagging
```

3. Assessment Reporting

Advanced Analytics Dashboard:

- **Performance Heat Maps**: Visual representation of Excel skill proficiency across different areas
- **Learning Trajectory Visualization**: Graphical display of skill development during assessment
- **Comparative Analysis**: Benchmarking against role-specific competency standards
- **Predictive Insights**: ML-based predictions of job performance based on assessment results

Detailed Competency Reports:

- **Skill Gap Analysis**: Precise identification of knowledge gaps with remediation suggestions
- **Growth Potential Assessment**: Evaluation of learning ability and improvement trajectory
- **Role Readiness Scoring**: Specific scoring for different business analyst, data analyst, or finance roles
- **Custom Report Generation**: Tailored reports for different stakeholders (HR, hiring managers, candidates)

Technical Improvements

1. AI Model Advancement

Multi-Model Integration:

- **Gemini 2.5 Pro –** Advanced Multimodal Reasoning model for agents with complex tasks.
- **Gemini 2.5 Flash –** Efficient Performance with Balanced Cost and Latency.
- **Gemini 2.5 Flash Lite –** Lightweight Model for High-Volume, Low-Latency Applications.

Advanced Reasoning Capabilities:

- **Chain-of-Thought Prompting**: Enhanced reasoning for complex Excel scenarios
- **Few-Shot Learning**: Better question generation using exemplar-based learning
- **Multi-Agent Collaboration**: Enhanced coordination between different AI agents

2. Extended Skill Coverage

Microsoft Office Suite Expansion:

- **PowerPoint Assessment**: Presentation design and data visualization skills
- **Word Assessment**: Document formatting, mail merge, and template creation
- **Outlook Assessment**: Email management and productivity features
- **Power BI Integration**: Advanced analytics and business intelligence capabilities

Advanced Excel Features:

- **Power Query Assessment**: Data transformation and ETL capabilities
- **VBA Evaluation**: Basic programming and automation skills
- **Excel Add-in Knowledge**: Assessment of third-party tool integration
- **Advanced Charting**: Sophisticated data visualization techniques

Conclusion

Smart-Sheet represents a significant advancement in skills assessment technology, combining adaptive AI, real-time question generation, and contextual data creation to provide unprecedented personalization in Excel proficiency evaluation. The multi-agent architecture ensures scalable, maintainable, and extensible system design while the dual LLM strategy optimizes both performance and capability.

The planned enhancements, particularly voice-enabled assessment and Excel proctoring, position Smart-Sheet as a comprehensive solution for modern talent evaluation needs. The technical infrastructure supports both current requirements and future growth, making it suitable for enterprise deployment and continued innovation.

This design documentation serves as the foundation for continued development, integration planning, and system evolution as the platform adapts to changing business needs and technological capabilities.