



MORGAN STANLEY

# Automated KYC/AML regulatory reporting program

A SUBMISSION FOR THE CFG FINAL PROJECT

*ALEJANDRA CUADROS RIVAS  
ANA CAROLINA BERGAMASCO PEREZ  
BRUNA ROCHA  
CHIAMAKA MARIA-GORETTI ADINNU  
CLARA MARIA BARBOSA TEODORO  
LILLA HORVÁTH*

*ARTHUR KALULE  
PROJECT MARKER*

*DECEMBER 12, 2024*

# WHAT WE BUILT

WE BUILT AN **AUTOMATED KYC/AML REGULATORY REPORTING PROGRAM** TO EVALUATE IF STARTUPS MEET MORGAN STANLEY'S INCLUSIVE VENTURE LAB (MSIVL) REQUIREMENTS AND COMPLY WITH ANTI-MONEY LAUNDERING (AML) REGULATIONS. THE PROGRAM WILL PRIMARILY AUTOMATE ELIGIBILITY AND COMPLIANCE CHECKS, FOCUSING ON KEY ASPECTS:

- DATA EXTRACTION, TRANSFORMATION, AND VALIDATION FOR REGULATORY COMPLIANCE
- KYC FUNCTIONALITIES: VERIFYING UNDERREPRESENTED FOUNDERS, TECHNOLOGY FOCUS, REVENUE GENERATION, AND ALIGNMENT WITH SUSTAINABILITY THEMES, CHECK STARTUPS AGAINST AML AUTHORITIES TO ENSURE THERE ARE NO EXISTING ISSUES THAT COULD POSE RISKS.

WITH THIS SOLUTION, MORGAN STANLEY CAN STREAMLINE ITS DUE DILIGENCE, IDENTIFYING STARTUPS THAT MEET REGULATORY AND INVESTMENT STANDARDS, PARTICULARLY IN SOCIAL IMPACT AND MONEY LAUNDERING COMPLIANCE.

# WHAT WE SOLVE

- **AUTOMATED COMPLIANCE VERIFICATION:** CHECKS IF STARTUPS ALIGN WITH MSIVL'S REQUIREMENTS, INCLUDING DIVERSITY IN LEADERSHIP, TECH FOCUS, AND SUSTAINABILITY GOALS.
- **AML AUTHORITY SCREENING:** CONFIRMS IF STARTUPS HAVE A CLEAN RECORD WITH AML AUTHORITIES, MITIGATING POTENTIAL RISKS ASSOCIATED WITH MONEY LAUNDERING.
- **IMPACT AND SUSTAINABILITY ASSESSMENT:** VALIDATES STARTUP CONTRIBUTIONS TO ENVIRONMENTAL AND SOCIAL GOALS, ENSURING ALIGNMENT WITH MSIVL'S MISSION.
- **EFFICIENCY AND RELIABILITY:** RIGHT NOW THE PROCESS DEMANDS A HUMAN WORKFORCE, WITH THIS PROGRAM THE COMPLIANCE CHECKS ARE AUTOMATED, IMPROVING SPEED AND ACCURACY.

# KEY FEATURES

- **ELIGIBILITY VERIFICATION MODULE:** CHECKS STARTUPS FOR REQUIRED CRITERIA LIKE UNDERREPRESENTED FOUNDERS, TECH FOCUS, REVENUE GENERATIONS, AND SUSTAINABILITY ALIGNMENT.
- **AML COMPLIANCE SCREENING:** CONDUCTS BACKGROUND CHECKS AGAINST AML AUTHORITIES TO ENSURE STARTUPS OR THEIR FOUNDERS/MANAGERS ARE NOT FLAGGED FOR MONEY LAUNDERING RISKS.
- **SUSTAINABILITY ASSESSMENT:** EVALUATE STARTUPS ON MSIVL'S FOCUS AREAS (E.G., CLIMATE CHANGE, CARBON CREDITS, WASTE MANAGEMENT).
- **REAL-TIME COMPLIANCE DASHBOARD:** SHOWS COMPLIANCE STATUS AND AML-RELATED ISSUES.
- **AUTOMATED REPORTING:** CREATES REPORTS SUMMARISING THE STARTUP'S ELIGIBILITY AND AML COMPLIANCE FOR MSIVL APPLICATIONS.

## FUNCTIONAL REQUIREMENTS

### DATA INPUT:

1. THE SYSTEM MUST ALLOW THE INPUT OF DATA FOR CLIENTS, FOUNDERS, AND STARTUPS.

### DATA VALIDATION:

1. THE SYSTEM MUST VALIDATE ALL INSERTED DATA TO ENSURE ACCURACY AND INTEGRITY.
  - DURING THE VALIDATION PROCESS, THE SYSTEM MUST VERIFY THE FOUNDERS' RECORDS AND THE ORIGIN (COUNTRY) OF THE ASSOCIATED COMPANIES.

### DATA VISUALIZATION FOR ANALYSTS:

1. THE SYSTEM MUST PROVIDE A DATA VISUALIZATION INTERFACE ENABLING MORGAN STANLEY ANALYSTS TO REVIEW COMPANIES' DATA.

### DATA EXPORT:

1. THE SYSTEM MUST SUPPORT EXPORTING DATA FROM THE ANALYST DASHBOARD INTO PDF AND CSV FORMATS.

## NON-FUNCTIONAL REQUIREMENTS

### USABILITY AND INTERFACE:

1. THE SYSTEM SHOULD PRESENT AN INTUITIVE AND USER-FRIENDLY INTERFACE, ENSURING THAT
2. NON-TECHNICAL USERS (E.G., ANALYSTS) CAN EASILY NAVIGATE AND OPERATE THE PLATFORM.

### TECHNOLOGY STACK - BACKEND AND DATABASE:

1. THE SYSTEM'S DATABASE MUST BE IMPLEMENTED USING MYSQL.
2. THE DATA HANDLING AND PROCESSING SHOULD BE DONE USING PYTHON SCRIPTS.

### FRONTEND FRAMEWORK:

1. THE SYSTEM'S FRONTEND SHOULD BE BUILT USING THE STREAMLIT LIBRARY, PROVIDING A STRAIGHTFORWARD AND INTERACTIVE WEB-BASED INTERFACE.

### API AND CONNECTIVITY:

1. THE SYSTEM SHOULD ESTABLISH CONNECTIONS TO THE DATABASE AND OTHER BACK-END SERVICES USING THE REQUESTS LIBRARY TO ENSURE EFFICIENT COMMUNICATION AND DATA RETRIEVAL.
2. THE PEP/POI API SHOULD BE ACCESSED VIA SECURE HTTP REQUESTS TO ENSURE RELIABLE AND TIMELY RESPONSES FOR DATA VALIDATION.



Analyst dashboard screen.

## FUNCTIONAL REQUIREMENTS

### DATA INPUT:

THE SYSTEM MUST ALLOW THE INPUT OF DATA FOR CLIENTS, FOUNDERS, AND STARTUPS.

### DATA VALIDATION:

1. THE SYSTEM MUST VALIDATE ALL INSERTED DATA TO ENSURE ACCURACY AND INTEGRITY.

DURING THE VALIDATION PROCESS, THE SYSTEM MUST VERIFY THE FOUNDERS' RECORDS AND THE ORIGIN (COUNTRY) OF THE ASSOCIATED COMPANIES.

### DATA VISUALIZATION FOR ANALYSTS:

THE SYSTEM MUST PROVIDE A DATA VISUALIZATION INTERFACE ENABLING MORGAN STANLEY ANALYSTS TO REVIEW COMPANIES' DATA.

### DATA EXPORT:

1. THE SYSTEM MUST SUPPORT EXPORTING DATA FROM THE ANALYST DASHBOARD INTO PDF AND CSV FORMATS.

## NON-FUNCTIONAL REQUIREMENTS

### USABILITY AND INTERFACE:

1. THE SYSTEM SHOULD PRESENT AN INTUITIVE AND USER-FRIENDLY INTERFACE, ENSURING THAT

2. NON-TECHNICAL USERS (E.G., ANALYSTS) CAN EASILY NAVIGATE AND OPERATE THE PLATFORM.

### TECHNOLOGY STACK - BACKEND AND DATABASE:

1. THE SYSTEM'S DATABASE MUST BE IMPLEMENTED USING MYSQL.

THE DATA HANDLING AND PROCESSING SHOULD BE DONE USING PYTHON SCRIPTS.

### FRONTEND FRAMEWORK:

1. THE SYSTEM'S FRONTEND SHOULD BE BUILT USING THE STREAMLIT LIBRARY, PROVIDING A STRAIGHTFORWARD AND INTERACTIVE WEB-BASED INTERFACE.

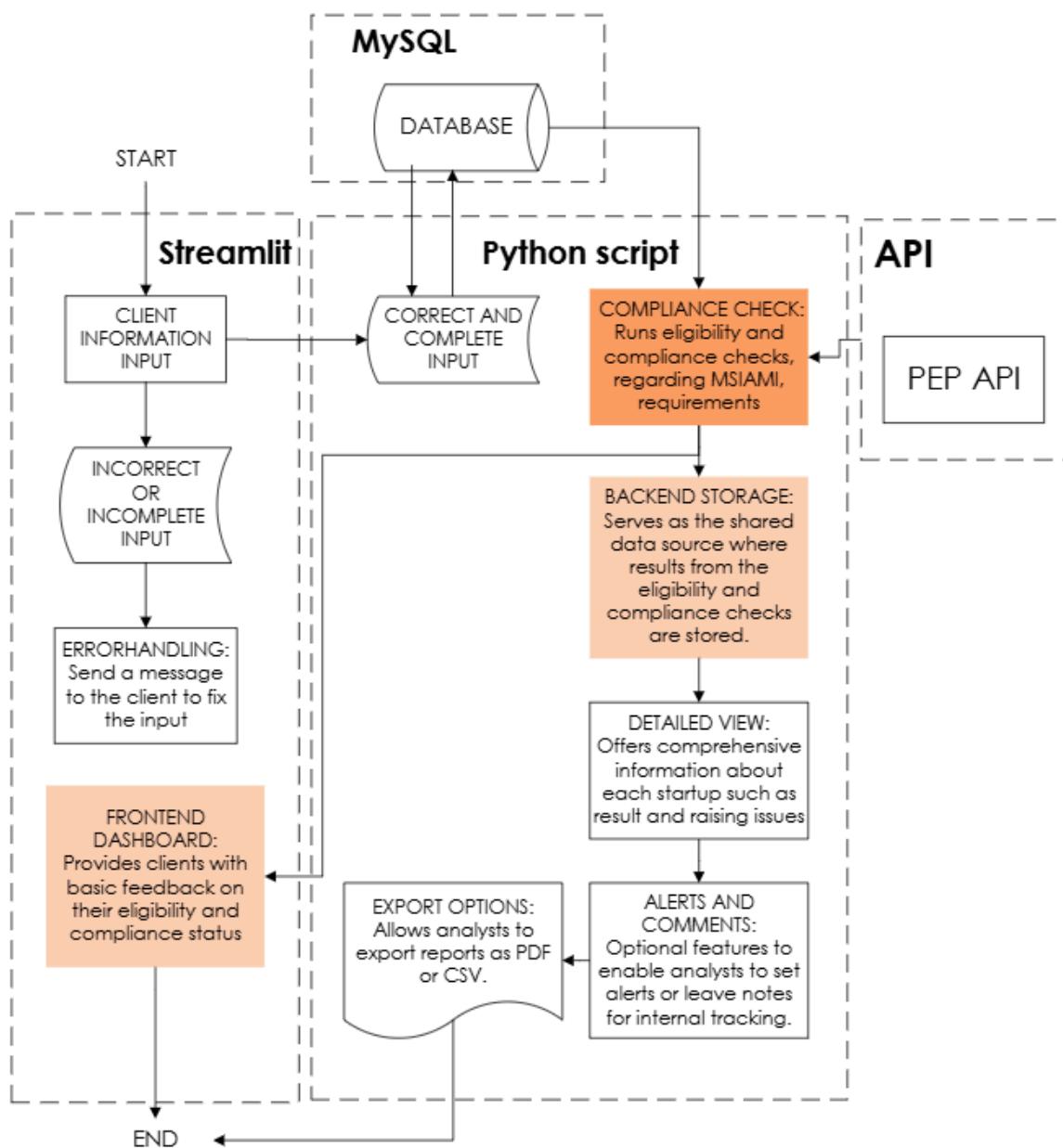
### API AND CONNECTIVITY:

1. THE SYSTEM SHOULD ESTABLISH CONNECTIONS TO THE DATABASE AND OTHER BACK-END SERVICES USING THE REQUESTS LIBRARY TO ENSURE EFFICIENT COMMUNICATION AND DATA RETRIEVAL.

2. THE PEP/POI API SHOULD BE ACCESSED VIA SECURE HTTP REQUESTS TO ENSURE RELIABLE AND TIMELY RESPONSES FOR DATA VALIDATION.

Must Have	Should Have	Could Have	Won't have
✓ PEP Screening	◆ Enhanced Risk Scoring	◆ Multi-Language Support	✗ Manual Document Review
✓ High-Risk Country Check	◆ Advanced Fraud Detection	◆ Self-Service Portal	✗ Customizable Risk Rules
✓ User Input Validation	◆ Notifications for Errors	◆ Multi-Factor Authorization	✗ Manual Audits
✓ Automated Report Generation	◆ PDF and CSV formats	◆ Machine Learning-Based Behavioral Analytics	✗ Offline Processing Capabilities

# SYSTEM WORKFLOW



## DEVELOPMENT APPROACH AND TEAM MEMBER ROLES

ROLES WERE ASSIGNED BASED ON EACH TEAM MEMBER'S AVAILABILITY AND PREFERENCES. IF SOMEONE COULD START A TASK IMMEDIATELY, THEY WERE ASSIGNED EARLIER-STAGE WORK; OTHERWISE, THEY TOOK ON TASKS SCHEDULED TO BEGIN LATER. THIS APPROACH ALLOWED FOR FLEXIBLE RESOURCE ALLOCATION AND ENSURED CONTINUOUS PROGRESS. TASK MANAGEMENT AND ORGANIZATION WERE HANDLED USING THE JIRA PLATFORM.

## TOOLS AND LIBRARIES

THE PRIMARY LIBRARIES EMPLOYED INCLUDED:

- PANDAS FOR DATA MANIPULATION AND ANALYSIS
- STREAMLIT FOR BUILDING THE INTERACTIVE FRONTEND
- MYSQL-CONNECTOR-PYTHON FOR DATABASE CONNECTIVITY
- REQUESTS FOR API INTERACTIONS

THE MAIN TOOLS USED WERE:

- PYTHON AS THE CORE PROGRAMMING LANGUAGE
- MYSQL FOR DATABASE MANAGEMENT
- VISUAL STUDIO CODE (VSCODE) AS THE INTEGRATED DEVELOPMENT ENVIRONMENT (IDE)

## IMPLEMENTATION PROCESS AND AGILE DEVELOPMENT

THE TEAM FOLLOWED AN ITERATIVE APPROACH, REFLECTING ELEMENTS OF AGILE METHODOLOGIES. WE UTILIZED JIRA FOR TASK TRACKING, HELD FREQUENT, FOCUSED MEETINGS WITH RELEVANT TEAM MEMBERS FOR COLLABORATIVE PROBLEM-SOLVING, AND CONDUCTED INFORMAL CODE REVIEWS. THESE PRACTICES HELPED US MAKE TIMELY ADJUSTMENTS, REFINE OUR CODE, AND MAINTAIN CLARITY ON RESPONSIBILITIES AND PROGRESS.

## IMPLEMENTATION CHALLENGES

A SIGNIFICANT CHALLENGE WAS LIMITED AVAILABILITY AND COMPETING COMMITMENTS. MANY TEAM MEMBERS WERE JUGGLING TIGHT DEADLINES, PURSUING MASTER'S DEGREES, BALANCING WORK RESPONSIBILITIES, AND MANAGING PERSONAL ISSUES. THIS CONSTRAINT REQUIRED ADAPTABILITY, CLEAR COMMUNICATION, AND FLEXIBLE TASK SCHEDULING TO ENSURE THE PROJECT STAYED ON TRACK.

## TESTING STRATEGY

THE ORIGINAL INTENTION WAS TO WRITE UNIT TESTS, INTEGRATION TESTS AND END-TO-END TESTS AS WELL IN ORDER TO COVER ALL POSSIBLE SCENARIOS AND TESTS EVERY POSSIBLE OUTCOME. UNFORTUNATELY, DUE TO THE COMPLEXITY OF THESE TESTS AND THE LIMITED TIME, ONLY THE UNIT TESTS WERE FINISHED. SOME OF THE UNIT TESTS USED MOCKING TO MIMIC INTERACTION BETWEEN THE API AND THE DATABASE, USING THE UNITTESTS.MOCK PYTHON LIBRARY. OTHER UNIT TESTS USED THE UNITTESTS LIBRARY. ALL TESTS WERE RUN IN THE VENV VIRTUAL LIBRARY IN PYCHARM.

THE TESTING PROCEDURE STARTED SOON AFTER THE FIRST TWO BRANCHES CONTAINING PYTHON FILES WERE MERGED ON THE MAIN BRANCH OF THE GROUP'S GITHUB REPOSITORY. THE TESTS WERE CONTINUOUSLY IMPROVED TO ENSURE RELIABILITY. THE TEST RESULTS WERE SHARED WITH THE REST OF THE GROUP IN ORDER TO IMPROVE THE TESTED FILES AND CORRECT ANY BUGS OR OTHER SORT OF UNEXPECTED BEHAVIOUR. THIS PROCESS WAS REPEATED AFTER EACH UPDATE.

## FUNCTIONAL AND USER TESTING

DUE TO LACK OF RESOURCES, ESPECIALLY TIME, ONLY FUNCTIONAL TESTS WERE MADE DURING THE TESTING PROCESS. ALL THE FUNCTIONAL TESTS WERE UNIT TESTS, SOME OF THEM WERE MOCK TESTS, AS MENTIONED ABOVE. IF THERE WAS MORE TIME FOR THE PROJECT, USER TESTS WOULD ALSO HAVE BEEN MADE IN ORDER TO ENSURE THAT THE PROGRAM WORKS PROPERLY WHEN REAL USERS USE IN REAL WORLD SCENARIOS.

## SYSTEM LIMITATIONS

SINCE UNIT TESTS ARE NOT COMPLEX ENOUGH, THEY CAN ONLY BE USED TO TEST WHETHER FILES ARE WORKING SEPARATELY. IN ORDER TO CHECK THE WHOLE PROGRAM, INTEGRATION TESTS AND END-TO-END TESTS ARE ALSO NEEDED. UNFORTUNATELY, THE LACK OF TIME PROHIBITED THE CREATION OF MORE COMPLEX TESTS. IF I HAD MORE TIME FOR THE PROJECT, I WOULD HAVE CREATED AT LEAST INTEGRATION TESTS TO ENSURE RELIABILITY AND TO VERIFY THAT THE PROGRAM FILES WORK CORRECTLY IN INTEGRATION AS WELL. WITH MORE TIME, I WOULD ALSO LIKE TO INCLUDE USER TESTING AS A FINAL STEP FOR THE TESTING PROCESS.

- STREAMLINED KYC/AML COMPLIANCE FOR STARTUPS:

THE PLATFORM PROVIDES A SIMPLIFIED, END-TO-END COMPLIANCE SOLUTION TAILORED TO STARTUP REQUIREMENTS, ENSURING EFFICIENT KNOW YOUR CUSTOMER (KYC) AND ANTI-MONEY LAUNDERING (AML) PROCESSES.

- FULLY AUTOMATED WORKFLOW:

THE SYSTEM AUTOMATES EVERY STEP—FROM INITIAL DATA INPUT AND SCREENING (INCLUDING PEP/POI CHECKS AND HIGH-RISK COUNTRY VALIDATIONS) THROUGH TO GENERATING A RISK SCORE AND FINAL COMPLIANCE REPORT—MINIMIZING MANUAL INTERVENTION.

- ENHANCED ANALYST CAPABILITIES FOR MORGAN STANLEY:

COMPREHENSIVE DATA VISUALIZATION AND EXPORT OPTIONS ENABLE MORGAN STANLEY'S ANALYSTS TO MAKE INFORMED DECISIONS QUICKLY AND CONFIDENTLY, SIGNIFICANTLY REDUCING MANUAL EFFORTS AND EXPEDITING THEIR DUE DILIGENCE TASKS.

- SCALABLE AND ADAPTIVE DESIGN:

WHILE MEETING THE CURRENT REQUIREMENTS, THE PLATFORM'S MODULAR ARCHITECTURE ALLOWS FOR FUTURE ENHANCEMENTS, SUCH AS ADVANCED FRAUD DETECTION OR MACHINE LEARNING INTEGRATIONS, ENSURING LONG-TERM ADAPTABILITY.

- UAT-READY FOR MORGAN STANLEY REVIEW:

THE SYSTEM IS PREPARED FOR USER ACCEPTANCE TESTING (UAT) BY MORGAN STANLEY REVIEWERS, REFLECTING READINESS FOR PRACTICAL EVALUATION, FEEDBACK, AND FINAL REFINEMENT BEFORE FULL-SCALE DEPLOYMENT.