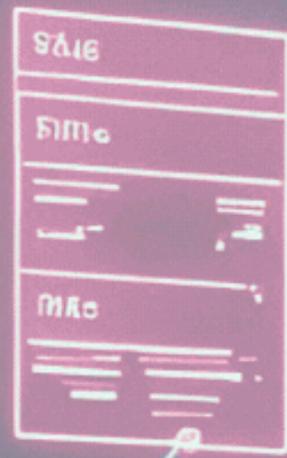
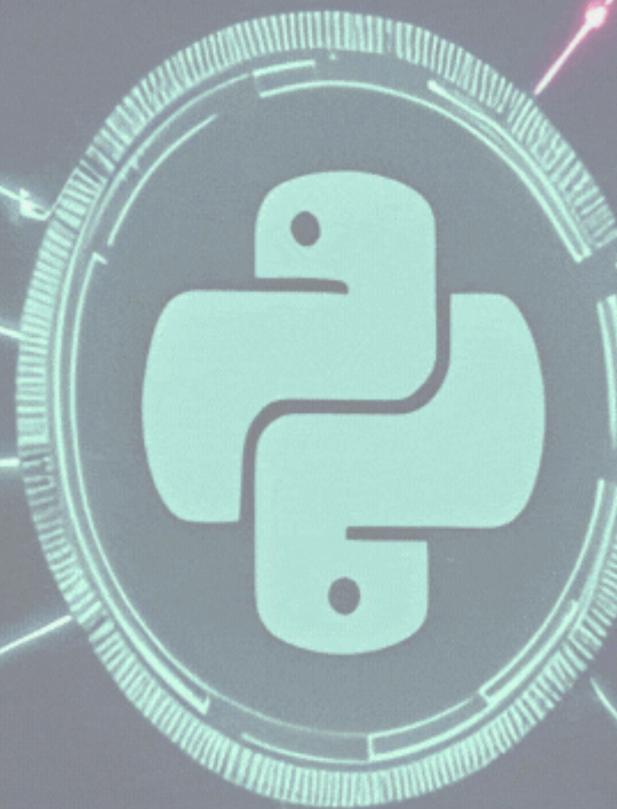


Process executed data extraction



# TINDER

Profile, message, like, swipe

# PROFILE

# SCRAPE

R7



# Project Work Plan

Company:	Excel Work House (Data Processor)
Department:	Programming & Data Management
Project Manager:	<b>Md Rehmanul Alam</b>
Date:	<b>7th March 2025</b>

## Introduction

### a. Project Overview:

The Tinder Profile Scraper project aims to develop an automated system to collect and label female Tinder profiles for research purposes. The system will collect 25,000 profiles with 5+ images each, implement location rotation for diverse sampling, and provide a comprehensive labeling interface for categorizing profiles according to specific attributes.

### b. Objectives:

1. Develop a system to collect 25,000 female Tinder profiles with 5+ images each.
2. Implement location rotation to ensure diverse profile collection across 20+ global cities.
3. Create a labeling interface for categorizing profiles according to 19 specific criteria.
4. Integrate with Google Sheets for data storage, logging, and reporting.
5. Complete the project in 5 milestones of 5,000 profiles each.

### c. Scope:

The project encompasses the development of a web-based application for scraping Tinder profiles, downloading and organizing images, labeling profiles, and generating statistics and reports. It includes integration with Google Sheets but excludes computer vision model training and synthetic image generation using the collected data.

## II. Project Scope

### a. Deliverables:

- Functional web application with scraping, viewing, and labeling capabilities.
- Automated scraping system with location rotation functionality.
- Image downloading and storage system with proper naming conventions.
- User-friendly labeling interface with all required attribute categories.
- Integration with Google Sheets for data storage and logging.
- Statistical dashboard for tracking progress and data distribution.

**b. Inclusions:**

- Development of Python-based web application and scraping modules.
- Integration with Tinder via automated browser interaction.
- Google Sheets API integration for data storage.
- Automated location rotation functionality.
- Image downloading and storage system.
- User interface for profile labeling.
- Logging and reporting functionality.

**c. Exclusions:**

- Training of computer vision models using the collected data.
- Generation of synthetic profiles using the collected data.
- Integration with other dating platforms besides Tinder.
- Ongoing maintenance after project completion.
- Mobile application development.

**d. Assumptions:**

- Availability of a valid Tinder Premium account with Passport feature.
- Sufficient server resources for continuous operation.
- Stable internet connection for scraping operations.
- Google Cloud Platform access for API integration.

## **III. Project Schedule**

**a. Milestones and Timeline:**

Milestone	Timeline
Project Setup	Weeks 1-2
Core Development	Weeks 3-6
UI Development	Weeks 7-9
Testing and Deployment	Weeks 10-12
First 5,000 Profiles Milestone	Week 5
Second 5,000 Profiles Milestone	Week 7
Third 5,000 Profiles Milestone	Week 9
Fourth 5,000 Profiles Milestone	Week 11
Final 5,000 Profiles Milestone	Week 12

### b. Dependencies:

- Tinder account setup must be completed before scraping begins.
- Google Cloud project setup required before Google Sheets integration.
- Core scraper development must be completed before UI development.
- Testing must be performed on small sample before full-scale scraping.

///{The 12-week timeline was developed based on several factors related to the project's complexity and data collection requirements. However, you're right to question whether this could be optimized. Let me explain the reasoning and potential adjustments:

### Reasons for the 12-Week Timeline

1. **Data Collection Volume:** Collecting 25,000 profiles with 5+ images each represents a significant volume of data. This requires not just technical development but also actual runtime for the scraper to collect profiles at a sustainable rate.
2. **Technical Complexity:** The project involves multiple complex components: Selenium-based web scraping, automated browser control, image downloading and storage, Google Sheets integration, and a full web application interface.
3. **Rate Limitations:** To avoid detection and account blocking, the scraper needs to operate at a measured pace with human-like behavior patterns and appropriate delays between actions. This extends the collection timeframe.
4. **Location Rotation:** Rotating through 20+ global locations adds complexity and time, as each location change requires adjustment periods and may affect profile availability.
5. **Quality Assurance:** Testing and refinement are essential to ensure reliable data collection and minimize errors, especially given Tinder's likely anti-scraping measures.

## Potential Timeline Optimization

If you need to accelerate the timeline, these adjustments could be considered:

1. **Parallel Development:** Increase development resources to work on multiple components simultaneously (scraper, UI, database integration).
2. **Multiple Accounts:** Use several Tinder Premium accounts simultaneously to increase collection rate.
3. **Simplified First Version:** Develop a minimal viable product first, then add additional features while data collection is already running.
4. **Reduced Scope:** Consider whether 25,000 profiles is necessary for initial purposes - starting with 10,000 would significantly reduce the timeline.
5. **Cloud Resources:** Utilize cloud infrastructure for faster processing and parallel operations.

With these optimizations, the timeline could potentially be reduced to 6-8 weeks, though this would require additional resources and might introduce more risk of detection by Tinder's systems.

Would you like me to propose a revised timeline based on any of these optimization strategies?}///

## IV. Resource Planning

### a. Human Resources:

- Project Manager
- Backend Developer (Python/Selenium expertise)
- Frontend Developer (HTML/CSS/JavaScript)
- QA Tester

#### **b. Roles and Responsibilities:**

Role	Responsibilities
Project Manager	Overall project coordination, timeline management, stakeholder communication.
Backend Developer	Core scraper development, Google Sheets integration, image processing functionality.
Frontend Developer	Dashboard creation, profile viewer, labeling interface, statistics visualization.
QA Tester	System testing, performance validation, error identification and reporting.

#### **c. Staffing Requirements:**

- Full-time commitment required for Backend and Frontend Developers.
- Part-time allocation for Project Manager and QA Tester.
- Additional resources may be required during peak development periods.

## **V. Risk Management**

#### **a. Risks:**

1. Tinder UI changes affecting scraper functionality.
2. Account blocking due to automated activity detection.
3. Rate limiting impacting data collection speed.
4. Data privacy concerns related to profile information.
5. Resource constraints affecting system performance.

#### **b. Mitigation Strategies:**

- Implement adaptive selectors and regular monitoring for UI changes.
- Use human-like behavior patterns, random delays between actions, and multiple accounts if necessary.

- Implement exponential back off strategies and optimize request timing.
- Ensure proper anonymization of profile data in storage and reporting.
- Utilize cloud resources with scaling capabilities as needed.

## VI. Communication Plan

### a. Stakeholder Communication:

- Regular project updates through email and bi-weekly status meetings.
- Demonstration of progress at each 5,000 profile milestone.
- Technical documentation shared through secure project repository.

### b. Reporting Frequency:

- Weekly progress reports on scraping statistics and development status.
- Bi-weekly stakeholder meetings for comprehensive project review.
- Ad-hoc reporting for critical issues or significant achievements.

### c. Escalation Procedures:

- Any project issues beyond the Project Manager's control will be escalated to the Project Sponsor.
- Technical challenges that impact timeline will be documented and reported immediately.
- Budget or resource constraints requiring adjustment will trigger stakeholder consultation.

## VII. Monitoring and Control

### a. Performance Metrics:

- Profiles scraped per day/week
- Images collected per profile
- System uptime percentage
- Error rate during scraping operations
- Labeling completion percentage

### b. Project Tracking:

- Weekly progress reviews to ensure adherence to project schedule.
- Daily monitoring of scraping operations and error logs.

- Regular code reviews to maintain quality and identify improvements.

**c. Change Management:**

- Formal change request process for any modifications to the project scope or schedule.
- Impact assessment required for all proposed changes.
- Approval required from Project Sponsor for significant adjustments.

**Prepared by:**

Md Rehmanul Alam, Project Manager

Date: 7th March 2025

by:

Rehmanul



**T** TEMPLATE.NET X