**1. Overview & Purpose**

**Objective**: Enable users to enrich a CSV dataset of soil sampling locations (with pointid, lat, lon, plus any custom columns) by fetching relevant geospatial variables from Google Earth Engine (GEE).

**Key Features**:

* CSV import and flexible column mapping for pointid, latitude, and longitude.
* Dynamic user selection of Earth Engine layers via simple search.
* Batch data retrieval using Python Earth Engine (or similar) libraries.
* Seamless authentication via a configured Google service account—users are not required to authenticate manually.
* Built-in handling for GEE rate limits and quotas, including batching logic.
* Real-time progress feedback during enrichment.

**2. User Flow / Functional Requirements**

1. **CSV Upload & Column Mapping**
   * User uploads a CSV.
   * Tool previews sample rows and prompts the user to confirm which columns correspond to pointid, lat, and lon.
2. **Layer Selection Interface**
   * User can search and select from available Earth Engine layers (e.g., TerraClimate, SST, NDVI).
   * Option to specify variables or temporal date ranges.
3. **Authentication Setup (Admin Side)**
   * Administrator (tool owner) configures:
     + A Google Cloud project registered with Earth Engine.
     + Creation of a service account with appropriate roles (e.g., Earth Engine Resource Viewer/Writer, Storage roles if needed), and JSON key setup [docs.tethysplatform.org+13Google for Developers+13Google for Developers+13](https://developers.google.com/earth-engine/guides/app_engine_intro?utm_source=chatgpt.com)[docs.tethysplatform.org+10Google for Developers+10Google for Developers+10](https://developers.google.com/earth-engine/guides/service_account?utm_source=chatgpt.com)[Google for Developers+4docs.tethysplatform.org+4Stack Overflow+4](https://docs.tethysplatform.org/en/stable/tutorials/google_earth_engine/part_3/service_account.html?utm_source=chatgpt.com).
   * JSON credentials stored securely, perhaps via environment variables or secrets management.
   * Tool initializes Earth Engine using those credentials so end users aren't required to authenticate.
4. **Batch Fetch Execution**
   * Internal batching mechanism (e.g., default batch\_size=500) to respect GEE’s per-request limits and prevent memory/timeouts [Google for Developers](https://developers.google.com/earth-engine/guides/usage?utm_source=chatgpt.com).
   * Users can optionally adjust batch size.
   * Leverage similar logic as your reference snippet: use ee.ImageCollection, compositing selected layers, and reduceRegions() for efficient lookup.
5. **Progress & Feedback**
   * Real-time progress bar indicating batch progression, estimated time remaining, success/fail alerts.
6. **Output Delivery**
   * Enriched table downloadable as CSV (or other formats: Excel, JSON).
   * Option to preview the first few rows within the tool.

**3. Technical Architecture & Considerations**

* **Backend Layer**:
  + Python-based processing leveraging the earthengine-api (ee) library.
  + Authentication using ee.ServiceAccountCredentials(...) and ee.Initialize(...) with JSON key and project ID [Google for Developers+1](https://developers.google.com/earth-engine/guides/quickstart_python?utm_source=chatgpt.com)[Google for Developers+14Google for Developers+14gee-documentation.readthedocs.io+14](https://developers.google.com/earth-engine/guides/service_account?utm_source=chatgpt.com).
  + Option to also support environment variable patterns such as GOOGLE\_APPLICATION\_CREDENTIALS for flexibility [gee-documentation.readthedocs.io](https://gee-documentation.readthedocs.io/en/latest/authentication/index.html?utm_source=chatgpt.com).
  + Service account must be registered in an Earth Engine-enabled project and have required IAM roles [Stack Overflow+14Google for Developers+14gee-documentation.readthedocs.io+14](https://developers.google.com/earth-engine/guides/service_account?utm_source=chatgpt.com).
* **Rate Limits & Resource Management**:
  + Respect GEE's limits: request payload size (≈10 MB), computed result size (100 MiB), concurrent aggregations, task queue limits [docs.tethysplatform.org+4Google for Developers+4Stack Overflow+4](https://developers.google.com/earth-engine/guides/usage?utm_source=chatgpt.com).
  + Implement batching, error retries, and fallback strategies (e.g., split large jobs into smaller time/date ranges).
* **UI/Frontend Layer**:
  + Intuitive CSV uploader with auto-detection and manual override of column mapping.
  + Layer search UI integrating Earth Engine's catalog or a curated list.
  + Progress bar with batch number, rows processed, estimated time, and error logs.
  + Result preview and download.

**4. Non-Functional Requirements**

* **Usability**: Simple interface requiring minimal technical knowledge.
* **Security**: Secure handling of service account keys; no exposure to end users; audit logging of accesses.
* **Scalability**: Handle large point datasets via batching and queueing.
* **Reliability**: Graceful handling of network/timeouts; retries; clear error reporting.
* **Maintainability**: Modular codebase separating auth, data retrieval, batching, and UI.

**5. Admin & Deployment Workflow**

1. Register Cloud Project with Earth Engine and enable Earth Engine API.
2. Create service account and grant roles (e.g., Earth Engine Resource Viewer, Storage Viewer/Creator if exporting).
3. Generate and securely store JSON key.
4. Deploy tool with credentials configured (e.g., via env vars or secure vault).
5. Optionally rotate credentials and manage access logs.

**6. Milestones / Development Phases**

| **Phase** | **Description** |
| --- | --- |
| 1: Prototype | CSV import, column mapping, sample GEE retrieval using service account auth. |
| 2: Layer Selector & Batching | UI for layer search + implement batching logic and progress tracking. |
| 3: Quota Handling | Integrate GEE limit guards (batch size cap, retries, error detection). |
| 4: Output & UX | Table preview, downloads, UI polish. |
| 5: Security & Admin Setup | Secure key handling, service account setup documents, demos. |
| 6: Testing & Documentation | Unit/integration tests, admin instructions, user docs. |

**7. Risks & Mitigations**

* **Quota Exceeded / Memory Limits**  
  *Mitigation*: Use batching, limit concurrent aggregations, and provide guidance to users.
* **Authentication Errors or Key Misconfig**  
  *Mitigation*: Validate service account at setup; test initialization on deployment. Provide clear admin troubleshooting steps [Google for Developers+1](https://developers.google.com/earth-engine/guides/quickstart_python?utm_source=chatgpt.com)[gee-documentation.readthedocs.io+3Google for Developers+3Google for Developers+3](https://developers.google.com/earth-engine/guides/service_account?utm_source=chatgpt.com)[gee-documentation.readthedocs.io+2Google for Developers+2](https://gee-documentation.readthedocs.io/en/latest/authentication/service-account-auth.html?utm_source=chatgpt.com)[gee-documentation.readthedocs.io](https://gee-documentation.readthedocs.io/en/latest/authentication/index.html?utm_source=chatgpt.com)[Google for Developers](https://developers.google.com/earth-engine/guides/usage?utm_source=chatgpt.com)[GIS Stack Exchange+1](https://gis.stackexchange.com/questions/350527/authentication-issue-earth-engine-python-using-ee-serviceaccountcredentials?utm_source=chatgpt.com).
* **Exporting Data to Wrong Drive (Service Account vs User)**  
  *Mitigation*: Be aware that service account exports (e.g., to Drive) go to service account’s drive unless impersonation or proper drive access is managed [Stack Overflow](https://stackoverflow.com/questions/71834208/why-is-the-export-to-the-drive-empty-when-using-google-earth-engine-in-python-w?utm_source=chatgpt.com).

**8. References**

* **Service Account Authentication Guide**: Steps to create, register, and initialize with service account credentials in Python [GIS Stack Exchange+11Google for Developers+11gee-documentation.readthedocs.io+11](https://developers.google.com/earth-engine/guides/service_account?utm_source=chatgpt.com).
* **Earth Engine Quotas & Batch Limits**: Payload size caps, memory/computation limits, task queue limits [Google for Developers](https://developers.google.com/earth-engine/guides/usage?utm_source=chatgpt.com).
* **Common Pitfalls**: Register service account for Earth Engine; CLI vs Python auth differences; drive export behavior [Google for Developers+2docs.tethysplatform.org+2](https://developers.google.com/earth-engine/guides/auth?utm_source=chatgpt.com).