Mapping Spreadsheet Sprawl

General Assembly — Session 3 reference handout

When a folder has multiplied into drafts, extracts, old versions, dashboard exports, meeting notes, and AI summaries, the first job is not to read everything. The first job is to make a roadmap: what exists, what matters, and who relies on it.

Human version

1. Cluster by purpose
Group near-duplicates and abandoned drafts by what they were trying to support. Use filename patterns, folder location, owner clues, and modification age before opening every file.

2. Inventory the shape, not the contents
Capture filenames, locations, owners, last-modified dates, file sizes, labels, and obvious type or status. Do not turn the cleanup into a manual reading project.

3. Identify what matters
Find the handful of files that are still doing work. These should become governed datasets or governed artefacts with owners, as-of dates, review points, and clear caveats.

4. Draft the folder roadmap
Create a first index of important folders, live sources, drafts, archive material, and open questions. It does not have to be complete or perfect to be useful.

5. Let the roadmap be updated
If people disagree about sources of truth, write that down. A visible disagreement is better than a hidden filing system that everyone interprets differently.

6. Find your upstream users
People who depend on your outputs should be invited to document that dependency in the roadmap. This turns the index from a tidy filing note into operational context.

What the root index should make visible

Area	What to record
Authoritative sources	Which files, folders, systems, or reports should be treated as the source of truth.
Working material	Drafts, scratch analyses, prompt experiments, and temporary exports that should not be mistaken for governed outputs.
Archive material	Old versions kept for reference, audit, or recovery.
Known conflicts	Competing files, unclear owners, stale outputs, or places where teams disagree.
Dependencies	People, teams, dashboards, reports, scripts, or AI workflows that rely on the output.

What to put in a folder guide

A useful guide is not a complete catalogue. It is a short navigation layer that helps a new staff member, manager, analyst, or approved assistant avoid guessing.

Guide field	What to capture
Purpose	The purpose of this folder or knowledge area.
Audience and handling	Who the material is for, plus classification, sensitivity, or handling limits.
Read first	The most important files to read first.
Source status	Source-of-truth files, working files, and records.
Key spreadsheets	Important spreadsheets and what they are used for.
Recurring outputs	Reports, dashboards, meeting packs, or other repeated products.
Definitions	Terms that matter, plus known stale, retired, duplicate, or disputed material.
Contacts	Owner, steward, approver, and support contact.
Review and retention	Review date and retention or disposal expectation.
AI-use notes	What an approved assistant may summarise, what it must not use, and what requires human or source-owner confirmation.
Related sources	Other folders or data sources to look at.
Uncertainty comments	Notes that say "I am not sure, but I think ..." instead of hiding uncertainty.

Why this matters

The folder guide is a lightweight control surface. It tells people what to trust, what to check, what to leave alone, and when an AI-generated summary needs human confirmation.

Prompt to give a bot

Use this as the instruction block for an assistant that is allowed to inspect the folder or repository and update a root-level index file.

```
You are updating the root index for this repository or shared folder.  Create or revise
the existing root-level index or instruction file, following the local convention if one
already exists:  AGENTS.md, CLAUDE.md, FILE_INDEX.md, INDEX.md, or README.md.

Do not move, delete, rename, or rewrite source files.  Work from paths, filenames,
headings, existing index files, metadata, recent modification dates, and light sampling
of representative files.  Do not expose sensitive contents; summarize structure and purpose
instead.

Please do the following:
1.  Read the existing root index, README, AGENTS.md, CLAUDE.md, and any nearby instruction
files.  Preserve mandatory instructions exactly.
2.  Map the tree by purpose, not only by path.  Group near-duplicates, drafts, exports,
generated outputs, source documents, scripts, and archive material.
3.  Inventory the shape of important areas:  folder purpose, likely owner or team if
visible, freshness signals, generated-vs-source status, sensitivity or handling clues,
and obvious dependencies.
4.  Identify the files and folders that appear to be sources of truth.  If there are
competing sources, stale copies, or unclear owners, document that as an open question
rather than guessing.
5.  Add or update a concise root index with these sections:  Purpose; How to Navigate;
Authoritative Sources; Working Material; Generated Outputs; Archives; Dependencies;
Maintenance Notes; Open Questions.
6.  For each important folder or knowledge area, include the folder-guide fields that
can be inferred safely:  purpose; audience and handling level; files to read first;
source-of-truth files; working files versus records; key spreadsheets and their use;
recurring reports, dashboards, and meeting packs; important definitions; stale, retired,
duplicate, or disputed material; owner, steward, approver, and support contact; review
date; retention or disposal expectation; AI-use notes; related folders or data sources; and
uncertainty comments.
7.  Keep the result useful for the next human and the next assistant.  Use concrete paths.
Say what each area is for.  Mark uncertain entries as uncertain.  Keep operational warnings
visible.
8.  Finish by listing what changed in the index and any follow-up questions that require a
human decision.
```

Good output standard

The finished index should help someone answer three questions quickly: **Where should I start? What can I trust? Who or what depends on this?**