Professional Work Tasks

Dataset specification

GR-LLM-04 · LLM · Early access · updated 2026-11-04

1 - Overview

Expert-graded finance, legal and office tasks with real deliverables.

Open-ended professional tasks that end in real work products - spreadsheets, memos and slide decks. Each task bundles the source documents, a reference deliverable and a detailed rubric, so models can be trained and scored on the work analysts, lawyers and office professionals actually do.

- **Domains:** Finance · Law · Office
- **Deliverables:** XLSX · DOCX · PPTX
- **Rubric items per task:** 27–65
- **Languages:** EN · ZH
- **Modalities:** Documents, Spreadsheets, Slides, Rubrics

2 - Use cases

- Reinforcement learning on long-horizon office work
- Evaluation of agents that produce documents, spreadsheets and decks
- Rubric and judge-model development

3 - Task families

- **Finance:** Capital, margin and ratio analyses from public filings; spreadsheet model plus written memo (English)
- **Law:** Simulated case files based on landmark matters; memo, control matrix and board briefing (Chinese)
- **Office:** Training and budget decks built from public source documents (Chinese)
- **Occupational:** Multi-file tasks by sector and occupation, such as compliance and healthcare management (English)

4 - Each task includes

- **Brief:** Task instructions, constraints and acceptance criteria
- **Inputs:** Source documents: PDF, DOCX, XLSX, CSV and HTML
- **Reference deliverable:** Gold output in the requested formats
- **Rubric:** 27–65 criteria per task, with points, numeric tolerances and partial-credit rules
- **Metadata:** Domain, difficulty tier and targeted skills

5 - Provenance

- **Inputs:** Public filings and documents, or synthetic material; no personal health data
- **Model trajectories:** Not included
- **Volume:** Confirmed per order

6 - Data schema

- `instruction.md` **(markdown):** Task brief given to the model.
- `task.toml` **(toml):** Domain, difficulty and skill tags.
- `input_files/` **(directory):** Source documents for the task.
- `golden_output/` **(directory):** Reference deliverables.
- `rubric` **(json / toml):** Criteria with points, tolerances and grading notes.

7 - Directory layout

```
professional-work-tasks/      # indicative; final layout confirmed per delivery
`-- tasks/<domain>/<task_id>/
    |-- instruction.md
    |-- task.toml
    |-- input_files/
    |-- golden_output/
    `-- tests/
        `-- rubric.toml
```

8 - Evaluation sample

- **Availability:** On request
- **Size:** On request
- **Formats:** XLSX, DOCX, PPTX, PDF, JSON
- **Contents:** Tasks from each domain with inputs, reference deliverables and rubrics

9 - License and contact

- **Graviti Commercial Data License:** Commercial training and evaluation rights. Volume and exclusivity agreed per order.
- **Dataset page:** https://www.graviti.com/datasets/professional-work-tasks/
- **Sales:** sales@graviti.com