Biosyn■KB: End■to■End Setup & Usage Guide

This document describes how to run the multi■agent biosynthesis knowledge■base pipeline and the web chat UI on your machine.

Prerequisites

- macOS with Homebrew (or Linux with Python 3.11)
- SerpAPI API key (or Brave key) in apis.yaml
- Ollama installed locally with a model (e.g., gpt-oss:20b)

Project Layout

All files are under the biosyn-kb/ folder. Artifacts (DB/JSONL), pages, and work directories are stored there.

Install & Environment

Create Python 3.11 environment, install project and CrewAI (this installs web UI deps too):

```
cd biosyn-kb make crewai
```

Configuration

Set search keys and LLM defaults: copy examples and fill values.

```
cp apis.yaml.example apis.yaml   # set SERPAPI_API_KEY or BRAVE_API_KEY cp llm.yaml.exa
mple llm.yaml    # optional: set provider/model/base_url/temp/top_p/seed
```

Web Chat UI (preferred)

Launch the web interface and open http://localhost:8000/chat:

```
.venv311/bin/biosyn-kb-web
```

In the chat UI: enter a query (e.g., 'biosynthesis of artemisinin'), choose provider/model, submit. A primer appears first, then structured evidence populates as background jobs finish. View job history at /status.

One■shot Agent (CLI)

Run the full pipeline from a conversational query with minimal flags:

```
.venv311/bin/biosyn-kb agent-query   --query "biosynthesis of artemisinin"   --provider
 serpapi --max-results 3 --exclude-domain wikipedia.org   --api-config apis.yaml   --sa
ve-html-dir pages/artemisinin --work-dir work/artemisinin   --db artifacts/artemisinin.
db --out artifacts/artemisinin_summaries.jsonl   --limit 2
```

Outputs: pages saved under pages/artemisinin/, JSONL summaries in artifacts/, and SQLite DB with pages and metrics.

Q&A; (RAG over stored pages)

```
.venv311/bin/biosyn-kb qa   --db artifacts/artemisinin.db   --question "What organism a
nd enzymes are key in artemisinin biosynthesis and what titers are reported?"   --k 5 -
-model "gpt-oss:20b"
```

Queue & Worker (scaling)

Enqueue an agent job and watch the worker process it:

```
.venv311/bin/biosyn-kb queue-add-agent   --queue-db artifacts/queue.db   --query "biosy
nthesis of isobutanol"   --provider serpapi --api-config apis.yaml   --max-results 3 --
exclude-domain wikipedia.org   --save-html-dir pages/agent
.venv311/bin/biosyn-kb queue-worker-watch --queue-db artifacts/queue.db --sleep 2
```

Run Unit Tests

```
make test
```

Troubleshooting

- Missing keys: set SERPAPI_API_KEY (or BRAVE_API_KEY) in apis.yaml

- Slow first QA: ONNX MiniLM downloads on first build; retry QA

- LLM slow on long pages: jobs run in background; use queue-worker-watch

- Forms error: install python-multipart (make crewai already installs it)