

Lab 3: FastAPI

CSI403 - Full Stack Development

Faculty of Information Technology

Sripatum University

Weight: 3.75%

Objectives

- ✓ Study api.py from Generic-RAG
- ✓ Run FastAPI server
- ✓ Test API endpoints

Repository: <https://github.com/amornpan/Generic-RAG>

Task 1: Study api.py

Review the structure of `api.py`:

- **FastAPI app** - Application instance
- **Pydantic models** - Request/Response schemas
- **Endpoints:**
 - `/health` - Health check
 - `/search` - Search documents
 - `/query` - RAG query

Task 2: Run Server

```
cd Generic-RAG  
conda activate rag_env  
python api.py
```

Server runs on <http://localhost:9000>

Task 3: Test Endpoints

Health check:

```
curl http://localhost:9000/health
```

Search:

```
curl -X POST http://localhost:9000/search \  
-H "Content-Type: application/json" \  
-d '{"query": "test", "top_k": 5}'
```

Task 4: Access Swagger Docs

Open in browser:

<http://localhost:9000/docs>

- Interactive API documentation
- Test endpoints directly in browser
- View request/response schemas

Task 5: Add New Endpoint

Add /documents endpoint to list all documents:

```
@app.get("/documents")
async def list_documents():
    # Query OpenSearch for all documents
    response = client.search(
        index="documents",
        body={"query": {"match_all": {}}}
    )
    return {"documents": response["hits"]["hits"]}
```

Deliverables

Item	Check
Server running	<input type="checkbox"/>
Endpoints tested	<input type="checkbox"/>
New endpoint added	<input type="checkbox"/>
Screenshots	<input type="checkbox"/>

Deadline: Sunday 23:59