Lab 1: Git + Python Fundamentals

CSI403 - Full Stack Development

Faculty of Information Technology

Sripatum University

Weight: 3.75%

Objectives

- ✓ Fork and clone Generic-RAG repository
- ✓ Practice basic Git workflow (clone, add, commit, push)
- ✓ Review Python fundamentals (OOP, Type Hints)

Repository: `https://github.com/amornpan/Generic-RAG`

Task 1: Fork Repository

1. Go to https://github.com/amornpan/Generic-RAG
2. Click "**Fork**" button (top right)
3. Select your account
4. You now have your own copy!

Result: https://github.com/YOUR_USERNAME/Generic-RAG

Task 2: Clone to Local

```
git clone https://github.com/YOUR_USERNAME/Generic-RAG.git
cd Generic-RAG
```

Verify remote:

```
git remote -v
# origin https://github.com/YOUR_USERNAME/Generic-RAG.git
```

Task 3: Setup Environment

```
# Create conda environment
conda create -n rag_env python=3.10 -y

# Activate environment
conda activate rag_env

# Install dependencies
pip install -r requirements.txt
```

Task 4: Study Python OOP Files

Review files in `src/week02-git-python/`:

File	Content
`document.py`	Document class with OOP
`rag_config.py`	Dataclasses for configuration
`utils.py`	Utility functions
`main.py`	Entry point demo

```
cd src/week02-git-python
python main.py
```

Task 5: Create Your Document Class (1/2)

Create `src/week02-git-python/lab01_example.py`:

```
"""
Lab 01: My Document Class
Student Name: YOUR_NAME
Student ID: YOUR_ID
"""
from typing import List, Dict, Optional
from datetime import datetime

class Document:
    def __init__(self, title: str, content: str,
                 source: Optional[str] = None):
        self.title = title
        self.content = content
        self.source = source
        self.created_at = datetime.now()
        self.word_count = len(content.split())
```

Task 5: Create Your Document Class (2/2)

Add methods to Document class:

```
    def get_summary(self, max_length: int = 100) -> str:
        if len(self.content) <= max_length:
            return self.content
        return self.content[:max_length] + "..."

    def get_word_count(self) -> int:
        return self.word_count

    def to_dict(self) -> Dict:
        return {
            "title": self.title,
            "content": self.content,
            "source": self.source,
            "word_count": self.word_count
        }
```

Task 5: Add Helper Functions

```
def process_documents(documents: List[Document]) -> Dict:
    total_documents = len(documents)
    total_words = sum(doc.word_count for doc in documents)
    average_words = total_words / total_documents if total_documents > 0 else 0
    return {
        "total_documents": total_documents,
        "total_words": total_words,
        "average_words": average_words,
        "titles": [doc.title for doc in documents]
    }

def search_documents(documents: List[Document],
                     query: str) -> List[Document]:
    query_lower = query.lower()
    return [doc for doc in documents
            if query_lower in doc.content.lower()]
```

Task 5: Test Your Code

```
cd src/week02-git-python
python lab01_example.py
```

Expected output:

```
================================================
Lab 01: Document Class Demo
================================================
Documents:
   Document(title='RAG Introduction', words=6)
   Document(title='OpenSearch Guide', words=7)
Search 'RAG':
   - RAG Introduction
   - OpenSearch Guide
Statistics:
   Total: 2 docs, Words: 13, Average: 6.5
```

Task 6: Commit and Push to Master

```
# Go to repository root
cd ../..

# Check status
git status

# Add your file
git add src/week02-git-python/lab01_example.py

# Commit
git commit -m "Lab␣01:␣Add␣Document␣class␣example"

# Push to master
git push origin master
```

Task 7: Verify on GitHub

1. Go to your GitHub repository
2. Navigate to `src/week02-git-python/`
3. Check that `lab01_example.py` is there
4. **Take screenshot** of the file on GitHub

Important: Screenshot must show your GitHub username and the file content visible.

Grading Criteria

Task	Points
Fork & Clone	20
Environment Setup	10
Document class with methods	30
process_documents function	20
search_documents function	20
Total	**100**

Deliverables

Item	Check
Forked repository	☐
Cloned to local machine	☐
Environment setup completed	☐
`lab01_example.py` created & tested	☐
Pushed to master branch	☐
Screenshot of file on GitHub	☐

Deadline: Sunday 23:59