**1. AutoSDT-Search**
Crawl relevant GitHub repositories

- Seed keywords
- Expand keywords
- Query GitHub and Papers with Code
- Filter research repos
- Crawled repos
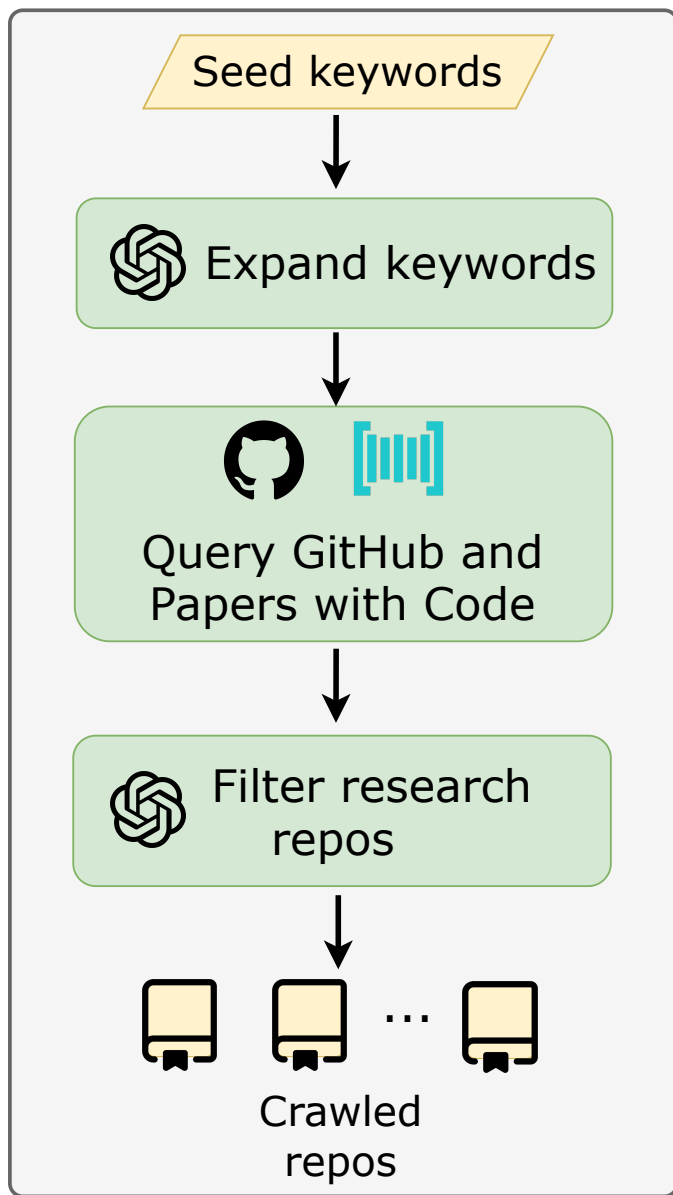
**2. AutoSDT-Select**
Select candidate programs

- Crawl Python files
- Filter data-driven scientific code
- Extract dependency files
- Programs and dependency folders
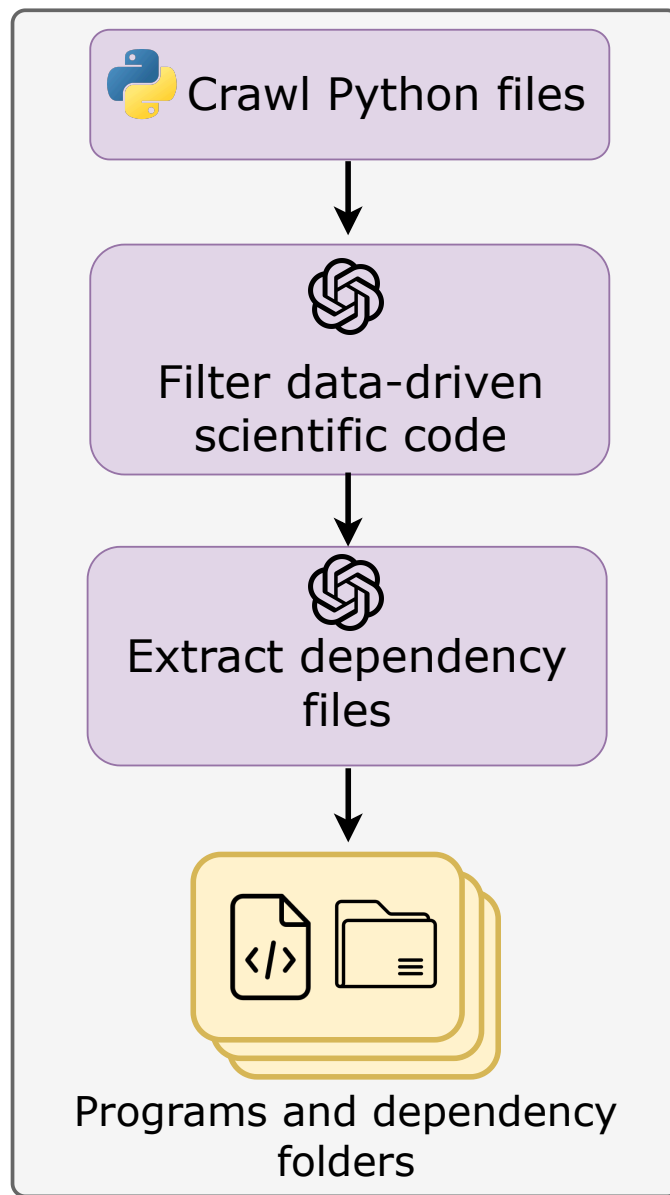
**3. AutoSDT-Adapt**
Adapt programs into finalized tasks

- Program
- Dependency folder
- Adapt program for standalone executability
- Successful execution
- Error
- Generate task instruction

**AutoSDT-5K**

- Task instruction
- Code solution