JabRef CLI

Software Composition Seminar Project
The Goal
Implement a CLI for JabRef
What is JabRef?

- Reference manager for BibTeX libraries
- Full-text downloader for bib entries
- Open source, on GitHub
- Written in Java
What is JabRef?
Looking Up Full Text Documents
@InProceedings{Meri17a,
    author   = {L. Merino and M. Ghafari and C. Anslow and O. Nierstrasz},
    title    = {CityVR: Gameful Software Visualization},
    booktitle = {2017 IEEE International Conference on Software Maintenance and Evolution (ICSME)},
    year      = {2017},
    month     = {Sept},
    pages     = {633-637},
    doi       = {10.1109/ICSME.2017.70},
    file      = {Meri17a-CityVR__Gameful_Software_Visualization.pdf},
    keywords  = {Games;Software;Software engineering;Tools;Urban areas;Visualization;3D visualization},
}

@Comment{jabref-meta: databaseType:biblatex;}
Why do we want a CLI?

- GUI is not intuitive and has screen scaling issues
- CLI is scriptable
- CLI can be much faster to work with
Our Approach

1. Fork project on GitHub
2. Define a command line interface
3. Implement command line parser
4. Wire up CLI to existing logic
   - Load library
   - Find BibTeX entry
   - Find PDF source
   - Download PDF
   - Format entry, add file reference
   - Save changes to BibTeX database
Our CLI

Database Handler

(1) load DB

(2) search for entry

Query Engine

Full-text crawler / downloader

(3) find PDF

(4) download PDF

jabref-cli -key Meril7c -c get-pdf
How it actually was...

Our CLI

Database Handler

Query Engine

FindFullTextAction

DownloadExternalFile

`jabref-cli -key Meril7c -c get-pdf`
Our CLI

Database Handler

Query Engine

FindFullTextAction

FullTextFetchers

DownloadExternalFile

How it actually was...

$\texttt{jabref-cli -key Meri17c -c get-pdf}$
How it actually was...

```
jabref-cli -key Meri17c -c get-pdf
```

- Database Handler
- Query Engine
- FindFullTextAction
  - FullTextFetchers
    - IEEE
    - Google Scholar
    - ArXiv
    - Springer
- DownloadExternalFile
How it actually was...

```
jabref-cli -key Meri17c -c get-pdf
```

- Our CLI
  - Database Handler
  - Query Engine
  - FindFullTextAction
    - FullTextFetchers
      - IEEE
      - Google Scholar
      - ArXiv
      - Springer
    - ?
      - GUI
  - DownloadExternalFile
How it actually was...

```
jabref-cli -key Meri17c -c get-pdf
```

**Our CLI**

**Database Handler**

**Query Engine**

**FindFullTextAction**

**FullTextFetchers**

- IEEE
- Google Scholar
- ArXiv
- Springer

**GUI**

**DownloadExternalFile**

???
How it actually was...

```
jabref-cli -key Meri17c -c get-pdf
```
How it actually was...

```
jabref-cli -key Meri17c -c get-pdf
```
How it actually was...

```
jabref-cli -key Meri17c -c get-pdf
```

Database Handler

Our CLI

Query Engine

FindFullTextAction

FullTextFetchers

EntryFetchers

IEEE

Google Scholar

ArXiv

Springer

DownloadExternalFile

GUI

SidePane

???
How it actually was...

```
jabref-cli -key Meri17c -c get-pdf
```

**Our CLI**

- Database Handler
- Query Engine
- EntryFetchers
- FindFullTextAction
- FullTextFetchers
- DownloadExternalFile
- GUI
- SidePane
But we did it!
Current Features

- Find a paper by key in the BibTex library and download the PDF
- Find and download PDFs by URL (ArXiv, ACM, IEEExplore, doi.org, Springer) and add entry
- Insert $file = \{\ldots\}$ in BibTeX file
Lessons learned

- Spent way more time reading code than writing
- Well-maintained doesn’t mean well-engineered
- Accidental complexity?
- “Hacking” does the job, but the solution will hardly be sustainable
Don’t judge a book by its cover...
...or a project by its package structure
Future work

- Batch mode (run it on entire SCG bibliography)
- Handle multiple search results
- Refactoring
Applications

- SCG bibliography
- CiteWise
- Integrated Literature Research Environment
- Have a single command during daily work: get entry, get PDF, format entry, add file link