Thomas Jäger thomas.ralf.jaeger@univie.ac.at

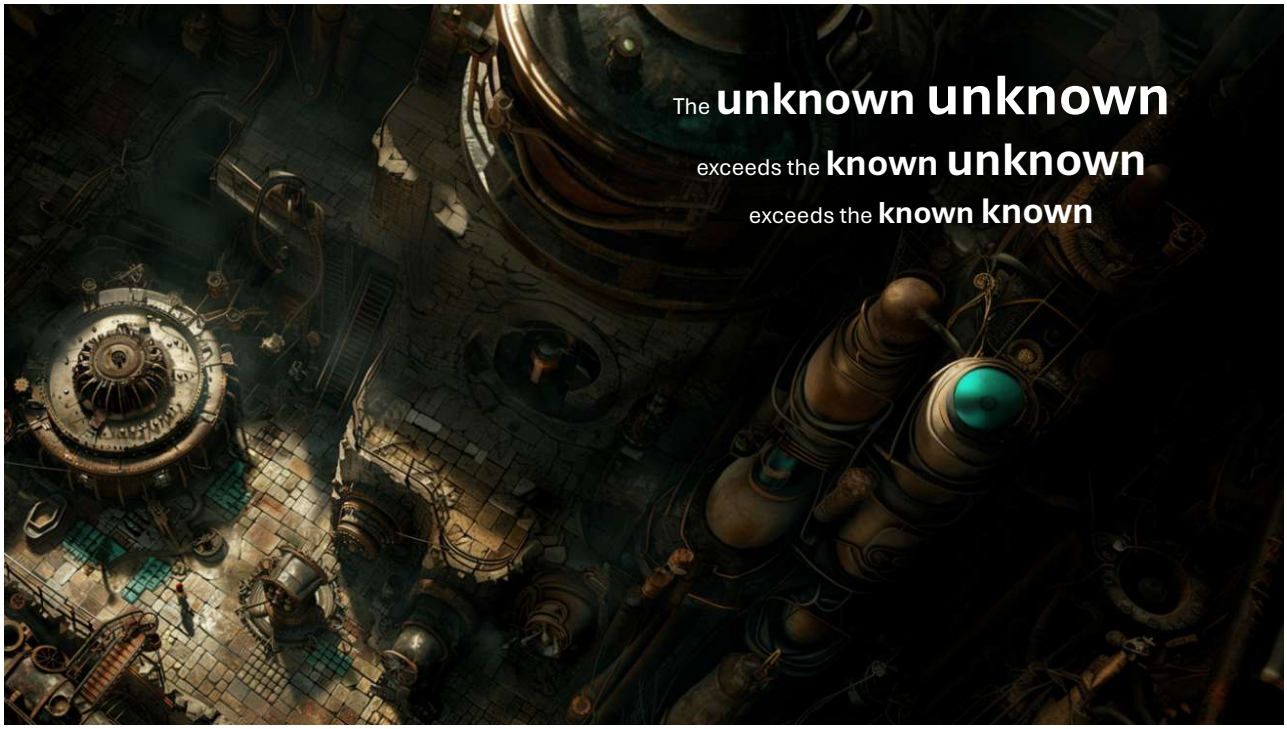
Intermediate

Classic DH-Problems
Short Pieces in Python

For Beginners

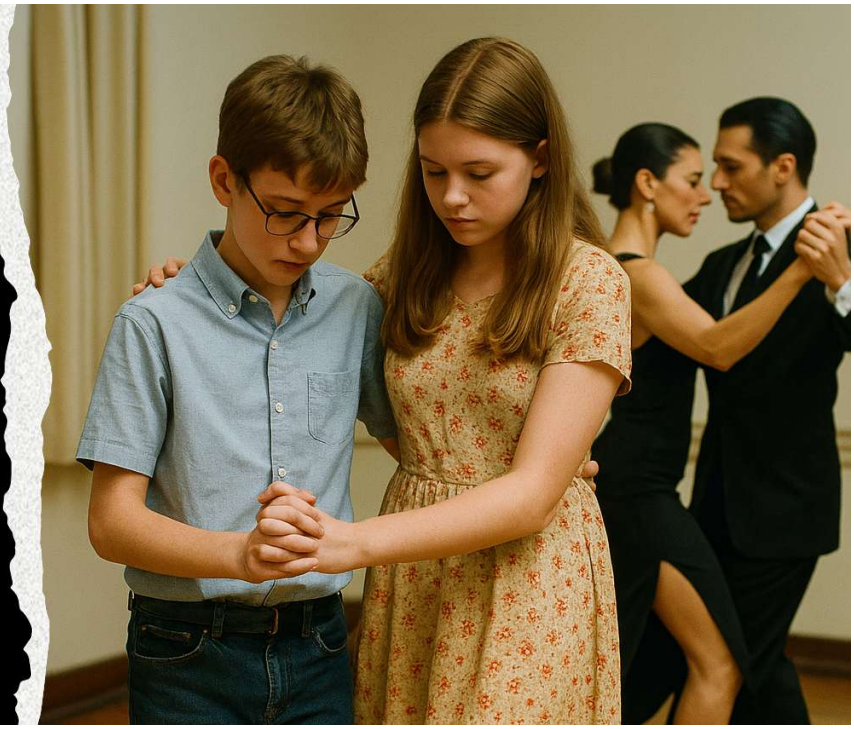
Learn Python like a Native

The imposter syndrom

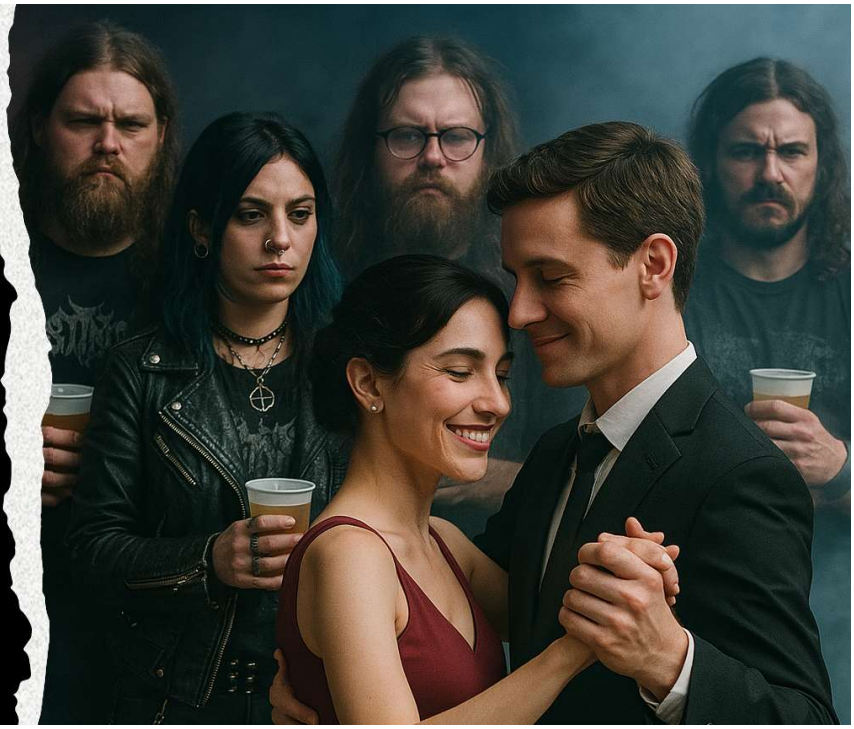


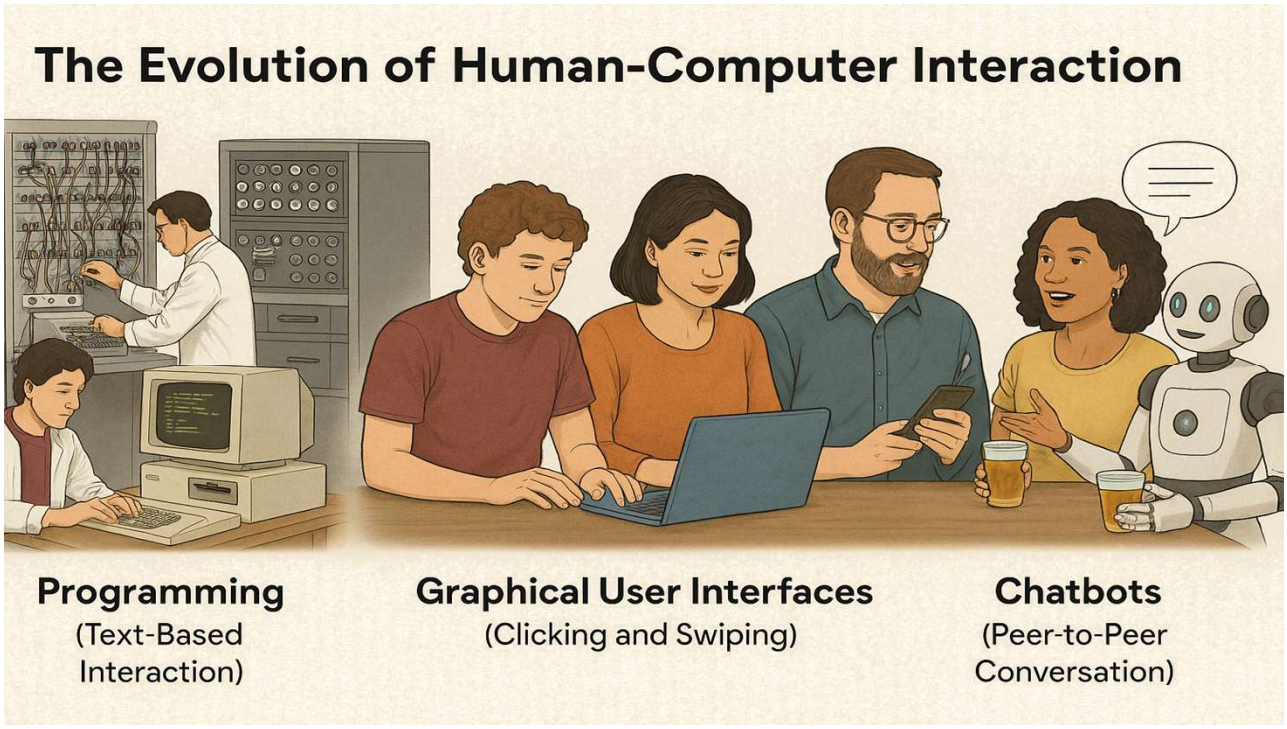
Coding is like dancing?

- 🌀 **At first, you count every step**
(Just like learning dance moves)
- 🤯 **It feels clunky and complicated**
- 🔁 **Practice builds confidence**
(Repetition leads to rhythm)
- 🧠 **Eventually, you feel the flow**
(Syntax and logic become natural)
- 🎵 **Then it starts to feel fun and effortless**
- 💃 🕺 **Coding becomes a form of creative expression**



But – what, if you hate dancing?




The Evolution of Human-Computer Interaction
Programming
(Text-Based Interaction)
Graphical User Interfaces
(Clicking and Swiping)
Chatbots
(Peer-to-Peer Conversation)


Reading Your Way Into a New Language!

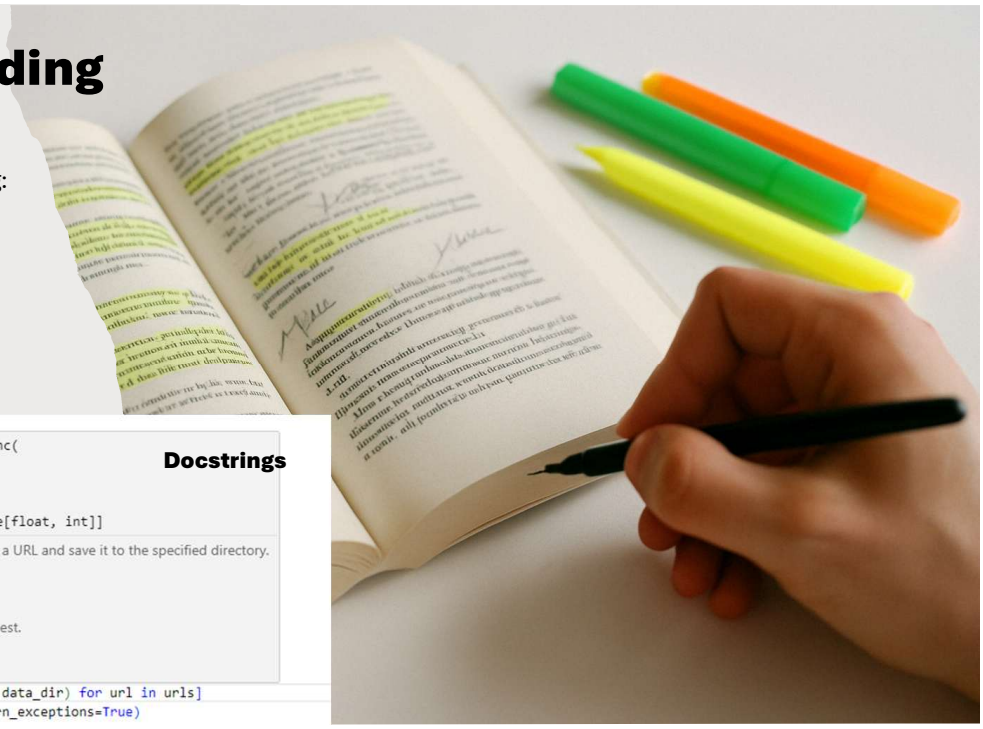
Reading Boosts Language Learning

- Expands vocabulary in authentic context
- Reinforces structural and syntactic knowledge naturally
- Improves comprehension and fluency over time
- Encourages intuitive language acquisition through exposure

Especially effective when…

- Texts are *didactically prepared* for language learners
- A *tutor provides guidance* and explains challenging parts
- Reading is part of a *collaborative group activity* with discussion

Active Reading



Three-Pass Code Reading:

1. exploration phase
2. consolidation phase
3. documentation phase

Docstrings

```
(function) def download_image_async(
    session: Any,
    url: Any,
    data_dir: Any
) -> CoroutineType[Any, Any, tuple[float, int]]
```

Asynchronously download an image from a URL and save it to the specified directory.

Args

session : *aiohttp.ClientSession*
The aiohttp session to use for the request.

url : *str*
The URL of the image to download.

```
= [download_image_async(session, url, data_dir) for url in urls]
ts = await asyncio.gather(*tasks, return_exceptions=True)
```

Programming:
- Concurrent Threads
- Crawler
- Word2Vec

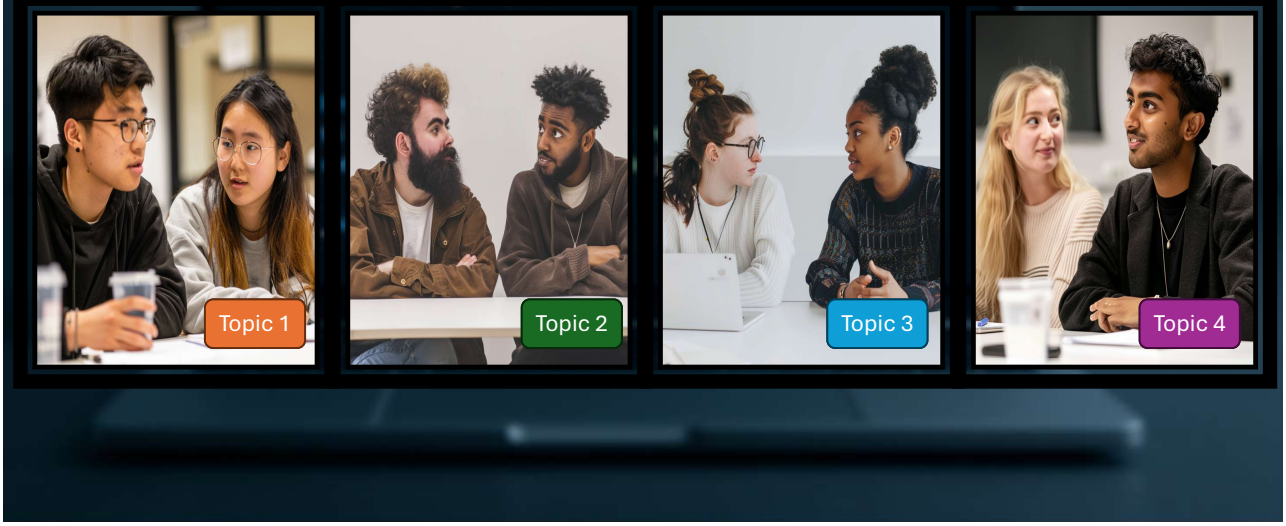
Developing:
- Deal with unknown Code
- Documenting learnings
- Evaluating
- Optimizing
- Refactoring

Professional Knowledge:
- Pair Programming

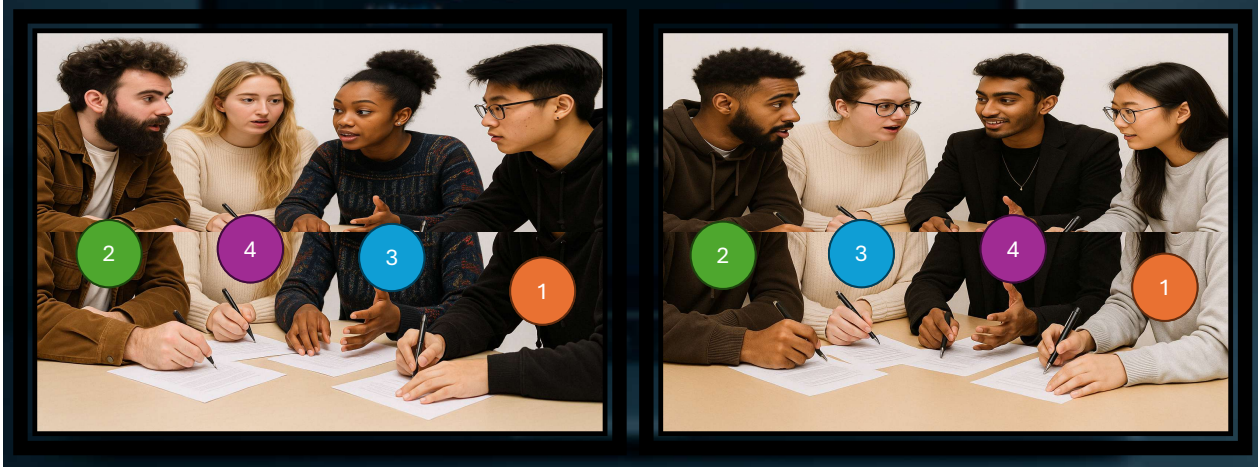
Naming Conventions

- Good names replace lengthy comments
- Make code more readable
- Consistent patterns prevent confusion

A) Working in Pairs
Each pair of students develops and masters a specific topic together.



B) Expert Groups
For a more complex topic, one “expert” from each pair joins a larger group to share their specialized knowledge.




Topic 1
Topic 2
Topic 3
Topic 4
2
4
3
1
2
3
4
1


Like Father like daugther
Family-Likeness

Judging a Word by the company it keeps or King:Queen = Man:Woman.

Word2Vector

Men don‘t cook. They ROAST!

Similiar words:
bake_Verb

British National Corpus

1. cook VERB 0.66
2. preheat VERB 0.64

English Wikipedia

1. cook VERB 0.65
2. roast VERB 0.60
3. fry VERB 0.55
4. bread VERB 0.54

Men are rogues. Women are childish.

man_NOUN → woman_NOUN

Word frequency

☑ High ☑ Medium ☐ Low

British National Corpus

1. childish ADJ 0.44
2. insincere ADJ 0.43
3. vacuous ADJ 0.42
4. patronizing ADJ 0.42
5. spiteful ADJ 0.41

mischievous_ADJ → ???

1. playful ADJ 0.71
2. cheerful ADJ 0.66
3. clumsy ADJ 0.64
4. cute ADJ 0.63
5. childlike ADJ 0.61
6. silly ADJ 0.61
7. arrogant ADJ 0.61
8. amiable ADJ 0.61
9. evil ADJ 0.60
10. obnoxious ADJ 0.60

woman:man = pleasant:creepy??

woman_NOUN → man_NOUN

Word frequency

☑ High ☑ Medium ☐ Low

British National Corpus

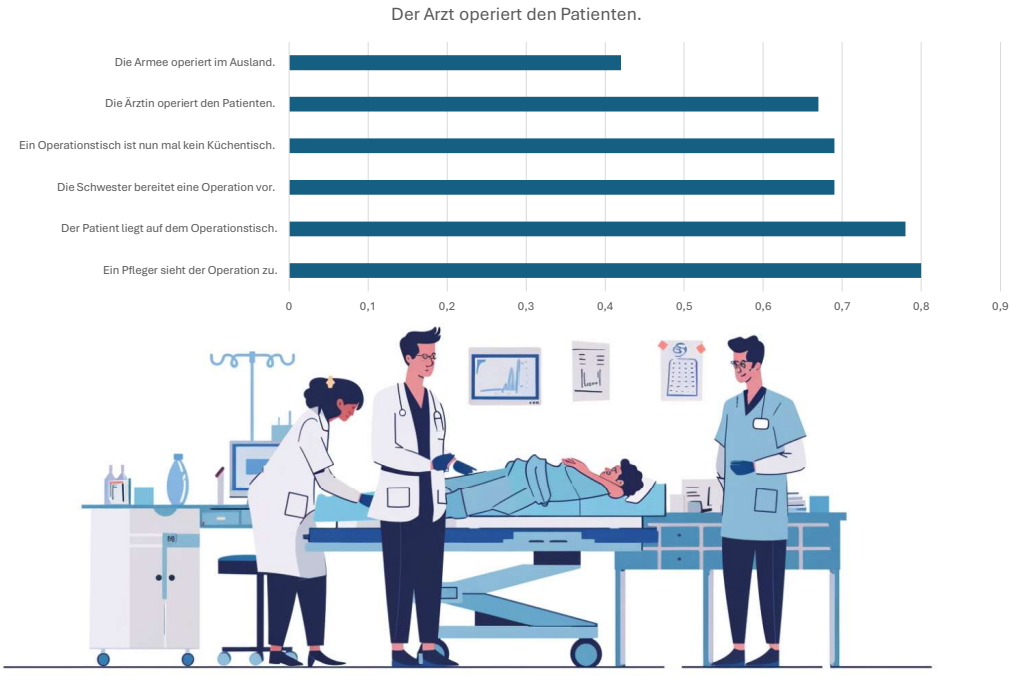
1. agreeable ADJ 0.51
2. lovely ADJ 0.51
3. quiet ADJ 0.51
4. delightful ADJ 0.51
5. creepy ADJ 0.50

pleasant_ADJ → ???

British National Corpus

1. delightful ADJ 0.63
2. agreeable ADJ 0.62
3. lovely ADJ 0.61
4. quiet ADJ 0.58
5. charming ADJ 0.58
6. beautiful ADJ 0.57
7. pleasing ADJ 0.57
8. attractive ADJ 0.57
9. enjoyable ADJ 0.56

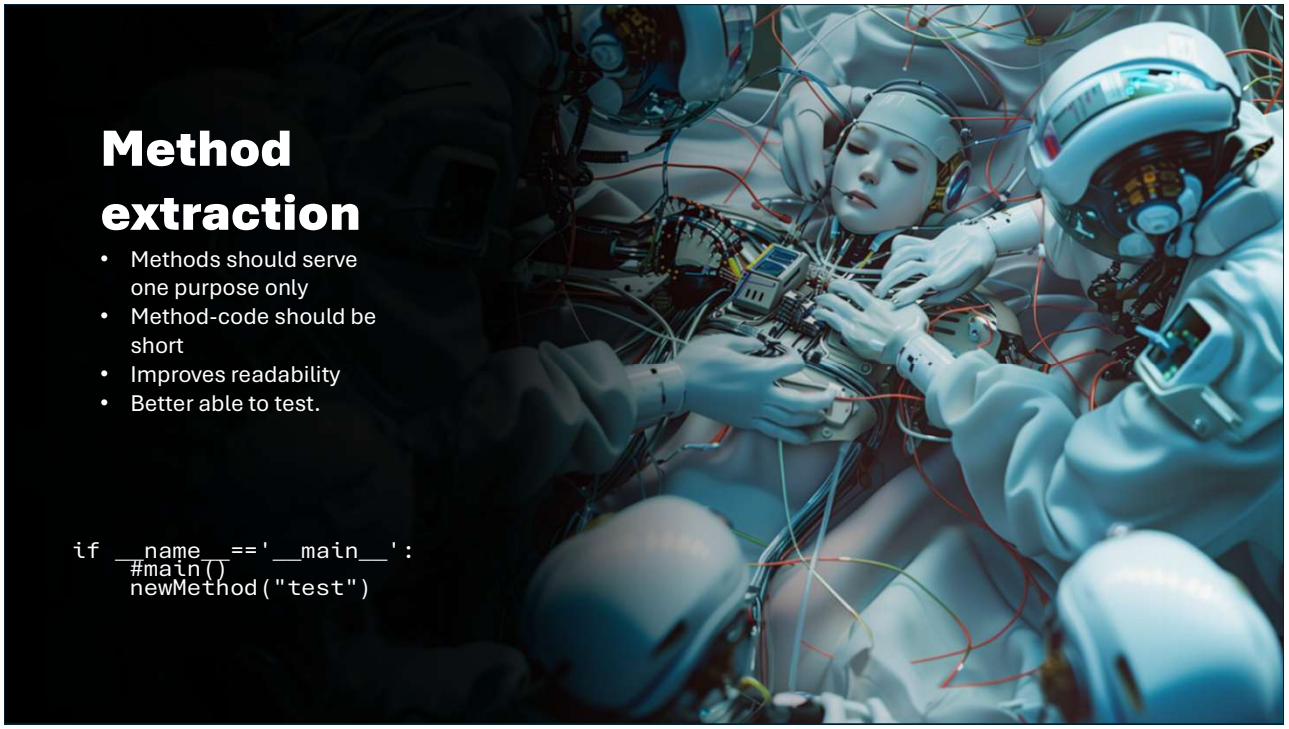
A sentence with a female is not a sentence with a male.



Method extraction

- Methods should serve one purpose only
- Method-code should be short
- Improves readability
- Better able to test.

```
if __name__=='__main__':
    #main()
    newMethod("test")
```




Training
WORMS


Worming

Crawling the internet

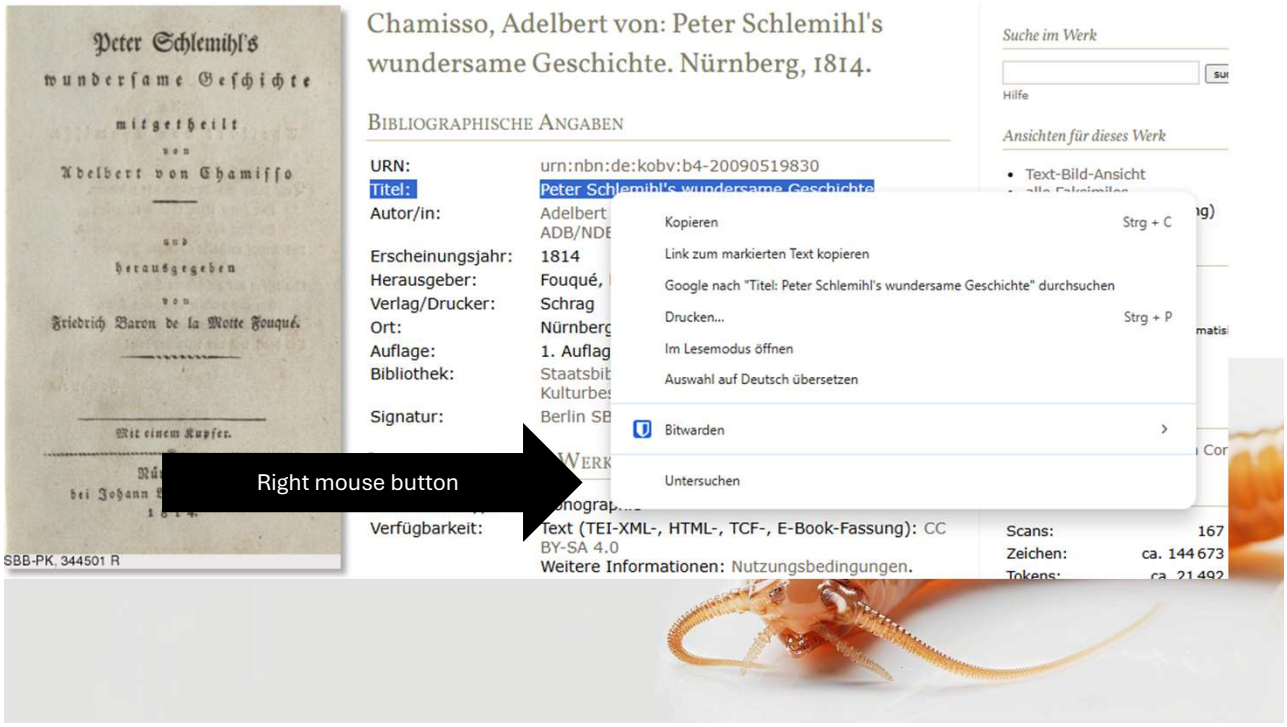
Extracting content via automated programs that

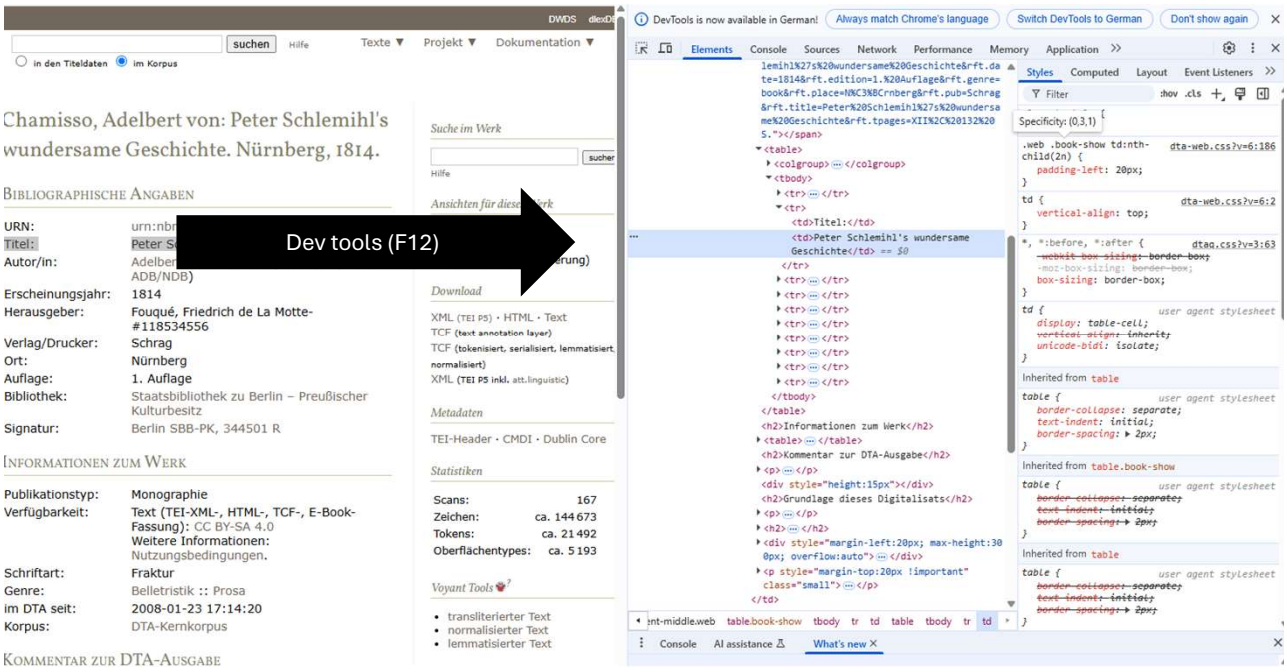
- systematically browse the web
- by following hyperlinks to
- discover
- retrieve,
- and index content

for data analysis.

Crawling the internet

1. Identify pages with content you want to retrieve
2. Analyse structure
3. Find patterns
4. Write programs using xpath or regex to exploit the patterns


Peter Schlemihl's
wundersame Geschichte
mitgetheilt
von
Adelbert von Chamisso
und
herausgegeben
von
Friedrich Baron de la Motte Fouqué.
Mit einem Kupfer.
SBB-PK, 344501 R
Chamisso, Adelbert von: Peter Schlemihl's wundersame Geschichte. Nürnberg, 1814.
Bibliographische Angaben
URN: urn:nbn:de:kobv:b4-20090519830
Titel: Peter Schlemihl's wundersame Geschichte
Autor/in:
Erscheinungsjahr: 1814
Herausgeber: Fouqué,
Verlag/Drucker: Schrag
Ort: Nürnberg
Auflage: 1. Auflag
Bibliothek:
Signatur: Berlin SB
Verfügbarkeit: Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung): CC BY-SA 4.0
Weitere Informationen: Nutzungsbedingungen.
Kopieren Strg + C
Link zum markierten Text kopieren
Google nach "Titel: Peter Schlemihl's wundersame Geschichte" durchsuchen
Drucken... Strg + P
Im Lesemodus öffnen
Auswahl auf Deutsch übersetzen
Bitwarden
Untersuchen
Right mouse button
Suche im Werk
Hilfe
Ansichten für dieses Werk
Text-Bild-Ansicht
Scans: 167
Zeichen: ca. 144 673


DWDS
suchen
Hilfe
Texte
Projekt
Dokumentation
in den Titeldaten
im Korpus
Chamisso, Adelbert von: Peter Schlemihl's wundersame Geschichte. Nürnberg, 1814.
Bibliographische Angaben
URN:
Titel:
Autor/in:
Erscheinungsjahr: 1814
Herausgeber: Fouqué, Friedrich de La Motte- #118534556
Verlag/Drucker: Schrag
Ort: Nürnberg
Auflage: 1. Auflage
Bibliothek: Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
Signatur: Berlin SBB-PK, 344501 R
Informationen zum Werk
Publikationstyp: Monographie
Verfügbarkeit: Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung): CC BY-SA 4.0
Weitere Informationen: Nutzungsbedingungen.
Schriftart: Fraktur
Genre: Belletristik :: Prosa
im DTA seit: 2008-01-23 17:14:20
Korpus: DTA-Kernkorpus
Kommentar zur DTA-Ausgabe
Dev tools (F12)
Suche im Werk
Ansichten für dieses Werk
Download
XML (TEI P5) · HTML · Text
TCF (text annotation layer)
TCF (tokenisiert, serialisiert, lemmatisiert, normalisiert)
XML (TEI P5 inkl. att.linguistic)
Metadaten
TEI-Header · CMDI · Dublin Core
Statistiken
Scans: 167
Zeichen: ca. 144 673
Tokens: ca. 21 492
Oberflächentypes: ca. 5 193
Voyant Tools
transliterierter Text
normalisierter Text
lemmatisierter Text
DevTools is now available in German!
Always match Chrome's language
Switch DevTools to German
Don't show again
Elements
Console
Sources
Network
Performance
Memory
Application
Styles
Computed
Layout
Event Listeners
<td>Titel:</td>
<td>Peter Schlemihl's wundersame Geschichte</td> == $0
Inherited from table
Inherited from table.book-show
What's new

Pattern: with <a> inside, href contains url

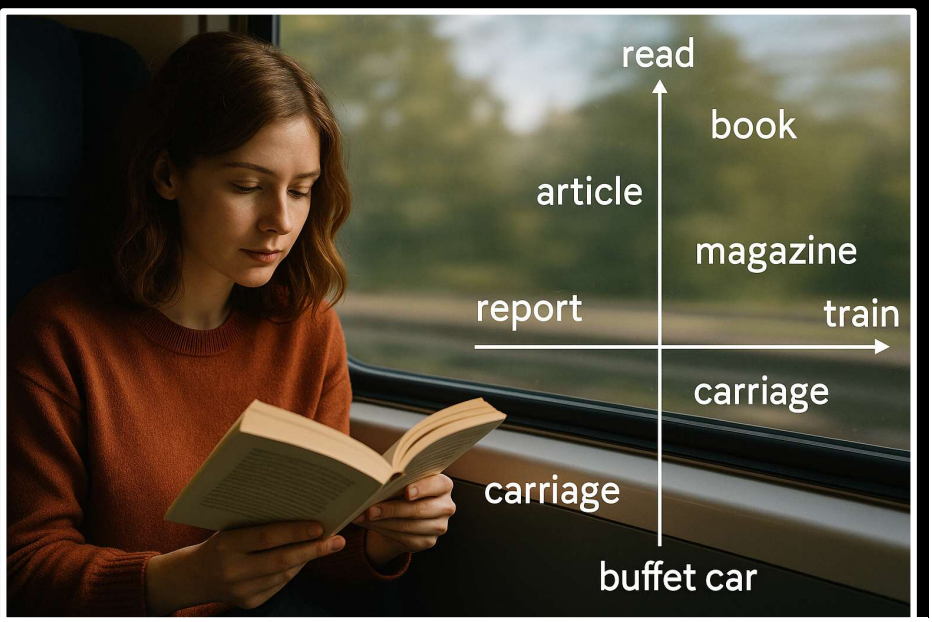
```
<div style="margin-bottom:10px; text-indent:-1.5em; padding-left:1.5em">
  <a href="https://www.deutschestextarchiv.de/book/show/buechner_werke_1879">…</a> == $0
</div>
<div style="margin-bottom:10px; text-indent:-1.5em; padding-left:1.5em">…</div>
<div style="margin-bottom:10px; text-indent:-1.5em; padding-left:1.5em">…</div>
```

Pattern: The url to download the text file

```
<a href="https://www.deutschestextarchiv.de/book/download_txt/buechner_werke_1879">Text</a> == $0
<br>
```

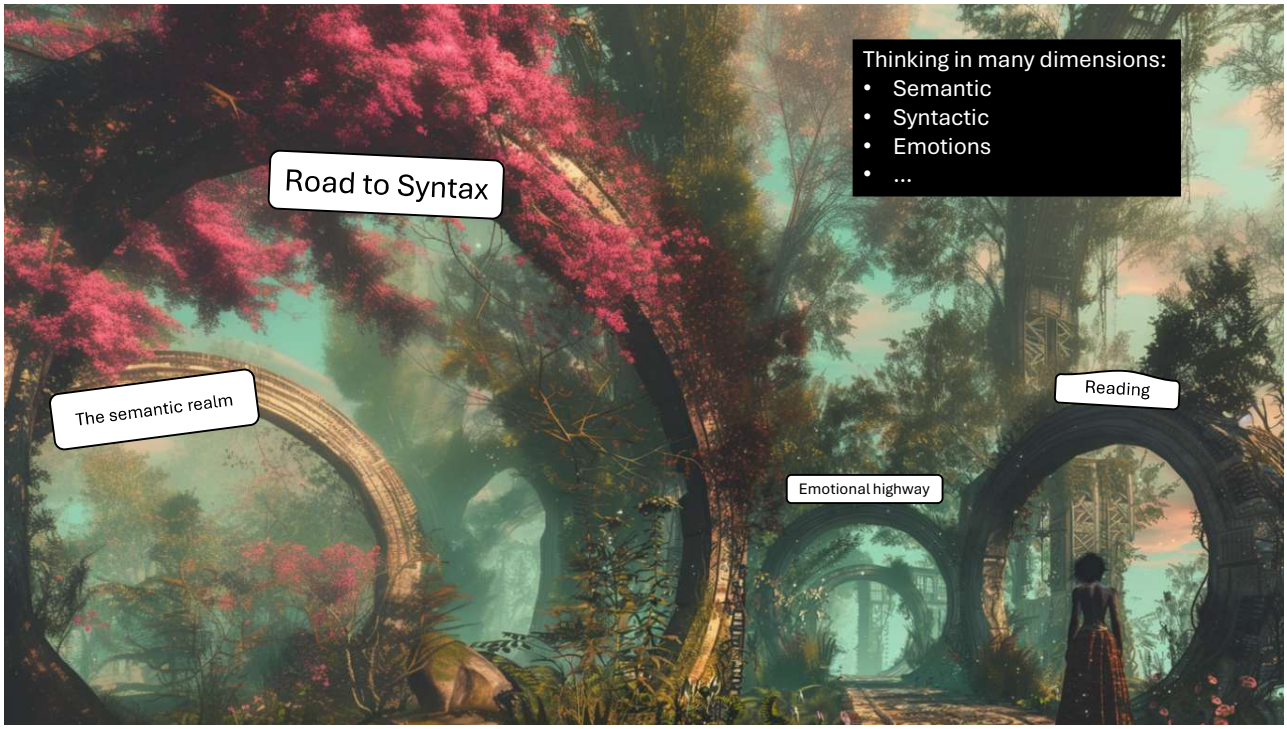
The plausibility of a filling for a blank in a sentence is influenced by the words close to the blank.

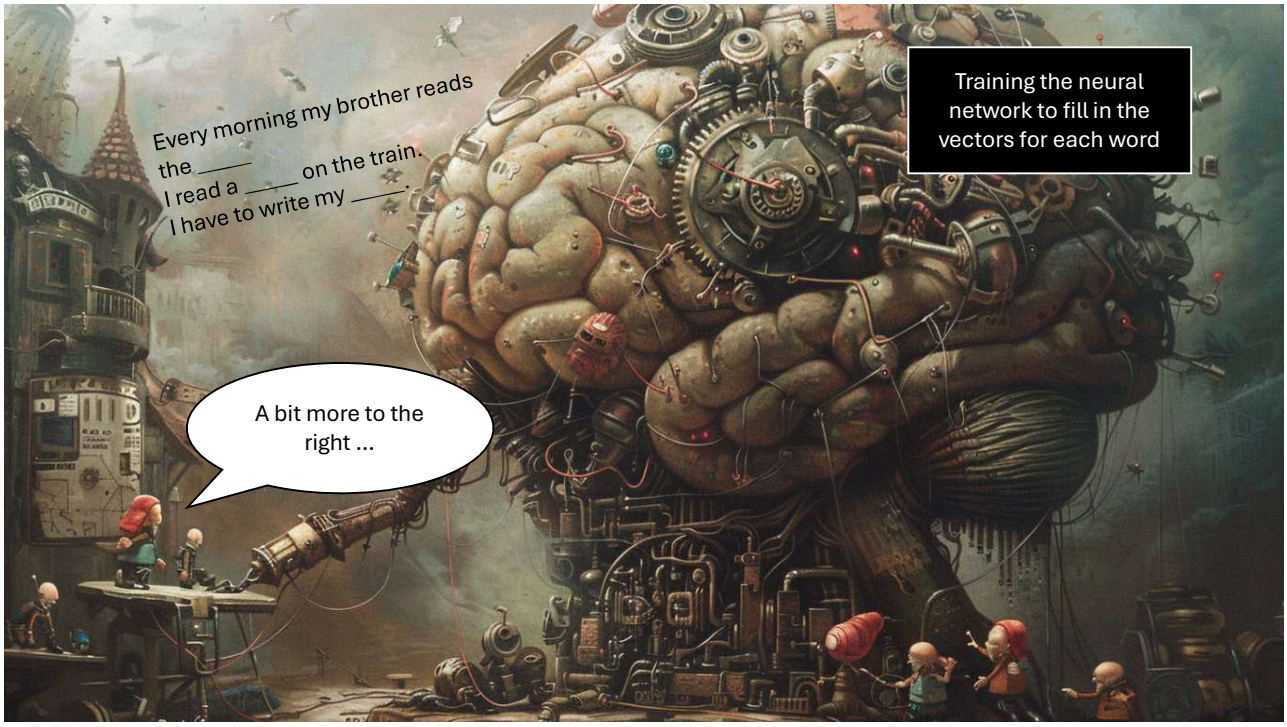
On the train I read a very interesting _____



Every morning my brother likes to read the _____

The plausibility of a filling for a blank in a sentence is influenced by the words close to the blank.




Every morning my brother reads the _____
I read a _____ on the train.
I have to write my _____.
Training the neural network to fill in the vectors for each word
A bit more to the right ...