A Little Help From a Friend

Nikola is a student assistant at a small research project at the University. He and his colleague Anna are both studying **Digital Humanities**.
As part of the project, Nikola was asked to write a small program that would "extend" gendered nouns in German texts by adding an alternative gender form.
Nikola wrote the code and was convinced that his approach was good. However, the results are... odd.
Frustrated, Nikola meets up with Anna at a café. He shows her his Python script and a few examples. “I really don’t get it,” he says. “The logic makes sense, but it’s not doing what I expected.”
Anna agrees to take a look — and that's where you come in.

What is the program supposed to do?
- Based on the output and the code, what do you think the intended functionality is?
- How is the gender transformation achieved?

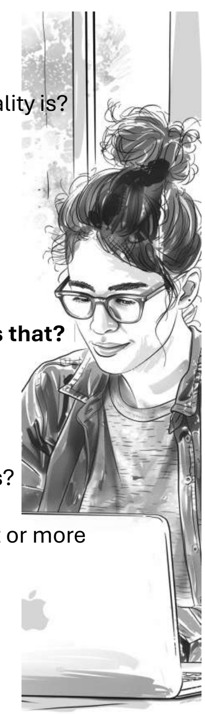
2. What works — and what doesn’t?
- Which sentences are transformed in a meaningful way?
- Which transformations are incorrect, confusing, or unnecessary?

Why are neutral nouns also transformed?
Can you identify the part of the code responsible?
How could this be prevented?

Some masculine nouns are not extended correctly (e.g., *Gärtner*). Why is that?
Take a look at the function that selects the alternative form.
What might cause a failure to find a matching female equivalent?

Why are animals like *Hund* and *Katze* included?
Is that desirable?
How could the code be adjusted to exclude non-human professions or roles?

Make Suggestions
- Suggest at least three changes that would make the output more consistent or more useful.
- Bonus: Can you improve how the alternative form is chosen?



Nicola *(leaning in, voice soft but fervent)*: You know, Marius, it's maddening—some people swim in obscene wealth, while others barely scrape by. And yet power doesn't always follow the money.

Marius *(sips his ale, nods)*: Exactly. For the seminar work, I've mapped the semantic patterns in newspaper coverage—what jobs are tagged as "powerful," who's "influential." Wealth shows up clearly, but influence? It's more subtle.

Nicola *(eyes bright)*: But what about gender? Women hardly ever land in those same semantic clusters. Even when they're shaking up the system.

Marius *(shrugs, half-smile)*: That's the glitch. My program can flag wealth and power mentions—implicit and explicit—but it's not tuned to gender bias yet.

Nicola *(slams her glass on the table)*: It should be! Imagine visualizing how media sidelines women's clout, even when they're running the show behind the scenes.

Marius *(leans back, contemplative)*: That's exactly what I want to highlight in my thesis – what the program is for: A tool that doesn't just count dollars and titles, but surfaces the hidden dynamics—gendered, racial, structural.

Nicola *(grinning)*: Now that's a revolution I'd subscribe to. Sign me up.

1. **What is the program supposed to do?**
   - Based on the output and the code, what do you think the intended functionality is?
2. **Rename if needed**
   - Some of the naming in the code is a bit hard to understand. Rename variables or methods to make them more understandable.
3. **Apply for axis gender**
   - Replace one of the axes with a male-female direction.
   - Which axis do you choose? What do you expect and what is the outcome?
4. **Discuss!**
   - To what extent does the visualization reflect objective reality, and to what extent does it mirror only the patterns learned from the training texts?
   - Which word associations surprised you, and which did you find predictable? Why?
   - How might results differ if we trained the model on literary texts instead of newspaper articles?
   - What risks arise when using a model like this—without accounting for its implicit biases—for real-world tasks (e.g., résumé screening or candidate ranking)?
   - How could biased semantic associations reinforce existing inequalities in hiring, lending or other decision-making processes?

[BBC Archive Newsreel, London, circa 1965]
[Rain patter, intro sting]
HOST (under umbrella):
"Good evening, London. Tell us—what's a man's job?"
GENTLEMAN (in bowler hat):
"Engineer."
HOST:
"And the female equivalent?"
GENTLEMAN (hesitating):
"Engineeress?"
[Quick cuts of passersby: "Doctor" → "Doctoress?", "Shopkeeper" → "Shopkeepress?"]
HOST (voice-over):
"Asking for the 'her' to every 'him' reveals our hidden assumptions."
[Cut to YOUNG WOMAN:]
HOST:
"A 'proper' women's job?"
YOUNG WOMAN:
"Nurse."
HOST:
"And the man's version?"
YOUNG WOMAN (smiles):
"Nurser—doesn't sound quite right, does it?"
[Fade to BBC lion roar over rain sound]

1. **Code Comprehension & Annotation**
   - Describe in your own words what the program does from start to finish.
   - inline comments to the code and rename variables/functions for clarity.
2. **Interpreting the Visualization**
   - Explain what each node and edge represents in the network graphs.
   - When a counterpart exists, compare the left- and right-hand graphs. How do they relate?
3. **Understanding the Bias Score**
   - What does a positive bias value signify? What about a negative value?
   - Choose three input words and report their bias scores along with the ♀ or ♂ symbol.
4. **Missing Counterparts**
   - Some words yield no counterpart and only a single plot. Propose two reasons why.
   - Hypothesize whether increasing num_similar could help. Double it, rerun, and report any new counterparts.
5. **Threshold Sensitivity**
   - Not all input words appear under the default neutral_threshold. Change it (e.g. to 0.01 or 0.05).
   - Which additional words now show up? Which drop out? Describe the effect of threshold changes on "gendered" words.
6. **Linguistic Reflection**
   - English has few explicitly gendered nouns. Why do you think explicitly feminine words show weak bias scores?
   - Can adjectives or verbs carry gender bias?
7. **Model Considerations**
   - The FastText model was trained on noisy news data. How might misspellings, domain shifts, or style affect results?

Anja: Just a moment, Grandma—I still have to finish this home assignment. I'm supposed to read these 18th-century letters for my paper, and sometimes I really struggle to follow what they're saying.
Grandmother: Is the handwriting too tricky?
Anja: Not exactly. I'm using a printed version. It's the words themselves. I tried translating them into modern German expressions, but it just doesn't fit.
Grandmother *(raising an eyebrow)*: Modern expressions? What do you mean?
Anja: Well, I thought using today's vocabulary would help—but then I realized I was imposing a modern perspective. For example, they use *Muße* all the time, and I translated it as "leisure," but it feels more nuanced in context.
Grandmother: Fascinating. Words really did carry different shades of meaning back then. You know, your great-grandmother used to say the same thing: "In my day, we used the same words—but we meant something completely different by them."
Anja *(smiling)*: Exactly! It's not just that words evolve; they reflect whole worlds. The ideas they express are worlds apart from ours.

1. **Setup**
   - Download the file.
   - Download and unpack the data files into the same folder.
2. **Inspect & Rename**
   - Run the script and note console messages.
   - Read each function—what does it do?
   - Rename methods/variables if names aren't self-explanatory.
3. **Experiment**
   - Test at least three other German words (e.g. "Zeit," "Buch," "Liebe") and observe differences.
4. **Discuss**
   - **Model Differences:** How do neighbour lists and scores diverge between a 19th-century literary model and a news-trained model? Why?
   - **Corpus Impact:** How would more literary or genre-specific texts (fantasy, historical novels) change the modern model?
   - **Practical Issues:** What pitfalls arise when using a 19th-century model for modern tasks, or a modern news model on historical texts?