Open Science in the Humanities: Galaxy as a Flexible and yet Reproducible Platform

Daniela Schneider[*] and Jan Leendertse[+], **and the European Galaxy Team**
[*] *Freiburg Galaxy Team, Bioinformatics Group, Department of Computer Science, University of Freiburg, Germany*
[+] *Central Data Facility, Computer Center, University of Freiburg, Germany*

Galaxy EUROPE

Challenges for digital Research

- Installing programs only with Admin access
- Low programming skills
- Software and hardware costs
- Where do I start?
- How do I share my data?

Galaxy[1,2] – data analysis and workflow

- Unified and intuitive web User Interface for **FAIR** data analyses
- Extensive **Open Educational Resources** for **self-study**
- 3000+ tools from standard to scientific software
- Developed by a large research community
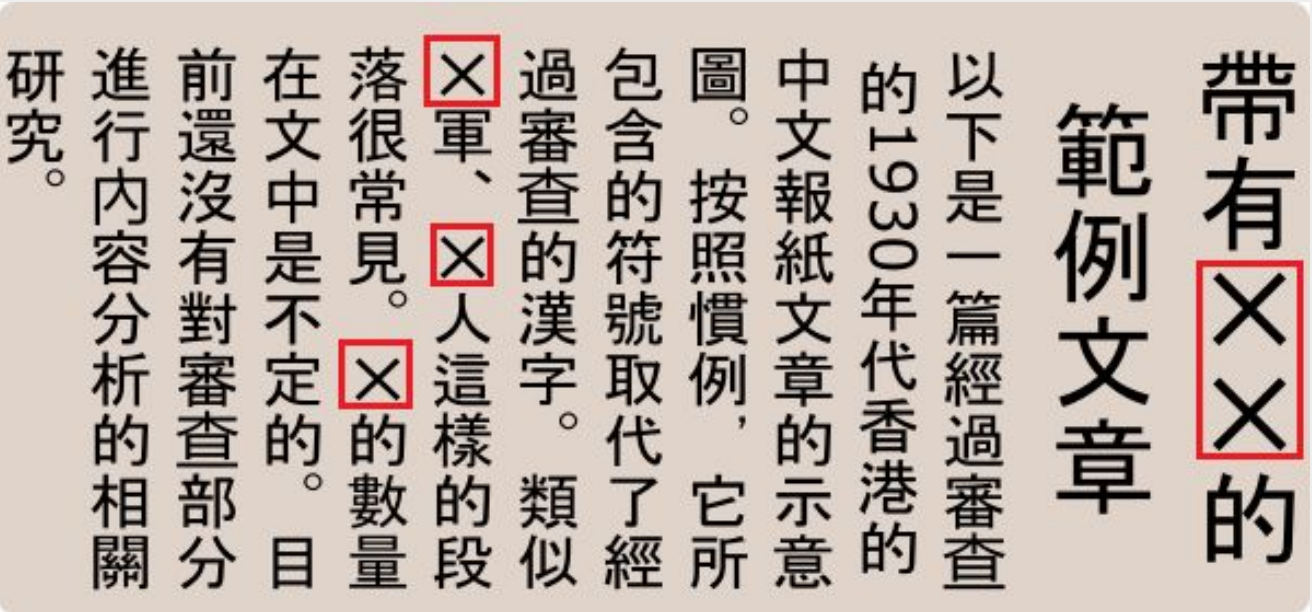- Contributions from many scientific disciplines

Reuse · Plan · Collect · Process · Analyse · Preserve · Share [6]

Wissenschaftliche Integrität
GO FAIR [4,5]

Overwhelmed?[7]

`diff` as single-user CLI

Example - uncovering censorship in Chinese newspapers

What content was censored in a Chinese newspaper? How can I find out without automatically?

Example of a censored Chinese article from the 1930s. "x" marks cut-out characters.[3]

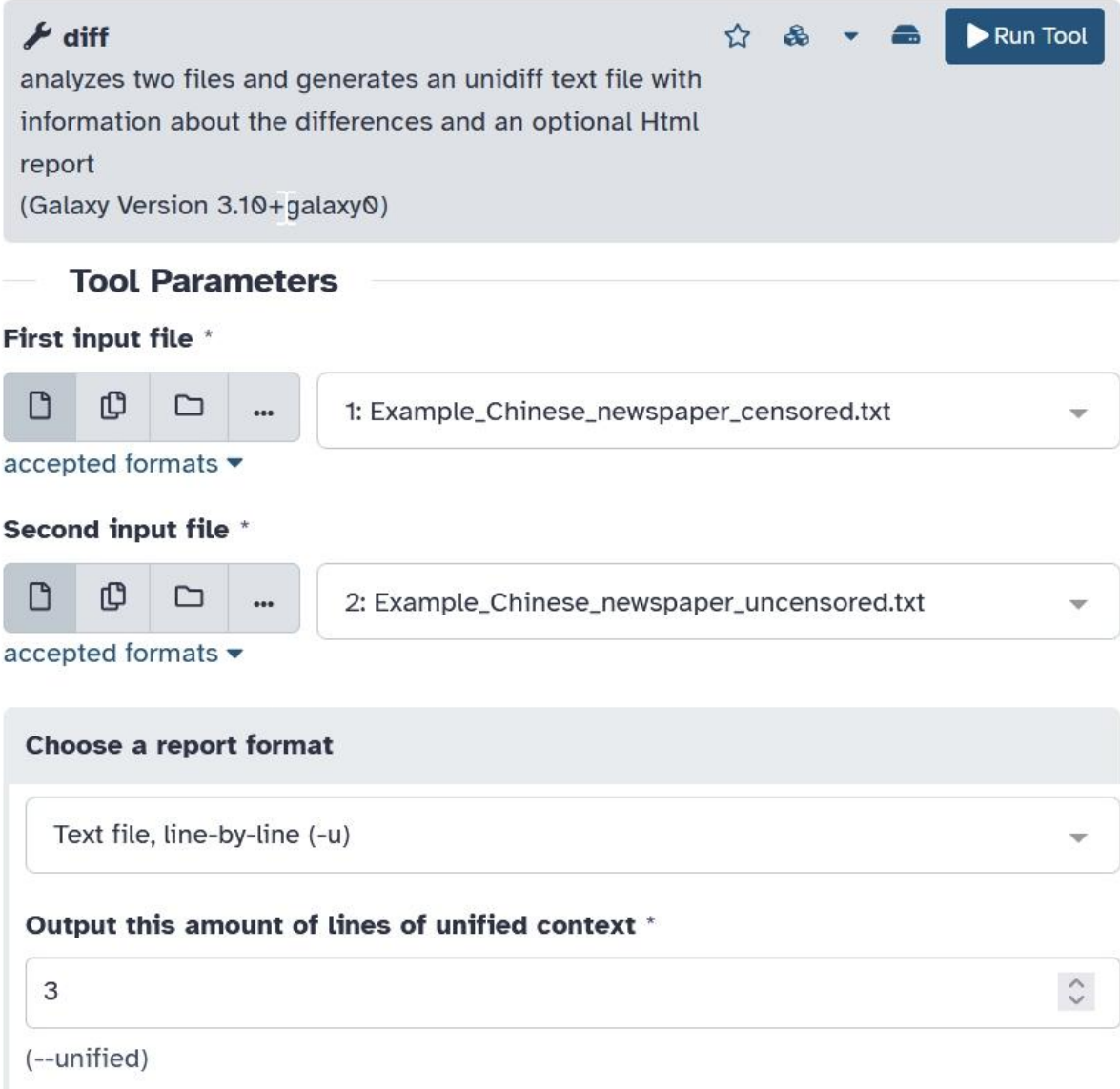We uncovered a second, uncensored version of the same article.

↓

Text cleaning

↓

Text comparison with `diff`-tool. The censored text is on the left, the uncensored one on the right. "x" marks derogatory terms for Japanese (enemy, brave, etc.)

↓

The last step extracts the censored parts of the article. One possible visualisation is a word cloud.

Galaxy Training Network

Training material available on **training.galaxyproject.org**[8]

`diff` as GUI in Galaxy

Finished OER text mining[9]

References

[1] The Galaxy Community. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update, Nucleic Acids Research, 2024;, gkae410, https://doi.org/10.1093/nar/gkae410

[2] https://usegalaxy.eu

[3] Adapted from: Schneider, D., 2024 Passed by the Censor. Press Censorship in the British Empire With a Special Focus on Hong Kong, 1938–1941 [Unpublished Dissertation]: Albert Ludwigs Universität Freiburg im Breisgau, p. 19.

[4] https://wissenschaftliche-integritaet.de/

[5] https://www.go-fair.org/resources/go-fair-workshop-series/gofair_logo_rgb/

[6] https://github.com/elixir-europe/rdmkit/blob/master/LICENSE

[7] Image generated with the prompt "Create a picture of a desperate researcher in the humanities. On the desk are large piles of books. The researcher is exhausted and overwhelmed with all the Chinese texts in front of her." by OpenAI, ChatGPT, 03/2025 (https://chatgpt.com/).

[8] Hiltemann, Saskia, Rasche, Helena et al., 2023 Galaxy Training: A Powerful Framework for Teaching! PLOS Computational Biology 10.1371/journal.pcbi.1010752

[9] Schneider, D., Text-Mining Differences in Chinese Newspaper Articles (Galaxy Training Materials). https://training.galaxyproject.org/training-material/topics/statistics/tutorials/text_mining_chinese/tutorial.html Online; accessed Tue Mar 04 2025.

Reusable workflow

Contact