Seamless Integration of Continuous Quality Control and Research Data Management for Indigenous Language Resources

Anne Ferger & Daniel Jettka | Paderborn University, Germany
Seamless integration of Continuous Quality Control and Version Control into existing workflows enhances the acceptance of the measures.
Continuous Quality Control for research data is crucial for better research data.

We noticed a big difference after employing continuous quality control mechanisms in the project INEL in the time needed for manual fixing of inconsistencies and the time needed to prepare the data for publication.
Seamless integration of new workflows or tools is eased using a VCS.

Combining a simplified solution using git (with the tool Lama) and automatic quality control fixes and checks (corpus services with automatic git scripts) doesn't change the existing data creation workflows too much.
Further information and tools:

Simplified git solution for users: **Lama**

Quality checks and automatic fixes for linguistic corpora: **corpus services**

Automatic git solution for corpus services: **cubo**
Contact

https://github.com/digitalhumanists

Anne Ferger
anne.ferger@uni-paderborn.de
https://twitter.com/anneferger1

Daniel Jettka
daniel.jettka@uni-paderborn.de
https://twitter.com/DJettka