

Student/Master Project: Linking to Python: Making the R-based HarmonizR package available to a wider audience

Description: To account for internal biases and shifts in measurements within biomedical data, the HarmonizR framework [1] has been developed to adjust for these so-called 'batch effects' while tolerating missing data points. HarmonizR is an 100 % percent R-based package and therefore only available for users using the R programming language. While the majority of the targeted audience - being the life sciences - uses R, the potential audience could be increased by providing the package in Python, a very popular programming language.

Prerequisites: Knowledge about the Python programming language and at least basic knowledge about the R programming language.

Contact: Chair for High Performance Computing
Simon Schlumbohm, schlumbohm@hsu-hh.de
Prof. Dr. Philipp Neumann, philipp.neumann@hsu-hh.de

References:
[1] Voß, H., Schlumbohm, S., Barwikowski, P. et al. HarmonizR enables data harmonization across independent proteomic datasets with appropriate handling of missing values. Nat Commun 13, 3523 (2022). <https://doi.org/10.1038/s41467-022-31007-x>