OpenITI release 2023.1.8

The 8th version (2023.1.8) of the [OpenITI](https://github.com/OpenITI) corpus is now available at [Zenodo](https://zenodo.org/records/10007820). The release is open access and is also accessible through our [GitHub repository](https://github.com/OpenITI/RELEASE). The primary version of the corpus, which contains a single digital version for each text in a flat folder structure, is also available at [Zenodo](https://zenodo.org/records/10021513). The primary versions of each text are marked as "primary" (PRI) in the corpus metadata.

This release features 8,600 unique books written by 3,254 authors and including all the different versions of the texts, the corpus now comprises 13,158 text files, which contains a total of 2,275,216,771 words. The new additions are available in the `ids.csv` file in the [release notes](https://zenodo.org/records/10007820/files/release_notes.zip?download=1) accompanying the corpus release.

A major change to this corpus is that the development team has removed editorial parts from all primary versions of the texts. These texts do not have editorial introductions, footnotes, indices or any other modern additions to the text and have the tag "CLEANED\_VERSION" in the metadata.

In addition to the new texts, annotation work has been carried out and all together we currently have 1,373 structurally annotated text from which 561 texts are vetted by a second annotator (extension: `.mARkdown`) and 801 books that have gone through one round of annotation (extension: `.completed`). For a full list of the new annotation work, see the `annotation\_update.csv` file in the [release notes](https://zenodo.org/records/10007820/files/release_notes.zip?download=1) in this release.

Moreover, for an overview of the modified URIs as well as the statistics on the corpus, please see the release notes and the corresponding csv file (modified\_uris.csv) in this release.

In order to cite this version, please use the following information (for the bibliographical export, see the [publication page](https://zenodo.org/records/10007820)):

Nigst, L., Romanov, M., Savant, S. B., Seydi, M., & Verkinderen, P. (2023). OpenITI: a Machine-Readable Corpus of Islamicate Texts (2023.1.8) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10007820