




python-ags4: A Python library to read, write, and validate AGS4 geodata files

Asitha I. Senanayake¹, Roger J. Chandler², Tony Daly³, and Edward Lewis⁴

¹ Fugro USA Marine, Inc. ² Bentley Systems ³ Association of Geotechnical and Geoenvironmental Specialists ⁴ British Geological Survey, Keyworth, UK

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: 

Submitted: 16 April 2022

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

Data gathered from geotechnical, geoenvironmental, and geophysical investigations can be broadly described as “geodata”. The AGS4 data format ([Association of Geotechnical and Geoenvironmental Specialists, 2017, 2021](#)) is one of the most widely used data transmittal formats for geodata and is used across the world. It is a plain text format consisting of multiple tables of comma-separated values, tied together with a robust data schema and a comprehensive suite of validation rules.

python-ags4 is a Python library that provides functionality to read, write, and validate AGS4 geodata files. It provides users with a gateway to access the full power of the Python ecosystem to explore, analyze, and visualize geodata. Pandas DataFrame ([The pandas development team, 2020](#)) is the primary data structure used within the library, therefore it can handle relatively large datasets reasonably fast. The data validation module checks the file for compliance with the validation rules and provides a detailed error report.

Statement of Need

This library fulfills the following needs of the engineering and scientific community that uses AGS4 geodata files:

- Provide a transparent and easily accessible tool to validate AGS4 geodata files.
- Provide access to the Python ecosystem to users of AGS4 geodata.
- Provide a cross-platform tool to work with AGS4 geodata files.

Having a free an open-source tool that can validate AGS4 files will help producers and receivers of data check that the geodata file is of the required standard. It also lowers the barrier to entry for those who are not familiar with this data format and fosters data sharing and collaboration. python-ags4 has already been adopted by the British Geological Survey to work with its large repository of publicly available AGS4 geodata files.

Python has a rich ecosystem of packages that can be utilized to analyze and visualize data. The ability to easily import geodata from AGS4 files will greatly enhance the ability of engineers and scientists to interpret data and improve knowledge of ground conditions. A Jupyter Notebook with examples of how to import data to Pandas and GeoPandas ([Jordahl et al., 2020](#)) for statistical and spatial analyses and visualize using Matplotlib ([Hunter, 2007](#)) is provided in the git repo of this library. We provide an example for users in the git repo - <https://gitlab.com/ags-data-format-wg/ags-python-library/-/blob/main/examples/ags.ipynb> ([Lewis, 2021](#)).

The command-line interface included with the library is a convenient and easy to use tool to

work with AGS4 geodata files. It is cross-platform and has been tested in Linux, Windows, and Mac environments.

Ongoing Projects

python-ags4 has been implemented in two free and open-source software projects, making the library more accessible to non technical users.

- Windows application offering the functionality of the library (<https://gitlab.com/ags-data-format-wg/ags-checker-desktop-app>)
- API and GUI (<https://agsapi.bgs.ac.uk/>) offering the functionality of the library on the web. (<https://github.com/BritishGeologicalSurvey/pyagsapi>)

It has also been used in commercial projects. (e.g. <https://digitalgeotechnical.com/apps-home/>)

Acknowledgements

The authors acknowledge the work done by the AGS Data Format Working Group to develop the AGS4 data format and the support it has extended to develop this library.

References

- Association of Geotechnical and Geoenvironmental Specialists. (2017). *Electronic transfer of geotechnical and geoenvironmental data*.
- Association of Geotechnical and Geoenvironmental Specialists. (2021). *Electronic transfer of geotechnical and geoenvironmental data*.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Jordahl, K., Bossche, J. V. den, Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur, A., ... Leblanc, F. (2020). *Geopandas/geopandas: v0.8.1* (Version v0.8.1) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.3946761>
- Lewis, E. (2021). *Example jupyter notebook*. <https://gitlab.com/ags-data-format-wg/ags-python-library/-/blob/main/examples/ags.ipynb>
- The pandas development team. (2020). *Pandas-dev/pandas: pandas* (latest) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.3509134>