

Open Humans - connecting, sharing and analyzing personal data that enables community-driven research

Bastian Greshake Tzovaras^{1,2}, Mad Price Ball²

¹Lawrence Berkeley National Laboratory, Berkeley, CA, USA, Email: bastian@openhumans.org

²Open Humans Foundation, Boston MA, USA

Project Website: <https://www.openhumans.org>

Source Code: <https://github.com/OpenHumans>

License: MIT

Open Humans is a community-driven platform that allows its members to upload, connect, and privately store a variety of personal data, including – but not limited to – genetic, activity, GPS, or social media data. *Open Humans* puts the individual in control of their data. Each *Open Humans* member can choose if and how to donate their data: While they can decide to share individual datasets publicly, they can also join and contribute to individual research projects.

To facilitate the creation of new research projects, *Open Humans* offers a rich toolbox to efficiently ask an engaged audience of participants to join and contribute to research projects. Projects can either rely on *Open Humans'* APIs to interact with its participants and get access to their data or run projects right on the site, using the website and a simple command-line interface to interact with them.

Both project participants and project leaders benefit from the community aspect of *Open Humans*. While participants can anonymously join a research project, project leads still retain a way to message their participants through *Open Humans*, enabling further data collection and interactions. Individual participants gain through the possibility of project leads to deposit new or enriched data back into their private *Open Humans* data storage.

This modular nature of *Open Humans* makes for a thriving, decentral community. Since its launch in 2015, over 5,000 people have joined *Open Humans*, uploaded a total of 32,546 data sets and created 18 projects and studies. Additionally over 50 projects are being under development – many of them driven by the community itself.

In our presentation we will highlight some of the functionalities of the *Open Humans* platform along with updates of what our community has build around these features. Examples include community-driven projects that add new data sources like activity tracking and social media data; add data-analysis features like genotype imputation and genetic genealogy.

Additionally we will present our latest project – *Personal Data Notebooks*. This integration of *JupyterHub* into *Open Humans* allows people to analyse their personal data in *Jupyter Notebooks*. The resulting notebooks can be shared with other users who can then run them on their own personal data, without having to give away control over it.