As a practicing computer scientist, when I applied to Library School in the mid 80’s, they were a bit shocked, but I explained that these areas were about to merge, given the coming Internet. That merge has now happened, but most Universities do not yet educate students with this in mind.

The Internet Archive, a non-profit digital library, has 20 petabytes of data in thousands of collections, programmer interfaces, and millions of users every day. We are starting to get digital humanitarians and researchers coming expecting to be able to use these collections.

Unfortunately, the tools, our staff, and our users are struggling to keep up. We would like to hire your graduates, we would like to serve your graduates as researchers. There are things a University can do to help.

Towards that goal, I would like to outline some of challenges/opportunities we see on the interface of big data and people: rights and privacy issues, tool building, programming for data analysis.

Dr. Brewster Kahle, is a passionate advocate for public Internet access and a successful entrepreneur who has spent his career intent on a singular focus: providing Universal Access to All Knowledge. He is the founder and Digital Librarian of the Internet Archive, one of the largest libraries in the world. Soon after graduating from the Massachusetts Institute of Technology where he studied artificial intelligence, Kahle helped found the company Thinking Machines, a parallel supercomputer maker. In 1989, Kahle created the Internet’s first publishing system called Wide Area Information Server (WAIS), later selling the company to AOL. In 1996, Kahle co-founded Alexa Internet, which helps catalog the Web, selling it to Amazon.com in 1999. The Internet Archive, which he founded in 1996, now preserves 20 petabytes of data - the books, Web pages, music, television, and software of our cultural heritage, working with more than 400 library and university partners to create a digital library, accessible to all.