Data Science - Cover Letter

I am Computer Science PhD student at Stanford specializing in **Data Science** and HCI, aiming to start full-time opportunities in September. I am experienced in using **PyTorch** for deep learning, and can conduct data analyses in **Python** and **R**. My research experience is in conducting **large-scale data science** experiments (during my PhD work at Stanford using Coursera's datasets, as well as on my own HabitLab system), modeling user behavior using **deep learning** (during my PhD work at Stanford on my HabitLab system), as well as building **machine learning-powered systems**, including **natural-language processing** for review quality prediction (during my internships at Google), and **reinforcement learning** (during my PhD work at Stanford on my HabitLab system). I am a US citizen and would prefer to work in the San Francisco Bay Area. My portfolio is at https://www.gkovacs.com and my resume is at https://www.gkovacs.com/resume.pdf

During my PhD, I have run a number of data science and machine learning experiments on HabitLab (https://habitlab.stanford.edu/), a behavior change system I built that provides users with personalized interventions learned through reinforcement learning. It is an in-the-wild **large-scale A/B testing** platform that has been installed by over **300,000 users**, has over **12,000 daily active users**, all organic installs achieved without paid advertising, from which I have derived a dataset of over **800 million browsing sessions** which I analyze in my research studies. I have done a variety of data science and machine learning work with this platform, including predicting user preferences for interventions over time (classification using **recurrent neural networks — deep learning**), personalizing interventions to maximize effectiveness (**reinforcement learning** using multi-armed bandits) predicting time spent on sites (regression using **random forests — machine learning**), as well as a number of analyses such as predicting effects of intervention rotation on intervention effectiveness and user retention (**cox regression** and **linear mixed models — data science**), and modeling externalities of interventions on time spent (**linear mixed models — data science**). My publications listed at https://www.gkovacs.com and https://www.gkovacs.com/resume.pdf describe these various data science and machine learning projects in additional detail.