



Seattle, December 29, 2013

Letter of Commitment for NSF Proposal: **CHS: Small: Interactive Machine Learning for Text Analysis**, by
Jeffrey Heer

Dear Colleagues,

I am writing to document my collaboration and provide very strong commitment for Professor Jeffrey Heers's NSF proposal entitled "**CHS: Small: Interactive Machine Learning for Text Analysis**."

This is an ambitious proposal addressing crucial fundamental research in building holistic intelligent machine learning systems that include humans in the loop. Much research has been devoted to the development of machine learning techniques in the text domain. However, in practice, much of the actual human's effort is devoted to iterating over the data, refining features, finding "modeling bugs" and defects in the data. These "human in the loop" problems have been largely ignored by my community. Professor Heer's proposal brings together a unique combination of data visualization, machine learning and HCI techniques to provide a system to tackle this challenge head on. If funded, this project could change how we interact with data.

As a point of introduction, I'm the Amazon Professor of Machine Learning at the Computer Science & Engineering Department of the University of Washington. I'm also a co-founder and CEO of GraphLab Inc., focusing large-scale machine learning and graph analytics. Previously, I held positions that include the Finmeccanica Associate Professor at Carnegie Mellon University and senior researcher at the Intel Research Lab in Berkeley. My research focuses on large-scale machine learning, with a variety of applications, including in text data.

Since Prof. Heer moved to UW, we have started collaborating on building a system for large-scale machine learning with humans in the loop. In particular, an exciting aspect of our collaboration is the combination of our complementary backgrounds, with our desire to actually build systems that impact a large number of people. We have recently started a joint project in this regard, with a PhD student in our department. If Prof. Heer's NSF proposal is funded, my group will continue to be an unfunded partner, thus amplifying the funding. In particular, we will collaborate on investigating active learning methods for text analysis, e.g., posterior regularization for feature-space constraints, where the user introduces constraints on the solution rather than simply labeling documents. Later in the project, when scalability becomes challenging, we will provide support for usage, implementation and debugging of his proposed systems and algorithms by building on our open-source machine learning system, GraphLab. My group will be excited to transition Prof. Heer's work into GraphLab, and we will also adapt future versions of GraphLab to better-fit Prof. Heer's needs. Additionally, GraphLab now has a



UNIVERSITY *of* WASHINGTON

COMPUTER SCIENCE & ENGINEERING

broad user base in industry and academia, receiving many thousands of downloads. If this project is funded, we will distribute any jointly developed open-source system as part of the standard GraphLab distribution. Thus, Prof. Heer's project will immediately have visibility and attention in this growing machine learning community, in addition to his own tremendous visibility in the visualization arena.

Please feel free to contact me if you require any more information.

Sincerely,

A handwritten signature in blue ink, appearing to read 'Carlos Guestrin'.

Carlos Guestrin, Associate Professor
Amazon Professor of Machine Learning

Carlos Guestrin, Associate Professor, Amazon Professor of Machine Learning
Paul G. Allen Center for Computer Science & Engineering
Box 352350 Seattle WA 98195-2350
Tel: 206-685-0486 Email: guestrin@cs.washington.edu