

Evaluation as Enabling Tool for Research and Development

Daniel Marcu
ISI, USC

Abstract

Students of natural language often perceive evaluation as “the thing one should do in order to get a paper published”. Seasoned researchers understand that evaluation is a crucial step in the research and development process loop; the better one understands how to evaluate a system, the higher the chance to build a good one. Funders and policy makers understand that evaluation is a means to further progress in a field. In this talk, I first explain why building summarization programs and properly evaluating them is a difficult enterprise that gives headaches even to seasoned researchers and funders. I then discuss pitfalls of previous evaluations of summarization systems and show how we attempt to avoid them in the context of a new series of Document Understanding Conferences.