Moving NLP out of the lab

Experimental practices in NLP have made unrealistic assumptions about the nature of data (e.g., i.i.d-assumptions) and the availability of resources. This has led to interest in cross-domain adaptation and cross-language adaptation. We argue that most of this work still makes unrealistic assumptions about data and resources. More realistic assumptions will introduce interesting new challenges, we argue, but also force us to explore new opportunities, bringing us closer to getting robust NLP for the world’s languages and all their varieties.