# Migrating ACL Anthology Network / "All About NLP"

Presented by Jay Hou and Clark Xie. Additional group members Dan Keller and Andre Monteiro.

Department of Computer Science, Yale School of Engineering and Applied Sciences, New Haven, CT

LILY

## Introduction

- ACL Anthology Network / "All About NLP"
- Collection of NLP (Natural Language Processing) papers, tutorials, and resources
- More than 21k papers and 17k authors
- Hosted on http://tangra.cs.yale.edu/index.php

## Motivation and Goals

1. For learning:
- Easily find resources on a given topic
- NLP topics and related topics taxonomized
- Lectures and tutorials in addition to papers

2. For research:
- Find related publications and leading papers
- Find authors and colleagues along with the number of papers published, number of times cited, h-indexes, and more
- Visualize credibility of papers and authors

## Technical work

- Moving the codebase and database from University of Michigan to Yale on tangra
- Updating from php5 to php7
- Fixing bugs and manual QAing to ensure consistency between sites
- Updating the database by uploading papers from recent conferences through admin site
- Integrating PDF box to perform automatic text extraction and reference building
- Developing a tool for manually adding PDFs and online resources

**Figure 1:** The index page of the website



**Figure 4.** Ability to list and sort authors



## Future Features

- Single document citation-based summary
- Build paper prerequisite chains
- Personalized learning progress tracking and recommendation
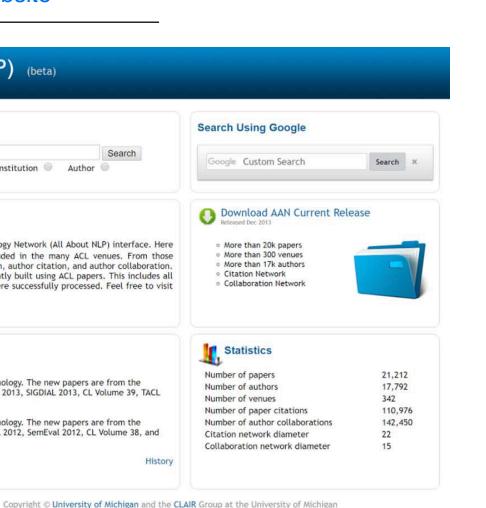- Identify hot, outdated, and developing topics for research

**Figure 2.** Tools for learning



**Figure 3.** Tools for researching