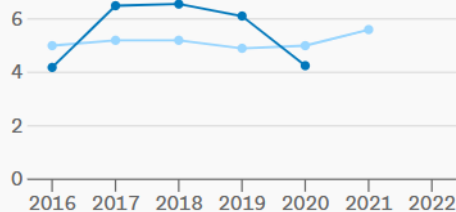


The Center for Investigative Reporting (CIR) is a non-profit news organization affiliated with Reveal News and Mother Jones, both of which focus on investigative journalism. CIR sued the government to make company-level injury and illness data public. Following the successful suit, CIR partnered with the University of Chicago's Data Science Clinic to improve a data

Injury Rate

Injuries per 200,000 work hours (or 100 full-time employees)

■ AMAZON.COM.DEDC LLC: reported General Warehousing and Storage
■ Industry average



pipeline in two ways: better location geocoding and increased industry matching accuracy. This builds on previous work from the clinic that created a website allowing users to compare specific company data with broader industry data, both in table and graphic formats, as shown in Figure 1.

The team worked to reduce false or missing values in the data displayed on the website by using North American Industry Classification System codes to hierarchically match industry and company-level data through the shared classification. However, there is still work to be done in reducing missing and unmatched data.

Figure 1: Reveal News Injury Data Search

More accurate geocoding through Bing Maps has significantly reduced the number of missing location values to negligible levels, as illustrated in Figure 2. This effort enhances the reliability and accuracy of the data available on the website, providing users with a clearer picture of workplace safety and industry standards.



Figure 2: Proposed Bing Maps Geocoding