

The Bluebook as Code: A Deterministic Engine for U.S. Legal Citation

Eric Martínez

Texas A&M University School of Law
ericmart@law.tamu.edu

Matthew Dahl

Yale Law School
matthew.dahl@yale.edu

Abstract

Legal citation is the bedrock of legal practice and scholarship, but formatting citations correctly requires navigating complex manuals such as the *Bluebook*, the notoriously exacting 500-page rulebook governing U.S. citation form. We present a deterministic execution engine that verifiably implements the rules of the *Bluebook* for the first time, covering 113 citation types across the manual’s practitioner and academic regimes, including jurisdiction-specific conventions for the U.S. federal and state systems and 44 foreign jurisdictions, as well as the required italics and small capitals. Its rule-mirrored architecture and facts-only input schema make its decisions transparent and allow it to serve as a reusable component of legal NLP systems. We show that the engine reproduces 99.1% of the manual’s printed examples and 97.9% of a held-out sample of 962 *Harvard Law Review* citations under exact string matching with typeface enforced. Given gold source information for the same inputs, six frontier language models average only 33.9% on the same task, with reasoning yielding minimal improvement and in-context rule text closing only part of the gap. We release our engine as a public web application and API for programmatic integration.