IN KEEPING with its statutory powers to independently enforce federal laws forbidding employment-related discrimination based on race, color, religion, sex, or national origin, the Equal Employment Opportunity Commission (EEOC) typically files several hundred cases annually against private employers.

The EEOC’s employment-related civil rights cases are an especially important part of the federal district courts’ total civil docket. Whether adjudicated or settled, they directly implement federal anti-discrimination policy. However, they have rarely been systematically studied, says Pauline Kim, the Charles Nagel Professor of Law and a renowned employment law scholar.

This EEOC enforcement litigation takes place in the U.S. federal district courts. The trial court judges who staff these courts differ from their counterparts in the appellate courts and the U.S. Supreme Court in significant ways, yet research on their activities and decision-making processes has been relatively sparse, and many of the empirical studies are methodologically flawed, Kim observes.

All that is changing—thanks to the vision of Kim and two other law professors, along with their exacting use of technology and quantitative methods. In October 2012, faculty in the law school’s Center for Empirical Research in the Law (CERL) released a new online resource, the EEOC Litigation Project (eeoclitigation.wustl.edu). The database is the work of Kim; Andrew Martin, vice dean, the Charles Nagel Chair of Constitu-
The EEOC database contains extensive quantitative data about a randomly selected stratified sample of more than 2,300 enforcement lawsuits from the EEOC’s federal court docket of 4,000 cases from 1997 through 2006. The project is primarily of interest to scholars for two reasons—1) it allows systematic study of the EEOC’s enforcement activities; 2) it provides data that can be used to study the litigation process in federal district courts more generally, by analyzing the subset of cases involving EEOC-initiated litigation.

“The unique dataset permits rigorous empirical study of the entire litigation process from filing through the decision to appeal—along with insight into litigation dynamics and the interaction between litigant and judicial decision-making,” Kim says.

General advantages to the database’s focus on EEOC litigation include:

- Details of case outcomes are usually available as matters of public record;
- In a single issue area, trends in judicial decisions, settlement, and party win rates can be systematically analyzed; and
- Because one party’s identity can be held constant, how characteristics of the other party influence process and outcome can be studied.

The database information is lucidly parsed, remarkably detailed, and accessible to the public. It has been assembled into “bricks” of data that can be downloaded for use. The database also includes information about each case’s substantive allegations, principal actors, and litigation processes.

In particular, the database offers details on the amount and types of relief the EEOC obtained, with emphasis on the terms of any injunctive relief. It also includes details about all types of court decisions—published and unpublished, final and non-final, and written and summary. Similarly, it contains all kinds of outcomes—default, settlement, pretrial adjudication, and judgment after trial, as well as litigation events prior to resolution. A downloadable, 96-page codebook lists the variables collected and how each was coded.

One unusual feature is that while most databases are flat, presenting data in a two-dimensional grid of rows and columns, this database incorporates multiple tables capturing as many case events as occurred, ranging from zero to thousands.

“To build the database, we obtained a list of all EEOC cases brought in federal court; it included considerable basic information,” says Martin, who at CERL has developed a dozen projects for Washington University faculty.

Next came the scientific case-sample selection. Then the team used district-court docket numbers from the EEOC to gather extensive additional data from the federal courts’ web-based docketing system, PACER (Public Access to Court Electronic Records). The massive amount of coding for the project was accomplished with what Martin calls “an army” of graduate students who are pursuing law degrees or PhD degrees in political science, as well as Arts & Sciences undergraduates, most of whom are either political science majors or in American Culture Studies.

Kim and Schlanger are currently using the database to examine the injunctive relief ordered or agreed to in resolved EEOC cases to analyze the broader impact on the workplaces concerned. Other scholars have begun to use the data to pursue additional inquiries.

One indicator of the interest in the project are the invited scholars in the fields of law, political science, and economics who attended a closed conference at the law school in spring 2013. “All are doing empirical work on employment discrimination or the litigation process,” Kim says. “After learning more about the data, they began formulating projects to pursue and will be producing related scholarly papers in the fall.”