



# Getting to CSV: Unlocking Public Data in the Civil Justice Space

Dan Bernstein,  
Research Analyst





# Overview

- Why do we care about court data?
- What is civil court data?
- How are we getting court data?
- What lessons can we learn from this project?



# What is the Legal Services Corporation (LSC)?

- A congressionally established corporation created to fill “the justice gap”
- The largest funder of civil legal aid for low-income Americans in the United States
- Provided over \$400 million in 2018 in funding to 132 grantee organizations in the 50 states and territories



# “The Civil Court Data Project”

- Our goal is to gather **state court data** at the **state and local level** to:
  - Answer fundamental questions about the experience of low-income individuals in the civil justice system
  - Develop tools to promote equality in the civil justice system
  - Develop relationships with local organizations who can utilize court data and analysis



# Garnishment in Tennessee during COVID

- Advocates wanted:
  - TN Supreme Court to stop debt collectors from garnishing stimulus checks
  - asked for any recent data on debt collection and garnishment in April 2020
- We determined:
  - nearly 2500 instances of garnishment activity since courts closed
  - \$200,000-\$400,000 collected in garnishment during COVID
- Our partners submitted the analysis to the Tennessee Supreme Court to inform next steps

# What does court data look like?

unique case number with the court system

Case:

Case type, date filed, and court location

Parties:

full name, street address, attorney name ('pro se' means no legal representation)

Judgment:

Outcomes, money owed, interest rates, other notes

Events: original filing, payments and fines, hearings

Skip to Main Content | Logout My Account | Search Menu | New Civil Search Refine Search Back Location : All JP Courts Help

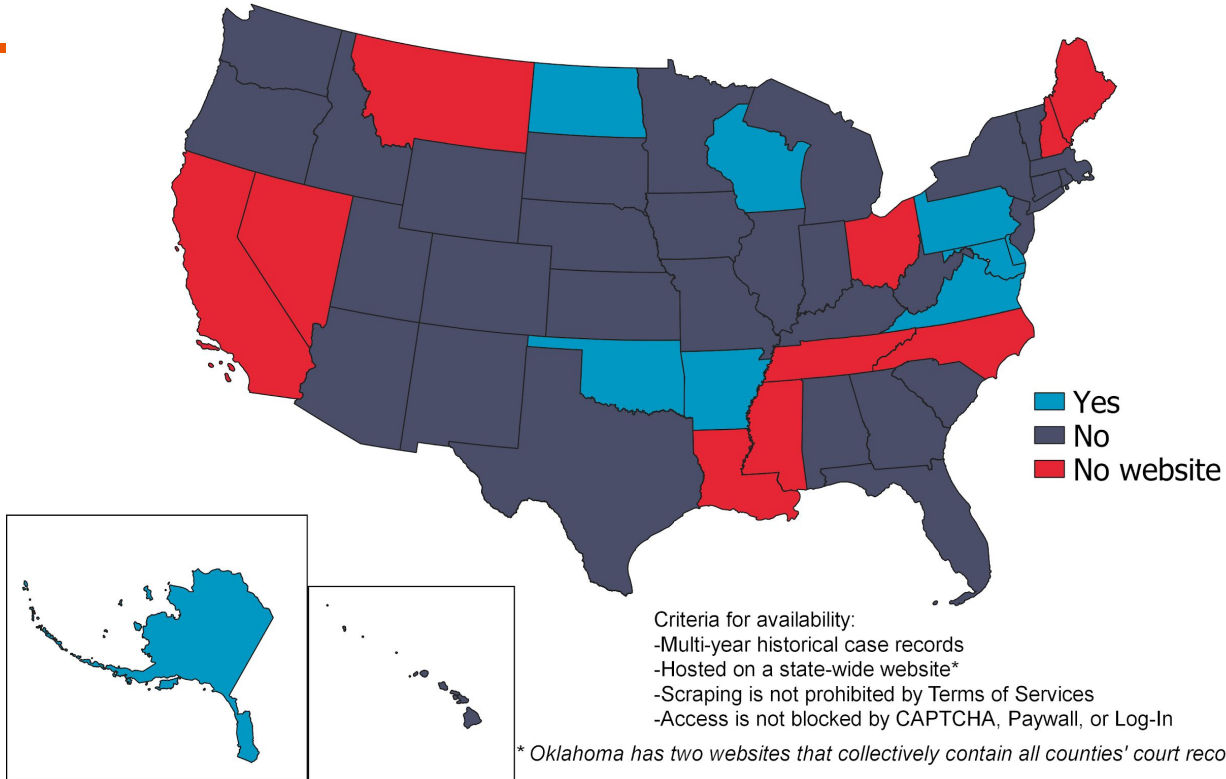
REGISTER OF ACTIONS	
CASE No. JP02-F00021966	
HURST INN APARTMENTS vs JOHN ERIK KOEPERNIK & ALL OCCUPANTS	§ § § § §
Case Type: Evictions Date Filed: 01/03/2000 Location: JP No. 2	
PARTY INFORMATION	
Defendant	KOEPERNIK, JOHN ERIK 122 HARMON RD #25 HURST, TX 76053 Lead Attorneys Pro Se
Plaintiff	HURST INN APARTMENTS 122 HARMON RD #208 HURST, TX 76053 Pro Se
EVENTS & ORDERS OF THE COURT	
01/13/2000	<b>DISPOSITIONS</b> <b>JUDGMENT</b> (Judicial Officer: Blerman-Davis, Linda) Judgment of \$480.00 Awarded To: HURST INN APARTMENTS Awarded Against: JOHN KOEPERNIK Comment: Cross Action: N PJ Atty Intrst: 0.00 Interest Set Date: 01/13/2000 Trial Date:
01/03/2000	<b>OTHER EVENTS AND HEARINGS</b> <b>Case Filing</b> CASE FILING   CASE FILING
01/03/2000	<b>Payment Receipt</b> RECEIPT OF PAYMENT   PAYMENT RECEIVED
01/03/2000	<b>Paper</b> ISSUE PAPER
01/03/2000	<b>Attorney Appointed to Defn Pursuant to SSRA</b> FORC   FORCIBLE DETAINER CITATION
01/03/2000	<b>Historical Comments</b>

Tarrant County, TX Justice of the Peace Court.



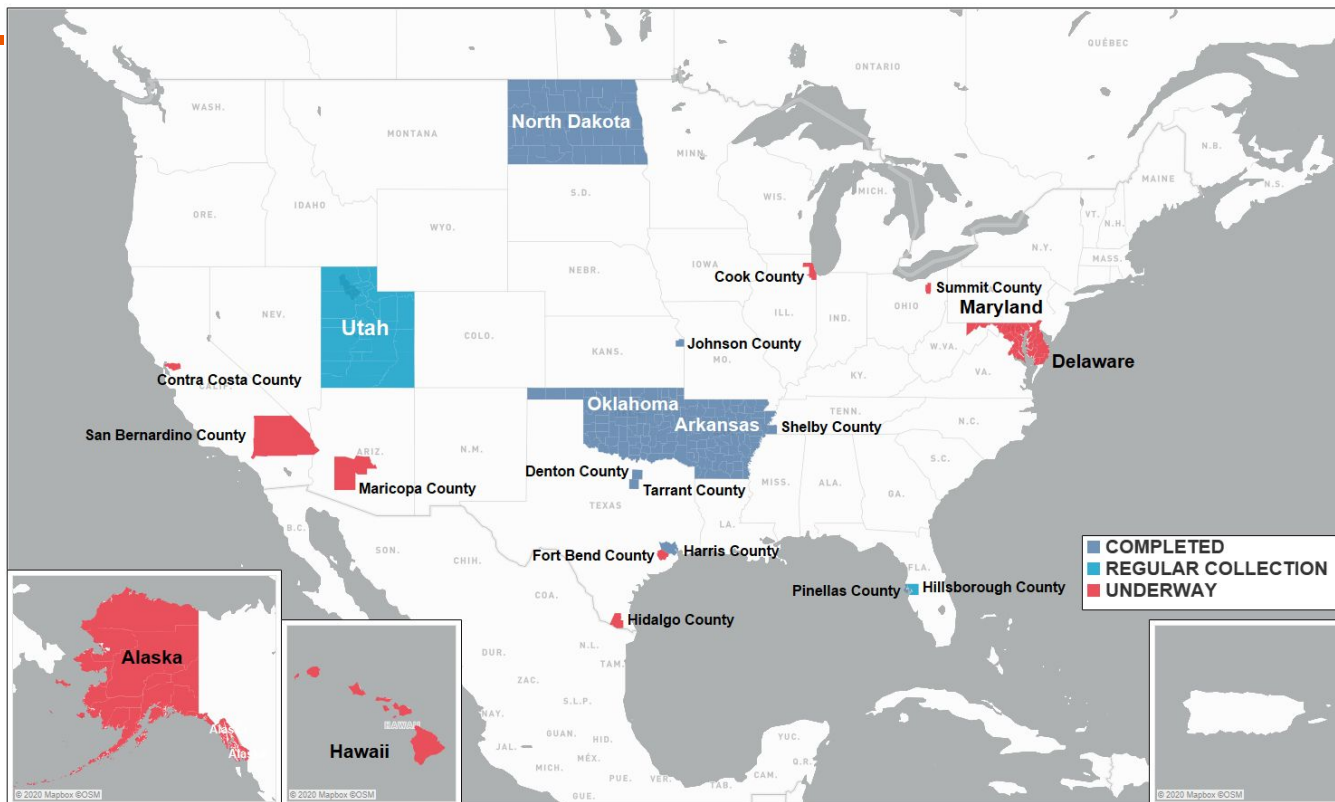
**So what's the problem?**

## Court Data Availability via Statewide Portals, 2020





# Data Collection (April 2020)



# Court data comes in all shapes and sizes



Home Search Results Log on

**1JU-01-00886C1 Estate of marguerite, Scott vs. Satal, Mark PBF**

Case Type: Eviction District Court (1JU) Action: Eviction - District Court  
 Case Status: Closed Case Judge: Froehlich, Peter B  
 File Date: 06/12/2001 Next Event:

All Information Party Docket Disposition

**Party Information**  
 Estate of marguerite, Scott - Plaintiff

DOB: [Field] Alias: [Field]  
 Disposition: [Field] Party Attorney: [Field]  
 Disp Date: [Field]

Satal, Mark - Defendant

DOB: [Field] Alias: [Field]  
 Disposition: [Field] Party Attorney: [Field]  
 Disp Date: [Field]

**Docket Information**

Date Docket Text

06/12/2001 All Records Prior to 06/24/10 are Converted Records. For please come to the Clerks Office with the case number or mail inquiries to: Jurnieu Trial Courts, 123 4th Street, Jurnieu, AK 99801. Please note that some converted data, especially dates, may differ from the original data.  
 Filed Date: Jun 12 2001

06/12/2001 Forcible Entry & Detainer Prior to Conversion Date: 06/16/2001 Time: 09:00 AM Judge: Peter B Froehlich Location: 1JU Event Text: FED (15 MIN)

06/12/2001 \*\* CONVERTED PAYMENT HISTORY AS OF 06/24/10 Receipt Date : 06/12/01 Receipt Number : 238740 Party from : LAW OFF OF PATRICK COHNDY.

**Case 95CV10831 Caption NEIGHBORS ET AL vs. FRONT ROW VIDEO ET AL**

**Chapter 61 Nature RE**

**Status TERMINATED Judge DANIEL W VOKINS**

**Division M4**

Search Case History (ROA) Plaintiff/Defendant Court Events Other Cases Accounting

Sort By Ascending Order Sort by Descending Order Print Friendly

10/11/1999	CASE CONVERTED TO JIMS
10/30/1995	FILE STAMP 10/25/95, MICR # 00199, SHERIFF'S RETURN FROM JO. CO., KS. SERVED LARRY D. LOCHNER, JR. 10-13-95 C/S
10/05/1995	FILE STAMP 10/05/95, MICR # 00115, WRIT OF RESTITUTION AND EXECUTION AND RETURN;REQUEST AND SERVICE INSTRUCTION FORM FILED WRIT OF RESTITUTION & EXECUTION ISSUED TO SHERIFF OF JO CO KS
09/29/1995	FILE STAMP 09/26/95, MICR # 00319, SERVICE RETURNED FROM SPECIAL PROCESS SERVER SERVED LARRY LOCHNER 9-11-95 C/S
09/26/1995	FILE STAMP 09/21/95, MICR # 00240, JOURNAL ENTRY OF JUDGMENT FOR PLTF AGAINST DEFTS FOR##\$3819.00 & COSTS & POST JDMT INT## CASE STATUS HAS BEEN CHANGED FROM PENDING TO TERMINATED
09/08/1995	FILE STAMP 09/05/95, MICR # 00697, PETITION (61) (FORCIBLE DETAINER). REQUEST AND SERVICE INSTRUCTIONFORM FILED. SUMMONS AND PETITION (1) ISSUED TO PLAINTIFF'S ATTORNEY FOR SERVICE BY PROCESS SERVER. ANSWER DATE 9-21-95.

**Docket Report Results**

Case ID: 778798  
 Docket Start Date:  
 Docket Ending Date:

**Case Description**  
 Case ID: 778798 - THE TIMBERS vs BAMFORD, TRACY -Non-jury Trial  
 Filing Date: Monday, January 04, 2000  
 Type: 16 - FED - OTHER  
 Status: none

**Related Cases**  
 No related cases were found.

**Case Event Schedule**  
 No case events were found.

**Case Parties**

Seq #	Assoc	Expn Date	Type	ID	No
37					TR
509					SA
53					GC

Go to Main Content Local My Account Search Menu News-Civil Search Refine Search Help Location / All JP Courts Help

**REGISTER OF ACTIONS**  
 CASE No. JP02-F00021966

Case Type	Evictions
HURST INN APARTMENTS vs JOHN ERK KOEPEKNIK & ALL OCCUPANTS	01/03/2000 JP No. 2

Plaintiff Enumeration

**CaseSearch** District Court of Maryland

Go Back Now

**Case Information**

Court System: District Court For Baltimore County - Civil  
 Location: Towson  
 Case Number: 08-04-0000217-2000  
 Title: COMCAST CABLEVISION OF MD, LP vs SMITH, ANN  
 Case Type: Contract  
 Filing Date: 01/04/2000  
 Case Status: Closed

**Causes Information**

Remedy Type	Amount	Comment
Not Applicable	\$273.40	CONTRACT ; DISMISSED (RULE 3-507)

File Date: 01/04/2000  
 Cause Description: 001 ; REGULAR CLAIM  
 Filed By: Plaintiff  
 Filed Against: Defendant

**Involved Parties Information**

**Defendant**  
 Name: SMITH, ANN  
 Address: 112 MARYBIN CIR  
 City: OWINGS HILLS State: MD Zip Code: 21117

**Plaintiff**  
 Name: COMCAST CABLEVISION OF MD, LP  
 Address: 8031 CORPORATE DR  
 City: BALTIMORE State: MD Zip Code: 21236

**Attorney(s) for the Plaintiff**  
 Name: KODECK, NED S  
 Appearance Date: 01/04/2000  
 Address Line 1: STE 203 B RESERVOIR CIR  
 City: Pikesville State: MD Zip Code: 21208

**Judgment Information**

Case Judgment Comment History

Amount: 0.00 Interest: 0.00 Atty\_Fee: 0.00 Judg\_Cost: 0.00  
 Other\_Amount: 0.00 Poss\_of\_Prop: 0.00 Damages\_Pop: 0.00  
 Value\_of\_Prop: 0.00 Damgs\_Val\_of\_Prop: 0.00 Repl\_Detrn\_Amnt: 0.00

Lead Attorneys  
 Pro Se

Pro Se

Set Date: 01/13/2000 Trial Date:

# What I think my workflow will be...



Evaluate website for data quality, quantity, and “scrape”-ability

↓

Scrape all case numbers

↓

Scrape case records using case numbers

↓

Parse sections from HTML case records

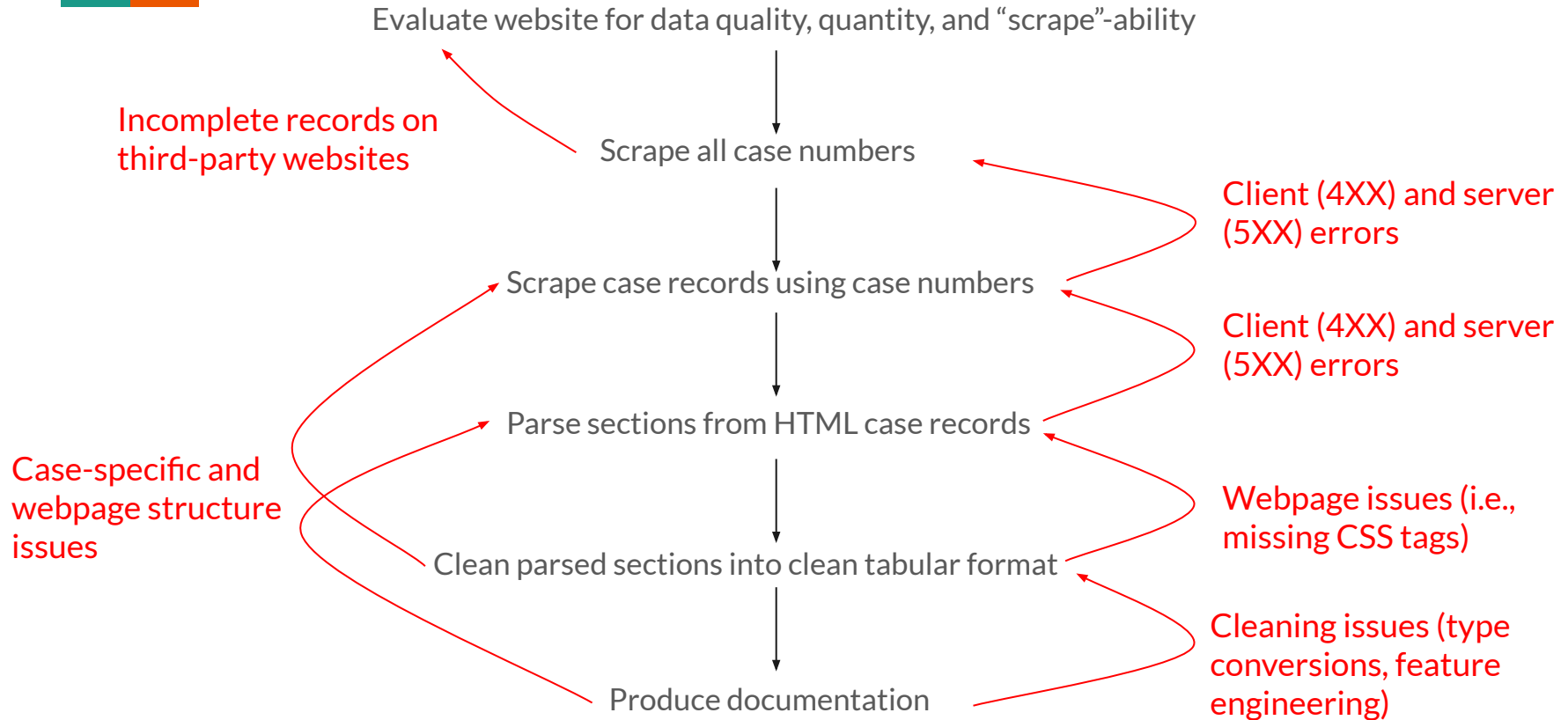
↓

Clean parsed sections into clean tabular format

↓

Produce documentation

# What my workflow actually is...



# Lessons Learned: Web scraping

---

- Be respectful
  - Monitor website response times and adjust your requests accordingly.
  - Plug: Urban Institute's [SiteMonitor](#) makes this easy and pain-free
- Be resilient
  - Find ways to “break” the website to understand error patterns: try big queries, small queries, clearly erroneous queries
  - Add explicit verification steps that confirm the expected inputs and outputs of scraping, parsing, and cleaning
- Be resourceful
  - Understand the network activity required for accessing web data.
  - Plug: [Chrome DevTools](#) provides methods for inspecting

# Lessons Learned: Creating datasets

---

- Be clear about the intention of your final dataset
  - Do you strive to preserve the structure of the raw data?
  - Do you need to standardize the data model across different sources?
- Understand how users will access the data
  - Will there be live connections to dashboards?
  - If data will need to be updated, what data formats makes this easy?

# I want to hear from you!



For questions, collaboration, and data sharing, please reach out:

Dan Bernstein

[bernsteind@lsc.gov](mailto:bernsteind@lsc.gov)

@danbernstein18