



A user-centered approach to developing an AI system analyzing U.S. federal court data

Rachel F. Adler^{1,2} · Andrew Paley¹ · Andong L. Li Zhao¹ ·
Harper Pack¹ · Sergio Servantez¹ · Adam R. Pah¹ · Kristian Hammond¹ ·
SCALES OKN Consortium¹

Accepted: 29 May 2022
© The Author(s) 2022

Abstract

We implemented a user-centered approach to the design of an artificial intelligence (AI) system that provides users with access to information about the workings of the United States federal court system regardless of their technical background. Presently, most of the records associated with the federal judiciary are provided through a federal system that does not support exploration aimed at discovering systematic patterns about court activities. In addition, many users lack the data analytical skills necessary to conduct their own analyses and convert data into information. We conducted interviews, observations, and surveys to uncover the needs of our users and discuss the development of an intuitive platform informed from these needs that makes it possible for legal scholars, lawyers, and journalists to discover answers to more advanced questions about the federal court system. We report on results from usability testing and discuss design implications for AI and law practitioners and researchers.

Keywords User-centered design · Notebook interface · Data analytics · Natural language processing · Visualization

1 Introduction

The judiciary carries the responsibility of interpreting laws and deciding our legal disputes, yet this system is far from transparent. In the United States, the federal court system handled approximately 425,000 cases in 2020 alone (U.S. Courts 2020), which are maintained on PACER (Public Access to Court Electronic

✉ Rachel F. Adler
r-adler@neiu.edu

¹ Northwestern University, Evanston, IL, USA

² Northeastern Illinois University, Chicago, IL, USA

Records), and available online to the public for a fee. PACER, though, has a non-intuitive user interface with differing standards for recording information across the 94 regional trial courts (Martin 2018; Alexander and Feizollahi 2020).

While many argue that there should be greater levels of government transparency, in practice data is often not well structured or easy to analyze (Weerakkody et al. 2017). Despite the need to conduct large-scale systematic analyses across multiple legal subject areas, legal research has not benefited from the methodological advances in big data and analytics as much as in other fields (Howe et al. 2008; González-Bailón 2013; Khoury and Ioannidis 2014; Carter et al. 2015; Obermeyer and Emanuel 2016).

Even if we had fully available and standardized data, domain experts, such as legal scholars, lawyers, and journalists, rarely have the technical expertise to conduct complex data analyses. There is a gap between open data and accessible information for all. Therefore, domain experts are either unable to answer more complicated questions or must carry the significant costs of hiring a data scientist to perform these tasks. There are many questions lawyers, legal scholars, or journalists may want to ask of the data, including understanding the ruling patterns of a particular judge or how general court trends have changed over time. While there are some systems out there, such as LexisNexis and Westlaw, which provide an easier to use interface with analytics capabilities, the cost of these systems are prohibitive to many users and are not intended to answer a broad range of questions of the data at a systemic level.

Furthermore, artificial intelligence (AI) solutions in law have gravitated towards the development of mostly algorithmic and computational solutions (Curtotti et al. 2015; Zhong et al. 2020; Di Porto 2021), with an emphasis on search engine interfaces (Sekaran et al. 2020) and automating processes for people who already know what they are doing, such as a data analyst (Wang et al. 2019). However, as we integrate AI into more human interactions, it is imperative to develop additional aspects of intelligence for AI systems that allows them to move from the automation of tasks for experts to the automation of expertise for non-technical users.

A number of advances across statistical and symbolic techniques under the big tent of AI have opened up a variety of new types of system architectures and user experiences. By building AI systems that know how and when to do the appropriate analyses to answer questions (e.g. determining the appropriate visualization and statistical techniques), we can democratize access to legal information to non-expert users. However, without including users in the process of AI systems, there is a risk to the user comprehension of that system (Holzinger 2013).

Therefore, we used a user-centered approach to inform the development of an AI system that provides information transparency in law. Rather than looking to automate processes for data analysts/technical experts, we look to automate the nature of their expertise (via an ontology-driven AI reasoning engine) and their interactions with the true end-users of the data science pipeline (by designing the AI system based on user-centered practices). In this paper we report on the process of surveying, interviewing, and observing legal scholars, lawyers, and journalists, focusing on their needs rather than solutions (Patnaik and Becker 1999), to discover what

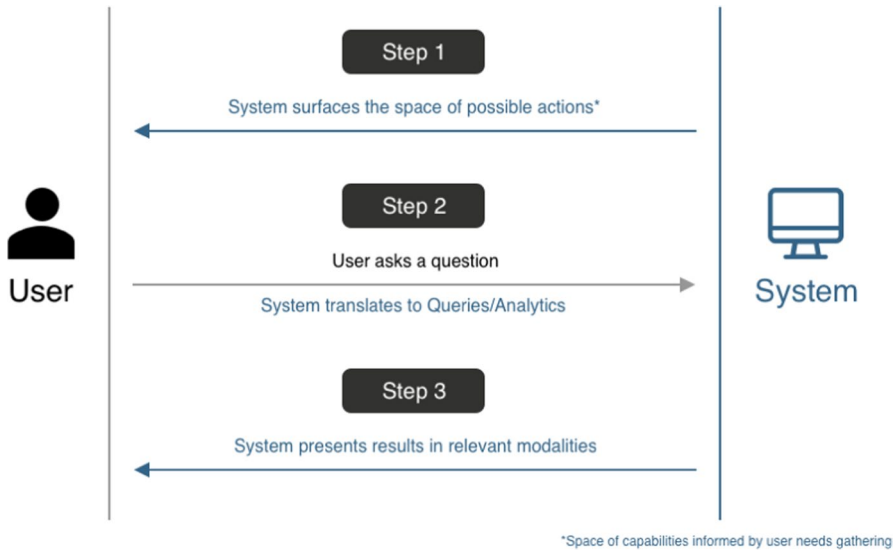


Fig. 1 A conceptual map of the division of responsibilities between user and system in running analysis and generating results

questions potential users would like to ask an AI legal system and map those questions to information needs.

We then describe the design of a system that translates domain expert questions on federal court records and information needs to queries and analytics and communicates the results back to the users in a clear and concise method (Fig. 1) and report on its usability. Given a simplified ontology of domain semantics mapped to data as well as an analytics ontology, the legal AI system understands how to generate relevant analysis plans (including SQL queries) and their natural language representations such that novice users can interact with the system without assistance from technical experts. Our ultimate goal is to put the burden on the machine, leaving users to simply assert “I’m interested in information about judges” and the system responds, “given the data I have on hand and what I know about judges, here’s what I can do for you.” Furthermore, as we design the system we are looking at scalability by increasing the number of questions that users can receive answers to (e.g. outcomes and motions) through classifying legal dockets using machine learning.

Our main research questions are as follows:

R1. What types of questions would non-technical legal scholars, lawyers, and journalists like to ask of the data?

R2. Can we design a system that permits non-technical legal experts to ask questions about the legal system that would normally require the skills of a data scientist to answer?

R3. How can we validate that our solution is working well?

2 User-centered design in AI and law

In order to design an AI system that efficiently utilizes federal court data to answer questions using appropriate analytical techniques and communicates those answers effectively, we wanted to understand the target users and their needs. Therefore, we incorporated user-centered design strategies.

User-centered design (Norman and Draper 1986) is the process of focusing on users and their needs throughout the stages of the design process. User-centered design has expanded to applications in numerous areas, such as health (Harte et al. 2017), education (Ebner and Holzinger 2007), business modeling (Arar et al. 2018), and journalism (Aitamurto et al. 2019). Studies using this method have included feedback from users, designers, and even community leaders (Iacobelli et al. 2018). In the book *Change by Design*, Tim Brown from the firm IDEO describes how design methods can be expanded to numerous areas such as the organizations themselves, hospitals, and universities (Brown 2019).

Few studies focus on using user-centered design in the legal realm to make legal applications usable for users, such as legal scholars, journalists, and the public. Jackson (2016) proposes that law schools teach human-centered design in addition to technology courses. This initiative is already happening in some universities (Hagan 2020). Hagan (2014) mentions that design thinking is particularly suitable for legal practitioners due to the potential to better solve problems, manage information, and provide more positive experiences overall for both lawyers and clients.

Quintanilla (2017) proposes a human-centered civil justice design to discover how people respond to features of the civil justice system. There is a push for court systems to have ‘a culture of usability testing and feedback’ to find user frustrations and implement improvements (Hagan 2018). User-centered, or citizen-centered, approaches can be applied to better include users in the development of online government services (Bertot et al. 2008). This can be done by including users during multiple phases of the development cycle through surveys, interviews, and usability testing (Jaeger 2008; Verdegem and Verleye 2009).

Understanding who your user is can help create legal systems that are efficient for them. For example, non-lawyers report having twice as many usability difficulties than lawyers when completing information retrieval tasks on legal databases (Newman and Doherty 2008). Non-lawyers have also reported wanting legal help sites to have clarity, open access, authority, comprehensiveness, modern design, and conversation (Hagan 2016). In order to make contracts readable to non-lawyers, studies have found that including visualizations of text-based contracts can improve participants’ answering speed and accuracy (Passera 2012; Passera and Haapio 2013; Passera et al. 2017). People benefit from systems that automatically present data visually (Mackinlay et al. 2007; Wongsuphasawat et al. 2017), however, it is important to choose appropriate principles for selecting visual representations that support the data and analytical tasks (Cook and Thomas 2005). For example, line graphs are a common method for depicting trends over time (Evergreen 2019). As more legal data become available for the public, data analytics will continue to be a common technique for analyzing legal judgements and laws (Park et al. 2021). Therefore, it is

important to determine ideal methods to display this data and legal analytics back to the users in a way that they can easily understand (Lettieri et al. 2018).

There have been recent advances in AI interfaces for automated analyses and visualizations, partly drawing from work in visual analytics. Visual analytics is a decision making approach that integrates visualization, human factors and data analysis through identifying the best automated analysis, appropriate visualization, and interaction techniques (Keim et al. 2008). Visual analytics can benefit from the inclusion of more user-centered approaches and evaluation (Scholtz 2017). Additionally, there has been recent work in integrating natural language interfaces with automated visualizations. For instance, one approach explores how understandable visualizations can be automatically generated through natural language queries supplied by the user, where the complexities of the natural language processing techniques used to parse and semantically understand the query are abstracted away from the user (Narechania et al. 2020).

There is a line of research arguing to include more user-centered concepts within AI systems (Augstein et al. 2020). Our research expands on this work by focusing on federal litigation data and creating an easy-to-use AI interface with a built-in system that integrates natural language techniques and makes analytical decisions based on large amounts of data and dynamically chooses appropriate visualizations to better help legal scholars, journalists, and lawyers understand and answer questions they otherwise would not be able to answer. To do this we include a user-centered design process to discover user needs and ensure we have a system that works for all our target users.

3 Generating user needs

3.1 Interviews

In order to learn more about how legal scholars and journalists typically find answers to questions regarding court records, we first conducted user interviews (both in person and over the phone). In-person interviews generally included observations first on how users used their current system for their legal research.

We conducted 28 sets of interviews on a total of 38 people (25 male and 13 female). Sixteen sets were in-person, with three out of the sixteen interviews containing three people, two containing two, and the remaining 11 being individual interviews. We recorded in-person semi-structured interviews at the participant's location with no compensation. To balance perspectives and time requirements, some junior and senior members of the same project were interviewed together. Twelve sets of interviews were over the phone with two interviews having two people on the call and the remaining ten phone interviews being individual interviews. For the interviews, our inclusion criteria were participants who could be potential users of the system, such as legal scholars, lawyers, and journalists. For example, a legal scholar may want to examine court data for correlations or trends over time, a journalist may wonder whether a specific attribute of a judge (age, race, etc.) would impact the outcome of a case, and a lawyer may want to research similar cases out

there or perhaps previous cases by a specific judge in order to impact their current cases. The individuals chosen either responded to a direct email solicitation or contacted us because of publicity on social media. In-person interviews were approximately an hour and phone interviews 30 min, as in-person interviews generally contained observations as well. Those interviewed included faculty in law, sociology, and economics; lawyers; reporters; a journalism student; and directors of civil and criminal justice centers and commercial legal companies.

Our interview guide had 4 open-ended questions such as “What questions would you want to ask with a resource like this?” to help us understand participant’s analytical questions. During our phone and in-person interviews, we asked all participants questions relating to the types of resources they currently used, what would be beneficial for them, and the types of research questions they would like to ask. During in-person interviews, we also conducted observations where we asked participants to show us the current tools they use for their research. Notes from each interview were examined in order to identify common themes.

Participants generally accessed and analyzed legal data through advanced Google searches, PACER, Westlaw, Caselaw Access Project, and Bloomberg Law. One participant reported using Python or STATA for analysis and another participant had a custom-developed application that was created for her own research purposes.

Common themes that emerged when analyzing interview results were in terms of wanting structured data and answers to questions. First, many participants mentioned having the federal data accessible is not enough as it is too overwhelming, needs to be structured, and linked with other datasets. In terms of answers to questions, participants felt they were limited by the tools they were currently using and wanted to ask questions of the data that they currently could not answer such as regarding outcome, motions, and finding cases by specific judges or judge demographic. For example, one participant reported wanting to find out information such as Plaintiffs argue X issue in Y motion and it is successful Z% of the time. Another participant mentioned wanting to be able to search in different ways based on topics, answers, and opinions (a written statement by a judge explaining a case decision and the underlying reasoning), and answer questions such as finding copyright cases that have an opinion. Another mentioned wanting visualizations of the data, such as timelines. One participant said he would only use a system that was similar to what he knew in terms of advanced searching. Only one participant mentioned not trusting others to do the data manipulations and wanting to do it himself.

3.2 Surveys

In order to reach a larger pool of participants, we sent a 35-question survey to our legal contacts and networks. While our interview targets were legal scholars, lawyers, and journalists, we targeted our survey through email and social media to anyone interested in court records. The survey was separate from the interview process described above, though since the survey is anonymous and was posted on our social media it is possible there is some overlap. Participants were not compensated and the mean duration was 5.6 min. Fifty one respondents completed the survey

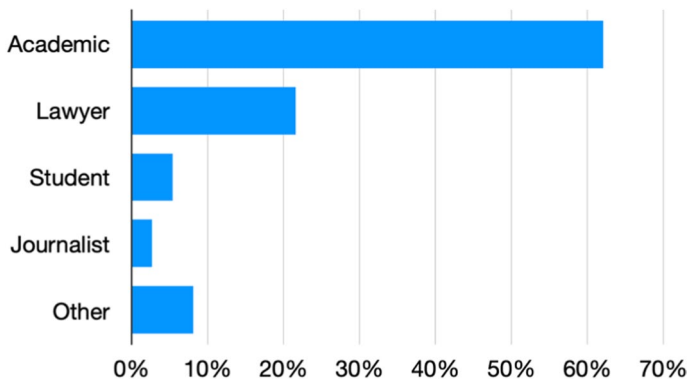


Fig. 2 Breakdown of survey participants by profession

Table 1 First four survey questions gauging respondents’ legal and technical skills

Q1 Which of the following best describes your level of legal understanding?
Legal expertise in multiple areas
I have some specialized knowledge about at least one area of law
I am generally familiar with how the US justice system works
I am not familiar with US law or the US justice system
Q2 Which of the following best describes your practical experience with litigation?
I am a civil and/or criminal litigator
I study or report on civil and/or criminal litigation
I have a general understanding of civil and/or criminal procedures
I am not familiar with litigation or the US justice system
Q3 Which of the following best describes your experience searching for information?
I have advanced coding skills and can search for and construct complex queries
I can conduct advanced searches to acquire data for my research
I am comfortable with keyword search using a search engine (e.g. Google, Bing, Westlaw, Lexis)
I am not comfortable with searching for information in the Internet
Q4 Which of the following best describes your experience analyzing data?
I can build programs (in R or Python for example) to extract information from data
I can use statistical tools (e.g., Tableau, Power BI, Stata, SPSS) to build visualizations
I am comfortable reading a quantitative graph, table, or chart
I am not comfortable with data analysis or representation

with 37 respondents providing occupational information. Of those respondents, 62% reported being academics with 17% of academics from law schools (Fig. 2). Examples of areas of legal expertise of our respondents included intellectual property law, immigration, race and ethnicity, criminal procedure, and health.

The goals of the survey were to determine user skills and research questions. We wanted to determine what questions users wanted to answer but currently were not able to, in order to ensure we incorporated the correct analytics into the system. Our survey had four main questions with response-based follow-ups, and choices to these questions ranged from not familiar/comfortable, generally familiar/comfortable, advanced, and expert (see Table 1).

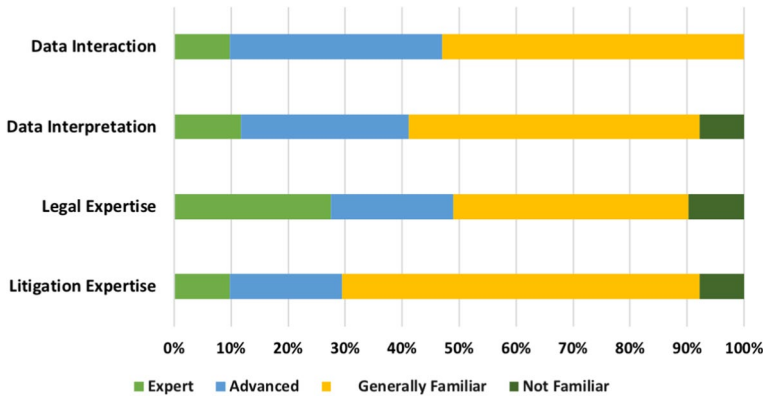


Fig. 3 Participants' self-ranked levels of data and legal expertise

Our survey showed that the majority of our respondents do not have the necessary technical skills to conduct complex data analysis. In terms of searching for information, we found that 53% of the respondents were only comfortable with general keyword searching, which we denote as Generally Familiar, with 37% conducting advanced searches (Advanced), and only 10% had advanced coding skills to search for and construct complex queries (Expert). See Fig. 3. Similarly, 51% of respondents were comfortable reading graphs (Generally Familiar), 30% were able to use statistical tools (Advanced), and only 12% were able to build programs to extract information from data (Expert). Note that 8% of the respondents reported not being comfortable with data analysis (Not Familiar).

Most of our potential users have a general understanding of the U.S. justice system. In terms of their legal knowledge, 10% were not familiar with the justice system (Not Familiar) while 41% were generally familiar with how the U.S. justice system works (Generally Familiar). The remaining 22% (Advanced) and 27% (Expert) had legal expertise in one or more areas respectively. With regard to their litigation expertise, 8% were not familiar with litigation (Not Familiar), 63% of the respondents had a general understanding of civil and/or criminal procedures (Generally Familiar), 20% studied litigation (Advanced), and 10% were litigators (Expert).

We conducted analysis across all 4 dimensions, mapping responses to ordinal values 0 to 3, where 0=Not Familiar and 3=Expert. Figure 4 shows a scatterplot of the survey results. For simplicity, we only include analysis across legal expertise and data analytical expertise which show that few participants had both advanced legal and data analytical skills and in fact the majority of participants had low analytical skills. Based on our results, we generated four types of personas:

- (1) **Basic Users- Low Legal and Technical Expertise (Lower Left Quadrant):** 13 participants (25.5%) had low legal and low technical skills. We call these users *Basic Users*. These users were at most generally familiar with the US justice system and comfortable reading a quantitative graph, table, or chart.

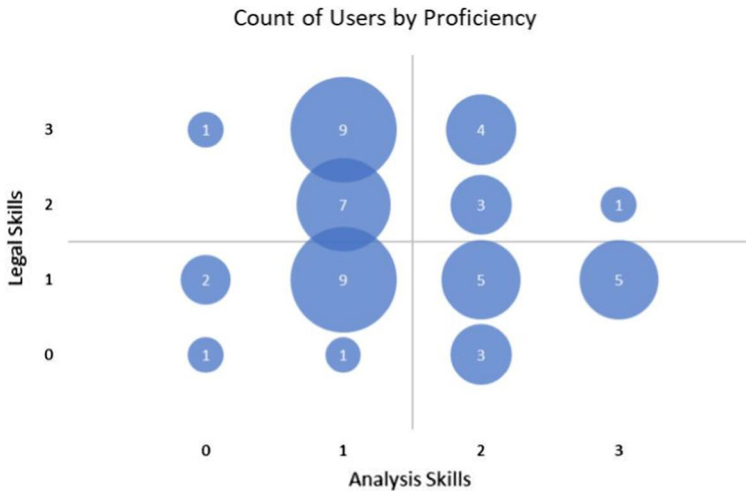


Fig. 4 Scatterplot of Participants' legal and data analytical skills

- (2) **Domain Experts-** Legal Expertise with Low Technical Skills (Upper Left Quadrant): 17 participants (33.3%) had high legal expertise with low technical skills. This is our largest quadrant and we therefore name them our *Domain Experts*. These users have the domain knowledge with legal expertise in one or more areas of law, but lack statistical or coding skills.
- (3) **Advanced Users-** Technical Experts with Low Legal Expertise (Lower Right Quadrant): 13 participants (25.5%) had high technical but low legal expertise. We name these our *Advanced Users*, however, we note that only five of them had coding skills. In other words, while they had statistical skills to build visualizations, few could build programs (in R or Python for example) to extract information from data.
- (4) **Legal Analysts-** High Legal and Technical Expertise (Upper Right Quadrant): 8 participants (15.7%) had both high legal and technical expertise. We call these users *Legal Analysts* as they have legal expertise in one or more areas of law in addition to advanced data interpretation skills such that they could use statistical tools to build visualizations or even code. The smallest number of participants fell into this quadrant, as the majority of our participants with legal expertise did not have advanced technical experience. In fact, only one participant in this category reported being able to program to extract information from data, with no participant reporting having a high level of Legal Expertise as well as coding experience (the high/high category).

In order to find out more about their experiences with data, we also asked them questions pertaining to what types of questions they were currently trying to answer, what obstacles they encountered, and what questions they would like to explore but are unable to. Some examples we found for what participants were currently asking with the data, included: "Who was judge? Who was refugee in asylum case (name

and other PII)?”; “Is there a relationship between characteristics of sentencing and the reentry process, including monetary sanctions and probationary requirements?”.

Sample responses for what questions users would like to ask but are currently not able to include:

“I would love to assess how predictive a person’s criminal record/history, especially juvenile history, is of future gun offenses. I would also love to know whether sending someone to prison (for a gun offense) has both an individual and general deterrence effect.”

“I don’t have specific questions, but as a journalist [I] would love simple access to legal cases. I have tried (and been confused by) PACER; I cannot afford and don’t have training in Lexis searches.”

In terms of challenges they experienced, many shared that they had limitations regarding their lack of data skills, some examples include: “Unexpected data format while parsing data. Personal knowledge limitations.” and “lack of skills.” Many of the survey respondents reported seeking assistance when they could not do the analysis themselves.

3.3 User needs

Based on the user interviews, observations, and surveys, below we discuss three user needs: Intuitive Search Interfaces; Answers to Research Questions; and Access to Raw and Clean Data.

Intuitive Search Interfaces: Many users mentioned wanting better searching options/filters to limit cases of interest. They wanted an easier-to-use user interface for searching. Some users mentioned wanting to have key cases appear, similar cases appear, and the ability to search by specific entities they were interested in such as a judge or location. In addition, they wanted to have graph visualization of this data.

Answers to Research Questions: Many of our users mentioned that they would benefit from a system that was able to answer more advanced research questions through data analysis that they were incapable of doing themselves. Furthermore, they wanted results to be in terms of visualizations to make it easier for them to obtain and view answers to questions they were unable to answer themselves.

Access to Raw and Clean Data: While the majority of our users were not comfortable with code, some participants with high data analytical skills wanted to have access to the raw data to perform their own analysis. While these participants mentioned not trusting analysis unless they did it themselves, they did want a system that would process and standardize the data in order to make it more consistently structured, for example, tracking cases across different courts.

The user feedback also led to the development of three main use cases for our system:

1. Users who want only a way to search and find relevant case documents, e.g. search fee waivers granted.

2. Users who want more targeted analyses, such as with cases regarding a particular judge, e.g. comparing if one judge takes longer than another with cases pertaining to Habeas Corpus.
3. Users who want more advanced analyses on the system, such as scrutinizing past trends, e.g. has the average number of cases per judge gone up in recent years for different districts.

While data scientists may be able to run these analyses themselves, it would take considerable time. An automated system would save time for those who know how to do their own data analysis and provide transparency and answers to questions for people who do not have a strong technical background.

4 SCALES OKN: a legal analytics platform

To address the needs described in the previous section, particularly pertaining to search and question-answering leveraging data analysis outside the capabilities of the users, we developed our own end-to-end platform consisting of an underlying AI engine and a natural language notebook. These together comprise an intricate AI platform that is able to understand users' intent and automatically determine the appropriate analytics given that intent and the context. We utilize a database of court records imported from PACER to conduct the legal analyses. The work discussed below was introduced in Paley et al. (2021).

4.1 AI system

The platform's AI system utilizes a domain ontology, an operation space ontology, and an inference engine. The domain ontology is a configuration that specifies semantic information about the domain and maps these to a standard SQL database schema. The operation ontology details the space of possible operations depending on the domain semantics. By leveraging both of these, the engine automatically generates available frontend mechanics (e.g. appropriate filters for different underlying data types) and their corresponding implementations (e.g. query builders). Furthermore, it pulls from those domain semantics and analytic ontology to derive all available and appropriate analytic plans, their corresponding natural-language representation for user interaction (e.g., *Average Case Duration Year-Over-Year grouped by District*), and their answer with appropriate visual representation (e.g. line graph, bar graph).

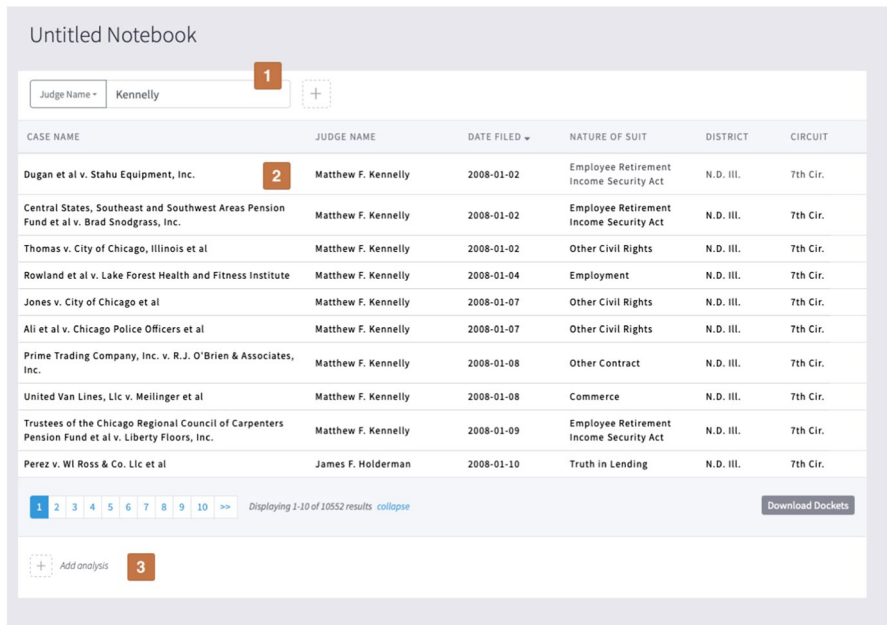


Fig. 5 Interface from the point of view of 1) Search (the ability to filter a set by entities and free text), 2) Context (the collection of cases in our focus, defined by the filters), and 3) Questions (the ability to ask questions about the context and elements within it)

4.2 Natural language notebook user experience

The core interface of our system is based on the ‘notebook’ platform pioneered by Mathematica and popularized by the Jupyter notebook project (Ragan-Kelley et al. 2014; Kluyver et al. 2016). In our previous work, we coined this the *natural language notebook* user experience, where users can drive data analysis and exploration without the need to code or write SQL.

In Fig. 5, we show our web-based interface along with a high-level depiction of the interactions between user and system across these views in Fig. 6. Users can interact with this system via filters or asking questions. The former allows users to segment the space, the latter allows users to invoke analysis even when they themselves don’t know what those analytics are. A user can add filters to narrow the displayed results (for example, specifying the district or name of the judge) and click on a row to view the original case document. If a user enters a keyword search (one of the filter options, ‘Docket Entries’), those keywords are highlighted directly in the docket.¹

Users can also do different kinds of analysis via basic and complex aggregations. Basic aggregations include questions such as determining the average duration of a

¹ Note that some functionality is still in the prototype phase. For example, the Download Dockets button on the bottom right of Fig. 5 is currently a placeholder and will be implemented in a future iteration.

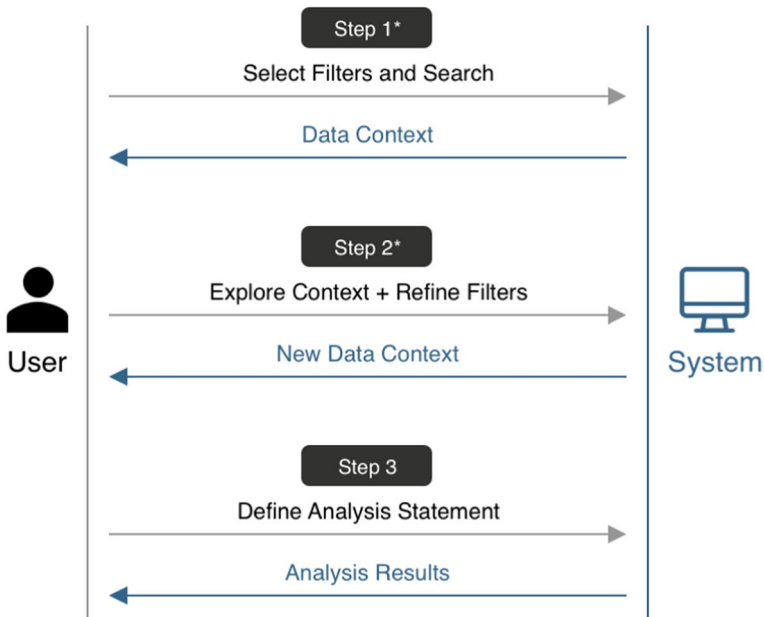


Fig. 6 A high-level user flow of notebook interactions, with Steps 1 and 2 being optional in support of driving analysis via Analysis Statements

case (Fig. 7), which can be refined by searching within it, such as for the Northern District of Illinois (N.D. IL) as each notebook's search context is connected to the analyses of that cell. Complex aggregations can include asking questions such as how the number of habeas corpus cases (legal actions which determine whether the detention of an individual is lawful) per capita varies by jurisdiction. Users can also do time-series analyses of the search results such as graphing how the number of immigration cases changes over time for a given period. Simple aggregations return answers in natural text whereas more complex ones return appropriate visualizations to the user.

Through these functionalities, users can do legal analysis such as looking at the evolution of case duration over time by looking at how it changes year over year (Fig. 8a) or the relationship between whether a party has legal representation or is self-represented ("pro se") and the likelihood that a request for a fee waiver will be granted (Fig. 8b). The analyses depicted are both run in the data context of cases in the Northern District of Illinois (per applying the district filter "N.D. IL"), but given the system's mechanics, these same analyses could be run against any filtered data context.

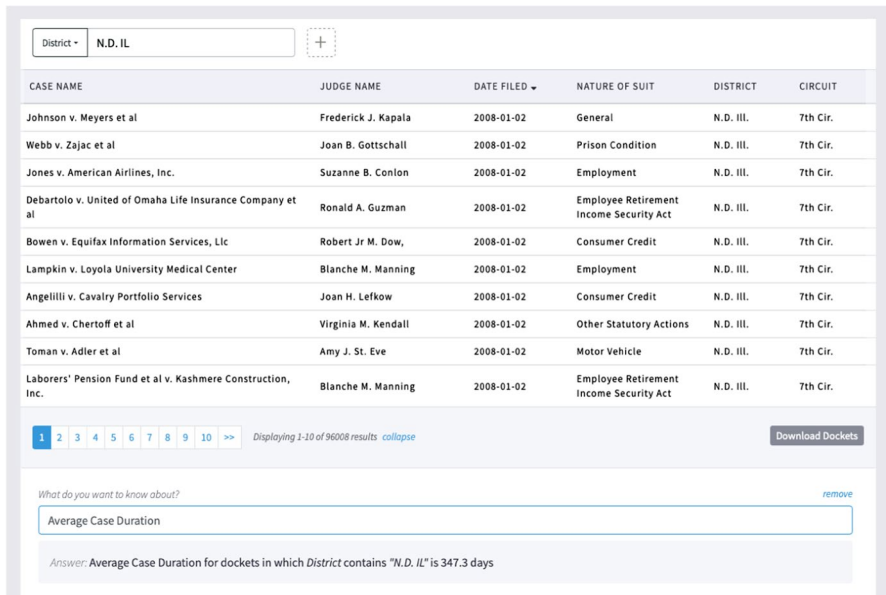


Fig. 7 Asking questions

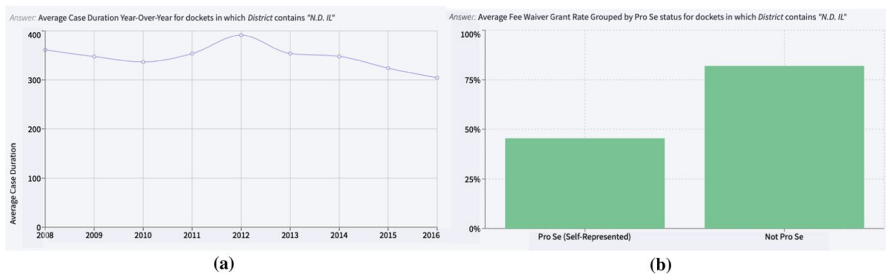


Fig. 8 Asking questions: **a** Case duration year-over-year **b** Fee waiver (a request to the court to waive court fees based on inability to pay) grant rate based on whether a party has legal representation or is self-represented ("pro se")

5 Results from usability testing

To test the viability of our design, participants were recruited for usability testing through our network of legal scholars, lawyers, and journalists. Usability testing included a total of 15 participants: 13 from law (lawyers, legal scholars, a legal assistant, and law librarian), a JD/PhD candidate in sociology and law, and a journalist. Seven participants overlapped in interviews and usability testing (surveys were anonymous so overlapping participants from surveys and usability testing is unknown). Each session was 45 min and was conducted over Zoom with two interviewers and one participant. Sessions were recorded for further analysis.

We intentionally introduced the UX without a tutorial first in order to determine how intuitive it would be for participants to learn without guidance, however we tested participants in five specific tasks which were in ascending order of difficulty to help them gradually become comfortable with the interface, these were: search, analysis returning a number, analysis returning a chart, analysis based on a filter, and analysis based on multiple filters. These include: “Conduct a search for all judges with *Kennelly* in their name. How many results appear?” “What is the average case duration?”, “Find the average case duration by year. Which year has the lowest in our dataset?”, “For all cases in the ‘N.D. IL’ district, which year had the highest average case duration?”, “For all habeas cases in the 11th circuit, how often on average are fee waiver requests granted?” We briefly provided help if needed. To measure the efficiency of the system, we examined the number of attempts participants took to answer each of the 5 scenarios. To measure the effectiveness of the system we examined the percent of completion of the tasks without help.

After concluding the scenarios, participants were given time to try out their own scenarios that they would be interested in searching for and conducting analysis on. We asked participants to *think aloud*, a method in usability testing (Lewis 1982; Nielsen 1994), in order for us to hear what they were trying to do and how they thought to do it.

All participants were presented with a survey to complete at the end of the session. We used the System Usability Scale (SUS), which is a 10 question tool using a 5-point Likert scale ranging from 1, strongly disagree, to 5, strongly agree for assessing a system’s usability (Brooke 1996). We used the modified SUS by Bangor et al. (2008) which uses the more recognizable word “awkward” instead of “cumbersome.” To calculate the SUS scores, 1 is subtracted from the raw score of the odd-numbered items (those items phrased in a positive way) and the raw score of the even-numbered items (those items phrased in a negative way) is subtracted from 5. The sum of the scores are multiplied by 2.5 to reach a “standardized SUS Score” which is out of 100.

Our goal during post-session analysis was to understand if the system was easy to use, would be used frequently, and potential improvements. Therefore, we also asked them to comment on features they liked, didn’t like, and questions they would have liked to ask the system.

5.1 Task completion

We found that participants were generally in three categories:

Category 1: Completed the scenario in 1 attempt without help.

Category 2: Completed the scenario in more than 1 attempt without help.

Category 3: Completed the scenario in more than 1 attempt with help.

Two coders calculated the number of participants who fell into each category for each scenario with a 93% agreement. When there were differences, we defaulted to the higher, more critical category. Out of the 75 instances (15 participants X 5 scenarios), 46 were in Category 1 (61%). In over a majority of cases, participants were

able to figure out where to go relatively quickly without any prior knowledge and without any help. 11 (15%) instances fell into Category 2, where participants did not need help and were able to complete the scenario on their own. In all except one of these cases, participants took two attempts to complete the task with the remaining participant completing it in 3 attempts. Last, 18 participants (24%) were in Category 3 and would not have been able to complete the task without guidance.

We found that Category 3 generally occurred for the following four reasons:

1. Filter Dropdown: Participants used the default ‘docket entries’ filter without selecting the appropriate filter.
2. Add Analysis: Participants did not realize there was an analysis component below without guidance.
3. Filter/Analysis: Participants did not realize the relationship between filtering and then analyzing those results.
4. Update Results Button: Participants had trouble remembering to press the update results button when starting a filter before beginning an analysis.

5.2 SUS results

We normalized our SUS scores using the formula above to obtain a value out of 100, where 68 is considered an average usability (Sauro and Lewis 2016) and good usability is at 71.4 (Bangor et al. 2009; Lewis 2018). Our participants’ average SUS score was computed as 72.83 which is, therefore, considered good usability. Features participants liked included both the searching and analysis capabilities. Some sample comments from the survey include:

“I liked the ability to search entire dockets and to pull up pre-defined analytics. I could see these features being very useful.”

“the ability to add analysis on top of a search that drilled down by lawyer etc. seems useful”

“Great variables (both filter and analysis). Loved the charts.”

“Very few steps required to get the data – instant results.”

“I liked that you could add multiple features and then ask a questions [sic] based on those results.”

Features participants did not like included comments on the user experience, missing features, and performance. Below we summarize their concerns in those three categories.

- UX flow comments:
 - The initial + sign for filtering was unclear.
 - Add pre-populated responses when using the filter.
 - The analysis menu was too long to scroll through.
 - The relationship between the top (search) and bottom (analysis) was unclear.

- Missing features:
 - They wanted an export option for their results.
- Performance-Related:
 - Full docket search was slow.

We also asked participants what questions they would like to answer regarding legal analyses in order to improve the analysis component in future design iterations. Some comments included analysis pertaining to outcomes, motions, unsealed dockets, average time to disposition of a motion, and how frequently cases are appealed or went to a jury trial in a particular district.

6 Discussion

Our main research question was whether we could successfully develop an AI system using user-centered design principles to allow academics, journalists, or the public—particularly those without technical or data analytical skills—to be able to ask questions of federal court data and receive answers from our system that abstracts away the underlying data analytics involved. Results from usability testing showed that participants liked the interface, found it to be usable, and felt they would use it frequently. Participants found it easy to navigate and even when needing help were able to apply that knowledge to the future tasks.

Our user interviews, observations, and surveys highlighted users' two main needs: (1) intuitive search interfaces and (2) answer to research questions. We chose a notebook format and focused its capabilities on search and analysis, abstracting away complex analytics to address the first need. The user's search choices specify an information context that determines which of their research questions can be answered, thus moving some cognitive load from the user to the platform. We addressed the second need by focusing on natural language queries that more directly link analysis statements with the questions that users have using an answer modality that is dynamic.

A unique aspect of our research design is being able to answer questions without manually conducting data analysis. Judicial Analytics is defined as “big data meets court dockets,” and referred to as “the next wave in legal research” (Bissett and Heinen 2017). Judicial analytics can lead to transparency into the work of judges as well as discovering potential biases (Chen 2019; McGill and Salyzyn 2021). Our goal supports this effort as we design an interface for asking legal questions and seamlessly obtaining analytical answers.

6.1 Limitations

There are limitations to our research, which include:

- The state of the available data: Our sample draws on more than a quarter-million court dockets acquired through purchase and batch downloading from PACER. As data ingestion itself is a work in progress, the version of the application we tested

with users was connected to a database with ten years of data (2007 through 2016) for the Northern District of Illinois court, but only one year of data (2016) for all 94 district courts across the U.S. Thus, the exploratory component of our tests was more engaging for users for whom the Northern District of Illinois was a subject of interest (and thus yielded more feedback), and less so for others.

- A U.S. data focus: We are eager to explore the possibility of bringing our approach to legal datasets sourced from other countries. In subsequent iterations, we will explore sourcing data from a second country and look to test user flows with relevant domain experts.
- Similarly, the lack of localization to support non-English language interactions: The majority of the language presented to users through our platform's UI is pulled from the configuration. Thus, the path towards localization is relatively easy to contemplate, but not yet implemented – our initial user tests were all in English and limited to English speakers. And so, in addition to exploring datasets from other legal systems, we also intend to test the capabilities and utility of our language-based approach in a second language in subsequent iterations.
- “Possible” vs “Domain Relevant”: Our overall model is in pursuit of leveraging configuration and data semantics to simplify and constrain the space of available analysis based on what's possible and what's domain relevant. In our current iteration, we have achieved the former – based on data types and relationships, we generate a set of analysis statement candidates that only includes those that can be run given the available data, ensuring that any user selection will successfully generate results. And while the majority of this set are also domain relevant, the ability to guarantee true relevance in a given domain – ensuring that any analysis statement is guaranteed to map to a domain-relevant question – is a matter of ongoing design and future iterations of our configuration will support mechanisms by which technical users can further control the space of analysis. For example, we compute *judge tenure*, a numeric value that represents a given judge's time on the bench. Because our system recognizes judge tenure as numeric, it knows it can support users asking for the average judge tenure, maximum judge tenure, and minimum judge tenure. Further, it supports exploration of other metrics grouped by judge tenure (via statements like “Average Case Duration grouped by Judge Tenure”). However, based on the knowledge that judge tenure is a numeric value, the system also provides access to *total judge tenure*, a sum of all derived judge tenures associated with all cases in the data view. No reasonable exploration of the domain would include such analysis despite the fact that it is supported at the data level.

6.2 Future work

Based on results from the usability testing, we plan to make modifications to our user interface to improve the user experience in terms of Search and Analysis, Transparency, and Scalability.

Search and Analysis: In terms of the search features, we plan to integrate auto-complete in the filter textbox and show all possible options that are available. For example, when entering a nature of suit value, users will be able to see all the

possibilities instead of having to remember the options or look them up. Similar to the “Add Analysis” text near the analysis filters, we plan to include “Add a filter” message near the search filter, as some participants had some trouble finding the search box. Other changes include not having “docket entries” filter as the default option, improving the speed of free text search by implementing elasticsearch, including clarifying text making the connection between search and analysis more apparent, and allowing users to perform the analyses, using the information selected in the above filters, regardless of whether a user pressed the Update Results button.

We also plan to collapse the search view initially, so that the analysis is higher up and easier to find. While users want to have the option to verify the data that results from the search, some users did not notice the analysis option on the bottom until we pointed it out. Further, we plan to reorganize the analysis dropdown by categories and make it more obvious that they could type text in addition to scrolling through options.

A few of the more technically savvy users requested access to the data in order to conduct further analyses. These users mentioned not trusting a system. Research has shown that with the expansion of big data and visual analytics, users have uncertainties and a lack of trust in artificial intelligence and visual analytics (Sacha et al. 2015; Siau and Wang 2018). Therefore, after assisting the user in filtering the court records using a dynamic set of search parameters, we plan to allow the user to (1) take a snapshot of the filtered docket set using a download button and (2) export to csv their analysis results. While we are implementing the download docket and export to csv buttons as an incremental addition, our goal is to discover what they plan to do with that data in order to further refine our system in the future. This is an iterative user-centered design process which will lead to long-term improvements to the user experience of the platform.

Transparency: We plan to include more transparency on the completion of the dataset, how we calculated fields (e.g. case duration, judge tenure), and clarification on entities displayed, for example explaining which judge we display on cases with multiple judges.

Scalability: We plan to incorporate additional dockets and other datasets in order to answer more types of questions. In addition, we will keep track of users’ search history, include the ability to add collaborators to a Notebook, and create a summarization of the analysis. Addressing summarization in the future will allow us to better meet the needs of our identified user personas who lack domain knowledge. Furthermore, we created a separate web-based tagging tool where lawyers and law students tagged motions in dockets and classified them using machine learning. Next, we plan to tag outcomes as well. We also plan to create a similar generic tagging tool, which will allow users to tag any item of interest in the docket entries, such as orders, notices and affidavits, and use machine learning to recognize these tagged items so that they can also be included as searchable items in the future. Future work also involves abstracting the system so that it can be used for data analysis when given datasets in other domains, such as education.

6.3 Implications for designing an AI system

Through the process of designing the SCALES Open Knowledge Network, we identified some key design implications for user-friendly AI systems.

Data analytics for non-technical users requires a different level of abstraction: Computer science is well known for leveraging different levels of abstraction for types of users and applications (Te'eni and Sani-Kuperberg 2005; Te'eni 2017). However, as we abstract away some of the details, we lose some of the expressiveness; for instance, programming in a higher-level language gives users less control than coding in assembly. Similarly, our system abstracts away various steps of the analytical pipeline (e.g. data parsing and processing, visualization generation, available operations). However, our abstraction level is still more than sufficient to meet users' needs and answer their questions, since all but two users (87%) said that they would use our system frequently (one of them was neutral and the other did not use docket sheets in their research). Furthermore, our design allowed many non-technical users to do analysis that was unavailable to them in platforms that had a larger space of possible analyses due to higher barriers to entry. This shows that there are more conducive levels of abstraction for data analytics that allow a wider set of users to use these interfaces.

Building trust in AI is also a design problem, not just a technical one: As we move to integrate AI more into different tools, we need to take deliberate design considerations to build trust between users and AI platforms. The issue of building trust for AI has been explored in other work (Sacha et al. 2015; Siau and Wang 2018) and we saw it be a significant issue in our design validation. For instance, many users were more confident in using the system due to the capability of our system to show the original docket sheet data. Some users wanted even more information, with questions about how the data was processed, when the data was last updated, and the scope of the data (e.g. from which jurisdictions did we have data). And there were even some technically savvy users who did not trust the system at all and would have preferred to do all the analyses themselves. This indicates that building trust between users and AI is not merely a matter of model performance; rather, it will require design considerations that meet users' needs for transparency.

Integrating a user-centered approach in developing an AI system can help with usability: As more AI systems get developed, it is important to keep in mind the users of the system by keeping them involved throughout the development lifecycle. By engaging with our users, both in the early phases (interviews and surveys) and in the later stages (usability testing) we were able to keep on top of what users' needs were and what they still needed from our system. AI should expand beyond the computational domain and integrate more human approaches through the inclusion of research teams composed of both AI and UX researchers (Margetis et al. 2021). Designing AI systems with a user perspective can help reduce potential biases in the system and algorithms and make them more transparent, which can be particularly useful in designing legal systems (Augstein et al. 2020).

7 Conclusion

Based on user-centered design research, we interviewed, observed, and surveyed potential users and designed an AI system that provides nontechnical users with a method for searching and finding answers to questions on federal court data. Legal scholars, journalists, policymakers, and social scientists can use the system to answer questions they have about the U.S. legal system even if they lack data analytics skills. Future research will expand on the interface to include additional datasets and custom tagging of data.

Acknowledgements This material is based upon work supported by the National Science Foundation Convergence Accelerator Program under grant no. 1937123 and grant no. 2033604.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aitamurto T, Ananny M, Anderson CW, Birnbaum L, Diakopoulos N, Hanson M, Hullman J, Ritchie N (2019) HCI for accurate, impartial and transparent journalism: Challenges and solutions. In: Extended abstracts of the 2019 CHI conference on human factors in computing systems. pp 1–8
- Alexander CS, Feizollahi MJ (2020) On dragons, caves, teeth, and claws: legal analytics and the problem of court data access. In: Computational legend studies, Promises challenges data-driven res 95
- Arar R, Ren G, Jiang S, Becker V, Huang L, Liu EY (2018) Applying user-centered design to business modeling: CBM.next as a case study. In: 2018 IEEE 20th conference on business informatics (CBI), pp 164–169
- Augstein M, Buschek D, Herder E, Loepp B, Yigitbas E, Ziegler J (2020) UCAI 2020: workshop on user-centered artificial intelligence. Mensch und comput 2020-workshopband
- Bangor A, Kortum PT, Miller JT (2008) An empirical evaluation of the system usability scale. *Int J Hum Comput Interact* 24:574–594
- Bangor A, Kortum P, Miller J (2009) Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud* 4:114–123
- Bertot JC, Jaeger PT, McClure CR (2008) Citizen-centered e-government services: benefits, costs, and research needs. In: Proceedings of the 2008 international conference on digital government research. pp 137–142
- Bissett J, Heinen M (2017) Beyond court decisions—dockets, documents, and analytics. *Colo Law* 46:59
- Brooke J (1996) SUS-A quick and dirty usability scale. In: Usability evaluation in industry, pp 189–194
- Brown T (2019) Change by design: how design thinking transforms organizations and inspires innovation. HarperBusiness
- Carter DR, Asencio R, Wax A, DeChurch LA, Contractor NS (2015) Little teams, big data: big data provides new opportunities for teams theory. *Ind Organ Psychol* 8:550–555. <https://doi.org/10.1017/iop.2015.82>
- Chen DL (2019) Judicial analytics and the great transformation of American Law. *Artif Intell Law* 27:15–42

- Cook KA, Thomas JJ (2005) Illuminating the path: the research and development agenda for visual analytics
- Courts US (2020) Federal Judicial Caseload Statistics 2020. <https://www.uscourts.gov/statistics-reports/federal-judicial-caseload-statistics-2020>. Accessed 14 Dec 2021
- Curtotti M, McCreath E, Bruce T, Frug S, Weibel W, Ceynowa N (2015) Machine learning for readability of legislative sentences. In: Proceedings of the 15th International Conference on Artificial Intelligence and Law, pp 53–62
- Di Porto F (2021) Algorithmic disclosure rules. *Artif Intell Law*. <https://doi.org/10.1007/s10506-021-09302-7>
- Ebner M, Holzinger A (2007) Successful implementation of user-centered game based learning in higher education: an example from civil engineering. *Comput Educ* 49:873–890
- Evergreen SDH (2019) Effective data visualization: the right chart for the right data. SAGE publications, Thousand Oaks, CA
- González-Bailón S (2013) Social science in the era of big data. *Policy Internet* 5:147–160
- Hagan M (2016) The user experience of the internet as a legal help service: defining standards for the next generation of user-friendly online legal services. *Virginia J Law Technol* 20:394–465
- Hagan MD (2018) A human-centered design approach to access to justice: generating new prototypes and hypotheses for intervention to make courts user-friendly. *Indiana J Law Soc Equal* 6:199–239
- Hagan M (2020) Design comes to the law school. In: Denvir C (ed) *Modernising legal education*. Cambridge University Press, Cambridge, England, pp 109–125
- Hagan M (2014) Design thinking and law: a perfect match. *Law Pract Today*
- Harte R, Glynn L, Rodríguez-Molinero A, Baker PMA, Scharf T, Quinlan LR, O’Laighin G (2017) A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. *JMIR Hum Factors* 4:e8
- Holzinger A (2013) Human-computer interaction and knowledge discovery (hci-kdd): what is the benefit of bringing those two fields to work together? In: International conference on availability, reliability, and security, pp 319–328
- Howe D, Costanzo M, Fey P, Gojobori T, Hannick L, Hide W, Hill D, Kania R, Schaeffer M, St Pierre S, Twigger S, White O, Rhee S (2008) Big data: The future of biocuration. *Nature* 455:47–50
- Iacobelli F, Adler RF, Buitrago D, Buscemi J, Corden ME, Perez-Tamayo A, Penedo FJ, Rodriguez M, Yanez BR (2018) Designing an mHealth application to bridge health disparities in Latina breast cancer survivors: a community-supported design approach. *Des Heal* 2:58–76. <https://doi.org/10.1080/24735132.2018.1452871>
- Jackson D (2016) Human-centered legal tech: integrating design in legal education. *Law Teach* 50:82–97. <https://doi.org/10.1080/03069400.2016.1146468>
- Jaeger PT (2008) User-centered policy evaluations of section 508 of the rehabilitation act: Evaluating e-government web sites for accessibility for persons with disabilities. *J Disabil Policy Stud* 19:24–33
- Keim D, Andrienko G, Fekete J-D, Görg C, Kohlhammer J, Melançon G (2008) Visual analytics: Definition, process, and challenges. *Information visualization*. Springer, Berlin, Heidelberg, pp 154–175
- Khoury MJ, Ioannidis JPA (2014) Big data meets public health. *Science* 346:1054–1055. <https://doi.org/10.1126/science.aaa2709>
- Kluyver T, Ragan-Kelley B, Pérez F, Granger BE, Bussonnier M, Frederic J, Kelley K, Hamrick JB, Grout J, Corlay S (2016) Jupyter Notebooks—a publishing format for reproducible computational workflows. In: *ELPUB*, pp 87–90
- Lettieri N, Guarino A, Malandrino D (2018) E-science and the law. three experimental platforms for legal analytics. *Legal knowledge and information systems*. IOS Press, Amsterdam, pp 71–80
- Lewis C (1982) Using the "thinking-aloud" method in cognitive interface design. IBM TJ Watson Research Center Yorktown Heights, NY
- Lewis JR (2018) The system usability scale: past, present, and future. *Int J Hum Comput Interact* 34:577–590
- Mackinlay J, Hanrahan P, Stolte C (2007) Show me: automatic presentation for visual analysis. *IEEE Trans vis Comput Graph* 13:1137–1144
- Margetis G, Ntoa S, Antona M, Stephanidis C (2021) Human-centered design of artificial intelligence. *Handb human factors ergonomics*, pp. 1085–1106

- Martin PW (2018) District court opinions that remain hidden despite a longstanding congressional mandate of transparency – the result of judicial autonomy and systemic indifference. *Cornell Leg Stud Res Pap* No 17–38
- McGill J, Salyzyn A (2021) Judging by the numbers: judicial analytics, the justice system and its stakeholders. *Dalhousie LJ* 44:249–284
- Narechania A, Srinivasan A, Stasko J (2020) NL4DV: a toolkit for generating analytic specifications for data visualization from natural language queries. *IEEE Trans vis Comput Graph* 27:369–379
- Newman DR, Doherty U (2008) Making the law accessible to non-lawyers: effects of different kinds of expertise on perceived usability of online legal information services. *Behav Inf Technol* 27:423–437. <https://doi.org/10.1080/01449290601111002>
- Nielsen J (1994) Usability engineering. Morgan Kaufmann
- Norman DA, Draper SW (1986) User centered system design. Lawrence Erlbaum, Hillsdale, N.J., USA
- Obermeyer Z, Emanuel EJ (2016) Predicting the Future—Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med* 375:1216–1219. <https://doi.org/10.1056/NEJMp1606181>
- Paley A, Zhao ALL, Pack H, Servantez S, Adler RF, Sterbentz M, Pah A, Schwartz D, Barrie C, Einarsson A, Hammond K (2021) From data to information: automating data science to explore the US court system. In: *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*, pp 119–128
- Park S-H, Lee D-G, Park J-S, Kim J-W (2021) A survey of research on data analytics-based legal tech. *Sustainability* 13:8085
- Passera S, Haapio H (2013) Transforming contracts from legal rules to user-centered communication tools: a human-information interaction challenge. *Commun Des Q Rev* 1:38–45. <https://doi.org/10.1145/2466489.2466498>
- Passera S, Kankaanranta A, Louhiala-Salminen L (2017) Diagrams in contracts: fostering understanding in global business communication. *IEEE Trans Prof Commun* 60:118–146
- Passera S (2012) Enhancing contract usability and user experience through visualization - an experimental evaluation. In: *2012 16th International conference on information visualisation*, pp 376–382
- Patnaik D, Becker R (1999) Needfinding: the why and how of uncovering people’s needs. *Des Manag J Former Ser* 10:37–43
- Quintanilla VD (2017) Human-centered civil justice design. *Penn State Law Rev* 121:745–806
- Ragan-Kelley M, Perez F, Granger B, Kluyver T, Ivanov P, Frederic J, Bussonnier M (2014) The Jupyter/IPython architecture: a unified view of computational research, from interactive exploration to communication and publication. In: *AGU fall meeting abstracts*. pp H44D-07
- Sacha D, Senaratne H, Kwon BC, Ellis G, Keim DA (2015) The role of uncertainty, awareness, and trust in visual analytics. *IEEE Trans vis Comput Graph* 22:240–249
- Sauro J, Lewis JR (2016) Quantifying the user experience: practical statistics for user research. Morgan Kaufmann, Cambridge, MA
- Scholtz J (2017) User-centered evaluation of visual analytics. *Synth Lect vis* 5:i–71
- Sekaran K, Chandana P, Jeny JRV, Meqdad MN, Kadry S (2020) Design of optimal search engine using text summarization through artificial intelligence techniques. *Telkomnika* 18:1268–1274
- Siau K, Wang W (2018) Building trust in artificial intelligence, machine learning, and robotics. *Cut Bus Technol J* 31:47–53
- Te’eni D (2017) the strange absence of abstraction levels in designing HCI. In: *International conference on group decision and negotiation*. pp 15–29
- Te’eni D, Sani-Kuperberg Z (2005) Levels of abstraction in designs of human–computer interaction: the case of e-mail. *Comput Human Behav* 21:817–830
- Verdegem P, Verleye G (2009) User-centered E-Government in practice: a comprehensive model for measuring user satisfaction. *Gov Inf Q* 26:487–497
- Wang D, Weisz JD, Muller M, Ram P, Geyer W, Dugan C, Tausczik Y, Samulowitz H, Gray A (2019) Human-ai collaboration in data science: exploring data scientists’ perceptions of automated AI. *Proc ACM Hum Comput Interact* 3:1–24
- Weerakkody V, Irani Z, Kapoor K, Sivarajah U, Dwivedi YK (2017) Open data and its usability: an empirical view from the Citizen’s perspective. *Inf Syst Front* 19:285–300. <https://doi.org/10.1007/s10796-016-9679-1>

- Wongsuphasawat K, Qu Z, Moritz D, Chang R, Ouk F, Anand A, Mackinlay J, Howe B, Heer J (2017) Voyager 2: augmenting visual analysis with partial view specifications. In: Proceedings of the 2017 CHI conference on human factors in computing systems, pp 2648–2659
- Zhong H, Xiao C, Tu C, Zhang T, Liu Z, Sun M (2020) How does NLP benefit legal system: a summary of legal artificial intelligence. Preprint at <https://arxiv.org/abs/2004.12158>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.