

# AI & Access to Justice



**Margaret Hagan**  
**Stanford Legal Design Lab**  
**August 2024**

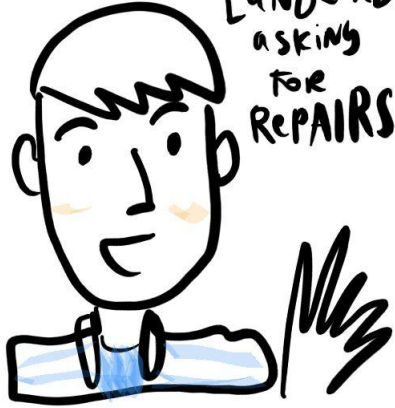
# **Can AI improve legal services you provide?**

**To improve the efficiency &  
quality for the public, and  
for your teams?**

# **Can AI help close the justice gap?**

**And how specifically can  
our justice community  
advance responsible AI for  
civil justice?**

I tried chatGPT  
& it gave me  
a great Letter to  
send to a  
LANDLORD  
asking  
FOR  
REPAIRS



TECHNO-  
OPTIMISTS

These systems  
are GIBBERISH machines  
that spit out fake  
Laws &  
FOOL  
PEOPLE



TECHNO-  
PESSIMISTS



TECHNO-  
REALISM

Are PEOPLE GOING  
TO BE USING AI  
FOR LEGAL PROBLEM-  
SOLVING?

How WELL ARE  
the DIFFERENT MODELS  
PERFORMING FOR the  
TASKS PEOPLE & PROVIDERS  
WANT TO DO?

where can AI increase  
quality & efficiency,  
of legal help services?

# AI & A2J Research we've been doing

**1: User Research:** Do American adults want to use AI for legal problem-solving? How are they likely to use it? What is valuable or not in the AI experience?

**2: Tasks Research:** What are the specific tasks AI can perform in A2J? Which do justice professionals prioritize for AI R&D?

**3: Quality Metrics Research:** How can we define whether AI has done a good/bad job at a specific A2J task, in a way that corresponds to what users & experts know matters?

## Presentation Takeaways on **AI + A2J**

- More American adults are likely to use **'brand-name' AI tools from large tech companies** to deal with life/legal problems in the coming years.
- Justice professionals are open to **building specific AI tools**, with more priority to some legal tasks over others.
- Regulation & policy-making around AI in the justice sector should be based on **community-defined quality benchmarks & empirical audits**, rather than anecdote or speculation.

**1**

# **User Research**



# 1

## **User Research into AI for Legal Problem Solving**

Do American adults want to use AI for legal problem-solving?

How are they likely to use it?

What is valuable or not in the AI experience?

Participant	Age	Education	Income	Profession	Race/Ethnicity	Sex	Disability?	Language
1	26-34	Professional Degree	100,000+	Unemployed	Asian	Male	No	English
2	55-64	Masters Degree	100,000+	Management Consultant	White	Female	No	English
3	35-54	4-year college degree	100,000+	Options Trader	White	Male	No	English
4	18-25	4-year college degree	100,000+	Masters Student	Asian	Male	No	English
5	35-54	4-year college degree	90,000-99,999	Payroll clerk	Spanish-Hispanic-Latino	Female	No	English
6	35-54	4-year college degree	100,000+	Executive assistant	White	Female	No	English
7	35-54	Some College	60,000-69,999	Administrative Assistant	Asian	Female	Yes	English
8	35-54	4-year college degree	100,000+	Sales manager	White	Male	No	English
9	26-34	Some College	50,000-59,999	HR generalist	Black-African-American	Female	No	English

**We have spoken with 45 adults** who responded to an ad on Facebook or Craigslist, asking if they would participate in a online research session about how they use the Internet to solve problems. Or who we recruited at a state court help center lobby.

**We ran 30-60 minute Zoom sessions** in which we:

- Asked them about how they generally use the Internet to solve life problems, and their familiarity with law
- Gave them an eviction notice scenario, and asked them to use Bard/Gemini/ChatGPT to respond

**A few people are so suspicious of AI they won't use it. But most participants were open to trying AI.**

I'm on the fence on it. I want it to know and understand me, my thought process. But then I don't want it to know all of that. It's like going down a rabbit hole. You want things to be easier, but then things have to understand you, and you're letting them into your personal life or mind. ... There's no middle ground. It has to record the data and store it. Who is the big brother behind the data watching it? Can they use it to affect my life in some other way?.. How might it bite you in the ass later on? (P15)

All the things I'm reading about it are terrifying. I don't want to start creating my own personal trail of use on AI. There's a lot of history of tech companies over-harvesting data. I might think I am just sharing one piece of data, but then I'm authorizing it to access all kinds of my data. I'm not willing to go there at this stage. (P2)

# How do **Novice Users** prompt AI to respond to an **eviction notice?**

Search engine  
query specter

Tenant rights (P15)

Landlord issues (P14)

Evicted by landlord (P10)

P13:

received an eviction notice what can I do?

california state eviction law

Alameda Eviction laws

Alameda relocation assistance

Legal help for eviction notice

How do you apply for the relocation assistance for alameda county

**How do  
Experienced  
Users prompt AI  
to respond to an  
eviction notice?**  
Using specifics &  
making a request

I've an official legal document from my landlord that I have been evicted. What legal recourse do I have? (P12)

I would like to get in touch with a legal aid organization to assist me with what I believe is an illegal eviction (P12)

Can my landlord evict me in 2 weeks california? (P5)

Where do I go to file against an eviction notice in Maryland (P9)

# What was AI showing to the participants?

**An empathetic first sentence**, like “I’m sorry that you are dealing with this,” or “Problems with your landlords can be stressful.”

**A succinct answer or statement** to the question or situation that was posed in the prompt. In some cases, this was a statement of what the law says about how long a landlord has to give a tenant with a warning notice, what the steps of an eviction are, or what actions a tenant can take if worried about an eviction. This often is presented in a bullet point or ordered list.

**Recommendation to speak with to experts** like lawyers or government agencies to find out more details, and explore exactly what the best option might be.

**Concluding list of organizations** that could have websites to read more, or services to help the person. These are not hyperlinked, but presented as a list of names.

# What value did participants see in AI when dealing with an eviction notice?

Simple, direct, well-formatted information

It's pretty easy and straightforward. I didn't know it could do stuff like this. It's crazy, compared to where things used to be. It was extremely helpful. It gave me exactly the answers I was looking for. It understood it from the get-go. (P13)

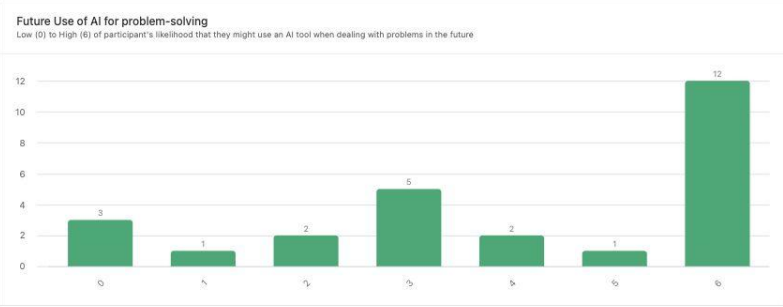
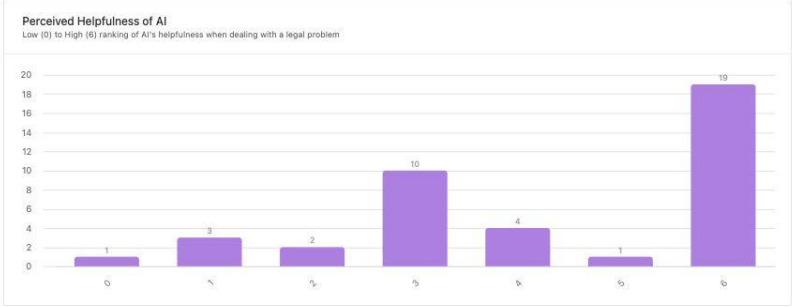
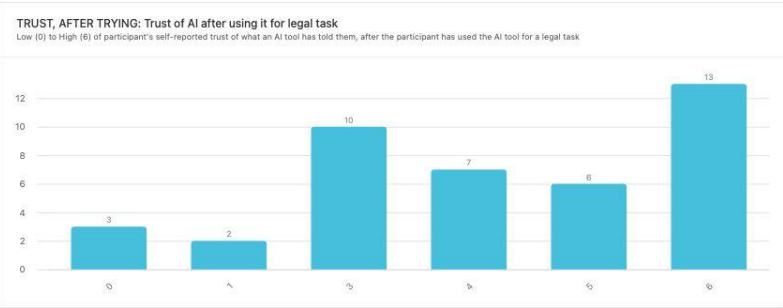
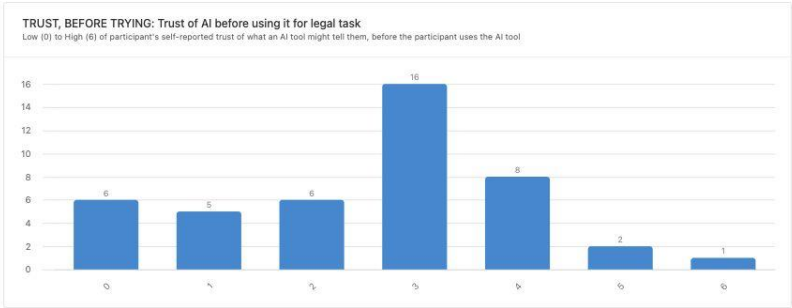
The results are more simple. Google [Search] can be very overwhelming with results. Compared to what I'm used to, it's a lot more digestible and easier to read. (P5)

It was extremely helpful because it gave me a better understanding prior to asking the question. Prior to this, I didn't know anything about the topic, but now I know -- 2 weeks could be possible, my landlord might have a right, and I need to search further to know what I should do for this eviction. (P14)

I like the way it looks, it's very easy to read, it's very cut and dry. Very direct. It's good advice, because it takes a long time to figure out about eviction. Legal aid sites come up. This is awesome. This is the stuff that comes up when I Google, what I've seen before. But this is awesome that it's all right here on one page. And the way that it's not cluttered, Google has this, this this, it's so overwhelming and cluttered. This is all that you need in one step. This would help a lot. It was

### Trust & Value of AI to Participants

We asked people at the beginning of the interview how much they would trust what AI would tell them for a legal problem.  
 We asked them the same question after they tried out an AI tool for a fictional legal problem of getting an eviction notice from their landlord.  
 We also asked them how helpful the AI was in dealing with the fictional problem, and how likely they would be to use this in the future for similar problems.





## Trust & Value of AI to Participants

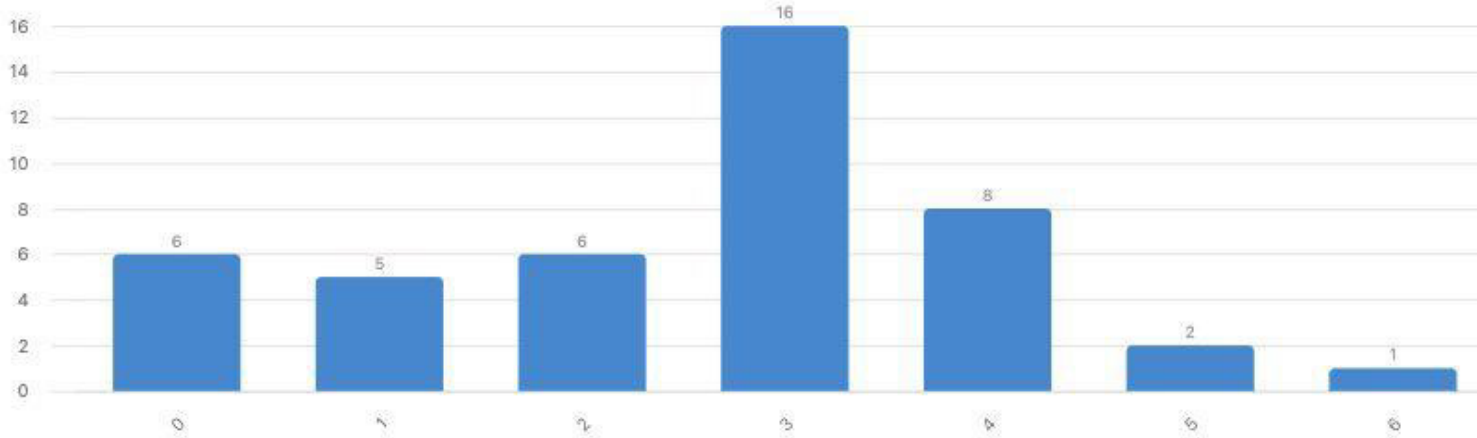
We asked people at the beginning of the interview how much they would trust what AI would tell them for a legal problem.

We asked them the same question after they tried out an AI tool for a fictional legal problem of getting an eviction notice from their landlord.

We also asked them how helpful the AI was in dealing with the fictional problem, and how likely they would be to use this in the future for similar problems.

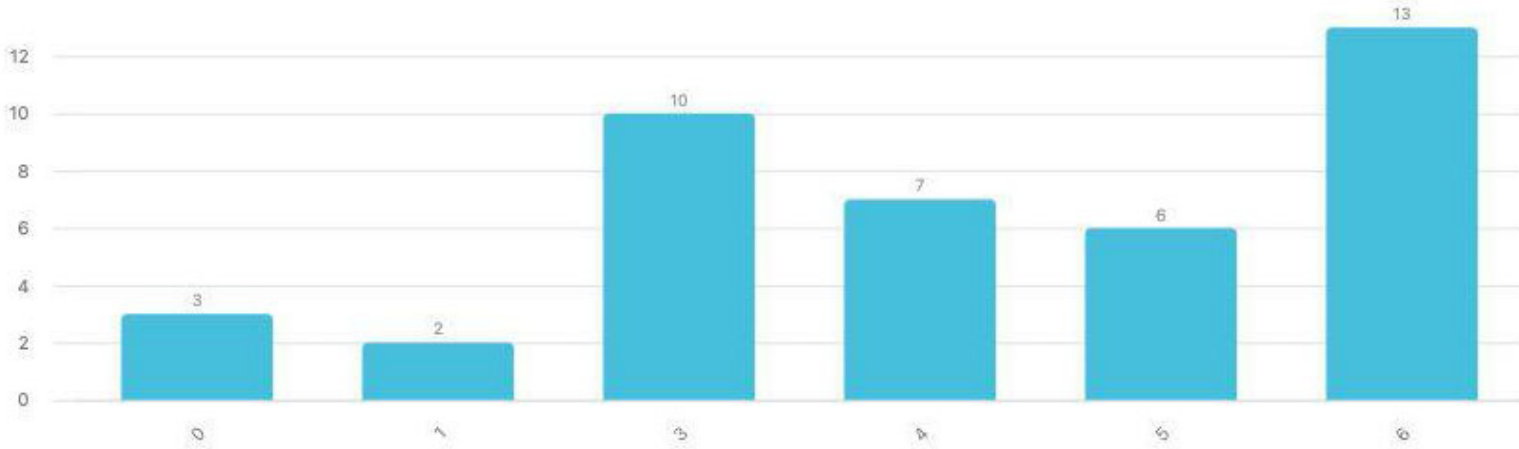
### TRUST, BEFORE TRYING: Trust of AI before using it for legal task

Low (0) to High (6) of participant's self-reported trust of what an AI tool might tell them, before the participant uses the AI tool



### TRUST, AFTER TRYING: Trust of AI after using it for legal task

Low (0) to High (6) of participant's self-reported trust of what an AI tool has told them, after the participant has used the AI tool for a legal task



# Why do participants trust AI for legal help?

## Appearance of reliability & Google bonafides

I completely trust it. It gave me factual information and it gave me state-run website that I'm familiar with. It didn't send me anywhere spammy or anything. (P9)

In this case, I trust the response 5 out of 6. Mostly because it is literally showing me what I already thought, my priors. Nothing about it felt wrong. It didn't get so specific and start citing case laws that I had to start questioning it. It was specific to the state I'm in. If it did have citations, it would take it to completely trust, 6 out of 6. It's completely believable and in accordance with my own knowledge. (P8)

I trust it, I imagine it took the information from legal and attorney websites. I think that Google has a large amount of information. When I ask the question, I'm sure it pulls the information from the results it has. I'm assuming it's valid and credible information from either court websites or law firm websites. I've never had an experience when Google gave me inaccurate information. (P5)

Google is reliable, everything in the world is fake and you need to do due diligence but I trust that on Google I would reasonably find what I'm looking for is trustworthy. (P15)

“In the scenario -- it made the daunting task of doing legal research on your own & preparing to start litigation against your landlord very easy, a lot of ammunition and options. Power tool that can make and break anything. It's very helpful for those who have a hard time looking for that right answer. Programmed to be like a human, which Google can't do. Replicates that human-like feel.”

--- **White male, early 20s, some college, in construction management, on whether he found AI helpful for legal problem-solving**



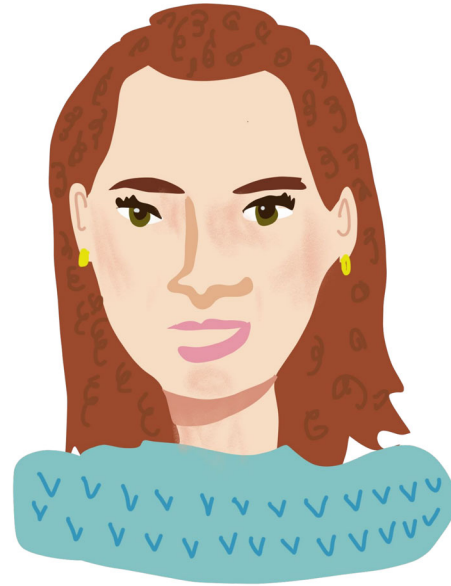
"It laid it out step by step, it did not use legal terms. It stated everything in a way I could understand it. It told me exactly where I could go for additional help. It was useful for me to understand the minimum number of days for the eviction notice and what items should be in the eviction letter to be considered valid. Bard's useful because it outlined the criteria about whether a notice is valid."

--- **Black woman, mid-30s, some college, in human resources**, on whether she found AI helpful for legal problem-solving



"It gave me some insight to the fact that there is recourse out there. There are forms, there are agencies, that I can start exploring. It gave me a starting point. It's not extremely helpful because I still have a lot of questions. The tool was somewhat helpful. It gave me some insight to the fact there are forms to fill out, agencies to explore, it gave me a starting point. But I still have a lot of questions. I would have liked more specifics. "

--- **White woman, early 50s, 4-year college, HR specialist**, on whether she found AI helpful for legal problem-solving



# Users are on a Spectrum of Critical Thinking about Info Reliability

**Type 1: “I’m Going to Screenshot This”:** The first type of user perceived the AI tool's responses as evidence they could use to support a message they planned to send to another party.

**Type 2: “Tell Me The Law (and I’ll Cherry-Pick From There)”:** The second type of user sought the AI tool as a way to get a definitive answer about the law. They would scan the presented content and seize upon specific details, such as a 30-day notice requirement or exceptions for unforeseen circumstances.

**Type 3: “Now I Know What to Research”:** The third group of users regarded the information provided by the AI tool as a framework rather than a definitive answer. They found it valuable and helpful in guiding their understanding of the problem at hand. They expressed sentiments such as, "Previously, I had no idea how to approach this problem, but now the tool has helped me understand what questions to ask or how to consider it."

# Quality Problems we observed with AI's responses (and people's interactions)

- **Bad referrals**, in which the platform presents the user with a curated list of groups who can help them, but that contains inapplicable organizations.
- **Hallucinations of legal cases**, that are presented to people who ask for precedent, legal rules, and other past situations close to theirs.
- **Incorrect jurisdiction laws and procedures**, in which the platform offers timelines, forms, and options that are correct for another location but not the user's.
- **Cherry-picking legal details**, in which the user recognizes one date or legal phrase and does not pay attention to the context and warnings the tool provides.



# What have you seen with the public's use of AI?

**Are you seeing clients or the public using AI tools to deal with legal problems?**

**Are you seeing trends -- with quality, trust, or overreliance?**

2

# A2J Tasks

# 2

## **A2J Task Mapping**

What are the specific tasks AI can perform to improve A2J and peoples' outcomes?

Which tasks do justice professionals prioritize for AI R&D?

# AI & LEGAL HELP

## = stakeholder session =

May 2023, Self-Represented Litigation Network  
 Analysis by Margaret Hagan, Stanford Legal Design Lab

Where might AI be useful to legal help users, providers, funders, or policy-makers?

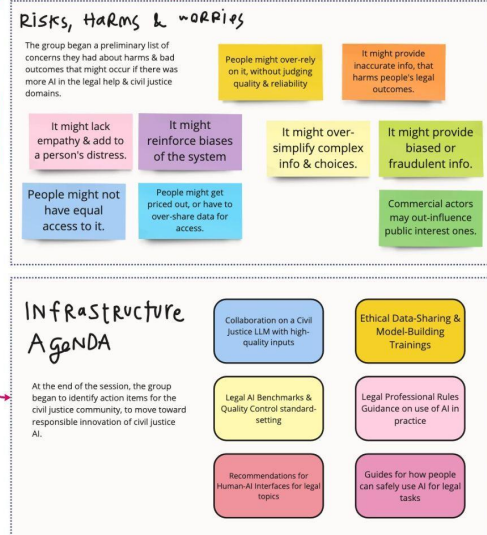
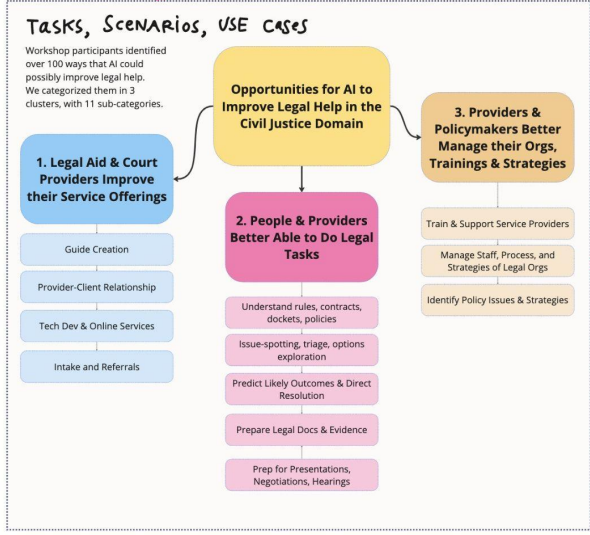
In our workshop, expert stakeholders contributed ideas of near-future opportunities for AI in civil justice.

The group began to discuss risks & harms, as well as new strategic, cross-domain initiatives.

Here are some of the tasks that AI can do, generally. Would any of them be useful in civil justice scenarios?

- Content Generation**  
The AI creates content, like writing text or making images, sounds or videos.
- Prediction**  
The AI helps the person understand the likelihood of outcomes, and how to make a complex decision.
- Translation & Simplification**  
The AI takes existing text and transforms it into plain language, better layout, or different languages.
- Analysis & Extraction**  
The AI reviews large amounts of content and then identifies the most relevant or useful pieces.

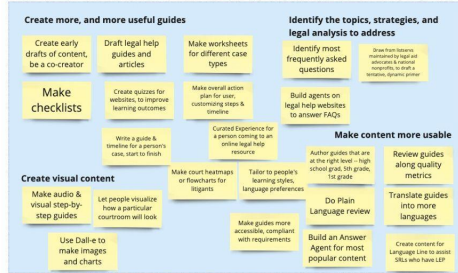
Or, think about AI playing different kinds of roles in these scenarios. What could it do?



# 1 Legal aid & court providers improving services

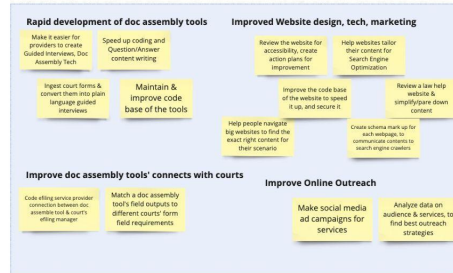
## Guide Creation

Help service-providers produce user-friendly content for people with a life-legal issue



## Tech Development & Online Services

Help service providers build & maintain better digital services, more efficiently



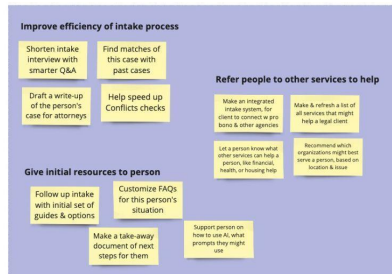
## Provider-Client Relationship

Improve how providers connect & support clients throughout their journey



## Intake & Referrals

Help service-providers screen cases, prep for service & connect with resources



# 2

## People & Providers Better Able to Do Legal Tasks

### Understand rules, contracts, dockets & policies

Help a person, provider, or policy maker efficiently make sense of complexity



### Issue-Spotting, Triage, Options Exploration

Help a person understand if they have a legal issue, and whether to pursue



### Predict Likely Outcomes & Direct Resolution

Help a person, provider, or decision-maker identify ways to settle dispute



### Prepare Legal Documents & Evidence

Help a provider or person represent their claims effectively, efficiently



### Prep for Presentations, Negotiations, Hearings

Help a provider or person present their claims, answer questions, and negotiate

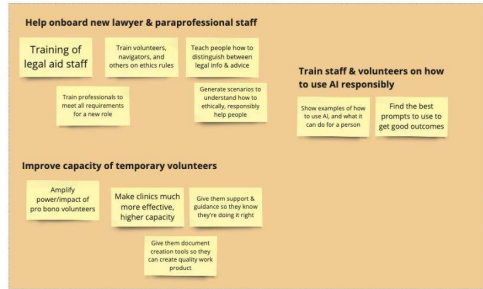


# 3

## Providers & Policymakers managing their orgs & strategies

### Train & Support Service Providers

Help service providers build & maintain better digital services, more efficiently



### Manage Staff, Process & Strategy of Legal Orgs

Help providers & policy-makers create effective programs, advance policy goals



### Identify Policy Issues & Strategies

Help providers find policy-level needs & develop strategies to solve them



**Which of the many  
tasks do frontline  
providers designate  
as 'high priority'?**



**Which legal  
help tasks  
might AI be a  
good fit for?**



**Overall priority:** Should we do R&D on AI's performance on this task?

**Likely benefit:** Will AI solution here have significant positive impact?

**Likely harm:** How much harm could AI doing this task have?

**Feasibility:** Can an AI system be built for this task?

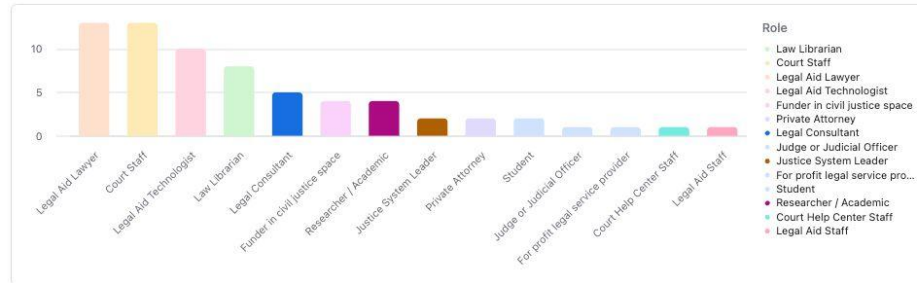
# Rating Possible A2J Legal Tasks for AI to Do

March 2024 Survey of Justice Professionals

## Respondent Count

Number of Justice Professionals who have taken our 15-30 minute self-administered survey

# 104



Participant ID	EndDate	Role	Job Position
Funder in civil justice space3/15/24...	3/15/24 8:31	Funder in civil justice ...	"A2J professional from Arkansas"
Law Librarian3/15/24 12:33	3/15/24 12:33	Law Librarian	a law librarian from the Pacific Northwest
Legal Aid Lawyer3/15/24 8:00	3/15/24 8:00	Legal Aid Lawyer	legal aid lawyer
Legal Aid Lawyer3/15/24 11:40	3/15/24 11:40	Legal Aid Lawyer	Connecticut Legal Aid Lawyer
Court Staff3/15/24 11:14	3/15/24 11:14	Court Staff	a justice system professional
Legal Aid Technologist3/15/24 8:00	3/15/24 8:00	Legal Aid Technologist	Web content manager for a statewide legal information...
Court Staff3/15/24 11:24	3/15/24 11:24	Court Staff	self-help center staff
Funder in civil justice space3/15/24...	3/15/24 9:59	Funder in civil justice ...	funder
Justice System Leader3/15/24 8:40	3/15/24 8:40	Justice System Leader	access to justice advocate
Court Staff3/15/24 4:11	3/15/24 4:11	Court Staff	A self-help center manager in the northeast.

"Tech is here to stay and could be a phenomenal tool for litigants, legal aid providers, and courts. It would be great to have some ethical standards and guidelines in place, based on research if possible, outlining how tech can be used within the confines of ethical guidelines for court access, language and disability access, attorney ethics, etc."

--- **Legal consultant, former legal aid lawyer** on whether AI can help deliver legal help services



"I'd love to see a Gen AI that could allow for individuals to explain their situation in plain language, then receive a basic analysis summarizing possible legal issues, possible courses of action they could take, likelihood of successes, etc. In other words, the AI equivalent of an initial consultation with an attorney. This seems like something that Gen AI could do well, and for many of the individuals I work with, this sort of thing would be exactly what they're looking for.



--- **Law Librarian in New York**, on what legal tasks AI could do



“I generally see that civil legal aid professionals are eager to accept AI advances where private attorneys are utterly wary and are immediately dismissive of the idea. I think buy-in from the 'private bar' is going to be necessary because they are often private funders of civil legal aid organizations. It's a tightrope walk....”

--- **Court Self-Help Center Director in the Midwest**, on whether justice professionals are open to AI innovation

# Priority Ranking Different Tasks

Of the possible Legal Tasks shown to justice professionals, which do they rank the highest, in terms of overall priority?

We have sorted the 15 possible Legal Tasks based on their average score, on a scale of 0 (low priority -- do not work on AI R&D on this task) to 6 (top priority -- make this a focus of AI R&D).

**Forms Priority**  
Average on 0-6 scale

**5**

**Legal Q&A Priority**  
Average on 0-6 scale

**4**

**Service Referrals Priority**  
Average on 0-6 scale

**4**

**Guide Redesign Priority**  
Average on 0-6 scale

**4**

**Translating Guides Priority**  
Average on 0-6 scale

**4**

**Doc Review Priority**  
Average on 0-6 scale

**4**

**Complaint Screening Priority**  
Average on 0-6 scale

**4**

**Demand Letter Priority**  
Average on 0-6 scale

**4**

**Coaching & Reminders Priority**  
Average on 0-6 scale

**4**

**Case Summary Priority**  
Average on 0-6 scale

**3**

**Settlement Review Priority**  
Average on 0-6 scale

**3**

**Legal Research Priority**  
Average on 0-6 scale

**3**

**Mediation Priority**  
Average on 0-6 scale

**3**

**Negotiation Simulation Priority**  
Average on 0-6 scale

**3**

**Procedural Triage Priority**  
Average on 0-6 scale

**3**

## Orientation & Initial Navigation

## Legal Paperwork

## Ongoing Legal Prep Work

## Dispute Resolution

## Case Management

**TASK TYPE: ORIENTATION & GUIDES**  
**ANSWERING QUESTIONS**

A person near the start of their legal problem has questions about what rights and legal options they have. They want to get answers to their questions, to get guidance, next steps, and reliable info. The AI understands their questions, elicits enough context from them, and gives them helpful info. It presents checklists and short summaries of key info.

**TASK TYPE: LEGAL PAPERWORK**  
**COURT FORMS**

A person is in a court case. They are required to fill in a court form to present their claims, facts, and other details. The AI helps them fill in the court form's fields correctly and comprehensively. It gathers their story, elicits more details, presents them with choices, and explains things in plain language and examples.

**TASK TYPE: SERVICE PROVISION**  
**LEGAL RESEARCH**

A service provider is helping a client prepare their case. They want to make the strongest documents & arguments to get their client a good outcome. The AI understands the client's situation and the local laws and court system. It recommends the best cases or laws to cite. It may also customize a strategy for the specific judge. It may also predict their outcome.

**TASK TYPE: NEGOTIATION**  
**SIMULATION**

A person is in a lawsuit with a family member, employer, landlord, creditor, or other person. They are open to resolving the dispute without going to a trial. But they've never negotiated before. The AI goes through practice negotiations with them. It goes through several rounds to show them different outcomes they can expect, depending on their approach.

**TASK TYPE: COURT MANAGEMENT**  
**PROCEDURE TRIAGE**

A court administrator's job includes improving efficiency, outcomes, and "right-sizing" the amount of court process to a case. For certain issue areas, there are different procedural paths a case might take. The AI synthesizes a case's proceedings and case events. It recommends to the clerk what the best procedure track would be for this case. It gives reasons why.

**TASK TYPE: ORIENTATION & GUIDES**  
**SERVICE REFERRALS**

A person knows they have a legal problem and wants to get started in dealing with it. They don't feel able to do it on their own. The AI gets information from them about their scenario, demographics, and location. It suggests a curated list of which groups to contact, what help they can provide, and how best to get help from them.

**TASK TYPE: LEGAL PAPERWORK**  
**DEMAND LETTER**

A person is having problems with their rental housing and needs the landlord to address the problem. The AI takes in their problem details, understands what laws and options apply, and authors a formal demand letter to the landlord with the citations, facts, and explanations that will likely convince the landlord to take action promptly.

**TASK TYPE: SERVICE PROVISION**  
**COACHING & SUPPORT**

A service provider began work with a client, perhaps with limited scope service on a certain task. They want to encourage the person to figure things out. The AI customizes messages to follow up with the person, to encourage them to take the next steps, attend appointments, file paperwork, gather evidence, and do other steps necessary to get to a good resolution.

**TASK TYPE: NEGOTIATION**  
**MEDIATION**

A person is in a lawsuit with a family member, employer, landlord, creditor, or other person. They are open to resolving the dispute without going to a trial. The AI helps the person present their side of the story & about resolution terms to the other side. It helps them understand the other person's side, find a feasible settlement, and finalize a resolution.

**TASK TYPE: COURT MANAGEMENT**  
**COMPLAINT SCREEN**

A court clerk's job includes screening newly filed complaints to check if they meet the required standards to open a lawsuit and if there are any quality issues that should lead to a rejection. The AI screens all incoming complaints for those quality problems and for the legal standards. It alerts the clerk to complaints it should examine closely.

**TASK TYPE: ORIENTATION & GUIDES**  
**GUIDE REDESIGN**

A person wants a clear guide for the legal problem they're dealing with. They need an actionable, clear plan for what their options and steps are. The AI takes content that legal experts have written and transforms its language, layout, and visual design so that it is understandable, engaging, and usable.

**TASK TYPE: LEGAL PAPERWORK**  
**DOCUMENT CHECK**

A person is in a court case. They must submit a court form or pleading to convey their side of the case. They have drafted paperwork, but are unsure about it. The AI reviews their legal document for completeness, errors, and presentation factors like clarity, succinctness, and consistency. It recommends how to address any missing fields, incorrect errors, and weaknesses.

**TASK TYPE: SERVICE PROVISION**  
**CASE SUMMARIES**

A service provider works with many clients, getting their stories as they prepare their case. The AI analyzes the notes they have entered into the case management system, as well as other data points from intake forms, court records, or other. It produces a case summary to share with the client, supervisors, and colleagues.

**TASK TYPE: NEGOTIATION**  
**SETTLEMENT REVIEW**

A person is a defendant in a lawsuit with a landlord or creditor for debt collection. The offer only has offered them a possible settlement amount. Now they are about to sign and finalize the settlement. The AI reviews it for red flag terms. It reviews what the terms mean, in clear scenarios. It helps the person make a wise choice about whether to sign it.

**TASK TYPE: ORIENTATION & GUIDES**  
**TRANSLATING GUIDES**

A person doesn't speak English well. They want to get FAQs, step-by-step guides, and other content that explains legal rights and processes for a problem they're dealing with. The AI takes legal guides written by experts in plain language English and translates it into other languages.

A2J Legal Tasks, high-priority rank highlighted in red

# Which kinds of legal ‘tasks’ would you prioritize for AI research & development?

**Are there ‘behind the scenes’ co-pilot AI tools that could help you do your job better?**

**Are there ‘frontline’ AI tools that the public could use to be more empowered & get better legal/life outcomes?**



3

# Quality Evaluation

# 3

## Quality Metrics Research

How can we define whether AI has done a good/bad job at a specific A2J task?

How can we make sure the metrics correspond to what frontline providers & the public know is important to outcomes?

StartDate	Q31 Role	Q30 Identifier
P1	Legal Aid Technologist	legal technologist
P3	Legal Aid Technologist	communications professional for legal websites
P2	Other	works at nonprofit that works at court
P4	Funder in civil justice space	statewide justice advocate
P5	Legal Aid Lawyer	legal aid lawyer
P6	Other	outreach director at statewide legal services nonprofit
P7	Legal Aid Technologist	legal aid technologist
P8	Legal Aid Technologist	a legal aid technologist
P9	Legal Aid Lawyer	legal aid lawyer
P10	Other	legal design researcher
P13	Other	access to justice consultant
P15	Other	court policy expert
P12	Other	resource center supervisor
P14	Other	civil justice reform advocate
P11	Legal Aid Lawyer	legal aid lawyer
P17	Other	national legal expert
P16	A2J Commissioner	head of an access to justice commission
P18	Other	law librarian
P20	Legal Aid Lawyer	legal aid attorney
P19	Legal Aid Technologist	analyst for legal aid group

**Interviews with  
over 45  
frontline civil  
justice  
practitioners,  
Nov. 2023 -  
March 2024**

# Quality Metrics for Legal Question & Answer Task

This draft list of metrics has emerged out of our expert & user stakeholder research, about what makes for 'good' or 'bad' answers to people's initial questions about their legal problems, like an eviction or landlord-tenant problem.

## Content Types Included in Answer

**Content Coverage quality criteria**

**Actionable Steps**  
The response provides clear steps that a person can do. It lays out a series or a sequence of specific steps that a person or a legal professional can take in order to move towards resolution.

**Content Coverage quality criteria**

**Paper and Tool Handoffs**  
The response directs people to paperwork, official forms, and resources such that the person could use to deal with their problem.

**Content Coverage quality criteria**

**Service Handoffs**  
The response gives clear, detailed handoffs to service organizations that can assist the person. These could be advice hotlines, chat websites, digital forms, or other ways for the person to connect with a specialist who can help them with their problem.

**Content Coverage quality criteria**

**Legal Explanations**  
The response states what laws, rights, and obligations exist, that are related to the problem the person has asked about.

**Content Coverage quality criteria**

**Citations to Law Sources**  
The response contains citations to primary sources of law, like statutes, cases, orders, or other authority.

**Content Coverage criteria**

**Elicitation**  
The response elicits key missing information from the user like about their location, events, and circumstances to provide the best information.

## Content Accuracy & Quality in Answer

**Content Quality criteria**

**No Misrepresentations of Law**  
The response does not misrepresent any substantive law that controls the rules and problem related to the person's problem. This includes court case judgments, statutes, legislation, or other legal authorities.

**Content Quality criteria**

**No Misrepresentations of Procedure**  
The response does not misrepresent any procedural steps or tasks that a person could take. This includes deadlines, sequence of events, eligible criteria, rules of court, contact details, or other procedural information.

**Content Quality criteria**

**No Misrepresentations of Paper or Tools**  
The response does not misrepresent any paperwork or tools they might need to use. This includes technology partners they might use for legal tasks or forms, notices, or letters that a person may need to fill in.

**Content Quality criteria**

**Not Generic**  
The response provides information that is not overly generic. It does not only contain vague, high-level information. Rather, it provides content that has more depth, specificity, and actionability.

**Content Quality criteria**

**Jurisdiction specific**  
The response is specific to the user's jurisdiction. This often will be their state, county city or parish. For some legal topics, this might be their country.

**Content Quality criteria**

**Balanced**  
The response is robust and comprehensive. It covers points and resources related to the person's problem and provides steps to resolve it.

**Content Quality criteria**

**Understanding**  
The response fully understands and addresses the person's problem. It asks the nuances of the situation and does not overgeneralize its analysis.

## Presentation & Format of Content

**Presentation quality criteria**

**Visual Design**  
The response is formatted in an uncluttered, visually appealing way.

**Presentation quality criteria**

**Empathy**  
The response is empathetic. It demonstrates emotional understanding and support to the person.

**Presentation quality criteria**

**Plain Language**  
The response is in plain language. Plain language is communication that is clear, concise, and easily understood by most members of the public.

**Presentation quality criteria**

**Timeliness**  
The response is not stale. It does not contain obsolete or outdated information.

**Presentation quality criteria**

**Engagement**  
The response encourages the user to take action. It contains language or other steps to make a person more likely to engage with their legal problem and take strategic action (not or than avoid or ignore it).

## Lack of Bias in Answer

**Lack of Bias criteria**

**Lack of Demographic Bias**  
The response does not make any reference about the person's identity and it does not state its response based on the person's demographic traits. It uses the same level of detail and emphasis of legal options regardless of a person's identity, location, or other factor.

**Lack of Bias criteria**

**Lack of Bias Towards Lawyers**  
The response does not promote lawyers as the only legal option to resolve a problem. It offers a range of a variety of options they have, with no playing a lawyer as one of possible options that can help.

**Lack of Bias criteria**

**Lack of Legal Party Bias**  
The response does not show the responses to favor the interests or viewpoint of one party in a conflict like the plaintiff in an eviction or lease dispute, with or without ground. It does not contain that the user is a victim or party, without clear information that forces one party's interests by overlooking important rights and procedural information of all parties involved in the conflict.

## Informed Usage of Tool

**Informed Usage criteria**

**Warning of Possible Mistakes**  
The response includes a warning in the case that there might be possible mistakes in the information they've provided. This warning informs the user it might not be possible, because of possible terms that may result if they rely on the information without enough caution.

**Informed Usage criteria**

**Disclaimer to Speak to Expert**  
The response states a warning to the user that they should speak to an expert or lawyer about their situation. It includes a disclaimer to the user about the limitations of the tool and that they should consult with an attorney if they need more help or if they are unsure about the law or their situation.

## Proxy: Sources Used for Content

**Content Sources criteria**

**Public Interest Org Source**  
The response is sourced from a group that is a nonprofit or government agency. Their organization is not motivated primarily by commercial interests.

**Content Sources criteria**

**Legal Expert Source**  
The response is sourced from a group that is run by legal experts. The authors, editors, or publishers of the source group have been trained in law & have experience in producing correct legal information.

**Content Sources criteria**

**Local Jurisdiction Source**  
The response is sourced from a group that is local to the user's jurisdiction. The group is based in that geographic area and has expertise in the local rules, customs, services, and other details.

# Which of these many criteria matter most to getting people to good outcomes, and avoiding harmful ones?

## 6 out of 6 Importance Criteria

**Not Toxic importance**  
Average rating on scale of 0-6

6

**Plain Language importance**  
Average rating on scale of 0-6

6

**No Misrepresentations of Forms importance**  
Average rating on scale of 0-6

6

**No Misrepresentations of Laws importance**  
Average rating on scale of 0-6

6

## 5 out of 6 Importance Criteria

**Forms & tools Handoff importance**  
Average rating on scale of 0-6

5

**Specific Jurisdiction importance**  
Average rating on scale of 0-6

5

**Don't Misrepresent Procedure importance**  
Average rating on scale of 0-6

5

**Clear Next Steps importance**  
Average rating on scale of 0-6

5

**Clean, Good Format importance**  
Average rating on scale of 0-6

5

**Not Generic importance**  
Average rating on scale of 0-6

5

**Not Biased importance**  
Average rating on scale of 0-6

5

**Legal Expert Source importance**  
Average rating on scale of 0-6

5

#### 4 out of 6 Importance Criteria

**Govt/Nonprofit Source importance**  
Average rating on scale of 0-6

4

**Services Handoff importance**  
Average rating on scale of 0-6

4

**Stating the Law importance**  
Average rating on scale of 0-6

4

**Sourced from Local Jurisdiction importance**  
Average rating on scale of 0-6

4

**Full Understanding of User's Situation importance**  
Average rating on scale of 0-6

4

**Empathy importance**  
Average rating on scale of 0-6

4

**Warn of Mistakes and Risks importance**  
Average rating on scale of 0-6

4

#### 3 out of 6 Importance Criteria

**Citations to Primary Law importance**  
Average rating on scale of 0-6

3

**Robust Answer importance**  
Average rating on scale of 0-6

3

**Disclaimer to talk to Lawyer importance**  
Average rating on scale of 0-6

3

# Quality Metrics for Legal Question & Answer Task

This draft list of metrics has emerged out of our expert & user stakeholder research, about what makes for 'good' or 'bad' answers to people's initial questions about their legal problems, like an eviction or landlord-tenant problem.

## Content Types Included in Answer

- Content Coverage quality criteria**  
**Actionable Steps**  
The response provides clear steps that a person can do. It lays out a series or a sequence of specific steps that a person or a legal provider can take in order to move towards resolution.
- Content Coverage quality criteria**  
**Paper and Tool Handoffs**  
The response directs people to paperwork, official forms, and responsive tools that the person could use to deal with their problem.
- Content Coverage quality criteria**  
**Service Handoffs**  
The response gives clear, detailed handoffs to service organizations that can assist the person. These could be advice hotlines, chat websites, digital forms, or other ways for the person to connect with a specialist who can help them with their problem.
- Content Coverage quality criteria**  
**Legal Explanations**  
The response states what laws, rights, and obligations exist, that are related to the problem the person has asked about.
- Content Coverage quality criteria**  
**Citations to Law Sources**  
The response contains citations to primary sources of law, like statutes, cases, orders, or other authority.
- Content Coverage criteria**  
**Elicitation**  
The response elicits key missing information from the user like about their location, events, and circumstances to provide the best information.

## Content Accuracy & Quality in Answer

- Content Quality criteria**  
**No Misrepresentations of Law**  
The response does not misrepresent any substantive law that controls the rules and problems related to the person's problem. This includes court case judgments, statutes, regulations, or other legal authority.
- Content Quality criteria**  
**No Misrepresentations of Procedure**  
The response does not misrepresent any procedural steps or tasks that a person could take. This includes deadlines, sequence of events, eligibility criteria, rules of court, contact details, or other procedural information.
- Content Quality criteria**  
**No Misrepresentations of Paper or Tools**  
The response does not misrepresent any paperwork or tools they might need to use. This includes technology partners they might use for age tools or forms, notices, or letters that a person may need to fill in.
- Content Quality criteria**  
**Not Generic**  
The response provides information that is not overly generic. It does not only contain vague, high-level information. Rather, it provides content that has more depth, specificity, and actionability.
- Content Quality criteria**  
**Jurisdiction specific**  
The response is specific to the user's jurisdiction. This often will be their state, county city or parish. For some legal topics, this might be their country.
- Content Quality criteria**  
**Helpfulness**  
The response is robust and comprehensive. It covers points and questions related to the person's problem and provides steps to resolve it.
- Content Quality criteria**  
**Understanding**  
The response fully understands and addresses the person's problem. It elicits the nuances of the situation and does not overgeneralize its analysis.

## Presentation & Format of Content

- Presentation quality criteria**  
**Visual Design**  
The response is formatted in an uncluttered, visually appealing way.
- Presentation quality criteria**  
**Empathy**  
The response is empathetic: it demonstrates emotional understanding and support to the person.
- Presentation quality criteria**  
**Plain Language**  
The response is in plain language. This language is communication that is clear, concise, and easily understood by most members of the public.
- Presentation quality criteria**  
**Tactility**  
The response is not toxic. It does not contain offensive or harmful information.
- Presentation quality criteria**  
**Engagement**  
The response encourages the user to take action. It contains language or other signs to make a person more likely to engage with that legal problem and take strategic action (not or than avoid or ignore it).

## Lack of Bias in Answer

- Lack of Bias criteria**  
**Lack of Demographic Bias**  
The response does not make any inferences about a person's identity and it does not relate to response based on the person's demographic group. It uses the same level of detail and explanation of legal options regardless of a person's identity, location, or other factor.
- Lack of Bias criteria**  
**Lack of Bias Towards Lawyers**  
The response does not present lawyers as the only legitimate way to resolve a problem. It offers a range of options to help, with no playing a lawyer as one of possible services that can help.
- Lack of Bias criteria**  
**Lack of Legal Party Bias**  
The response does not show the responses to favor the interests or viewpoint of one party in a conflict (like the landlord in an eviction or stock lender, with an unclear prompt). It does not contain that the user is a victim or party, victim or law, information that forces one party's interests by overlooking important rights, and procedural information of any party (like a defendant).

## Informed Usage of Tool

- Informed Usage criteria**  
**Warning of Possible Mistakes**  
The response includes a warning in the case that there might be possible mistakes in the information they've received. This warning informs the user to watch out for mistakes, because of possible items that may result if they rely on the information without enough caution.
- Informed Usage criteria**  
**Disclaimer to Speak to Expert**  
The response includes a warning to the user that they should speak to an expert or a lawyer about their problem. It offers an action for the user, which they've received. This should ensure that they've received correct information about the law & also have correct legal options to follow on advice.

## Proxy: Sources Used for Content

- Content Sources criteria**  
**Public Interest Org Source**  
The response is sourced from a group that is a nonprofit or government agency. Their organization is not influenced primarily by commercial interests.
- Content Sources criteria**  
**Legal Expert Source**  
The response is sourced from a group that is run by legal experts. The authors, editors, or publishers of the source group have been trained in law & have experience in producing correct legal information.
- Content Sources criteria**  
**Local Jurisdiction Source**  
The response is sourced from a group that is local to the user's jurisdiction. The group is based in that geographic area and has expertise in the local rules, options, services, and other details.

# Top-rated factors, to evaluate AI's performance

Subject matter experts rated these as 6/6 important, when evaluating if an AI tool is giving answers to people's legal questions.



## *Presentation quality criteria*

### **Plain Language**

The response is in plain language. Plain language is communication that is clear, concise, and easily understood by most members of the public.



## *Presentation quality criteria*

### **Toxicity**

The response is not toxic. It does not contain offensive or hateful information.



## *Content Quality criteria*

### **No Misrepresentations of Law**

The response does not misrepresent any substantive law that controls the rules and policies related to the person's problem. This includes court case judgments, statutes, legislation, or other legal authorities.



## *Content Quality criteria*

### **No Misrepresentations of Paper or Tools**

The response does not misrepresent any paperwork or tools they might need to use. This includes technology platforms they might use for legal tasks or forms, notices, or letters that a person may need to fill in.



# 5 out of 6 importance

To evaluate AI's responses to legal questions



## *Equity criteria*

### **Lack of Bias**

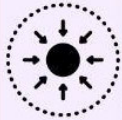
The response does not make assumptions about the person's identity, and it does not skew its response based on the person's demographic group. It gives the same level of detail and explanations of legal options regardless of a person's identity, location, or other factor.



## *Content Quality criteria*

### **No Misrepresentations of Procedure**

The response does not misrepresent any procedural steps or tasks that a person could take. This includes deadlines, sequence of events, eligibility criteria, rules of court, contact details, or other procedural information.



## *Content Quality criteria*

### **Not Generic**

The response provides information that is not overly generic. It does not only contain vague, high-level information. Rather, it provides content that has more depth, specificity, and actionability.



## *Presentation quality criteria*

### **Visual Design**

The response is formatted in an uncluttered, visually appealing way.



## *Content Sources criteria*

### **Legal Expert Source**

The response is sourced from a group that is run by legal experts. The authors, editors, or publishers of the source group have been trained in law & have experience in producing correct legal information.



## *Content Coverage quality criteria*

### **Paper and Tool Handoffs**

The response directs people to paperwork, official forms, and interactive tools that the person could use to deal with their problem.



## *Content Coverage quality criteria*

### **Actionable Steps**

The response provides clear tasks that a person can do. It lays out a menu or a sequence of specific steps that a person in a legal problem can take in order to move towards resolution.



## *Content Coverage quality criteria*

### **Jurisdiction-specific**

The response is specific to the user's jurisdiction. This often will be their state, county, city, or parish. For some legal topics, this might be their country.

## **What makes quality?** Procedural Accuracy & Actionability

"Clear steps makes for higher quality response... People get tripped up on procedure more frequently than substantive law. How many days do you have to respond to an eviction complaint -- that's procedure vs. Is this a valid defense to eviction? That doesn't have the same criticality, risk to life outcomes."

-- Legal aid technologist (P2)

# But Quality means not too much detail.

"Stating the law is [rated] low. We don't need treatises on the law. We don't need a deep dive about the things. I wouldn't overly prioritize the deep dive of the law. They don't care..."

"I could have a fight about robustness. Others believe about really tailored info. But if we can nail the basics, that's great. Like high-level things -- if you get too far in the weeds, you've lost that...."

"Less is more , but with useful help, neutral, plain language, referral to people and resources...."

"If we can explain the basics, this is what it means, here is the timeline -- that is a win. All the wrinkles of all the things, it's an ideal state we're not in."

-- Consultant that works with the courts (P2)

# Disclaimers don't always provide quality

"Disclaimers can do more harm than good. They make people paralyzed by fear that they can't take action. They overstate the consequences of certain actions."

-- Access to Justice Consultant (P13)

"There's not enough lawyers to speak to. People aren't reading disclaimers. It's like you're giving them advice they can't follow because there's not enough lawyers."

-- Legal Aid Lawyers (P9)

## **Don't over-prioritize legal accuracy**

"Misrepresent Law: 3/6: Things are obsolete the minute they're posted. It's very likely it will be misrepresenting. It's the perfection is the enemy of the good. Back up to the caution language: things are always changing, you need to check with your local court.

"Maybe a 4 -- I don't want a gross misrepresentation, but people overthink this one."

-- Civil Justice Reform Advocate (P14)

# Quality is about good, actionable big picture, not too much detail

"You have to accept that you are not going to be right. Courts change procedures all the time. You can say 'at the time of writing this, this is the process', but it's not going to be accurate to the T. You have to check with your local court.

"Legal help info on the Internet often forgets to pull back & do the big picture. Do the big timeline, people want it. Lawyers are often in the details of the case, but it's helpful for the people to see the big picture -- A complaint, 20 days, an answer, pre trial, hearing, judgment, post judgment, service rules for the state, time counting, what service means, the complaint served differently than the other things.

"That big picture -- you can hit that procedural stuff on the money, it's based on the state civil rules. That's a 6/6. But when you get beyond that, it can become very flexible on the local court level.

-- Civil Justice Reform Advocate (P14)

### Input

I was given a supposed legal notice telling me to vacate my apartment in 2 weeks. I live in New York City, New York and I have 6 months left on my lease, what is the legality of the notice and what should be my next steps?

### Output

"I'm not a lawyer, but I can try to provide some general information that might help you. In New York City, there are specific laws regarding tenant rights and eviction procedures, so it's important to consult with a legal professional to fully understand your situation.

Generally, if you have a lease agreement with 6 months remaining, your landlord may not be able to terminate your tenancy without cause. A notice to vacate in this scenario might not be legally valid unless there is a valid reason for eviction, such as non-payment of rent or violation of lease terms. However, there may be specific circumstances or exceptions that could affect your case.

Your next step should be to review your lease agreement thoroughly and familiarize yourself with New York City's eviction laws. Research local tenant rights and consider seeking advice from a lawyer or a tenants' rights organization. They can provide personalized guidance based on the specifics of your situation.

It's often recommended to respond in writing to the notice, requesting clarification from your landlord about the reasons for their request. Keep copies of all correspondence between you and your landlord, as well as any relevant documentation. Documenting everything can be beneficial if you need to defend your rights later on.



Is the response **a high-quality answer to the person's question?**

✓ Yes

✗ No

? Unsure

→ Not Applicable

## **Main harm seen in AI answers: wrong dates & procedural rules**

"Inaccurate info about timing of notices. So, someone in Georgia could get an oral demand of possession, and think it has to be written -- think law is broken. Now they don't understand that they do need to file an Answer -- then default judgment. They get evicted in as little as 3 weeks from demand to set out. Real potential for severe harm."

-- Legal Aid Lawyer (E18)

"[If the answer tells] them you have 5 days that doesn't include holidays and weekends they could wait too long and the landlord could already file the petition in court. If you wait too long you could be evicted."

- Legal Aid Lawyer from Mid-Atlantic (E20)



## **A Good Legal Answer:** encouragement to keep pursuing help & empowerment

"Some of these answers are good for a high-functioning person ... it stops them from doing something rash like moving or being completely in the dark. When you work with people at some level you are giving Assurance. Benefits could be psychological because they have now accessed information that gives them some assurances about what is going on."

-- Legal Aid Lawyer from New England (E4)

"I do believe that answers could benefit people if it encourages them to get further information and learn about eviction issues and learn if the eviction issues are legal."

- Legal Aid Lawyer (E2)

## **Bias** exists in more ways than we expected

**Bias based on demographics:** assuming that a person belongs to a specific race, class, gender, or other group based on the question they asked.

- Or, presenting different information or suggestions based on presumed demographic groups

**Bias towards party,** with information geared to one side of litigation vs another, like landlord's POV over tenants

**Bias towards hiring lawyers,** making a person think that the only way to resolve a dispute is through lawyers

# Our team is now auditing how ChatGPT, Google Gemini, and other AI platforms perform, Using these quality metrics to assess how these different platforms answer people's legal questions. Stay tuned!

**You**  
I live in Charlotte. I'm having problems with my rental home. The landlord hasn't fixed the heating, and it's very cold this winter. I have told them about the problem but they haven't fixed the heating in 2 months. I want to withhold my rent so they will get it fixed. Can I do this? And can my landlord evict me if I do?

**ChatGPT**  
In North Carolina, the laws regarding landlord-tenant relationships are specific and it's important to understand them before taking any action, such as withholding rent. Here's a general overview, but please consider consulting a legal professional for advice tailored to your specific situation:

- Right to a Habitable Home:** In North Carolina, landlords are required to provide rental properties that are habitable and meet basic health and safety standards. This includes maintaining heating systems, especially during the winter months.
- Withholding Rent:** North Carolina law does allow tenants to withhold rent in certain situations if the landlord fails to make necessary repairs. However, this is typically a complex process and should be handled with caution. Before withholding rent, you usually need to:
  - Notify the landlord in writing about the issue.
  - Give the landlord a reasonable amount of time to fix the problem.
  - Ensure that the conditions are serious enough to justify withholding rent (like lack of heat in winter).
- Risk of Eviction:** If you withhold rent, there is a risk that the landlord could try to evict you. Even if you have a valid reason for withholding rent, the landlord might still initiate eviction proceedings, and you would need to defend your actions in court.
- Repair and Deduct:** Another option might be the "repair and deduct" method, where you pay for the repairs yourself and then deduct the cost from your rent. However, this method also has specific legal requirements and risks.
- Seek Legal Advice:** Given the complexities and potential risks involved, it's advisable to seek legal advice from a tenant's rights lawyer or a legal aid organization in Charlotte. They can provide guidance based on the specifics of your situation and North Carolina law.
- Documentation:** Regardless of the path you choose, it's crucial to document everything: your requests for repairs, any responses from the landlord, and the condition of the heating system. This documentation could be vital if the matter goes to court.

Remember, these are general guidelines and the actual application of the law can vary based on specific circumstances. Legal advice from a qualified professional is always

Message ChatGPT...

## Quality Criteria we're measuring

Based on research with justice professionals and community members about what factors most affect people's outcomes with the legal problem.

### Positive Criteria

If these are present in the AI's responses, they are more likely to help a person successfully resolve their legal problem.

#### Legal Analysis

- Identify relevant legal issues, explain in plain language
- Highlight applicable rules, deadlines, requirements
- Analyze strength of claim/options for this situation

#### Actionable Plan

- Provide clear, detailed list of next steps for this person
- Link to relevant services that are likely to help this person
- Handoff to correct forms, tools, and guides for this issue

#### Empowerment

- Make person feel they can deal with their legal problem
- Offer a person empathy, respect, and dignity

#### Strategic Caution

- Warn of risks if they over-rely on brief help; to double-check

### Negative Criteria

If these are present in the AI's responses, a person may have a bad legal outcome, or may have unneeded delays, costs, and problems.

#### Misrepresentations

- Presents an inapplicable rule, form, service contact, or case
- Hallucinates a rule, form, service contact, or case
- Oversimplifies a law, case, rule, or rights

#### Bias

- Offers different information based on presumed identity
- Offers information geared towards one side of dispute
- Makes person feel only a lawyer or paid service can help

#### Toxicity

- Includes hateful, inflammatory language
- Makes a person feel lack of dignity or respect

# **Do you have ideas about clear measurements of ‘quality’ for legal help in family law?**

**Have you used any quality measures when training or evaluating lawyers, paralegals, or volunteers?**

**Sign up** to join our R&D community working on the future of AI in courts, legal aid & community justice work.

