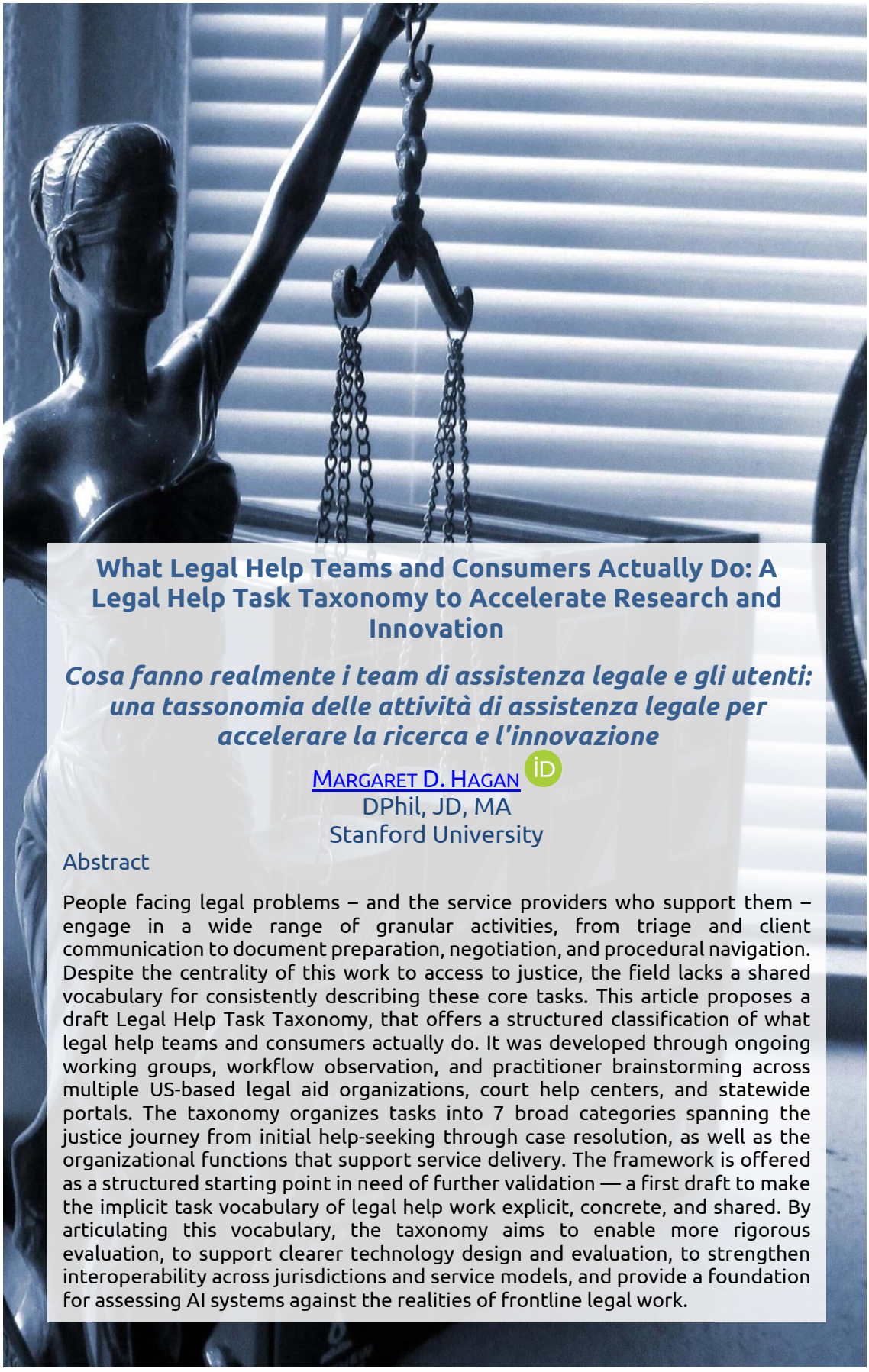
Hagan MD, 'What Legal Help Teams and Consumers Actually Do: A Legal Help Task Taxonomy to Accelerate Research and Innovation' (2025) Special Issue 2 EJPLT.
DOI: https://doi.org/10.57230/EJPLT25SIJMDH
Corresponding author. E-mail address: mdhagan@stanford.edu (M.D. Hagan)
Crown Quadrangle no. 559, Nathan Abbott Way, Stanford, CA (US)

EJPLT
ISSN 2704-8012



What Legal Help Teams and Consumers Actually Do: A Legal Help Task Taxonomy to Accelerate Research and Innovation

Cosa fanno realmente i team di assistenza legale e gli utenti: una tassonomia delle attività di assistenza legale per accelerare la ricerca e l'innovazione

MARGARET D. HAGAN
DPhil, JD, MA
Stanford University

Abstract

People facing legal problems – and the service providers who support them – engage in a wide range of granular activities, from triage and client communication to document preparation, negotiation, and procedural navigation. Despite the centrality of this work to access to justice, the field lacks a shared vocabulary for consistently describing these core tasks. This article proposes a draft Legal Help Task Taxonomy, that offers a structured classification of what legal help teams and consumers actually do. It was developed through ongoing working groups, workflow observation, and practitioner brainstorming across multiple US-based legal aid organizations, court help centers, and statewide portals. The taxonomy organizes tasks into 7 broad categories spanning the justice journey from initial help-seeking through case resolution, as well as the organizational functions that support service delivery. The framework is offered as a structured starting point in need of further validation — a first draft to make the implicit task vocabulary of legal help work explicit, concrete, and shared. By articulating this vocabulary, the taxonomy aims to enable more rigorous evaluation, to support clearer technology design and evaluation, to strengthen interoperability across jurisdictions and service models, and provide a foundation for assessing AI systems against the realities of frontline legal work.

© The author(s) 2025, published by Suor Orsola Benincasa Università Editrice.
This contribution is licensed under a Creative Common Attribution 4.0 International Licence CC-BY-NC-ND, all the details on the license are available at:
https://creativecommons.org/licenses/by-nc-nd/4.0/

Abstract
Le persone che si trovano ad affrontare problemi legali – e gli operatori che le assistono – svolgono una vasta gamma di attività, che vanno dallo smistamento dei casi e dalla comunicazione con i clienti alla preparazione dei documenti, alla negoziazione e alla gestione delle procedure. Nonostante la centralità di questo lavoro per l'accesso alla giustizia, il settore non dispone di un vocabolario condiviso che consenta di descrivere in modo coerente questi compiti fondamentali. Il presente articolo propone una bozza di "Tassonomia delle attività di assistenza legale", che offre una classificazione strutturata di ciò che i team di assistenza legale e gli utenti svolgono effettivamente. È stata sviluppata attraverso gruppi di lavoro continui, l'osservazione dei flussi di lavoro e sessioni di brainstorming con professionisti provenienti da diverse organizzazioni di assistenza legale con sede negli Stati Uniti, centri di assistenza giudiziaria e portali statali. La tassonomia organizza le attività in 7 categorie generali che abbracciano l'intero percorso giudiziario, dalla richiesta iniziale di assistenza fino alla risoluzione del caso, nonché le funzioni organizzative a supporto dell'erogazione dei servizi. Il quadro di riferimento viene offerto come punto di partenza strutturato che necessita di ulteriore convalida: una prima bozza per rendere esplicito, concreto e condiviso il vocabolario implicito delle attività dell'assistenza legale. Articolando questo vocabolario, la tassonomia mira a consentire una valutazione più rigorosa, a supportare una progettazione e una valutazione più chiare della tecnologia, a rafforzare l'interoperabilità tra giurisdizioni e modelli di servizio e a fornire una base per valutare i sistemi di intelligenza artificiale alla luce delle realtà del lavoro legale in prima linea.

Keywords: Access to Justice; Legal Aid; AI; Taxonomies

Summary: 1. Introduction. – 2. Background on Taxonomies in Law. – 2.1 Mapping Legal Tasks for AI development. – 2.2 Evaluation and Policy Discussions of Specific Legal Tasks & AI. – 2.3 Building from Other Legal Taxonomies. – 2.4. European Legal Ontologies and Access to Justice Frameworks. – 2.5 Learning from Task Standardization Across Domains. – 2.6 Gaps and Contribution. – 2.7 Taxonomy Development Process. – 3. Proposed Legal Help Task Taxonomy. – 4. Regulatory Implications for European AI and Data-Protection Frameworks. – 5. Preliminary Coding Exercise. – 5.1 Vignette Development. – 5.2 Coding Procedure. – 5.3 Results of Task-Level and Category-Level Agreement. – 5.4 Disagreement on Coding. – 5.5 Multi-Coding one Vignette with task labels. – 5.6 Taxonomy Gaps. – 5.7 Coder Background, Prior Exposure, and Agreement. – 5.8 Discussion of Coding Results. – 6. Future Evaluation and Validation Plan. – 6.1 Subject Matter Expert (SME) Review Across Diverse Organizations. – 6.2 Cognitive Task Walkthroughs and Workflow Replays. – 6.3 Cross-Jurisdiction and Cross-Organization Consistency Testing. – 6.4 Metrics for Validation. – 6.5 Iterative Revision and Publication. – 7. Conclusion.

1. Introduction.

Around the world, civil legal problems affect billions of people, yet the majority never receive meaningful help.[1] Legal aid organizations, court self-help centers, hotlines, and pro bono clinics work to get people advice, assistance, and representation, most often in manual and bespoke ways. This day-to-day legal help work is complex, heterogeneous, and poorly documented. The tasks and work being done with clients is not structured or documented in a way that can make it easily automated. That said, as interest grows in applying artificial intelligence (AI) to legal help and court services, the field could use a common understanding of the specific tasks that frontline professionals and clients perform across the justice journey. Without such a shared vocabulary, researchers cannot compare interventions, technologists cannot build aligned tools, and policymakers cannot evaluate what types of support are most needed.

Socio-legal research has described in detail how legal aid and court help systems operate,[2] but these accounts of what lawyers do to help people remain narrative and organization-specific. Existing legal ontologies – such as the Legal Issues Taxonomy (LIST) [3] or Akoma Ntoso – [4] focus on legal problems documents, not on the underlying work being done. Similarly, broader classification systems in other fields – like in health, human services, and AI research – demonstrate that if the field can produce a shared task taxonomy, they can better coordinate, benchmark, and develop responsible technology.[5] But in the legal field, we have not identified an openly available, cross-organizational taxonomy focused specifically on the granular tasks of providing frontline civil legal help.

This paper proposes a new contribution to address this gap: the Legal Help Task Taxonomy, a proposed classification of the tasks performed by legal aid staff, court navigators, community justice workers, legal clinic teams, and the people they serve. The task taxonomy was developed through ongoing working groups with legal aid organizations (including workflow demonstrations, practitioner walkthroughs, sustained design collaboration, and hands-on

---

[1] RL Sandefur, *Accessing Justice in the Contemporary USA: Findings from the Community Needs and Services Study* (American Bar Foundation 2019); World Justice Project, 'Measuring the Justice Gap' (2019) <https://worldjusticeproject.org/our-work/research-and-data/measuring-justice-gap> accessed 2026.

[2] R Engler, 'Connecting Self-Representation to Civil Gideon: What Existing Data Reveal About When Counsel is Most Needed' (2010) 37 Fordham Urban Law Journal 37; BH Barton and S Bibas, *Rebooting Justice: More Technology, Fewer Lawyers, and the Future of Law* (Encounter Books 2019); M Hagan, 'A Human-Centered Design Approach to Access to Justice: Generating New Prototypes and Hypotheses for Intervention to Make Courts User-Friendly' (2018) 6(2) Indiana Journal of Law and Social Equality.

[3] M Hagan and Stanford Legal Design Lab, 'Legal Issues Taxonomy (LIST): A Shared Vocabulary for Civil Legal Problems' <https://taxonomy.legal> accessed 2023.

[4] M Palmirani and F Vitali, 'Akoma-Ntoso for Legal Documents' in G Sartor and others (eds), *Legislative XML for the Semantic Web* (Springer Netherlands 2011) 75.

[5] World Health Organization, 'International Classification of Health Interventions (ICHI)' <https://www.who.int/standards/classifications/international-classification-of-health-interventions> accessed 2021; Open Referral Initiative, 'Human Services Data Specification (HSDS) v3.0' <https://openreferral.org/hsds> accessed 2020.

engagement with legal help processes) and synthesis of existing descriptive accounts in the socio-legal literature. The process was qualitatively rich but informal. Accordingly, this paper presents the taxonomy as a structured proposal, not a validated standard. In this paper, I focus on articulating the task taxonomy framework, the rationale behind its construction, and its potential uses for AI development, evaluation, and regulatory compliance.

The taxonomy aims to support a more interoperable, evidence-based foundation for the future of legal help innovation. A shared task vocabulary – even a provisional one – can help the field be more transparent about what legal help organizations and users do. It can support cross-jurisdiction collaboration, and provide more precise language and targets for evaluating AI systems. Ideally a task taxonomy aligns the work of researchers, technologists, and service providers. The contribution is also timely in light of European regulatory frameworks, and the demands they make of AI in the legal sector. The EU Artificial Intelligence Act classifies certain justice-sector AI deployments as high-risk depending on their intended purpose, institutional user, and decision context. This risk classification imposes system- and use-specific obligations that make it valuable to decompose complex legal workflows precisely into specific tasks, users, and decision contexts. These requirements make a structured task taxonomy for legal help work potentially valuable for compliance analysis and documentation.

2. Background on Taxonomies in Law.

The problem of defining what legal help professionals and consumers actually do has long been a subject of socio-legal research, but this day-to-day work has not been formalized in computational or interoperable ways. Foundational empirical studies have documented the practical and emotional labor of legal aid lawyers, court self-help staff, and community advocates, showing that their work extends well beyond doctrinal analysis to include triage, counseling, translation, and procedural navigation.[6]

A substantial body of descriptive and evaluative work details the frontline activities in court help centers and legal aid programs. Court navigator studies synthesize qualitative interviews and program descriptions across multiple states, identifying routine tasks such as orienting self-represented litigants, wayfinding, procedural guidance, referrals, document help, and in-court support.[7] On the funding and reporting side, the Legal Services Corporation's Case Service Reporting (CSR) framework defines service levels and closure categories (e.g., information, counsel and advice, limited action, extended/full representation) that implicitly describe bundles of different tasks that lawyers

[6] RL Sandefur, 'Elements of Professional Expertise: Understanding Relational and Substantive Expertise through Lawyers' Impact' (2015) 80 American Sociological Review 909; Engler (n 2); Barton and Bibas (n 2); Hagan (n 2); DJ Greiner and C Pattanayak, 'Randomized Evaluation in Legal Services: What We Have Learned and Why It Matters' (2019) 128 Yale Law Journal 1632.

[7] ME McClymont, 'Nonlawyer Navigators in State Courts: An Emerging Consensus' (National Center for State Courts/SRLN 2019) <https://www.srln.org/node/1403/reportnonlawyer-navigators-state-courts-emerging-consensus-mcclymont-2019>; ME McClymont, *Nonlawyer Navigators in State Courts: Part II—An Update* (Georgetown Law 2023).

and paralegals do, and the outputs that they create. [8] These sources demonstrate that moving past a hazy term of 'lawyering' to a concrete, observable, specific task list is both possible and necessary for improvement.

Still, these studies, frameworks, and accounts provide qualitative descriptions of the work lawyers do rather than structured, technology-friendly representations of tasks. As a result, researchers, professionals, and technologists lack a consistent way to compare interventions, to measure performance of the tasks, or design and share consistent, interoperable tech systems that align with real service delivery.

2.1. Mapping Legal Tasks for AI development.

Recent work on generative AI in law has begun to move towards an articulation of the concrete tasks that different legal professionals might delegate to or share with AI systems. Rather than treating "AI for legal work" as a single undifferentiated task, researchers, bar associations, and policy bodies are breaking apart the concept of 'legal work; into specific activities---such as legal research, document drafting, client communication, triage, translation, and data extraction. They do this to identify opportunities and to ask which of these are appropriate, effective, and ethical for AI interventions.

Survey- and interview-based studies of lawyers in private practice and in-house roles have enumerated a core cluster of AI-enabled tasks that happen in law practice: legal research, document review, contract analysis, drafting, and summarization of lengthy materials. The 2025 Future of Professionals report from Thomson Reuters, based on over 2,000 professionals, explored legal professionals' expectations of AI tools and pulled apart legal workflows to focus on high-opportunity tasks like legal research and drafting work, strategy, client counseling, and negotiation.[9] Similarly, legal industry analyses and bar association reports describe adoption of generative AI, and in their reports, decompose legal work into tasks like first-draft briefs, discovery responses, knowledge-management queries, and matter-intake support. [10] This line of work is largely descriptive, cataloging how lawyers say they are using AI and which tasks they perceive as lower-risk. While it is decomposing the broad category of 'legal work' into more specific sub-categories, it tends to group activities into broad categories (research, drafting, review, due diligence) rather than constructing a fine-grained task taxonomy of highly specific 'sub-sub-' categories.

In the civil legal aid context, Chien and Kim's 2025 field study of generative AI use in legal aid organizations offers one of the most detailed accounts of

[8] Legal Services Corporation, 'Case Service Reporting (CSR)' <https://www.lsc.gov/i-am-grantee/grantee-guidance/lsc-reporting-requirements/case-service-reporting> accessed 2024.
[9] Thomson Reuters, 'Future of Professionals Report 2025' (Industry Report 2025); M Richter, 'How AI is Transforming the Legal Profession' (*Thomson Reuters Legal Blog* 2025); J Feldman, 'The AI Adoption Divide: The 2025 Future of Professionals Report' (*Attorney at Work* 2025).
[10] American Bar Association, 'The Legal Industry Report 2025' (*ABA Law Practice Division* 2025); R Levine, '3 Innovative Ways AI May Be Used in Legal Practice' (Law360 Pulse 2024); Economics Observatory, 'How is Generative Artificial Intelligence Changing the Legal Profession?' (*Economics Observatory* 2024).

tasks that frontline advocates envision for AI support.[11] Drawing on interviews and a pilot deployment, the field study documents over 100 AI use cases in legal organizations, including: drafting demand letters and advice emails; summarizing case files; explaining legal documents in plain language; triaging and categorizing incoming intakes; creating checklists and timelines; translating client communications; and generating outreach materials. This field study gets into meaningful, specific detail of tasks that happen in civil legal aid offices, moving the field towards the kind of task taxonomy that can be useful for research and development.

Close by to civil legal aid work, empirical work in the public defender context has begun to map out specific legal tasks done by criminal legal help teams, in order to explore where AI might assist and where there should be boundaries. Cheong et al. conducted semi-structured interviews with fourteen public defenders across the United States to examine their experiences with AI, anticipated applications, and ethical concerns.[12] They found that defenders perceived the greatest potential in information-intensive, back-end tasks such as evidence investigation (e.g., processing large volumes of body-worn camera footage, jail calls, or digital discovery), targeted legal research, and certain forms of drafting and summarization. In contrast, they see far less scope for AI in tasks that turn on contextual judgment, trust, and relational work, such as courtroom advocacy, case strategy, and core client counseling.[13] Their study contributed a task-level map of public defense work, including the delineation between tasks that are more about information-processing and versus judgment-application.

This emerging literature on AI for legal services has moved the field toward thinking about legal work in specifics, with task descriptions and categorizations. But these surveys and field studies that break the concept of lawyering into specific tasks have so far remained tool- or project-specific. What the field lacks is a shared, cross-organizational, jurisdiction-agnostic taxonomy that systematically breaks the work that legal teams do into very specific, measurable tasks of work.

2.2. Evaluation and Policy Discussions of Specific Legal Tasks & AI.

Another area of work – that of evaluating and policy-making around AI in legal services – has also led to articulation of specific legal tasks. In legal AI evaluation, the LegalBench project and its follow-ons contribute collaboratively built task suites for legal reasoning and retrieval-augmented generation.[14] This work by LegalBench advances comparable measurement of

[11] CV Chien and M Kim, 'Generative AI and Legal Aid: Results from a Field Study and 100 Use Cases to Bridge the Access to Justice Gap' (2025) 57 Loyola of Los Angeles Law Review 903.
[12] I Cheong and others, 'How Can AI Augment Access to Justice? Public Defenders' Perspectives on AI Adoption' (arXiv 2025) arXiv:2510.22933.
[13] Cheong and others (n 12).
[14] N Guha and others, 'LegalBench: A Collaboratively Built Benchmark for Measuring Legal Reasoning in Large Language Models' (arXiv 2023) arXiv:2308.11462; N Guha and collaborators, 'LegalBench Datasets' (*Hugging Face* 2023); N Pipitone and GH Alami, 'LegalBench-RAG: A Benchmark for Retrieval-Augmented Generation in the Legal Domain' (arXiv 2024) arXiv:2408.10343.

technical systems across organizations and researchers. Kapoor, Henderson, and Narayanan categorized legal AI tools into groups by tasks, including (i) information-processing tasks (e.g., summarization, retrieval, classification); (ii) tasks involving creativity, reasoning, or judgment (e.g., drafting filings, negotiation, adjudication); and (iii) predictive tasks (e.g., risk assessments, outcome prediction).[15] This categorization then points to different levels of difficulty in evaluation, possibility of errors, or the stakes of deploying automated systems in these types of tasks. Simshaw, in his analysis of AI use in legal work, delineated tasks including triage, routing, document completion, and guidance, highlighting them as requiring special care to avoid embedding structural inequities.[16]

Courts and regulators, in their policy-making functions, have issued task-specific guidance about AI use. Judicial guidance in England and Wales identifies potential specific legal tasks for AI in judicial teams, such as summarization and administrative assistance, while still emphasizing that judicial officers must take particular care concerning legal accuracy, confidentiality, and the limitations of public AI systems.[17] The Federal Court of Australia has issued a practice note addressing the responsible use of generative AI in court-related work, including practitioners' duties concerning accuracy, verification, confidentiality, and disclosure in specified circumstances.[18] These policy discussions break the legal work that happens in court-based legal teams into sub-categories, while they are breaking apart this work In order to distinguish between lower-risk and higher-ones.

Taken together, this literature shows that the conversation about AI in legal services and court services is moving the field towards thinking about lawyer and judicial work in terms of tasks. Researchers and practitioners, as they are scoping out AI use cases, have been shifting from discussing 'legal work' as a whole, and moving into specific sub-categories of work like into for research, drafting, document explanation, translation, triage, intake, trend analysis, and more. In practice, public defender offices and legal aid groups are scoping out AI tools scoped to specific workflows. In policy-making, courts and regulators are examining the types of work that go on in their organizations, and then are drawing task boundaries around what is acceptable. These growing explorations of 'what legal teams do' often occur in project-specific, jurisdiction-specific, or sector-specific silos. That is what is missing – a unified, reusable, formal classification of legal tasks that could empower even more high-quality research and development of technology for frontline civil legal help.

---

[15] S Kapoor, P Henderson and A Narayanan, 'Promises and Pitfalls of Artificial Intelligence for Legal Applications' (arXiv 2024) arXiv:2402.01656.

[16] D Simshaw, 'Interoperable Legal AI for Access to Justice' (2025) 134 Yale Law Journal Forum 795.

[17] Courts and Tribunals Judiciary, 'AI Guidance for Judicial Office Holders' (December 2023, updated October 2025) <https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-guidance-for-judicial-office-holders/>.

[18] Federal Court of Australia, 'Generative Artificial Intelligence Practice Note' (GPN-AI, April 2026) <https://www.fedcourt.gov.au/law-and-practice/practice-documents/practice-notes/gpn-ai> accessed 2026.

This paper begins to address this gap, with a proposal for a Legal Help Task Taxonomy. Though still in a conceptual form, awaiting further validation, this first draft of a structured, cross-context vocabulary of tasks performed by legal aid teams, court help staff, and the people they serve, can lead to a foundation on which future sustainable, strategic AI work can build. A task taxonomy can help clarify which tasks are candidates for AI support, which require heightened safeguards, and how different organizations can compare and coordinate their AI development work. In doing so, this task taxonomy connects emerging discourse on categories of legal work and task-focused AI projects and policy, with a shared conceptual framework that supports responsible, interoperable innovation in justice ecosystems.

2.3. Building from Other Legal Taxonomies.

Standardized classifications have already been developed in many other areas of legal research and work. The Legal Issues Taxonomy (LIST), maintained by the Stanford Legal Design Lab, provides a hierarchical vocabulary of legal problems and sub-issues that ordinary people face, and it is used by statewide portals and chatbots.[19] Akoma Ntoso (LegalDocML) establishes XML standards for legislative and judicial documents to enable interoperability and provenance tracking, so that the substance of law can be more structured and easy to build on top of.[20] In law firm and corporate practice, the Uniform Task-Based Management System (UTBMS) supplies billing task codes and activities for budgeting and analytics.[21] The Standards Advancement for the Legal Industry (SALI) Alliance's Legal Matter Specification Standard (LMSS) is a comprehensive and growing ontology that contains over 10,000 machine-readable tags for classifying legal matters, services, forums, documents, and entities. SALI provides a shared, interoperable vocabulary that many law firms, vendors, and enterprise clients have used to consistently the corporate legal work that they perform.

2.4. European Legal Ontologies and Access to Justice Frameworks.

The European legal informatics tradition has produced ontological work that both informs and contextualizes the vision of a central task taxonomy of legal work. The Legal Knowledge Interchange Format (LKIF), developed in the EU ESTRELLA project,[22] provides a core ontology of legal concepts – norms, roles, actions, documents – designed to enable a cross-jurisdiction interchange between legal knowledge systems within the EU. The European Legal Taxonomy Syllabus (ELTS), developed at the University of Turin,[23] takes a

[19] Hagan and Stanford Legal Design Lab (n 3).
[20] Palmirani and Vitali (n 4).
[21] American Bar Association Section of Litigation, 'Uniform Task-Based Management System (UTBMS)' <https://www.americanbar.org/groups/litigation/resources/uniform-task-based-management-system/> accessed 2025.
[22] R Hoekstra and others, 'The LKIF Core Ontology of Basic Legal Concepts' in P. Casanovas and others (eds), *CEUR Workshop Proceedings* (Vol. 321, 2007) 43.
[23] G Ajani and others, 'The European Legal Taxonomy Syllabus: A Multi-Lingual, Multi-Level Ontology Framework' (2016) 11(4) Applied Ontology 325.

different approach: a lightweight, multi-lingual, multi-jurisdictional terminological vocabulary designed for practitioners to use in doing comparative law work. ELTS explicitly rejects an approach based on strong formalization and structure, in favor of flexibility and practitioner legibility. This lightweight design philosophy that resonates with the approach taken here, with the proposed Legal Help Task Taxonomy. Akoma Ntoso (LegalDocML), discussed above, likewise originates from the European legal-informatics community. These projects establish a European tradition of building interoperable legal knowledge infrastructure, though their focus has not yet been on the work activities of legal help teams. The present taxonomy is complementary to these ontologies: it addresses what people do rather than what legal concepts or documents exist.

European lawmakers have also pointed for the need of more specific discussions of legal work, and introduced laws and norms that encourage this standardized approach. The Council of Europe's Commission for the Efficiency of Justice (CEPEJ) adopted the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems in December 2018,[24] which establishes five principles (i.e., respect for fundamental rights, non-discrimination, quality and security, transparency and impartiality, and user control) that are directly relevant to AI deployments for specific legal tasks. Directive 2003/8/EC establishes minimum standards for cross-border legal aid within the EU,[25] with a category-specific framework of eligibility determination, application processing, and service delivery.

It is important to note that civil law jurisdictions organize legal aid differently from the US model that informed this Legal Help Task Taxonomy taxonomy's development. Systems such as Germany's Prozesskostenhilfe, France's aide juridictionnelle, and the Netherlands' Raad voor Rechtsbijstand operate through court-administered or state-funded mechanisms, unlike the US's non-profit legal-services organizations, though perhaps with similarities to court-based self-help centers. The proposed taxonomy has not been tested against these civil law systems. However, the taxonomy's categories describe work at a functional level – eligibility screening, intake, document preparation, procedural navigation, case coordination – that aims to be analogous across procedural traditions or financial models. Whether the specific task definitions transfer cleanly to civil law settings from the common law one is an empirical question that future cross-jurisdictional validation can resolve.

2.5. Learning from Task Standardization Across Domains.

If we turn to other disciplines, groups that have established task and workflow taxonomies have enabled scaling data collection, benchmarking, and automation. The World Health Organization's International Classification of Health Interventions (ICHI) and related service taxonomies describe very

[24] CEPEJ, 'European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment' (Council of Europe, adopted 3–4 December 2018).
[25] Council Directive 2003/8/EC of 27 January 2003 to improve access to justice in cross-border disputes by establishing minimum common rules relating to legal aid for such disputes [2003] OJ L26/41.

specific provider activities in clinical and community settings, thus promoting consistent measurement and research across jurisdictions.[26] The Civic Services Data Standards Alliance and Open Referral specifications [27] define interoperable schemas for human services directories of 'which organizations provide what services and events', paralleling legal aid's challenge of mapping heterogeneous local programs, and helping to build more interoperable, higher-coverage service directories and referral systems.

2.6. Gaps and Contribution.

As discussed, there are legal standards and ontologies have emerged to help bring together people across organizations to work better, more consistently together. The existing literature on information-infrastructure design emphasizes that shared ontologies and metadata standards are preconditions for reproducible research, policy coordination, and innovation. [28] But this standardization work is yet to capture the granular categories of what goes on in legal help services work. The absence of a shared task-level schema limits how much case management systems, analytics, evaluation, and tool development. inhibits interoperability between case management systems, analytics, evaluation frameworks, and emerging AI tools. This project's new Task Taxonomy addresses this gap by synthesizing descriptive socio-legal research with practitioner-grounded qualitative work to propose this specific categorizations of the work that legal teams do.

2.7. Taxonomy Development Process.

The taxonomy presented here emerged from qualitative, practice-grounded work conducted at the Stanford Legal Design Lab over several years, rather than from a single formal study. It reflects multiple streams of sustained engagement with frontline legal help practitioners.

The primary source of the taxonomy was an 18-month cohort convened by the Lab, comprising representatives from statewide legal help websites, state bar access to justice programs, legal aid organizations, legal content managers, law librarians, and legal strategists across 7 states. Over the course of the cohort, in facilitated workshops, participants inventoried the tasks their staff and users perform. They refined descriptions through successive working sessions, and identified tasks that had been overlooked or mislabeled. The taxonomy's categories and task definitions were iteratively developed through these conversations. The conversations were not based just on 'what do you want AI to do' but also what are the workflows that you and your team members do regularly?

A second source was biweekly collaboration with two legal aid organizations – the Legal Aid Foundation of Los Angeles (LAFLA) and Legal Aid Services of Oklahoma (LASO) – during 2025 and 2026 on justice AI copilot design for

[26] World Health Organization (n 5).
[27] Open Referral Initiative (n 5).
[28] GC Bowker and SL Star, Sorting Things Out: Classification and Its Consequences (MIT Press 1999).

housing and reentry services. During these interviews and conversations, the Lab team worked with legal team members to profile and pull apart the tasks that they do in serving their users and where technology could possibly assist them. These close working relationships provided detailed, ongoing visibility into the teams' granular frontline workflows including intake, document preparation, client coaching, evidence gathering, and case management.

A third source of task cateogries was a working session convened with the Self-Represented Litigation Network in March 2023, subsequently documented in a public report,[29] in which over 100 justice professionals – court staff, legal aid attorneys, navigators, technologists, and administrators – brainstormed a wide range of tasks that AI could support in legal help settings. The session was framed as 'where could AI help in the work that you do'? The session produced a broad inventory of activities spanning public-facing help, case work, organizational operations, and technology development.

Additional inputs about what specific tasks lawyers do came from the Lab's presentations and feedback sessions at the Equal Justice Conference (EJC) and the Innovations in Technology Conference (ITC), as well as the structured conversations with three legal aid organizations that partnered with the Lab in the project-based class AI for Legal Help – the Legal Aid Society of San Diego (LASSD), the Legal Aid Society of San Bernardino (LASSB), and LASO – in which legal team staff met with students and Lab staff every two weeks for 18 weeks, describing their service delivery models, workflows, and task structures in detail. These class discussions were about scoping exactly what the legal teams currently do, and where technology could assist them.

The Lab team then consolidated the brainstormed categories of legal work tasks through informal coding: grouping functionally equivalent tasks across organizations, splitting compound ones, and clustering tasks into categories that reflected the patterns practitioners described. The 7 top-level categories correspond roughly to who is typically doing this task (the person or the provider, though this changes with consumer behavior); the phases of a person's justice journey (initial help-seeking, service entry, substantive case work, ongoing case coordination; and the organizational functions that support service delivery (content production, administration, technology).

Admittedly, this process of identifying specific legal help tasks was qualitatively rich but informal. It did not follow a predetermined protocol, employ structured coding schemes, or produce inter-rater reliability statistics. It emerged out of work that was about scoping out a common legal help AI agenda and specific projects; the Lab team did not set out to build a taxonomy, but this opportunity emerged out of our design and research work. Therefore, this initial version of the help task taxonomy should be understood not as a validated classification but as a structured proposal. It's a first attempt to make the implicit task vocabulary of legal help work explicit and shared. It has the potential to serve as a foundation that future empirical work can test, refine, and extend, to become a validated and useful taxonomy.

[29] M Hagan, 'Opportunities & Risks for AI, Legal Help, and Access to Justice' (2023) Legal Design and Innovation <https://medium.com/legal-design-and-innovation/opportunities-risks-for-ai-legal-help-and-access-to-justice-9c2faf8be393> accessed 2026.

This flexible, gradual approach resembles established ontology-engineering methodologies – including the iterative cycles of METHONTOLOGY,[30] and the emphasis on reusing existing resources found in the NeOn methodology[31] – but it did not adhere formally to these. It is closer in spirit to the European Legal Taxonomy Syllabus project, which adopts a lightweight, practitioner-oriented approach that emerged organically rather than through a formally axiomatized ontology process.

3. Proposed Legal Help Task Taxonomy.

The current Legal Help Task Taxonomy breaks the work that legal help providers and consumers do into 7 top-level categories. Each category contains multiple tasks that represent work done across legal aid, courts, pro bono projects, and community justice organizations. Every task includes (1) a name, (2) a brief description, (3) its category, and (4) a unique Task ID for interoperability and future benchmarking.

This section presents these 7 main categories of tasks; see the full first version of the taxonomy in Appendix A with codes and definitions. The goal was to design a shared vocabulary that could generalize across jurisdictions, and also accommodate different laws, procedures, and organizational models. As

We assembled this taxonomy with the aim of generalizability across jurisdictions to account for local variation in law, procedure, and service delivery models. It is not intended as a static list; rather it is offered as a starting point for empirical validation with subject matter experts and for future extensions as legal needs evolve and new technologies emerge. The initial draft of this taxonomy was shared at the Jurix 2025 conference; since then 4 tasks have been added based on continued practitioner engagement – Law Summarization (TS-02-07), Signatures (TS-04-10), Outcome Prediction (TS-04-12), and Appointment Booking (TS-05-07) – and one task has been renamed to better reflect practitioner usage (Data Anonymizer became PII Masking, TS-06-06). The full taxonomy as of this writing contains 50 tasks across the same seven categories. The current version (and new versions) are maintained at justicebench.org/task.[32]

Category	Summary Description
Getting Brief Help	Front-door informational and diagnostic tasks that help people understand their legal issues, identify risks, choose next steps, calculate deadlines, understand documents, and find the right resources or forms.

[30] M Fernandez-Lopez, A Gomez-Perez and N Juristo, 'METHONTOLOGY: From Ontological Art towards Ontological Engineering' (1997) AAAI Spring Symposium Series.
[31] MC Suarez-Figueroa, A Gomez-Perez and M Fernandez-Lopez, 'The NeOn Methodology for Ontology Engineering' in MC Suarez-Figueroa and others (eds), *Ontology Engineering in a Networked World* (Springer 2012).
[32] Stanford Legal Design Lab, 'Legal Help Task Taxonomy' (JusticeBench) <https://www.justicebench.org/task> accessed 2026.

Providing Brief Help	Tasks enabling legal experts to create, update, translate, review, or enhance public-facing legal information, ensuring accuracy, clarity, and accessibility at scale.
Service Onboarding	Tasks that gather, structure, analyze, and verify information to onboard a client into a service–covering intake, eligibility, conflict checks, payment processing, and preliminary legal analysis.
Work Product	Tasks involved in drafting, preparing, filing, researching, negotiating, and organizing the substantive materials that make a case–documents, narratives, filings, evidence, and legal arguments.
Case Management	Ongoing coordination tasks that move a case forward, including scheduling, procedural routing, filing compliance checks, service verification, reminders, and transcription of proceedings.
Administration & Strategy	Organization-level tasks supporting sustainability, funding, data management, trend analysis, knowledge systems, and anonymization for safe and effective operations.
Tech Tooling	Tasks for creating, maintaining, testing, and administering digital tools, platforms, forms, interviews, and websites that power modern legal help ecosystems.

Table 1: Category-Level Overview of the Legal Help Task Taxonomy

The taxonomy was drafted to be at a functional abstraction level, so it could apply to organizations in different regions, models, and staffing, but still be specific enough for technologists and researchers to work clearly on. The task categories describe, like a documentary film, what legal help teams, tools, and consumers do – gather information, assess eligibility, prepare documents, navigate procedures, manage cases – rather than how specific procedural systems organize these activities. Empirical validation across European jurisdictions is a priority for future work, and the taxonomy's modular design can accommodate jurisdiction-specific extensions for tasks tied to particular legal settings.

The 7 categories are pragmatic workflow groupings rather than mutually exclusive branches derived from a single ontological dimension. The distinction between Getting Brief Help and Providing Brief Help reflects a deliberate design choice grounded in the taxonomy’s focus on both users doing legal work and providers doing this work Getting Brief Help describes tasks from the consumer's perspective (where legal teams might not necessarily be involved): the person facing a legal problem receives information, understands a document, or is routed to a resource. Providing Brief Help describes the corresponding provider-side tasks: a legal expert writes the help guide, reviews it for accuracy, translates it, or monitors legal changes that affect its currency. These may look similar on paper but in reality they operate quite differently – they involve different actors who have different levels of emotions, expertise, and standards; they may require different skill sets; they likely need different quality standards; and there are very different failure modes. Separating these 2 sets of tasks was done to better build and evaluate consumer-facing and provider-facing workflows independently and to assess AI suitability

differently for each. Also worth noting, the taxonomy's task-level codes are designed to be composable rather than mutually exclusive: a single interaction may include tasks in multiple categories.

Tech Tooling occupies a different level from the other 6 categories, as it describes back-office, infrastructure tasks rather than direct legal help delivery. Its inclusion reflects a pragmatic design decision, born out of the interviews and workshops. These back-end technology creation and maintenance tasks still consume significant legal team staff time, shape the quality of all other task categories, and represent a primary site of AI development and experimentation. Excluding them would produce a taxonomy that omits a substantial portion of the work that legal help teams actually do, and our goal is to reflect the lived realities of the teams we work with.

- Getting Brief Help

The Getting Brief Help category captures the set of informational and diagnostic tasks that support a person's earliest engagement with the legal system. These tasks may be done by the person by themselves, or with a friend, advocate, navigator, or other provider. These tasks help individuals understand their situation, recognize whether their problem has legal implications, identify urgent risks, and map out early steps they should take. This category includes core public-facing tasks such as Legal Q&A, Document Explainer, Issue-Spotting, Deadline Calculator, Form Selection, Referral Routing, and Action Plan generation.

These tasks are typically brief, high-volume, and performed across many different platforms and services – websites, hotline interactions, walk-in clinics, chatbots, informal conversations, or community navigator consultations. Across legal aid organizations and court help centers, these tasks often are described as the "stage one" or the "front door" of legal help: they respond to people's immediate confusion, triage complexity, provide clarity, and empower users with accurate, actionable information. In AI planning workshops, many participants zone in on this set of tasks first, often because this category of tasks seem more straightforward, and there is such a high need and impact of these 'front door' activities when people are at the start of their justice journey. Good services in this zone of categories can shape whether a person takes the next step toward resolving their issue – or does not proceed and likely does not get a just result.

- Providing Brief Help

The Providing Brief Help category still focuses on this 'front door', 'start of the justice journey' phases, but it looks at them from the perspective of the provider who is asked to produce, maintain, verify, and improve public-facing legal information that's being used in the 'get help' categories. In this second category, the delineated tasks include Help Guide Writing, Content Review, Legal Help Translation, Law Watching, Media Creation, Law Summarization, and tools that support professionals during brief consultations. As our providers discussed, these tasks require domain expertise, accuracy checking, and clear communication. They also include Q+A Supporter tools that augment professionals during brief advice consultations. Whereas the Getting Brief Help

category focuses on supporting the public directly with orientation and direction, the Providing Brief Help category focuses on supporting legal professionals as they prepare and deliver authoritative public-facing guidance. These Providing Help tasks involve expert judgment, accuracy checking, and clear communication. Providers highlighted how essential these tasks are, to ensure the quality and accessibility of the information ecosystem that powers search, AI responses, hotline conversations, volunteers and navigators training, and other brief help interactions. As organizations increasingly delegate to navigators, justice workers, and community members, this Providing Help category of tasks ensure that there is a trustworthy backbone for scalable, high-quality public legal education.

- Service Onboarding

Service Onboarding refers to the tasks performed by service providers, as they are transitioning an individual from general information-seeker to an enrolled or screened client of a legal or court service. These tasks, often performed by intake workers, clerks, hotline operators, or paralegals, require structured information-gathering, risk identification, and eligibility determination. This is crucial for grant compliance, triaging and resource allocation, and liability protections. The tasks that occur in service onboarding include Intake Interviewing, Document Issue-Spotting, Data Research, Legal Analysis, Conflicts Checking, Eligibility Verification, and Payments. These tasks are typically where organizations decide whether and how they can help. They carry out these tasks so they can determine the procedural posture of a case, the level of assistance appropriate for the client, the eligibility for this or other organizations' services, and the internal routing within the organization. Onboarding tasks must balance efficiency, accuracy, compliance, and trauma-informed interaction. Teams noted that these tasks often require specialized training to perform, but they also are increasingly being build on top of – and to feed back into – structured data systems. Service onbarding tasks must be carried out so that organizations can deliver services in fair, ethical, and compliant ways – and ideally to help the person in crisis get to the right help as soon as possible.

- Work Product

The Work Product category moves into the substantive lawyering tasks involved in creating legal documents, arguments, narratives, filings, negotiations, and evidence preparation. These activities tend to operationalize a user's or attorney's legal strategy, take more time, and involve more complex judgment. This category spans simple standardized, templated document completion (Standard Document Filler) to highly technical drafting (Expert Document Drafter), as well as Narrative Drafting, Contract Review, Document Draft Checking, Electronic Filing, Legal Research, Negotiation Support, Service of Documents, and the newly added Evidence Preparation. Participants both wanted to highlight the importance of these tasks, because of their central impact on case resolution and life outcomes – but also warn against blithe AI use in them. This category's tasks combine complex reasoning, jurisdiction-specific rules, storytelling skills, and attention to detail. Many litigants struggle

in this stage of the justice journey, when they are tasked with creating legal work product or proceeding through a negotiation. This burden makes the Work Product category a core area in which to explore technology systems. For legal professionals, the category raises huge concerns around quality, consistency, and efficiency. For consumers, performance on these tasks will shape whether they can meaningfully participate in their case or get a fair outcome.

- Case Management

The Case Management category moves to the procedural and coaching tasks that go into resolving a legal problem, including the ongoing coordination, monitoring, and follow-through needed to move a case forward. The Case Management tasks include Support Coaching, Smart Scheduling, Procedure Triage, Filing Screening, Service Verification, and the new task Transcribe Proceedings. Legal team members wouldn't necessarily describe these tasks when asked to describe what lawyers do, but upon more reflection they emerged as invisible yet critical to case progress. These tasks of getting deadlines met, people to come to meetings on time, checking documents for acceptable compliance, the scheduling of hearings, and the correct service of documents on the other side all have major impact on case and life outcomes. For litigants, these tasks can prove to be the difference between participating fully or experiencing default judgments, missed filings, or procedural barriers. Effective case management can be an efficiency win for legal aid and court teams, and can help litigants get out of the justice system more quickly with better outcomes. Many high-volume court teams struggle with procedural bottlenecks and case management issues that stem from failures in this category, making it a prime area where AI is being explored.

- Administration & Strategy

The Administration & Strategy category moves to the back-office of legal teams. It covers organizational-level tasks that ensure legal aid, courts, and other justice programs operate sustainably and strategically. These tasks include OCR/Data Extraction, Grant Reporting, Knowledge Management, Service Directory Maintenance, Trends Spotting, and PII Masking. These tasks are performed to ensure that data is entered and processed correctly, that funds are spent in compliance with all requirements, and that the organization is learning and acting strategically. While these tasks are often performed by operations, IT, or supervisory staff rather than frontline legal teams, they are fundamental to the long-term impact of justice organizations. Strong performance on Administration & Strategy tasks enable evidence-based program design, cross-organizational interoperability, and the creation of maintainable, trustworthy data pipelines. They also enable organizations to produce the structured, compliant data needed for audits, evaluation, tech development, and AI-assisted service delivery.

- Tech Tooling

The Tech Tooling category emerged out of conversations with those in charge of creating, maintaining, testing, and improving technology systems

used by legal help organizations and court service programs. We can expect this category to grow in importance and tasks as more AI-powered programs enter use in the field. The category's tasks include Form Creation, Site Administration, Interview Creation, and User Testing. These tasks are increasingly essential as legal aid organizations deploy a mix of digital forms, guided interviews, chatbots, case management integrations, and self-help tools. Many legal help organizations lack dedicated software engineering teams, and have vendors or lawyers doing these tasks, making them high-leverage opportunities for AI-enabled automation. Tech Tooling tasks support efforts to make better digital systems, and to develop new tools that are user-friendly, secure, and well maintained.

4. Regulatory Implications for European AI and Data-Protection Frameworks.

Before moving to discussion of the taxonomy's validation, it is worth seeing how a help task taxonomy can be of use to emerging AI regulatory frameworks, which aim to manage the safe deployment of technology in the justice sector. In the European context, three instruments are particularly relevant: the EU AI Act, the GDPR, and the Law Enforcement Directive.

The EU Artificial Intelligence Act (Regulation (EU) 2024/1689)[33] classifies AI systems used in the justice sector as being 'high-risk', and deserving of special treatment, if they fall in certain task categories. For example, under Annex III, category 8(a) -- for AI systems intended to assist judicial authorities in researching and interpreting facts and law, or to be used similarly in alternative dispute resolution. Note that comparable systems used solely by independent legal aid providers would require separate analysis, about their risk level. The Act's Annex III category 5(a) covers AI systems used to evaluate eligibility for essential public services, which calls attention to the taxonomy's tasks of Eligibility Verification (TS-03-06) and Referral Routing (TS-01-02). These tasks may implicate this provision when operated by or on behalf of public authorities administering benefits programs. With high-risk obligations taking effect from August 2026, a task-level vocabulary like that provided with this proposed taxonomy could help organizations identify which specific activities involve AI and map them to the risk-management, data-governance, transparency, and human-oversight requirements of Articles 9 through 15. We do not claim the taxonomy was designed for this purpose, but we note the alignment – and likely the use in future regulations that are task-based and risk-based.

The General Data Protection Regulation (GDPR)[34] imposes requirements on tasks that involve processing personal data. Article 22 may apply when these

[33] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L 2024/1689.

[34] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) [2016] OJ L119/1.

tasks are implemented through decisions based solely on automated processing that produce legal effects or similarly significantly affect an individual – for example, a fully automated denial of service eligibility. Tasks like human-reviewed intake assistance or preliminary conflict flagging (both in the Service Onboarding category of the taxonomy) may not necessarily trigger Article 22, although other GDPR requirements would continue to apply. Articles 5 and 6 (lawfulness, purpose limitation, data minimization) apply across the taxonomy wherever personal data is collected or processed. Article 9, which generally prohibits processing specified special categories of personal data unless an applicable exception is available, is relevant to tasks where health information, racial or ethnic origin, or other special-category data may be disclosed – particularly during tasks in Service Onboarding and Work Product creation, though it might happen in Getting Brief Help tasks as well.

The PII Masking task (TS-06-06) warrants particular attention as it applies to regulatory requirements. The GDPR draws a sharp distinction between anonymization, which places data outside the Regulation's scope, and pseudonymization, which does not (Article 4(5)). The taxonomy's listed PII Masking task – which involves transforming confidential data into usable datasets by removing identifiers – must be understood in light of this distinction. Whether a particular masking operation produces truly anonymous data (outside the GDPR's scope) or merely pseudonymized data (still subject to the full GDPR framework) has significant legal consequences, and because of this European deployments related to this task must engage in this careful distinction. Although the taxonomy currently uses the term PII Masking, European deployments should interpret the task in terms of personal data, pseudonymization, and anonymization under the GDPR. This is also a point of potential further taxonomy refinement, to make it more usable for clear regulatory work.

Another European regulation, the Law Enforcement Directive (Directive (EU) 2016/680), [35] applies to tasks that cross into personal-data processing by competent public authorities for criminal-law purposes. The taxonomy's tasks like Evidence Preparation (TS-04-11), Data Researcher (TS-03-03), and Transcribe Proceedings (TS-05-06) are likely to fall outside the LED if they are done by independent defense attorneys or legal help organizations. If these same tasks are done by public authorities, the regulation may apply. This signals that the taxonomy of tasks must be combined with other contextual factors, like the exact organization type or setting to determine a policy's applicability.

Some other taxonomy tasks interact with normative guidance, like the CEPEJ Ethical Charter's five principles. The Charter's 'user control' principle aligns with the category of Case Management tasks. The Charter's 'quality and security' principle points to the importance of tasks like Content Reviewer (TS-02-04) and Document Draft Checker (TS-04-05) as safeguards against AI error.

[35] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data (Law Enforcement Directive) [2016] OJ L119/89.

For this policy area, the task taxonomy can be a way to make concrete plans for how an organization can put higher-level principles into practice.

As shown with this analysis of European policy and the task taxonomy, a consistent vocabulary offers regulators, technologists, and evaluators a practical starting point for mapping AI involvement to regulatory obligations and principles. Whether the taxonomy's current level of granularity is sufficient for formal conformity assessment, especially in a European context when it emerged out of largely US-based practitioner qualitative work, is a question that deserves future exploration.

5. Preliminary Coding Exercise.

Is the taxonomy clear and comprehensive enough to move towards to further validation? After drafting the initial version out of our qualitative work and receiving preliminary feedback, we then ran a test of an inter-rater coding exercise to begin the next, upcoming round of validation. Our team used activity vignettes derived from the legal help literature and from practitioner engagement, and had members of our team attempt to label them with the taxonomy's terms.

5.1. Vignette Development.

For the preliminary inter-rater coding exercise, we developed 62 activity vignettes, each describing a concrete activity or interaction in a legal help setting. Each vignette had 2 to 6 sentences of narrative prose, describing a scene of people doing some activity. The vignettes were drawn from 2 sources. The first source was the literature's published empirical accounts of legal help workflows, including McClymont's studies of court navigator programs, which document tasks such as helping litigants complete paperwork, explaining court procedures, physical wayfinding in the courthouse, emotional support during proceedings, making referrals, and translating documents; Chien and Kim's field study of generative AI in legal aid, which catalogues over 100 use cases including document summarization, intake triaging, plain-language translation, and content creation; Cheong et al.'s study of public defender workflows, which describes evidence processing (including body-worn camera footage and digital discovery), legal research, and drafting; and the LSC Case Service Reporting framework, which defines service levels from brief information to extended representation. This thorough documentation and listing of different activities helped us to then generate activity vignettes which told stories of very concrete actions being taken by legal teams, people with legal problems, and others in their orbit.

The remaining vignettes were developed by the research team based on activities directly observed or described by practitioners during the qualitative working sessions. We removed the names of providers or any other identifiable detail that was shared with us, and made fictional versions of the specific work actions that participants had described. Each vignette represents a composite

activity type documented in one or more of these sources. No vignette describes a specific client interaction, provider action, or identifiable individual.

We checked that the vignette activities spanned all 7 taxonomy categories and a range of legal issue areas including housing, debt collection, family law, public benefits, employment, immigration, consumer protection, reentry and expungement, elder law, disability, and foster care. We did not try to make vignettes that cleanly matched to a give task taxonomy category, instead we allowed activity vignettes that might be ambiguous or multi-task scenarios — for example, a single interaction in which a worker explains a document, identifies deadlines, and selects the appropriate form. We wanted to have these vignettes correspond to how practitioners describe the units of their work, to see if we could apply taxonomy codes in consistent ways to the descriptions future practitioners and technologists are likely to have, when thinking through what exact category of work they are talking about.

We ensured that no vignette included a job title that would telegraph the taxonomy answer (such as 'knowledge manager' or 'intake specialist'). Instead we used random first names or neutral descriptions about staff members, people, or tools. The vignettes were presented in a randomized order unique to each coder.

5.2. Coding Procedure.

Three coders on the Legal Design Lab team independently labeled each of the 62 vignettes using a web-based tool. First, the coder saw the vignette story. Then the coder was asked: 'What task is happening here? Choose the label that best fits.' The tool let the coder select 2 primary task code from the taxonomy (46 tasks across 7 categories at the time of the exercise, plus a 'no matching task' option for taxonomy gaps). The 4 tasks added subsequently – Law Summarization, Signatures, Outcome Prediction, and Appointment Booking – were not included in the coding exercise. After selecting a primary code, coders could continue to add additional task codes if they felt more tasks were being done beyond the first one. Coders could also leave an optional note explaining why they chose the codes they did. If they didn't think any code applied, they could select "No Matching Task." No discussion between coders occurred during the exercise. Each coder's completion of the labeling took between 20 and 40 minutes.

Agreement analyses used each coder's primary task selection only. We analyzed the optional additional task separately. We assessed pairwise agreement using raw percentage agreement and unweighted Cohen's kappa. At the task level, the possible response categories consisted of the 46 task codes and the "no matching task" response, which was treated as an additional category. For the category-level analysis, each primary task selection was mapped to its corresponding top-level category ("no matching task" remained a separate response category). This category mapping was deterministic because each task code belonged to one of the 7 top-level taxonomy category. Because some coders submitted revised entries, the analysis used the most recent response recorded for each coder and vignette.

The 3 coders on the team had different professional backgrounds and levels of familiarity with the taxonomy. The first author (C1) developed the taxonomy and has been immersed in the practitioner engagement from which it emerged. A second coder (C2) is a lawyer with extensive social science research training, who participated in some of the working group sessions and practitioner conversations that informed the taxonomy but was not involved in its formal construction. A third coder (C3), an undergraduate student research assistant, has worked on legal service directory data and legal help projects but has no formal legal training and had not previously studied the taxonomy in detail.

5.3. Results of Task-Level and Category-Level Agreement.

After the 3 independent coding sessions, we analyzed agreement at both the individual-task and top-level-category levels. Agreement statistics were calculated using each coder's primary task selection only.

Pairwise agreement on primary task codes varied among the three coder pairs. C1 and C2 agreed on 29 of 62 activities (47 percent), with an unweighted Cohen's κ of 0.45. C1 and C3 agreed on 21 activities (34 percent), with κ = 0.32, while C2 and C3 agreed on 23 activities (37 percent), with κ = 0.35. Across all three coders, unanimous agreement on the same primary task code occurred for 16 of the 62 activities (26 percent).

Agreement was higher at the 7-category level. C1 and C2 agreed on the top-level category for 42 of 62 activities (68 percent), with an unweighted Cohen's κ of 0.61. C1 and C3 agreed on 32 activities (52 percent), with κ = 0.41, while C2 and C3 agreed on 36 activities (58 percent), with κ = 0.48. Across all three coders, unanimous category-level agreement occurred for 28 of 62 activities (45 percent).

The difference between task-level and category-level agreement indicates that coders more often converged on the broad functional area of what was happening in an activity scene, and then disagreed in what exact task was primarily happening in the scene. These observed differences among coder pairs may reflect the differences in their legal-domain knowledge, familiarity with the taxonomy, prior participation in its development, differences in how task definitions were interpreted, or some combination of these factors.

Sixteen activities produced unanimous agreement, where the coders all felt the same task was occurring. These included: a hotline worker answering a tenant's question (Legal Q&A), identifying multiple legal issues in a caller's story (Issue-Spotting), a conflict-of-interest check (Conflicts Checker), drafting a motion for protective order (Expert Document Drafter), website maintenance by a statewide IT manager (Site Administrator), translating help guides (Legal Help Translator), a client organizing evidence for a hearing (Evidence Preparation), a data presentation on caseload trends (Trends Spotter), a worker explaining multiple legal documents to a confused person (Document Explainer), transcribing a hearing recording (Transcribe Proceedings), pulling structured data from old paper intake forms (OCR/Data Extractor), testing a new online tool for bugs and accessibility (User Tester), explaining a food benefits denial letter (Document Explainer), explaining options to an elderly man's daughter about financial exploitation (Legal Q&A), a restaurant worker's

wage-theft situation being documented (Evidence Preparation), and a foster youth aging out of the system getting a step-by-step checklist (Action Plan).

These activity vignettes all shared a common feature: they describe a single, well-defined actor performing a discrete, clearly bounded action with an unambiguous purpose. A conflict check is a conflict check; translating a guide into another language is translating a guide; transcribing a hearing is transcribing a hearing. The taxonomy is clearest when the task is operationally specific and the vignette does not involve multiple overlapping functions – where the scene the person is hearing about or watching is relatively clear, without multiple things happening at once.

5.4. Disagreement on Coding.

Twenty-one activities produced complete disagreement – no 2 coders chose the same primary code. These fell into several instructive patterns.

A prime area of disagreement involved vignettes where a staff member is analyzing documents that have just been brought into their team. A court clerk reviewing a self-represented litigant's filing for completeness (a14) was coded as Filing Screener (checking filings for compliance before acceptance – the function the clerk is performing), Document Issue-Spotter (a different interpretation – the taxonomy intends this task for reviewing an opposing party's documents, but the label did not make that distinction clear to the coder), and Document Explainer (a task that the taxonomy intended to be about helping a person understand a legal notice they have been given). A paralegal catching errors in her own organization's filing before submission (a29) was coded as Document Draft Checker (reviewing your own team's output before filing), Document Issue-Spotter (which the taxonomy defines as reviewing an opposing party's documents for defects), and Issue-Spotting (which the taxonomy defines as identifying legal issues in a person's story). The taxonomy draws a distinction between reviewing your own work and scrutinizing the other side's documents, but the vignette's description of 'catching errors' did not cue all coders to that distinction. This points to a problem in the draft taxonomy's naming conventions, which aimed to be short enough to be usable but could then elide important context about the task that the term was meant to cover.

Another hot spot of disagreement about primary task codes emerged with vignettes where a person or provider was doing a mix of procedural and substantive activities at the same time. A mother getting help from a legal team member on how to enforce a child support order (a47) was coded as Support Coach, Legal Q&A, and Action Plan – 3 different framings of essentially the same activity: someone helping a person figure out what to do next, offering coaching, legal answers, and a step-by-step plan in one scene. A man seeking expungement of his criminal record (a48) was coded as Document Issue-Spotter (reviewing the criminal history for eligible convictions), Form Selection (identifying the right petition form), and Action Plan. These all could be correct, depending on which part of the scene a coder was focused on. An attorney explaining an order to split pension assets in a divorce (a54), an interaction involving a client receiving assistance after an adverse hearing outcome (a59),

and a hotline worker explaining tenant privacy rights (a62) each split three ways across Getting Brief Help tasks. These disagreements occurred precisely because the vignettes describe realistic, multi-step interactions where multiple tasks are plausibly present. In future applications of a task taxonomy, any person who has to use its terms to code a 30-minute meeting or even a 6-minute interaction will likely grapple with this challenge: there often is not a single primary task happening, but rather the provider and user are doing a mix of things in a short time.

Several technology-related vignettes also led to three-way disagreement splits. Building a document-assembly tool's guided interview (a18) was coded as Interview Creator (building the interview logic), Form Creator (the output is a form), and Q+A Supporter (a different interpretation – the taxonomy defines this task as provider-facing consultation support, but the coder may have read it more broadly as any tool that supports question-answering). Summarizing an administrative decision using a tool (a60) was coded as Legal Analyzer, Legal Researcher, and Q+A Supporter. Those first two tasks identify the analytical function while the third identifies the tool-support function. A know-your-rights presentation for immigrant families (a56) split between Help Guide Writer (creating the content), Media Creator (producing the presentation and materials), and Action Plan.

Other disagreements to note: the vignette of reviewing a settlement offer with a client (a37) was coded as Negotiation Helper (helping the client weigh options and model outcomes), Legal Analyzer (analyzing the legal implications), and Document Explainer (explaining the offer's terms). Contract Reviewer may also have been a plausible fit under the current definition, yet no coder selected it. This may indicate that the task label or description does not clearly signal its application to settlement review. The courthouse wayfinding vignette (a57) was coded as Referral Routing, Support Coach, and no matching task, suggesting that physical wayfinding support is not clearly captured by the current taxonomy. This was flagged as a necessary addition to a future version.

5.5. Multi-Coding one Vignette with task labels.

How often did coders use more than one code to label a vignette? Multi-coding frequency differed among the coders. C1 and C2 each added multiple codes for 26 of 62 activities (42 per cent), while C3 did so for 12 activities (19 per cent). This difference may reflect domain knowledge, familiarity with the taxonomy, prior participation in its development, coding style, or some combination of these factors.

Some of the coders' task-level disagreements resolved when additional codes were considered; they did not always match on the primary code but did match more frequently with additional codes. For activity a06 (an attorney reviewing the opposing landlord's documents and identifying defenses), C1 coded Legal Analyzer as the primary task with Document Issue-Spotter as additional – recognizing both the legal analysis and the scrutiny of the other side's documents for defects (the notice was served on a Sunday, the lease clause may have been violated). C2 coded Action Plan as the primary task with Legal Analyzer as additional. Both coders identified Legal Analyzer as present;

they disagreed only about whether the analytical or strategic-planning aspect was primary. This pattern – agreement on the set of tasks present, disagreement about which is most central – recurred across many activity vignettes. It supports a future taxonomy deployment plan where multi-label annotation scheme is the default (perhaps even without any primary code), rather than a forced single-choice approach.

5.6. Taxonomy Gaps.

Was the first draft of the taxonomy comprehensive enough to cover the activities that are happening in legal organizations? In this preliminary exercise, coders, flagged 2 taxonomy gaps – and likely there will be additional ones with future coding exercises. C2 found no matching task for the courthouse wayfinding vignette (a57), in which a worker physically guides a first-time litigant through the building. C3 found no match for the e-filing vignette (a24), though both other coders coded it as Electronic Filer. In this instance, the improvement might be retitling or defining the Electronic Filer task. But the wayfinding gap is substantive: physical navigation and orientation support is well-documented in the court navigator literature but was not included in the first version of the taxonomy. It was added in Summer 2026, as "Provide Court and Service Wayfinding", TS-05-08.

5.7. Coder Background, Prior Exposure, and Agreement.

How did the different coders interact with the coding platform? C1 and C2 agreed more frequently and used additional codes more often than C3. There are several possible explanations for this different behavior, including greater legal-domain knowledge, greater familiarity with the taxonomy, shared participation in its development, or differences in how the task descriptions cued the coders. The current preliminary exercise cannot distinguish among these explanations. Future validation should therefore include a larger group of independent coders. With the coders, validation studies should delineate their legal-domain expertise and taxonomy familiarity, and calibrate before going to independent coding.

5.8. Discussion of Coding Results.

Reflecting on this preliminary coding, there are several findings that can help create an improved draft of the taxonomy and plan for more robust validation and implementation work.

5.8.1. The Category Structure Shows Stronger Agreement.

Category-level agreement between the 2 experienced coders reached 68 percent, and even the least experienced coder agreed on category at least half the time with each of the others. Three-way category agreement was 45 percent – nearly double the three-way task agreement of 26 percent. The taxonomy's 7 category top structure may correspond to people's intuitions about the types of legal work, and how to name what is happening in a given legal help scene. This preliminary evidence indicates that those 7 categories

may be a good scaffold to keep building on top of, working on the granular, specific task descriptions inside of them.

5.8.2. Task-Level Disagreement Reflects Both Overlapping Practice and Unresolved Boundary Definitions.

Task-level agreement was lower than category-level agreement, particularly when vignettes described scenes in which multiple overlapping activities were happening in the interaction. This pattern may reflect both the integrated, complex nature of frontline legal-help work and ambiguity in the taxonomy's specific task boundaries. A hotline operator answering a caller's question may, within a single 2-minute exchange, identify the legal issues in the caller's story, explain what a court document means, calculate a filing deadline, identify the correct form, and sketch out a set of next steps to follow. From the worker's perspective, this is one fluid activity, one unit of legal work. But labeling the precise tasks going on in this 2-minute exchange turned out to be challenging to do consistently. In our exercises. What's happening? Is it Legal Q&A, Document Explainer, Deadline Calculator, or Action Plan? In lived practice, the boundaries between these tasks are inherently blurry because a skilled worker performs them in an integrated, overlapping way. But for the sake of technology development, evaluation, and regulation, we might treat these various tasks differently, and we need to break apart this 2-minute interaction in a consistent, clear way to do so.

The taxonomy purposefully is not primarily designed to describe human work as it is experienced by the worker or how it might be encoded for a billing or grant schematic (of 'brief help – 30 minutes' or 'legal consultation – 15 minutes'). The taxonomy aims to get more specific, decomposing that work into discrete units that can be delegated to tools, used to write product requirements for new technology, evaluated for quality, and assessed for AI suitability. When an organization builds a chatbot, it needs to know what that chatbot is going to do. Which tasks should it engage with and which should it refuse to do? Ideally a task list can help with his, deciding if the chatbot should answer general legal questions, explain specific documents, calculate deadlines, or select forms. Each of these tasks within the chatbot interaction will require different data, different evaluation criteria, and different risk profiles under frameworks like the AI Act. The taxonomy's task-level specificity exists to serve this decomposition. But it turns out to be hard, at least in this exercise with the first version of the taxonomy, to get coders to consistently decide which specific, decomposed task is happening.

5.8.3. Legal Q&A as a Container.

The coding exercise suggests that Legal Q&A functions in practice as a general container – the broad interaction mode in which a person receives brief legal help – rather than as one task among peers. In the back-and-forth of posing a question and receiving an answer, other Getting Brief Help tasks – Document Explainer, Issue-Spotting, Deadline Calculator, Form Selection, Action Plan, and Referral Routing – seem to be happening within them. These tasks may be better understood as sub-types or components within a Legal Q&A interaction. Coders frequently assigned Legal Q&A as a primary code with

a more specific task as additional, or vice versa: both were perceived as present, but coders differed about which was most central. Originally, the taxonomy included Q&A as a task where the system was giving the person a short, tailored statement of how the law applied to their situation. But it seems to often overlap and include these other parts of explaining a notice, calculating when forms are due, choosing what form is best, and other tasks. Future versions of the taxonomy can make the hierarchical relationship explicit: Legal Q&A might become the parent interaction type, and the remaining tasks describe the specific kinds of help delivered within it.

The same logic applies to other boundary areas. The distinction between Document Issue-Spotter (Service Onboarding) and Document Draft Checker (Work Product) is about the creator of the document that's being reviewed, and what the purpose of the review is. Is it about scrutinizing the other side's documents for defects and leverage, or is it about checking your own team's output for errors before filing? These look similar on the surface but serve entirely different functions and would require different tool architectures. The coding exercise surfaced this blurriness and suggests the taxonomy needs sharper language in the terms' names and descriptions to support coders' application of the terms.

The overlap and disagreement around the tasks of Law Watcher, Content Reviewer, and Knowledge Management reflects 3 distinct functions: detecting that the law has changed, checking whether specific public-facing content is still accurate, and maintaining the organization's internal system of policies, templates, and expertise. These correspond to different system capabilities, but a single person may perform all three in sequence. That is likely why coders who watched someone do all three in a vignette disagreed about which was primary.

5.8.4. Multi-Coding Supports Further Testing of Composability.

The high rate of multi-coding in our exercise – 42 percent of activities for both experienced coders – is an important finding. The frequent use of additional codes is consistent with the compositional design of the taxonomy, although this result should be interpreted in light of the deliberate inclusion of ambiguous and multitask vignettes, that were meant to reflect how practitioners describe their work. Some apparent disagreements were reduced when additional codes were considered. This suggests that future annotation guidance should explicitly support multi-label coding and that evaluation frameworks built on the taxonomy should accommodate activities that span multiple tasks. Developers and evaluators can still decompose 'legal work' to very specific tasks, but more work is needed to make sure that people agree on what the specific task is that they're talking about.

5.8.5. Coder Background and Agreement.

C1 and C2 agreed more frequently and used additional codes more often than C3. As discussed above, it is not clear exactly why this pattern emerged, because factors about legal knowledge or taxonomy familiarity were not formally measured. Future validation should use a larger and more independent coder pool and separately measure domain expertise and prior

taxonomy familiarity. These steps can ensure better analysis to get to the root of disagreement or patterns

5.8.6. Limitations.

Three coders is a small sample, all came from the same Lab team, and one of the coders (C1) is the taxonomy's developer. The vignettes are researcher-constructed composites, not transcripts or direct fictionalizations of real interactions. The vignette format compresses multi-step interactions into a few sentences of a scene, which may heighten ambiguity. A larger scale exercise with independent coders, calibration sessions, and real interaction data (like videos or transcripts of legal work) should be considered a priority for future work, to make the taxonomy as strong and usable as possible.

Despite these limitations, the coding exercise provides an initial demonstration about how independent coding is feasible, that there is category agreement about what task is happening, and unveiled gaps and blurriness that can help with immediate revisions of the task taxonomy's terminology, presentation, hierarchy and list of activities. Even in this short, preliminary round of validation, the taxonomy can advance slowly towards a standard that is easy to apply and covers the tasks at a level that will help great R&D and regulation happen.

6. Future Evaluation and Validation Plan.

6.1. Subject Matter Expert (SME) Review Across Diverse Organizations.

Building on the preliminary coding exercise reported in Section 5, the next phase of evaluation will involve more robust inter-rater coding exercises on a revised draft of the taxonomy, that has been improved based on this coding exercise. The coders will be drawn from various legal team roles and others who might have to delineate legal tasks in the future. The coding exercises will be built upon the insights already gathered about the background details we need to gather about the coders' expertise levels, and will ensure multi-coding ability is possible. This validation work will aim to find the taxonomy's performance on key standards, and ways to improve it. We hope to improve its performance on:

- Relevance: whether the taxonomy captures the tasks they perform or observe;
- Clarity: whether task labels and descriptions are understandable and actionable;
- Completeness: whether critical tasks are missing or under-specified;
- Granularity: whether tasks should be merged, split, or reorganized; and
- Cross-context applicability: whether tasks apply across jurisdictions or only within specific procedural environments.

6.2. Cognitive Task Walkthroughs and Workflow Replays.

Aside from coding exercises, we can continue to engage with more qualitative work. To verify that the taxonomy aligns with real-world workflows, evaluators can use cognitive task walkthroughs with practitioners. Participants would "walk through" recent cases or user interactions and map each action they took to one or more taxonomy tasks. This method tests whether:

- tasks are sufficiently inclusive to represent real work,
- categories reflect intuitive groupings used in practice, and
- the taxonomy can handle edge cases, multi-issue situations, and atypical workflows.

This form of qualitative validation would help assess whether the taxonomy reflects not only the nominal tasks in job descriptions, but also the practical micro-actions that shape user outcomes.

6.3. Cross-Jurisdiction and Cross-Organization Consistency Testing.

As mentioned throughout the paper, the taxonomy should operate across jurisdictions and organizations, and which have different regulatory compliance and principles. To evaluate consistency across jurisdictions, evaluators can take steps like:

- compare task applicability across at least 10 partner organizations in different jurisdictions;
- assess overlap and divergence in how organizations label and sequence tasks; and
- identify any jurisdiction-specific tasks that require optional framework extensions.

This analysis will help identify which tasks may be treated as "core" and which should be modular add-ons. Together, these evaluation methods will whether and how the taxonomy can be refined into a reliable and broadly usable standard.

6.4. Iterative Revision and Publication.

Throughout these evaluation rounds, the taxonomy will undergo iterative revisions and published as a versioned, open standard on JusticeBench.org. Regular updates here will incorporate feedback from additional jurisdictions, emerging legal issues, new technologies, and evolving service delivery models. The goal is to create a piece of well-maintained infrastructure that can provide a key coordination and enablement role in justice techonlogy's development.

7. Conclusion.

Frontline legal help work – whether performed by legal aid organizations, courts, community navigators, digital platforms, or users themselves – relies on a mix of overlapping, interdependent granular actions: interviewing, issue-spotting, triage, document preparation, negotiation, coaching, case monitoring, and more. Yet the legal field lacks a unified, openly available way

to spot and describe these specific activities. This absence of shared structure creates barriers to research, evaluation, interoperability, and (especially right now) the responsible deployment of AI in legal services.

This paper offers an initial draft of the Legal Help Task Taxonomy, that aims to be the structured, cross-organizational vocabulary for the tasks that occur during a person's civil justice journey and a provider's day-to-day of providing this justice help. The taxonomy emerged from sustained qualitative engagement with legal aid practitioners, court help staff, and technology teams – working groups, workflow demonstrations, practitioner walkthroughs, sustained design collaboration, and open-ended brainstorming over several years. It draws on prior socio-legal scholarship, existing legal ontologies, and standardization efforts in adjacent domains to provide conceptual scaffolding for understanding and comparing how legal help is delivered across service models.

We are clear that the taxonomy, as it stands in mid-2026, is a starting point, not a finished standard. It was developed primarily through engagement with US-based organizations, and its categories and task definitions have not yet been subjected to large-scale formal validation or tested for reliability across European or other civil-law jurisdictions. The initial development methodology was qualitatively rich but informal, and the framework should be read accordingly – as a structured proposal that makes the implicit task vocabulary of legal help work explicit and available for scrutiny, refinement, and empirical testing. The preliminary inter-rater coding exercise described here demonstrates that independent coding is feasible and finds greater agreement at the category level than at the task level, while surfacing several task boundaries and hierarchical relationships that require updates to the taxonomy's terminology and organization.

Even in its draft status, the taxonomy can be of practical use. It can inform regulators' or funders' assessments of AI systems by clarifying which specific tasks are candidates for automation and which require heightened safeguards. In the European context, its task-level specificity could support AI Act risk-classification and help controllers determine whether a GDPR data-protection impact assessment is required. Moreover, it can guide the development of interoperable data schemas across courts, legal aid organizations, and technology partners. Future researchers can use a task taxonomy to produce workload measurements, quality determinations, and the mapping of access to justice pathways across jurisdictions.

As AI tools gradually enter the justice system's day-to-day, linking these innovations to accurate, consistent representations of work categories is critical. A shared task taxonomy – even a provisional one – can ensure that new technologies are built for actual frontline practices, and that they can be measured, compared, and spread across organizations and jurisdictions. By offering this structured foundation, we hope to build greater collaboration and higher quality standards among the research and practitioner community working on AI for access to justice.

Appendix A: Full Legal Help Task Taxonomy.

Task ID	Name	Brief Description	Category
TS-01-01	Legal Q&A	Answer people's questions about their situation, giving brief information to help them understand issues, weigh options, connect to resources, and choose next steps.	Getting Brief Help
TS-01-02	Referral Routing	Match people with appropriate legal help providers, tools, and information sources through structured triage and eligibility analysis.	Getting Brief Help
TS-01-03	Document Explainer	Help people understand legal documents – such as notices, summons, and complaints – by summarizing meaning, implications, and required next actions.	Getting Brief Help
TS-01-04	Action Plan	Lay out a step-by-step roadmap tailored to a person's situation, outlining available options and required tasks.	Getting Brief Help
TS-01-05	Deadline Calculator	Determine jurisdiction-correct cutoff dates with plain-language explanations of timing rules and required steps before the deadline.	Getting Brief Help
TS-01-06	Form Selection	Match users to the exact jurisdiction-specific forms needed for their issue and procedural stage, with explanations of purpose and filing steps.	Getting Brief Help
TS-01-07	Issue-Spotting	Identify the specific legal issues embedded in a person's story, situation, or document.	Getting Brief Help
TS-02-01	Q+A Supporter	Support providers in giving high-quality brief legal help with prompts, legal knowledge, and clear language.	Providing Brief Help
TS-02-02	Help Guide Writer	Help legal experts create clear, accurate, and actionable public-facing help guides.	Providing Brief Help
TS-02-03	Media Creator	Turn help guides into videos, audio, and visual media.	Providing Brief Help
TS-02-04	Content Reviewer	Scan public-facing content for accuracy, consistency, tone, and clarity, recommending improvements.	Providing Brief Help
TS-02-05	Law Watcher	Detect changes in law or policy and suggest updates to public materials.	Providing Brief Help
TS-02-06	Legal Help Translator	Translate legal help content into users' preferred languages clearly and accessibly.	Providing Brief Help
TS-02-07	Law Summarization	Generate short, structured summaries of laws, cases, or other legal authorities so that others can understand their relevance and meaning.	Providing Brief Help
TS-03-01	Intake Interviewing	Gather and structure client information to determine needs, risks, and next steps.	Service Onboarding
TS-03-02	Document Issue-Spotter	Review legal documents to spot defects, extract details, and flag urgent or substantive issues.	Service Onboarding
TS-03-03	Data Researcher	Gather and summarize contextual data from external datasets or public sources to support case understanding.	Service Onboarding
TS-03-04	Legal Analyzer	Analyze the legal situation and surface applicable laws, rights, defenses, and strategies.	Service Onboarding

TS-03-05	Conflicts Checker	Determine whether taking on a new client or matter creates a conflict of interest.	Service Onboarding
TS-03-06	Eligibility Verification	Verify whether someone qualifies for a service, program, or benefit by checking evidence against rules.	Service Onboarding
TS-03-07	Payments	Manage collection, processing, and confirmation of payments required for filings, services, or fees.	Service Onboarding
TS-04-01	Standard Document Filler	Help users complete standardized forms through data gathering, validation, and accuracy checking.	Work Product
TS-04-02	Expert Document Drafter	Draft complex legal documents (motions, complaints, discovery requests) tailored to case specifics.	Work Product
TS-04-03	Narrative Drafter	Help people tell their legal story clearly and effectively.	Work Product
TS-04-04	Contract Reviewer	Review proposed contracts or settlements for risks, alignment with goals, and compliance.	Work Product
TS-04-05	Document Draft Checker	Detect errors, omissions, or inconsistencies in drafted documents.	Work Product
TS-04-06	Electronic Filer	Support accurate, rule-compliant filing of documents with courts.	Work Product
TS-04-07	Negotiation Helper	Help people prepare for negotiation by clarifying goals, assessing trade-offs, and modeling options.	Work Product
TS-04-08	Legal Researcher	Conduct targeted, jurisdiction-specific legal research and synthesize findings.	Work Product
TS-04-09	Service of Documents	Serve court documents on the correct parties and generate valid proof of service.	Work Product
TS-04-10	Signatures	Ensure that important legal documents and agreements are correctly signed, witnessed, and validated – so they are legally complete and enforceable.	Work Product
TS-04-11	Evidence Preparation	Prepare, organize, and present case evidence – documents, photos, logs, messages – into admissible, structured packets and exhibits aligned to court rules.	Work Product
TS-04-12	Outcome Prediction	Estimate the likely outcomes of a legal action or decision based on case facts, procedural posture, and available data.	Work Product
TS-05-01	Support Coach	Help people stay on track through reminders, encouragement, and personalized follow-up.	Case Management
TS-05-02	Smart Scheduling	Schedule meetings, hearings, and events by balancing availability and logistics.	Case Management
TS-05-03	Procedure Triage	Route cases to the appropriate procedural track based on issue type, posture, and rules.	Case Management
TS-05-04	Filing Screener	Check filings for completeness, correctness, and compliance with rule requirements.	Case Management

TS-05-05	Service Verification	Verify whether documents were properly served under local rules.	Case Management
TS-05-06	Transcribe Proceedings	Convert hearings or meetings into structured, timestamped transcripts that are searchable and compliant with court norms.	Case Management
TS-05-07	Appointment Booking	Make and manage appointments by coordinating availability, user needs, and service requirements to ensure timely access to help.	Case Management
TS-06-01	OCR/Data Extractor	Automatically extract structured, actionable data from court forms, PDFs, contracts, and images to populate case management systems and support legal service workflows.	Administration & Strategy
TS-06-02	Grant Reporter	Compile data and narratives to produce accurate funder reports.	Administration & Strategy
TS-06-03	Knowledge Management	Maintain and organize internal policies, templates, and expertise.	Administration & Strategy
TS-06-04	Service Directory Maintenance	Maintain structured, accurate service listings for referrals and portals.	Administration & Strategy
TS-06-05	Trends Spotter	Analyze data across systems to identify trends in needs, strategies, and outcomes.	Administration & Strategy
TS-06-06	PII Masking	Help providers transform confidential data into usable, secure datasets by removing identifiers and possibly replacing them with synthetic but realistic fictional information.	Administration & Strategy

Table 2: Full Legal Help Task Taxonomy

Appendix B: Activity Vignettes Used in the Coding Exercise.

The following 62 activity vignettes were used in the preliminary validation exercise described in Section 5. Each vignette describes a composite activity drawn from published accounts of legal help workflows or from activities observed during practitioner working groups. No vignette describes a specific client interaction or identifiable individual. Job titles that would indicate the expected taxonomy code were replaced with first names or neutral descriptions. Vignettes were presented to each coder in a randomized order unique to that coder.

A01. Maria walks into the courthouse lobby clutching an envelope. She shows it to the help desk worker and says, 'I got this in the mail and I don't know what it means.' The worker reads it – it's a summons and complaint for an unpaid medical bill – and spends five minutes explaining in plain language what it says, what the plaintiff is asking for, and what happens if Maria doesn't respond.

A02. A man calls the housing hotline panicking. He says his landlord taped a paper to his door yesterday and he doesn't know what to do. The hotline worker asks a few questions – what does the paper say, how long has he lived there, is he behind on rent – and then explains that he likely has five business days to respond and should go to the self-help center to get the right paperwork.

A03. A woman calls legal aid and starts describing a messy situation: her ex won't return her kids, her car was repossessed last month, and she thinks her boss shorted

her paycheck. The person on the phone listens and then says, 'It sounds like there are actually three different legal issues here – a custody matter, a consumer debt issue, and a wage claim. Let me help you think about which one to deal with first.'

A04. A legal aid organization gets a call from someone who needs help with an eviction. Before anything else, the person who answers the phone goes through a set of screening questions: What county do you live in? What's your monthly household income? How many people are in your household? Do you have any pending cases? The worker types everything into the case management system.

A05. A paralegal pulls up the new client's name and the opposing party's name in their internal database. She checks whether anyone at the firm has ever represented the landlord or the landlord's property management company. She finds a match – another attorney did a brief consultation with the property manager two years ago – and flags it for the supervising attorney to decide.

A06. After reviewing the client's lease, the three-day notice, and the landlord's complaint, the attorney sits down and thinks through the legal angles. The notice was served on a Sunday. The lease has an automatic-renewal clause the landlord may have violated. There might be a habitability defense based on the mold the client photographed. She writes up a memo outlining three possible defenses.

A07. An older gentleman comes to the self-help center with a blank court form and his reading glasses. The center worker sits next to him and walks through the form line by line – writing in the case number from the summons, his full legal name, his address, and checking the boxes that correspond to his defenses. It takes about twenty minutes.

A08. The attorney spends two hours drafting a motion for protective order. She pulls language from a recent appellate opinion, adapts it to the facts of her client's domestic violence situation, attaches three supporting declarations, and formats everything according to the local court's rules.

A09. The client – a single mother facing eviction – needs to explain to the judge why she fell behind on rent. The advocate sits with her and helps her put together a written statement. They talk through the timeline – when she lost her job, when she applied for rental assistance, when the landlord first contacted her – and turn it into a clear, chronological narrative in the client's own voice.

A10. A small team at a statewide legal aid organization sits down to write a new guide: 'What to Do If You Get a Debt Collection Lawsuit in Oregon.' They interview three staff attorneys for the legal content, a plain-language editor simplifies the draft, and a designer lays it out with flowcharts and callout boxes. It'll go on the website and be handed out at court self-help centers.

A11. Every quarter, Rosa goes through all the self-help guides on the organization's website. She checks whether the law has changed, whether the court forms referenced are still current, whether any links are broken, and whether the tone is clear enough for someone reading at an eighth-grade level. She marks up three guides that need updates.

A12. Priya works at a statewide legal help organization and she has a long list. The content management system is three versions behind and needs migrating. The website's design hasn't been touched in four years and doesn't work well on phones. There are security patches overdue on the server. And leadership wants a new chatbot on the homepage that can answer basic questions. She needs to keep everything running seamlessly and securely while making these changes – and she's the only technical person on staff.

A13. Three days before the hearing, Diane, a staff member, sends a text: 'Hi David – just a reminder that your court date is Thursday at 9am, Dept. 5. Do you have a ride? Do you have your documents ready? Reply if you have questions.' David replies asking what to wear. Diane responds with practical advice.

A14. At the court filing window, the clerk looks through the paperwork a self-represented litigant just handed over. The proof of service is missing. The fee waiver form is unsigned. One of the exhibits is a photocopy so faded it's illegible. The clerk explains what needs to be fixed before the filing can be accepted.

A15. It's the end of the fiscal year. Keisha exports data from three different systems – the case management database, the phone log, and the outcomes tracker – and spends two days writing the narrative sections of the LSC grant report: how many people were served, what types of cases, what the outcomes were, what challenges the program faced.

A16. The organization serves a community where 40% of clients speak Vietnamese as their primary language. A bilingual staff member and an outside translation service work together to translate the top ten help guides – on eviction, debt, family law, public benefits – making sure legal terms are explained rather than just literally translated.

A17. A woman calls the hotline and says she was just served with an eviction lawsuit. She wants to know when she has to respond. The worker pulls up the statute, confirms the service method (personal service), and tells her she has five calendar days – but since day five falls on a Saturday, the actual deadline is the following Monday. The worker explains each step she needs to take before then.

A18. A statewide legal help website team – two developers and a legal aid attorney – is building a new guided interview in Docassemble. A user will answer questions about their situation: Are you a tenant or a homeowner? Did you receive a written notice? How many days ago? Based on the answers, the system will select the right response form, pre-fill it with the user's information, and generate a PDF ready to print and file. The attorney keeps revising the branching logic because the rules are different in three of the state's judicial districts.

A19. The client has a court hearing next week about conditions in her apartment. Over the weekend she goes through her phone: screenshots of texts to the landlord asking for repairs, photos of the broken heater and leaking ceiling, a video of mice in the kitchen. She prints everything out at the library, labels each page with a date, and puts them in a folder in chronological order.

A20. Janet at the organization notices that the governor just signed a bill changing the notice period for no-fault evictions from 30 days to 90 days. She sends an email to the whole legal staff with a summary, flags the three website guides that reference the old rule, and starts drafting updated language.

A21. A man at the self-help center says, 'I just want to know what my options are.' The navigator pulls out a pad of paper and sketches out a timeline: here's where you are now, here's the deadline to respond, here are three things you could do – file an answer, try to negotiate with the plaintiff, or file for a continuance. For each option she explains what it involves and what comes next.

A22. A staff member at the legal aid organization produces a three-minute TikTok-style video called 'Got Served? Don't Panic.' It walks viewers through what an eviction notice looks like, what it means, and three things to do right away. A staff attorney reviews the script for accuracy before filming.

A23. The paralegal has been trying to get the client in for a preparation meeting for two weeks. The first appointment was cancelled because the client's car broke down. The second was rescheduled when the client got called into a shift at work. Trial is in nine days. The paralegal calls again, offers three different times including an evening slot, and arranges for the meeting to happen at a community center closer to where the client lives.

A24. The attorney finishes the reply brief at 11pm. She logs into the court's e-filing system, selects the case, uploads the PDF, adds the proof of service as an attachment,

pays the $20 filing fee with the office credit card, and downloads the confirmation receipt.

A25. A woman drives to the defendant's apartment, knocks on the door, and hands the papers to the person who answers. Later that day she fills out a proof-of-service declaration – noting the date, time, address, and a physical description of the person served – and sends it back to the attorney's office.

A26. The organization's leadership sees the intake data as a gold mine – years of records that could reveal patterns, inform new programs, or help build better technology. But nobody on staff has the skills to do that kind of analysis. They want to share the data with a university researcher and a tech consultant, but the records are full of names, addresses, case details, and sensitive personal information. Before anything can leave the building, someone has to strip it down enough that no individual could be re-identified.

A27. Marcus gets an email that a domestic violence organization in the next county has changed its phone number and now offers services in Tagalog. He updates the listing in the organization's resource directory, notes the new hours, and adds the language capability. He also removes a listing for a clinic that closed last month.

A28. The attorney is preparing for a debt collection case. She pulls the plaintiff's business registration from the Secretary of State's website, checks whether the debt collector is licensed with the state regulator, downloads three years of complaints filed against the company, and searches for any relevant enforcement actions.

A29. Before the response goes to the court, Elena reads through every page. She catches that the client's name is misspelled on page three, a paragraph references 'Exhibit C' but the exhibits only go to B, and the certificate of service lists the wrong mailing address. She fixes all three and sends it back for the attorney's final sign-off.

A30. At the quarterly all-staff meeting, Sonia presents a chart: over the past six months, the number of clients coming in with wage-theft issues has doubled, and 70% of them work in the restaurant industry. She recommends the organization consider a targeted outreach campaign and partnership with the workers' center downtown.

A31. A woman walks into the self-help center holding a stack of papers. The worker looks through them – there's an eviction summons, a notice from the power company threatening shutoff, and a letter from a debt collector. The worker helps her understand all three documents but also explains that these are three different legal problems requiring different responses and different deadlines.

A32. An attorney on the brief-advice line takes a call from a tenant who wants to know if she has grounds to break her lease early because of ongoing pest problems. While listening, the attorney pulls up a reference tool that shows relevant statutes and recent case law for the caller's county. He uses the tool to confirm the habitability standards and tells the caller what steps to document.

A33. A court navigator is sitting with a self-represented litigant before a hearing. She checks the court's online docket and notices that the other side filed a proof of service – but it says the papers were left with 'Jane Doe, co-occupant' and her client lives alone. The navigator flags this for the client to raise with the judge.

A34. At the court clerk's window, a man needs to pay a $200 filing fee before his petition can be accepted. The clerk processes the payment, generates a receipt, stamps the filing as received, and hands back the conformed copy. The man didn't realize he'd need exact payment and only has a credit card – the clerk explains which forms of payment the court accepts and where the nearest ATM is.

A35. Tomás, a staff member at the legal aid office, listens back to a recording of last week's administrative hearing. He types up a word-for-word transcript with timestamps every thirty seconds, labels who is speaking (judge, respondent, petitioner's counsel), and saves it in the case file.

A36. A person visits a legal aid website and types into a chatbot: 'I make $2,400 a month and I have three kids. Can I get free help?' The chatbot walks them through a series of questions and then says they appear to meet the income guidelines. It gives them a phone number to call and a list of documents to have ready.

A37. The debt collector's attorney sends over a settlement offer: the client can pay 60% of the original debt in monthly installments over two years, and the collector will dismiss the lawsuit. The legal aid attorney goes through it with the client – explaining what they'd pay total, what rights they'd waive, what happens if they miss a payment – and helps the client weigh it against the alternative of going to trial.

A38. A developer at a legal aid organization needs to build a new online court form for self-represented litigants. It has to ask the right questions in plain language, validate the answers – no future dates in the 'date of incident' field, no letters in the ZIP code box – calculate whether the user meets the income threshold for a fee waiver, and generate a PDF that matches the court's official form layout exactly. She spends three days getting the conditional logic right because the form branches differently depending on whether the user is responding to a lawsuit or initiating one.

A39. Linda is responsible for making sure that every new attorney, every volunteer, and every intern is working from the same playbook. She maintains the internal resource library – guides, templates, checklists, sample motions – and updates them as the law changes, as court procedures shift, as better approaches emerge. This quarter she needs to retire three outdated templates, train a new cohort of pro bono volunteers on the current materials, and figure out whether the team is actually using the resources or just winging it.

A40. A man comes to the self-help center saying he can't afford the filing fee. The worker checks his income against the fee waiver guidelines, confirms he qualifies, pulls up the fee waiver form, helps him fill it out, and then walks him to the filing window to submit both the fee waiver and his answer at the same time.

A41. A woman comes into the courthouse with paperwork from her landlord. She's not sure whether her case belongs in small claims court or in the regular civil division – the amount in dispute is right at the threshold. The person at the help desk looks at the complaint, checks the dollar amount and the type of relief being requested, and tells her which courtroom and which track her case will be assigned to, and what the different procedures are for each.

A42. The organization has a filing cabinet full of old intake forms – hundreds of paper documents from the past three years. A staff member has been asked to pull out specific data points from each one: the client's ZIP code, the type of legal issue, whether they were referred to an attorney, and the outcome. She scans the forms and uses a tool to extract those fields into a spreadsheet, cleaning up handwriting-recognition errors as she goes.

A43. A colleague at the organization just finished building a new online tool that's supposed to help people figure out if they have a defense to an eviction. Before it goes live, someone else on the team goes through it pretending to be six different users – a monolingual Spanish speaker on a phone, someone with a screen reader, a person who enters contradictory answers, a person who leaves fields blank. She documents every place the tool breaks, gives confusing instructions, or produces a wrong result.

A44. A man walks into the self-help center and says he wants to file something about his custody situation but he's not sure what. The worker asks a series of questions – has a case already been filed? Is he responding to something or starting something new? Does he want to change an existing order or get one for the first time? Based on the answers, the worker pulls three different forms from the rack, explains what each one is for, and tells him which one matches his situation.

A45. An attorney is staffing the legal aid hotline and a caller asks a complicated question about whether their wages can be garnished for a medical debt in their state.

While the caller is talking, the attorney opens an internal tool that lets her search the organization's knowledge base by topic and jurisdiction. The tool surfaces the relevant statute, a summary written by a colleague last year, and a checklist of exemptions. She uses this to give the caller a more precise answer than she could have from memory alone.

A46. A client has a mediation session next week with her former landlord over a security deposit dispute. A staff member sits down with her beforehand and helps her think through what she wants to get out of the mediation, what she's willing to compromise on, what her best alternative is if the mediation fails, and how to present her case. They role-play the opening statement together and make a list of the key documents she should bring.

A47. A mother comes in saying her ex-husband hasn't paid child support in four months. She has a court order but doesn't know how to enforce it. A staff member explains the contempt process, helps her figure out how much is owed, and walks through what paperwork she'll need to file to get the court to act on the existing order.

A48. A man was released from prison eight months ago and wants to apply for jobs, but his record keeps showing up on background checks. He comes to a legal aid clinic and someone reviews his criminal history, identifies two old convictions that are eligible for expungement under a recent state law, and starts putting together the petition paperwork.

A49. A woman's SNAP benefits were cut off without explanation. She got a letter in the mail but it's full of codes and abbreviations she doesn't understand. Someone at the legal aid office reads the letter with her, explains that the agency says she missed a recertification deadline, and helps her figure out whether she can get the benefits reinstated by requesting a fair hearing.

A50. An elderly man's daughter is worried that her father is being financially exploited by a caregiver. She calls a legal aid hotline and describes the situation – large checks being written, changes to his bank accounts, the caregiver discouraging family visits. The person on the phone helps her understand the legal options: adult protective services, a restraining order, and possibly a conservatorship proceeding.

A51. A restaurant worker says she's been working 50-hour weeks but only getting paid for 40. She doesn't have pay stubs because her boss pays her partly in cash. A staff member helps her reconstruct a record of her hours from text messages, calendar entries, and co-worker statements, and explains how to file a wage claim with the state labor agency.

A52. A woman's asylum application has a filing deadline in three weeks. She speaks limited English and the supporting declaration needs to describe what happened to her in her home country in specific detail. A staff member works with her through an interpreter over two sessions to draft the declaration in her own words – getting the chronology right, naming the people involved, and describing the threats she received.

A53. A man bought a used car that broke down two weeks later. The dealer won't return his calls. He comes to the self-help center wanting to sue but isn't sure if he has a case. Someone asks him questions about the sale – was there a warranty? Did the dealer say anything about the car's condition? Did he get anything in writing? – and helps him think through whether what happened might violate the state's consumer protection or lemon law statutes.

A54. A couple is going through a divorce and they've agreed on most things – the house, the savings, who the kids live with during the week. But they can't figure out how to divide a retirement account without triggering tax penalties. A staff member looks at the account statements and explains that they'll need a specific court order – a QDRO – and walks them through what information they'll need to give the plan administrator.

A55. A teenager in foster care is about to turn 18 and will age out of the system in a few months. A staff member at a legal aid organization sits down with her to map out everything she needs to get in order before that happens: applying for Medicaid continuation, securing her birth certificate and Social Security card, understanding her rights to extended foster care benefits, and figuring out her housing options. They make a checklist and set dates for each step.

A56. A community organization that serves immigrant families asks a legal aid group to come do a 'know your rights' presentation. Two staff members put together a thirty-minute presentation in Spanish covering what to do if ICE comes to your door, how to make a family safety plan, and where to find legal help. They print wallet cards with emergency phone numbers and create a simple one-page handout with the key points.

A57. It's the client's first time in the courthouse and she's overwhelmed. She doesn't know which floor to go to, where to check in, or where the right courtroom is. A worker at the information desk walks her through the building – past the metal detector, up to the second floor, shows her the clerk's office, points out the restroom, and walks her to Department 12 where her hearing will be. He tells her where to sit and what will happen when her case is called.

A58. A public defender's office receives 40 gigabytes of body-worn camera footage, jail phone recordings, and cell phone extraction data as part of discovery in a felony case. A staff member has to go through all of it – flagging the segments that are actually relevant, noting timestamps, identifying speakers, and pulling out the pieces the attorney needs to build the defense. It takes days.

A59. A client just came out of a hearing where the judge ruled against her. She's sitting on a bench in the hallway crying. A staff member sits down next to her and talks her through what just happened, what it means, what her options are for next steps, and reminds her that she can still appeal. She makes sure the client has a way to get home and gives her a phone number to call on Monday to talk about what comes next.

A60. A legal aid attorney has a case involving a client who was denied unemployment benefits. The state sent a twelve-page decision full of administrative code citations. The attorney needs to quickly understand what the agency found, what rules they applied, and where the reasoning might be wrong. She feeds the decision into a tool that produces a structured summary – pulling out the key findings, the legal standards cited, and the specific evidence the agency relied on.

A61. A disabled veteran is trying to get his VA benefits increased but the paperwork is overwhelming – medical records from three different providers, service records, a previous denial letter, and a new diagnosis. A staff member helps him organize everything into the order the VA application requires, writes a cover letter summarizing the claim, and makes sure each piece of supporting evidence is referenced in the right section of the form.

A62. A tenant calls a legal aid hotline after discovering that her landlord has been entering her apartment while she's at work. She's scared and wants to know her rights. The person on the phone explains that unauthorized entry may violate the state's landlord-tenant statute, walks her through how to document each incident in writing, and explains the options – a written demand letter, contacting code enforcement, or seeking a restraining order if the behavior escalates.