Proceedings of CONF-SPML 2027 Symposium: Intelligent Network Security and Machine Learning for Communication Systems
DOI: 10.54254/2755-2721/2027.MELB37599

In Her Purse: Designing a Benefits Navigation Platform for Incomplete and Sensitive Disclosures

Shannon Zhang

Lynbrook High School, San Jose, USA
shannonzhang2020@gmail.com

Abstract. Public assistance is organized by program, while people seeking help often describe overlapping needs without knowing which agency or eligibility category applies. In Her Purse is a deployed web platform that accepts a plain-language description and optional profile information to search approximately 4,489 official U.S. assistance listings. This paper examines its Eligibility Matcher as a design for navigation across housing, income, education, employment, and other support domains. The matcher combines semantic retrieval with language-model analysis, presents criteria that appear satisfied separately from those requiring clarification, and requires neither an account nor a complete profile. A descriptive evaluation used thirty written personas situated in economic abuse, including five Spanish- or Chinese-language inputs. The system returned 297 listing instances representing 131 distinct programs, with a reported mean of 3.9 need domains per persona. The evaluation also exposed organizational funding notices among individual-facing results and lower confidence labels for incomplete and non-English inputs. These results establish output coverage in the tested scenarios, not eligibility accuracy, user benefit, or safety. The paper contributes an implemented navigation approach and an account of the design problems that remain when broad program retrieval must support decisions made with limited information.

Keywords: Benefits Navigation Platform, Administrative Burden, Economic Abuse, Information Retrieval, Privacy

1. Introduction

A request for help rarely arrives in the language of an assistance directory. Someone may need rent money, food for children, and a way to continue training, yet know none of the programs that address those needs. Identifying relevant support requires learning what exists, translating circumstances into eligibility categories, and deciding which application is worth attempting. Administrative-burden research describes this work in terms of learning, compliance, and psychological costs [1, 2]. Evidence on benefit non-take-up shows why discovering a program is a substantive part of access, even when assistance is formally available [3, 4].

Economic abuse makes the cost of that discovery especially consequential. A partner may control household money, withhold documents, exploit credit, or interfere with work [5, 6]. The person seeking assistance may therefore lack both the information requested by a form and the privacy

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2027.MELB37599

needed to supply it. Research on intimate partner surveillance documents the use of shared accounts, ordinary device features, and access to personal information to monitor survivors [7, 8]. A navigation service must account for the circumstances in which a question can be answered, as well as the information an eligibility assessment would ideally contain.

In Her Purse addresses this problem through a web platform whose Eligibility Matcher accepts free text and any completed portion of a sixteen-field profile. The deployed implementation, at https://inherpurse.org, retrieves official assistance listings and explains which criteria appear to fit the information supplied. Users do not have to select a benefit category, create an account, or complete every field before receiving results. The platform covers several kinds of support; employment is one domain within this broader navigation task.

The design draws on trauma-informed computing's emphasis on choice and control [9, 10]. It also uses retrieval-augmented generation to connect generated explanations to source material [11]. Retrieval constrains the set of programs that can be returned, but it does not establish that an explanation is correct or that a user can apply. Language-model errors remain possible within a retrieved document set [12].

We examine two questions. First, how can a benefits navigator accept incomplete descriptions and present useful next steps without treating missing information as a disqualification? Second, what coverage and failure patterns appear when that design is exercised with written scenarios? We contribute the implemented matcher, its intake and explanation rationale, and a descriptive evaluation of thirty personas. The evaluation is specific to economic-abuse scenarios. The architecture supports navigation across need domains, but its usefulness for other populations remains an open empirical question.

2. Background and design rationale

2.1. Moving program knowledge into the service

An assistance listing tells a reader about one program. A person seeking help must first determine whether that program belongs in the search at all. This distinction matters when one situation crosses agency boundaries. A short description of unstable housing may also contain information relevant to income support, health coverage, or education. Requiring the user to choose the correct category before searching transfers this interpretation back to the person least likely to know the directory [1, 2].

Reducing learning costs does not remove the administrative burden of applying. Verification, deadlines, local delivery arrangements, and documentation can still prevent access. Eubanks's account of automated public services cautions against assuming that additional computation necessarily reduces burdens for claimants [13]. We accordingly treat navigation as a bounded task: identify possible routes, preserve access to authoritative sources, and state what is still unresolved.

2.2. Disclosure under economic abuse

Economic control can remain unnamed during help-seeking, and its financial consequences can continue after a relationship ends [14, 15]. An intake process that requires users to identify themselves as survivors may miss people who describe the effects without using that term. Free text allows a user to describe a partner taking wages or withholding money without first adopting a diagnostic category.

DOI: 10.54254/2755-2721/2027.MELB37599

The ability to type such a description does not mean doing so is safe. Studies of technology-mediated abuse describe risks distributed across devices, accounts, relationships, and services [16, 17]. Expert-supported security consultations demonstrate the importance of examining a person's actual technical circumstances [18]. A benefits interface cannot replace that work. Requiring no account removes one source of persistent user-facing information; it does not establish that browser traces, network records, caches, or third-party processing leave no record.

2.3. Explanations as a record of available evidence

A retrieved listing and a personalized eligibility judgment are different objects. The first establishes a source; the second interprets that source against incomplete user information. The design keeps official links visible and distinguishes apparent matches from criteria that need checking. This separation follows the practical logic of retrieval-augmented generation [11], while recognizing that grounding alone cannot prevent mistaken interpretation [12]. The user should be able to inspect the basis for a suggestion rather than rely on an unexplained score.

3. Platform and matcher implementation

In Her Purse includes an Eligibility Matcher, a Learning Hub about economic abuse, a Resource Library, Advisor Matching, and a Community space. The implementation and evaluation reported here concern the matcher. No claims about use or effectiveness of the other components follow from the persona tests.

3.1. Intake

The matcher offers a free-text description and sixteen optional structured fields. The fields cover demographic and household circumstances, location, citizenship, education, income, employment, financial aid, veteran status, disability, and parental status. Figure 1 shows part of the intake interface. The description can supply circumstances that a structured field would miss, while completed fields provide additional information for interpreting eligibility criteria.

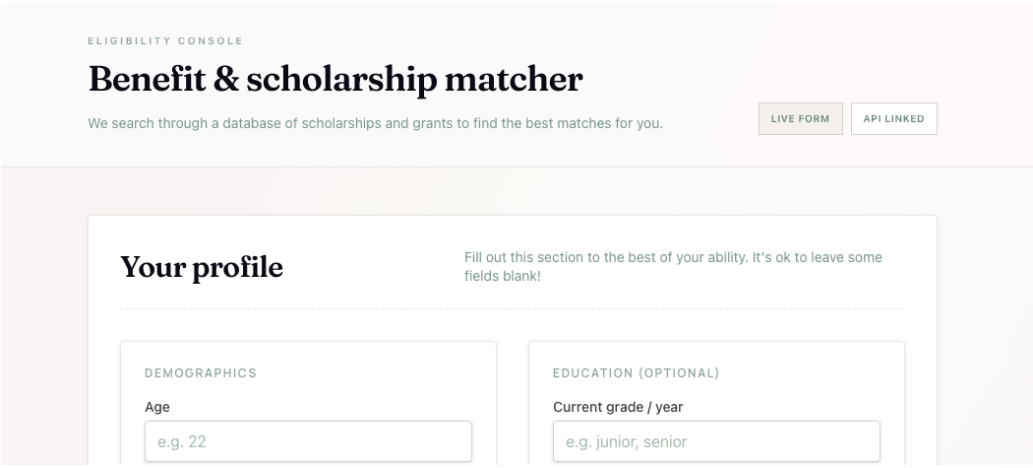

Figure 1. Part of the In Her Purse intake screen

The profile prompt permits fields to be left blank. The evaluation concerns the matcher, not the other platform components.

DOI: 10.54254/2755-2721/2027.MELB37599

Four decisions organize the interaction. First, the user can begin with a situation rather than a program name. Second, a blank field is treated as missing information rather than a negative answer. Third, a result distinguishes criteria that appear satisfied from criteria requiring confirmation. Fourth, a match can be requested without an account or mandatory identity information. Together, these decisions aim to make an initial search possible before a person has assembled a complete application profile.

The distinction between unknown and ineligible is particularly important. If citizenship or income is omitted, the system should not invent the answer. Keeping a candidate visible with an explicit unresolved criterion may be useful for exploration, but it also creates a responsibility to avoid presenting the candidate as a confirmed entitlement.

3.2. Corpus and retrieval

The evaluated corpus contains approximately 4,489 official assistance listings. It draws on SAM.gov Assistance Listings, the successor to the Catalog of Federal Domestic Assistance [19], with additional records from Grants.gov and the California Grants Portal. The corpus size describes the implementation used for the reported evaluation; it is not an estimate of the number of benefits available to an individual. The collection is federal-heavy and has only partial state coverage.

Each record retains an official program title, administering agency, eligibility text, and source link. Records are stored in Firebase Firestore, with embeddings held in a Pinecone index. The completed fields and free-text description are combined into one query, embedded using text-embedding-3-small, and compared with listing embeddings by cosine similarity. Approximately twenty-five nearest listings are retrieved as candidates.

A GPT-based analyst then considers the description alongside each retrieved listing. It assigns a Full, Strong, or Partial match label, identifies apparently satisfied and unresolved criteria, and supplies application guidance. The analyst can reject a candidate but is constrained to the retrieved program set. Up to ten results are displayed. This restriction limits which program identities can appear; it does not independently validate eligibility explanations, deadlines, or application instructions.

3.3. Results and implementation boundaries

The application uses a Next.js front end and a FastAPI back end. Result cards show the official title and agency, the match label, criteria that appear satisfied, criteria requiring further checking, application steps, and a source link. Repeated queries are cached. The evaluation does not include a storage audit, cache-retention assessment, or analysis of third-party data handling, so the absence of a required account should not be interpreted as a guarantee of anonymous or nonpersistent processing.

The implementation report records an automated structural check informed by WCAG 2.2 [20]. That check found a declared page language, heading structure, labels for the sixteen fields, a named submit button, and landmarks; it also reported working zoom. Manual assessment of contrast, keyboard operation, focus management, and screen-reader announcements for dynamic results was incomplete. These checks do not establish full accessibility conformance.

DOI: 10.54254/2755-2721/2027.MELB37599

4. Descriptive evaluation

4.1. Scenarios and procedure

We used thirty written personas, P01–P30, to examine system output. Personas provide a way to articulate situations for design and testing [21]; they are not substitutes for survivor participation. The scenarios were informed by the literature and were not transcripts or accounts collected from participants. No survivors were recruited for this evaluation.

The set includes seven pre-recognition personas, eight currently living under economic abuse, eight who had recently escaped, and seven in longer-term recovery. Ages range from 19 to 68. The scenarios vary in state, citizenship and immigration circumstances, disability, veteran status, and parental status. Each contains a short description naming no program and a structured profile with some fields deliberately blank. Five descriptions are in languages other than English: three Spanish and two Chinese.

Each payload was submitted once to the live matcher, in a single batch, through the same processing path used by the interface. No persona was rerun or selected after reviewing its output. The available evaluation record does not specify a dated corpus snapshot, an exact GPT model version, or generation settings. These omissions limit reproducibility and prevent attributing differences to any single pipeline component.

4.2. Measures and coding

We summarize whether each run returned results, the number and distinct identity of returned listings, the system's confidence labels, and approximate coverage across need domains. A returned instance is one listing appearing in one persona's results; the same program can appear in several runs. Across the evaluation, 297 returned instances represented 131 distinct programs.

One coder categorized the distinct programs using keyword rules applied to titles. The domains were Housing, Employment, Education, Income, Health, Food, Legal/Safety, Aging, Disability, Family, Financial capability, and Other. Listings were also categorized according to whether their apparent applicant was an individual or an organization seeking funding. These are descriptive title-based classifications. They were not checked by a second coder or validated against complete program rules.

No benefits counselor supplied a gold-standard relevance or eligibility judgment. Consequently, result counts measure output coverage, domain labels approximate topical breadth, and confidence labels describe the system's behavior. None is a measure of eligibility accuracy. Table 1 reports the stage summaries retained from the evaluation.

Table 1. Output summaries by persona stage

Stage	n	Returned instances	Mean per run	Full + Strong labels	Mean domains	Individual-facing
Pre-recognition	7	69	9.9	27	3.1	91%
Currently in	8	80	10.0	6	4.2	69%
Recently escaped	8	80	10.0	19	4.1	78%

DOI: 10.54254/2755-2721/2027.MELB37599

Table 1. (continued)

Long-term recovery	7	68	9.7	15	4.0	75%
All	30	297	9.9	67	3.9	78%

Note. Returned instances count repeated appearances across personas. Full and Strong are generated labels, not independently confirmed matches. Domain and applicant-type summaries use title-based coding; percentages and means are rounded.

4.3. Coverage across needs

All thirty runs returned results. The total was 297 listing instances, or 9.9 per persona against the ten-result display cap. Every non-English persona also received a full result set. Near-cap output shows that the pipeline completed these requests, but cannot distinguish relevant retrieval from a system that fills available slots with weak candidates.

The reported mean number of need domains was 3.9 per persona. Housing accounted for 23.9% of returned instances, Employment for 22.6%, Education for 21.9%, Income for 11.1%, Health for 5.4%, and Legal/Safety for 4.7%. Seventeen personas received employment or training listings alongside housing or income support. This co-occurrence illustrates retrieval across categories; it does not show that the programs could be combined or that the persona would qualify for them.

Table 2 presents selected outputs from the original scenario set. Olivia's description included a long absence from work and a need for income and training. The matcher returned unemployment insurance and workforce training alongside housing-related listings. That breadth is also a reason for caution: a request for income does not establish an unemployment-insurance entitlement, and an official housing program may require access through a provider. Grace's outputs similarly show how disability and financial control can retrieve several forms of support without establishing the eligibility of any particular match.

Table 2. Illustrative persona inputs and returned listings

Persona and stage	Scenario summary	Representative returned listings
Grace (P11, currently in, disability)	works, partner takes her paycheck, needs housing and health care she can access alone	Supportive Housing for Persons with Disabilities; Social Security Disability Insurance; Section 811 Project Rental Assistance
Olivia (P21, recently escaped)	left after 20 years; no recent work history or savings; wants income support and job training	Unemployment Insurance; WIOA Adult Program; Family Unification Program; Supportive Housing for Persons with Disabilities
Lucia (P08, currently in, Spanish)	needs food and housing for two children; no access to her husband's money	Domestic Violence Assistance Program RFP; Human Trafficking Victim Assistance RFP; Transitional Housing Program RFP — all three are organizational funding notices

Descriptions are scenario summaries, not participant quotations.
Note. Program names reproduce the recorded outputs. Inclusion does not establish eligibility, a direct application route, or an appropriate recommendation.

DOI: 10.54254/2755-2721/2027.MELB37599

4.4. Confidence and applicant mismatch

The analyst assigned 23 Full, 44 Strong, and 230 Partial labels. The eight currently-in personas received six Full or Strong labels in total, compared with twenty-seven among the seven pre-recognition personas. Because the persona groups also differ in the circumstances and information supplied, these counts do not establish that abuse stage caused the difference. They raise a design question about whether a prominent summary label communicates uncertainty about evidence as if it were uncertainty about the person's eligibility.

Approximately 22% of returned instances were classified as organizational rather than individual-facing listings. Lucia's Spanish-language scenario makes the issue concrete: the results included funding requests for organizations operating services. Such notices can be semantically close to a person's need while providing no direct application route for that person. Applicant type must therefore be distinguished from topical relevance.

All fifty results for the five non-English personas were labeled Partial, whereas 67 of the 247 English-language results were Full or Strong. Three non-English personas also belonged to the currently-in group, so language and scenario composition are confounded. Without equivalent scenarios tested across languages, the evaluation cannot isolate a language effect or establish translation quality.

5. Implications for benefits navigation

5.1. Separate a program from a usable route to help

An official listing can describe funding that reaches individuals only through an intermediary. Treating such a listing as a direct application option shifts another search task onto the user. The next implementation should record applicant type before ranking and distinguish direct individual applications from provider-mediated access. Organizational notices may still be useful when the service can identify a relevant provider; that provider mapping was not evaluated here.

Eligibility and practical access also need separate treatment. A program can be relevant yet require an address, appointment, or contact method that the person cannot currently use. Future cards should state verified application conditions alongside formal criteria. Those annotations must come from program documentation or reviewed provider information, rather than inference presented as fact.

5.2. Report uncertainty at the criterion level

Full, Strong, and Partial labels compress several different conditions: omitted information, uncertain interpretation, and a requirement that may not be met. A criterion-level presentation can distinguish them. For example, a card might report that location appears compatible, household income is unknown, and a specific document will be required. This gives the user an actionable description of the evidence without treating a blank answer as a personal deficiency.

This is a proposed revision, not a feature validated by the current evaluation. Its value should be tested through comprehension tasks: whether people can explain why a listing appeared, identify what remains unresolved, and decide when an official source or adviser is needed.

DOI: 10.54254/2755-2721/2027.MELB37599

5.3. Test multilingual access and privacy directly

The non-English outputs justify a controlled multilingual evaluation, not a claim of multilingual effectiveness. Parallel versions of the same scenarios would help separate language effects from differences in profile content. Review should examine program relevance, criterion interpretation, and application instructions, rather than simply counting returned results or comparing generated confidence labels.

Privacy also requires evaluation beyond the intake screen. Optional fields reduce required disclosure, but the free-text box may contain identifying or sensitive details. An implementation review should trace those details through retrieval, model processing, caching, logging, and any subsequent support referral. A proposed session-level choice of usable contact channels would need the same scrutiny before it could be described as protective.

6. Limitations and next evaluation

The study is a single descriptive exercise with thirty authored scenarios. It provides no evidence about actual benefit receipt, time saved, user comprehension, trust, or real-world safety. The title-based coding is approximate, and the absence of independent relevance judgments makes claims of retrieval precision or eligibility accuracy unwarranted. Stage and language comparisons are based on small, uneven groups. The corpus has limited state and local coverage, while absent version details make the exact run difficult to reproduce.

The next evaluation should begin with benefits-counselor review of persona–program pairs, including applicant type and application route. Comparisons with keyword search or a directory interface would then help identify whether semantic retrieval improves discovery. Advocate review can examine whether scenarios and explanation wording reflect plausible constraints. Any subsequent survivor study would require appropriate ethics review and support arrangements, with choice and control carried into recruitment, data collection, and withdrawal [9, 10].

The generalizable design proposition is modest: a service can accept incomplete descriptions and search across assistance categories before demanding a full profile. Whether that proposition improves access for other populations, or for survivors using the system under actual constraints, remains to be established.

7. Conclusion

In Her Purse implements benefits navigation around a person's description rather than an agency category. Its matcher searches official listings, permits incomplete input, and exposes unresolved criteria. Thirty persona runs showed broad output coverage and revealed problems that program retrieval alone does not solve: indirect application routes, confidence labels that obscure missing evidence, and uncertain performance across languages. The next step is to establish which suggestions are actionable and understandable, and to examine the data practices on which safe use would depend.

References

[1] Moynihan, D., Herd, P., & Harvey, H. (2015). Administrative burden: Learning, psychological, and compliance costs in citizen-state interactions. *Journal of Public Administration Research and Theory, 25*(1), 43–69.

[2] Herd, P., & Moynihan, D. P. (2018). *Administrative burden: Policymaking by other means*. Russell Sage Foundation, New York, NY.

DOI: 10.54254/2755-2721/2027.MELB37599

[3] Currie, J. (2006). The take-up of social benefits. In *Poverty, the distribution of income, and public policy* (pp. 80–148).

[4] Bhargava, S., & Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *American Economic Review, 105*(11), 3489–3529.

[5] Adams, A. E., Sullivan, C. M., Bybee, D., & Greeson, M. R. (2008). Development of the scale of economic abuse. *Violence Against Women, 14*(5), 563–588.

[6] Postmus, J. L., Plummer, S.-B., & Stylianou, A. M. (2016). Measuring economic abuse in the lives of survivors: Revising the scale of economic abuse. *Violence Against Women, 22*(6), 692–703.

[7] Matthews, T., et al. (2017). Stories from survivors: Privacy & security practices when coping with intimate partner abuse. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2189–2201.

[8] Freed, D., Palmer, J., Minchala, D., Levy, K., Ristenpart, T., & Dell, N. (2018). "A stalker's paradise": How intimate partner abusers exploit technology. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13.

[9] Chen, J. X., McDonald, A., Zou, Y., Tseng, E., Roundy, K. A., Tamersoy, A., Schaub, F., Ristenpart, T., & Dell, N. (2022). Trauma-informed computing: Towards safer technology experiences for all. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, Article 544, 1–20.

[10] Substance Abuse and Mental Health Services Administration. (2014). *SAMHSA's concept of trauma and guidance for a trauma-informed approach*. U.S. Department of Health and Human Services.

[11] Lewis, P., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems*, 9459–9474.

[12] Ji, Z., et al. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), 1–38.

[13] Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press, New York, NY.

[14] Postmus, J. L., Hoge, G. L., Breckenridge, J., Sharp-Jeffs, N., & Chung, D. (2020). Economic abuse as an invisible form of domestic violence: A multicountry review. *Trauma, Violence, & Abuse, 21*(2), 261–283.

[15] Sharp-Jeffs, N. (2022). *Understanding and responding to economic abuse*. Emerald Publishing, Bingley, UK.

[16] Freed, D., Palmer, J., Minchala, D., Levy, K., Ristenpart, T., & Dell, N. (2017). Digital technologies and intimate partner violence: A qualitative analysis with multiple stakeholders. *Proceedings of the ACM on Human-Computer Interaction, 1*(CSCW), 1–22.

[17] Tseng, E., et al. (2020). The tools and tactics used in intimate partner surveillance: An analysis of online infidelity forums. In *Proceedings of the 29th USENIX Security Symposium* (pp. 1893–1909).

[18] Havron, S., Freed, D., Chatterjee, R., McCoy, D., Dell, N., & Ristenpart, T. (2019). Clinical computer security for victims of intimate partner violence. In *Proceedings of the 28th USENIX Security Symposium* (pp. 105–122).

[19] U.S. General Services Administration. (n.d.). *Assistance listings*. SAM.gov.

[20] World Wide Web Consortium. (2023). *Web content accessibility guidelines (WCAG) 2.2*. W3C Recommendation, 5 October 2023.

[21] Pruitt, J., & Grudin, J. (2003). Personas: Practice and theory. In *Proceedings of the 2003 Conference on Designing for User Experiences (DUX '03)* (pp. 1–15).