An Architecture for Grant Discovery and Grounded Application Drafting

Uthman Olomide
uo@grantivo.com

Technical White Paper
July 2026

Abstract

Grant seekers face a discovery problem: tens of thousands of private foundations give billions of dollars each year, little of it through open calls, yet the usual tools only index the calls. Grantivo is built on a different premise: the most informative signal about a funder is its record of past giving, disclosed in annual public filings. It structures those filings into a queryable model of funder behavior, more than 140,000 foundations and over 7.5 million grants, and puts it to work in two ways. In discovery, organizations are matched to funders by demonstrated giving, not keyword overlap, and each candidate resolves to a research profile: giving history, the people who run it, and its stated application guidance. In writing, an assistant working from the target funder's priorities and the organization's own materials both drafts applications and audits them for fit to the funder, a judgment that a general-purpose model cannot reliably make. One property of the design matters beyond its parts: because the writing happens inside the system, it observes the applicant-side text and, over time, the outcomes that the filings omit, supervision no public dataset contains. Structured discovery, evidence-backed drafting, and draft audit operate now; the learned ranking and outcome-driven supervision that let the system improve with use are its next stage, with the governance that using customer applications requires.

1 Introduction

An organization seeking philanthropic funding faces a search problem, but not the one most tools solve. The usual framing asks to what open opportunities to apply to now; the larger question is which funders are likely to support work like this and how to approach them. Private foundations rarely advertise, and much giving flows through relationships, invited proposals, and recurring support that never reaches an opportunity board, so a tool that indexes only posted calls sees a small and unrepresentative fraction of funding activity.

Funders do, however, leave a public trail. Each year,, every private foundation discloses its finances and an itemized list of the grants it made, with recipient, geography, and amount. Aggregated across years and across the whole population of filers, this is a record of who funds what, and it reflects a funder's actions, not its self-description. Grantivo is built on the premise that this behavioral record, not the set of currently open calls, is the most informative signal for matching an organization to a funder, and the rest of the design follows from it.

The design is based on three contrasts with existing tools. Keyword portals index opportunities; Grantivo models funders, turning a sparse feed of announcements into a standing map of who supports work like yours. Generic AI helps you write; Grantivo first works out who is worth writing to because a fluent draft sent to the wrong funder is wasted effort. And most systems lose the information the moment an application is submitted; because the writing happens inside the product, applications and, with consent, their outcomes accrue as data no public dataset holds, so the system is designed to improve the more it is used. Model funders, not opportunities; target before drafting; learn from the work itself. The rest of the paper is how (Figure 1), with

Section 6 on the property we find most interesting: coupling discovery with drafting allows the system to observe data the source filings do not contain.

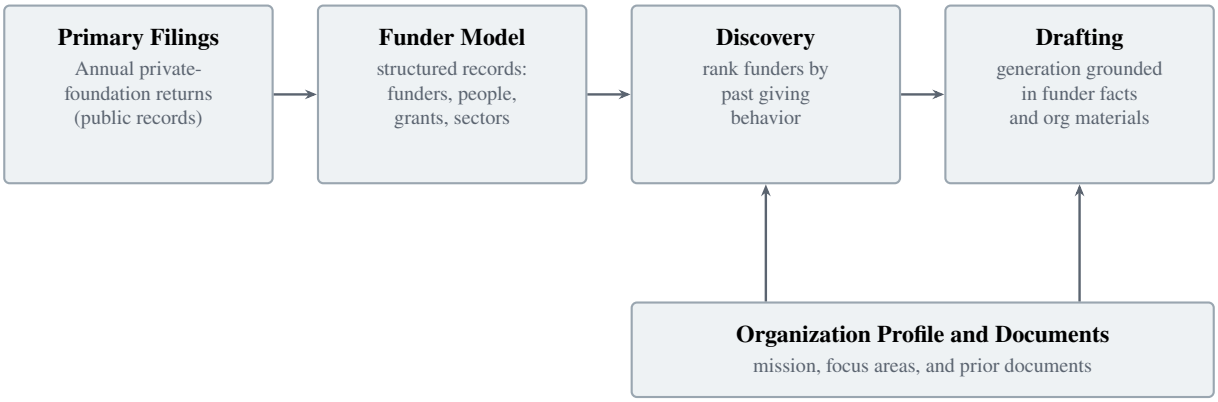

Figure 1: The path from data to drafting. A structured funder model built from primary filings supports behavior-based discovery and grounded drafting; the organization's profile and documents inform both discovery and drafting.

A worked example (synthetic). Consider an invented applicant: River's Edge Literacy, a small Dayton, Ohio nonprofit seeking $25,000 for after-school reading programs. Structured filters (education, Ohio, grant size $10k–50k, accepts unsolicited requests) narrow the funder universe to a few dozen candidates; the top-ranked is a staffed Ohio funder with recurring $15k–30k grants to literacy nonprofits, each supporting fact drawn from its filings and shown to the user. Drafting then grounds a letter of inquiry in those facts and the organization's own materials, and the audit scores the result on writing quality and funder fit separately, flagging, for example, that no concrete outcome figure is present and declining to score impact evidence until the author supplies one. The rest of the paper describes each of these stages.

---

2 The Funder Model

The system begins with the annual returns private foundations file (Form 990-PF), released as public records (Internal Revenue Service, n.d.). These give population-scale statutory coverage of U.S. private-foundation filers, not a self-selected subset that posts opportunities, though that coverage is bounded by filing compliance, publication lag, amendments, and extraction limits; and they are behavioral, itemizing grants actually made. These arrive as bulk, semi-structured documents that vary in completeness and format across filers and years. An ingestion pipeline extracts the relevant entities and normalizes them conservatively, cleaning identifiers, reconciling formats, and resolving dates and amounts, so a funder can be followed across years, which is a record-linkage problem (Christen, 2012). The result is a relational model of three record types: funders, with identity, location, financial scale, sector, staffing, and any application guidance they disclose; the people who run them; and the grants they have made, itemized by recipient, geography, and amount. Because a foundation files every year, a funder is a sequence of filings over time, and discovery uses the most recent complete filing while retaining the history. Records are also enriched with a standard sector taxonomy (National Center for Charitable Statistics, n.d.), so each funder has a consistent field of activity such as arts, education, or health.

Scale of the current dataset. The model spans the population of U.S. private foundations that file: more than 140,000 distinct foundations, roughly 1.8 million associated officers, trustees, and staff, and over 7.5 million itemized grant records, drawn from hundreds of thousands of annual filings from 2017 onward, refreshed as new filings are released. Distinct foundations are counted by EIN and grants and filings per record, from a mid-2026 snapshot of ingested returns; the counts exclude nonfilers and returns not yet published or successfully parsed.

Three properties of the source shape how it is used. Fields differ in reliability: amounts, geography, and identity are recorded consistently, while the free-text purpose attached to individual grants is often too terse to carry signal, so structured fields drive filtering and a funder's longer-form application guidance, where present, informs thematic judgment. Filings arrive on a lag of a year or more, so the model captures a funder's established pattern, not its latest move. And machine-readable coverage becomes materially complete only after the Taxpayer First Act's electronic-filing mandate (tax years beginning after July 2019), leaving earlier years sparser to extract.

3 Discovery and Funder Research

This model is what discovery searches over. Today, discovery is structured and explainable: an organization narrows the funder universe by reliably disclosed attributes, geography, sector, typical grant size, financial scale, whether the funder is staffed, and whether it accepts unsolicited requests. Because grant history is part of the record, these are behavioral rather than purely textual filters, expressing questions such as which staffed education funders in this state give at this grant size; and because each is explicit, a user can see why any funder appears. A keyword portal cannot answer a query of that shape, because it turns on giving behavior the portal never indexes. Discovery works from each funder's most recent filing to avoid duplicate and stale entries, and it relates funders to one another by shared geography and field of activity, forming a navigational similarity network the user can traverse from one relevant funder to similar ones, a browsing aid, not a learned graph model.

Finding a funder is only the first half; the second is knowing enough to approach it well. Each funder resolves to a research profile assembled from its filings: financial scale and giving totals across years, the itemized history of every grant it has made (recipient, amount, year), and the people who run it, drawn from roughly 1.8 million officer, trustee, and staff records. The people records separate a professionally staffed grantmaker from a family-run vehicle with no staff, and they surface the program officer an applicant would write to. The profile also carries whatever the funder discloses about how to approach it, deadlines, restrictions, and whether it accepts unsolicited requests, plus related funders in the same geography and field. This is the diligence a grant professional would otherwise assemble by hand, and because every number traces back to a specific line in a public filing, it is checkable, not asserted.

Three qualifications bound the matching. The strongest form of behavioral matching would profile a funder's past grantees and surface funders whose recipients resemble the user; grant records name recipients without describing them, a recipient-side entity-resolution problem (Christen, 2012) the system today approximates with sector, geography, and grant-size patterns. Approachability is a first-class filter, not an afterthought: among filers whose policy is determinable, only about one in five accepts unsolicited requests, so this signal alone removes much of the field. And past giving is read as historical propensity, not current willingness to fund; a prior grant may reflect a multiyear commitment, a board relationship, or a since-departed strategy.

What structured filters do not capture is thematic fit: whether an organization's work matches the language a funder uses to describe what it funds. Closing that gap follows a well-understood path, lexical matching plus dense embeddings, a second-pass reranker, and a learned ranker such as LambdaMART fusing these with the structured behavioral features (Robertson and Zaragoza, 2009; Karpukhin et al., 2020; Nogueira and

Cho, 2019; Burges, 2010). Adopting this hybrid stack while keeping the structured predicates as explainable features is the main line of the discovery roadmap (Section 8); Section 9 traces the lineage of these methods.

4 Drafting: Retrieval-Augmented Generation

Discovery ends with a shortlist of funders; writing to them is the harder half. Instead of asking a general-purpose model to describe a funder from memory, the drafting stage conditions generation on a bounded, verifiable context: the structured facts about the funder in question, the organization's profile, and the document in progress. This is retrieval-augmented generation (Lewis et al., 2020): grounding a language model in retrieved, authoritative context to keep output specific and reduce fabrication. Grounding this way makes the evidence available and traceable, though it does not by itself guarantee that every sentence is entailed by a source. The same mechanism serves both drafting a section and revising a single paragraph.

The organization's own materials are the second source of evidence. Organizations accumulate reusable text, past proposals, program descriptions, evaluation reports; the system lets them build a private store, encoded with a general-purpose sentence embedding model (Reimers and Gurevych, 2019) so that the most relevant passages can be drawn into the drafting context and a draft reflects the organization's real programs instead of invented detail. The assistant itself is built on general-purpose language models chosen by capability, principally long-context comprehension and reliable instruction-following, not by vendor, and given the same verifiable context in every request.

5 Draft Audit

Generation produces a draft; the real test is whether that draft will win over the funder it targets, and this is where a general-purpose model is weakest. Ask a generic assistant to review a grant application and it returns generic writing advice, tighten this, expand that, because it has nothing specific to judge the draft against. Grantivo's audit starts from exactly what generic feedback lacks: the demonstrated priorities of the target funder, read from the funder model, together with the organization's own verified profile. It reads the actual draft in place and scores it on two axes it keeps separate, the quality of the writing and its fit to the funder, so that a polished but off-target application and a well-aimed but poorly written one are told apart, not blurred into a single grade. Each axis is a rubric of separately scored criteria, not one opaque number, and every suggestion cites the funder criterion or applicant source it rests on, carries a confidence level and an explicit missing-information flag, and abstains when the evidence is too thin to support a judgment.

Feedback is returned as discrete suggestions, each anchored to a span of the draft and typed by what it asks of the writer: a direct edit the system can apply, or a prompt for substance only the author can supply, such as a concrete outcome, a figure, or a named partner. That distinction is deliberate. It keeps the tool from inventing evidence it does not have while still doing the mechanical work it can, and the aim is to read the way a program officer does, weighing whether the application answers this funder's criteria. Analysis is iterative, each pass aware of the last, so a draft and its critique improve together across revisions. Because the judgment is anchored to a real funder and a real organization, not to generic best practice, the suggestions are specific enough to act on, which is the difference between feedback that changes a draft and feedback that merely comments on it.

Two caveats temper this. The scoring is performed by a language model acting as judge (Zheng et al., 2023), a class of scorer with documented biases toward fluency, length, and answer position, so its scores are treated as directional and calibrated against real funding decisions as outcome data accrues (Section 6). And the fit score is only as good as what a funder discloses: where guidance is thin, fit rests on coarser signals such as sector and grant sizes, and is weaker for it. That the same system drafts and then judges an application is a tension we manage by keeping the judge's context and criteria explicit and inspectable, so a user sees

what the score is based on.

6 The Missing Side of the Data

There is a gap at the center of this design, between what the source data contains and what the task needs. The filings describe funders and only funders: their money and the grants they made. They say nothing about the organizations that apply, how those organizations describe their work, or which applications succeed, yet that applicant-side information, and outcomes above all, is what a matching or generation model would most benefit from. A tool that only searches a funder database never sees it. A tool that also hosts the writing does.

Every application drafted in the workspace is an example of how one organization frames its work for a specific funder, and every reported result attaches an outcome (Figure 2). This is the raw material of supervision: outcomes furnish weak, noisy relevance signals for learning-to-rank (Burges, 2010), a reported rejection being far from a clean negative label, and the applicant-side text furnishes reference material for generation. Scarce early labels can be bootstrapped the way production ranking systems are cold-started, by distillation (Hinton et al., 2015) and synthetic queries (Liu et al., 2024), with progress measured by standard ranking metrics (Järvelin and Kekäläinen, 2002). Discovery and drafting are therefore one system, not two products: the coupling is what makes the product improve as it is used.

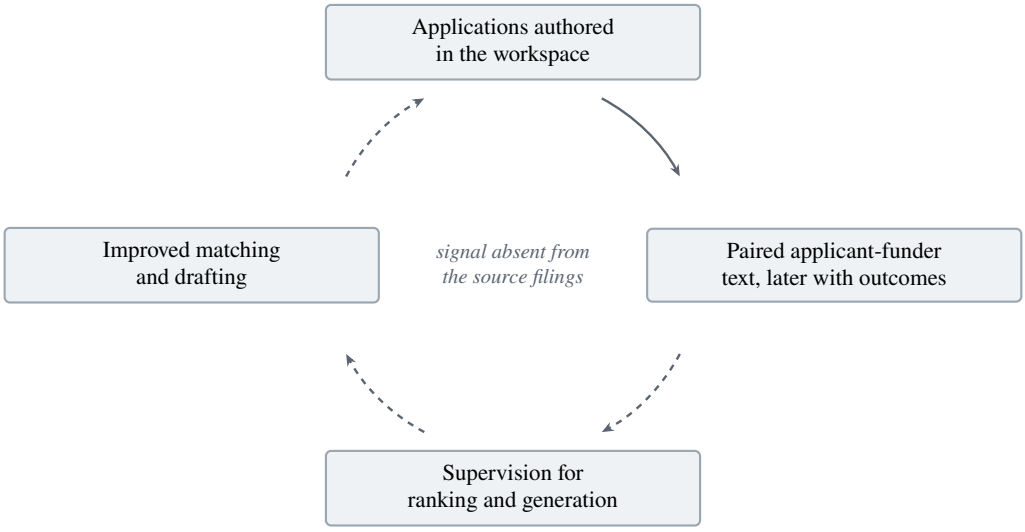

Figure 2: The workspace supplies the missing side. Applications authored in the workspace contribute applicant-side text, and later their outcomes, that the funder filings do not contain, providing supervision for matching and generation. Solid: live today. Dashed: planned, depending on outcome capture and learning.

None of this is free. The outcome label is difficult: decisions take six to eighteen months, so the signal is slow; applicants may never report them, so it is sparse and self-selected; and a rejection can reflect timing, a closed budget cycle, or a prior relationship as much as fit, so it is noisy. The loop is also reflexive, since recommending a funder concentrates applications there and the resulting data inherits the system's own choices, and because the workspace helped write the applications, the corpus carries model-generated text back toward the model, a known path to degradation if left unchecked. These are reasons to treat learned ranking as an augmentation to the structured, explainable discovery of Section 3, not a replacement, and to hold outcomes as one signal among several while the human-legible filters remain the backbone.

The durable advantage is not the 990-PF data, which is public and available to any team willing to parse it, and it is not the language models, which are commercially available to everyone. It is the dataset that only the integration produces. Because discovery, drafting, and audit run in one workspace, the system accrues paired

examples of how organizations describe their work, how they frame it to particular funders, and, in time, which of those framings succeeded. Every individual component described in this paper could be rebuilt by a competitor; the corpus their combination generates cannot be, because it exists nowhere else.

7 Data Governance

Because that advantage comes from customer applications, data handling is part of the design. Applications and outcomes are confidential and competitively sensitive; nonprofits pursuing the same funder are competitors. Three commitments follow. Participation is by consent: an organization's material improves shared models only if it opts in, while its own documents remain available for its own drafting either way. Tenancy is isolated: the system learns aggregate patterns and never surfaces one organization's text, figures, or identifying detail in another's draft; what is reused is learned regularity, not verbatim content (Figure 3). And retention is bounded and revocable, with the officer- and trustee-level records aggregated and access-controlled, not exposed. These commitments rest on per-tenant isolation and encryption, audit logging, training-data lineage, versioned consent, deletion that propagates from storage, and provenance tracking so model-generated text can be held out of the supervision corpus. One limit is stated plainly: removing a document from storage does not remove its influence from already-trained weights, so opt-out governs future training, memorization tests check that rare names or figures are not reproduced, and model unlearning remains an open problem we do not claim to have solved.

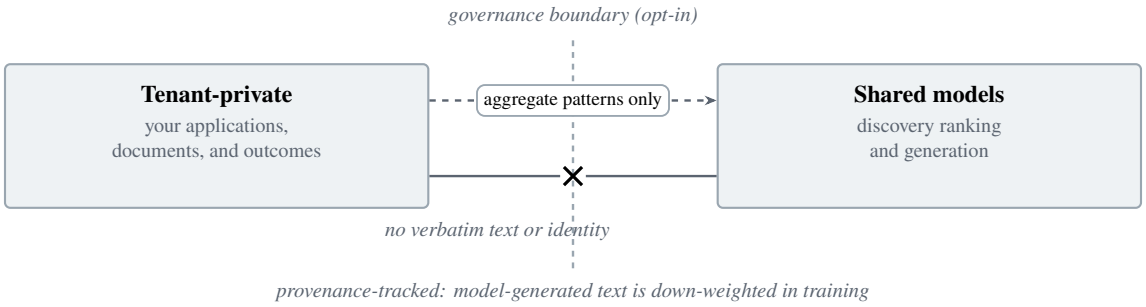

Figure 3: The governance boundary. An organization's applications, documents, and outcomes stay tenant-private and power only its own drafting; only aggregate learned patterns cross into the shared models, never verbatim text, figures, or identifying detail. The crossing is dashed because it is conditional on opt-in and on the outcome-driven learning still being built.

8 Status and Roadmap

The sections above mix what runs today with what is planned. Table 1 shows where each stands.

Live today	On the roadmap
990-PF ingestion and funder model	Recipient-level resolution (validated, not yet shipped)
Structured, behavior-based discovery	Document-store retrieval into drafting (maturing)
Funder research profiles (giving history, people, guidance)	Semantic matching and learned ranking
	Outcome-driven learning
Evidence-backed drafting	Formal evaluation
Draft audit and analysis	
Organizational document vault	

Table 1: Current status of each component of the architecture.

Discovery gains the hybrid retrieve-and-rerank stack and the recipient-level resolution described in Section 3; drafting gains claim-level citation and unsupported-claim detection, so faithfulness becomes measurable rather than assumed; coverage widens across more filing years and adjacent grantmakers, such as community foundations that file under different forms; and a formal evaluation framework, with ranking metrics, faithfulness checks, and tenant-boundary tests against baselines, is being established. We report no results yet. When results come, success should be judged on four user-visible measures: whether recommended funders convert to funding at a higher rate than self-directed prospecting, whether funder research that took days takes minutes, whether audited drafts are measurably stronger than unaudited ones, and, ultimately, whether organizations using the system win more of the funding they pursue. The largest item on the roadmap serves all four: closing the loop of Section 6 by instrumenting outcome capture, so that recorded results become supervision that sharpens matching and drafting over time.

9 Related Work

This design combines three established lines of work rather than extending any of them. Public disclosure of nonprofit filings has long supported philanthropic research and a family of funder databases, such as Candid's Foundation Directory and Instrumentl, built on the same primary sources; we differ in pairing that structured data with an integrated matching and drafting workspace. The discovery roadmap adopts the modern retrieval consensus: lexical and dense first-stage retrieval (Robertson and Zaragoza, 2009; Karpukhin et al., 2020; Reimers and Gurevych, 2019), stronger second-stage rerankers (Nogueira and Cho, 2019), and gradient-boosted learning-to-rank to fuse text signals with structured features (Burges, 2010). Drafting rests on retrieval-augmented generation (Lewis et al., 2020), and the plan for cold-starting rankers before outcome data accrues follows standard practice, distillation (Hinton et al., 2015) and synthetic supervision (Liu et al., 2024), evaluated with established metrics (Järvelin and Kekäläinen, 2002). The contribution is not a new algorithm but an arrangement of these components around a specific data source and a product position, the drafting workspace, that generates the supervision the components need.

10 Conclusion

Grantivo treats grant seeking as a question about funder behavior, not open listings. It structures primary-source filings into a model of what funders have actually funded, uses that model for explainable discovery, and holds a writing assistant to funder facts and an organization's own materials. Its most distinctive feature is structural, not algorithmic: placing the drafting workspace in the path of the work lets the system observe the applicant-side text and outcomes the filings omit, which are among the signals a matching or generation model would most benefit from. The discovery, drafting, and audit it offers today stand on their own. But the reason to prefer this architecture over its alternatives is what it becomes: a funder database can be replicated by anyone with the filings, and a writing assistant by anyone with a model, but a system that sits in the path of the work compounds, because every application it helps write and every outcome it records makes the next match and the next draft better. Tools that only search or only draft have no equivalent mechanism, and the gap widens with use.

References

Burges, C.J.C. (2010) *From RankNet to LambdaRank to LambdaMART: an overview*. Technical Report MSR-TR-2010-82. Microsoft Research.

Christen, P. (2012) *Data matching: concepts and techniques for record linkage, entity resolution, and duplicate detection*. Berlin: Springer.

Hinton, G., Vinyals, O. and Dean, J. (2015) 'Distilling the knowledge in a neural network', *NIPS Deep Learning Workshop*.

Internal Revenue Service (n.d.) *Tax Exempt Organization Search (TEOS) bulk data downloads*. Washington, DC: U.S. Department of the Treasury.

Järvelin, K. and Kekäläinen, J. (2002) 'Cumulated gain-based evaluation of IR techniques', *ACM Transactions on Information Systems*, 20(4), pp. 422–446.

Karpukhin, V., Oğuz, B., Min, S. et al. (2020) 'Dense passage retrieval for open-domain question answering', *Proceedings of EMNLP 2020*, pp. 6769–6781.

Lewis, P., Perez, E., Piktus, A. et al. (2020) 'Retrieval-augmented generation for knowledge-intensive NLP tasks', *Advances in Neural Information Processing Systems 33*.

Liu, R., Wei, J., Liu, F. et al. (2024) 'Best practices and lessons learned on synthetic data for language models', *Proceedings of COLM 2024*.

National Center for Charitable Statistics (n.d.) *The National Taxonomy of Exempt Entities (NTEE) classification system*. Washington, DC: Urban Institute.

Nogueira, R. and Cho, K. (2019) 'Passage re-ranking with BERT', *arXiv preprint* arXiv:1901.04085.

Reimers, N. and Gurevych, I. (2019) 'Sentence-BERT: sentence embeddings using Siamese BERT-networks', *Proceedings of EMNLP 2019*, pp. 3982–3992.

Robertson, S. and Zaragoza, H. (2009) 'The probabilistic relevance framework: BM25 and beyond', *Foundations and Trends in Information Retrieval*, 3(4), pp. 333–389.

Zheng, L., Chiang, W.-L., Sheng, Y. et al. (2023) 'Judging LLM-as-a-judge with MT-Bench and Chatbot Arena', *Advances in Neural Information Processing Systems 36*.