Automating Systematic Literature Reviews (SLR) in Oncology: Development and Performance of a Proprietary Data Management and Maintenance System (DMS)

Rozee Liu[1], Oscar Correa[2], Sergio Zapatel[2], Evgeniya Correa[2], Anna Forsythe [1]

[1]Oncoscope-AI, Miami, FL, USA; [2.] Eviviz Inc., Vancouver, BC, Canada

BACKGROUND

- The accelerating volume and complexity of oncology evidence have made traditional manual SLRs increasingly difficult to conduct, maintain, and update
- Living SLR is a common concept of continuously updating an SLR as evidence becomes available. However, there is no standardization regarding the frequency of SLR updates in order to be qualified as a living SLR
- We developed a new concept of Real-Time AI-assisted Living Systematic Literature Review (REAL-SLR), where the SLR libraries are updated daily with the support of an agentic AI system and a data management system (DMS)

OBJECTIVES

→ This study describes the design, implementation, and performance of a proprietary DMS developed to automate the most resource-intensive SLR components while preserving methodological rigor, transparency, and auditability

METHODS

- The DMS architecture is built on a NoSQL database with a data editor supporting full CRUD (Create, Read, Update, Delete) operations (**Figure 1**)
- Automated components include structured search execution in PubMed + conferences, deterministic rule-based deduplication, and automated master records (MRs) creation for each unique study
- An integrated agentic large language model (aLLM) system with retrieval-augmented generation (RAG) architecture supports screening + extraction, with human quality control and strategic data aggregation
- Each master record is cross-linked to external regulatory, guideline, and reimbursement resources

RESULTS

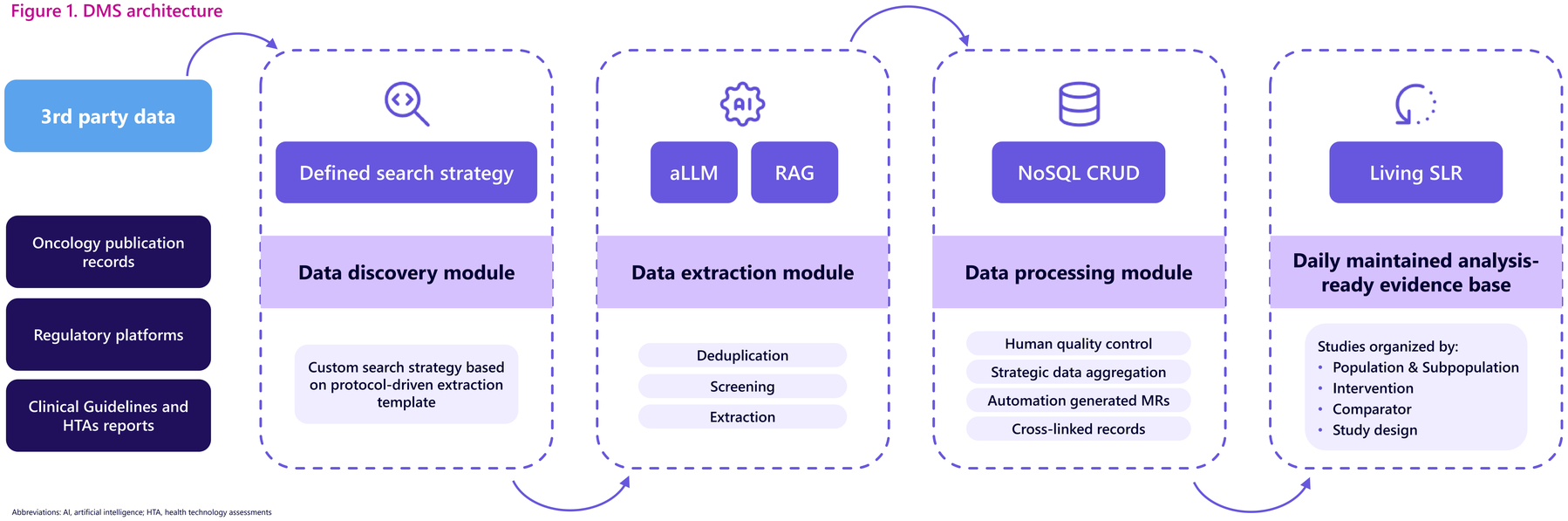
- System implementation resulted in three major efficiency gains. First, automated search execution, deduplication, screening, and data extraction are pre-scheduled and completed in ~5 minutes, compared with 6 hours required for a manual daily update (98.6%-time reduction) (**Figure 2**)
- Second, the DMS automatically generates MRs that aggregate study/publication-level data while prioritizing recent/relevant efficacy; human review of a pre-generated MRs requires ~5 minutes, compared with 15 minutes for manual aggregation (66.7%-time reduction)
- Third, the DMS functions as a centralized data governance framework, passively maintaining data hygiene, versioning, and linkage of publications to unique studies (~20% reduction vs. ongoing project management and quality control effort)
- The system currently supports Living-SLRs containing 3,898 unique studies across four cancers: 1,527 non-small-cell lung, 924 breast, 738 prostate cancer, and 709 multiple myeloma, with new evidence incorporated daily

CONCLUSIONS

→ A purpose-built DMS substantially improves the scalability, efficiency, and sustainability of oncology SLRs

→ Combining programmatic workflows, rule-based automation, and aLLM/RAG-assisted data review and extraction enables creation and maintenance of a true Living SLR, thereby supporting faster HTA and regulatory decision making in rapidly evolving cancers

Figure 1. DMS architecture



Abbreviations: AI, artificial intelligence; HTA, health technology assessments

Figure 2. Time savings comparing DMS vs. human researchers