

Title: *From Data Entry to Discovery: AI-Driven Transformation in Research Information Management*

Abstract Summary

The landscape of Research Information Management (RIM) is undergoing a significant paradigm shift driven by the integration of Artificial Intelligence (AI). Institutions and national consortia face persistent challenges in maintaining complete researcher profiles, reducing administrative burden, and effectively communicating research impact to external audiences. This presentation explores how AI technologies are being deployed within Current Research Information Systems (CRIS), specifically Symplectic Elements, to address the "manual entry barrier" and the accessibility of complex research outputs.

We will examine the evolution of workflow automation, from the current capability of using Large Language Models (LLMs) to extract structured metadata from free-form text snippets, such as CV entries, award letters, and pre-publication manuscripts. Furthermore, we will analyse best practices for intelligent metadata enrichment, demonstrating how AI-generated plain-language summaries can maximize the discoverability of research by non-specialists. Finally, the session will outline a rigorous framework for "Responsible AI," focusing on human-in-the-loop governance, duplicate detection, and resource management.

Detailed Session Overview

1. The Challenge: The Manual Entry Barrier and Data Completeness

Despite the maturation of CRIS infrastructures, institutions struggle with a fundamental problem: the reliance on manual data entry for non-traditional outputs. While STEM publications often benefit from automated harvesting via DOIs, researchers in the Arts, Humanities, and Social Sciences frequently rely on manual data entry for creative works and non-standard outputs.

Furthermore, critical institutional data regarding "Professional Activities" (e.g., committee memberships, editorial boards) and "Teaching Activities" (e.g., mentoring, course leadership) are rarely captured in external databases. These activities often exist only on a faculty member's CV, leading to significant information gaps in institutional reporting and faculty assessment. This session will quantify the administrative burden of this manual entry, citing user research indicating that some institutions spend up to 20 hours per new faculty member to manually curate profiles.

2. The Technological Solution: Intelligent Metadata Enrichment & Decision Support

- The presentation will introduce advancements in AI-Assisted Data Entry developed by Symplectic Elements. This technology leverages LLMs to parse unstructured text and convert it into structured institutional records, e.g. workflow automation, scope of application, data quality via duplicate detection.

- Impact: it demonstrates how intelligent metadata enrichment encourages exploratory search behaviors and guides users toward relevant expertise, effectively turning a CRIS from a passive platform to collect outputs into an active decision-support tool.

3. The Gains: From Snippets to Bulk data processing

Building on these foundational capabilities, we will preview upcoming advancements in bulk data processing designed to support high-volume administrative tasks, such as onboarding new faculty, preparing for tenure and promotion (P&T) reviews, or managing national assessment exercises where historical data completeness is critical.

We will discuss the ability to ingest and process complete historical documents in a single workflow, allowing institutions to populate extensive lists of outputs and activities efficiently while maintaining strict accuracy standards.

4. Best Practices: Ethical and Responsible Use of AI

A central theme of the presentation is that AI must be implemented with strict governance safeguards. We will outline the "Digital Science Approach to Responsible AI".

- Human-in-the-Loop: The cardinal rule of this infrastructure is that no AI-generated data is saved or modified without explicit user confirmation. The AI acts as a drafter, not a decision-maker. Users must review extracted values and have the ability to edit them (via a pre-populated manual form) before they become part of the institutional record.
- Administrative Control & Configuration: Institutions require granular control. Best practices include the ability to toggle AI features on/off globally or for specific data categories (e.g., enabling it for Publications but disabling it for Grants)
- Data Privacy: The presentation will clarify that client data remains secure and is not used to train or fine-tune third-party public models, addressing a primary concern for research institutions.

Keywords: Artificial Intelligence, Research Information Management (RIM), Data Quality, Workflow Automation, Responsible AI, Research Discoverability, Current Research Information Systems, CRIS.

Target Audience: Research Administrators, Library Directors, CRIS Managers, Research Strategy Officers, institutional leadership involved in research strategy and assessment, and Technologists in Higher Education.

Shorter Abstract:

The landscape of Research Information Management (RIM) is undergoing a significant paradigm shift driven by the integration of Artificial Intelligence (AI). Institutions and national consortia face persistent challenges in maintaining complete researcher profiles, reducing administrative burden, and effectively communicating research impact to the public and industry. This presentation explores how AI technologies are being deployed within Current Research Information Systems (CRIS), specifically Symplectic Elements, to address the "manual entry barrier" and the accessibility of complex research outputs.

We will examine the evolution of workflow automation, beginning with the current capability to utilise Large Language Models (LLMs) for extracting structured metadata from free-form text snippets, such as CV entries, award letters, and pre-publication manuscripts. By automating the capture of "challenging" data categories like professional activities, teaching, grants and Non traditional research outputs (NTRs) institutions can ensure a more comprehensive record of researcher activity.

Building on these foundational capabilities, we will preview upcoming advancements in bulk data processing designed to support large researcher onboarding and assessment preparation. We will discuss the ability to ingest and process complete historical documents in a single workflow, allowing institutions to populate extensive lists of outputs and activities efficiently while maintaining strict accuracy standards.

Finally, we will address the ethical and responsible use of AI in RIM infrastructures. The session will outline a governance framework that prioritises the "Human-in-the-loop" principle, ensuring no AI-generated data is committed to the institutional record without user confirmation. Attendees will leave with a comprehensive understanding of how responsibly implemented AI can drive data completeness, enhance institutional reporting, and maximise the societal impact of research.

Keywords: Artificial Intelligence, Research Information Management (RIM), Data Quality, Workflow Automation, Responsible AI, Research Discoverability, CRIS.

Target Audience: Research administrators, CRIS managers, library professionals, and institutional leadership involved in research strategy and assessment.