26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

# Do androids dream of workflow provenance?
## An intelligent interface to the Provena metadata system

**Peter Baker [c], Thomas Shewan [d], Jonathan Yu [b], Ben Leighton [b] and Ross Petridis [a]**

[a] *CSIRO NCMI, Hobart, Tas, Australia,* [b] *CSIRO Environment, Clayton, Vic, Australia,*
[c] *CSIRO Environment, Black Mountain, ACT, Australia,* [d] *Western Sydney University, NSW, Australia*

**Abstract:** Verifying data applicability, veracity and integrity often requires scrutinising the holistic process that produced it—its provenance. While mature automated workflows can generate provenance through artifacts like logs and metadata, scientific modelling is frequently heterogeneous, ad-hoc, and incompletely documented, leaving provenance patchy or absent. Modelling teams face difficult choices: accept gaps in provenance, manually curate metadata, or invest significant effort uplifting and retrofitting models into structured workflows that emit automated provenance artifacts. We propose, and implement an early demonstrator of, an AI-enhanced read-write interface to the Provena metadata system, leveraging Large Language Models and the Model Context Protocol to reduce the intellectual labour required to create and understand workflow provenance. This conversational interface would bridge the gap between unstructured modelling practices and formal provenance systems, enabling researchers to focus on environmental science rather than metadata management.

**Keywords:** *Data provenance, Large language models, Metadata, Workflows, Model context protocol*

## 1. BACKGROUND

Data provenance is critical for environmental decision-making, providing the documented history that rationalises decisions and strengthens their defensibility. Current approaches to provenance creation fall into two categories: manual curation (time-consuming, error-prone, inconsistent) and automated generation through workflow orchestration systems or similar approaches (requires significant engineering expertise and infrastructure investment) (Petridis et al., 2025).

The Reef Restoration Adaptation Program (RRAP) experiences these challenges. Ongoing efforts, since 2020, have been made to achieve operational-readiness in modelling including robust provenance, facilitated through the Provena metadata system (Yu et al., 2023). However, barriers include building sufficient adoption for users to see benefit, expertise and resourcing gaps across specialised engineering skillsets, challenges encoding ad-hoc modelling practices into structured records, and underlying engineering limitations that make automation difficult.

These challenges create an unforgiving landscape where modelling teams lacking technical expertise and maturity are forced to either manually create provenance records or attempt costly uplifts of existing models into workflow automation systems—neither approach scaling effectively across diverse research contexts.

## 2. AI-ENHANCED PROVENA INTERFACE

Large Language Models present opportunities for augmentation across the entire modelling and provenance lifecycle, from conceptualisation through deployment and provenance registration and interpretation. We focus specifically on leveraging the Model Context Protocol (MCP) (Anthropic, 2025)—a standardised interface allowing servers to expose tools, resources, and prompts to AI clients—to create an intelligent interface to Provena's APIs through the Provena Python Client (Baker et al., 2025).

The MCP specification enables interoperable AI access to tools (functions for database queries and API calls), resources (contextual data like documentation and schemas), and prompts (curated template prompts for specific interaction patterns). For Provena, exemplar use cases include step-by-step registration wizards, workflow lineage exploration, report generation and system documentation access and summarisation. LLM interfaces also offer new input modalities such as multi-media and speech-to-text, which are worth further exploration.

This approach proposes to address Provena's complexity by providing a more informed, flexible and natural conversational interface, capable of reasoning (or ‘dreaming’) about provenance. Currently, registering a model run requires understanding dependency chains (e.g. model run requires a workflow template requires a model and dataset templates) and the subsequent proper sequencing of registration steps. An MCP-powered interface can guide users through these processes naturally, reducing cognitive overhead while maintaining data quality.

Our early demonstrator uses the Python tool FastMCP, with the Provena Python Client to build a simple MCP server. Early testing shows promise, demonstrating the ability for tool understanding and multi-step sequencing with minimal prompting. The conversational interface can guide a user through dataset registration processes while providing contextual help, automation and validation (such as in finding related records and formatting dates).

## 3. IMPLEMENTATION CONSIDERATIONS

Operational deployment requires addressing several architectural challenges including security and user session management, cost management and configurable LLM model selection. Customised user interfaces may be necessary to provide optimal steering through system prompts, contextualised task flows, and Provena-specific interface elements.

Ethical considerations mandate explicit user consent for AI data processing, updated terms and conditions, and careful handling of data sovereignty requirements given that the Australian government and the general public prioritise Australian data localisation (Economics References Committee, 2023). Trust, liability and correctness in AI augmented systems are critical challenges in a rapidly moving legal landscape (Novelli et al., 2024).

Success criteria and metrics for such a system remain an open research question requiring stakeholder consultation. Potential measures include registration time reduction, metadata completeness or quality improvement, user satisfaction ratings, error frequency, and confidence and correctness in provenance interpretation. The challenge lies in balancing efficiency gains with record quality and determining which success criteria drive solution development in alignment with the end-goal of supporting the defensibility of decision-making.

## 4. CONCLUSION

The integration of AI into structured metadata systems like Provena represents a promising approach to bridging the gap between ideal comprehensive provenance and practical modelling realities. By providing intelligent, conversational interfaces to complex registration and exploration processes, we can potentially reduce barriers to adoption while maintaining or improving metadata quality.

However, success depends on thoughtful implementation addressing data sovereignty, user consent, liability frameworks, and clear value propositions. The path forward requires careful validation through controlled testing, stakeholder engagement, and iterative refinement of AI steering approaches to ensure genuine utility rather than additional complexity.

Do androids dream of workflow provenance? Perhaps they should. AI-augmented systems bringing a human touch to the structured complexity of metadata systems shows promise in democratising access to workflow provenance across the full spectrum of modelling teams, ultimately enabling more transparent and defensible decision-making for critical environmental challenges like Great Barrier Reef restoration and adaptation.

## 5. ACKNOWLEDGEMENTS

This work was funded by Reef Restoration and Adaptation Program (RRAP), which is funded by the partnership between the Australian Government's Reef Trust and the Great Barrier Reef Foundation.

## 6. REFERENCES

Anthropic. (2025). *MCP Specification*. Model Context Protocol. https://modelcontextprotocol.io/specification/2025-06-18

Baker, P., Kulkarni, P., Petridis, R., & Yu, J. (2025). *Provena Python Client v0.25.0*. https://doi.org/10.25919/xvrn-gz28

Economics References Committee. (2023, November). *Economics References Committee: Influence of international digital platforms* (Australia) . https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Economics/Digitalplatforms/Report/Chapter_5_-_Data

Novelli, C., Casolari, F., Hacker, P., Spedicato, G., & Floridi, L. (2024). Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity. *Computer Law & Security Review*, *55*, 106066. https://doi.org/10.1016/j.clsr.2024.106066

Petridis, R., Yu, J., Baker, P., Munnerly, J., & Mirza, F. (2025, September). *Towards seamless provenance in collaborative modelling: Integrating Airflow with Provena*. To appear: MODSIM 2025.

Yu, J., Baker, P., & Cox, S. J. D. (2023). Provena: A provenance system for large distributed modelling and simulation workflows. *MODSIM 2023*. https://doi.org/10.36334/modsim.2023.yu90

 