



Contribution ID: 444

Type: Talk

Intelligent Metadata: Harnessing AI for Discovery, Visibility, and Preservation in the Digital Library

The University of Pretoria's Department of Library Services (DLS) has developed an inhouse Artificial Intelligence (AI) Toolkit to enhance metadata creation and enrich digital collections. Aligned with the theme "From Data to Decisions: Leveraging Cyber Infrastructure," this initiative uses AI to strengthen discovery, visibility, and preservation across institutional repositories and special collections.

The toolkit integrates multiple AI capabilities—image and video analysis, scene detection, emotion recognition, transcription, keyword generation, Sustainable Development Goals (SDG) classification, and object detection—to automate and improve metadata quality. Through scene and object detection, audiovisual materials are segmented and described according to identifiable elements such as people, settings, and activities. Emotion recognition provides contextual insight into affective dimensions in recordings, while speech-to-text transcription and keyword extraction enable automated summaries and thematic indexing, reducing manual cataloguing effort and improving search accuracy.

The inclusion of SDG classification models, trained on BERT-based architectures, links research and digital outputs to global sustainability goals, supporting institutional reporting and impact measurement. The toolkit's modular architecture ensures integration with existing library systems and adherence to privacy and ethical standards while promoting scalability across various workflows, from digitisation to research dissemination.

By embedding AI into metadata workflows, the DLS transforms static data into actionable insight, aligning digital preservation with decision-making and strategic research visibility.

The outcomes include:

Page 2 of 2

- Improved discoverability and interoperability of digital assets.
- Increased efficiency in metadata generation.
- Enhanced contextual understanding of multimedia collections.
- Strengthened alignment with global open scholarship and sustainability frameworks.

This initiative demonstrates how AI-driven metadata enrichment can empower libraries to move beyond traditional curation, positioning them as key actors in data-driven innovation and digital scholarship within the modern academic ecosystem.

Presenting Author

Mr Isak van der Walt

Email

isak.vanderwalt@up.ac.za

Student or Postdoc?

Institute

University of Pretoria

Registered for the conference?

Yes

CHPC User

Yes

CHPC Research Programme

Primary author: Mr ISAK, van der Walt

Presenter: Mr ISAK, van der Walt

Session Classification: DIRISA

Track Classification: DIRISA