Information Research - Vol. 31 No. iConf (2026)

Integration of generative artificial intelligence (GenAI) in metadata training

Oksana L. Zavalina

DOI: https://doi.org/10.47989/ir31iConf64272

Abstract

Introduction. GenAI tools are adopted by information professionals globally, including GenAI incorporation into metadata workflows in archives and libraries. Researchers and practitioners emphasise the critical need for GenAI learning integration into coursework to prepare iSchools graduates for this emerging professional demand. While reports on GenAI content integration in iSchools curricula are increasing, there is a lack of publications about GenAI integration in courses focusing on digital repository metadata.

Method. University of North Texas recently integrated GenAI practical metadata learning in an advanced graduate course. This paper describes this integration and reports results of its testing.

Analysis. The author performed qualitative and quantitative analyses of practical assignment submissions in which students used GenAI tools for generating metadata records and evaluated the process and outcomes.

Results. AI-generated metadata varied in completeness depending on the GenAI tool, its version, and prompt used, while typically lacking accuracy. Students expressed appreciation of this practical experience, reported developing greater confidence in using GenAI tools and understanding their advantages and disadvantages in metadata work, demonstrated development of analytical and troubleshooting skills.

Conclusion. This paper is expected to be useful for iSchools metadata educators, researchers, and practitioners.

Introduction

Descriptive metadata is crucial for information discovery and reuse, and libraries, archives, and museums (LAMs) have been responsible for descriptive metadata creation for a long time, following the metadata standards developed by this professional community. iSchools help future information professionals develop the needed knowledge and skills in this important area through library cataloguing and digital repository metadata coursework.

Generative Artificial Intelligence (GenAI) tools are rapidly evolving, with new versions and improvements spurred by fierce competition among the major players of GenAI market. Integration of GenAI into metadata workflows holds promises of accelerating metadata production (for example, Gordon, 2025). Multiple recent publications detailed results of testing GenAI tools in individual metadata-related tasks, including image captioning (for example, Torabi, Emdad, & Pashootanizadeh, 2025, etc.) and generation of subject terms and classification codes to represent aboutness of information resources (for example, Bodenhamer, 2023; Chou & Chu, 2022; Chow, Kao & Li, 2024; Ganadi et al., 2023; Martorana et al., 2024; Zhang, Wu & Zhang, 2023). The reports also emerge that present outcomes of applying GenAI tools to generating entire metadata records in different metadata schemes. These experiments (for example, Brzustowicz, 2023; Zavalin & Zavalina, 2025) emphasise the need for quality evaluation of GenAI output by human metadata experts and the current suitability of GenAI tools for assisting human metadata creators similar to other metadata automation tools of the past rather than for taking over the metadata creation tasks.

In publications that review directions of GenAI integration in metadata (for example, Chen & Li, 2024; Moulaison-Sandy & Coble, 2024; Sussmeier & Henry, 2025), interviews with metadata experts (for example, Alemu & Tamarro, 2025), and panel discussions (for example, Hoffman et al., 2024), an important theme is the need for iSchools metadata curricula to keep up with all these rapid developments. While LAM educators are increasingly integrating GenAI learning in their courses (for example, Kizhakkethil & Perryman, 2024; LaDell-Thomas, 2025), so far, there are no published reports on such integration in digital repository metadata courses. Yet, the need for exchanging such curriculum innovation experiences and collaborative development of best practices is pressing. This paper begins addressing this need, by sharing early results of GenAI learning in digital repository metadata courses at the University of North Texas.

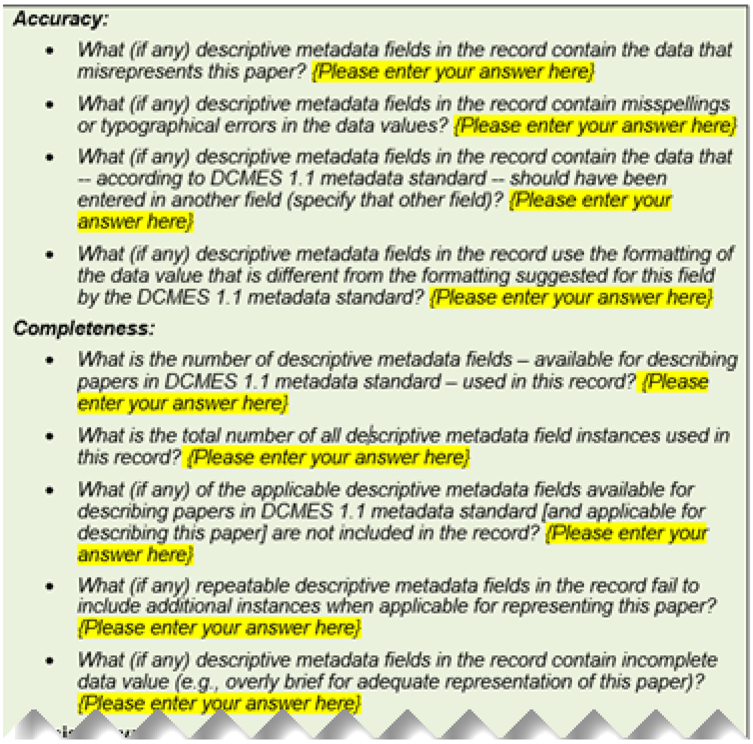
Method

At our iSchool, digital repository metadata quality evaluation is taught in the graduate course INFO 5224, which is offered in 16-week-long semesters once a year and has 2 prerequisites that focus on organising information and digital library metadata. In this advanced graduate metadata course, the GenAI practical work is included since Spring 2025 semester. It is integrated with activities of the second learning module, which is devoted to metadata quality control. In the practical assignment, students first collect small random samples of metadata records from two same-repository digital collections with different types of information resources and comparatively evaluate quality of metadata records in these samples. To generate a metadata record with the GenAI tool of their choice, students are asked to pick one of the readings – papers published in journals and conference proceedings – that they became familiar with in discussion assignments. To ensure that in the metadata creation process, the GenAI tool has access to the full text of information resource, students are instructed to select a paper that is freely available online without logging-in.

In the practical assignment, students are asked to generate XML-encoded Dublin Core (DCMES 1.1) metadata record and are provided with the following instructor-constructed prompt to use:

'Create metadata record to represent this information resource: [insert here the URL that leads to the paper]. Follow the Dublin Core Metadata Element Set 1.1 metadata standard: [*https://www.dublincore.org/specifications/dublin-core/dces/*](https://www.dublincore.org/specifications/dublin-core/dces/)*. Encode the metadata record in* XML.'

They are responsible for revising this initial prompt if it does not produce expected results. Also, graduate students are required to assess resulting metadata record's accuracy, completeness and consistency and answer 3-5 specific questions per metadata quality criterion using the template shown in Figure 1.



Accuracy:

- *What (if any) descriptive metadata fields in the record contain the data that misrepresents this paper? {Please enter your answer here}*
- *What (if any) descriptive metadata fields in the record contain misspellings or typographical errors in the data values? {Please enter your answer here}*
- *What (if any) descriptive metadata fields in the record contain the data that -- according to DCMES 1.1 metadata standard -- should have been entered in another field (specify that other field)? {Please enter your answer here}*
- *What (if any) descriptive metadata fields in the record use the formatting of the data value that is different from the formatting suggested for this field by the DCMES 1.1 metadata standard? {Please enter your answer here}*

Completeness:

- *What is the number of descriptive metadata fields – available for describing papers in DCMES 1.1 metadata standard – used in this record? {Please enter your answer here}*
- *What is the total number of all descriptive metadata field instances used in this record? {Please enter your answer here}*
- *What (if any) of the applicable descriptive metadata fields available for describing papers in DCMES 1.1 metadata standard [and applicable for describing this paper] are not included in the record? {Please enter your answer here}*
- *What (if any) repeatable descriptive metadata fields in the record fail to include additional instances when applicable for representing this paper? {Please enter your answer here}*
- *What (if any) descriptive metadata fields in the record contain incomplete data value (e.g., overly brief for adequate representation of this paper)? {Please enter your answer here}*

Figure 1. Instructor-provided template for assessing GenAI-generated metadata record's accuracy and completeness by graduate students.

Anonymised student submissions of the practical assignment's GenAI component were analysed by the author of this paper (who developed the course and assignment and was also the course instructor in Spring 2025 semester) quantitatively and quantitatively for the following indicators:

- distribution of student chosen GenAI tools and versions
- frequency of initial prompt revisions, and
- number and types of student-identified mistakes in GenAI-created metadata.

In addition, the overall student perception on GenAI experience in the practical metadata exercise was evaluated based on the free-text comments. This includes the university-administered anonymised standard survey for assessing course effectiveness. Although survey results were available for the Spring 2025 semester (with 87% response rate), the survey did not contain specific questions about the GenAI content. In addition to objective measurements based on instructor-provided analysis template, some students also chose to include with their assignment submission the comments on their perception of GenAI metadata generation, which have been analysed as context information to help interpret the objective metadata quality analysis findings.

In line with ethical research requirements, the findings presented here show the quantitative data analysed and aggregated by the author (developer and instructor of the course), as well as a small selection of anonymised illustrative quotes from student answers to some of the assignment questions. The author confirmed that no formal Institutional Review Board approval is needed for such analyses and dissemination of research results and ensured compliance with university policy's requirements to meet all 5 criteria when student work is used: (1) the work is used only once, (2) the work is not used in its entirety, (3) use of the work does not affect any potential profits from the work, (4) the student is not identified, and (5) the work is identified as student work.

Some of the findings of the content analysis of student submissions from the practical assignment integrating GenAI-assisted metadata creation tasks and evaluation of outputs are presented and discussed in the next section of this paper.

Findings and discussion

In the 16-week Spring semester (January – May 2025) in which GenAI learning was first incorporated into the INFO 5224 course, a total of 14 students submitted the Module 2 practical assignment which included the GenAI-focused tasks. One student failed to generate the metadata record and to indicate the GenAI tool used, so the data in Table 1 is based on 13 submissions. One more student did not identify any accuracy or completeness errors (out of several present in the Gemini-generated record), so the numbers of identified errors (shown in Table 1) exclude the data from that student's submission.

Most of our graduate students (almost 70%) relied on ChatGPT in metadata generation: either version 4.0, also referred by students to as 4 or 4o (38.46%) or unspecified version (30.77%). Microsoft Copilot and Google Gemini were used by 2 students each. It is a somewhat unexpected observation, given how ubiquitous and easily accessible are Gemini and Microsoft Copilot, available for free on one's computer or in Google. We attribute the overall student preference for ChatGPT in student metadata generation to the name of this GenAI tool being immediately recognisable due to its continuous significant media and social media coverage since approximately 2022.

Five graduate students out of 13 (38.46%) reported needing to reformulate the initial prompt. It took them between 2 and 12 prompt reformulation attempts to achieve expected results. For example, in one of the two reported cases of using Google Gemini (50%), this GenAI tool did not produce metadata record with expected level of detail or for the current information resource, and the student (after also unsuccessfully trying Claude), moved to using ChatGPT. In part due to previous experience with human-created metadata evaluation preceding the GenAI-produced metadata evaluation in the same assignment, all graduate students (100%) identified some errors in the metadata records created by GenAI tools.

GenAI tool (version)	% of students (n=13)	Student-reported accuracy errors (n=12)	Student-reported completeness errors (n=12)
ChatGPT (unspecified)	30.77%	3-6	1-10
ChatGPT (4.0 or 4o or 4)	38.46%	1-7	0-5
Google Gemini (unspecified)	15.38%	8	11
Claude.ai	7.69%	no results generated	no results generated
Microsoft CoPilot (1.25014.121.0)	7.69%	5	4
Microsoft CoPilot (unspecified)	7.69%	11	11

Table 1. GenAI tool and version selection, accuracy and completeness of metadata generation output as reported by students.

The number of accuracy errors reported by students ranged from 3 to 11 per metadata record (with the highest reported for Microsoft CoPilot). Our graduate students observed some non-Dublin-Core metadata fields (for example, Advisor) in GenAI-generated records. It is important to note that most of our students, who are only starting to gain experience as metadata evaluators, did not notice some additional accuracy errors in the GenAI-created metadata records beyond the ones they identified.

ChatGPT and Google Gemini were generally found to produce less errors, some of records generated by them had significant flaws. For example, the record shown in Figure 2 and created by ChatGPT had a number of quality issues. This included not following the DCMES 1.1 Dublin Core metadata scheme in 3 metadata elements, and lack of a required component of XML encoding (XML declaration), without which machines cannot process this XML -encoded metadata record for information retrieval.

The highest number of completeness errors reported by graduate students was 11. This level of errors was observed for Microsoft CoPilot. Two student submissions reported no completeness errors (n=0): one student who used ChatGPT version 4o did not find any completeness errors, and another (using the same tool and version) mistakenly restated the accuracy mistakes in the assignment section intended for documenting completeness errors. This does not mean however that these ChatGPT-generated descriptive metadata records contained no completeness errors at all: just like with accuracy errors, our students for most of whom this was only the second experience in metadata evaluation (after assessing the quality of human-created metadata) tended to miss one or more completeness errors in GenAI-generated descriptive metadata. Also, for one of these two students, evaluating completeness was meaningless as the ChatGPT GenAI output was for the wrong information resource.

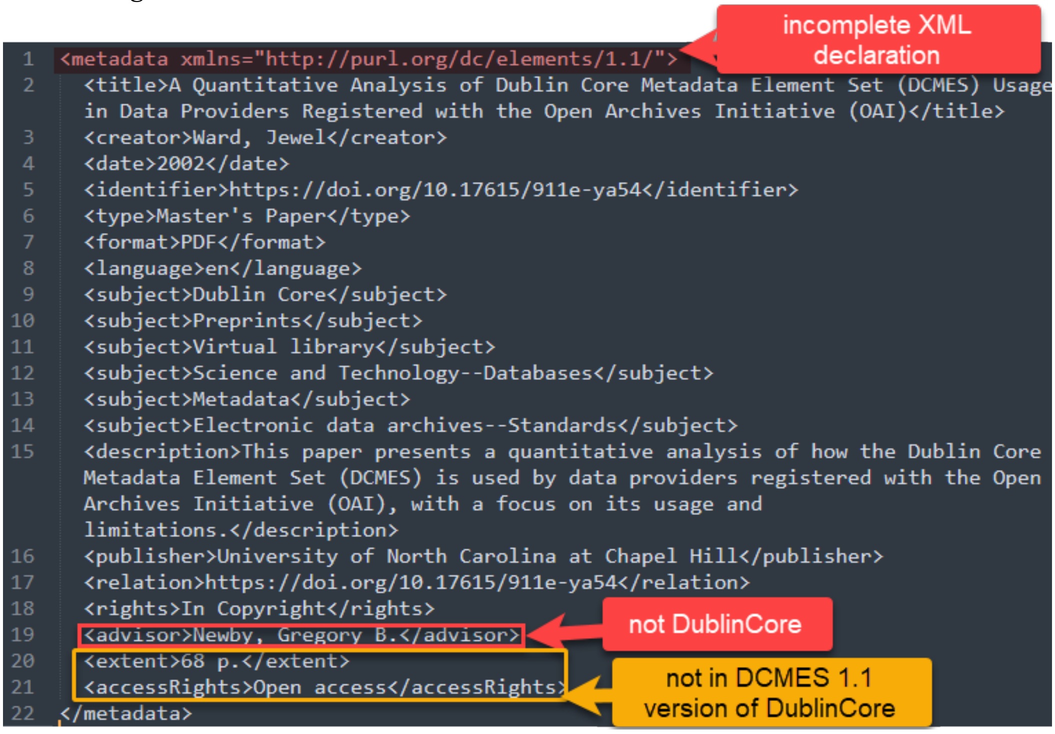


Figure 2. Example of DCMES 1.1 record with significant flaws: generated by ChatGPT-4 in response to student's prompt.

Three students (21.43%) chose to share detailed feedback about the experience generating Dublin Core metadata with GenAI tools and evaluating its quality. They reported that the trial-and-error process taught them a lot about importance of prompt engineering and made them more confident in using the GenAI tools in general and specifically for assistance with metadata creation tasks, including application of controlled vocabularies. For example:

> This is the first time I have really worked with AI, so I spent about 30 minutes playing with it to see how much tweaking and explaining it would take to get Gemini to give me a high-quality metadata record. Below is the best result I could get after about 12 adjustments to the initial prompt. The record is more complete, and it appears mostly accurate […]. I was able to prompt Gemini to use controlled vocabularies. It did properly apply LCNAF to creator and contributor fields, MIME to the format field, and ISO 8601-1 to date fields. (Student 10).

Graduate students also shared how they creatively overcome the technical barriers that were preventing GenAI tools from producing expected output, for example:

> It is possible ChatGPT was having issues accessing all the URLs I was sending it to, even though I verified that the resource was accessible without having to use a login. So, I decided to upload the resource in the link in my prompt above directly to ChatGPT as a PDF […]. Here is the prompt I used for the record created from the uploaded PDF: Create a metadata record to represent the attached PDF and follow the Dublin Core metadata element set 1.1 metadata standard found here https://www.dublincore.org/specifications/dublin-core/dces/ and encode the metadata record in XML. (Student 9).

Overall, our findings demonstrate that the practical assignment which includes GenAI metadata generation and evaluation tasks helps students develop not only analytical skills but also troubleshooting skills related to integration of GenAI in metadata workflows.

Conclusions

This paper presents a selection of early results of the ongoing project, based on the initial small dataset. Nevertheless, it is important to share these early results with colleagues – metadata researchers and educators at iSchools worldwide, as well as practicing information professionals – to begin the conversation and collaborative development of best practices in these time-sensitive iSchool curriculum enhancement efforts. The author hopes that this paper will stimulate exchange of ideas that would help the iSchools to continue meeting the rapidly evolving demands of the information profession.

In addition to expected results of our curriculum experiments – greater level of comfort with GenAI and understanding of strengths and weaknesses of GenAI in metadata generation – an added value outcome was observed. While due to the course design, prerequisites, and other academic experiences in iSchool programs, graduate students in our course have substantial levels of traditional descriptive metadata creation skills and experience with metadata quality evaluation before evaluating GenAI-created descriptive metadata, the levels of previous exposure to GenAI and its use in metadata generation do not depend on these variables. However, engagement with GenAI metadata generation and evaluation in practical assignments was found to spur students' critical thinking and technical creativity.

The more in-depth analyses on the data collected in the first semester (16-week Spring 2025) in which we tested GenAI integration in the advanced graduate metadata course at our iSchool are ongoing. In addition, starting 2025, we integrated GenAI-assisted metadata creation tasks and evaluation of outputs in the introductory undergraduate course in 16-week Fall 2025 and 3-week Spring 2025 and Spring 2026 semesters. These tasks are required in long semesters and optional for extra credit in short semesters. Just like in the INFO 5224 graduate metadata course, our undergraduate metadata students also create Dublin Core DCMES 1.1. XML-encoded metadata records for text resources (in the case of undergraduate course: the same children's eBook for entire class). We are in the process of analysis of GenAI learning results in the 3 semesters of our undergraduate metadata course, with the plans to compare findings to those for GenAI-assisted DCMES 1.1. metadata generation for research publications. These comparative analyses will produce more robust and generalisable results. We are planning to use comparative analysis findings not only for research purposes but also importantly for refining practical assignments to optimise the development of these important skills.

Our GenAI curriculum development efforts also expand the integration of GenAI metadata creation tasks beyond digital repository metadata courses to include traditional library cataloguing curricula at University of North Texas, starting with the final learning module in the advanced cataloguing and classification course INFO 5220 of the Fall 2025 semester. In that course, students use GenAI tools to generate descriptive metadata that follows a much more robust metadata standard than DCMES 1.1: Machine Readable Cataloging (MARC). Instead of text information resources, this assignment focuses on using GenAI tool in representing two-dimensional and three-dimensional visual resources: maps, postcards, digitised photographs, and video games. Comparing the findings of GenAI-created MARC metadata from this course to the findings reported in this paper will allows for comparative evaluation of GenAI tools performance in representing different types of information resources with different metadata schemes in response to prompts engineered by learners of GenAI among future information professionals.

About the author

Oksana L. Zavalina is a Professor in Department of Information Science, University of North Texas, USA. Her research interests focus on metadata in both library catalogues and digital repositories hosted by libraries, archives, and museums. She can be contacted at Oksana.Zavalina@unt.edu.

References

Alemu, G., & Tammaro, A. M. (2025). Navigating the artificial intelligence frontier on cataloguing and metadata work in libraries: An interview with Getaneh Alemu. Digital Library Perspectives, 41(3), 587-592. https://doi.org/10.1108/DLP-08-2025-208

Bodenhamer, J. (2023). The reliability and usability of ChatGPT for library metadata. https://hdl.handle.net/11244/339626

Brzustowicz, R. (2023). From ChatGPT to CatGPT: The implications of Artificial Intelligence on library cataloging. Information Technology and Libraries, 42(3), 1-22. https://doi.org/10.5860/ital.v42i3.16295

Chen, S., & Li, M. (2024). AI for cataloging and metadata creation: Perspectives and future opportunities from cataloging and metadata professionals. Technical Services Quarterly, 41(4), 317–332. https://doi.org/10.1080/07317131.2024.2394919

Chou, C., & Chu, T. (2022). An analysis of BERT (NLP) for assisted subject indexing for project Gutenberg. Cataloging & Classification Quarterly, 68 (8), 807-835. https://doi.org/10.1080/01639374.2022.2138666

Chow, E. H. C., Kao, T. J., & Li, X. (2024). An Experiment with the Use of ChatGPT for LCSH Subject Assignment on Electronic Theses and Dissertations. Cataloging & Classification Quarterly, 62(5), 574–588. https://doi.org/10.1080/01639374.2024.2394516

Ganadi, A.E., Vigliermo, R.A., Sala, L., Vanzini, M., Ruozzi, F., & Bergamaschi, S. (2023). Bridging Islamic knowledge and AI: inquiring ChatGPT on possible categorisations for an Islamic digital library. Proceedings of the 2nd Workshop on Artificial Intelligence for Cultural Heritage (IAI4CH 2023) co-located with the 22nd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2023), November 6, 2023, Roma, Italy. https://ceurws.org/Vol-3536/03_paper.pdf

Gordon, A. (2025, August 20). CRL Pilots AI to Accelerate Metadata Production. Retrieved from https://www.crl.edu/crl-pilots-ai-accelerate-metadata-production

Hoffman, G., Miksa, S., Snow, K., Kizhakkethil, P., & Zavalin, V. (2024, October). Truth & information? Preparing information organisation & retrieval students for an AI future. In, J. Krammer & M.S. Park (Eds.), The Ethics and Evolution of Truth and Information: Proceedings of the Association for Library and Information Science Education Annual 56Conference: ALISE 2024. https://doi.org/10.21900/j.alise.2024.1723

Kizhakkethil, P., & Perryman, C. (2024). Student perceptions of generative AI in LIS coursework. Proceedings of the Association for Information Science and Technology, 61 (1), 531-536. https://doi.org/10.1002/pra2.1053

LaDell-Thomas, J. (2025). Balancing act: Ethical exploration of AI tools to support student learning in a Library and Information Science program. The IAFOR International Conference on Education – Hawaii 2025 Official Conference Proceedings (pp. 667-676) https://doi.org/10.22492/issn.2189-1036.2025 .

Martorana, M., Kuhn, T., Stork, L., & Ossenbruggen, J.V. (2024). Zero-Shot Topic Classification of Column Headers: Leveraging LLMs for Metadata Enrichment. International Conference on Semantic Systems. arXiv-Computer Science-Databases. https://doi.org/10.48550/arXiv.2403.00884

Moulaison-Sandy, H., & Coble, Z. (2024). Leveraging AI in cataloging: What works, and why? Technical Services Quarterly, 41(4), 375–383. https://doi.org/10.1080/07317131.2024.2394912

Sussmeier, S., & Henry, J. A. (2025). Mind the gap! How do we ethically bridge the divide between the cataloging/metadata community and the world of AI? Journal of Library Metadata, 25(4), 245-259. https://doi.org/10.1080/19386389.2025.2525720

Torabi, M., Emdad, A., & Pashootanizadeh, M. (2025). From captions to catalogues: A systematic review of GenAI models for image captioning and their role in image cataloguing. Cataloging & Classification Quarterly, 63 (6/7), 436-474. https://doi.org/10.1080/01639374.2025.2539792

Zavalin, V.I., & Zavalina, O. L. (2025). Are we there yet? Evaluation of AI-generated metadata for online information resources. In I. Sserwanga, M.R. Sanfilippo, C. Inskip et al. (Eds.), Living in an AI-gorithmic World: 20th International Conference, iConference 2025, Proceedings (pp.732-740). https://doi.org/10.47989/ir30iConf47215

Zhang, S., Wu, M., & Zhang, X. (2023). Utilising a large language model to annotate subject metadata: A case study in an Australian national research data catalogue. arXiv-CS-Computation and Language. https://doi.org/10.48550/arXiv.2310.11318