2025-26 HIVE Summer Internship Project

Telling the story of the WA library community through visualisations of the WAIN Email Archive.

Primary Academic Supervisor
A/Prof. Gaby Haddow

Project Background
The Western Australian Information Network (WAIN) is a mailing list utilised extensively by Western Australian library and information professionals. The list commenced in 1994 and has over 2,000 subscribers. It was established to allow dissemination of information about events, advertise vacant positions in the sector, and discuss issues of interest within the profession in Western Australia.

The WAIN Email Archive is a digital archive held by the Curtin Library containing over 23,000 email posts received from WAIN between 1994 and 2024.

This project will tell a story of the library community through visualisations. Drawing on the metadata and content of the WAIN Email Archive, the project will explore trends and changes relating to events, employment and topics of discussion over the period of the list. This story may include exploring the top contributors to WAIN over those years, the use of WAIN to advertise positions and promote events, and how these have changed over time, and the engagement of the library community across metropolitan and regional Western Australia through the list.

The project will investigate the suitability of different software, such as Large Language Model (LLM) and Retrieval-Augmented Generation (RAG) to process the Email Archive's data, and software and tools for creating the visualisations on the HIVE displays. The student will develop expertise in using the software and producing visualisations using the exported data.

Project Description, Expected Outputs, Possible Stretch Goals
The primary output will be visualisations of the Western Australian library community's interactions and points of interest over the period 1994-2024.

There has been no published or publicly available record of the WA library community's activities since 1990 and this project will fill that gap. If feasible, visualisations created by the project will be made available on the Curtin Library's website.

Importantly, the project will also contribute to a wider discussion about how email archives can be interrogated and visualised, the technological challenges encountered and ethical considerations.

Links to background reading and any relevant recent work in the field
Baker, F. (2014). Carcanet Press Email Preservation Project—Phases 2-3. Manchester, UK: The University of Manchester ; 2014.; The University of Manchester. https://www.escholar.manchester.ac.uk/uk-ac-man-scw:226625

Bartliff, Z., Kim, Y., & Hopfgartner, F. (2022). A survey on email visualisation research to address the conflict between privacy and access. Archival Science, 22(3), 345–366. https://doi.org/10.1007/s10502-022-09387-2

Email Archiving Tool. (n.d.). Retrieved 15 September 2023, from https://www.epaddproject.org/

Green, P. (2025). AI and the visualisation needs of researchers using email archives. AI & SOCIETY. https://doi.org/10.1007/s00146-025-02187-z

McKean, C., & Randall, C. (2025). Data analysis and network visualisation as tools for curating hybrid correspondence archives. AI & SOCIETY. https://doi.org/10.1007/s00146-025-02200-5

Rochester, M. K. (1990). Australian Library History Research. Libraries & Culture, 25(1), 115–129. https://www.jstor.org/stable/25542232

Schneider, J., Adams, C., DeBauche, S., Echols, R., McKean, C., Moran, J., & Waugh, D. (2019). Appraising, processing, and providing access to email in

contemporary literary archives. Archives and Manuscripts, 47(3), 305–326. https://doi.org/10.1080/01576895.2019.1622138

Sharman, R.C. & Clyde, L.A. (Eds.)(1990). Western perspectives: Library and information services in Western Australia. Australian Library and Information Association (WA Branch).

Smith, J., & Carlyle, P. (2022). Guide-for-Management-of-and-Access-to-Email-Archives.pdf. https://rylandscollections.com/wp-content/uploads/2022/02/Guide-for-Management-of-and-Access-to-Email-Archives.docx

Stanford University's Special Collections & University Archives. (2022). EPADD Discovery Module Collections. https://epadd.stanford.edu/epadd/collections

What type of visualisation will the student develop or produce?
The data contained in the email archive combine time, people, location, affiliation, and content. Visualisations which can draw on these elements will include 2 or 3 dimensional visualisations of a size to take advantage of the large, tiled or cylinder displays in the HIVE.

How will the visualisation contribute to your research outcomes?
Visualisation will be the central part of the narrative, enabling a view of the activity reflected in the WAIN Email Archive that would not otherwise be possible. Applying visualisations to an email archive is an emerging area and this project will be at the forefront of that work. The research outcomes will contribute to greater understanding of the benefits and challenges involved in working with email collections, including identifying effective forms of visualisations.

If the project is successful, where would you consider publishing the results?
A journal article would be produced to disseminate the project's work and outcomes.

With its focus on a community within the Australian library sector, the most suitable journal is: The Journal of the Australian Library and Information Association (a peer reviewed, quarterly publication for information science researchers, information professionals, related disciplines and industries).

There is also potential to publish an open access report about approaches taken to the data processing, and publishing the processed data for particular visualisations subject to privacy and copyright considerations being met.

Draft Project Timeline:

Week 1

Nov 10 - Full day HIVE induction
Nov 11 - Area and Project Induction with Primary supervisor
Develop project plan with HIVE and academic team
Literature and relevant project review

Meet with Library staff and settle into workspace.

Week 2

Develop understanding and expertise in using the software, and familiarity with the content of the WAIN Email Archive.

Week 3

Workshop the narrative and the data that would be visualised. Produce wireframe of the story. Document the way data from the WAIN Email Archive would need to be categorised and cleaned to enable visualisation. Create Data Management Plan and request R drive allocation.

Week 4

Prepare the exported data for use by the LLM and RAG. Export and clean the data.

Week 5

Use the LLM and RAG to explore the data and begin exploring visualisations.

Week 6

Use the LLM and RAG to explore the data and begin exploring visualisations.

Week 7

Refine the visualisations and develop the narrative.

Week 8

Complete the narrative and visualisations.

Week 9

Focus on report writing and presentation preparation

Supervisors to provide feedback on report and presentation.

Week 10

Focus on report writing and presentation preparation
30th Jan Final Presentation Showcase Day and final report due

Supervisors to provide feedback on report and presentation.

Student Experience and Supervision:

How often will you meet with the student over the 10-week period?

Weekly

Your work desk location and the location of student desk

B105.437

Student Attributes:

Please indicate any preference for student's academic discipline or field of study

The student will be engaged in a project that uses an archival data set to reveal the story of a community, working with different stakeholders and working to a deadline. They will be mentored throughout the project to develop both technical and research skills, as well as providing experience in the creation of scholarly written work and effective presentation skills. Some technical skills relating to the management of large data sets would be an advantage.

What competencies are required to start this project

Beginner - 3D modelling software (e.g. Blender, 3ds Max)
Beginner - Unity 2D/3D Artistry (assets, lighting, cameras, materials implementation)
Beginner - Unity Programming (C# coding, animation syntax, debugging, problem-solving)
Beginner - Unity Virtual Reality Development (rendering pipelines, scene content design, interaction)
Intermediate - Data structures, analytics, statistical modelling

Do you have any other student attributes you think are important to the project?