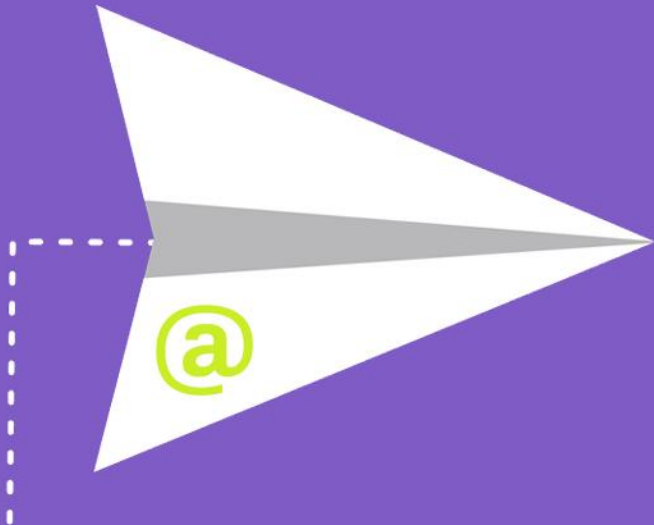


Public Record Office Victoria

Stage 3 Email Appraisal, Disposal and Preservation Project 2022 – 2024

Appraisal Approaches and Automated Processing



Project background and purpose

Problem Statement:

1. Online Emails: Government continues to use email in the Microsoft suite to conduct business, storing them in an online environment. Due to its semi-structured data, email is difficult to capture, store, preserve, make accessible and dispose of email records not required.
2. Backlog of Lotus Notes email archived on Linear Tape-Open (LTO) format
 1. 1993-2007 (Pre Cenitex)
 2. 2007-2023

Purpose Pilot solutions using automated techniques for online emails within a macro appraisal framework.

Outcome PROV has capacity to retain a core of de-duplicated official records from agencies and employees to which data visualisation and analysis tools may be applied by users in future.

Key findings from Stage 2

NSF



Lotus Notes NSF format is a very unsuitable format for access and preservation



Email appraisal should consider organisational recordkeeping practices



The cost of processing and analysing emails is expensive



Threading contextualises emails, providing a more understandable, complete and accessible record



IT regimes disturb the original order of organisational email



Access to sophisticated (AI/ML) tech is needed to apply functional appraisal



Direction for Stage 3

Moving forward, we have decided to:

- Continue to target online emails
- Reduce the scale of the problem through macro appraisal
- Develop automated, repeatable and efficient methods to retain the official core of agency records
- Prioritise de-duplication and removal of non-official and low value ephemeral records through automated processing
- Design a workflow for capture and processing that can be managed by agencies and be integrated into a scheduled transfer program and scaled up over time

Macro Appraisal



With over 300,000 Victorian Government employees and email remaining a key tool for communication, the volume of emails created and captured increases everyday.



Without an effective approach to appraisal and disposal, key evidence of government decisions could be lost within the unstructured data presented by email.



Taking examples from and discussing with other countries such as the US and UK, PROV has been exploring the concept of Capstone as developed by National Archives and Records Association (US).



To reduce the scale of the problem, the Capstone approach allows an agency to categorise and appraise email based on the work and/or position of the email account holder.



By prioritising accounts held by key decision makers in government positions, the volume of emails that need to be appraised reduces into a more manageable set of data.

Open-Source Tools for Appraisal and Disposal



Stage 2 highlighted that only a small percentage of emails (5-10% of total created) are classified as permanent value records. The process was not cost effective and was labour intensive.



To reduce the scale of the problem, PROV have shifted their focus to reducing the volume of emails by identifying non-official and low value ephemeral emails.



Research into the open-source market of tools for email appraisal and disposal has revealed an appetite across the world for tools that automate the process.



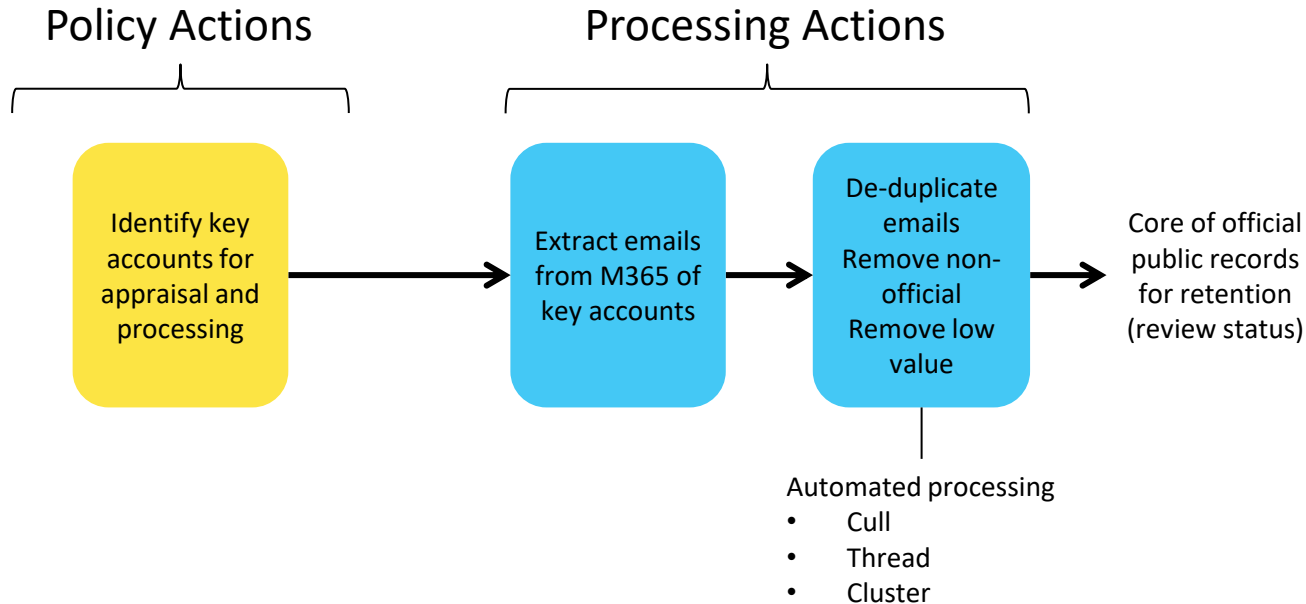
Open-source tools such as ePADD, RATOM and DArcMail will be explored by PROV to find automated, repeatable and efficient methods of deduplicating and identifying low value records.



Select tools will be tested to investigate the extent of their appraisal functionality and identifying their limitations. In doing so identifying future avenues of email preservation as technology advances.

Proposed Disposal Process

A process that allows for a broader, less ambitious solution aimed at capturing and retaining core official emails of value across Victorian Government:



Proposal for reducing email volume

The final step of the proposed disposal process requires automated processing to de-duplicate and:



Graphics for this presentation were sourced from Freepik.com

Why do this work and take this approach



A high risk of loss of email as an information resource for policy development, service delivery, and as a record of government decision making



A high risk that the backlog of email will grow



Scale necessitates the application of automated methods within a framework designed by macro appraisal so that only the highest priority are actively managed



Stage 3 Outcomes



A process to dispose of online emails in categories 1 and 2 by public offices researched, tested and trialed.



A process to capture and transfer to PROV online emails in category 3 researched, tested and trialed.



Advice on software tools and advice on email capture and disposal published by PROV for public offices.