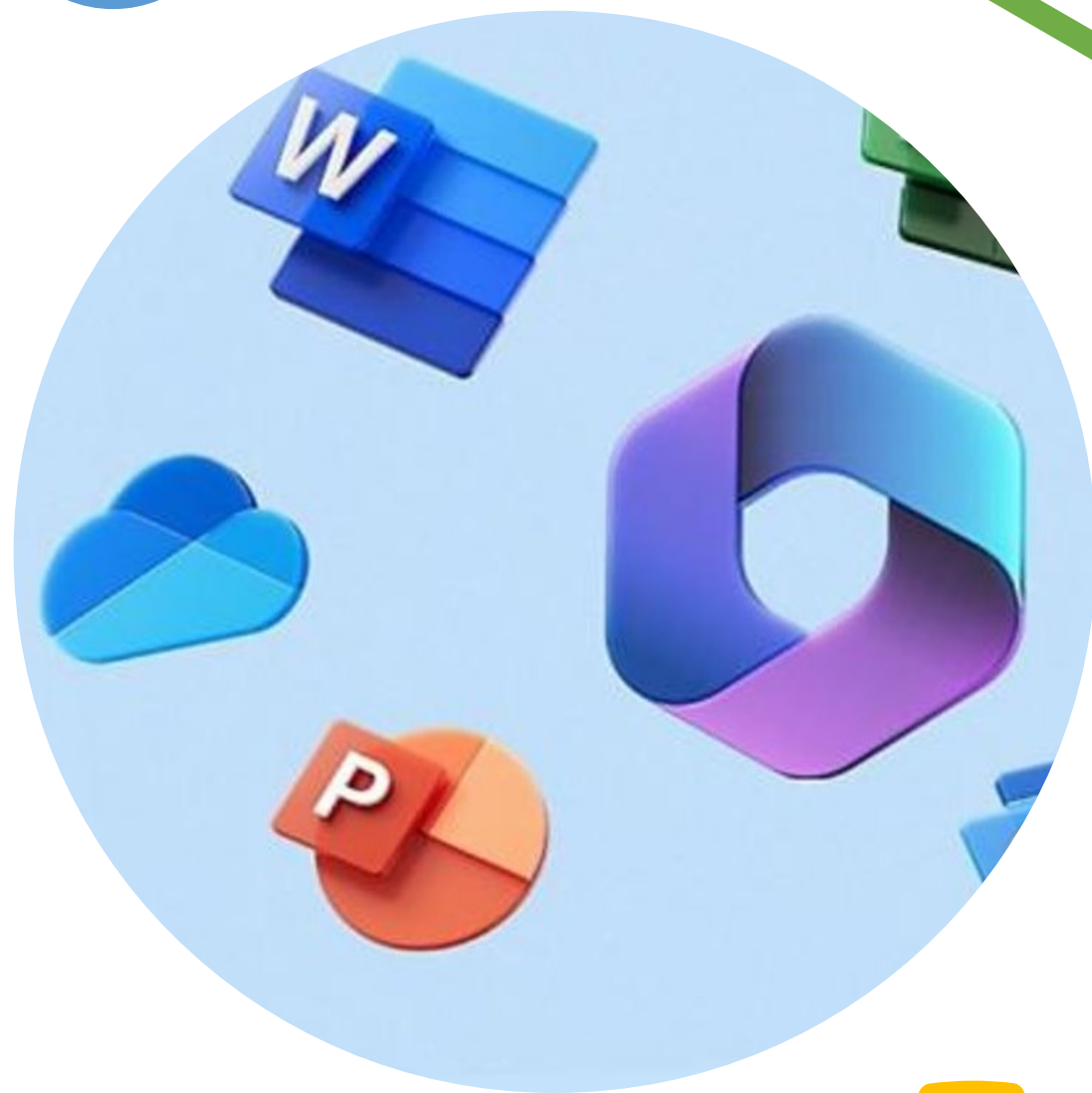




Your Partners in (Cloud) Preservation:

An Information Management Perspective on M365

Angela Beking

Privy Council Office

Government of Canada

Overview

Material to **preserve**

- “Static” files that are no longer static
- New apps and services create fluid content
- How to identify (and/or create?) the “record”/“preservation object”

Material to **delete**

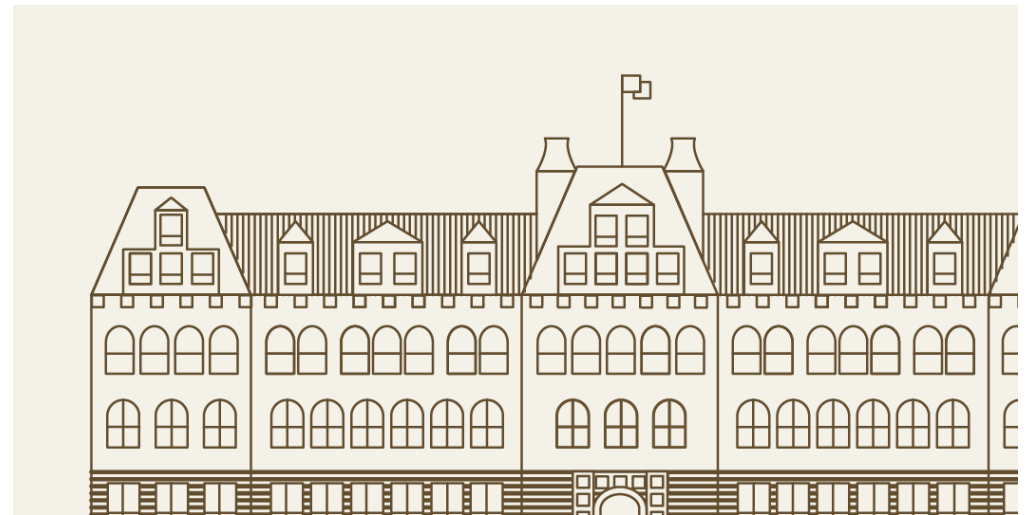
- How to identify and address the proliferation of information “noise” in cloud-based systems

Partnerships to **create**

- How information management and digital preservation professionals might tackle these challenges together

Policy context

- The Privy Council Office (PCO) supports the Prime Minister and Cabinet
- PCO helps the government in implementing its vision, goals, and decisions in a timely manner
- M365 is used as a collaborative workspace
- Per the *Library and Archives of Canada Act*, government records deemed to be archival must be transferred to Library and Archives Canada at the end of their retention periods, regardless of medium or form

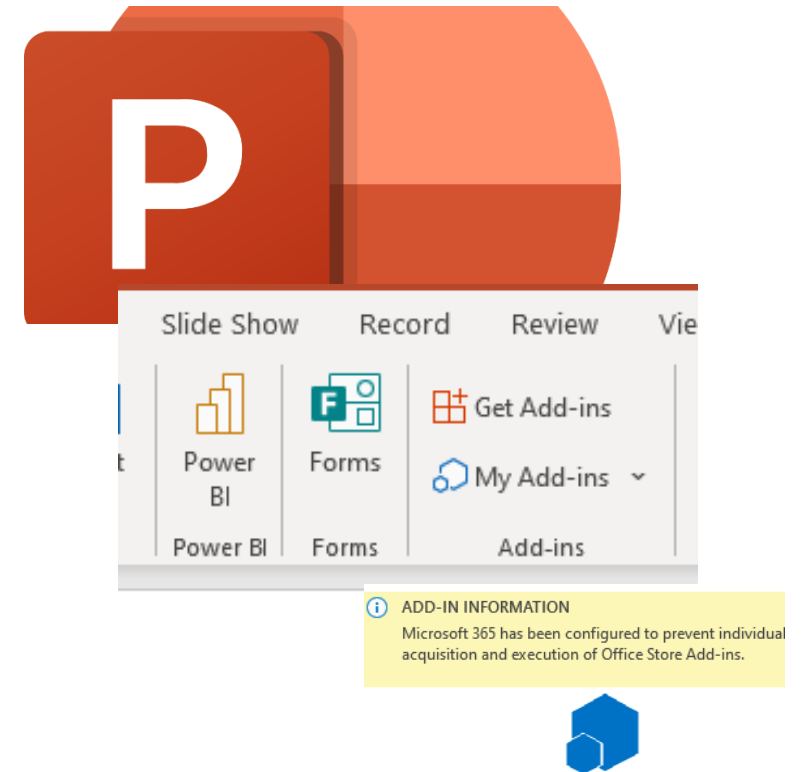


Material to preserve: (not so) static files

M365 files (*.docx, *.pptx, *.xlsx) are not always static

Integrated apps and services create **interdependencies** between structured data, files, and other types of content

- E.g., a *.pptx file has an integrated MS Form, which is stored in a separate data center
- Challenges emerging from this can sometimes be avoided with **system configuration** - but this limits functionality



Material to preserve: (not so) static files

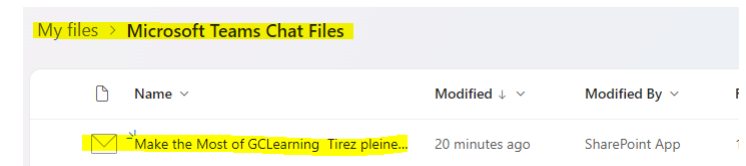
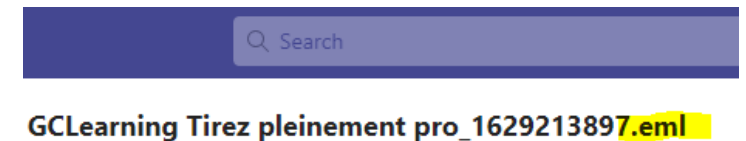
Integrated apps and services can also create **duplicate content**

- E.g., Exchange Online “Share to Teams” function creates a copy of my Outlook *.eml file in my OneDrive and/or my team’s SharePoint

At scale, it is difficult to:

- 1) Detect this duplication across apps and services
- 2) Determine potential value

The million-dollar question: How do we do this at scale, and **only** for the information of value?



Material to preserve: (not so) static files

Proliferation of versions

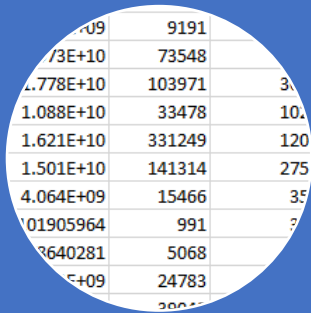
- If autosave is turned on, SharePoint will automatically save a version of the document **every few minutes**
 - At scale, how do we identify important and non-important versions?

Version history

[Delete All Versions](#)

No. ↓	Modified
374.0	10/24/2023 3:17 PM
373.0	10/23/2023 2:55 PM
372.0	10/23/2023 2:43 PM
371.0	10/23/2023 2:27 PM
370.0	10/23/2023 2:20 PM
369.0	10/23/2023 2:19 PM
368.0	10/23/2023 1:45 PM
367.0	10/20/2023 4:34 PM
366.0	10/20/2023 4:24 PM
365.0	10/18/2023 12:10 PM
364.0	10/17/2023 4:49 PM
363.0	10/17/2023 4:26 PM

Material to preserve: “non-static” content

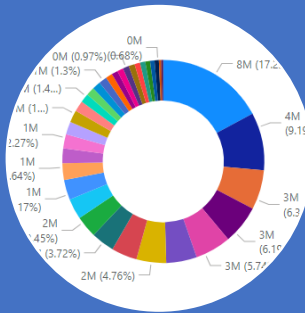


Data source

Connected online services **or** static file(s) (.csv, .xlsx)

Connected (no “file” exists), or saved locally, in OneDrive, or in SharePoint

Dependent on the stability of the connection



Power BI Desktop

.pbix file

Saved locally, in OneDrive, or in SharePoint

Dependent on the data source



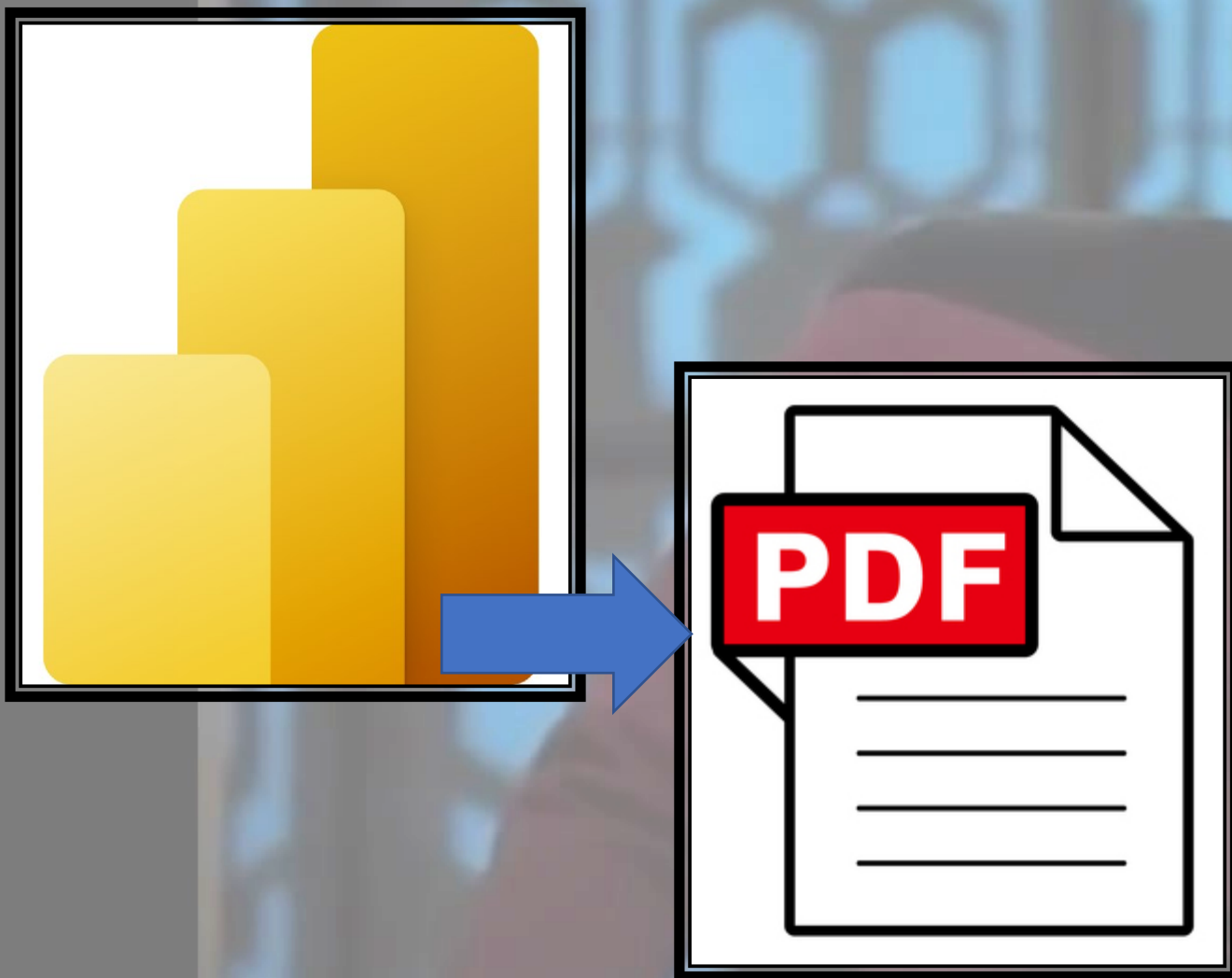
Power BI Service

Creates a copy of local files; syncs to OneDrive/SharePoint

Depends on data source, Power BI Desktop file



Material to preserve: Make it a PDF!



Material to delete: the “noise”

Information created in the back end that has no value

- System configuration may provide solutions

From: Beking, Angela <Angela.Beking@pco-bcp.gc.ca>

Sent: Thursday, September 14, 2023 3:06:45 PM

To: CSB Information and Data Management Policy (I&DM Policy) | SM Politique de gestion de l'information et des données (Politique GI&D) (PCO/BCP) <IDMPolicy-PolitiqueGID@pco-bcp.gc.ca>

Subject: Comments on task "test"

do we have an email site

Reply in Microsoft Planner

You can also reply to this email to add a task comment.

This task is in the [CSB Information and Data Management Policy \(I&DM Policy\) | SM Politique de gestion de l'information et des données \(Politique GI&D\) \(PCO/BCP\)](#) plan.



How cloud challenges us to think differently

- The M365 ecosystem is designed for real-time collaboration and reporting
 - It is **not** designed for digital preservation
- Unanticipated content is created that may have no value
- Need to think proactively and conscientiously of the impact of this on our core assumptions
 - Moving towards digital environments where there will be no more “naturally occurring” digital preservation objects

Partnerships to create

Information Management and Digital Preservation need each other to answer the **big questions**:

- What is a “record”?
 - Is this still a useful term? Do we need new terminology, or an entirely new framework of thinking?
- What is a “preservation object”?
 - How do we identify what the preservation object **is** (for objects that “exist”), or what it **could be** (for objects that must be actively created)?
 - Preferred and accepted file formats are a great start - but more broadly, what are we trying to document? How does that reconcile with technology opportunities and limitations?





Thank you

Questions?

