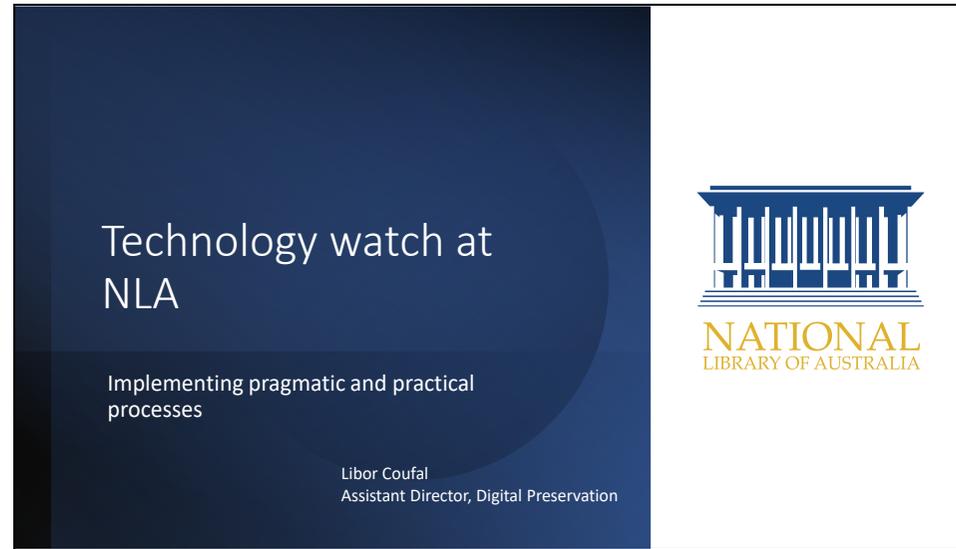


Slide 1



Thank you for the introduction and the opportunity to speak today! I am going to talk about the NLA's take on Technology Watch.

Slide 2



To give you some context, we are a small(ish) National Library, compared to the heavy-weight likes of the BL or LoC. Our Digital Preservation team has only three people, although we do have other staff contributing directly or indirectly towards digital preservation. We started to build our digital collections in the 1980s, and to date we have accumulated over 2.5PB or almost 12 bn files.

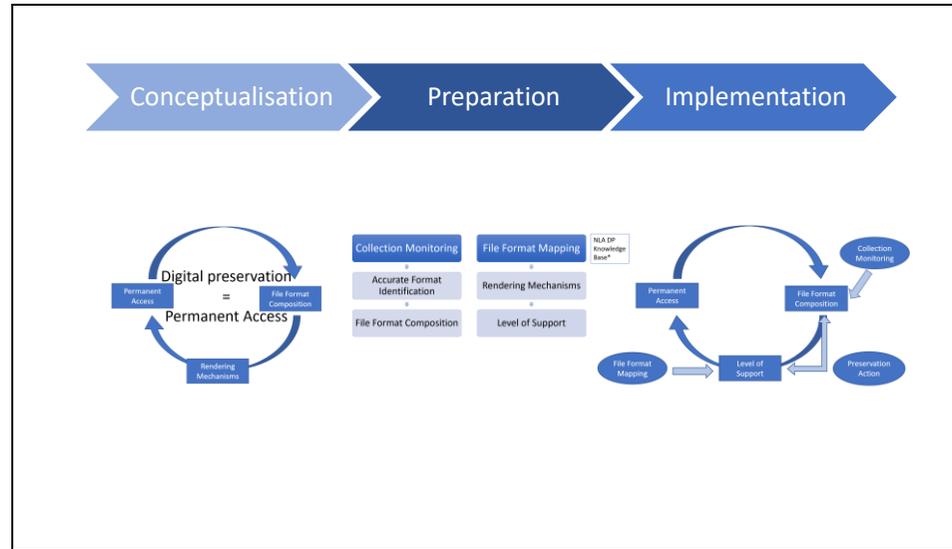
## Slide 3

- 
- 2003 - Automatic Obsolescence Notification System (AONS)
  - 2006- AONS II  
Pearson D, Webb C. Defining File Format Obsolescence: A Risky Journey. *International Journal of Digital Curation*. 2008; 3(1):89-106.
  - 2011 – File format registry
  - 2013 – NLA Digital Preservation Knowledge Base  
Kay G, Coufal L, Pearson M. Backing up digital preservation practice with empirical research: The National Library of Australia's Digital Preservation Knowledge Base. *Alexandria*. 2017;27(2):66-82.
  - 2014-15 – File format risk assessments

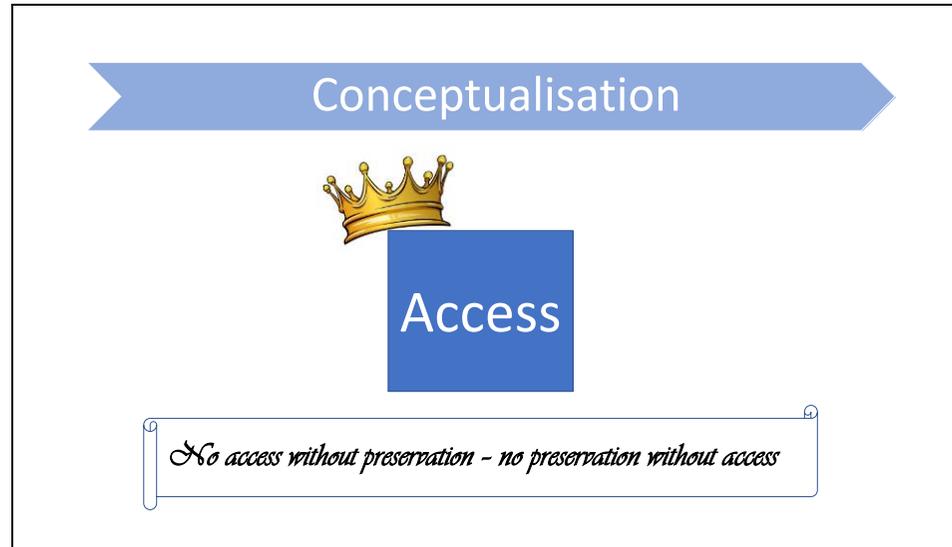
We have been in the digital preservation business since the 1990s and from very early on our focus has been on preservation planning and technology watch. This was most notably manifested in our initiating the Australian Partnership for Sustainable Repositories project in 2003, and its sequel in 2006. The partners set off to develop an automated system to monitor file format obsolescence (known as AONS), a goal that turned out to be overly ambitious and tricky to implement, to say the least. However, our thinking on preservation planning, as captured in the cited paper by Pearson and Webb, has stood the test of time and has been polished to its current shape.

After AONS we flirted with file format registries and risk assessments, which we found to be not very helpful and too onerous, so we pulled the plug on them. A project which did survive to date, and indeed turned into a cornerstone of our current technology watch processes, is the NLA Digital Preservation Knowledge Base. I will talk about how it fits into the processes but will not have enough time to go into much detail so I will refer you to the paper by Kay, Coufal and Pearson, for more information.

Slide 4

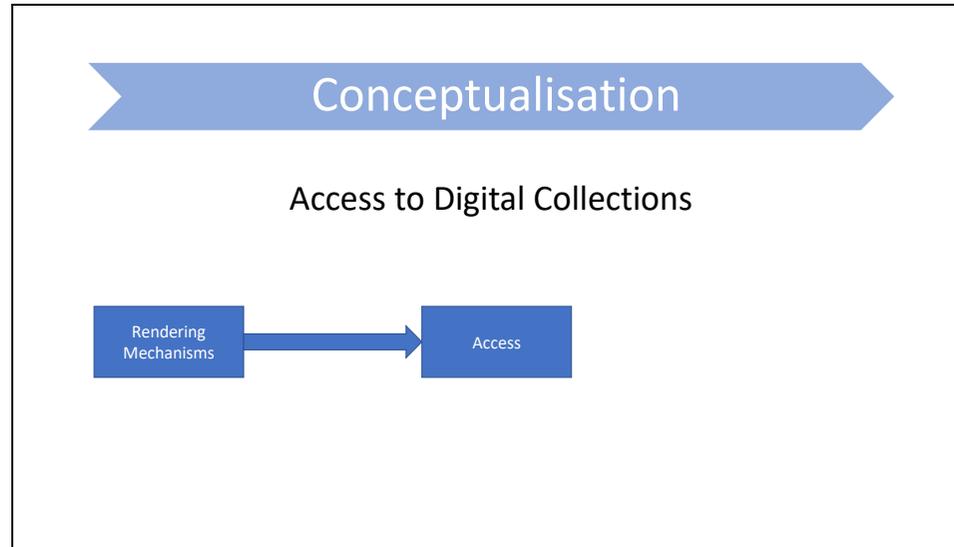


Today, I will talk about how we conceptualised technology watch, what critical building blocks we put in place, and how we are bringing it all together to implement practical and pragmatic preservation watch processes that are sustainable and scalable.



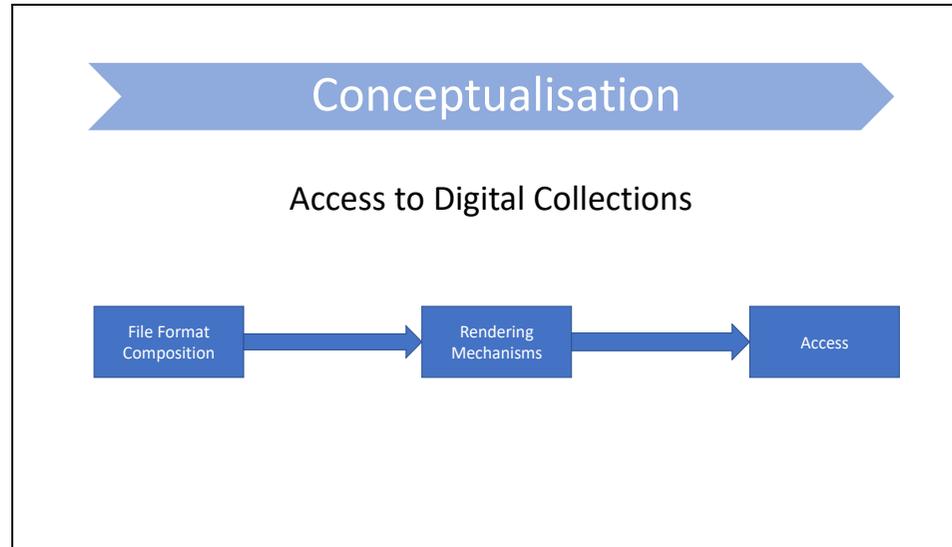
It meant going back to the drawing board and revisiting why we do digital preservation, which things matter and make difference and therefore must become our focus, and which do not and must be ignored.

With that in mind, let's start – at the end. What is our ultimate goal? We do not preserve, or indeed, collect for the sake of it. Everything that we do is governed by access!

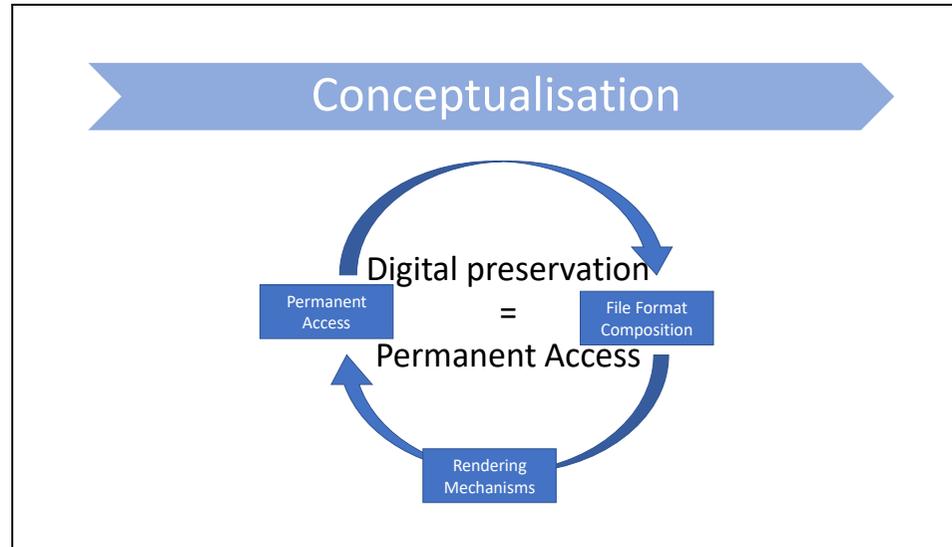


To provide access to our digital collections, we must have ways to render the digital objects in them.

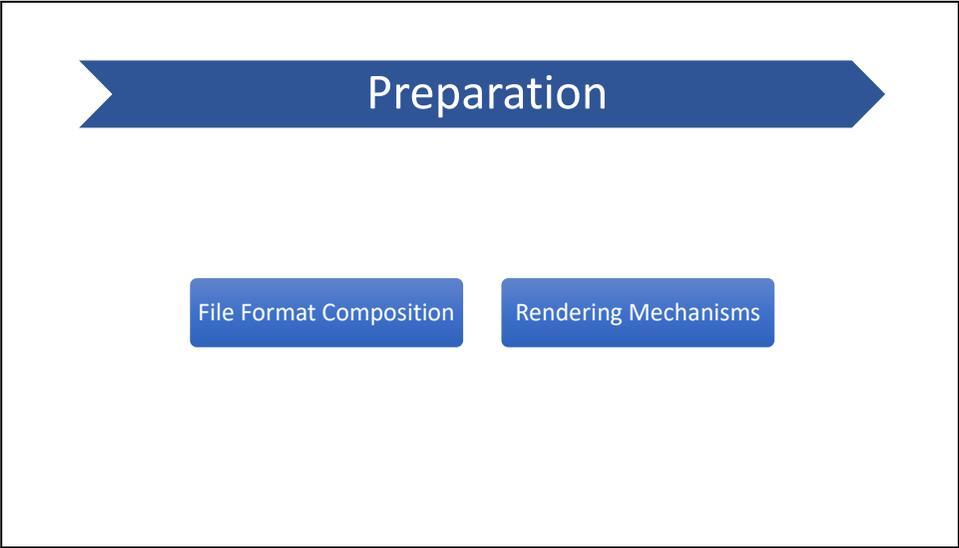
Slide 7



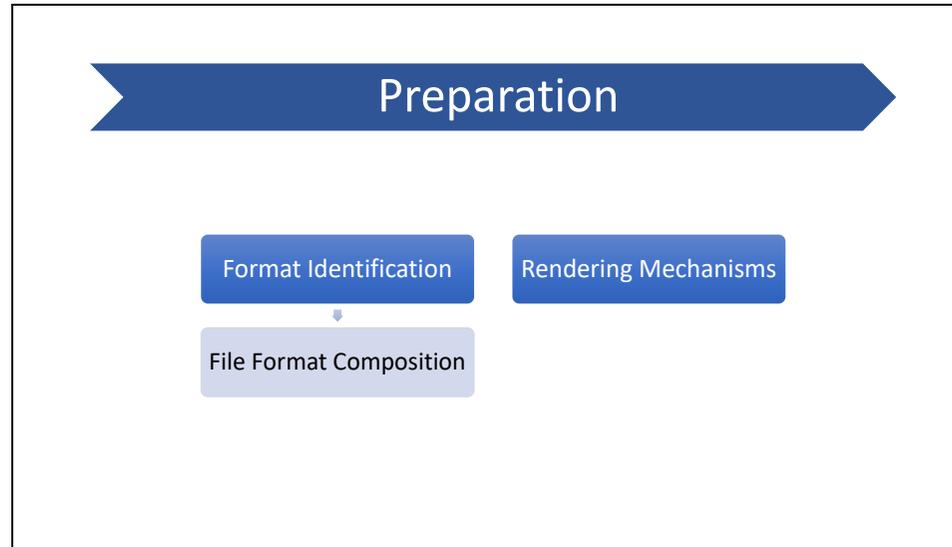
That means having SW which can render the file formats in the collections, and that in turns means we need to know what formats there are. Now, this isn't a once off or static process.



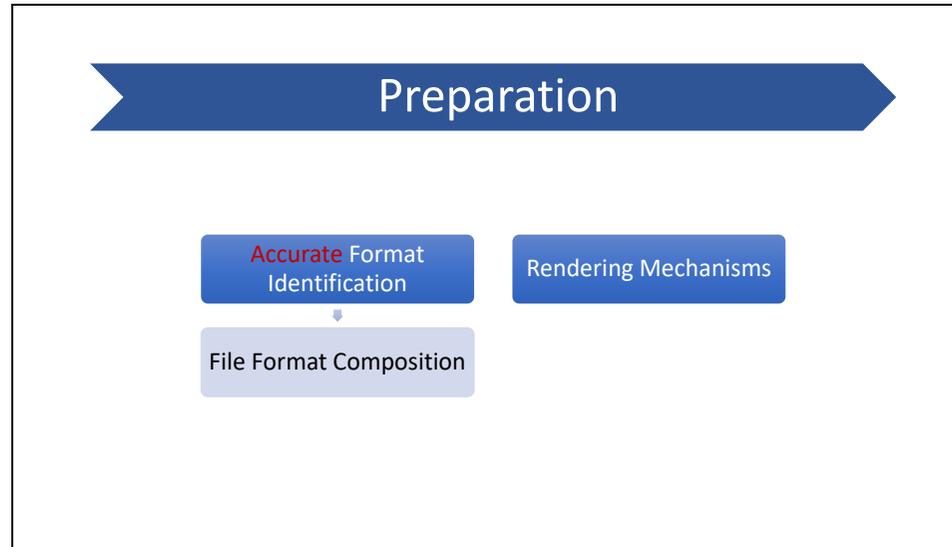
Both collection composition and rendering mechanisms will change over time, and so it is an ongoing, cyclical and dynamic process. If we can ensure access at any given time in the future, as file format composition and rendering mechanisms change, we will be on track to achieve permanent access – which is synonymous with digital preservation.



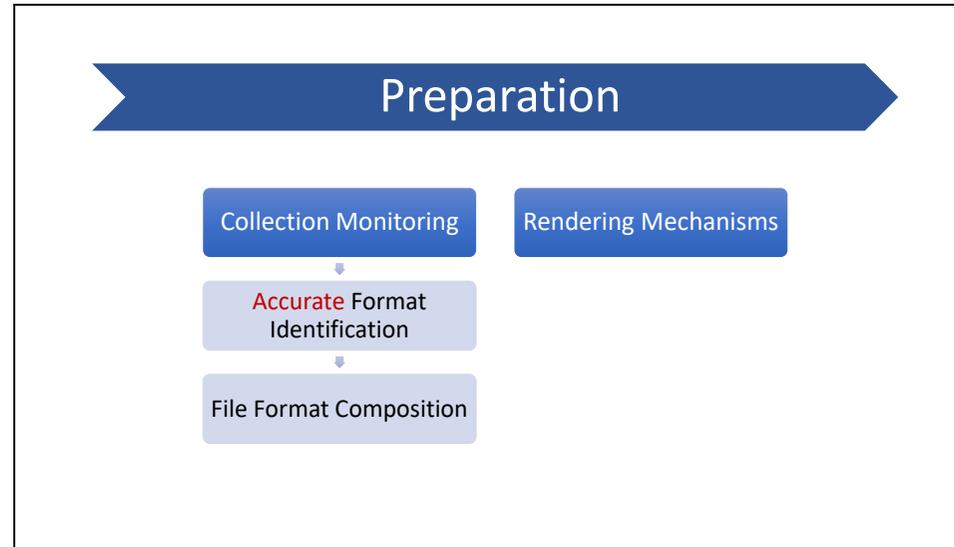
To implement the processes, we had to first put in place two important missing building blocks



The first helps us to report on file format composition. That means identifying file formats in the collections. Unfortunately, format identification isn't an exact science – in fact, it is very flawed: there are formats that the common identification tools will not identify, and formats which they will misidentify.

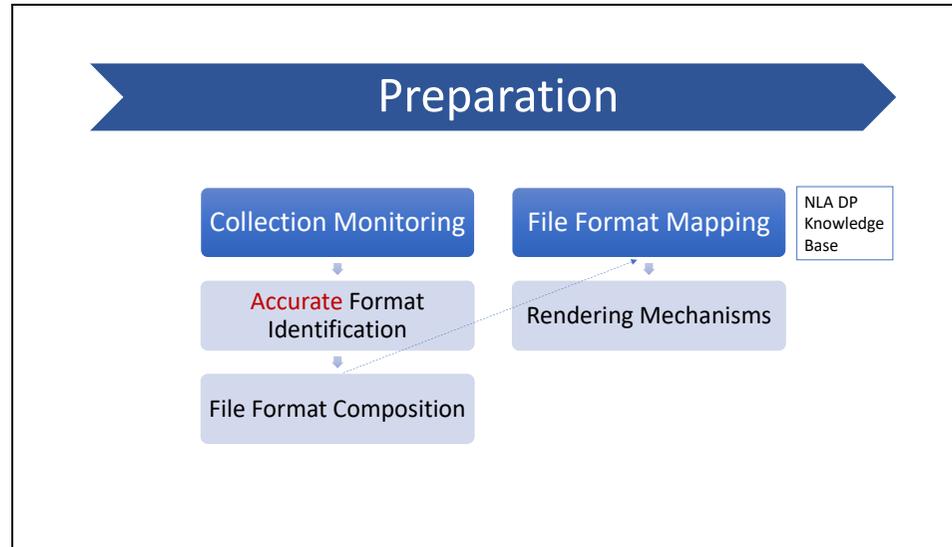


That is obviously a showstopper – if we think we have format A, but in fact have a completely different and unrelated format Z, then everything else which we base on this inaccurate assumption falls like a house of cards!

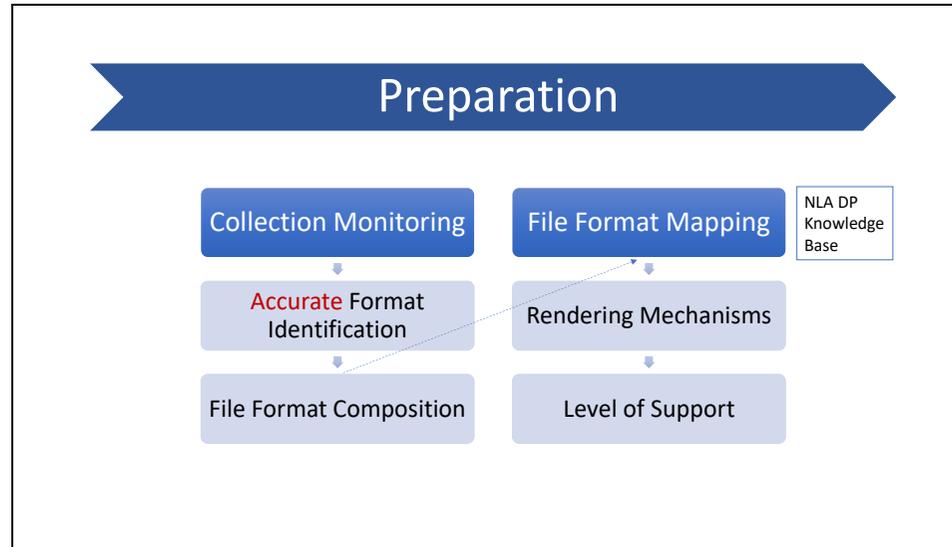


We implemented regular processes, that we call Collection Monitoring, where we weekly check and verify the identification results of all new content added to the repository the previous week. We add any missing identification and correct anything that is wrong. This allows us to report on file format composition with complete confidence.

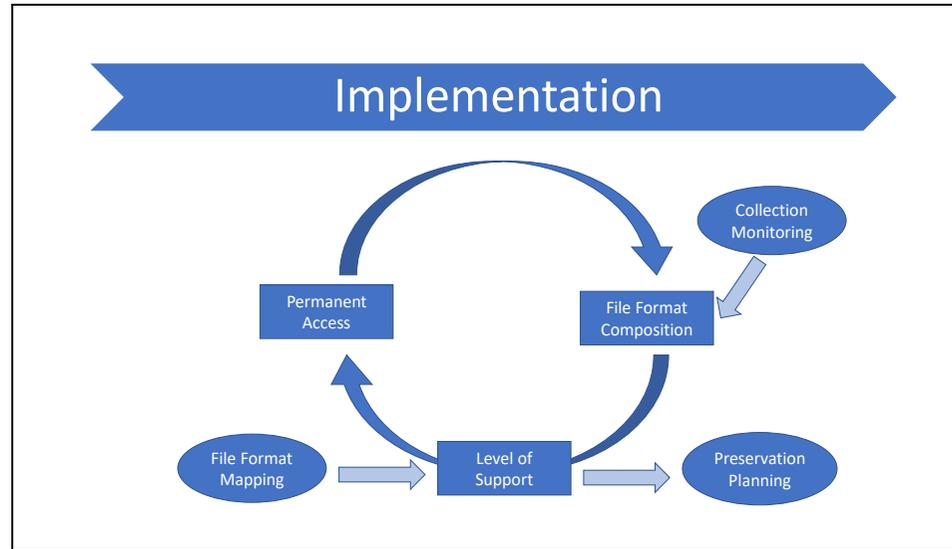
The second process we implemented is to do with rendering mechanisms.



It is a regular ongoing research, that we call File Format Mapping, where we map the formats we identified as present in our collections to SW applications which can render them, and confirm it, as well as the rendering quality, through testing. This process is part of building our DP Knowledge Base.

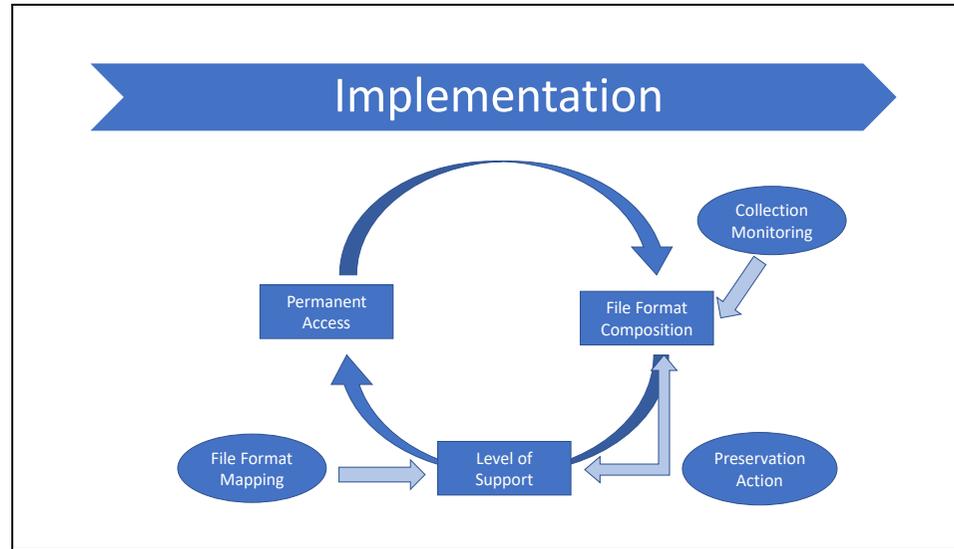


We then use this data to establish how well various file formats are supported in SW applications.

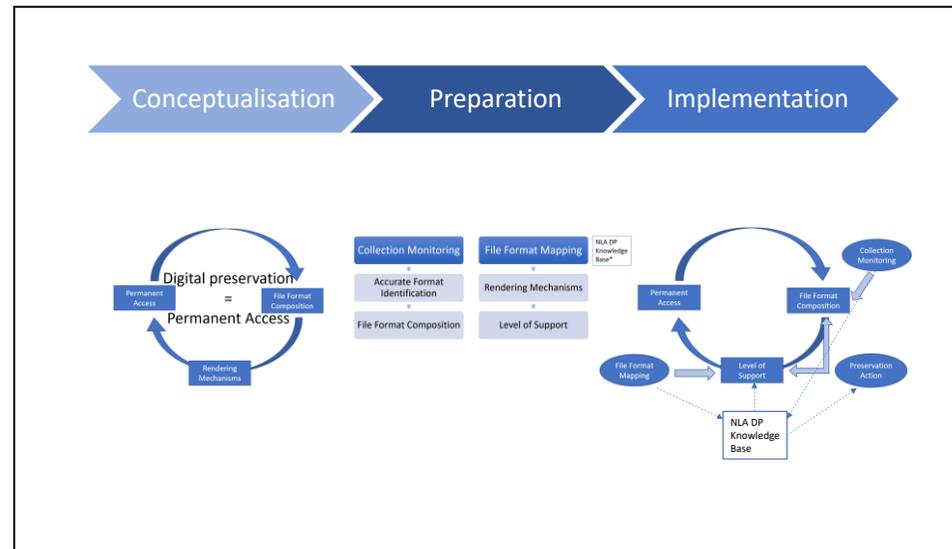


So, the weekly Collection Monitoring helps us report on the file format composition of our collections. We then add the File Format Mapping research to establish the Level of Support for each of the formats in our collections.

When we don't have support, or don't have sufficient support, or it is dropping, it will trigger preservation planning.



Which will eventually lead to a preservation action. The preservation action will either boost level of support (e.g. by adding a new renderer, or through emulation) or change the file format composition (through migration). That will renew access, the cycle will complete and the process will start anew.



We now have most of the Technology Watch elements in place. We are currently working on implementing the last missing bit – reporting on Level of Support. However, it wouldn't have been possible without the upfront investment in Collection Monitoring and building the DP Knowledge Base.

Once Technology Watch is implemented, it will allow us to move onto Preservation Planning. We are not aiming to create Preservation Plans in advance for all our formats – we will only create them as, and when, needed.

Slide 18



Thank you!

Libor Coufal  
lcoufal@nla.gov.au