Reliable Robust and Resilient Digital Infrastructure for Nuclear Decommissioning – Phase Two (November 2020 – February 2023)

Project Completion Report: February 2023

Table of Contents

1. Introduction and background

This report describes the work and deliverables of the second phase of the joint collaborative research project undertaken by the Digital Preservation Coalition and the Nuclear Decommissioning Authority. Phase One of the project completed in November 2020, when it was agreed to extend the project for a further two years. The broad aims of the project remained the same, namely to put in place the foundations for retaining and preserving data and systems critical to the long-term business needs of the NDA and develop detailed policy to steer NDA staff and contractors working on behalf of NDA to ensure that:

- legacy data and systems are managed to optimize accessibility and integrity in the long term;
- current data and systems will be accessible and dependable, now and in the future;
- future data and systems are commissioned with long term resilience from the outset.

Phase One of the project provided a good understanding of the digital preservation challenges facing the NDA, and supported the development of a new maturity model for digital preservation (DPC RAM) which was used to benchmark the NDA against the wider digital preservation community (see the Phase One project completion report produced in November 2020 for further details).

It was agreed that Phase Two of the project would build on the previous work, and proceed with three broad workstreams:

- **Innovation**: to create practical solutions and guidance around the NDA's challenges based on innovation and research, adding value therefore to the wider digital preservation community;
- **Capability**: to build capability within the NDA by interrogating state of the art research and 'best of breed' solutions from the digital preservation community and interpreting them to the NDA context, adding value to both as a result.
- **Management**: to provide accountability and quality assurance appropriate to the responsive, iterative project design.

This report summarizes the activities undertaken in Phase Two, including any activities that were carried over or extended from the previous Phase. Due to delays to negotiating the start of Phase Two and thus recruitment of staff, activities formally continued until early February 2023. However, some final deliverables (including post-project dissemination events) will take place in the months immediately following the formal conclusion of the project. A summary list of all the outputs of both Phases of the project is included in Appendix 1 of this report.

2. Project approach

For the duration of Phase Two of the project, it was agreed that the DPC would employ an additional member of staff to research and advise on good practice in digital preservation. In February 2021, Michael Popham was appointed as a Digital Preservation Analyst with the DPC, reporting directly to the DPC's Head of Good Practice and Standards (Jen Mitcham). Working 0.75FTE on the project, the Analyst's role was to support the NDA to achieve its documented goals for innovation and capability in digital preservation, and to document and share any lessons learned with the wider digital preservation community.

During 2021 there were also changes in NDA personnel attached to the project. Martin Robb, who had coordinated much of the NDA's involvement during Phase One retired and his role as project lead was taken over by Michelle Donoghue, who joined as Information Governance Manager for the NDA.

The project Steering Group continued to meet approximately every quarter to provide oversight and general direction to the project team and more regular 'keep in touch' meetings were held on a fortnightly basis to share information and discuss progress.

The project was not only focused on the project deliverables as described in the Phase Two project proposal but was also committed to providing ad hoc digital preservation advice and support to the NDA as required. In order to maintain lines of communication and knowledge sharing between the DPC project team and the wider NDA Group, DPC staff were able to regularly attend and contribute to key NDA meetings such as the IGO Forum and Nucleus Communications Group.

3. Summary of work carried out in Phase Two

The table below summarises the work carried out during Phase Two of the project (i.e. covering the period November 2020 to February 2023). Each of three main workstreams of Phase Two are listed, along with the particular aims and deliverables associated with that activity.

Workstream	Aim and summary of work carried-out
Innovation	**Aim: To create practical solutions and guidance for the NDA based on innovation and research, therefore adding value to the wider digital preservation community.** • **Preserving geospatial data** During Phase One of the project, this had been identified as a topic of particular interest to the NDA. Several meetings were held with key individuals from across the NDA, and an online briefing day was organized in March 2022. The DPC also commissioned a revision of an existing Technology Watch Report on Preserving Geospatial Data, which is due for release later this year. • **Database preservation** (with additional funding from Sellafield) The challenge of preserving databases had also been identified in Phase One as something which the NDA would like to address and efforts to commission a practical project to look at this issue had stalled due to the Covid-19 pandemic. It was agreed that this project would be revived during Phase Two working closely with Sellafield who would provide funding and case studies. Unfortunately the level of engagement that was needed from Sellafield in order to carry out this sub-project was not forthcoming, and the work was unable to proceed. The Project Steering Group agreed that internal work within the NDA should focus on establishing whether database preservation is a challenge that they will need to address. To this end, staff at the NDA have developed a survey to scope the scale of the challenge that they might face.

	• **Participation in archiving expert group of IDKM project** The DPC was available to support Gordon Reid (Nuclear Archivist) and his work with the Information, Data and Knowledge Management (IDKM) project. The DPC provided input and advice when requested, in particular on the development of a survey instrument. • **Innovation and Implementation Plan (Phase Three)** Although this workstream was initially included in Phase Two, the NDA need to time to reflect on the lessons learned from the project so far, and it was agreed that there is no immediate need to progress to a third phase of work.
Capability	**Aim: To build capability in the NDA by interrogating state of the art research and 'best of breed' solutions from the digital preservation community and interpreting them to the NDA context, adding value to both as a result.** • **Workforce development** In parallel with this Phase of the project the DPC developed a new Competency Framework which was designed to identify the key competencies required to undertake successful digital preservation, whilst also being flexible enough to suit a variety of purposes (e.g. recruitment, professional development etc.). To accompany the Framework, this project also supported the development of a Competency Assessment Toolkit (CAT) to help facilitate the audit of preservation skills at individual and organizational levels. Both the Framework and the CAT can be used by anyone involved in digital preservation, including practitioners, managers, students, or educators. Both resources were launched to DPC Members in June 2022, and NDA colleagues from Nucleus and Magnox volunteered to help pilot and test the CAT at their organization and have provided valuable feedback to the project team. • **Requirements for a digital archive** The DPC developed a set of "Core Requirements for a Digital Preservation System", which proposed a set of ten high-level functional requirements for a digital preservation system. This document was launched in May 2022 as part of the DPC's existing Digital Preservation Procurement Toolkit. It was well-received by the international community, and in February 2023 the German Federal Office for the Safety of Nuclear Waste Management (BASE) published a German translation. • **Assessment of current infrastructure against requirements** The "Core Requirements for a Digital Preservation System" were published with an accompanying spreadsheet which can be used as a simple tool to assess the functional capabilities of any given preservation system. Colleagues from the NDA used this tool to evaluate the NDA's current digital preservation infrastructure, looking in particular at the functionality of the AMS2. • **Checklist for procuring IT systems** This output focused on the broader topic of procuring IT systems, and how and why digital preservation requirements should be brought into consideration. "Digital preservation requirements for procuring IT

systems" provides typical requirements that practitioners can contribute to procurement exercises for non-digital preservation systems at their organization. These requirements relate primarily to ensuring that digital content can be extracted from the IT system in question, along with necessary metadata, in a form that can be reused without dependence on that system. Ultimately this work is about ensuring that IT systems purchased today have a digital content exit strategy and do not lock-in data when the system reaches end of life.

- **DPC RAM assessment**
  The publication of the DPC's Rapid Assessment Model was one of the significant successes of Phase One of the project, and uptake continued to grow throughout Phase Two, with an increasing number of DPC Members sharing their results each year. In April 2021, an updated version of the DPC's Rapid Assessment Model (RAM) was released, reflecting feedback from the community.

  Throughout Phase Two, the NDA continued to undertake an annual RAM exercise, and this helped identify areas of progress as well as those where further work would be required to reach the organization's desired levels of capability in digital preservation.

  The NDA's RAM results for 2022 showed that the NDA is performing above average in several areas (notably Organizational viability, Legal basis, Continuous improvement, Bitstream preservation and Metadata Management). A full analysis of the NDA's progress against the RAM framework is included in Appendix 2.

- **EDRMS preservation**
  The work of the EDRMS Preservation Taskforce, which was established during Phase One of the project, continued into the first half of 2021. In May 2021, an online briefing day was held to share the experiences of the EDRMS Taskforce and to highlight the forthcoming launch of the EDRMS Preservation Toolkit which was published in July 2021. Both the briefing day and the Toolkit were well-received by the wider digital preservation community.

- **Recovering data from portable media**
  The NDA had identified the need for advice and guidance on recovering data from portable media (e.g. floppy disks, CD-ROM, and portable USB drives). It was proposed that this would form one of the outputs of Phase Two, but this was overtaken by the decision to introduce Metadefender kiosks at key sites across the NDA. The DPC provided input into discussions around Metadefender including guidance on the importance of preserving the authenticity of data as digital materials are ingested into a preservation system, and offered to provide further advice at such point as kiosks were being installed and configured.

- **Defining designated community**
  In order for Nucleus to meet the needs of its users, and to fulfil the NDA's aspiration to align their digital preservation activities with the OAIS Reference Model, the DPC recommended that Nucleus should create and maintain a documented definition of the Designated Community that it serves. This documentation will be required should the NDA seek formal digital preservation certification in the future.

  Alongside wider work on this topic (a #DPClinic community discussion and the drafting of a Technology Watch Guidance Note), the DPC produced a

	succinct definition of the NDA's Designated Community, which was subsequently accepted by the project Steering Group.
Management	**Aim: To provide accountability and quality assurance appropriate to the responsive, iterative project design.** • **Community engagement** As in Phase One, community engagement throughout Phase Two of the project took a number of forms, notably blogging, themed webinars, and conference presentations. The focus of engagement during Phase Two was on the outputs delivered by the project, in particular key publications and reports and World Digital Preservation Day (in November of every year) provided an annual opportunity to talk about the project both inside and outside of the NDA Group. The international iPres 2022 Conference which took place in Glasgow 12th-16th September, was hosted by the DPC. This event presented an ideal forum to present many of the major outputs of the project to the digital preservation community (not least because this was the first major face-to-face conference following the Covid-19 pandemic). As well as a short paper summarizing the work of the project as a whole, workshop and paper sessions were dedicated to the use of the RAM and CAT tools to help organizations assess and develop their digital preservation capabilities. A complete list of the outputs delivered in both Phases is given in Appendix 1 of this report. • **Project dissemination** Originally commissioned during Phase One of the project, Phase Two saw the release of a set of short but informative Data Type Guidance Notes, published in the DPC's Technology Watch series. These were produced in collaboration with Artefactual Systems and made freely available to the wider digital preservation community. Each Guidance Note provides an introduction to the challenges presented by a given data type, and also offers some recommended courses of action. The titles in the series are: ○ Preserving Documents ○ Preserving Email ○ Preserving Spreadsheets ○ Preserving Databases ○ Preserving Raster Images ○ Preserving Moving Images ○ Preserving Audio ○ Preserving CAD ○ Preserving GIS ○ Preserving 3D In addition to these short Notes, several more substantial publications and reports were also produced during Phase Two and these have been detailed in the appropriate sections of this report. The dedicated project webpage on the DPC's website continued to serve as a single convenient location where all the publicly available outputs of the project can be found. • **Project management and reporting** The project management approach adopted in Phase One was carried over into the second Phase of the project. A joint Steering Group made up

	of representatives from both the NDA and DPC, met regularly to receive and review reports on progress, and to provide overall direction to the work of the Project. Regular reports received at this meeting included a running activity log and the project timeline.

Running alongside the areas of work described in the table above was a commitment to providing ad hoc support to the NDA as required. The DPC was on hand to advise the NDA on digital preservation issues over the course of the two years on request. A significant amount of staff time in the first 12 months of Phase Two was focused on discussions around the digitization of analogue documents, photographs and drawings which the NDA was undertaking as part of a major 'sift-and-lift' operation in collaboration with a commercial partner, Restore Ltd. The DPC provided extensive advice and guidance on the preservation implications of particular standards and specifications for digitized material, which had to be balanced against the technologies, time, and staff resources available to Restore.

4. Future work and recommendations

The conclusion of Phase Two of the project is a convenient time for the NDA to reflect upon and embed the good progress in digital preservation that the organization has made over the past four years. The NDA should take this opportunity to incorporate the advice and guidance produced by the project into its business-as-usual processes, and continue to use the tools developed by the project to inform and shape progress. Within a relatively short timeframe, the NDA has succeeded in its aspiration to occupy a leadership role within the international digital preservation community, especially when compared to its peers in the nuclear industry.

In order to consolidate the NDA's achievements and maintain the momentum of the project, the following recommendations indicate some important next steps:

- **Workforce:** The NDA should consider appointing someone within the organization to lead on digital preservation for the NDA and, specifically, Nucleus. It is important that whoever assumes this role should have the remit, authority, time, and skills to be effective. The post holder should be fully supported by and report to the NDA's Senior Information Risk Owner; and should be encouraged to engage across the global digital preservation community.
- **Skills:** The NDA should commit to using the Competency Assessment Toolkit (CAT) where it is useful to do so and act upon its findings. For example, incorporating the toolkit as a regular activity in sections of the organization where digital preservation skills are required should serve to provide an ongoing skills gap analysis. NDA management will then be better placed to consider how to address any shortfalls between current and desired capabilities.
- **Continuous improvement:** Undertaking an annual assessment of the NDA's digital preservation capabilities using DPC RAM, will help the NDA to identify its goals and monitor progress. Furthermore, sharing these annual findings with the DPC will help both the NDA and other DPC Member organizations to benchmark their capabilities and progress with their peers.
- **Digital preservation system requirements:** The NDA should revisit the Core System Requirements and associated assessment spreadsheet produced during this project. This will help identify and prioritize any further work that might be required to enhance the NDA's technical infrastructure for digital preservation.
- **IT procurement:** The NDA should ensure that digital preservation is considered when procuring any new IT systems that may store records and data with long term value. Typical requirements and principles that apply to procuring IT systems can be found in 'Digital preservation requirements for procuring IT systems'.
- **Digital preservation policy:** The draft Digital Preservation Policy produced for the NDA was a key deliverable that remains to be fully exploited. The NDA should consider how best communicate the organization's commitment, vision and strategic direction on digital preservation.
- **Future digital preservation needs:** The NDA should consider the types of digital content that it is required to preserve for the long term, particularly with regard to more complex data types which are widely used across the NDA (such as design and construction records, GIS data, BIM records and databases). The NDA's internal work to understand whether there is a

need to preserve databases is a useful first step. This work should be extended to include other types of data and to consider how NDA digital preservation infrastructure should adapt to respond to future needs.

- **Forward planning:** The NDA should consider creating a forward plan for digital preservation, setting some key targets for the next three to five years. This process may usefully be informed by the findings of RAM and CAT assessment exercises as well as other activities mentioned in this set of recommendations.
- **Community engagement:** The NDA should continue to engage actively with the wider digital preservation community to share experiences and learn from others. Digital preservation is a rapidly changing field, and the advice produced by the DPC over the course of this project will become out of date as community good practice evolves. Effective engagement with the community is one of the best ways to ensure that the NDA is prepared for the challenges ahead. Attendance at events such as the forthcoming iPres 2023 or PV2023 conferences, or the programme of events organized by the DPC, are good ways to learn from others in the field.
- **DPC support and guidance:** As a Full Member of the DPC, the NDA is entitled to ongoing support and advice from the DPC every year. Through this the NDA can continue to benefit from tailored DPC support and is encouraged to do so in order to maintain the momentum of digital preservation work to date.

5. Project legacy

A separate legacy plan was created during Phase One of the project, and has been updated to reflect the findings and outputs of Phase Two. It should be noted that some of the outputs of the project will be delivered outside the current timeframe of the project, such as the updated Technology Watch Report on Geospatial Data (due for publication later in 2023), and the programme of project dissemination events which will be used to promote the findings of the project to the wider community.

Appendix 1 - List of published resources (Phases One and Two)

- DPC's Rapid Assessment Model - a new digital preservation maturity model for organisations with a need to preserve digital content for the long term (first published September 2019, version 2 published March 2021).
- Data Type Series of Technology Watch Guidance Notes - a set of short but informative notes by Artefactual Systems and the DPC on preserving specific content types (published July 2021).
    - Preserving Documents
    - Preserving Email
    - Preserving Spreadsheets
    - Preserving Databases
    - Preserving Raster Images
    - Preserving Moving Images
    - Preserving Audio
    - Preserving CAD
    - Preserving GIS
    - Preserving 3D
- EDRMS Preservation Toolkit - developed by the EDRMS Preservation Taskforce (published July 2021).
- Preserving Born-Digital Design and Construction Records - a Technology Watch Report by A Leventhal and J Thompson on the preservation of 3D design and construction datasets (published December 2021).
- Core requirements for a digital preservation system - 10 core requirements for a preservation system that can provide a starting point for requirements development or be used as tool for assessing current systems (published May 2022)
- DPC Competency Audit Toolkit (CAT) – DPC CAT provides users with practical, structured processes for assessing the skill levels of those undertaking digital preservation work at their organization, to assess current skills and highlight gaps (published June 2022)
- Digital preservation requirements for procuring IT systems - this guide notes the requirements that should be considered when procuring an IT system (for example, an EDRMS, DAMS or GIS), that may ultimately contain at least some records or digital content that need to be retained beyond the life of the system (published January 2023)
- Preserving Geospatial Data - a revised Technology Watch Report on the preservation of geospatial data (forthcoming 2023).

Appendix 2 – NDA RAM results analysis 2022

DPC RAM analysis for NDA/DPC Project Steering Group meeting - September 2022

As part of Phase One of this project the NDA were keen to have a means of measuring their progress in digital preservation and benchmarking against others. The DPC's Rapid Assessment Model (DPC RAM)[1] was therefore created by the DPC in 2019 in conjunction with the NDA as an output of Phase One of the project. Version 2 of the model was released in April 2021. Users of this maturity model are guided through an assessment in which a score is assigned (from 0-4) for each of the eleven sections of the model. Users are asked to consider both where they are now and where they would like to be in the future. The DPC encourages members to carry out a self-assessment using DPC RAM on an annual basis and share their results with us. This enables us to understand digital preservation maturity across the membership and also allows organizations to benchmark against others.

This document reports on the latest DPC RAM assessment for the NDA and compares it with aggregated summary statistics from those DPC Members who shared their results this year.

NDA results for 2022

DPC RAM assessment for the NDA this year was carried out in May by Michelle Donoghue, Debra Kirkpatrick, Gordon Reid and Stephen Beck and shared with the DPC. The picture was quite different to last year with improvements in a few key areas and targets altered in several areas.

- **Organizational viability** saw a move from Basic to Managed level and a revision of the target up to Optimized. Much activity (around pilot projects, governance, the AMS2 and digitisation) had been carried out over the last year which fed into the revised score for current level.
- **Policy and strategy** had moved from Awareness to Basic level and the target had been revised downwards to Managed, thus narrowing the gap between current and target levels – the development of retention schedule and metadata standards led to the revision of current level.
- **Legal basis** has jumped from Basic to Optimized level.
- **IT capability** has dropped from Basic to Awareness with the target level of Optimized remaining in place. There was a recognition that there is much still to be done in this area but optimism that the NDA digital strategy will help move things forward.
- **Continuous improvement** saw both current and target levels moving up by one, with DPC CAT, AMS2 and NDA strategies that are in the pipeline all contributing to this improvement.
- **Community** scores remained static at Managed and Optimized, though there was a sense that engagement with the iPRES 2022 conference this year may help the NDA move forward.
- **Acquisition, transfer and ingest** current level remained at Awareness, but the target has been revised down to Managed. Several activities in the pipeline will help move this area forward (AMS2 launch, bulk upload, EDRMS roadmap).

[1] https://www.dpconline.org/digipres/implement-digipres/dpc-ram

- **Bitstream preservation** has made a big leap forward, from Awareness to Managed, with the target remaining at Optimized. Work has been undertaken on checksums and audit trail and will be delivered through the AMS2.
- **Content preservation** has seen a narrowed gap between current and target levels, with current level moving from Awareness to Basic and target moving down from Optimized to Managed.
- **Metadata management** has moved from Basic to Managed with target level remaining at Optimized. Standards have been agreed across the NDA Group and controlled vocabulary is being explored currently.
- **Discovery and access** has moved from Awareness to Basic level now find and retrieve is in place, with target remaining at Optimized. The online catalogue development is expected to help move the NDA forward with this section.

Results for the last three years are illustrated in Figure 1, 2 and 3 below and demonstrate strong improvement in many areas since the project began and a moving of the NDA closer to its digital preservation goals.

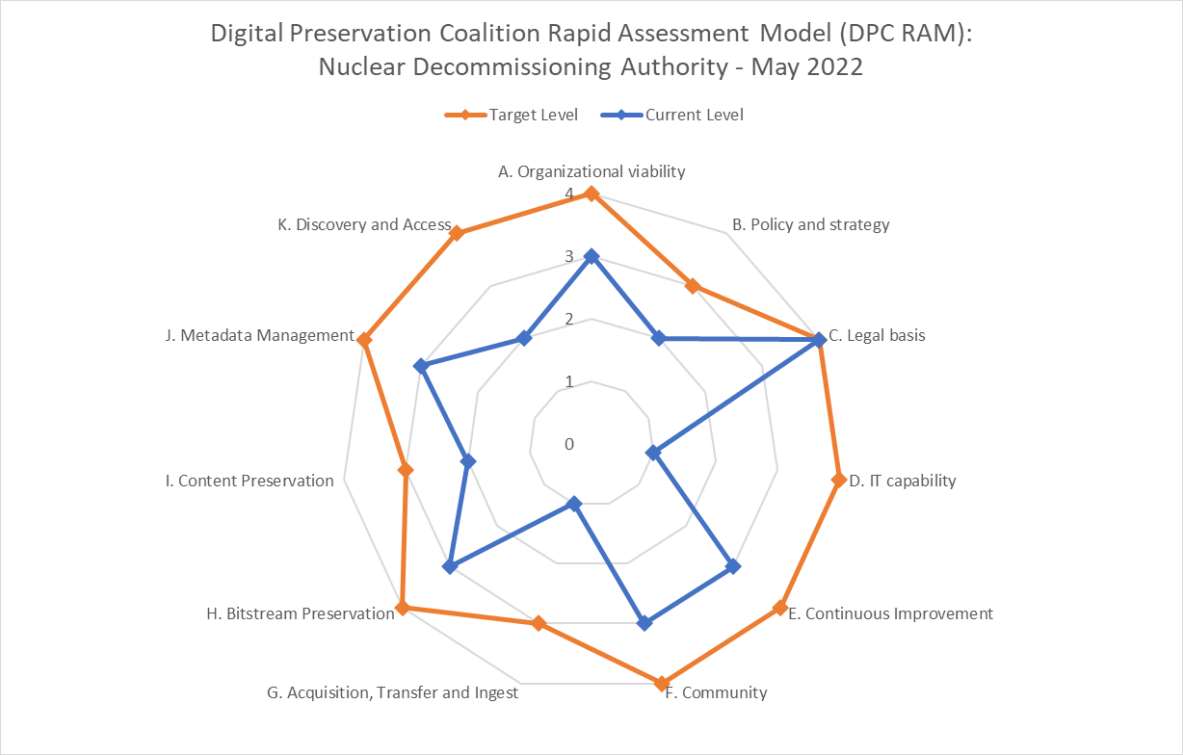

Figure 1: DPC RAM results from NDA – May 2022 (completed by Michelle Donoghue, Debra Kirkpatrick, Gordon Reid and Stephen Beck)

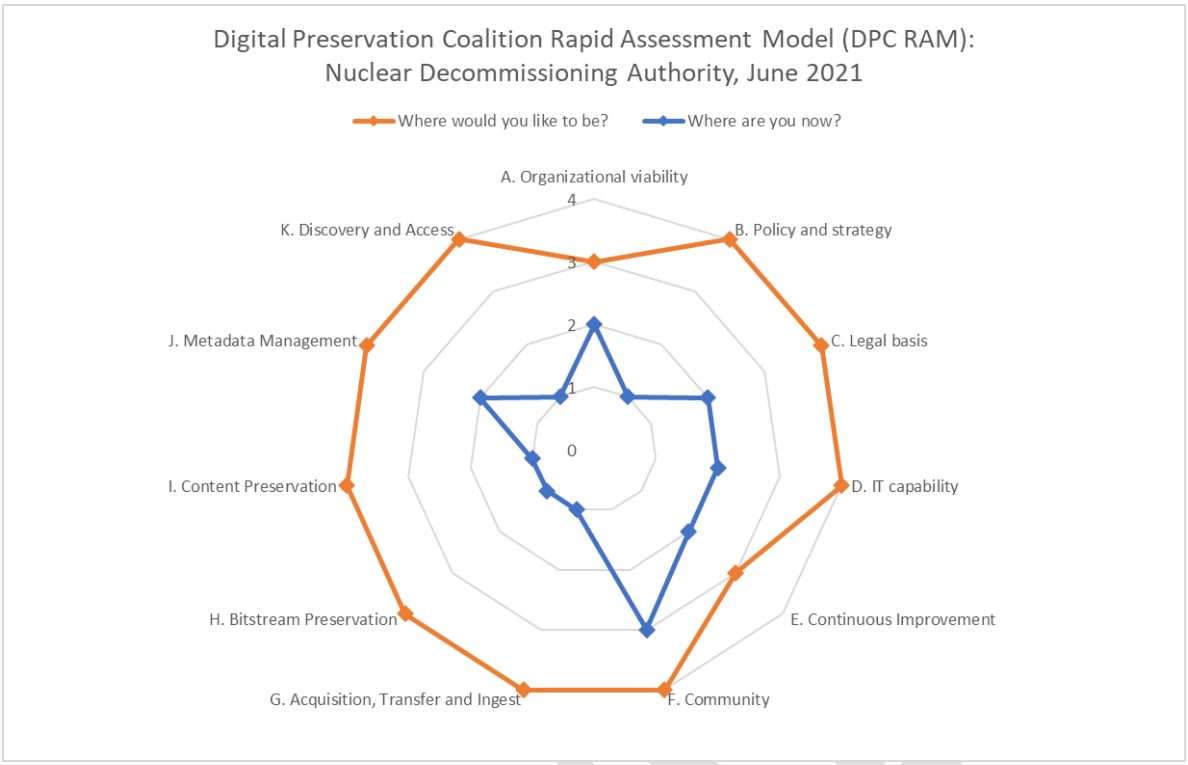

Figure 2: DPC RAM results from NDA – June 2021 (completed by Jenny Mitcham and shared with NDA for comment)

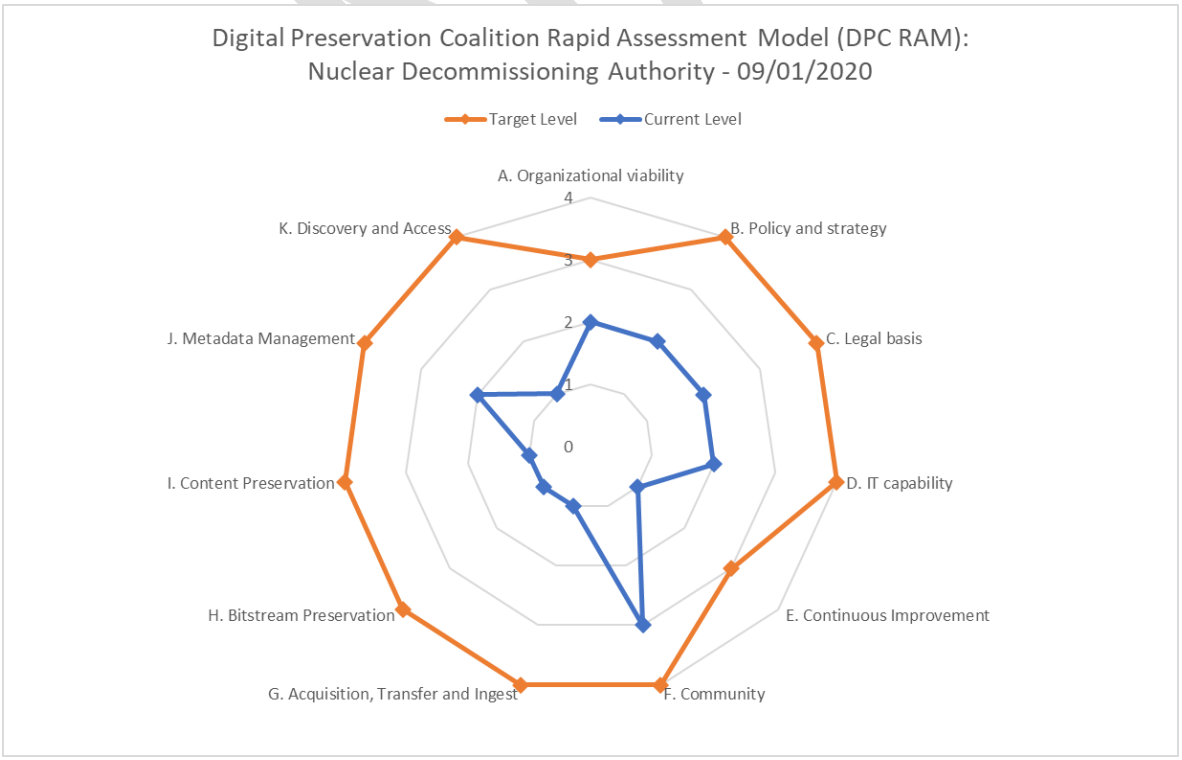

Figure 3: DPC RAM results for NDA – January 2020 (completed by Debra Kirkpatrick)

DPC Member results for 2022

Summary results for the 41 DPC Members who shared their results with us this year are similar to last year with a slightly higher set of scores for the organisational capabilities of

RAM (sections A to F), particularly the notable peak at 'Community', and lower scores for the service capabilities (sections G to K).

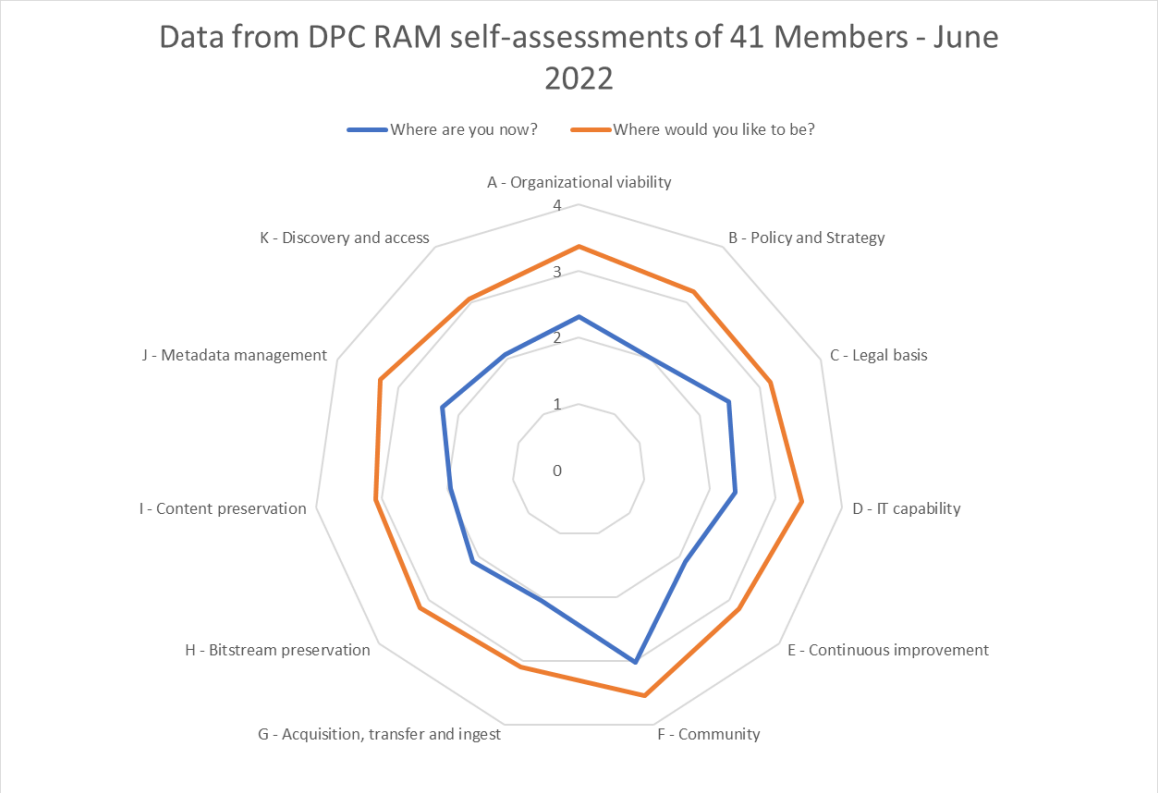

Figure 4: DPC RAM results for DPC Membership – June 2022

Comparison of 'Where are you now?' results

The following radar chart shows a comparison of the 'Where are you now?' results for the NDA and the rest of the DPC membership for 2022. It is clear that the NDA is now performing above average in several areas (notably Organizational viability, Legal basis, Continuous improvement, Bitstream preservation and Metadata management). In other areas, NDA scores are very much comparable to those of the DPC membership (Policy and strategy, Community, Content preservation and Discovery and access) and in two areas the NDA fall behind the average set by the DPC community (IT capability and Acquisition, transfer and ingest). Overall, this is an impressive and healthy result for the NDA. In particular it is encouraging to see how recent work on the AMS2 has helped move the NDA forward with some of the service capabilities of DPC RAM since last year.

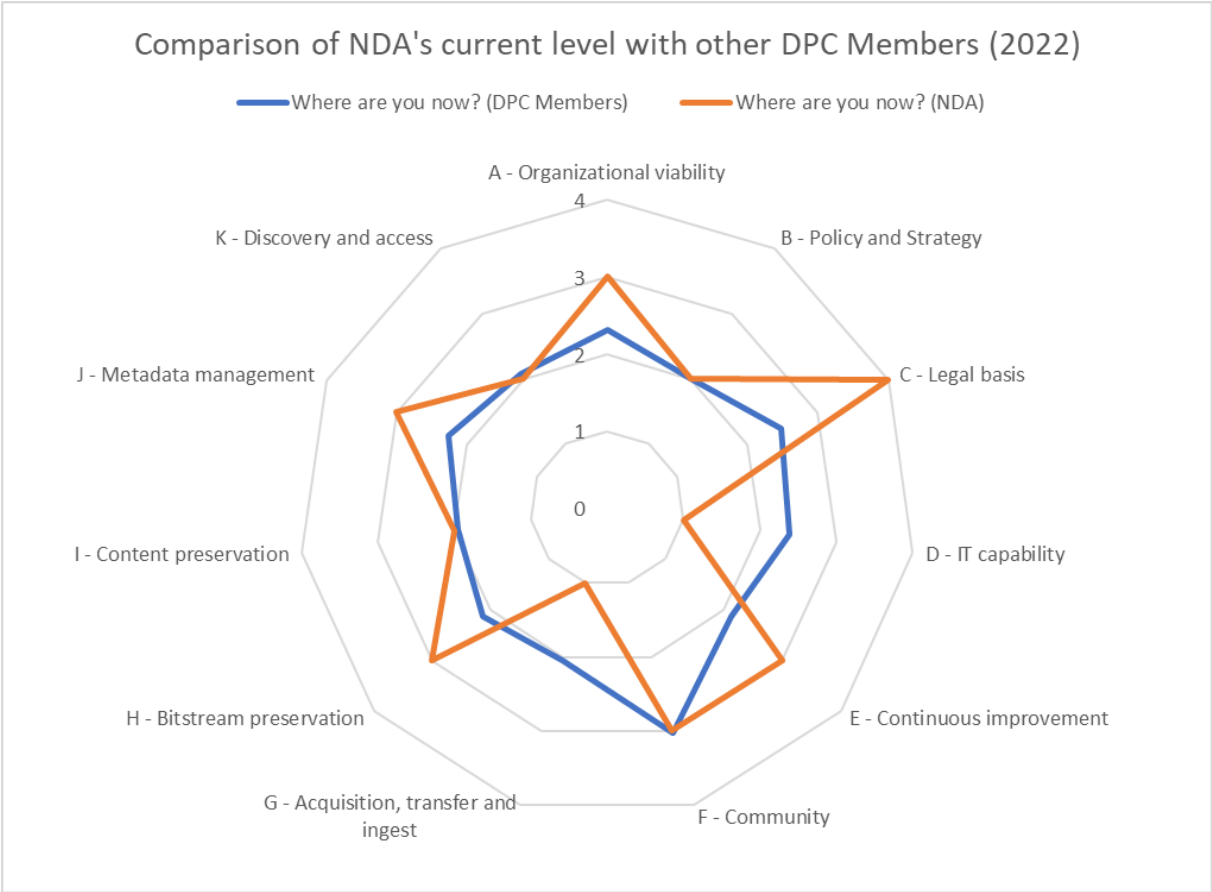

Figure 5: comparison of NDA's current RAM scores with a summary of DPC member scores for 2022

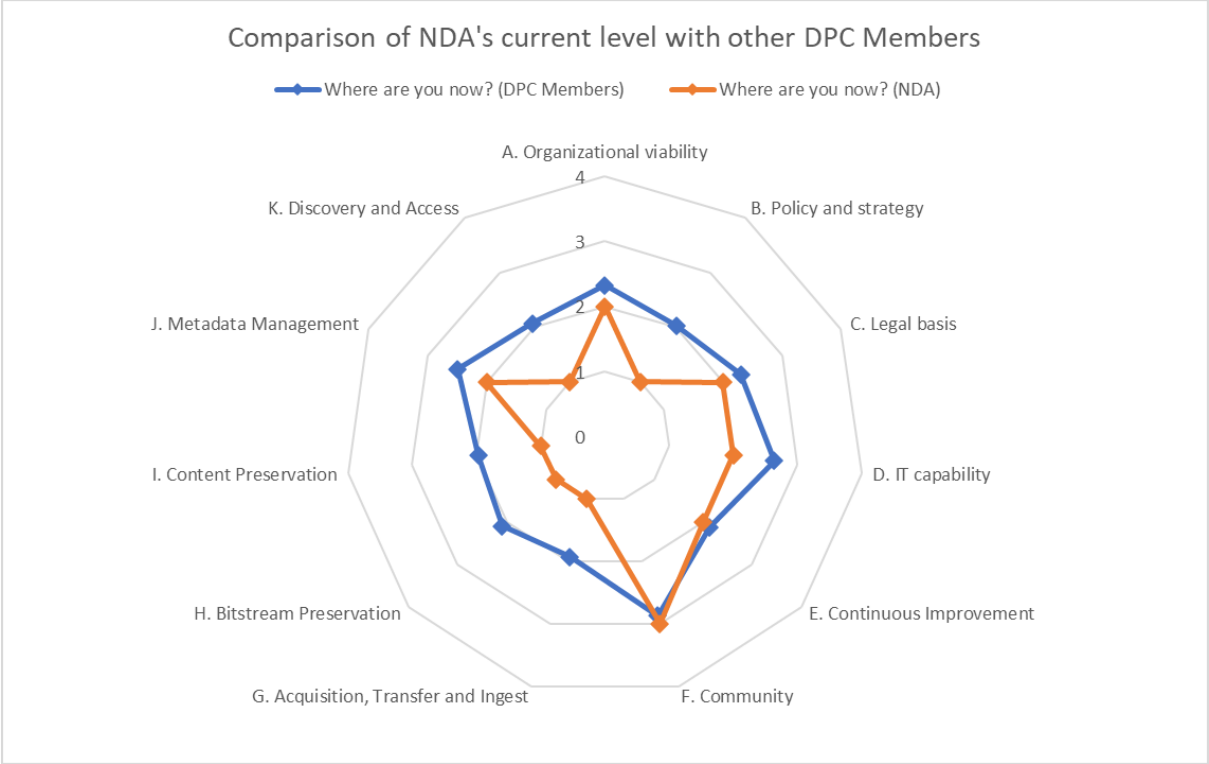

Figure 6: comparison of NDA's 2021 RAM scores with a summary of DPC member scores for 2021

Conclusion

At the start of Phase One of this project, the NDA were clear that they wanted a method for measuring their progress in digital preservation. DPC RAM was created to meet this aim and has had a big impact on the global digital preservation community as well as providing a tool for the NDA to use as the project progressed. Using DPC RAM for self-assessment of digital preservation capabilities and progress has been helpful in demonstrating the progress the NDA has made towards goals over the course of the last three years. It is recommended that the NDA continues to use DPC RAM on an annual basis to check in on progress and set future directions.