

# DiAGRAM Workshop Exercise

For the DPC Quantifying Digital Preservation Risk workshop 1

## Task 0 - set up

- a) Open the application. The URL will be <https://nationalarchives.shinyapps.io/DiAGRAM> suffixed by your assigned number e.g. <https://nationalarchives.shinyapps.io/DiAGRAM3>. Your facilitator will let you know what the assigned numbers for your breakout room are.
- b) Read the guidance on the 'Home' page.
- c) Go to the 'Definitions' tab and explore some of the 'nodes'. In particular, look at their definitions, node states and how they fit into the overall network structure.

## Task 1 - create model

Imagine yourself in the following scenario:

"I work at a well-established archive with a long history of managing and preserving analogue records. In fact, I would say that we are at the top of our game in terms of preserving analogue materials. Digital is another matter. We have had a surrogate collection for a while, but have started to receive a few born-digital record deposits and we don't know much about how to preserve these. The senior managers within the organisation would like us to improve our capability for preserving digital collections and have asked us to look at which options would be the most effective."

- d) Go to the '1. Create your model' tab and answer the questions using the answers below. When you have put in your answer, press 'Next' once and wait for the question to change.

Q	Node referred to	Input	Reasoning
1	Technical Skills	5%	We have no experience but are keen to learn.

2	System Security	70%	We have IT support for our cataloguing software and online catalogue and our surrogates are held on the organisation's corporate servers but there is no dedicated secure storage for archival records.
3	Checksum	0%, 0% 100%	We don't know what this is.
4	Information Management	50%	We are great at this for our analogue collections, but might need a bit more work to bring it to the same standard for digital records.
5	Digital Object	1%, 0%, 99%	At the moment almost all the digital records we have are surrogates, but we'd like to see what happens when we receive more in the future as a scenario.
6	Storage Medium	1%, 99%, 0%	The born-digital records are stored on stand-alone USB hard drives, the surrogates on corporate servers in house.
7	Replication and Refreshment	99%	We don't have copies of the content on the USB hard drives but the surrogates are backed up by the organisation's IT Department.
8	Operating Environment	100%	We have a lot of protections already in place on site to protect our analogue collections and corporate IT network.
9	Physical Disaster	High	We are based in an area with high flood risk from the sea.

e) At the end call you model 'Workshop' and press 'Save model'.



Please answer the following question:

**All questions answered. Please give model a name:**

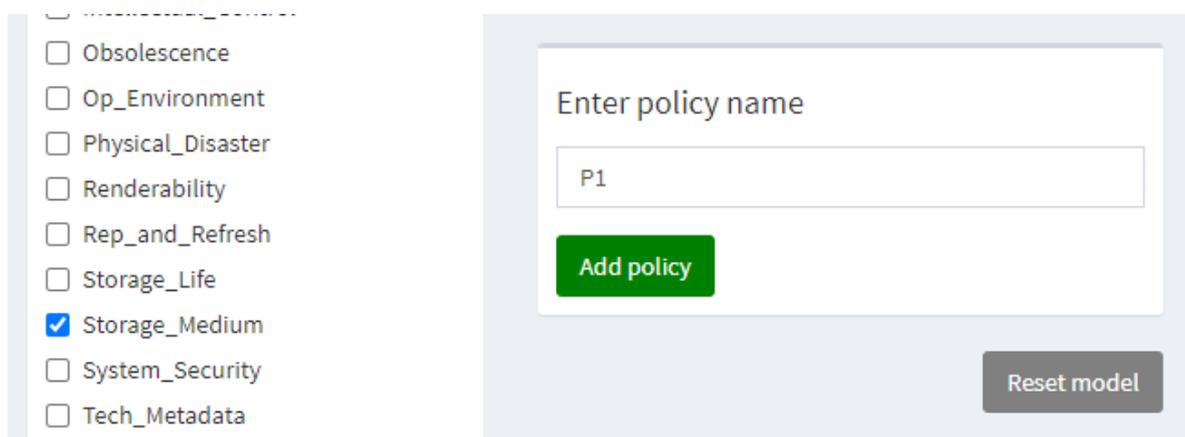


f) Scroll down and look at the 'Model risk score' plot. You should see a two coloured bar labelled 'Default' and another labelled 'Workshop' which is the one you have just created. The height of the bar represents how good each model is at mitigating against the digital preservation risks. The higher the score, the better. **How do these compare?**

- g) The purple (bottom) colour of each bar represents the score specifically for renderability, and the orange (top) colour represents the score for intellectual control. **How do the scores for each of these outcomes compare when looked at separately?**

## Task 2 - compare policies

- h) Now go to '2. Compare policies' and make sure you select 'Workshop' as the model.
- i) Policy 1: "What if we outsourced all our data storage?" This is equivalent to storage medium type 'C'. Select the node you want to change in the checklist on the left, in this case 'Storage\_Medium'. Update the sliders so that C is on 100% (the maximum it can go). Name the policy "P1" and press 'Add policy'. You should now see another bar appear on the right hand side. **Does this policy give a higher or a lower score compared to your original model?**



The screenshot shows a user interface for configuring a policy. On the left side, there is a vertical list of nodes with checkboxes: 'Obsolescence', 'Op\_Environment', 'Physical\_Disaster', 'Renderability', 'Rep\_and\_Refresh', 'Storage\_Life', 'Storage\_Medium' (checked), 'System\_Security', and 'Tech\_Metadata'. On the right side, there is a form titled 'Enter policy name' with a text input field containing 'P1' and a green 'Add policy' button. Below the form is a grey 'Reset model' button.

- j) Press the 'Reset model' button before trying another policy
- k) Policy 2: "What if we increase our technical skills to 60% and start self-generating checksums on receipt to increase this to 10%?" This time select 'Technical\_Skills' and 'Checksum'. For the checksum input, set the sliders to 10% 'Self\_Generated' and 90% 'No' leaving 'Yes' at 0%.. For technical metadata, set the slider to 60%. Name the policy "P2" and press 'Add policy'. You should now see another bar appear on the right hand side. **Does this policy give a higher a lower score compared to your original model? How about compared to policy 1?**
- l) Reset the model again. Now we are going to test a scenario rather than a policy.
- m) Scenario: "We expect our born-digital collection will increase more rapidly than the others." What would our risk be if they accounted for 50% of the overall collection? Change the Digital\_Object to be 50% born-digital and 50% surrogates. Name this 'S' and press 'Add policy'. **Does this policy give a higher a lower score compared to your original model?**

- n) Policy 3: “In this future scenario, what if we made sure we had the tools to render the material 95% of the time?” This time, do not press ‘Reset model’. Instead, select the ‘Tools\_to\_Render’ node and deselect ‘Digital\_Object’. Input 95% for rendering tools and save this policy as ‘S-P3’. **How does this compare to the scenario model?**

### Task 3 - report

- o) Go to ‘3. Report’ and make sure you select ‘Workshop’ again. You will see the same column chart on the right, but this time there are also numerical scores displayed in the ‘Summary’ box on the left. **What was the highest scoring policy and what was this score? Which policy would you want to start implementing straight away? What more policies would you want to test?**
- p) Download the plot. A zip folder should now have downloaded. Unzip this folder and save it somewhere on your PC.

### Task 4

Now it is over to you to try the model for yourself. Create a model based on your archive’s policies and practices, then test out a few different policies and see what the impact is. If you do not work at an archive, set the inputs however you like.

A facilitator will be on hand to help with any questions you may have, and you might find it helpful to discuss and compare with the other workshop participants also.

-----