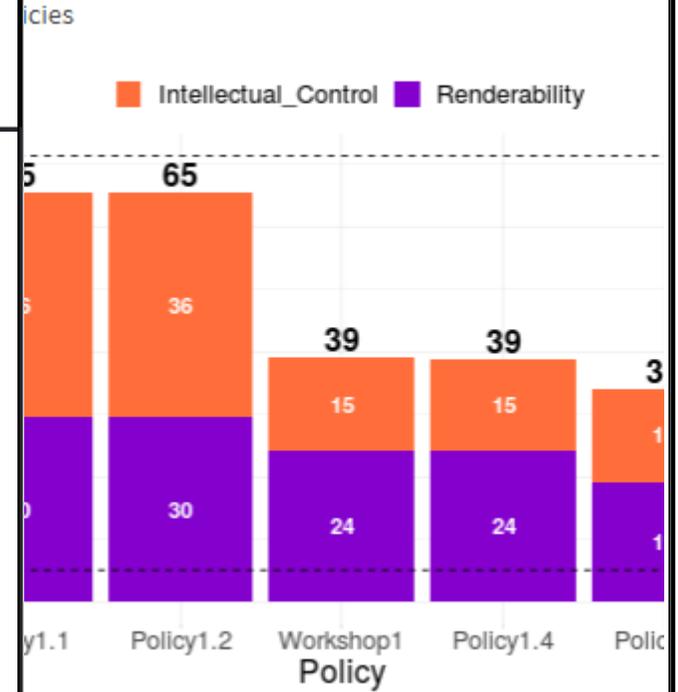


THE

NATIONAL

ARCHIVES

# Introducing DiAGRAM: The Digital Archiving Graphical Risk Assessment Model



Hannah Merwood and David Underdown

17 July 2020



Engineering and Physical Sciences Research Council













THE

NATIONAL

ARCHIVES



Live Demo

<https://nationalarchives.shinyapps.io/DiAGRAM-dev/>



### Important note: This model is still in development

There will be further user interface changes and additional functionality added as the project progresses. Any feedback to inform the future development would be welcome - please send your comments to a member of the project team.

about the project please see our [project page](#).

## DiAGRAM - The Digital Archiving Graphical Risk Assessment Model

Version 0.9.7 (Prototype)

This is the Digital Archiving Graphical Risk Assessment Model (DiAGRAM) built by the University of Warwick and The National Archives with support from the National Lottery Heritage Fund and the Engineering and Physical Sciences Research Council. For more information about the project please see our [project page](#).

Before using the tool for the first time, we would advise you to read the [presentations](#) from our online workshop with the Digital Preservation Coalition, where there is also an [exercise sheet](#) you can work through.

Create your model

Archivist has full intellectual control.

Data from April 2020 were used to inform the

### Introduction

This decision support tool enables users to score their archive's digital preservation risk and then explore how this would change under different policies and risk scenarios. The risk score is based on the proportion of files in the archive that are renderable and where the archivist has full intellectual control.

The underlying methodology used to create this model is based on a Bayesian network - a probabilistic graphical model that captures the conditional dependencies of risk events. When historical data were unavailable, data from an expert elicitation session conducted in April 2020 were used to inform the probabilities needed for this model.

This interface enables users to:

- Understand the risk definitions used in the model and how the risk events are linked together
- Create a model that reflects the policies and practices for their Digital Archive
- Test alternative policies to see how this impacts the risk score
- Download the model and a summary of the results
- Upload a pre-built model and continue exploring scenarios from there
- Update the probability tables for the model based on the user's own data or experience
- Create bespoke scenarios by directly manipulating the probabilities used in the model



Engineering and Physical Sciences Research Council

### Guidance

**Definitions :** This page has a visualisation of the underlying network of digital preservation risks and allows you to see the full definitions, states and data sources used for each 'node'.

**1. Create your model :** This goes through 9 questions to create a risk model and a score which is based on the user's archive and policies.

**Recommendations :** This page looks at the impact changing each of the answers to the input questions would have to the risk score.

**2. Compare policies :** Create and save different policies and see how the risk score changes.

**3. Advanced customisation :** This tab allows users to edit the marginal and conditional probabilities in the model directly. This allows for users to input their own data for any nodes within the model or create scenarios by altering conditional probabilities.

**4. Report :** This contains a summary and comparison of the policies for each model, and allows the model and plots to be downloaded.

DiAGRAM home page

NATIONAL

ARCHIVES



# Definitions

## Node selection

Please select a node from the menu to view its definition.

Checksum

## Node description

**Definition:** A unique numerical signature derived from a file that can be used to compare copies (definition from the DPC handbook: <https://www.dpconline.org/handbook/glossary>). A checksum is needed to ensure integrity of the digital object.

**Data source:** Loosely based on JISC digital skills survey results

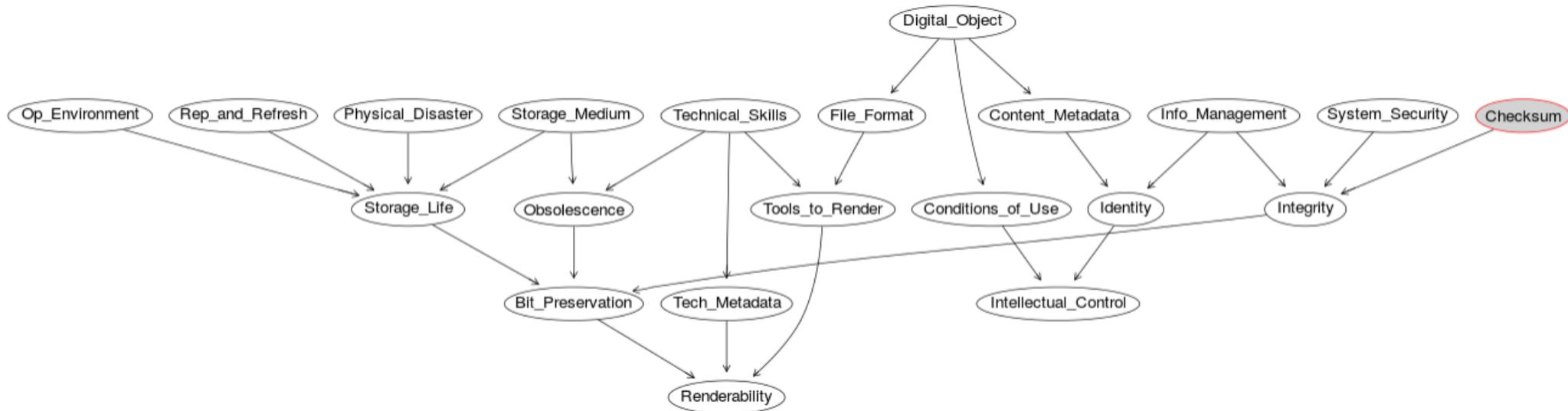
**Data collected:** 2019

### Node State

### State Definition

Yes	We have a checksum from the depositor.
Self_Generated	We do not have a checksum from the depositor but generated one ourselves instantly upon receipt.
No	We do not have a checksum or have only generated one post-ingestion.

## DIAGRAM structure



Definitions page, showing details of the Checksum node

**NATIONAL**

**ARCHIVES**

THE																	
NATIONAL		Modelling your archive															
ARCHIVES																	























**THE**

**NATIONAL**

**ARCHIVES**

Comparing policies













**THE**

**NATIONAL**

**ARCHIVES**

Advanced modelling



