

Preserving and sharing born-digital and hybrid objects from and across the National Collection

Decision Model Report: March 2022

Decision Model Report

Authors

Tom Ensom, Freelance Digital Conservator
Stephen McConnachie, British Film Institute

Acknowledgements

We are grateful to all interviewees and workshop participants for their substantial contribution to the research. We also recognise the critical and sustained feedback from our colleague Corinna Gardner, and the contribution to developing the case studies from Will Massa and Henrik Mauler. This research has been possible thanks to the AHRC grant AH/T01122X/1.

We would like to thank Kaja Marczevska, Claudia Heidebluth and Joanna Norman from V&A Research for their support on the project.

We would like to thank Rachel Boon, Ian Cooke, Corinna Gardner, Jenny Mitcham, Caylin Smith and Arran Rees for their valued critical feedback during the development of this model.

Publication design by *Supanaught*.

Lead research organisation



In partnership with



Funded by



Contents

1. Introduction
2. Methodology
3. Community Feedback
4. Discussion
5. Future Work
6. Decision Models

1. Introduction

When considering the complex challenges faced by cultural heritage organisations in collecting, preserving and sharing born digital and hybrid objects, it becomes clear that the process of defining solutions as a community of practice is in its early *probing* phase: characterised as tentative, exploratory, questioning, experimental. The workshops within this *Preserving and sharing born-digital and hybrid objects from and across the National Collection* project, which examined the case studies from multiple angles, yielded a richly discursive examination of the main considerations.

This *Decision Model* represents an attempt to create a structured representation of those main considerations and the discourse from the workshops, to codify the main decision-making processes that an organisation may go through when assessing an acquisition of such an object, categorised into high level areas. It attempts to create a traversable system that could be used by collections professionals in their work - policy makers, managers, collections management or digital preservation practitioners, conservators.

The scale and complexity of the challenge - the case studies represent a very broad range of object and content types, with a very large variation in preservation and access concerns - necessarily makes the decision model complex and in some areas difficult to navigate. Time permitting, a decision model could have been created for each case study, to represent the specific factors for each object-content type. However, this model attempts to achieve that for the general concept of a born-digital and hybrid object, and as such it risks forcing a huge range of complex factors into a simple model. We discuss the inherent limitations and issues of the model in some detail below, with attention to usability issues for

the community it addresses, and suggestions for future / alternative developments.

The practical usefulness of the traversable representation was tested by a group of community peers with experience of real-world born-digital and hybrid object acquisitions. Feedback from this process informed further refinements to the model and is discussed further below. The model could be further critiqued and developed by the community after the project, if it proves useful.

2. Methodology

The decision models were formulated based on a set of key questions in the acquisition, digital preservation and access provision of complex born-digital or hybrid digital-physical objects ¹. These questions were developed using insights from workshops, case studies and interviews undertaken during the *Preserving and sharing born-digital and hybrid objects from and across the National Collection* project, and the authors' own experiences working in collecting institutions. Key questions were grouped within higher level categories, which would define the scope of the individual models:

- **Technical Constraints:** Identify the types of components that constitute the object and assess how these will affect the collecting activity.
 - **Technical Constraints - Digital Data:** As above, but questions relating specifically to digital data objects.
 - **Technical Constraints - Software:** As above, but questions relating specifically to software objects.
 - **Technical Constraints - Web Content:** As above, but questions relating specifically to web content.
- **Collection Policy:** Assess the extent to which the object complies with existing collection policy and navigate the consequences of non-compliance.

- **Data Protection:** Identify and navigate data protection issues arising from supply of collection of personal data.
- **Intellectual Property Rights:** Identify and navigate intellectual property rights issues.

The decision was made to frame the models as tools which could be used when assessing the viability of an object for acquisition. This is a point at which navigating these questions is likely to be particularly important, and limiting scope to acquisition helped refine terminology and question wording.

Two assumptions are made regarding expectations for the models and their use. The first is that using them is likely to necessitate collaboration between various stakeholders, including the creator of the acquisition candidate and various individuals/teams within the collecting institution. Aspects of the decision-making process require an understanding of the object's meaning and context, its technical constituents and their production, and applicable institutional and legal structures. The second assumption is that these models will not yield a definitive set of actions for the collecting activity. Instead, they are formulated with the aim of exposing key questions, problems and responses that can inform this work.

The decision was made to use a flowchart-like visual diagramming approach to representing the decision models. The Miro diagramming tool was used to develop the diagrams, which offered a multi-user platform to simplify collaboration and suitable options for export. A diagramming notation

¹ Object is used in a loose sense here to refer to anything ranging from a single digital file to a set of physical and digital components.

was developing for the models based on the widely understood flowchart notation:

- **Rounded rectangle:** Colour coded terminal; Diagram Start (Green), End (Black)
- **Diamond (solid yellow fill):** Decision taking the format of a question followed by two branching paths (Yes or No).
- **Rectangle (solid pink fill):** Actions, processes and points for reflection. In some cases these are tagged with 'flags' to indicate greater than normal risk:
 - **1. "Critical Issue Flag":** Issue may significantly impact the viability of the acquisition candidate.
 - **2. "Elevated Risk Flag":** Issue may cause higher than usual risk of loss.
- **Square (solid green line):** Notes that provide further contextual information on a specific node.

Each model was developed as an initial draft, and then iteratively reviewed and further refined.

3. Community Feedback

We invited a group of eight digital preservation experts to review the decision models and provide comments and criticism. Each reviewer was sent a copy of the Miro board containing the decision models, which they could annotate directly with notes and comments. All feedback was reviewed and, where possible, suggestions were integrated into the decision models. Generally feedback was positive and reviewers felt that the models would be useful to the wider community. Some feedback was identified as out of scope for integration during the project and is summarised below.

Some reviewers felt that steps could be taken to make the decision models more user friendly. One suggestion was clearer signposting to guide navigating the models and the order in which to work through them. We decided not to add a single model linking together all other models so that each model could stand alone, but instead added a pathway through the models as an entry point for first time users. Other suggestions for usability improvements included providing a clearer means to record or summarise the actions resulting from traversing the model and worked examples to demonstrate how the models can be used. A few users found some of the terminology unclear, and one suggested that it might benefit from glossary. The term 'Creator / donor' terminology was singled out as not inclusive enough (e.g. where an acquisition candidate is a purchase, or where it is owned by a group or organisation).

There were several instances where reviewers found binary Yes/No decisions to be inappropriate to convey the uncertainty of the decision at hand. It was proposed that the model would benefit from additional ways of indicating uncertainty in these situations. There were also a few instances where reviewers disagreed with us (and sometimes each other) over the

assignment of the 'Critical Issue' and 'Elevated Risk' flags. Our impression is that this is likely to reflect different institutional contexts and perspectives, and is difficult to account for in a single model. As one reviewer pointed out, most objects are facing a degree of risk of loss. With that in mind, can risk be meaningfully classified? From this research, it seems that assignment of risk is not likely to be consistent among different practitioners.

Another reviewer suggested a need for care over framing of risks, as warnings could be off-putting to some users. The reviewer noted that in such situations, work towards collaborative change and collective stewardship might help mitigate these risks. Reviewers had very different attitudes towards acceptability of an experimental (i.e. higher risk) approach to acquisitions, which is factored into some parts of the model. Some loved this idea and felt it represented a challenge to notions of guaranteed preservation, while others felt it would not be possible at their organisation. This reflects a difficult balance between the need to support and encourage new forms of collecting, while also making those tasked with collection care aware of the very real risks of losing access to collection objects.

Finally, one reviewer suggested that the Data Protection model may benefit from being reviewed by someone with an archival background, so that more detail might be incorporated. Examples of embargo periods and public task/good arguments were given as areas which the model could be extended to cover.

4. Discussion

Further to the feedback provided by members of the digital preservation community, in this section we will critically reflect on the limitations and successes of the decision models. Overall, we think the models can provide a useful tool for those working in collection care roles, and help navigate the array of complex decisions they will have to navigate when bringing complex born-digital and hybrid objects into a collection. However, we have identified a number of limitations to the models in their current form.

Some limitations related to usability. These stem in large part from inherent challenges of converting nuanced decision-making into a navigable visual representation. The models force the user to choose a sequential route through the decisions. This contrasts with the reality of decision-making during acquisition, which can be non-linear and unfold in a variable order over long periods of time. These limitations are further exacerbated by the variability of complex/hybrid digital objects and the acquisition processes required to bring them into collections. We attempted to include broad questions in the diagrams which allow for a variety of contexts of application, but ultimately the flow-chart diagram approach still forces questions and actions to be codified as neat text boxes. Some of what we have modelled in general terms might be best modelled through the formalisation of workflows as a direct response to a specific institutional context, rather than be compromised through generalisation.

There are also limitations which arise during any kind of modelling activity due to the subjectivity of the modelling process. A two person team, even with extensive research to support them, will inherently bring their own perspectives and biases to the modelling process. Any resulting models therefore risk failing to take into account the multiple perspectives

that exist within and among collecting institutions and within the digital preservation community. We acknowledge that there is no one way of handling the acquisition of complex and hybrid digital objects, and therefore we can only represent a restricted view on this dense and challenging topic. One way in which we have attempted to address this limitation is through inviting community feedback. All reviewers highlighted different strengths and weaknesses in their feedback, and there would be value to extending review beyond what was achievable within the project (see below).

Finally, the appropriate level of depth and complexity to model was very hard to gauge during the modelling process, and may not be appropriate to all potential users. This may have been a consequence of the broad coverage of the models and the resulting lack of clear audience; there is no single role with responsibility for the activities the models cover and who does what within an organisation can vary considerably. The challenge is therefore finding an appropriate balance between conveying the complexity of the issues at hand and not slowing the user down with unnecessary detail. We have modelled at a level that felt appropriate based on our experiences, but initial testing has already shown this to diverge from expectations of some potential user groups.

Despite the clear limitations, there are insights to be gained from the process of developing the decision models. Modelling has drawn into focus some of the challenges faced by the cultural heritage sector in collecting and caring for complex born-digital and hybrid digital-physical objects. Firstly and most significantly is the incredible variety of object types which can be acquired, be they singular, multiple or just a part of a hybrid/aggregate object. This makes any attempt at standardisation/formalisation

very challenging and perhaps impossible. We might therefore expect further divergence in approaches and methodologies for different kinds of object in the future. Nonetheless, the categories developed for this model – digital data, software and web content – may be a useful baseline level of granularity at which to appraise hybrid/aggregate objects.

More generally, the modelling work has demonstrated that modelling can help us learn more about complex subjects, even if the outcomes are not concrete. In this case, it has revealed a number of interesting questions regarding uncertainty in process of acquiring and preserving complex and hybrid digital objects, including:

- What is a 'safe level' of care which can allow an acquisition to proceed? We flagged issues that would put the acquisition candidate at higher risk of loss, but the user was still able to continue to the end of the model. What are the implications of an acquisition proceeding despite elevated risk?
- How do we integrate complex acquisitions that contain an aggregate of physical and digital objects with the realities of institutional working? Acquisition of such objects is highly interdisciplinary and demanding on collection care teams. What kind of institutional resourcing, infrastructure and expertise is required to support this? Where can cross-institutional collaboration support this work?

5. Future Work

The decision models developed would benefit from further refinement and improvement to improve their coverage, ideally through the input of others in the digital preservation community. This might allow them to live beyond the project, gain broader relevance and become a useful tool moving forward. However, it may also be worth considering the more fundamental question of whether the flowchart-like diagram format is suitable for conveying the topics we have modelled. Alternatively, some of the ideas developed might be carried forward in a different format. A checklist might be a simpler, more digestible form in which to present many of the key questions/activities represented in the diagrams. By removing the directionality of the decision model, a checklist further accommodates non-linearity. However, it also removes the allowance for branching paths that a flowchart (or similar) accommodates and might become overwhelmingly large if not presented in a considered format.

6. Decision Models

These decision models are tools designed to help navigate decision-making for the acquisition and digital preservation of complex born-digital or hybrid digital-physical objects. Each model addresses a particular aspect of the acquisition process. These are presented in a non-linear form and can be navigated in parallel, as needed based on requirements:

- If you need to assess the digital constituents of the acquisition candidate and their implications for the collecting activity, traverse **Technical Constraints** and any relevant sub-models:
 - If the acquisition candidate incorporates digital *data* objects (e.g. video, audio, images), traverse **Technical Constraints: Digital Data** decision model for each type of data.
 - If the acquisition candidate incorporates software objects, traverse the **Technical Constraints: Software model** for each software program.
 - If the acquisition candidate incorporate web content, traverse the **Technical Constraints: Web Content** model for each type of web content.
- If you need to assess the acquisition candidates compliance with relevant Collection Policies, traverse the **Collection Policy** model.
- If you suspect any issues relating to Intellectual Property Rights, traverse the **Intellectual Property Rights** model.
- If the acquisition candidate involves supply or collection of personal data, traverse the **Data Protection** model.

Suggested route through models: **Technical Constraints** > **Intellectual Property Rights** > **Data Protection** > **Collection Policy**.

A few important notes:

- Working through the models is likely to require collaboration with other stakeholders. Before you start, identify those who will need to be involved in the acquisition process (e.g. creator, representatives from company/team, departments within organisation) and gather supporting information about acquisition candidate.
- These models will not yield a definitive, 'correct' set of actions for the collecting activity. Instead they are intended to expose (in a clear, structured form) the key issues that should be considered when approaching this collecting (e.g. planning, budgeting, policy making or strategic thinking).

Key to Diagram Notation



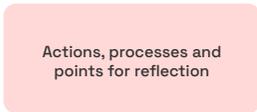
Diagram start (green)



Diagram end (black)



Question and decisions point, followed by two or more branching paths



Actions, processes and points for reflection. In some cases these are tagged with 'flags' to indicate greater than normal risk



Issue may significantly impact the viability of the acquisition

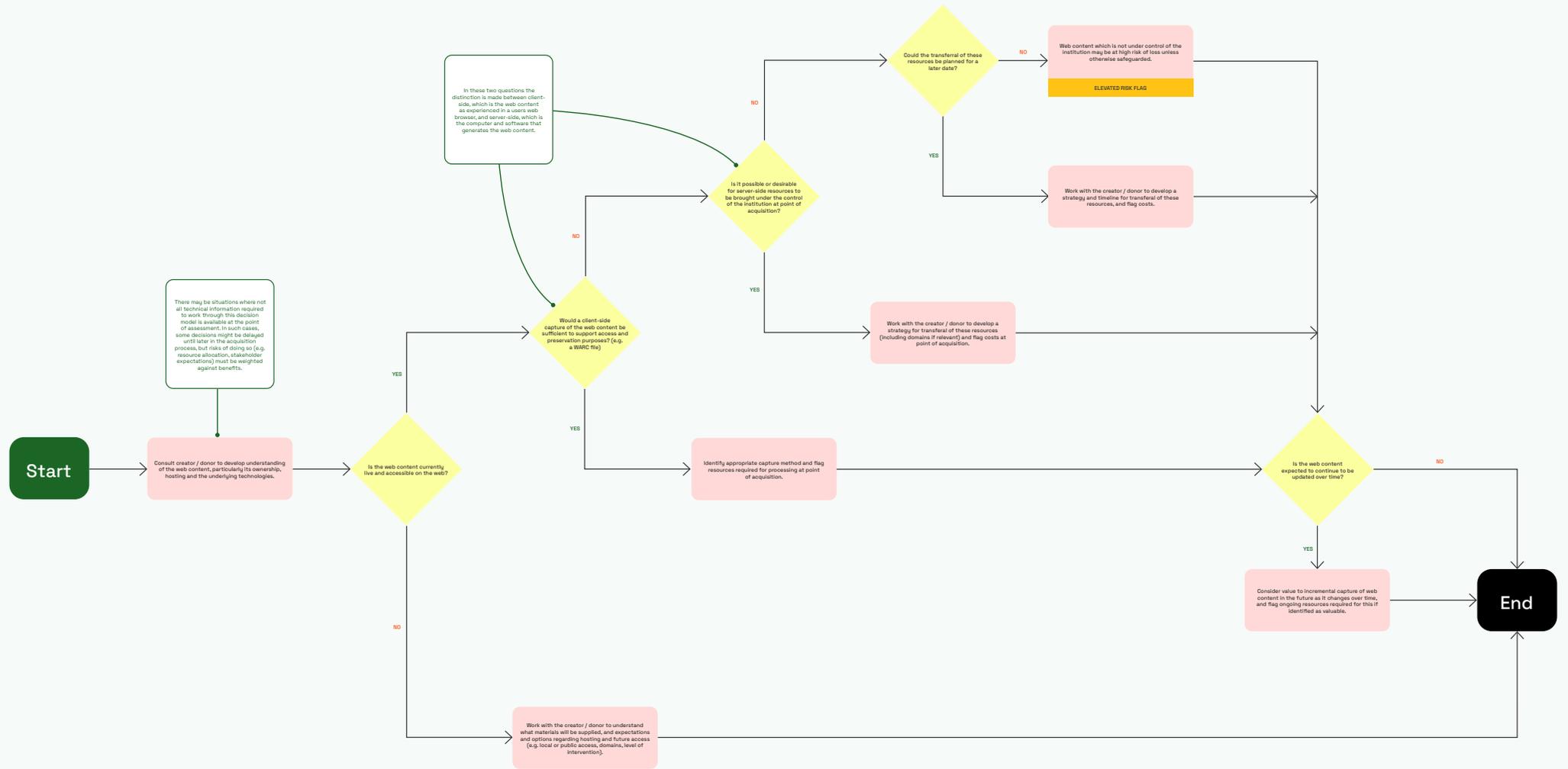


Issue may cause higher than usual risk of loss

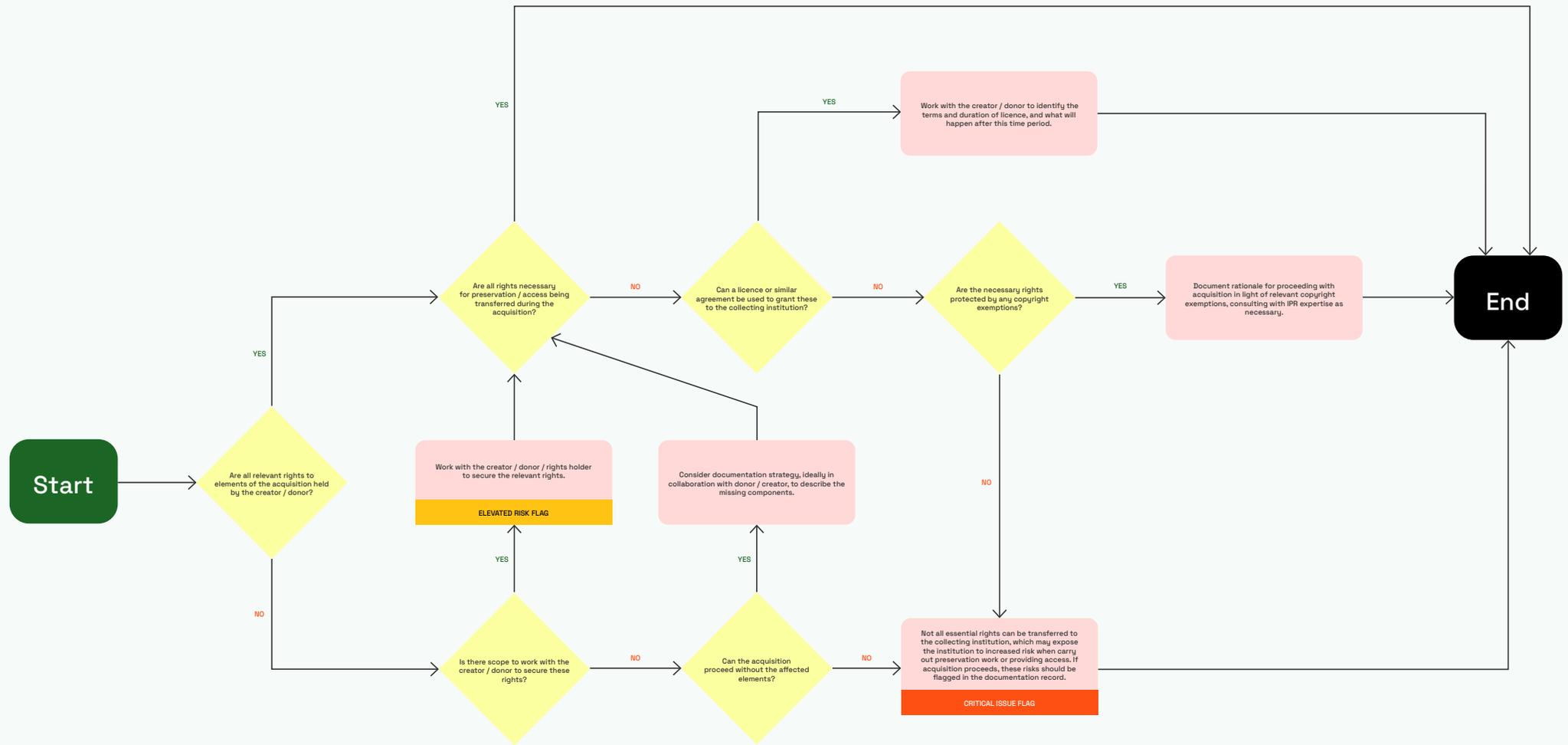


Notes that provide further contextual information on a specific node

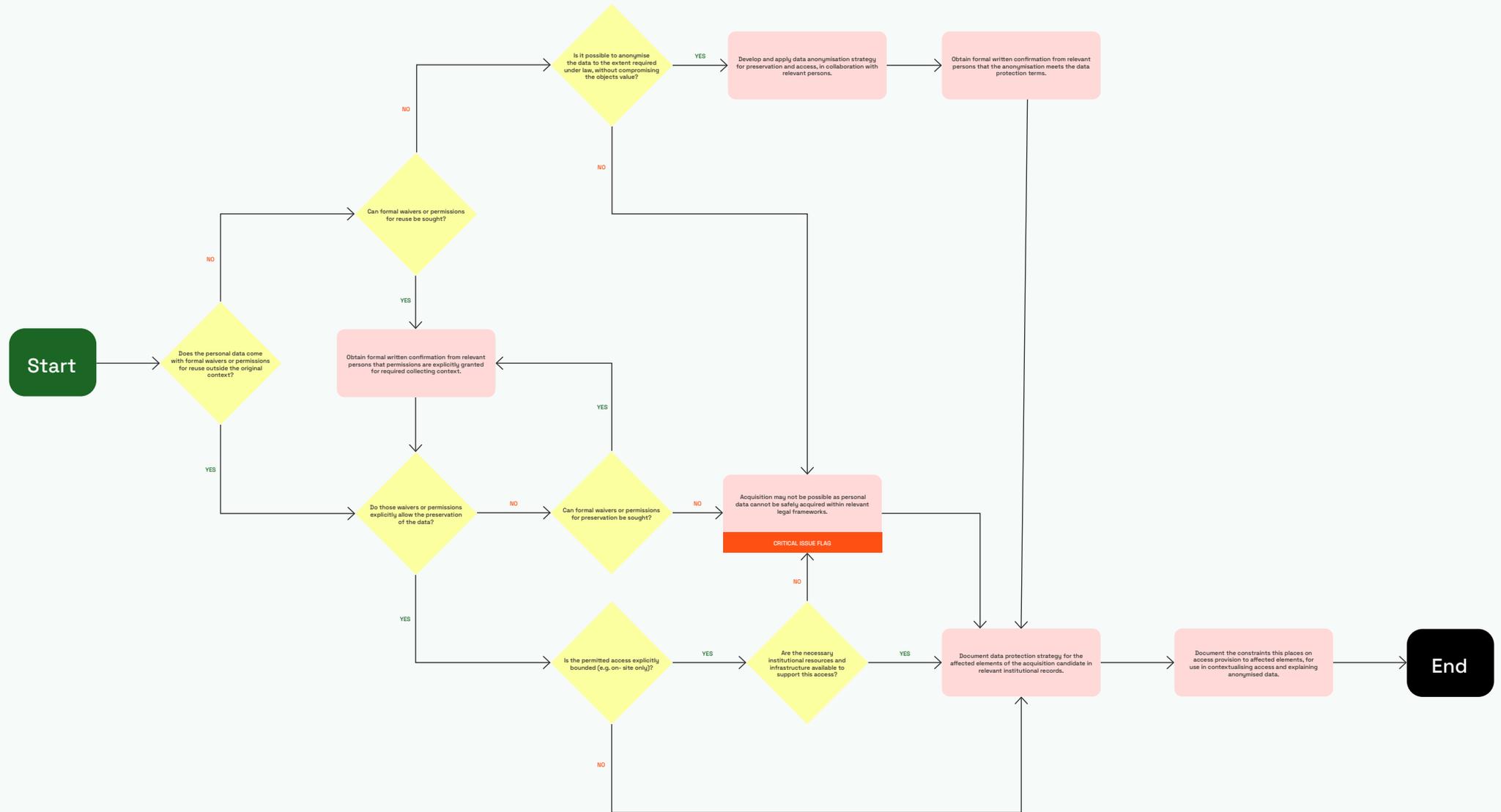
Decision Model: Technical Constraints - Web Content



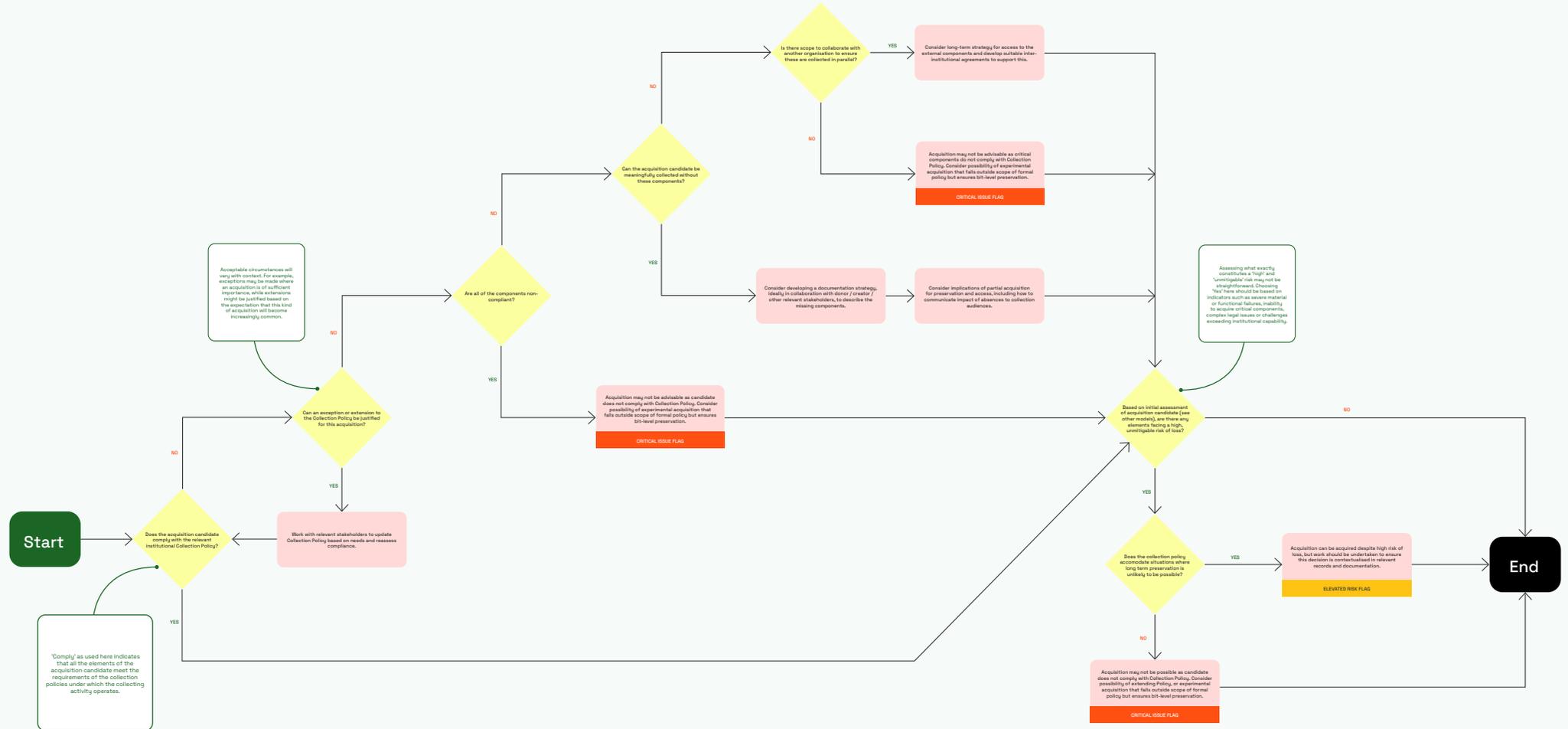
Decision Model: Intellectual Property Rights



Decision Model: Data Protection



Decision Model: Collection Policy Compliance



www.vam.ac.uk