

Modelling Archival Data And What It Can Do For You

Andre Hui

Jamie Lee

The National Archives of Singapore

ABSTRACT

Data modelling is the process of representing how different entities in a domain are related to each other so that both humans and systems can understand how to appropriately treat those entities. The description and arrangement of archival records is a core archival process with significant implications for how archives subsequently manage, and how users of the archives meaningfully access records. Therefore, how archives model their records and additional entities directly impacts how they are described and arranged; in other words, what we say about our records now will determine what and how we know about records in the future.

Archives and individual archivists regularly make decisions that are effectively trade-offs between resource constraints and the efforts required for quality archival description. The Government Linked Archival Data team from the National Archives of Singapore will present their experience with modelling government entities using the International Council on Archives' Records-in-Contexts Ontology (RiC-O) to construct a knowledge graph for a Linked Data proof-of-concept. Afterward, we will make the case that investing upstream effort into modelling entities in the archival domain—and training your archivists to do so—can effectively minimise technical debt and pay off dividends across a variety of downstream processes for your archives, regardless of whether or not your archives decides to use RiC-O to do so.

More specifically, they will share some possible ways of modelling the following entities:

1. Records
2. Authority records (e.g., of entities that created records) for:
 - a. Organisations
 - b. Individuals

Attendees will understand that deciding how to describe records is fundamental and of key importance as the flexibility and expressivity of the descriptive language used determines how much we can subsequently say.