

Global Format Registry Use Case: Digital Object Format Identification

Use Case ID	harvard-1
Description	Determine the representation format of a digital object.
Actors	Registry – a global format registry. Repository – a repository for digital objects.
Assumptions	The repository does not know the format of the object a priori and cannot determine it through some mechanism external to the object itself. The repository has defined matching criteria for the format, based on the identification attributes for the format in the registry.
Preconditions	The format of the object is registered in the registry.
Triggers	Any format-specific repository function applied against a digital object of unknown type.
Primary Scenario	<p>Step 1 The repository requests and receives the identification attributes, potentially including a copy of the authoritative specification, of all registered formats from the registry.</p> <p>Step 2 The repository examines the encoded content stream of the digital object with respect to the identification attributes of each format, looking for a match.</p> <p>Step 3 When the repository finds a match it must decide, on the basis of local policy, whether to accept the first match as the object's format or to continue to look for the best match.</p>
Primary Result	The format of the object is known.
Post-Conditions	
Non-functional requirements	The criteria for format identification matching are a matter of repository policy (e.g., file extension, magic number, internal syntactic signature, etc.), not registry policy. The registry merely makes available to the repository relevant identification information about specific formats that the repository can use in formulating its local policy. A repository may have different criteria for different formats.
Notes	The actions enumerated in the scenario can be encapsulated into tools for automated processing.
Issues	If the format of the object is not registered in the registry, the object will remain untyped.