Conversational CoreTrustSeal

Disclaimer

This working paper is not endorsed by the CoreTrustSeal Board, or indeed anyone else. It presents some thoughts about the CoreTrustSeal Requirements that don't quite fit into formal documents, project deliverables, blogs or slides. It presents some key concepts and provides some broad talking points around the CoreTrustSeal and Trustworthy Digital Repositories' data and metadata services in fewer than 3000 words. It's being shared because transparency and feedback are always worthwhile.

In cases of disagreement with this document the CoreTrustSeal Requirements[1], Extended Guidance[2] and Glossary[3] always take precedence.

Applicants for the CoreTrustSeal

If you take responsibility for digital objects and the actions that preserve them, then you can apply for the CoreTrustSeal. Others can help, and their role(s) must be acknowledged, but only the applicant that takes responsibility is certified as a Trustworthy[4] Digital[5] Repository (TDR).

Trustworthy Digital Repository (TDR)

We put things in repositories because we hope they will be stored/managed/cared for/curated in a reliable environment. In a trustworthy digital repository, we expect digital objects to be actively preserved so that they remain accessible, understandable and usable despite changes to the people, the processes or the technologies (the infrastructures) that store, curate and consume them.

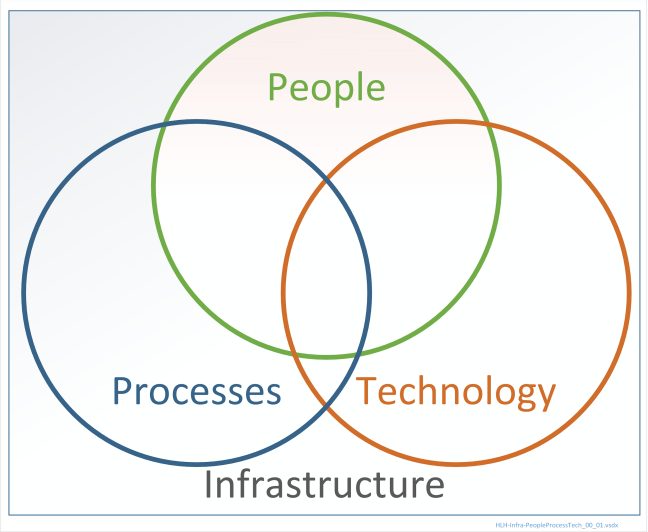

Diagram: Infrastructure. People, Processes, Technology

[1] https://doi.org/10.5281/zenodo.7051012
[2] https://doi.org/10.5281/zenodo.7051096
[3] https://doi.org/10.5281/zenodo.7051125
[4] Being trusted is not quite the same thing as demonstrating that you are trustworthy, but the terms are often used interchangeably.
[5] Digital objects can contain structured metadata and other documentation as well as data. The use of digital versus data is a matter of context or, sometimes, simple preference.

Audiences

The initial audience for a self-assessment against the CoreTrustSeal Requirements is internal to the applicant. The submission to the CoreTrustSeal reviewers and Board is a confidential process. Once certified, CoreTrustSeal assessments are public and can be read by anyone including peers, partners, depositors, users and funders.

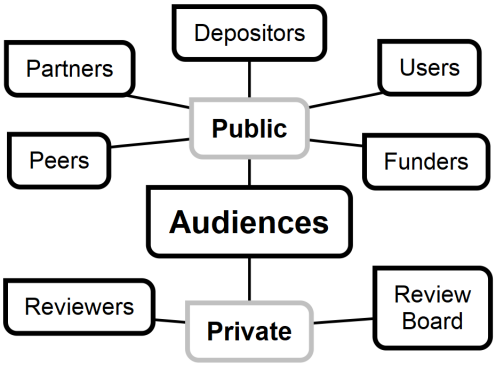

Diagram: Audiences

What Do You Do? Who Do You Serve?

The OAIS reference model[6] defines the mandatory repository responsibilities.

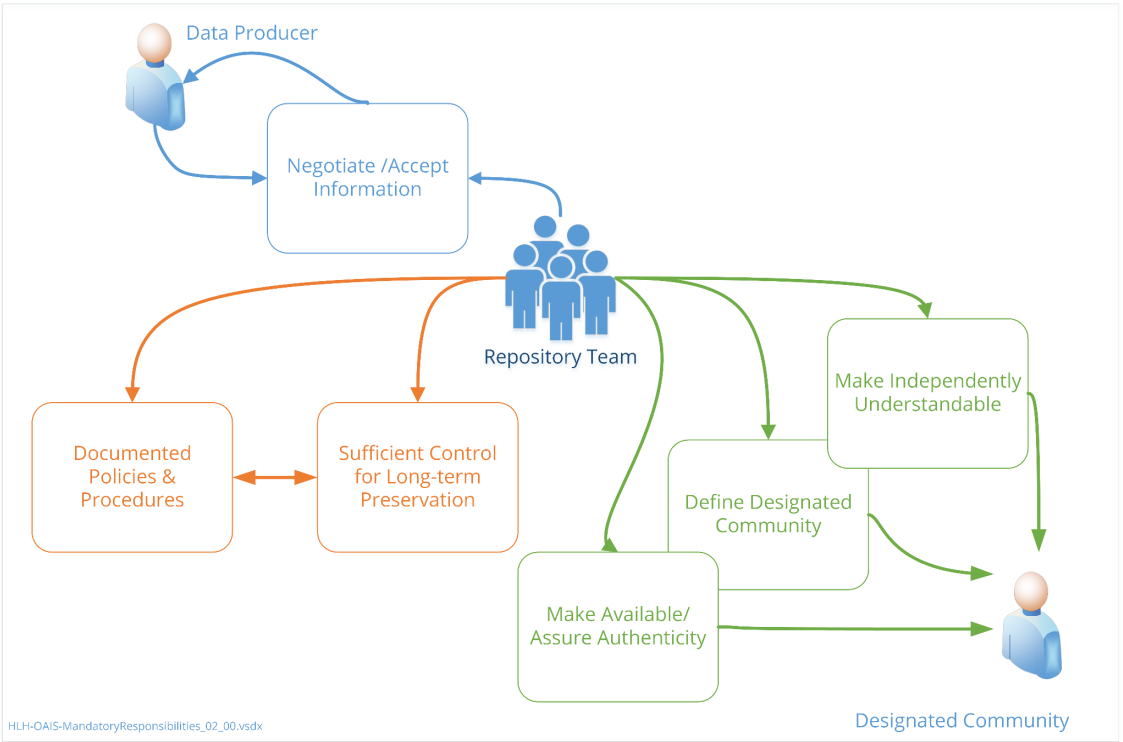

Diagram: a simplified overview of mandatory OAIS responsibilities

[6] https://public.ccsds.org/pubs/650x0m2.pdf

The CoreTrustSeal applies the OAIS definition of the designated community in their Glossary[7]. An applicant may serve multiple, complex communities of users but the CoreTrustSeal asks you to define one or more communities whose knowledge base, processes and technologies you understand. You should demonstrate that you know what they want, deliver what they need, and change as necessary over time.

Self-Assessments and Evidence

The CoreTrustSeal can be used as an internal self-assessment and communications tool without applying for certification. The self-assessment statements provided to the CoreTrustSeal are claims that must be supported by evidence. That evidence reflects the information that needs to be created and managed when running a quality, efficient, sustainable data service.

Policies, procedures and other business information is documented so that people know what we do, so that we can do it consistently, so that we manage change, and so that someone else can do it if we can't do it any more for some reason. We can also use this information as evidence for the CoreTrustSeal.

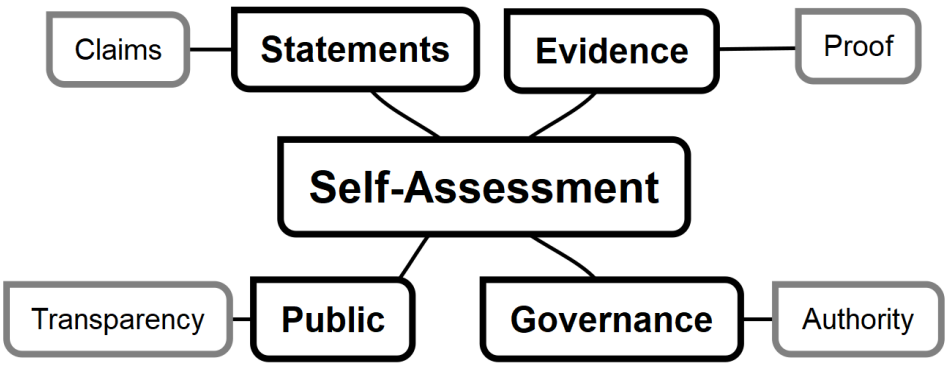

Diagram: Self-Assessment & Evidence

Clear authority for the creation and management of evidence demonstrates good governance. Public evidence demonstrates transparency.

A small organisation might rely on a limited number of very expert staff. Losing a member of staff is a big risk, so things should be written down. A larger organisation might have multiple staff doing the same job. Inconsistent work from different staff members is a big risk, so things should be written down.

Documentation reduces the risks to the applicant and to the digital objects they care for. Once documented a periodic review and continuous improvement process can begin within the repository.

[7] https://doi.org/10.5281/zenodo.7051125

CoreTrustSeal Requirements

R0. CoreTrustSeal Context

This information supports the interpretation of self-assessment and evidence for all 16 requirements. You will be asked to describe the repository, its type and its designated community (see above) and the levels of curation performed.
The applicant takes responsibility for meeting all the CoreTrustSeal Requirements, but they may be assisted by others that they cooperate with including host organisations, partners and other third parties. Make it clear what these organisations do for you and how you manage your relationship with them.

Organisational Infrastructure

Mission & Scope (R01)

- A TDR goes beyond the bits
- Access and Preservation are explicitly a part of your mission

Communicate (with the highest level of approval you can get) that data and metadata are assessed, curated, and made accessible and usable for the long term. Make it clear that you actively preserve digital objects for your community.

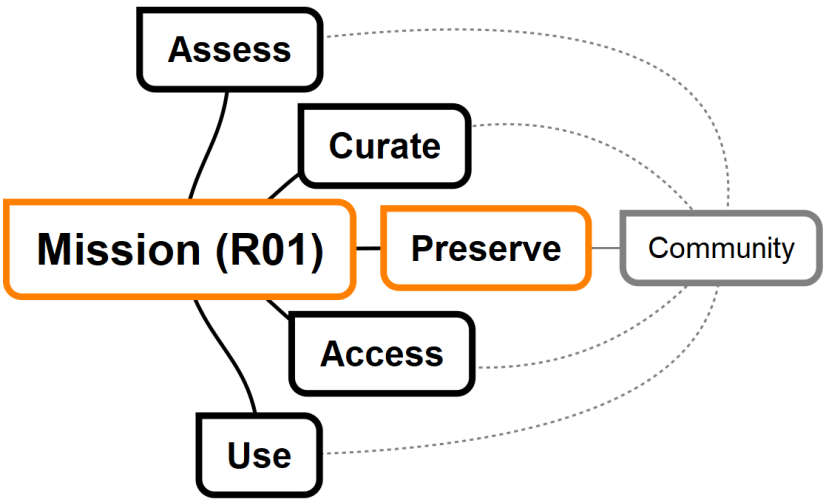

Diagram: Mission (R01)

Rights Management (R02)

- Listing licences by type and purpose
- Understanding objects' rights and rights holders
- Permission, prohibitions, obligations, constraints & duties

Rights management can be complex but being able to list any licences in place is an important first step. We should be clear on whether these are open and standard licences or if they are locally defined, and whether they refer to deposit, curation, preservation, access or (re)use of digital objects.

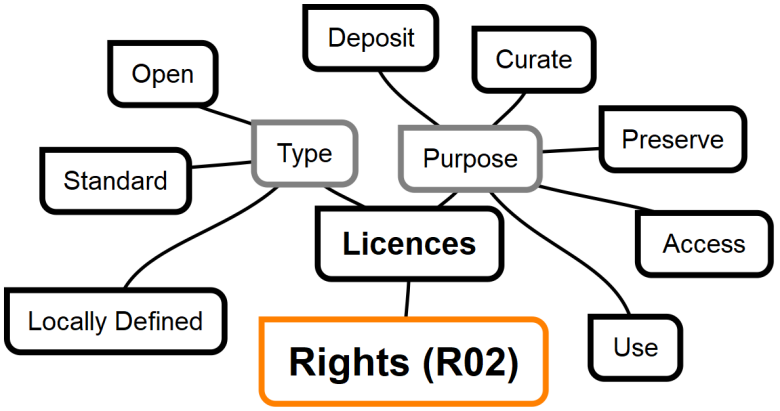

Diagram: Rights (R02). Listing Licences

More complex scenarios, including machine-actionability, depend on a more detailed understanding of rights (permissions, prohibitions, duties) and how they relate to actors (depositors, repository, users) and digital objects' data and/or metadata. Any measures in place to monitor compliance or address non-compliance need to be clear.

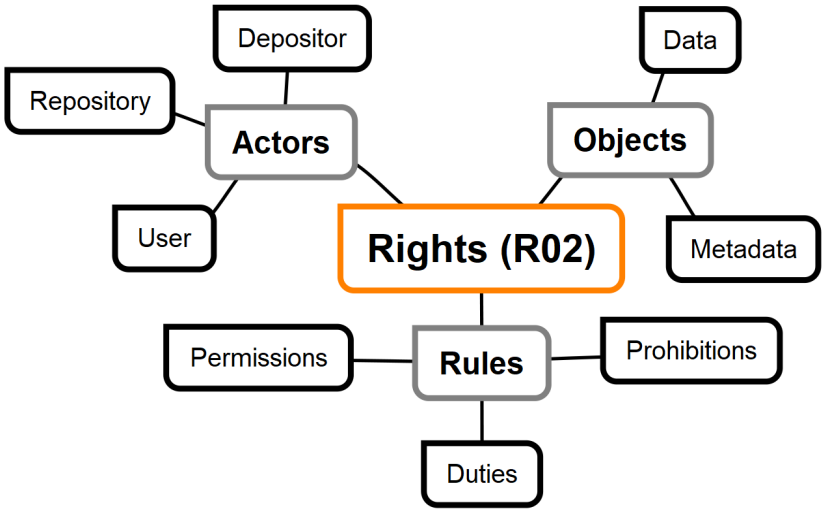

Diagram: Rights Management (R02)

Continuity of Service (R03)

- The organisation persists so the digital objects persist
- Business continuity, disaster recovery and succession

What could affect the services we deliver? What do we do if something reduces our services or stops them completely? How do we get back on track? In the worst-case scenario, could someone else offer the same level of service for the digital objects?

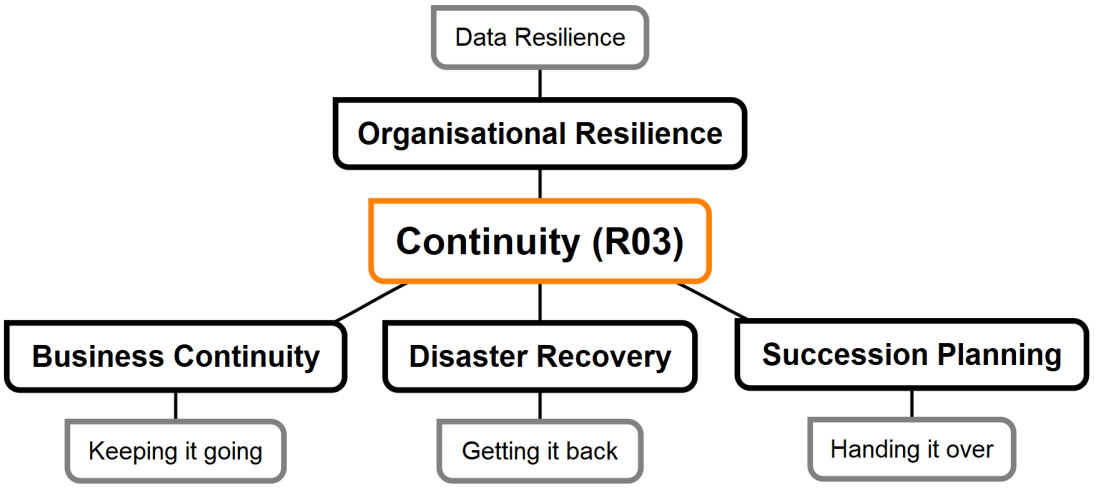

Diagram: Continuity of Service (R03)

Organisational resilience supports data resilience. Business continuity keeps things going in the face of adversity. Disaster recovery is about the processes and the time it takes to recover services back when they fail. Succession planning is about being able to transition your digital objects and services elsewhere.

Some organisations can tell you where their digital objects will go if they change their mission or cease to exist. A few can tell you what level of service they could hand over. The CoreTrustSeal is looking for evidence that you have considered the worst-case scenario. However, the CoreTrustSeal does not require that you provide a blueprint for replacing all of your complex, expert-driven data services.

Legal & Ethical (R04)

- Ethical research and the protection of sensitive data
- Understanding the legislative and policy context

Even when there are no ethics policies or standards review boards in place, any creation or collection of data can have ethical implications. Are you aware of the laws and other regulations about how depositors, repositories and users should handle data and metadata? If data is sensitive (e.g. about identifiable humans), what additional steps do you take?

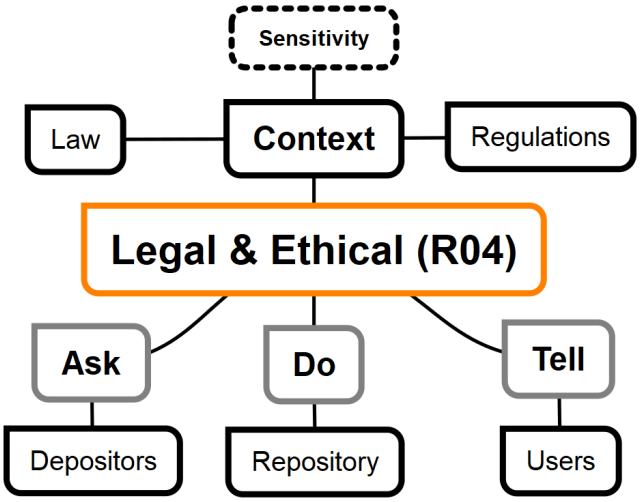

Diagram: Legal & Ethical (R04)

We need to demonstrate that we understand, apply and communicate the legal and ethical rules in place. What do you ask depositors about their legal and ethical approach? How do you act as a repository? What do you tell users about how they should act?

Governance & Resources (R05)

- Governance, structure and resources

How is the organisation that looks after the digital objects structured? What sections, departments, roles and hierarchies does it have? How are decisions made? How does it ensure it has the resources to deliver its services?

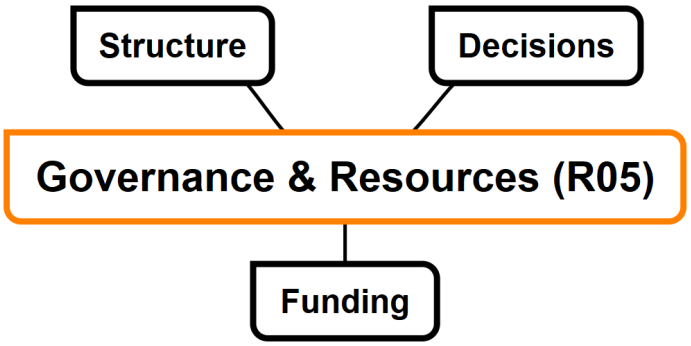

Diagram: Governance & Resources (R05)

Repository organisation and structure can be complex and varied. Describe your organisation structure, link to a diagram on the web, note relevant bodies that make decisions and briefly describe your funding model. Ideally, this information is public, which demonstrates transparency to all of your stakeholders.

Expertise & Guidance (R06)

- No organisation is an island.
- Communities of practice

No organisation can have all the skills it needs within their staff, especially if they are small or have a very wide-ranging mission. Describe the areas of knowledge you depend on and explain how you ensure you recruit and maintain internal expertise as well as engage with and participate in wider groups of experts. It's valuable to be a follower and a leader; often both in different areas.

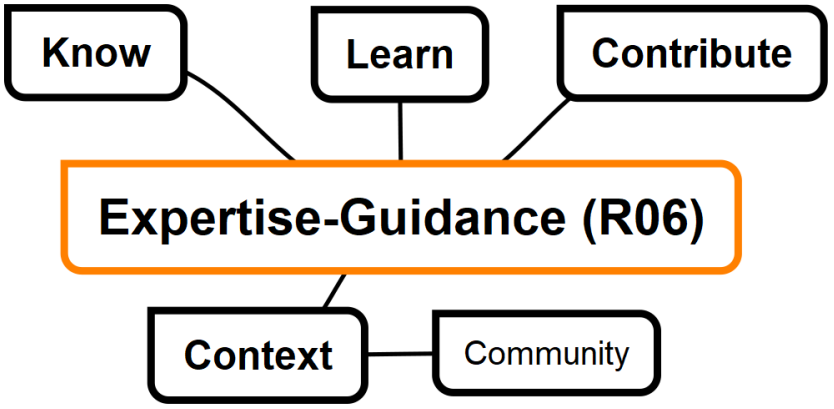

Diagram: Expertise & Guidance (R06)

Digital Object Management

Provenance & Authenticity (R07)

- Managing and communicating planned changes

How do you and your users know that a digital object is what it claims to be, and how it has been generated and changed over time? If you're changing something, or a copy of something, is it clear why? How do you record and communicate those changes? When and why do you create new versions?

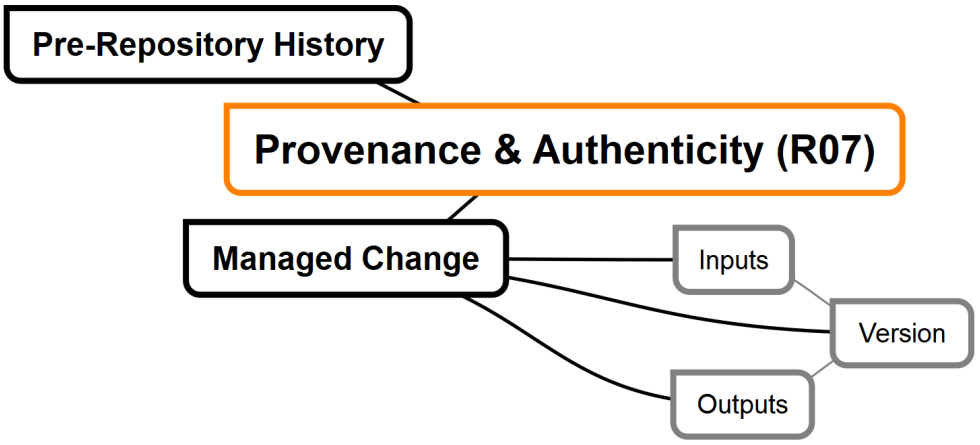

Diagram: Provenance & Authenticity (R07)

A clear chain of provenance defines the inputs and outputs for each action on an object and explains why actions were taken.

Deposit & Appraisal (R08)

- The gatekeeper to the repository
- Objects offered for deposit are evaluated
- Curation actions are defined

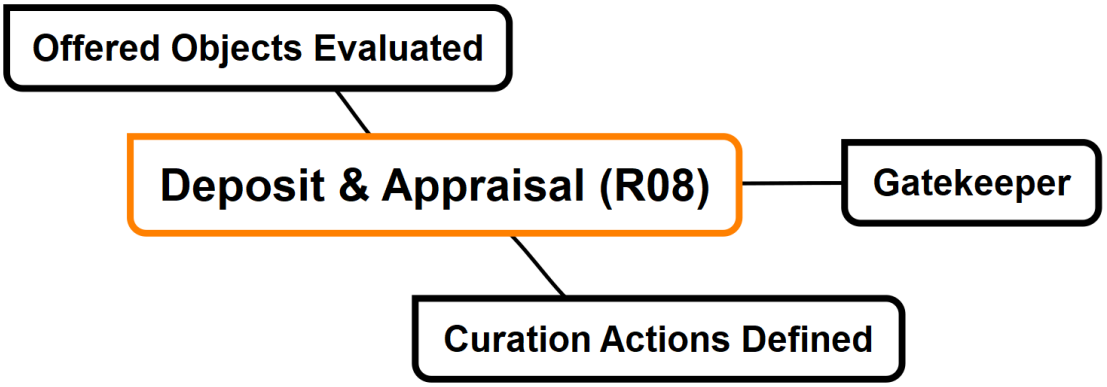

Diagram: Deposit & Appraisal (R08)

What rules do you use to decide what you will and will not accept to look after? Is it clear to the depositor what level of curation and preservation you will offer? How do you decide what steps you will take so that digital objects remain usable?

Preservation Plan (R09)

- Implement and communicate a clear level of preservation.
- If not preserved here, then where?

Preservation goes beyond basic curation. You must understand the needs of your designated community in accessing, understanding and using the digital objects. You also need to demonstrate that you can adjust to changes in that community and those needs over time.

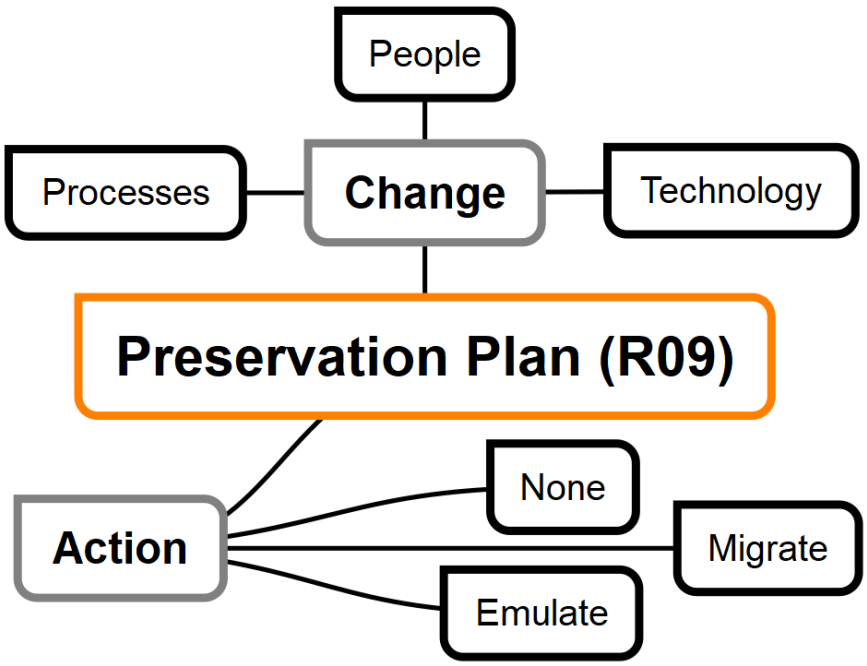

Diagram: Preservation Plan (R10)

If the needs of the community of people you serve, or their processes or technologies change then this may provide a reason to take a preservation action on the data, metadata or documentation of a digital object. That action may be to migrate to a new format, or to provide an emulation solution. You may decide that no action is required; the important thing is that your action or inaction is an informed decision.

Preservation doesn't have to mean 'forever'. However, it should be clear to your stakeholders what circumstances would cause you to change the level of preservation you apply to a digital object.

If you don't have some, or all, the responsibility for preservation it should be clear who does. If no one is taking responsibility for preservation, it is vital that this is clear to the stakeholders.

Are you taking the same level of care with every digital object you look after? If not, it should be clear what levels of curation and preservation you apply to which objects.

Quality Assurance (R10)

- Quality compliance expectations
- Assessing quality
- Curating and preserving for quality
- Communicating quality

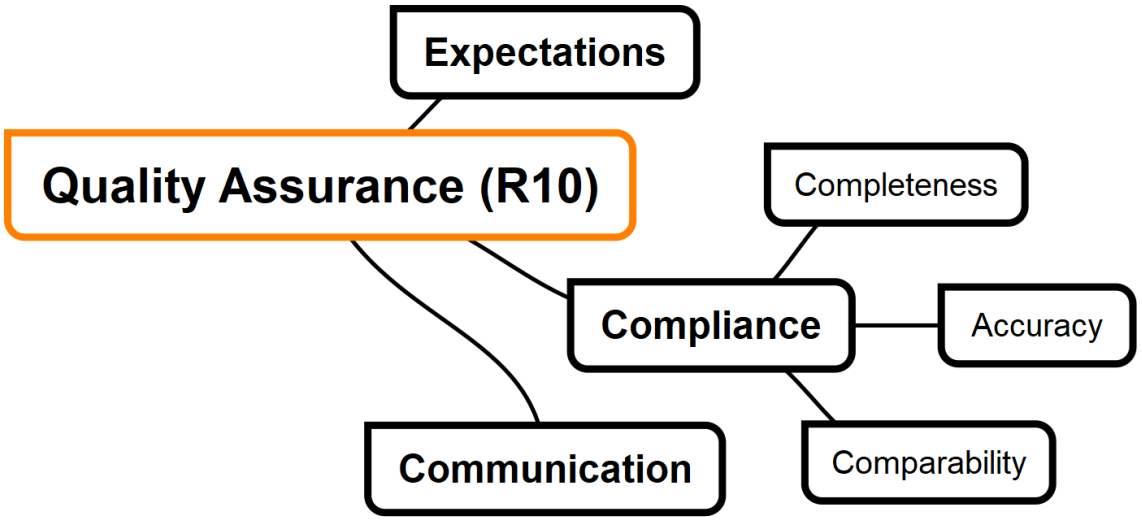

Diagram: Quality Assurance (R11)

Repositories are important in helping define and apply community quality expectations. But, in the end the scientific value aspects of quality are for the user to evaluate and decide.

The repository may make initial quality assessments at the point of appraisal (R08).

This requirement is about demonstrating that the repository understands community expectations and takes the necessary steps to ensure that the digital objects comply. This involves ensuring that digital objects are comparable to others and are complete and accurate based on some standard(s).

The repository communicates the measures taken, and the level of quality provided at the point of reuse (R13).

Workflows (R11)

- Defined curation actions
- Clear responsibilities

Taking actions in a clear, consistent way supports both quality and efficiency.

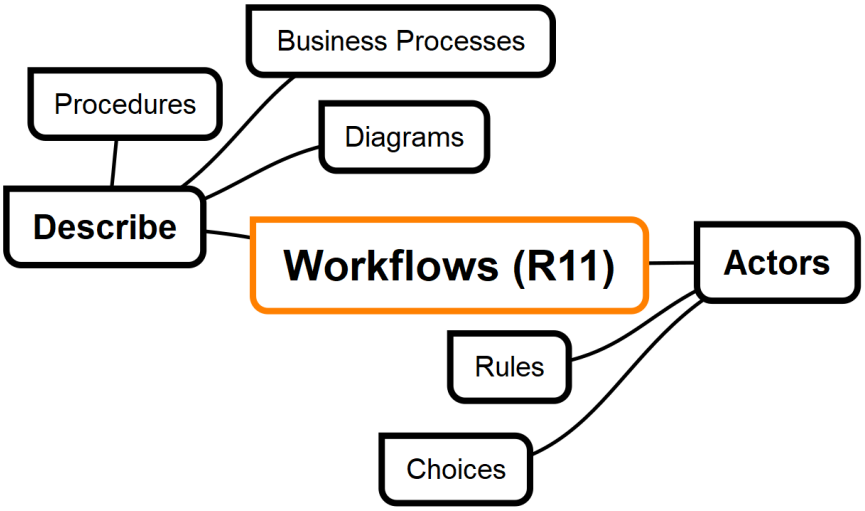

Diagram: Workflows (R12)

How are rules and choices defined? Do you describe the different functions and actions you undertake through business process descriptions, diagrams or standard operating procedures? How do you ensure that the appropriate actors take the right actions on the correct digital objects?

Discovery & Identification (R12)

- Providing the context for the discovery of specific objects
- Unique, persistent, resolvable identification

We need enough descriptive information to reach available digital objects and to give credit by acknowledging their use (citation). Each unique object must be associated with its own unique identifier that persists over time and helps us to reach the object (resolution).

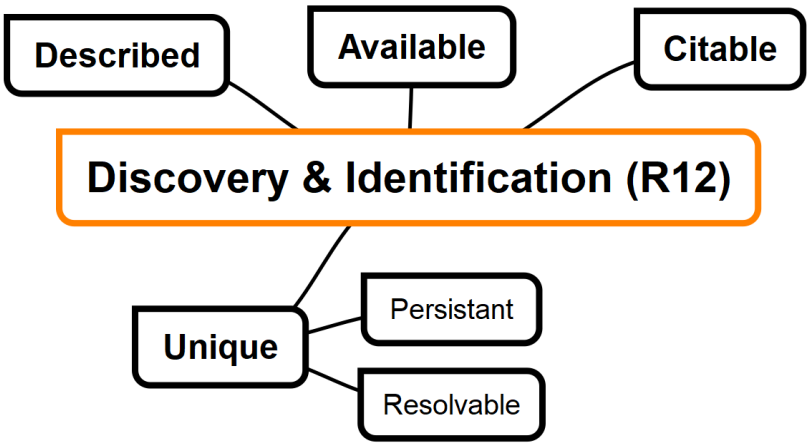

Diagram: Discovery & Identification (R12)

How do you assign identifiers to your digital objects and expose them and their metadata to systems that support searching? How do you ensure that users can find the search system itself? How do you support citing of data and metadata to ensure future provenance and the sharing of credit?

ReUse (R13)

- Integrating digital objects for re-use
- Understandable, actionable data and metadata
- Delivering impact through data

To support re-use we must design digital objects so that communities can understand them and take action on them by integrating them into future work. This may include re-running, repeating, reproducing or replicating previous work, or doing something entirely new. Through these actions, reuse delivers impact.

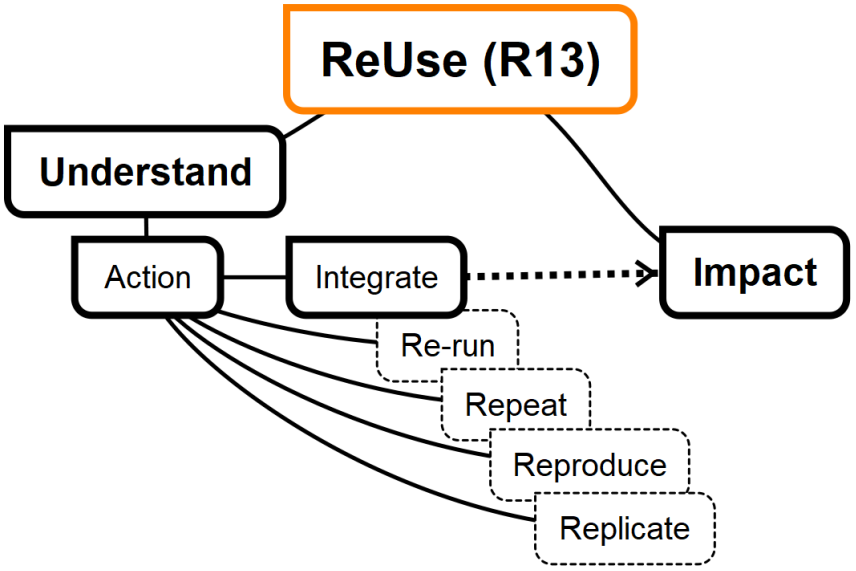

Diagram: ReUse (R14)

Providing information relevant to the knowledge and technology base of your intended users is critical if the re-use of digital objects is to be efficient, reliable and deliver impact. What steps do you take to ensure that users can re-use the digital objects?

Technology & Security

Storage & Integrity (R14)

- Location, location, location
- Protecting the bits
- Avoiding unintended change

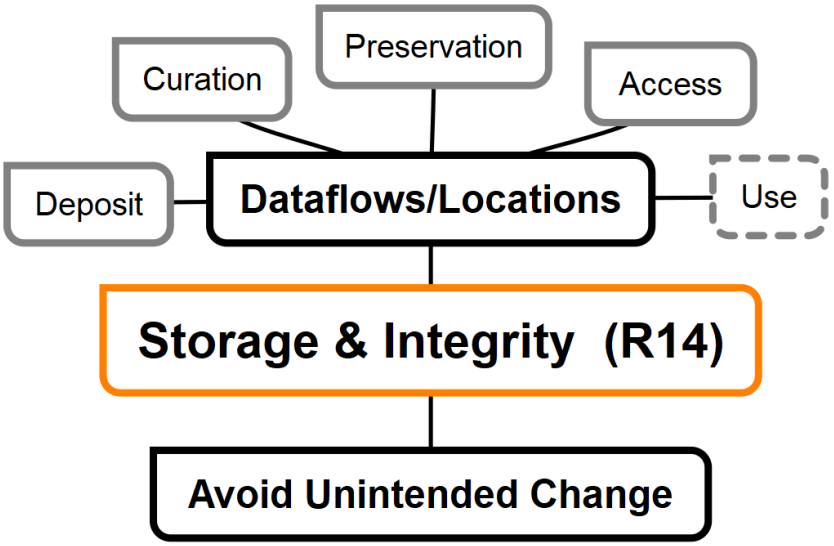

Diagram: Storage & Integrity (R14)

From the moment that you have custody of a digital object, you have responsibility for it. Can you describe where the data flows and how is it backed up and copied at each stage of processing and in each storage location? This includes the locations for deposit, curation, long-term preservation and access. It also includes any locations where repositories control data use (e.g. secure remote access). How much could be lost if something goes wrong? What integrity measures do you take to make sure all those copies stay the same?

Technical Infrastructure (R15)

- The tools of the trade
- Meeting the needs of repositories and users
- Planning for the future
- Recovering from the unexpected

How do you decide and provide the tools needed to meet your users' needs? How do you govern that technical system over time to manage the expected and the unexpected?

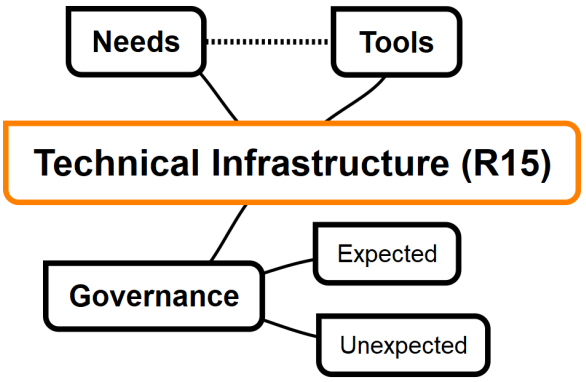

Diagram: Technology (R15)

Technology provides the environments and tools for managing, providing access to, and using digital objects. The technologies, standards and processes in place must meet the needs of stakeholders. There should be measures in place to ensure that the technical infrastructure remains fit for purpose over time and for responding to disasters or other business continuity issues.

Security (R16)

- Can you protect what you've been trusted with?

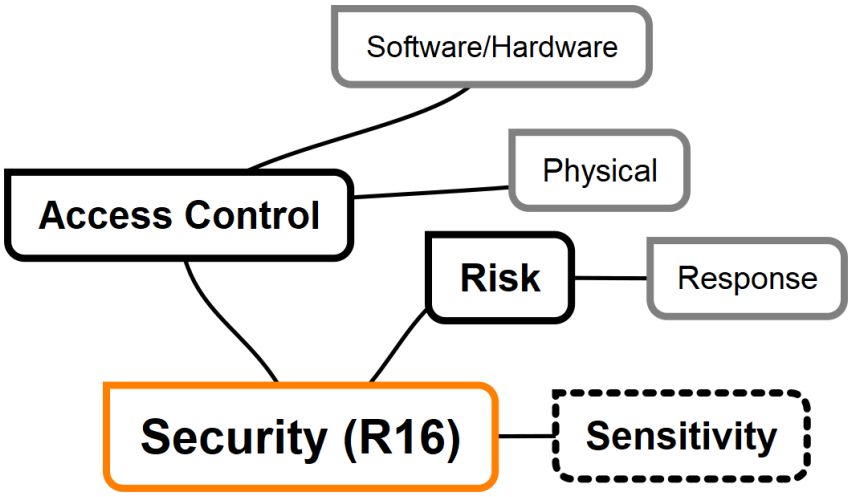

Diagram: Security (R16)

All digital objects need some degree of physical and technical (software/hardware) protection. The expectations increase when the data and/or metadata are sensitive for some reason. Access control and rights management measures help to avoid malicious actions and enable permitted actions. Risks should be evaluated based on likelihood and impact, and mitigated where possible. Measures should be in place to respond to any security incidents.

Version History

v01.00 Aligned with the CoreTrustSeal Requirements 2020-2022
v02.00 Aligned with the CoreTrustSeal Requirements 2023-2025