



Delft University of Technology

TU Delft Next Generation Repository Policy

Guicherit, Armand; Wahls, N.L.; Dunning, Alastair

Publication date

2024

Citation (APA)

Guicherit, A., Wahls, N. L., & Dunning, A. (2024). *TU Delft Next Generation Repository Policy*.

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

*This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.*



TU Delft Next Generation Repository Policy

March 2024

Authors: Armand Guicherit, Naomi Wahls, Alastair Dunning

Document history

Version	Date	Author	Change
0.1.	20 Nov 2023	Naomi Wahls	First draft made
0.2.	15 Dec 2023	Naomi Wahls	Added SLA Requirements, adjusted file format and content types based on questions. Added more questions.
0.3	14 Feb 2024	Armand & Naomi	Moved content to other policy, procedure, protocol documents, the about page, and Confluence Knowledge Base.
0.4	March 2024	Alastair	Comments added, changes made, switch to Google Docs
0.5	April 9 2024	Naomi and Armand	Changes made
0.6	April 10 2024	Naomi and Armand	Ready for MT Review
0.7	April 26 2024	Karin	Reviewed and gave input
0.8	April 30 2024	Armand & Naomi	Final Draft
0.9	May 28 2024	MT Library	Approved with minor changes
1.0	May 29 2024	Naomi	Minor changes

1. Introduction

The TU Delft Next Generation Repository policy outlines the objective, scope and intended usage of the institutional repository from 1 July 2024.

The Next Generation Repository will replace the repository that has been in place since 2007. From July 2024, the version 1 will be released with a new design and a new back-end technology infrastructure, but largely keep the functionality of the existing TU Delft Institutional Repository. Most existing content will be added to the new repository or temporarily kept in a dark archive until there is agreement on how best to deliver this content. From early 2025, new functionality will be introduced in version 2. Most significantly for version 2, TU Delft staff will be able to upload their

own content to the repository, and each domain (e.g. Research, Education) will have separate design views.

2. Objectives for TU Delft Next Generation Repository

The TU Delft Next Generation Repository will:

- Provide access to TU Delft outputs and collections so that these are findable and accessible to a wide audience.
- Provide reliable long-term storage of TU Delft outputs and collections and preserve its institutional memory and heritage.
- Foster collaboration and offer inspiration to researchers, staff, students and the public by providing relevant functionality via a range of accessible, intuitive interfaces.
- Enhance the TU Delft's reputation and visibility in the academic community, and beyond.

The repository aligns with the vision presented in the COAR (Confederation of Open Access Repositories) Next Generation Repository report.¹

2.1. Intended usage

The TU Delft Next Generation Repository's intended usage is to:

- Store TU Delft collections, outputs, digital assets, and related metadata securely and for the long term, including those that have not previously made available online
- To provide a variety of end-user interfaces meant to search and browse through, and to showcase and contextualise content.
- To provide a technical content platform, consisting of a range of flexible, well-documented APIs that are accessible to 3rd parties, to build different end-user interfaces based on output and collections stored in the repository.
- Provide an upload mechanism for TU Delft metadata and digital assets that can be used by TU Delft staff
- Support the validation and approval of uploaded content
- To provide a mechanism to share archive copies of the content with services such as the eDepot of the National Library of the Netherlands.
- To ensure indexation of TU Delft outputs by other relevant web services; to enhance discoverability, citation, and interoperability with harvesters, repositories and larger (global) infrastructures by employing standardised metadata formats such as Dublin Core, MODS and CERIF.

¹ Boliini, A., Knoth, P., Perakakis, P., Rodrigues, E., Shearer, K., Van de Sompel, & Walk, P. (2017). *Next Generation Repositories: Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group (2017) (Version 2)*. Zenodo. <https://doi.org/10.5281/zenodo.8077381>

- To support – in connection with the Current Research Information System (currently Pure) – the goals set out in [TU Delft's Rewards and Recognitions Programme](#).

2.2. Scope

Only outputs connected to the TU Delft are stored in the repository. By connected we mean content authored by current or former TU Delft researchers, students or staff members, or content that TU Delft has been made the custodian of.² The library facilitates information management of this content to digitally re-shelving as needed, to enable showcasing and grouping content in collections, and to archive when appropriate.

These can be imported / harvested from various resources or can be (from 2025 onwards) manually deposited in the repository. The Library is responsible for collecting the TU Delft research outputs from relevant sources, striving to make this as complete as possible. Research output metadata is validated and enriched by the Library and other core users across campus assist with validation of content in other categories beyond research output.

Not all content managed by the library ends up in the Next Generation Repository. For example, the following collections have different channels:

- Research Data and Research software collections go to the 4TU.ResearchData Repository.
- Data Management Plans go to DMP Online.

3. Domains

The TU Delft Repository stores and makes available content in four broad back-end domains - Research, Education, Heritage and Business. Each domain has its own specific collections within it (e.g. the collection of PhD Theses is part of the Research Collections). For the present document, only collections relevant to Research are considered in depth. This will expand as the ability of the institutional repository to deliver other domains increases.

3.1. Research Domain

The goal of the research domain is to create an accessible archive of past and present research outputs from TU Delft. To achieve this, the repository supports a broad spectrum of research outputs and file formats. New content categories/types to be added will be discussed in 2024.

Some of the existing types are:

² *In the current set up, new content is added via the TU Delft Current Research Information System (CRIS). As a result, the scope of what can be deposited is restricted by the content types available in the CRIS and the use of these content types as agreed with the faculties. With the Next Generation Repository, the scope of the content will be broadened.*

- Contributions to journals
- Book chapters
- Conference contributions
- Research reports
- Doctoral Theses
- Non-textual documents (images, designs, maps)

Full details are available in the appendix under 8.2.

3.2. Education Domain

Discussion over education collections and how they are to be made accessible via the Repository will happen in 2024. Full details will be available in the appendix under 8.2.

Connections to content from New Media Centre (NMC) will be discussed in 2024; some content from NMC falls under Research. Content from SURFShareKit and Open CourseWare will be discussed in 2024/2025.

Masters' Theses and Bachelor Student Reports are content types under a new collection called "Student Works". Lecture notes originally in the Research Domain are now considered Education.

3.3. Heritage Domain

Discussion over other academic and cultural heritage collections and how they are to be made accessible via the Repository will take place in 2024.

3.4. Business Domain

Other types of content developed by the university and related to its management can also be included in the repository. This can include faculty reports, campus policies, or other types of archived documents pertaining to the organisation.

Discussion over other organisational and business collections and how they are to be made accessible via the Repository will begin in the second half of 2024.

4. Depositing content

From early 2025 onwards, TU Delft authors will have the ability to deposit their own material in the repository at which time the Repository will replace the upload mechanism of Pure and the Repository will push metadata to Pure. See appendix 8.1 for details.

Those depositing outputs ("Depositors") must have the authority or permission to do so.

Faculty submitted batch upload requests must be in a digital format and adhere to the scope and intended usage.

4.1. License and Access Conditions

To determine how third parties can use outputs, Depositors can choose when and how an output is made available via access conditions and licences. Barring exceptional circumstances, the metadata about TUD outputs will be made available with a Creative Commons CC0 waiver, thus the files uploaded and metadata may have differing licence and access conditions.

4.1.1. File Access Conditions

Open Access – the output is made openly available for reuse according to the licence conditions chosen (for which see 5.1.2)

Embargoed Access – the output is only made available according to a specific open access licence after a specific time period named by the Depositor (maximum 2 years). The metadata, however, is made openly available (as CC0), unless specifically requested otherwise by the Depositor.

Closed Access/Request for Access - the output within the internal library processing space is only made available to specific users after permission of the Depositor. For the time being, the library staff will act as liaison between third-party user and Depositor, and forward the requested assets. Metadata (excluding any abstract) is still made available via CC0 waiver.

No access - the output, nor metadata, is not available to third-party users

4.1.2. Licence Conditions

For Depositors who choose Open Access, they can also choose from a range of Creative Commons licences to determine what type of reuse they will permit. The licences available are all the {{put on repository webpage}} The default licence is CC-BY.

Precise requirements for Depositors are in the Deposit Agreement {{add link when completed}}

4.2. Recommended file formats

Depositors are strongly encouraged to submit their outputs in open formats, to facilitate their long-term preservation. However, it is not always possible to export all outputs into proprietary open formats; therefore, outputs in file formats will also be accepted by the Repository.

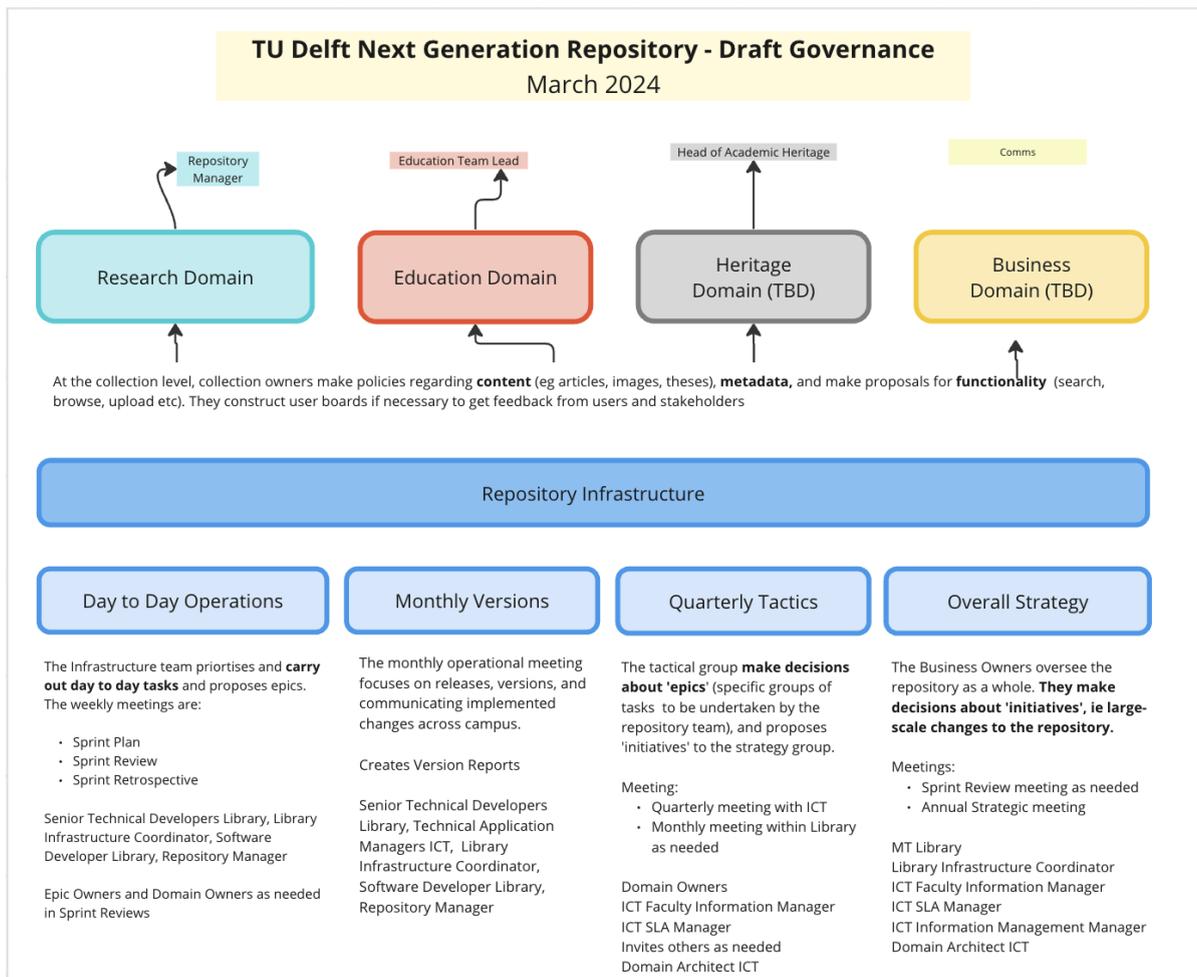
TU Delft repository currently follows the [recommended file formats](#). This will be reviewed in 2025.

5. Governance, roles and responsibilities

The TUD Repository has 1 infrastructure to support the 4 domains; this shared infrastructure allows TU Delft Library staff to manage the metadata and digital assets across all four domains. Each domain can have its own design and features as needed. Each Domain is assigned a Domain Owner.

The TUD Repository is governed at 3 levels: strategic, tactical, and operational. Operational is broken down into day-to-day meetings and the monthly version release meeting. The governance is also

described in the Service Level Agreement (SLA) with ICT and thus must be updated when adjustments are made in the SLA.



5.1. Day to Day Operations

The Infrastructure team within the Library is responsible for day-to-day operations of the TU Delft Repository. The team manages the repository service and technical aspects of the underpinning software and its integration with other University and Library systems.

In addition to service ownership, the team is also responsible for technical and software development, maintenance, and platform upgrades of the repository and associated integrations, together with ICT. ICT provides help with technical maintenance and will be needed together with the Library Infrastructure team for the expanded needs from the Education and Heritage Domains.

The Library Infrastructure team is responsible for:

- Technical & software development, maintenance and platform upgrades of TUD Repository and associated integrations
- User Support Tier 2: Provision of technical user support such as new feature requests, change requests, and bug reporting
- Management of the technical infrastructure for TUD Repository

- Monitoring system performance, and coordinating updates to the system
- Oversees the ingest of digital assets and metadata into systems and care

The Metadata and Collection Services team is responsible for:

- User Support Tier 1: Customer Service related to content
- Validation of the repository's content following validation protocol
- Provision of guidance and support to Repository depositors

The Domain Owners in consultation with Library Infrastructure team are responsible for:

- Management and curation of the repository's content

The Faculty Information Coordinators are responsible for:

- Approval of the publication of MA theses in the repository's content following approval protocol
- Provision of guidance and support to Repository depositors

The Graduate School is responsible for:

- Protocol for the publication of MA & PhD theses in the repository's content

5.2. Monthly Version Release

In the monthly version release meeting, the Repository Manager, Infrastructure Coordinator, ICT Storage Business Owner and ICT Applications Business Owner meet to discuss the version to release that month and how to communicate that across campus to various stakeholders. Optional attendees are ICT Change Manager, ICT Release Manager, Library Senior Technical Developers and Library Software Developer.

5.3. Quarterly Tactics

The Domain Owners, Repository Manager, and Infrastructure Coordinator meet quarterly or as needed to discuss ongoing discussions related to the repository and record their decisions in a log. The Epic Owners are invited to meetings when needed. Epic Owners are assigned by the Infrastructure Coordinator in collaboration with the Domain Owners. The ICT Faculty Information Manager and ICT SLA Manager join this meeting quarterly.

5.4. Overall Strategy

The Library and ICT Business Owners meet annually to discuss new initiatives, major change proposals, and policy. The Business Owners act as the final approvers on tactical decisions, allocate resources, and guide the portfolio value stream. Business Owners join a monthly sprint review or version release meeting as needed.

5.5. User Board(s)

A User Board for the Repository will be established later in 2024. This will allow the Infrastructure team to engage users about the use of the repository. Domain focuses may be considered at a later stage led by the Domain Owner.

6. Collection management

6.1. Collections & sets

Within the repository, records can be grouped and stored as a set or collection.

Sets are defined in the system based on the existing field values (metadata). Unlike collections, there is no curation process in which individual records are judged and selected: inclusion in a set is determined by system defined search and filter criteria alone.

Collections can either be created from the content in the repository or they can be 'pre-existing', in which case the collection existed (digital or in physical form) before it was accepted for the repository.

All collections must have a title and a description. The description must make clear what the purpose is of the collection, what content the collection contains and who the collection is aimed at.

6.1.1. Creation of new sets and collections

Sets/collections are created by the Repository Manager in collaboration with the Domain Owners. The creation of a collection may make it necessary to invite a curator(s) with expertise in the field of the subject. Proposals for new collections and the requirement of curators are reviewed by the Domain Owners.

All TU Delft staff are invited to do proposals for creating and highlighting a collection. Proposals are reviewed and approved by the Domain Owners following the collection procedure. The Domain Owners can also propose to create new collections.

6.1.2. Collection management and curation

The Domain Owners are responsible for the curation of collections that are part of their domain. If needed, the Domain Owners can invite a content expert (curator) to help manage a collection.

The Domain Owners decide:

- To highlight a set or collection (in collaboration with stakeholders and the other Domain Owners)
- The period during which the set or collection is featured.
- The level of visibility, findability and accessibility of a set or collection.
- If a pre-existing collection can be accepted into the repository.
- Whether a set or collection can be archived.
- Decide about the deletion or purging of (a part of) a collection.

The Domain Owners manage the collections that belong to their domain. This includes existing collections (see list in appendix 8.4).

6.1.3. Lifespan of a collection

Some sets or collections will be created with a specific 'life-span' in mind after which there is no need to keep the collection visible and findable.

Each collection, will be re-evaluated periodically and by which it will be decided if the collection needs to be kept in any way:

- Does the collection still need to be visible/findable?
- Does the collection need updating? (adding new, removing old)

6.2. Orphaned collections and other content

If the ownership of a collection comes into question, the Domain Owners will decide what needs to be done with the collection. Not all content will be contained in a defined set collection. The Repository Manager will periodically create a report in order to identify content that is not in any collection and has not been viewed or downloaded for a significant amount of time. The Document Management Archive team will provide Domain Owners advice for content to be deleted or purged. The Domain Owners will then decide if that content needs to remain visible/searchable or if it can be archived, deleted or purged.

6.3. Previous Repository Content

All pre-existing collections will be reviewed with the aim to decide if the collection needs to remain visible and findable or if they will be archived or deselected. (See appendix 8.3) Content not transferred to the repository has an end of life as of July 2025 and will no longer be supported, thus content must be reviewed and moved by the end of 2024.

7. Backup and Archiving

Digital file and metadata within the TU Delft Repository are backed up in accordance with the university policy, as mentioned in the Service Level Agreement between the Library and ICT Department. This creates three copies of the repository's metadata and digital assets, held in two distinct locations.

As mentioned in section 3.5, TU Delft repository currently follows these [recommended file formats](#). This will be reviewed in 2025.

8. Appendix

8.1. Content Categories and Types

TU Delft Repository content for version 1 has 2 data sources (see section 3.1):

- the old repository
- synchronised with the TU Delft CRIS system (Pure) for new content

The alignment of the current content categories and types between Pure and the Repository is listed here before the new repository was built and is effective until 1 July 2024: [Current Pure and Repository Content Categories and Types](#)

The [Proposed Pure and Repository Content Categories and Types](#) will be finalised before May 23rd to announce the changes across campus.

The content categories and types of the TUD Repository are listed in an excel file for internal alignment across multiple platforms: [Types for Lire.xlsx](#).

8.2 New content schedule aligned to data migration and upload ability

1. July 2024 - new content imports from Pure for Research Domain & Education Domain and metadata team acting as depositors for business domain; all other existing content from Pure in domains other than Research will be imported into appropriate domain or library internal space for processing
2. October 2024 - Once all content is in from Pure, depositor role can be activated and sync is adjusted to Repository -> Pure for Research Domain and possibly Education Domain to be discussed with Education Domain Owner. Content can be digitally re-shelved between domains starting from October 2024; some batches re-shelving can be done between July and October 2024.

8.3 Orphaned and Archived Collections

The prior content from the repository has been discussed at length. Here will be the [Orphaned and Archived Collections Decisions](#) pending [decision in Confluence](#).

8.4 Collections

This section will be added after collections have been set up in the system.