

RDA Data Foundation and Terminology

DFT 5: Term Tool Description

Thomas Zastrow, Gary Berg-Cross, Raphael Ritz

State: 7.12.2014

Version 1.0

The general outline of documents from DFT WG is as follows:

- DFT 1: Overview
- DFT 2: Analysis & Synthesis
- DFT 3: Term Snapshot
- DFT 4: Use Cases
- **DFT 5: Term Tool Description**

1. Introduction

The RDA Term Definition Tool (TeD-T)¹ of the DFT WG is a web application for collecting and discussing term definitions. The application is free available for read access and after a free registration, users are also able to edit existing content or create new entries. Existing terms can be browsed as alphabetical list or structured by the underlying hierarchy of TeD-T, a fulltext search completes the browsing functionality. A tutorial is available². The TeD-T platform is deployed and maintained at RZG. Figure 1 shows the start page of TeD-T.

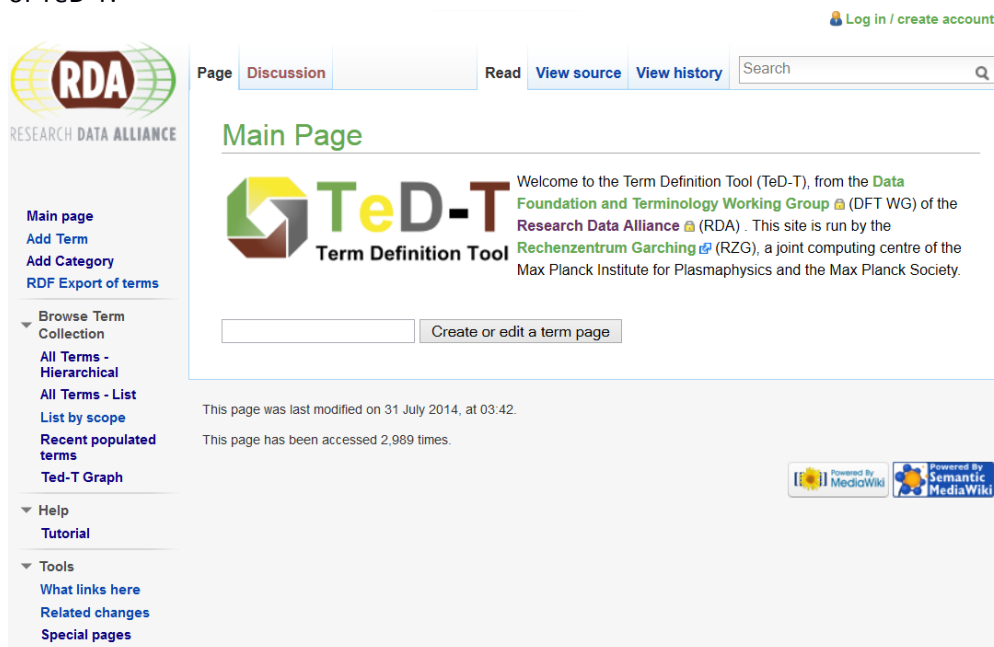


Figure 1: The start page of TeD-T

1 <http://smw-rda.esc.rzg.mpg.de/>

2 http://smw-rda.esc.rzg.mpg.de/index.php/RDA_Term_Collection_Tutorial

2. History

After the start of the DFT working group, several software solutions for creating and managing the DFT term collection were evaluated. Finally, the group decided to go with the Mediawiki³ software in conjunction with the Semantic Mediawiki Extension⁴. As a web application, this combination allows the worldwide distributed members of RDA to create, edit and discuss term definitions in a collaborative way.

A prototypical installation was deployed and configured at RZG. While entering first terms into the prototype, more requirements and feature requests came up. For example, other RDA working and interest groups showed up their interest in using the TeD-T tool for their group internal workflows. Taking these new feature requests and requirements into account, a productive version of TeD-T was deployed at RZG. This version of TeD-T was published and presented to a wider audience at the 4th RDA plenary in Dublin. The already existing content from the prototype was copied over to the productive deployment and the prototype was shut down.

Figure 2 shows the current data model of TeD-T: a term itself is always unique, but it can belong to several categories and can be described by several definitions.

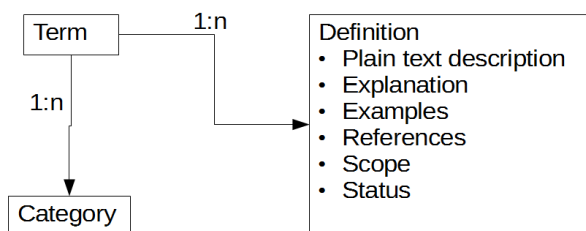


Figure 2: The underlying data model of TeD-T

2. Creating a new term

A registered user can create a new term by entering the name of the term on the start page of TeD-T (see Figure 3). If the term already exists, the system will redirect the user to the already existing page where the user can edit existing definitions or create new ones. If the term doesn't exist so far, a new entry will be created and the user can start entering definitions for that term. For every term, TeD-T adds a discussion page where users can talk about the current definitions. In addition, a version control system allows tracking back the history of a term definition or going back to a status in the past. Registered users can add term definitions to a watchlist, which informs the user about changes on that page via email.

In addition, there is a hierarchy of categories a term can belong to (Figure 4). Because a term can be member of more than one category, the whole content of TeD-T is forming a graph-like structure⁵.

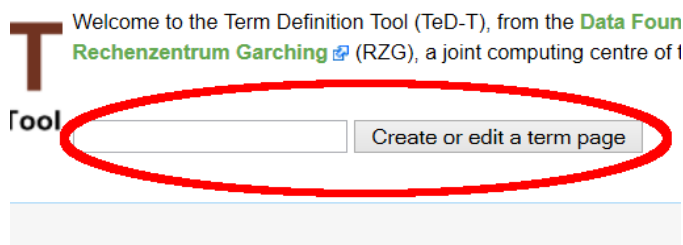


Figure 3: Creating a new term

3 <https://www.mediawiki.org/wiki/MediaWiki>

4 <https://semantic-mediawiki.org/>

5 A visualization of the current graph can be found here:

http://smw-rda.esc.rzg.mpg.de/index.php/Ted-T_Graph

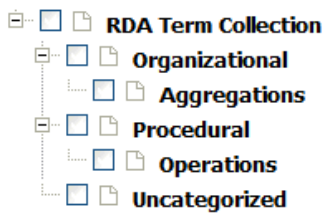


Figure 4: The current hierarchy of TeD-T

4. Creating Term Definitions

After a term was created, users can create an arbitrary number of definitions for it (Figure 5 shows the web form for entering a new definition of a term). Besides the textual description itself, a definition contains an explanation, examples, references, a scope and a status. Scopes are used to differentiate the working and interest groups of RDA – for example, the “Metadata Standards Directory Working Group” can have another definition of a specific term than the “RDA Metadata WG”. The status of a term definition is also a controlled vocabulary and identifies the current position of the definition in the workflow – new, in discussion, deprecated or confirmed.

Figure 5: Entering a new term definition

5. Scopes

Since there was much demand DFT opened the tool for various terminology purposes in RDA so that all groups can manage their own terms separately from the others. The following figure indicates that the term tool is currently used by 10 groups.

When creating a new term you specify the scope as indicated in the previous figure. A list of possible owners (groups of interests mostly being RDA WGs or IGs) is being shown from which you need to select an option.

The hierarchy as explained above is currently only valid for the DFT terms. We need to see how wishes will develop in future.

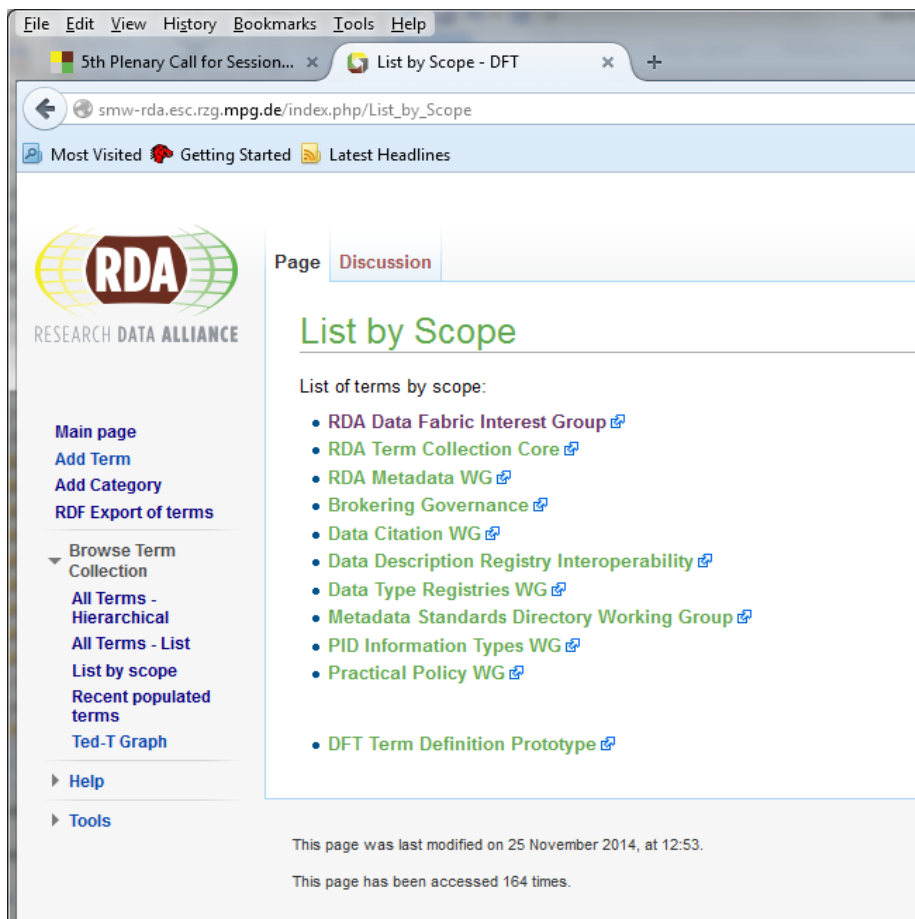


Figure 6: Selecting a scope

6. Current Status

In December 2014, the TeD-T tool contains 159 term definitions which can be browsed as an alphabetical list or ordered by the hierarchy of categories. These terms were entered by 18 active users. Many of the terms were discussed by several users on the attached discussion pages.

7. Further Work & Maintenance

Currently, TeD-T contains only rudimentary functionality for exporting its content into external data formats. It is planned to add a more exhaustive RDF respective OWL output to TeD-T in a further version. That would allow using well-known tools from the semantic web on the DFT term collection. Additionally, an export of selected terms into PDF files and more visualization options about the existing content is planned. In general we will need to see how the requirements will develop in future and in case of solid suggestions we will extend functionality.

RZG will continue to host and maintain the underlying software platform and maintain the TeD-T term wiki as long as there is interest from RDA WG/IG groups and as long as there are groups that take responsibility for the content. Any change in this respect will be first discussed with TAB and Secretariat.