

# TeD-T: Term Definition Tool http://dft-tool.rd-alliance.org

Raphael Ritz, MPCDF

Data Foundation and Terminology IG

research data sharing without barriers rd-alliance.org

# **Background and Motivation**

- Provide an online tool to coordinate and document community contributions to the development of the RDA DFT Vocabulary
- Software stack: Semantic Media Wiki plus Semantic Forms (think "enhanced wikipedia" software)
- · Customizations:
  - Pages have a substructure
  - Various overviews
  - Simple categorization







Main page
Add Term
Add Category
Stable Releases
Download in various
formats

Browse Term
Collection

All Terms Hierarchical

All Terms - List
List by scope
Recent populated
terms
Ted-T Graph

▶ Help

Tools

Page Discussion Read View source View history Search Q

#### Main Page

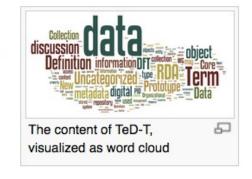


Welcome to TeD-T, the Term Definition Tool of the Data

Foundation and Terminology Interest Group ♠ (DFT IG) of the

Research Data Alliance ♠ (RDA).

The tool was originally put in place to support the DFT Working Group (2013-2015). The DFT WG task was to describe a basic, abstract data organization model which can be used to derive a reference data terminology that can be used across communities and stakeholders to better synchronize conceptualization, to



enable better understanding within and between communities and finally to stimulate tool building, such as for data services, supportive of the basic model's use. We assume that this abstract data organization model will focus on common building blocks and their characteristics, along with relevant protocols.

After termination of the Working Group the tool continues to support the then established DFT Interest

## Structure of a page

- Definition
- Explanation
- · Examples
- · References
- · Scope
- · Status

Possibly multiple times

Plus categorization



Page Discussion

Read Edit with form Edit



Search

#### Repository Registry

Definition	A repository Registries is a type of registry that collect useful information about repositories for human consumption in order that depositors and users can easily find where to go to for their data needs.		
Explanation			
Examples	re3data is an example		
References	https://rd-alliance.org/group/data-fabric-ig/wiki/data-fabric-ig-repository-registries.html		
Scope	RDA Data Fabric Interest Group		
Status	New		

Category: Infratructure

Main page

Add Term

**Add Category** 

**RDF Export of terms** 

**Browse Term** Collection

> All Terms -Hierarchical

All Terms - List

List by scope

Recent populated terms

**Ted-T Graph** 

▶ Help

**▼ Tools** What links here

Related changes

This page was last modified on 19 January 2016, at 22:30.

This page has been accessed 14 times.

## Simple categorization

## Contributing

- · First you need to register
- After approval you can log in and add new or edit existing terms, definitions, examples, explanations, references, etc.
- Content is versioned
- Content can be discussed ("Discussion" tab)



Raphael Talk Preferences Watchlist Contributions Log out



Read Edit with form Edit

Search

#### Edit RDA: Research data infrastructure

#### Main page

Add Term

**Add Category** 

Stable Releases

Download in various formats

**Browse Term** Collection

> All Terms -Hierarchical

All Terms - List

List by scope

Recent populated terms

Ted-T Graph

- ▶ Help
- Tools

#### Definition:

A research data infrastructure is a digital infrastructure organized to promote data sharing and consumption in support of research efforts.

#### Explanation:

Sometimes discussed as a vision of e-infrastrcture or cyberinfrastructure

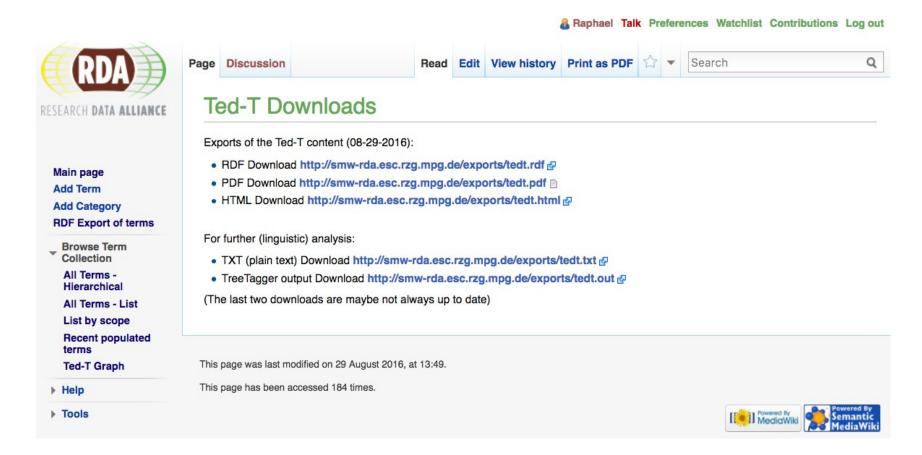
#### **Examples:**

Examaples include infrsstrucrtues made of high-throughput scientific instruments, telescopes, satellites, accelerators, supercomputers, sensor networks, and running simulations. https://www.eudat.eu/

#### References:

https://www.rd-alliance.org/vision-global-research-datainfrastructures.html

## Download HTML, PDF, RDF, ...



http://dft-tool.rd-alliance.org/index.php/Ted-T\_Downloads

## Releases



Main page

Add Term

**Add Category** 

Stable Releases

Download in various formats

Browse Term Collection

> All Terms -Hierarchical

All Terms - List

List by scope

Recent populated terms

**Ted-T Graph** 

Page Discussion

Read

View source

View history

Search

irch

#### **DFT Releases**

Short before an RDA plenary, we are exporting the content of TeD-T as stable snapshot. Handle PID(s) are assigned to these snapshots:

- Term Definitions Version 3.0 "Gaborone": https://hdl.handle.net/21.11109/dft-3.0 6
- Term Definitions Version 2.0 "Berlin": https://smw-rda.esc.rzg.mpg.de/dft-2.0.html 🙃
- Term Definitions Version 1.0 "Montreal": https://smw-rda.esc.rzg.mpg.de/dft-1.0.html
- Prerelease "Barcelona": https://smw-rda.esc.rzg.mpg.de//exports/tedt.html

This page was last modified on 3 September 2018, at 14:34.

This page has been accessed 248 times.





Q

- PIDs are assigned to snapshot releases of the entire vocabulary (an HTML export)
- Granularity
  - Full vocabulary: https://hdl.handle.net/21.11109/dft-3.0
  - Single term: https://hdl.handle.net/21.11109/dft-3.0@#
     Collection
  - Specific definition:
  - https://hdl.handle.net/21.11109/dft-3.0@#Collection\_2
- Versioning



- "Speaking PIDs"
  - human readable
- Structure inspired by URLs
  - familiar
- Includes version number
  - implicit "time stamp"
  - allows for evolution
  - plan: new release for every plenary
- Points to full vocabulary not terms, nor definitions, but ...



... terms or definitions can be referred to by appending a term as *part identifier;* plus number for a specific definition, e.g., https://hdl.handle.net/21.11109/dft-2.0@#Active\_Collection\_2



#### **Definition 2**

An Active Collection is a collection that is being generated dynamically by executable code.

Scope: RDA DFT Interest Group

Status: In discussion

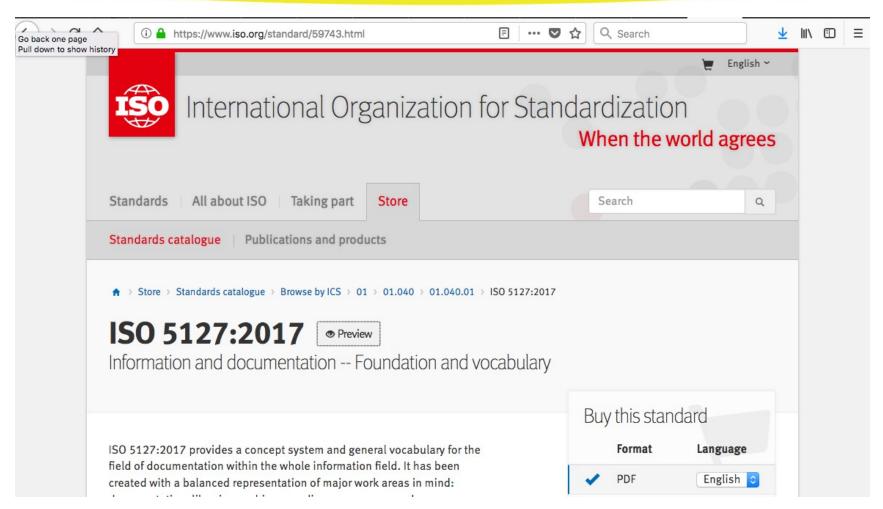
#### **Active Data**

#### **Definition**

Active data denotes virtual units of data objects that are created dynamically by executable code.

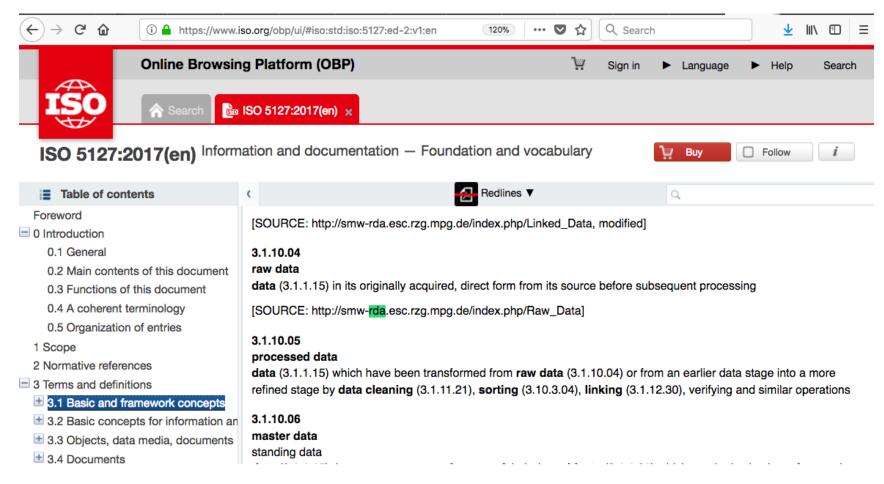
Examples: Here several possibilities can be imagined such as generating data tables from a relational database with the help of





https://www.iso.org/standard/59743.html





https://www.iso.org/obp/ui/#iso:std:iso:5127:ed-2:v1:en



- Work started at P8 in Denver
- Stefan Zednick provided a basic script to transform RDF to SKOS
- Remaining Issue: cast from terms to concepts

No progress since



- How to establish relationships across releases?
  - Collections?
  - PID-Versioning?
- Naming conventions OK?
- SKOS export any takers?
- Expanding scope?
- · ... your questions ...



#### Websites

- DFT IG: https://rd-alliance.org/groups/data-foundations-and-termig.html
- TeDT: http://dft-tool.rd-alliance.org
- People:
  - Gary Berg-Cross: gbergcross@gmail.com
  - thomas.zastrow@mpcdf.mpg.de
  - raphael.ritz@mpcdf.mpg.de



Backup slides



```
- <rdf:RDF>
- <rdf:Seq rdf:about="Data">
- <rdf:Sli>
- <rdf:Description rdf:about="Data">
- <rdf:Description rdf:about="Data">
- <tedt:text>
Data is a collection of datum. Data is a potential information bearer to a cognitive agent. (Prototype Wiki version)
</tedt:text>
- <tedt:explanation>
Data is the medium used to communicate and store information. It gives a concrete and persistent status to some information, but itis by itself (as a string of bits, for example) without context such the assumed formatting/representation information or reference systems without a single meaning. Thus, to extract the information intended in collections of data, an interpretation is necessary that assigns meaning to it using elements of this context such as assumed representation.
</tedt:explanation>
- <tedt:examples>
```

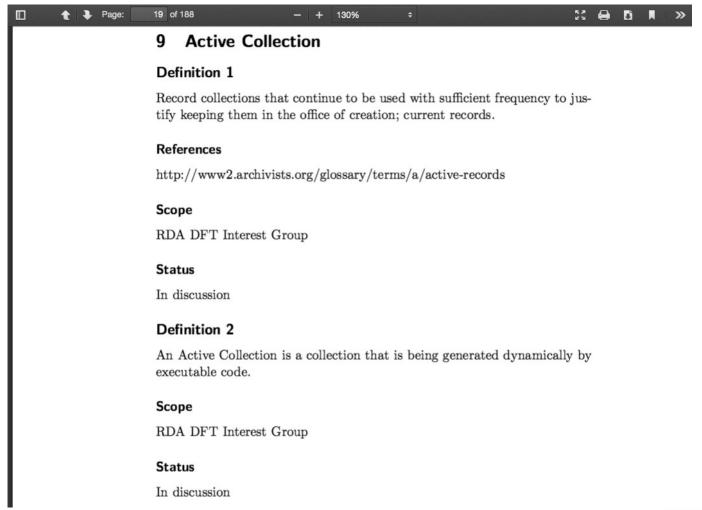
Data like a datum has a quantitative or qualitative value. Common types of data include content for sea surface temperature measurements, readings from monitoring equipment, user actions on a website, science funding projections, and demographic information. (after http://data-informed.com/glossary-of-big-data-terms/)The capital letter A is ASCII character 65, ut the bits to express this may be a pixel of some sound if the context is different. Likewise, the same information (content) may be expressed by very different looking data since it may be encoded in many different ways. A simple example is that the temperature of some object or region may be codes by degrees C or F (different reference systems). In this sense data may be considered analogous to syntax which helps communicate the semantics which is information.

-<tedt:explanation>

Note this is a general view and means that non-digital things are included such as analog data.

</tedt\*evnlanation>







#### **Term Definitions Version x.y**

Data Foundation and Terminology Interest Group (DFT IG) of the Research Data Alliance (RDA)

http://smw-rda.esc.rzg.mpg.de



API Consumer Layer Definition API Consumer Layer is a Digital Infrastructure layer that enables e-Infrastructure providers and university librarians to find connections across research data registries. See also Provider and Graph Creation Layer Data Description Registry Interoperability New Access Definition Data access typically refers to software and activities related to storing, retrieving, or acting on data housed in a some stored form In the context of data access refers to a user's ability to access or retrieve data digital objects and digital resources. The idea of access means that data content is available for use. Users who have data access can store, retrieve, move or manipulate data. Data can be stored on a wide range of devices \& technologies such as a database or repository. Related terms Access Control \& Access Control List. After http://www.techopedia.com/definition/26929/data-access Practical Policy WG In discussion Access Control Definition Access control refers to controlling access to information and systems. Access control policy provide details on the nature of controls placed on access via identification such as passwords. See Access control list. Having locked access to where the system is stored is a physical control, having access to files or software code is a digital access control. Practical Policy WG New Access Policy Definition Access Policy is a type of policy that expresses authorization rules, The granularity of access policies is typically coarse-grained making them suitable for matching up to a broad

http://smw-rda.esc.rzg.mpg.de/exports/tedt.txt



range ofservices.

API	NP	API	
Consumer		NP	Consumer
Layer	NP	Layer	
Definition		NP	<unknown></unknown>
API	NP	API	
Consumer	5	NP	Consumer
Layer	NP	Layer	
is	VBZ	be	
a	DT	a	
Digital	NP	Digital	
Infrastructure		NN	infrastructure
layer	NN	layer	
that	WDT	that	
enables	VBZ	enable	
e-Infras	structure	9	NP <unknown></unknown>
providersand		NN	<unknown></unknown>
university		NN	university
librarians		NNS	librarian
to	TO	to	
find	VB	find	
connecti	ions	NNS	connection
across	IN	across	
research	ı	NN	research
data	NNS	datum	
registries		NNS	registry
	SENT		
See	VB	see	
also	RB	also	
Provider		NP	<unknown></unknown>
and	CC	and	
Graph	NP	<unknown< td=""><td>1&gt;</td></unknown<>	1>
Creation	ı	NP	<unknown></unknown>
Layer	NP	Layer	
Data	NP	Data	

http://smw-rda.esc.rzg.mpg.de/exports/tedt.out



- DTR/PID Information Types (linking versions)
- Collections (grouping terms/definitions)
- Dynamic Data (tracking changes/additions)

What are your expectations, experiences, use cases?



http://dx.doi.org/20.500.11946/tedt-barcelona

One PID for a static HTML Snapshot of the tool's content.



https://smw-rda.esc.rzg.mpg.de/dft-1.0.html

A static HTML Snapshot of the tool's content with PIDs assigned to each **term**.

Example: 21.11109/dft/1.0/Collection



#### collection



#### Collection

#### **Definition 1**

A collection is an organized, systematic form of purposeful aggregation, grouping or arrangement of elements, that has an identity of its own separate from the identity of the elements.

**Explanation:** A collection's metadata should provide one or more reasons why this particular group of elements belongs together as part of a collection process. Collections are often associated with archives and repositories and their services.

**Examples:** A collection of books in a library. In this case the library serves as a repository for selected books and other media. Collections usually serve multiple functions, such as selection and collocation of related materials, narrowing of search scope, and clarification of information needs.

References: H. L. Lee, H. L. a???The concept of collection from the usera???s perspective.a??? Library Quarterly, 75 (1), 67-85. 2005.

Scope: RDA Term Collection Core

Status: New

#### **Definition 2**

Collection is defined as "a group of objects gathered together for some intellectual, artistic, or curatorial purpose."

References: "Representing Cultural Collections in Digital Aggregation and Exchange Environments"http://www.dlib.org/dlib/may14/wickett/05wickett.html Europeana Data Model (EDM)

Scope: RDA Term Collection Core

Status: New

#### **Definition 3**

A collection is a digital object which is identified by a PID and consists of a set or a list of PIDs/lds and a set of additional



https://hdl.handle.net/21.11109/dft-2.0

One PID for a static HTML Snapshot of the tool's content as part identifiers **can** be made to work, e.g.:

https://hdl.handle.net/21.11109/dft-2.0@# Active Collection





#### **Active Collection**

#### **Definition 1**

Record collections that continue to be used with sufficient frequency to justify keeping them in the office of creation; current records.

References: http://www2.archivists.org/glossary/terms/a/active-records

Scope: RDA DFT Interest Group

Status: In discussion

#### **Definition 2**

An Active Collection is a collection that is being generated dynamically by executable code.

Scope: RDA DFT Interest Group

Status: In discussion







#### Term Definitions Version 2.0 "Berlin"

### Data Foundation and Terminology Interest Group (DFT IG) of the Research Data Alliance (RDA)

This file file is an export of the TeD-T term collection of the DFT IG, done at Thu Mar 15 12:05:01 CET 2018

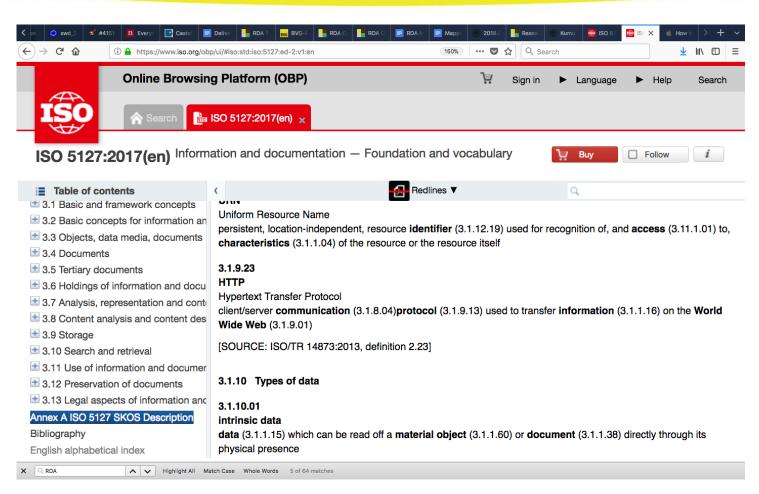
http://smw-rda.esc.rzg.mpg.de

A / B / C / D / E / F / G / H / I / J / K / L / M / N / O / P / Q / R / S / T / U / V / W

#### Α

API Consumer Layer	Definition	21.11109/dft-2.0@#API_Consumer_Layer
Access	Definition	21.11109/dft-2.0@#Access
Access Control	Definition	21.11109/dft-2.0@#Access_Control
Access Policy	Definition	21.11109/dft-2.0@#Access_Policy
Access Role	<b>Definition</b>	21.11109/dft-2.0@#Access_Role
Access Workflow	Definition	21.11109/dft-2.0@#Access_Workflow
Access a repository	Definition	21.11109/dft-2.0@#Access_a_repository





https://www.iso.org/obp/ui/#iso:std:iso:5127:ed-2:v1:en

