Technical documentation

D2.4 KPI Specification

Versions

Version	Date	Comment	Responsible
0.2	2016-11-25	Modified some statements to make them more precise, identified some areas where further elaboration is required.	Torill Hamre Øystein Godøy
0.1	2016-10-14	Added more details to document.	Øystein Godøy
0	2015-06-20	Initial draft	Øystein Godøy

Table of Contents

Index of Tables

1 Introduction

1.1 Background

The NorDataNet project aims at *providing the scientific community with an integrated, cost-efficient and sustainable infrastructure following established standards for data documentation, archiving, search and exchange*. Focus is on coordination of efforts and requirements, enhancing interoperability and user interfaces to existing facilities, as well as development of toolboxes for data documentation in order to reduce technical and governance obstacles. The infrastructure must provide online access to data and facilitate long term preservation, in order to maximise the benefit of public funds invested in the datasets. The intention is to link existing institutional and discipline specific systems to promote science regardless of geographical and institutional location.

1.2 Scope

The purpose of this document is to identify a number Key Performance Indicators that can be used as metrics for evaluation of the project. This evaluation shall cover both project progress and user uptake of the services developed.

1.3 Audience

This document if developed for both and internal and an external audience. The internal audience is the development team and the external audience is both the Research Council of Norway, the contributing data centres and the data providers.

1.4 Applicable documents

[1] Norwegian Scientific Data Network (NorDataNet) Project Plan, revision 19, NFR Contract 245967
[2] NorDataNet Concepts and acronyms, Version 0.1
[3] NorDataNet Use cases, Version 0.1
[4] NorDataNet System requirements, Version 0.1
[5] NorDataNet System design and interfaces, Version 0.1
[6] NorDataNet, System implementation and integration plan, Version 0.1

2 Key Performance Indicators definition

Key Performance Indicators (KPI) helps defining and measuring the progress towards the goals and objectives of the system under development. KPIs are a tool supporting this process. In this context KPIs should measure the uptake and relevance of the system developed in the user community as well as the timeliness and quality of the services provided to the user community.

Key characteristics of the KPIs considered in this context are:

- relevant and consistent in relation to overall objectives and goals

- representative

- realistic

- specific

- measureable

- trend related (i.e. consistent over time)

In order to support this KPIs are divided in two main categories according to the system performance they are measuring:

- user relevance, uptake and impact

- quality of services

When properly developed these KPIs are reported on a monthly basis and graphical representations of time series are made available on the project web page. No information related to a specific data centre, data provider or data consumer is made public available.

3 KPIs

3.1 User relevance, uptake and impact

3.1.1 Background

3.1.2 Number of users

Using the server logs, these are analysed for unique addresses accessing the service on a daily basis. The aggregated information (number of users, not individual users) is stored for future use. Server logs are rotated on a weekly basis. Users of a higher order services are assumed to register and will be incorporated in a more complicated KPI in the future. This KPI includes both users of the data search interface and the editorial information on best practises for scientists.

3.1.3 Number of datasets accessed

Collecting information from the contributing data centres, the number of datasets accessed (metadata viewed/downloaded/visualised/transformed) every day is stored as a time series along with the information from which data centre the data was collected.

This KPI will be modified/amended in the future to better reflect services through the central access point (which relies on the contributing data centres).

3.1.4 Number of integrated datasets served

For some datasets there is a possibility of generating new datasets which are combination of multiple datasets. This service is integrated in the central end point for the service, but relies on data access end points (OPeNDAP) at contributing data centres. A daily value is stored as a time series.

3.1.5 Number of publications or products

The number of publications or integrated products utilising a specific dataset is collected and reported. The temporal granularity of this KPI is monthly.

This KPI is not available until DOI is implemented by the contributing data centres and actively

used in the reporting through Cristin. Details needs to be discussed with other relevant national activities.

3.2 Quality of services

3.2.1 Background

3.2.2 Availability of end points

Several times a day the availability of end points for metadata (e.g. OAI-PMH) and data (OGC WMS, OPeNDAP, …) is checked and registered. This information is reported on a monthly basis. The availability is reported as a percentage measure.

The target for NorDataNet is 95% availability in office hours (08:00-17:00) Norwegian time.

3.2.3 Temporal frequency of dead links

All available metadata and editorial material is checked on a daily basis for broken links (internally or externally). The numbers of valid and broken links are reported on a monthly basis.

3.2.4 Availability of central access point

The availability of the central access point of the service is reported in a similar manner as for 3.2.2.

3.2.5 Response time for data discovery

The user experience of the central data discovery interface is registered through the response time of the search operations performed by users. This is collected for all search operations, but reported as mean numbers on a monthly basis with max and min values attached for the month.

The target is that no search request should take more than 30 seconds to complete.

3.2.6 Response time for data transformation

The response time for all data transformations is recorded. The KPI to be reported need further development and discussion within the project when transformations are developed. This will depend on the size of the data set, or sets if several datasets are to be combined.