



PROJECT DELIVERABLE REPORT

DELIVERABLE NUMBER	D1.2.1
TITLE	GOVERNANCE PLAN DATABASE
AUTHOR(S)	DR. J. BOUWMAN, PROF L DRAGSTED, PROF L BRENNAN
WORK PACKAGE	WP1
TASK	TASK 1.2.1
WP LEADER	DR. J. BOUWMAN
BENEFICIARIES CONTRIBUTING TO THE DELIVERABLE	TNO, UCC, UCPH
STATUS – VERSION	FINAL - VERSION 1.0
DELIVERY DATE (MONTH)	M24
SUBMISSION DATE	M30
DISSEMINATION LEVEL – SECURITY*	PU
DELIVERABLE TYPE**	R

* Security: PU – *Public*; PP – *Restricted to other programme participants (including JPI Services)*; RE – *Restricted to a group specified by the consortium (including JPI Services)*; CO – *Confidential, only for members of the consortium (including JPI Services)*

** Type: R – *Report*; P – *Prototype*; D – *Demonstrator*; - O - *Other*



INTRODUCTION

ENPADASI has delivered a database with 23 observational studies and 79 intervention studies (see deliverable 2.1.1). These studies can be integrated via the DASH-IN Web-based Interactive federated analysis system (deliverable 3.7.1). These systems are valuable and should be maintained for at least the coming five years.

The partners in ENPADASI that host servers that are part of DASH-In have indicated that they will keep the current system running as is for the coming 5 years on their own funds. However, the system will require updates with updates of web browsers, bug reports and new uploads of studies may result in new features requests. Moreover, nutritional researcher will require training (see also deliverable 1.2.3) and support for uploading of data. The specific needs for the maintenance of the systems are indicated below. Currently ENPADASI has not found funding for updates of the system, but has several links and opportunities that will be described below.

REQUIREMENTS

The DASH-IN (data sharing in nutrition) infrastructure has been developed in ENPADASI as a federated system. Data can be stored locally in either an instance of the Phenotype database or in Opal. In addition, TNO hosts an instance of the Phenotype database that can be used centrally (<https://dashin.eu/interventionstudies/>). Meta-data can be stored in any instance of the Phenotype database or can be store in the central instance of Mica (for Opal data: <http://mica.cloud.ba.infn.it/>). All these systems need support for the long term. If data can be anonymized and centrally stored, the central instance of the Phenotype is the system of preference as the system is most adhered to the FAIR principles, otherwise sharing of the meta-data is requested via the central system of Mica (for data that are stored in Opal) or an instance part of DASH-IN of the Phenotype database.

ENPADASI stimulates open sharing of data, however we have observed that sharing for ethical and/or legal issues can be restricted. By sharing only the meta-data, no actual data is revealed but the nutritional community is aware of the existence of the dataset. Sharing of data between specific partners can be resolved by signing a DTA (data transfer agreement) as is delivered in WP5 (deliverable 5.3). However, to prevent potential sharing difficulties a data access committee should be put in place and made sustainable. It will be in charge of the review of the data access form and to check if all the regulatory issues have been complied and the potential specific restriction of the data provider (see deliverable 1.2.3).

DASH-IN is relevant and useful for the entire nutritional community and the community is stimulated to use the infrastructure. However, as the maintenance of the system is not yet taken care of, storage of large studies may result in issues for the hosting parties in ENPADASI. Therefore, usage of the system cannot be guaranteed



Integration of data may result in large datasets that are too large to analyze on the hosting servers and may also become too big for local analysis on a PC. For this reason grid or cloud solutions may be needed. The ENPADASI partners can help with finding solutions for this purpose. A google mailing-list (enpadasi@googlegroups.com) is published on the ENPADASI website in order to communicate with the ENPADASI partners after the finalization of the project.

OPPORTUNITIES

A full support and maintenance (see also deliverables 1.2.3 and 1.2.4) of the system per year will require 100-500 kEuro per year, depending on the service level and the amount of data included in the database. This includes the server costs, salary of a programmer and support for the usage of the system. A minimal service level will include only the costs of the servers, online support and resolving of crucial bugs, whereas a higher level of service will include workshops on location and further development of the infrastructure based on feature request. Considering the initial investment and the willingness of the community to share data where possible we feel that it is imperative the funds are secured to harness the full benefits of the work performed within ENPADASI. Development of the analysis pipeline for integration, we expect that the main development will be part of projects that use the DASH-IN system.

The partners of ENPADASI are linking to ELIXIR in Europe and to local ELIXIR initiatives in order to retrieve some funding for this work. Moreover, the upcoming Horizon 2020 Foodcloud call may include some funding for this work. ENPADASI is closely linked to the research infrastructure initiatives around Food, Nutrition and Health (FNH) to build a collaborate research infrastructure for the Food and Nutrition initiative. With other FNH partners Jildau Bouwman has written an ELIXIR use case in order to link to the general 'data for life' initiatives.