



Call: PRIMA Section 1- Agri-Food Value Chain

Project ID: 2032

Project Acronym: FunTomP

Project Title: Functionalised Tomato Products

Date : 26/12/2021

Project Deliverable #/Name: D1.3 / Open Research Data Management Plan (ORDMP)

Work Package: WP8

Lead Beneficiary: METU

Dissemination level: CO



			HISTORY OF CHANGES
Version	Publication date	Prepared by	Changes
1.0	26.12.2021	Dr. Mecit Oztop	Initial version

Project Abstract:

Despite its anticipated benefits, recent studies show that the MedD is now progressively disappearing in Mediterranean countries, particularly amongst the youngest generations. In contrast; there is an increasing trend by that millennials and young population, that favor foods that are healthy, ethically and naturally sourced, and minimally processed. And this changing consumer tendency created a significant demand for increased plant-based protein product consumption. The objective of FunTomP is to reformulate traditional Mediterranean tomato products considering the current consumer trend of 'functional foods', using leaf proteins (by-products of sugar beet processing) and olive powder by using novel and eco-friendly processing technologies that will impact the nutrients minimally. Tomato will be transformed to different functional foods (juices, sauces, leather, bars, powder mixes) offering extra health benefits to satisfy the consumer demand while keeping a sustainable product and process cycle with the valorization of agricultural waste. Olives are going to be homogenized under high pressure and afterwards, will be dried by using an energy efficient drying methodology to obtain olive powder. Three different technologies will be evaluated on liquid products of which the benefits have been proved in the literature in terms of their effect on the microbial load and nutritional quality compared to conventional heating. These will be Microwaves (MW), UV-C light and High-Power Ultrasound (HPU). Sustainable production technologies will be preferentially selected for the production of the different products and sustainability approaches will be integrated to the processes in accordance with the 2030 Agenda for Sustainable Development Goals (SDGs). The selection of most environmentally friendly techniques will be assisted by Life Cycle Assessment tools. Waste of the tomato processing; that is mainly the skins and also beet processing waste (pulp and the green pellet) will also be utilized to recover the bioactive compounds. For the recovery of these valuable compounds, green technologies such as Supercritical Fluid Extraction (SFE), ultrasound (US) and High Voltage Electric Discharge (HVED) techniques will be used. FunTomP will also include consumer, market and sensory analysis to evaluate market receptivity for the developed products to predict the extent of the shift in the adherence to the Mediterranean diet consequent to its commercialization.

Partners and Work package info of the project can be accessed at <https://funtomp.com/>.

Contents

1. Data Summary	5
1.1 Purpose of the Data Collection/Generation	5
1.2 Relation to the Objectives of the Project.....	5
1.3 Types and Formats of the Data Generated/Collected	6
1.4 Re-use of Existing Data	6
1.5 Origin of the Data	6
1.6 Expected Size of The Data	6
1.7 Data Utility	6
2. FAIR data.....	6
2.1 Making data findable, including provisions for metadata	6
Discoverability of the data	7
Identifiability of data	7
Naming Conventions.....	7
Keywords	7
Metadata	7
2.2 Making Data Accessible	7
Openly Available Data and Availability of the Data	8
Tools to access the data	8
2.3 Making Data Interoperable:	8
2.4 Re-Use of the Data (Through Clarifying Licenses):	8
3. Other Research Outputs	9
4. Allocation of Resources	9
5. Data Security.....	9
6. Ethical Aspects.....	10
7. Other	11

1. Data Summary

Provide a summary of the data addressing the following issues:

- State the purpose of the data collection/generation
- Explain the relation to the objectives of the project
- Specify the types and formats of data generated/collected
- Specify if existing data is being re-used (if any)
- Specify the origin of the data
- State the expected size of the data (if known)
- Outline the data utility: to whom will it be useful

1.1 Purpose of the Data Collection/Generation

The objective of FunTomP is to reformulate traditional Mediterranean tomato products considering the current consumer trend of 'functional foods', using leaf proteins (by-products of sugar beet processing) and olive powder by using novel and eco-friendly processing technologies that will impact the nutrients minimally. Tomato will be transformed to different functional foods (juices, sauces, leather, bars, powder mixes) offering extra health benefits to satisfy the consumer demand while keeping a sustainable product and process cycle with the valorization of agricultural waste.

Purpose of the data collection/generation: Product formulation and characterization tests will be performed. Data will be collected throughout the experiments, surveys and literature analysis. The main data categories are listed as follows:

- 'Physical properties' experiment results of ingredients and products
- Consumer data obtained from surveys, sensory panels

The data will be made available to the users working in related areas to be used as 'example' data sets.

1.2 Relation to the Objectives of the Project

FunTomP is a project that will address and highlight the importance of Mediterranean Diet. New functional tomato products will be formulated using sustainable production technologies. The data that will be generated in line with the objectives of the project. The data collection and sharing will also enable researchers working in similar fields to analyse the data and use for their research. Work packages of FunTomP are given below. Except WP7 most of the data collected in the project will be 'experimental'.

WP #	WP Title	Lead Participant #	Lead Participant Short Name	Start Month	End Month
1	Management and Coordination	1	METU	1	48
2	Extraction and characterization of sugar beet leaf proteins	2	EgU	1	30
3	Production and quality evaluation of dried tomato products	1	METU	3	36
4	Formulation and quality evaluation of functional liquid tomato products	7	UALg	6	42
5	Nutraceutical and anticancer evaluation of the products	11	BIOMIC_AUTH	8	45
6	Production and characterization of bioactives from tomato skin & beet leaves	8	UoZ-1	25	48
7	Consumer and marketing and sensory analysis for the products	15	AJU	6	48
8	Dissemination & Training & Engagement with Stakeholders	1	METU	2	48

1.3 Types and Formats of the Data Generated/Collected

The experimental data generated will be mostly quantitative but qualitative data will also be collected in Work Package 7 where 'Consumer Analysis' on products will be performed. The data obtained will be in the following formats:

- .dat
- .csv
- .txt
- .xlsx
- .m
- .tiff
- .jpeg
- .docx
- .opj
- .raw

1.4 RE-USE of Existing Data

Literature data will be used in 'sustainability analysis' (in WP 2 and WP4). Protein databases for proteomics and bioinformatic analysis will also use the data in the relevant databases. Mostly, new data will be generated throughout the project. Data already published, or data collected by the consortium will be utilized to fulfill the objectives established in the different WPs. Data will be used in the scientific publications.

1.5 Origin of the Data

The data belong to experimental data of developed tomato and protein products and consumer analysis results of WP7. Commercial functional tomato and protein products that can be used for comparison purposes will also be tested to get data. Any other sample that is relevant to tomato products may also be examined under the project.

1.6 Expected Size of The Data

Data to be generated will vary greatly in size, from megabytes (Mb) (*relaxometry, spectroscopy, -omics data*) to larger sizes (gigabytes – Gb for *imaging and proteomics/metabolomics, spectroscopy data*).

1.7 Data Utility

The data generated in this project will be useful, first and foremost, to the scientific community around the world, and is relevant especially to research groups developing research and innovation on different functional food products. The data obtained will increase the awareness of the different companies and relevant sectors on the production methods used in the project.

2. FAIR DATA

2.1 Making data findable, including provisions for metadata

- Outline the discoverability of data (metadata provision)
- Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?
- Outline naming conventions used
- Outline the approach towards search keyword
- Outline the approach for clear versioning
- Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

DISCOVERABILITY OF THE DATA

The main strategy that will be followed to ensure that the data generated in **FunTomP** can be easily discovered by third parties will be through the publication of the scientific results in **open access journals** that are indexed in main peer-reviewed literature databases (SCOPUS, PUBMED, Web of Science). Open Research platform of the European Commission (<https://open-research-europe.ec.europa.eu/>) will also be a platform where publications will be submitted. In addition, when the publications are disseminated through Social Media, links to the open data set will also be provided. We will share our published dataset in the publicly accessible disciplinary repository Zenodo using descriptive metadata as required/provided by that repository. A '**Data Repository**' link will also be available in the web of page of the project.

IDENTIFIABILITY OF DATA

Dissemination of open access publications in the project web page and social media pages will enable third parties and users to become aware of the data, but also to locate and access project results and data. Datasets will be findable and identified, whether in Scientific Publications or Zenodo (a centralized repository supported by the EC), through a globally unique and eternally persistent identifier (e.g. DOI). Thereby dataset of FunTomP will get a unique Digital Object Identifier (DOI) provided by the publishers or other identifier (accession number or handle number) provided by the repository services.

NAMING CONVENTIONS

Data will be available through its publication. It will refer to the figures and table numbers used in the study so will be easy to identify the data. For the datasets in the repository ; files will be uniquely identifiable and versioned by using a name convention consisting of project name, dataset name, method used, ID, place and date.

KEYWORDS

Keywords of the publications will be relevant to the goals of the project.

METADATA

All data will be accompanied with metadata which will contain as a minimum:

- *Description of contents;*
- *Author/owner;*
- *Date of creation;*
- *Datatype;*
- *Terms of use;*
- *Keywords.*

2.2 Making Data Accessible

- **Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**
- **Specify how the data will be made available**
- **Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**
- **Specify where the data and associated metadata, documentation and code are deposited**
- **Specify how access will be provided in case there are any restrictions**

OPENLY AVAILABLE DATA AND AVAILABILITY OF THE DATA

All data that gave rise to the results published in the scientific publications, and is considered necessary for third parties to be able to reproduce and validate results will be openly accessible. These data sets will usually be available as supplementary materials. Datasets that are presented in publications, or as supplementary materials in publications will be available in the Google Drive folder of the project which will be accessed through the project web page. Users will need to fill a simple form to access the **Data Repository** of the project. Data not published anywhere will be deposited to Zenodo and also to Google Drive folder of the project.

TOOLS TO ACCESS THE DATA

All data will be presented in digital formats which are readable by most operating systems and respective standard software tools. For data that needs a special software; links will be provided for open access tool required to open the data. Data, associated metadata, documentation and codes (if available) will all be stored in the data repository of Zenodo. Zenodo has no restrictions of use and permissions. Data are granted to any user following a simple procedure of registration. In case it turns out that an institutional repository of a beneficiary will be used, the consortium will ensure that any user will be able to access the data. The consortium partners will opt for public domain licenses as these allow high reusability and place no restrictions upon access and use of data by third parties. In addition to Zenodo, in the web page of the project a 'Data Repository' tab has been added where links are given to the Google Drive Folder of the project data. Users will just fill a simple form on the web page to access to the dataset folder.

2.3 Making Data Interoperable:

- **Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**
- **Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

All partners will generate or will enable copies of datasets and results in formats compatible with most common and open source software applications, in order to allow data exchange and re-use between researchers, institutions, organizations, countries, as much as possible. These formats will enable transfer of information and also re-combinations with different datasets from different origins. The data formats that will be used in SuChAQuality are based on formal, accessible, shared, and broadly applicable language to allow inter-disciplinary interoperability. In the unlikely case that uncommon or project-unique specific ontologies or vocabularies have to be used, we will provide mappings to more commonly used standard vocabularies for all uncommon data types to allow maximum interoperability. Any assumptions made, and the mapping steps to more commonly used vocabularies will be documented in a readme file and stored on the Google Drive folder of the project or in the data repository (Zenodo).

The following formats will be used for the data acquired suggested:

- *.txt/.csv/.xml/. opj/. dat/. raw (for data)*
- *.docx/. doc/. pdf/. ppt (for reports/presentations)*
- *. jpg/. png/. gif/. tiff/. fig (for graphics)*

2.4 Re-Use of the Data (Through Clarifying Licenses):

- **Specify how the data will be licenced to permit the widest reuse possible**
- **Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**
- **Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**
- **Describe data quality assurance processes**
- **Specify the length of time for which the data will remain re-usable**

In Zenodo, the data will be licensed under Creative Commons CC BY 4.0 International. The open access approach will be extended to the scientific community only for non-confidential data. For confidential data or data which will be used for a patent application; IPR strategy written in the Consortium Agreement will follow.

All data acquired in the project will become available during the last three months of project end. Parts of the data can become available even before due to journal publications. There will be no embargo period. The data can be re-used by other scientists in the field of Physics, Chemistry, Statistics, Food Science, Mathematics, Consumer Sciences, Energy Research. Neighboring disciplines and interdisciplinary research groups might also be interested since FunTomP is an interdisciplinary consortium.

The data quality will be assured by each partner, that bears the responsibility of them. The tools necessary for describing and identifying the dataset and for preparing the metafiles. The consortium will follow the most relevant data quality standards set by peer review journals to secure the production of high-quality data to be shared with others.

As open formats are used for data archiving, the data will remain re-usable until the repository withdraws the data or goes out of business.

3. Other Research Outputs

In addition to the data set produced in the project; all scientific and non-scientific outputs of the project will be shared through the Social Media Channels of the project and also through the website of the project. In the webpage of the project there is already the 'Scientific Outputs' and 'Dissemination' tabs where users can be informed. To make these tools more accessible; research outputs will be highlighted in the forms of 'Tweets', 'Stories', 'Posts' in Social Media Accounts of Twitter, Instagram and LinkedIn. The related files will also be available in the Google Drive folder of the project.

4. Allocation of Resources

Explain the allocation of resources, addressing the following issues:

- **Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**
- **Clearly identify responsibilities for data management in your project**
- **Describe costs and potential value of long-term preservation**

The main costs of making FAIR are mainly associated with the fees of open access publishing of the scientific results. For the Open Research Platform of EC, these fees are also exempted. Other costs are negligible as the utilization of the majority of the repositories and software for data handling are free of charge. The publishing costs will be covered by funds from the project. These costs related to open access to research data are eligible and will be paid from the Indirect budget. In addition, there will be a monthly fee paid for the cloud service 'Google Drive'.

The implementation of the data management plan will be mainly responsibility of the Coordinator. Consortium will be updated on the plan regularly.

Long term storage will be secured with a cloud service and deposit of data in the Institutional Server @ OpenMETU (open.metu.edu.tr). Costs of this long-term storage are secured by Institutional funds. The data will be kept for 4 years after conclusion of the project, which is the period set in the Grant Agreement for beneficiaries to take measures aiming to ensure exploitation of its results.

5. Data Security

Address data recovery as well as secure storage and transfer of sensitive data

All data produced in the project, whether research data or other (including all documents produced from the management activities), will be centralized and stored in a commercial cloud service that has a monthly fee (Google Drive). All consortium members will have access to this service. Zenodo will also be used as the data repository for some data sets. A copy of all the original data will also be kept safe in the cloud service of METU which will provide the

basis for data recovery and secure storage in case the online platform fails. Transfer of sensitive data that have 'IPR' restriction will also be enabled by the institutional cloud service.

6. Ethical Aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former.

Every study involving humans that will be carried out during FunTomP project will comply with ethical principles and with applicable international, European and national law. The researcher will ensure the respect for people and for human dignity, fair distribution of research benefits, and the protection of the values, rights and interests of the research participants.

It should be mentioned that research involving children will only be carried out if it is the only way, or the most appropriate way, to obtain the information needed for the project, and also only if:

- Studies with adults would not be effective
- There is only minimal risk to the participants
- The results of the research will benefit the individual or the group represented by the participants

Special documentation will be provided to assist to FunTomP 's researchers in the issue of ethics during the process of studying human behaviour, paying special attention when children are involved in tasks researches. For the preparation of this documentation the following standards and directives, standards and publications will be followed:

- REGULATION (EU) 2016/679 (General Data Protection Regulation) 27th of April 2016
- ISO 9186-1 Graphical symbols – Test methods – Part 1: Method for testing comprehensibility. Second Edition 2014-03-15
- ESOMAR International code on Market, Opinion, Social Research and Data Analytics (2016)
- ESOMAR – World research codes and guidelines. Interviewing children and young people. Last version printed 2009
- ESOMAR - World research codes and guidelines. Passive data collection, observation and recording. Last revised 2009
- DIN 66399. Three-part standard regarding the seven security levels of paper shredders
- Guidance How to complete your ethics self-assessment. Version 4.01, October 2015. European Commission. Directorate-General for Research & Innovation. European Commission. Ethics for researchers. Facilitating Research Excellence in FP7. Directorate-General for Research and Innovation Science in society /Capacities FP7.

The main points that will be integrated in this document will be:

- Methodology description and protocols in research with humans
- Informed consent
- Details on the age range
- Details on child assistant procedures, parental consent and welfare of the minor.

Every test with consumers, will follow the following subtasks:

- Consumer test objectives, samples and methodology variables definition and scheduling. Sampling method by: Gender, social class and rural-urban representation and differences.
- Pilot preparation: Ensuring that all the requested project material will be available for the test with consumers.
- Questionnaire/ plot design.
- Questionnaire programming (just for quantitative studies).
- Consumer's recruitment.
- Fieldwork (carrying out the interviews).
- Statistics analysis.

- Report elaboration.
- Report presentation to FunTomP's design team

AJU, the FunTomP's partner that will coordinate the research with humans, has more than 30 years of experience in children and adult social research, and its ethical and data privacy procedures have been approved by the European Commission, the CEN in previous H2020 projects and European public tenders. Different variables will be analysed in each study as: Gender, social class and rural-urban representation and differences.

7. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

All necessary procedures have been described.