Deliverable 2.2
CHORIZO FLW "Insighter"

D 2.2

CHORIZO FLW "Insighter"

Document ID	D2.2
Due date	30th September 2023
Submission date	29th September 2023
Dissemination level	Public (PU)
Work Package	WP2
Author(s)	Capwell Forbang Echo, Samuel Bosch, Siham Koudi, Bart Van Droogenbroeck, Anna Twarogowska, Isabeau Coopmans, Jürgen Vangeyte, Mikayel Hovhannisyan, Efstathios Zavvos, Rosalyn Old
Document version	1.0

Grant agreement	101060014	**Duration**	36 months
Start Date	October 2022	**End date**	September 2025

Contributors

Name	Organisation
Capwell Forbang Echo	EV ILVO
Samuel Bosch	EV ILVO
Siham Koudi	EV ILVO
Bart Van Droogenbroeck	EV ILVO
Anna Twarogowska	EV ILVO
Isabeau Coopmans	EV ILVO
Mikayel Hovhannisyan	EV ILVO
Jürgen Vangeyte	EV ILVO
Efstathios Zavvos	VLTN
Chantal den Broeder	VLTN
Aristos Halatsis	VLTN
Rosalyn Old	CSCP
Nora Brüggemann	CSCP

Revision history

Version	Date	Reviewer	Modifications
V0.2	01/08/2023	Capwell Forbang Echo	Text and editing
V0.4	27/09/2023	Efstathios Zavvos	Text and editing
V0.6	28/09/2023	Rosalyn Old	Text and editing
V0.8	29/09/2023	Efstathios Zavvos	Minor editing
v1.0	29/09/2023	Capwell Forbang Echo	Final text

**Funded by
the European Union**

This project has received funding from the Horizon 2020 research and innovation programme under Grant Agreement No 101060014.

Disclaimer

The content of this document reflects only the author's view. Neither the European Commission nor REA are responsible for any use that may be made of the information it contains.

While the information contained in the documents is believed to be accurate, the author(s) or any other participant in the CHORIZO consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the CHORIZO Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the CHORIZO Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

Copyright message

Horizon
Europe

Executive summary

Food loss and waste (FLW) is a real and evident problem in the world today, with several driving factors. In order to mitigate this problem, information and data is crucial to provide relevant insights into the various context and drivers of FLW.

A systemic and multifaceted solution approach is necessary. In an attempt to provide part of the solution, the CHORIZO project has built a FLW data hub and "Insighter". This document accompanies the infrastructure as part of D2.2 deliverable to elaborate the potential and applicability of the CHORIZO FLW datahub / "Insighter" while giving user friendly directives on how to interact with the infrastructure, and use its resources to combat FLW.

This document starts with an introduction of the CHORIZO project, highlighting the critical links of the CHORIZO FLW datahub and "Insighter" to other work packages and tasks in the project. It further elaborates on the infrastructure's development approach undertaken by project partners to arrive at the final design, which addresses all the essential and expected functionalities. In the third chapter, the functional data layers making up the datahub and "Insighter" are explained, with directives on how to navigate the site to perform basic operations. This chapter also delves into the potential impacts of the infrastructure and briefly highlights its "future" and sustainability plan.

Finally, this write up ends with a conclusion and an appendix. The conclusion gives the potential next steps for the infrastructure, and the appendix contains material used for gathering information throughout the development process and more detailed procedure to perform certain actions on the infrastructure.

Table of Contents

List of Tables

List of Figures

Abbreviations and Acronyms

Acronym	Description
CHORIZO	Changing practices and Habits through Open, Responsible, and social Innovation towards ZerO food waste
HE	Horizon Europe
EU	European Union
RIA	Research and Innovation Action
FLW	Food Loss and Waste
FGI	Focus Group Interview
IDI	In-depth Interview
EC	European Commission
FAIR	Findable, Accessible, Interoperable and Reusable
OA	Open Access
CKAN	Comprehensive Knowledge Archive Network
UI	User Interface
DMP	Data Management Plan
WP	Work Package
CS(s)	Case Study(ies)
SP	Strategic Plan
API	Application Programming Interface
TRL	Technology Readiness Level
M (36)	Project Month (Month 36 of the project running)
ABM	Agent Based Model

1. Introduction

1.1. The CHORIZO project

Chorizo stands for "**C**hanging practices and **H**abits through **O**pen, **R**esponsible, and social **I**nnovation towards **Z**er**O** food waste". It is a Horizon Europe (HE), European Union (EU) Research and Innovation Action (RIA) type of project which principally aims to improve the understanding of how social norms and behaviours influence food loss and waste (FLW) generation across various areas of consumption. The knowledge generated in the project will be used to improve the effectiveness of decision-making and engagement of food chain actors towards zero food waste.

The CHORIZO project employs six real-life Case Studies (CSs) with partners spread across Europe (Figure 1) to serve three interlinked purposes;

Figure 1: CHORIZO Case Studies and partner locations

i. To provide information and data on the context and impact of previous FLW prevention/reduction actions undertaken by the Case Study members, thus enriching the evidence-based analysis on previous FLW actions;

ii. To generate new evidence on the interaction between social norms, behaviour and food waste, to feed into the project's FW models and innovation products;

iii. To validate the communication and science education packages of the project.

Data being an important prerequisite for attaining the projects objectives, the CHORIZO project will develop a Food Loss and Waste (FLW) Datahub and an FLW "Insighter". This tool will incorporate and store:

- The results of the evidence-based analysis of previous/ongoing FLW actions

- New empirical case study evidence on social norms, consumer behaviour, economic actor behaviour and charity (e.g. food banks) behaviour in relation to FLW, and

- Provide more timely and easily accessible evidence to inform FLW decision making, accompanied by a stakeholder engagement and governance model to ensure the datahub's evolution and sustainability beyond the project duration.

Building on the open science policy of the European Commission (EC), the CHORIZO datahub and FLW "Insighter" joins other existing research data management initiatives and international guidelines to ensure that research data is **F**indable, **A**ccessible, **I**nteroperable and **R**eusable (FAIR). CHORIZO datasets will be managed in line with the FAIR principles:

- Findability: For the data that can be made publicly available, a trusted OA repository will be used and a unique DOI to each dataset will be assigned. A naming convention and standard identification mechanism of the datasets for persistent identifiers was previously agreed upon by partners and elaborated in the project data protocol (D1.1) and data management plan (DMP; D7.2)

- Accessibility: Data that do not interfere with commercial and industrial interests will be made OA through open access publications and the research data associated with such publications.

- Interoperability: Accessed datasets are readable, and able to be processed with ordinary computer software. Concerning data generation from experimental activities, its format will ensure interoperability.

- Reusability: The datasets will be reusable as they will be findable, accessible and interoperable in the long term with sufficient documentation and metadata.

Overall, the data stored in the datahub will not only be FAIR for humans, but also for machines as it will enable automate search, linking and access to data. Important or key elements making CHORIZO datasets FAIR such as rich metadata and documentation, using open or standard file formats, and using licences for reuse are considered. This is crucial as it enhances possibilities to replicate research, a more and broader means to share knowledge and to refine previous research, higher visibility for researchers and greater efficiency by avoiding undesired duplication in the generation of data.

1.2. Critical links to other project work packages and tasks

The CHORIZO project is well structured with a coherent flow of data. Figure 2 below shows the inter-relationship between the various WPs and tasks in the project. The FLW datahub and FLW "Insighter" is a the core of the project as it plays a central role in the smooth linkage of data providers (case studies and other data providing tasks) to data data/insights users.
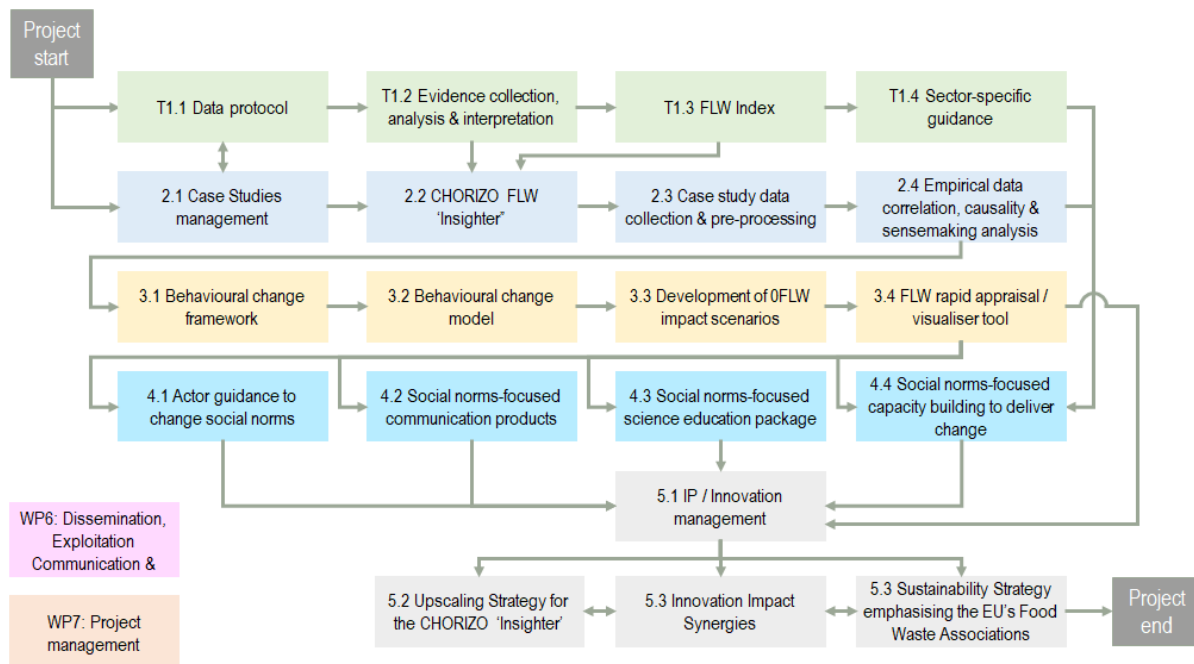

Figure 2: Link between Datahub and FLW "Insighter" and other project tasks

1.3. Deliverable overview and report structure

In the quest for making data FAIR, the project tasks amongst others began with elaborating a concise data protocol (deliverable D1.1): that defined (i) the specific data requirements of each Work Package (WP); (ii) the data collection methods; (ii) the format and naming of the datasets; (iii) the data definitions and descriptions (metadata). Aligned to D1.1, is the project's DMP (deliverable D7.2) which further describes the data and procedures to collect, process and store the data, thus the handling of the data during and after the project. In addition, a case studies' (CS) strategic plan (SP) (D2.1) which was aimed at redefining the objectives of the CSs, identifying and mapping out empirical data to be collected from six case studies, using well defined data collection techniques, and processed, paying relevant attention to meta data.

This deliverable has two main chapters, excluding the Introduction and Concluding chapters. The process, timeline and design (architectural) evolution that led to the conception of the CHORIZO datahub and FLW "Insighter" is first highlighted with basic user roles explained (Chapter two). The third chapter is dedicated to the FLW datahub, the FLW "Insighter" and the functionalities as provided by the tool with respect to datasets management and access to "insights" respectively.

2. Development approach

2.1. Development timeline

Table 1: Datahub and FLW "Insighter" development timeline

Month		M1-M3	M4 – M5	M6	M7 – M11	M12
	General plan	Initial discussions	Information collection and consultation	T2.2 discussions on design	Building Data hub / FLW "Insighter"	D2.2
Task participants **EV-ILVO** VLTN ICP CSCP	**Specific plan**	• Internal discussions at ILVO on possible structure of datahub • First tentative idea penned	• Prepare information tool and send out to partners • Meeting with ILVO colleagues to discuss output of survey on technical information collected from CSs and WPs • Prepare first proposal of datahub design	• Meeting with T2.2 members, WP leads to discuss/feedback on datahub proposed design • T2.2 lead works on feedback and update proposed design as necessary • T2.2 participants meet with DG Sante (EU FLW Prevention Hub). Email exchanges indicated a postponement of the meeting to a date TBD. • Discussions with T2.2 members to finalize design	• Implementation/building datahub according to final design • Test running the datahub • Continuous feedback and interaction with partners to improve datahub • Addressing any irregularities that may arise	• Internal quality review • D2.2 delivered

2.2. Design evolution

The CHORIZO datahub and FLW "Insigter" is expected to be developed and reach a maturity level of TRL5[1] (Technology Readiness Level of 5) at M36 of the project. The deliverable D2.2: CHORIZO FLW "Insighter" according to the project timeline is due for M12. Table 1 above details the process or phases taken within the project to arrive at M12.

Although the task was planned to start in M4 and end in M12 with D2.2, initial discussions to dissect the functional needs of the datahub/FLW "Insighter" (Appendix 5.7) as specified in the project grant agreement were held between M1 - M3 (phase 1).

Phase 2: M4 - M5 was an information collection and consultation phase with project partners. This phase saw the implementation of an information collection tool (appendix 6.1) to map out the understanding and expectations of CSs and WPs regarding the CHORIZO datahub/FLW "Insighter". The ideas gathered coupled with the initial thoughts of M1 – M3 were combined to produce a first draft design of the datahub / FWL "Insighter" (Figure 3).
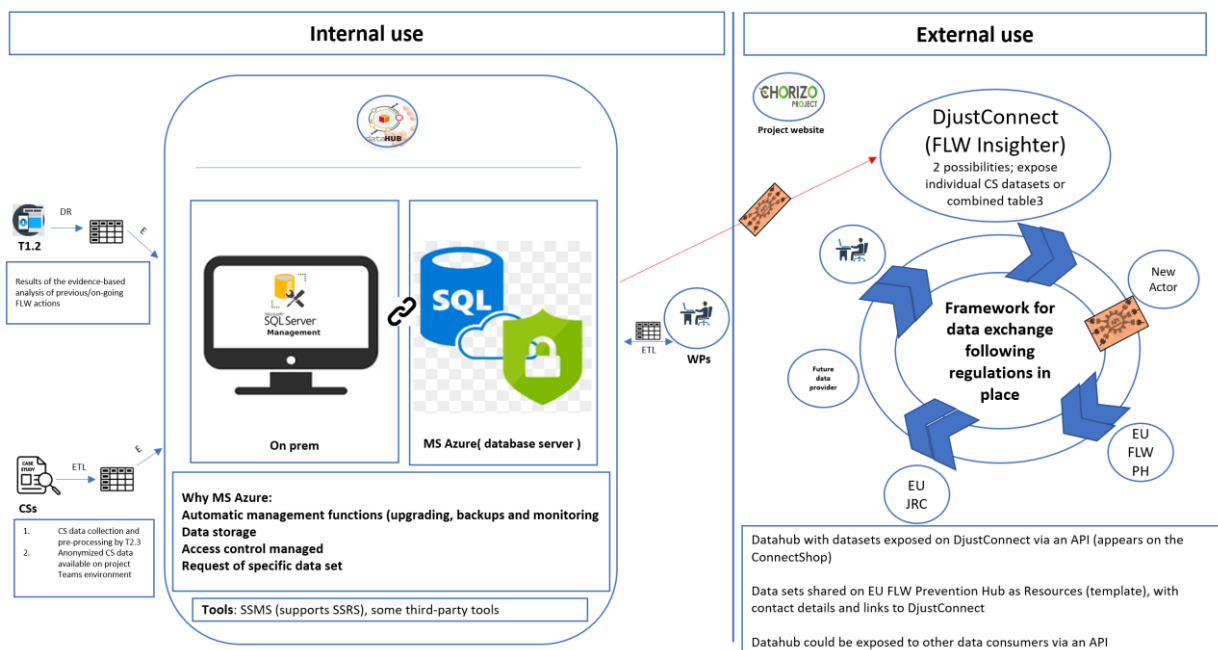

Figure 3: First proposed datahub and FLW "Insighter" design

Figure 3 above depicts the then proposed internal and external design, expected flow of data between various stakeholders, software involved, and processes to enable smooth flow of data. This design was based on MS Azure technology with other third party tools. The internal

---

[1] h2020-wp1415-annex-g-trl_en.pdf (europa.eu)

use mainly focused on the project and its duration while the external use looks at the infrastructure's "life" after the project.

Phase 3: M6 saw discussions and feedback on the first draft design (Appendix 5.8) of the datahub proposed to partners. Major feedback and action points on Figure 3 design were:

- Design should ensure datasets are exposed through APIs to other existing initiatives or tools such as the FLW Prevention Hub of the European Commission and the DjustConnect Platform. This will ensure we stick to the project grant agreement promises. An action to be followed up here after M12 is to liaise with the Directorate General for health and Food safety (DG SANTE) of the European Commission in charge of the EU FLW Prevention Hub to align on the technical possibilities of exposing datasets in the CHORIZO datahub through APIs on the formers' infrastructure.
- CHORIZO datahub and FLW "Insighter" will be an independent site with similar functionalities as the DjustConnect platform and not dependent on the later.
- Would an SQL-based tool best handle the qualitative (unstructured) datasets which area majority in the project?

The above feedback led to a revised design which better aligned with the needs and expectations of the project (Figure 4).
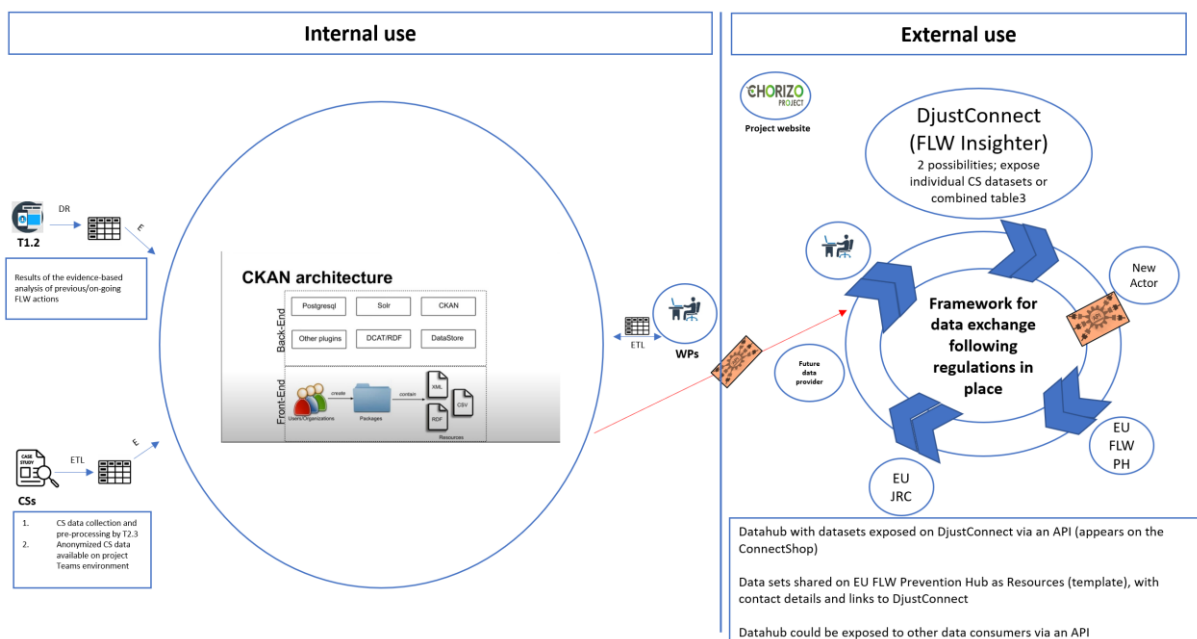

Figure 4: Final design of datahub and FLW "Insighter"

The final design, with a CKAN core architecture (see section 2.3.1 and 2.3.2 for elaboration on CKAN) still maintained an internal and external exploitation or usage pattern.

Internally (within the project until M36) three main data and insight providers or tasks can be elaborated; (i) results of the evidence analysis of previous and on-going FWL actions (which led to D1.2: Evidence-based analysis of FLW prevention actions). This is a comprehensive overview of past and current FLW prevention actions (interventions) identified across EU member states within task 1.2 of the CHORIZO project. The analysis explores various facets of the actions, including food waste prevention levels, implementation challenges, the broader social, economic, and environmental impacts, gender considerations, and to what extent the interventions are able to illuminate social norms at play that affect behaviour towards food waste; (ii) project CS data collected, pre-processed and analysed within tasks 2.3 and 2.4 of the CHORIZO project respectively. These data spanning various consumer context gives insights into how social norms influence FLW; and (iii) insights from analysed information from the CSs and T1.2 performed in the modelling WPs of the CHORIZO project.

Externally (M12 onwards and especially after M36), links (predominantly exposing the datasets through APIs) will be established between the CHORIZO datahub / FLW "Insighter" and other existing infrastructures such as the FLW Prevention Hub of the European commission and the DjustConnect platform. Other external stakeholders will be able to access open data and request other data sets either through the CHORIZO datahub or their APIs on other platforms. The project website will also have a re-direct link to the main CHORIZO datahub / FLW "Insighter" site.

2.3. CKAN

2.3.1. What is CKAN?

CKAN stands **C**omprehensive **K**nowledge **A**rchive **N**etwork. It is a python application or tool used for creating open data websites, capable of managing and publishing collections of data. It can be used by a wide range of stakeholders ranging from national and local governments, research institutions, and other organizations, projects that collect and handle huge amounts of data. Published data can be searched, previewed as tables, graphs, maps by various users

(CKAN contributors, 2023). The architectural design of CKAN showing components of the front-End and back-End can be seen below (Figure 6).

Front-End of CKAN

The front-end takes into account the specifications which results to an interactive and user friendly user interface (UI). The front-end consist of users/organizations (in our context of the CHORIZO datahub; Data layers), wherein packages are created to contain resources or datasets in various formats.

Users, organizations (data layers) and authorization

Users of CKAN-base data websites can register a user account in three major steps (Figure 5). With a user account, users can perform numerous data publishing functions: datasets can be created, edited, etc. depending on the level of permission granted. There are three user roles namely Admin, Editor and Member, with the corresponding privileges explained below;

- Admin: Can add/edit and delete datasets, as well as manage organization members.
- Editor: Can add and edit datasets, but not manage organization members.
- Member: Can view the organization's private datasets, but not add new datasets.

From an external stakeholder point of access, a user ID needs to be created.

To create a user ID:

- Step 1: Click on register on the top right corner of the homepage
- Step 2: provide the required information: a username, full name, e-mail address and a password. A profile picture can also be uploaded associated to the user. This is however optional.
- Step 3: Click on create account.

However, a login account is not required to search and subsequently download open datasets and other resources.
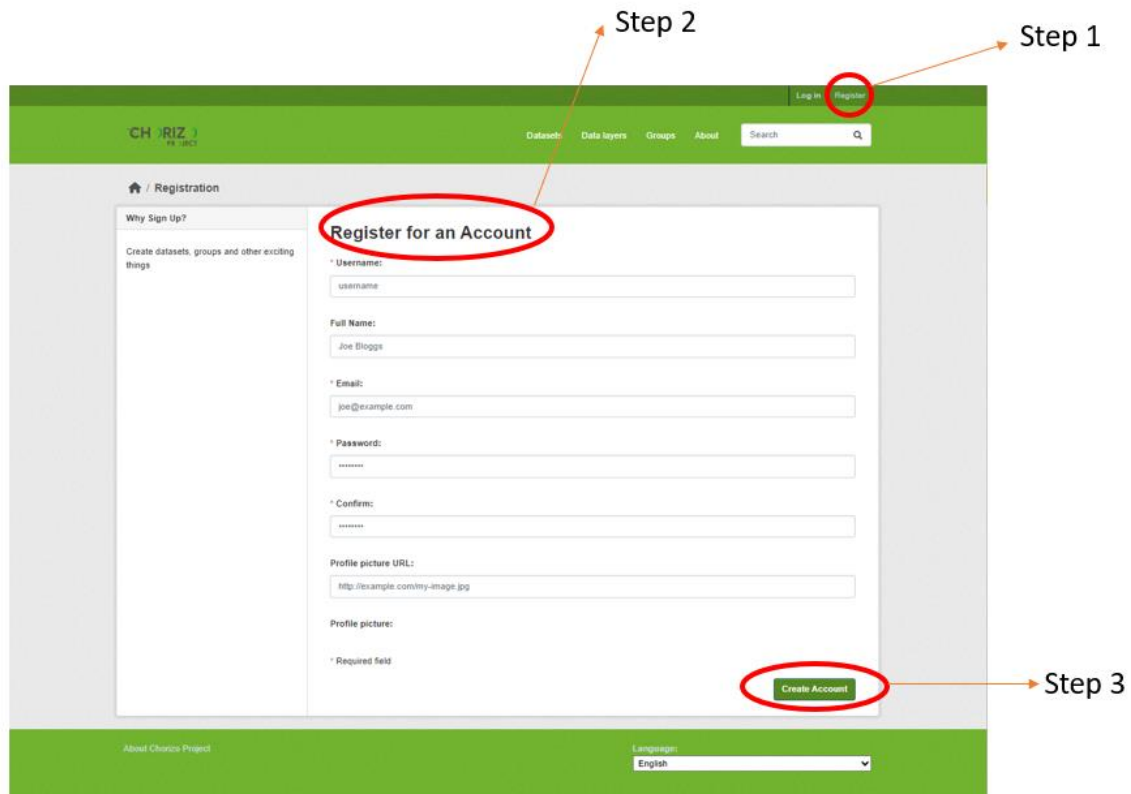
Figure 5: Registering an account of CHORIZO Datahub

"Organization" are considered "owners" of datasets. In the case of the CHORIZO datahub, data layers (section 3.1) is used in interchangeably with "organization". A CKAN instance can have any number of organizations. Each organization / data layer can have its own work flow and authorizations, allowing it to micro-manage its own publishing process. In the CHORIZO project, there are three data layers (see section 3.1). Admins of a data layer can add users to it, with different roles depending on the level of authorization needed. Datasets belongs to, and are normally created within a data layer.

Back-End of CKAN

The backend often called the "server-side" is the "invisible"[2] part of the website. The backend is essential for storing, organizing data, and ensuring everything at the client side works as expected. The backend, programmed in python, works in unison with the frontend by sending and receiving information that is displayed on every webpage. CKAN runs on a server and uses a *PostgreSQL* database which stores datasets and other resources in a structure that facilitates

---

[2] Part not seen from the user interface, but accessible to the developers or site admins

retrieval, organizing, editing, and saving. CKAN uses *Solr* as its search engine, and uses a customized *Solr* schema file that takes into account CKAN's specific search needs. *Solr* needs to be installed and configured after CKAN installation to enable its functionalities. The data catalog vocabulary (DCAT) is a resource description framework (RDF) vocabulary enhances interoperability of the site. The DCAT also describes the datasets in the catalog, enabling publishers to increase the discoverability of resources (datasets) and applications to properly consume metadata from multiple catalogs. CKAN backend also has possibilities of numerous plugins installed.
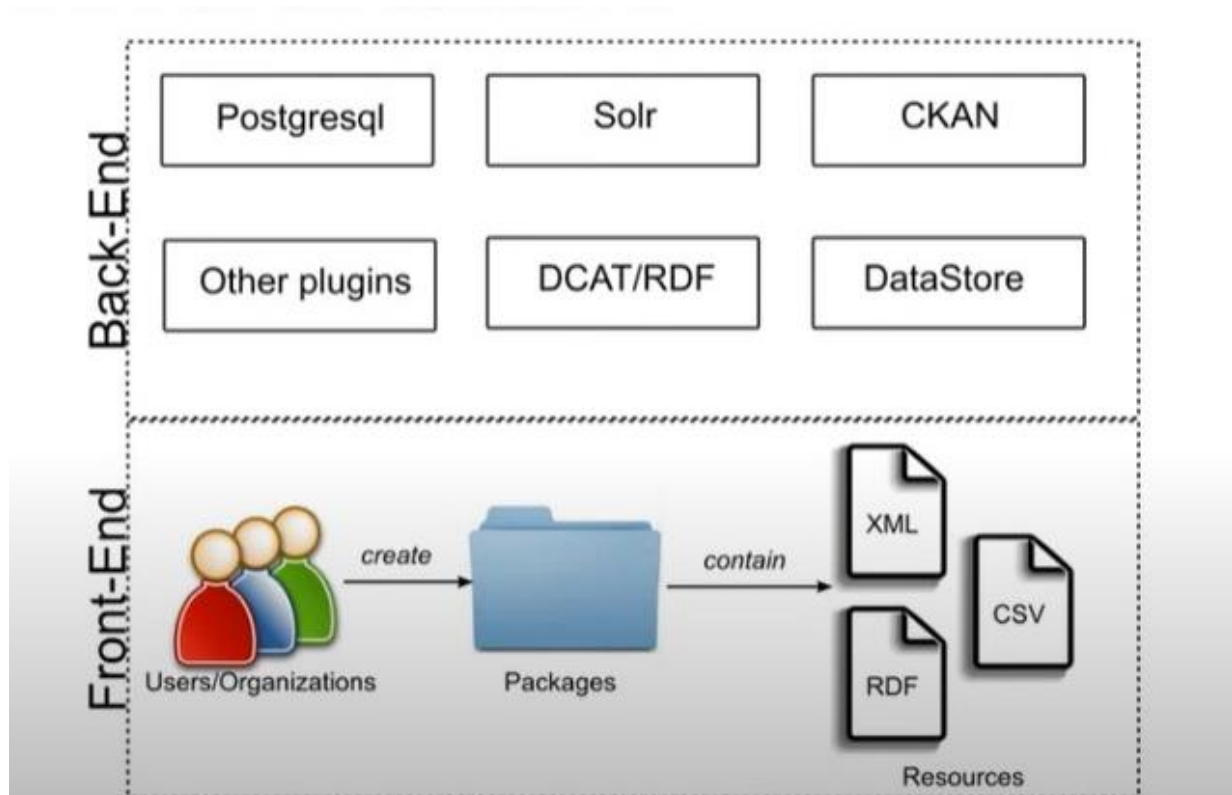
Figure 6: CKAN architecture

2.3.2. Why a CKAN based design?

In the consultation phase leading to the final design for the datahub, several possibilities were screened and a CKAN base design best suited the functional needs of the CHORIZO data hub and "Insighter". Table 2 below highlights the key CKAN features, their description and how they match the CHORIZO datahub and "Insighter" needs.

Table 2: Matching CKAN feature possibilities to CHORIZO Datahub / FLW "Insighter" needs

CKAN Feature	Description	CHORIZO Datahub / FLW "Insighter"
API	All of a CKAN-base website's core functionalities can, be used by external code that calls CKAN API	YES : ☒ NO : ☐
Datastore	The CKAN datastore extension provides an ad hoc database for the storage of structured data from CKAN resources.	YES : ☒ NO : ☐
Metadata	CKAN provides metadata by default, and additional attributes can be defined as needed	YES : ☒ NO : ☐
Manage data	CKAN provides an intuitive web interface that allows for easy registration, update and refine of datasets.	YES : ☒ NO : ☐
Search	CKAN provides a rich search experience which allows for quick 'google-style' keyword search and filtering.	YES : ☒ NO : ☐
Themeable	Customizable to fit needs	YES : ☒ NO : ☐
Visualization	CKAN creates data visualization views and exposes data as line, bar and pie charts	YES : ☒ NO : ☐
Federate	CKAN creates a federated network of data portals which share data between each other	YES : ☒ NO : ☐
Filestore	CKAN gives the possibility to upload media and image files. Stores files either on the server or in the cloud using extensions or links	YES : ☒ NO : ☐
Security	CKAN security practices and controls	YES : ☒ NO : ☐
Extend	CKAN allows you to pick and choose which essential features to use for the data portal. Also provides and extending guide to easily develop an extend to fit a particular context.	YES : ☒ NO : ☐
Geospatial	CKAN makes it possible to have advanced geospatial features, covering data preview, search and discovery.	YES : ☒ NO : ☐

Additionally, CKAN being an open-source and easily customizable tool gives a ser-friendly web-base interface for both administrators and end-users. This will be essentially for non-technically inclined users to contribute to the usage and growth of the CHORIZO datahub and "Insighter". Moreso, CKAN based site can be extended with a wide range of plugins and extensions to provide diverse tools and services enabling it to be versatile and creating potential for integration with other existing infrastructure. Furthermore, CKAN supports different data standards and compliance requirements, essential for stakeholders to adhere to specific regulations and has a active community of developers to provide support to users.

3. The CHORIZO datahub and FLW "Insighter"

The CHORIZO datahub and FLW "Insighter" is broadly a 2-in-1 infrastructure. As a datahub, it will efficiently catalogue and index large amounts of data, providing robust metadata management and comprehensive description of datasets. Datasets will come from within the project (Table 3) as well as outside the project from interested stakeholders tackling FLW. The stored datasets can be versioned, which is crucial for tracking changes and revisions to datasets. As a FLW "Insighter", the infrastructure will provide new insights- generated from data and FLW analytics on the levers to change social norms.

3.1. Infrastructural and functional layers

The CHORIZO datahub and FLW "Insighter" has three functional layers (Figure 7).

- Data layer 1: Multi-source evidence collection, analysis and interpretation
- Data layer2: New evidence generated from the project case studies on the interaction of social norms and FLW behaviour and
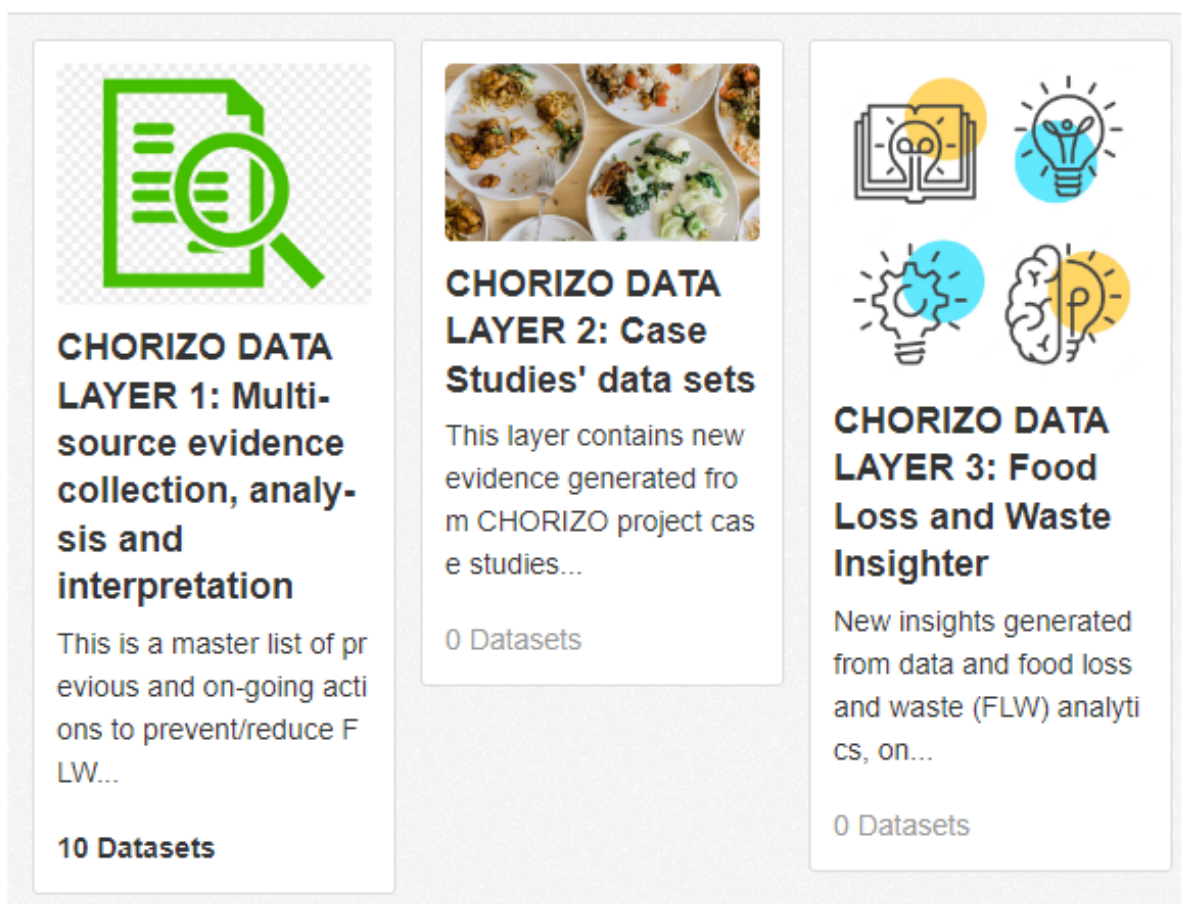- Data layer 3: The insights

Figure 7: CHORIZO Datahub and FLW "Insighter" layers

3.1.1. CHORIZO Datahub

Although the datahub attribute of the infrastructure touches all three layers, it is more relatable to the first two data layers.

Data layer 1: Multi-source evidence collection, analysis and interpretation. This layer contains previous and on-going actions to prevent/reduce FLW collected and assessed in terms of relevance. These actions tackle food loss and waste (FLW) in different countries of the European Union and the United Kingdom. Various stages of the food chain are tackled including:

- Primary production: the production, rearing or growing of primary products, including harvesting

- Processing and manufacturing including valorisation: The first processing and manufacturing of food after the primary production and before the retail and other distribution stage of the food supply chain

- Transportation: the transportation of food at any stage in the supply chain

- Retail: the handling of food and its storage at the point of sale or delivery to the final consumer - includes distribution terminals, shops, supermarkets distribution centres, wholesale outlets

- Redistribution: To redistribute surplus food fit for human consumption

- Food services: includes catering operations, factory and school canteens, institutional catering, restaurants, hotels, hospitals, cafes, and other similar food service operations

- Households: consumption of food in the household or small residential facilities

- General awareness-raising: focus on increasing overall, broad, general awareness about food loss and food waste - in the form of campaigns, forums, platforms, exchange of information/ideas; and

- The whole supply chain: addresses food loss and food waste along all stages of the supply chain.

Data in this layer are arranged in stages of the supply chain, searchable with relevant tags and will be open access with possibilities of accessing through APIs.

Data layer2: New evidence generated from the project case studies on the interaction of social This layer contains new evidence generated from CHORIZO project case studies on the interaction of social norms and food loss and waste behaviour. The case studies centred around:

- Case Study 1: Household food waste in and off crisis periods

- Case Study 2: Hospitality food waste

- Case Study 3: Food services food waste

- Case Study 4: School food waste and relation with obesity and malnutrition

- Case Study 5: Food waste in a food banks' mediated supply chain

- Case Study 6: Food waste in relation to date marking and sustainable and smart food packaging

Datasets in data layer 2 will be accessible to CHORIZO project partners for the execution of other project tasks, especially in the modelling work packages. For external stakeholders, the datasets in this layer will have visibility "private". Data sharing in this layer will be through a rules-based business model, yet to be developed.

3.1.1.1. Searching datasets

For data layer 1 and 2, datasets and resources (Table 3) can be conveniently searched and filtered by data layer, groups, tags, dataset formats and licences (Figure 8).

Table 3: CHORIZO datasets in data layer 1 and data layer 2

Type	Method	Data collection technique	CS / Task	Dataset name
Primary	Qualitative	FGI	CS1-BE	FGI.CS1-BE
Primary	Qualitative	FGI	CS4-DK	FGI.CS4-DK
Primary	Qualitative	IDI	T1.2	IDI.T1.2.Past.FLW.prevention
Primary	Qualitative	IDI	CS1-ES	IDI.CS1-ES
Primary	Qualitative	IDI	CS2-NO	IDI.CS2-NO
Primary	Qualitative	IDI	CS3-SI	IDI.CS3-SI
Primary	Qualitative	IDI	CS4-DK	IDI.CS4-DK
Primary	Qualitative	IDI	CS5-HU	IDI.CS5-HU
Primary	Qualitative	IDI	CS6-ES	IDI.CS6-ES.Consumers
Primary	Qualitative	IDI	CS6-ES	IDI.CS6-ES.Food.Industry
Primary	Qualitative	IDI	T4.1	IDI.T4.1.Challenges.changing.norms
Primary	Qualitative	IDI	T4.3	IDI.T4.3.Effectiveness.education.package
Primary	Qualitative	IDI	T6.2	IDI.T6.2.Communication.materials
Primary	Qualitative	workshop	CS6-ES	Wkshp.CS6-ES
Primary	Qualitative	workshop	T4.4	Wkshp.T4.4.Capacity.building
Primary	Qualitative	workshop	T6.2	Wkshp.T6.Dissemination
Primary	Qualitative	webinar	T4.4	Web.T4.4.Capacity.building
Primary	Quantitative	online survey	CS1-BE	Survey.CS1-BE
Primary	Quantitative	online survey	CS1-ES	Survey.CS1-ES
Primary	Quantitative	online survey	CS3-SI	Survey.CS3-SI
Primary	Quantitative	online survey	CS6-ES	Survey.CS6-ES
Primary	Quantitative	online survey	T4.1	Survey.T4.1.Challenges.changing.norms
Primary	Simulated	ABM	T3.3	ABM-sim.T3.3

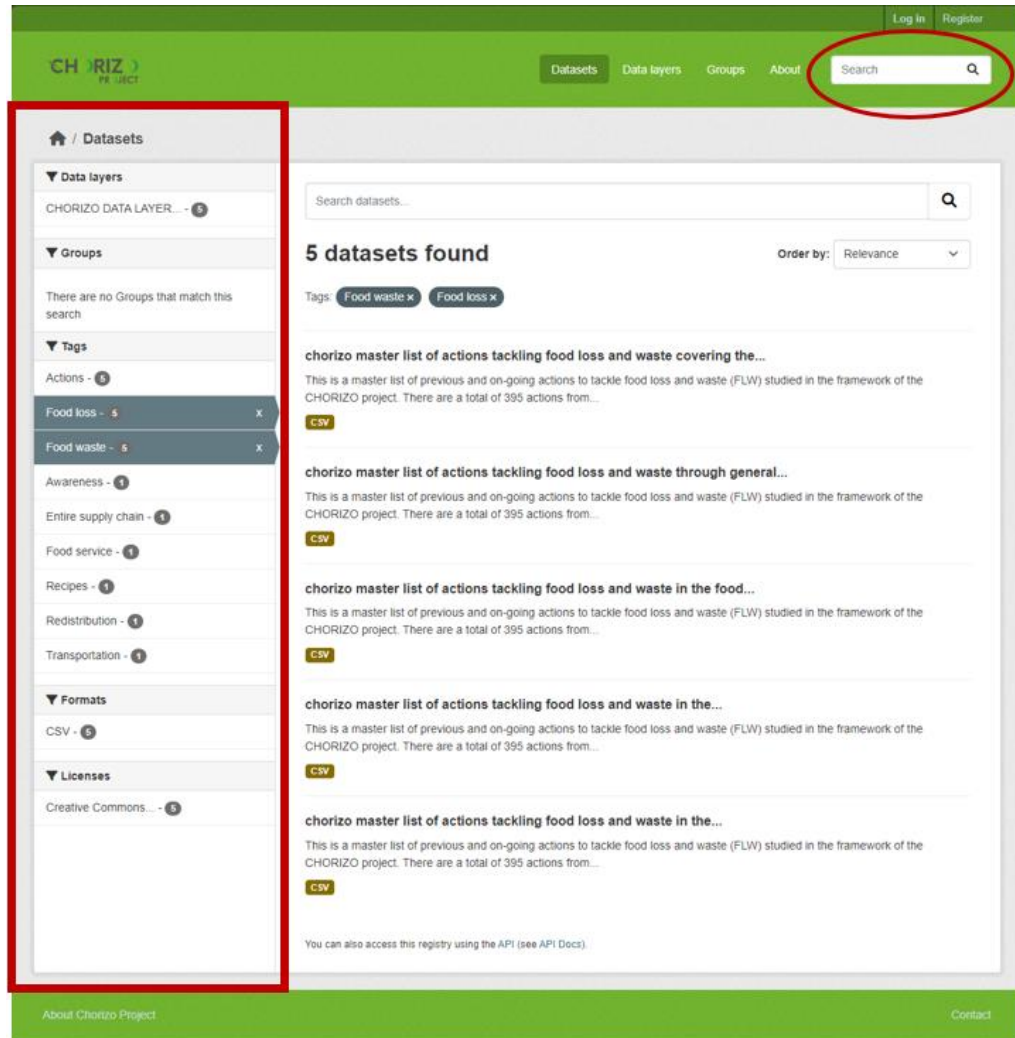| Secondary	Quantitative		T1.2	*395 actions on FLW findable by action name, stage in the food supply chain where the action was carried out

Figure 8: Searching datasets

3.1.1.2. Previewing and downloading data public datasets

Datasets in data layer 1, who's visibility is set to "Public" can be previewed through the "Explore" tab to see the content on a table. The columns and rows can also be filtered, with the possibility of produces line plots for quantitative datasets. These datasets can also be downloaded with different file formats possible (Figure 9; A and B). A login is not necessary for these actions on data layer1.
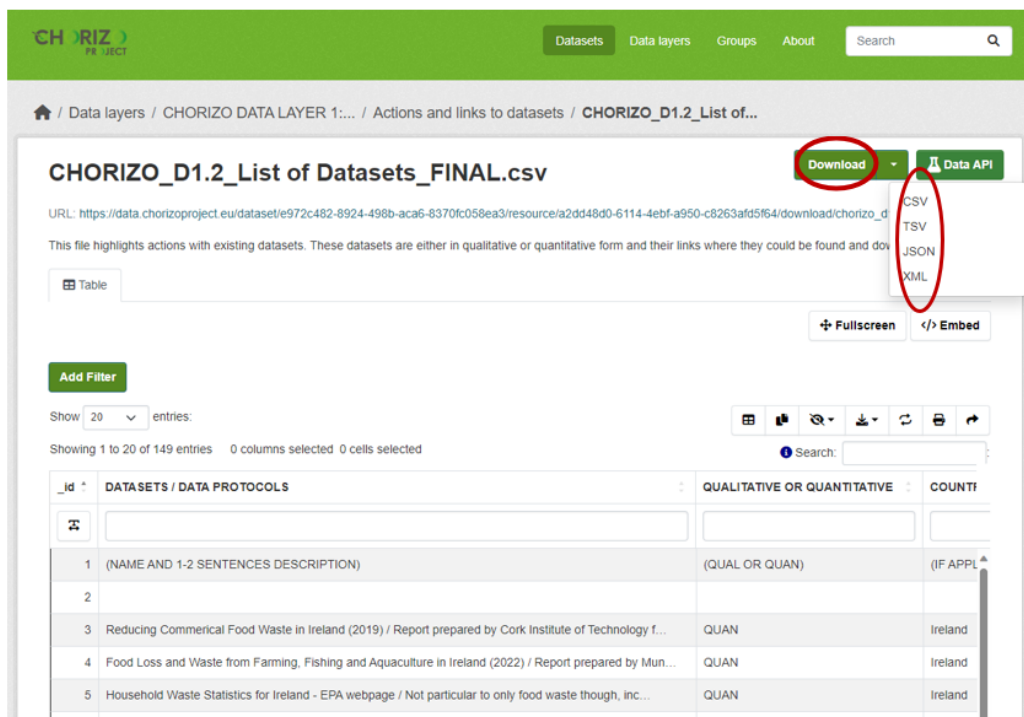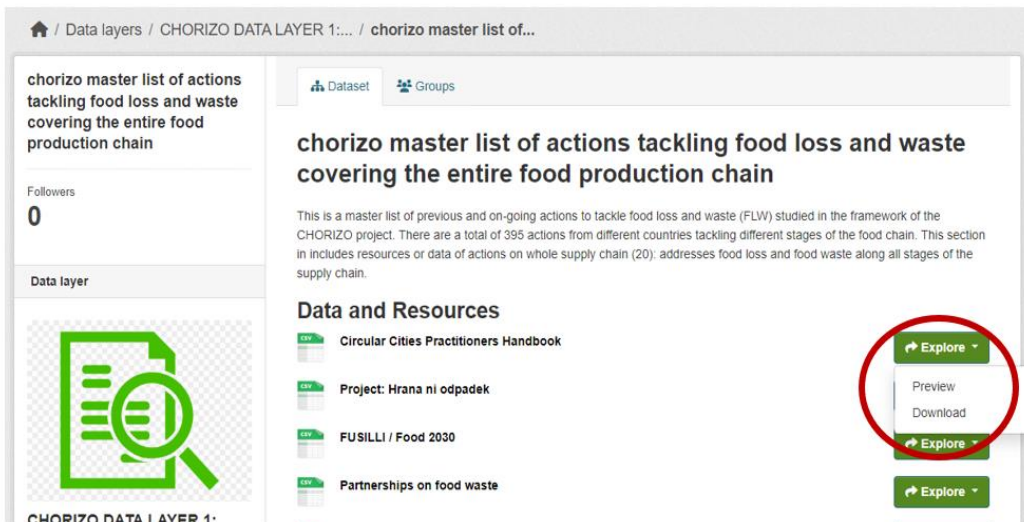
Figure 9: Previewing and downloading datasets

Accessing datasets through APIs

Datasets can also be accessed through APIs (Figure 10). The CKAN-based infrastructure automatically provides access APIs, and sample codes in various coding languages for external users. In the preview mode, at the top right corner, the tab "Data API" when clicked gives various API options. A user can select that which they are best knowledgeable with.
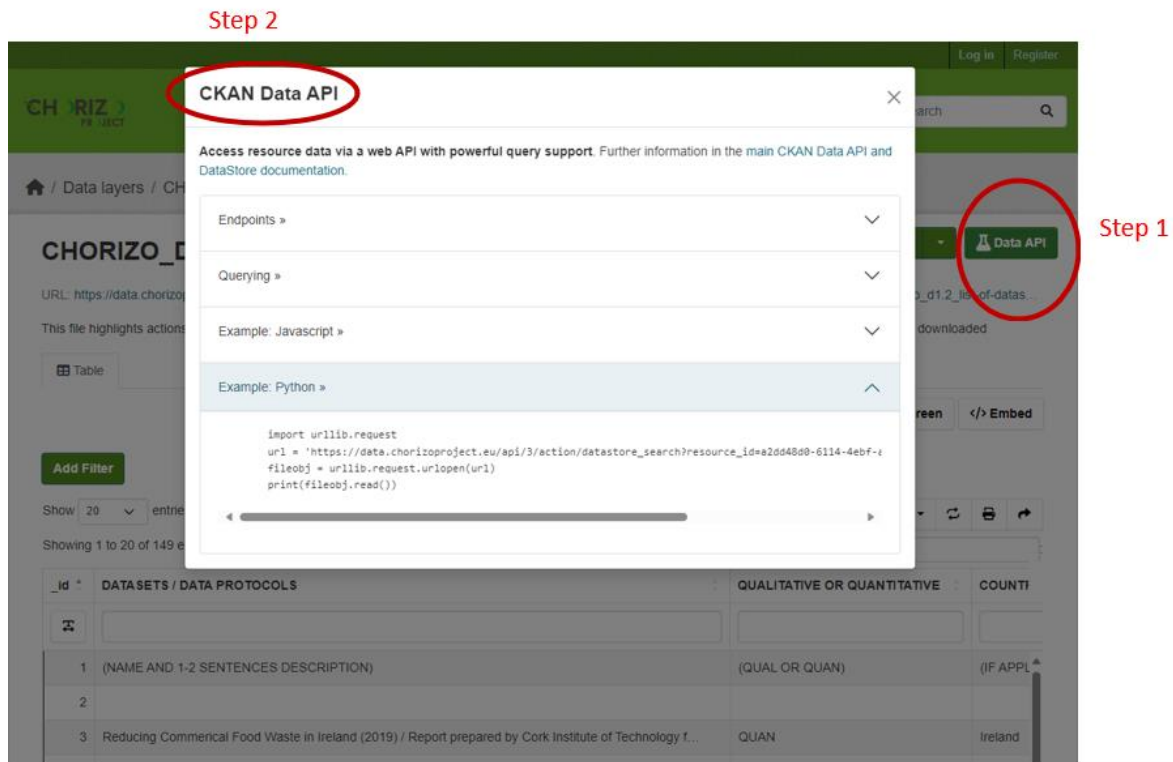
Figure 10: Accessing datasets using APIs

3.1.1.3.   Accessing "private" datasets

Datasets in data layer2 have a visibility status of "private". This implies they can neither be pre-viewed (on site exploration of content) nor downloaded by users from the public with no administrative role. They can however be seen as existing datasets. These manipulations are possible for the data owners within the project with administrative roles assigned. Access to these "private" datasets can be requested by contacting the site administrator by using the "contact" tab attributed to the "private" dataset. The requester will then be redirected to their mail page to send an informative email who's auto-generated subject contains the "private" dataset name asking access to these dataset. The site administrator will then channel this request to the dataset owner. A stakeholder agreement or business model pertaining to the specific "private" dataset can then be made, and the requester granted access.
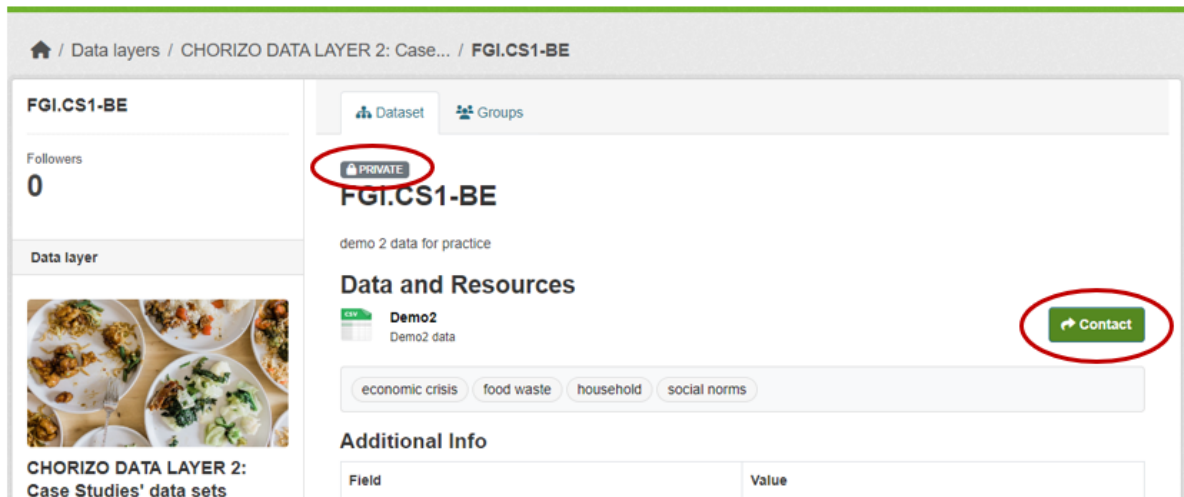
Figure 11: Requesting access to "private" datasets

3.1.2. CHORIZO FLW "Insighter"

The third layer of the infrastructure is the "Insighter" layer. New insights generated from data and food loss and waste (FLW) analytics, on the levers to change social norms. This is a service provision layer to the relevant corporate actors pursuing FLW prevention activities. As of this moment in the projects, the WPs and associated tasks involved in insights generated just started.

3.2. Impact and sustainability of the CHORIZ0 datahub and FLW "Insighter"

The impact and sustainability of the CHORIZ0 FLW datahub and "Insighter" could be seen in two directions; within the project for project partners and outside the project, taking into consideration all other stakeholders.

Within the project, project partners would benefit from the insights generated from the empirical data collected the project's case studies. Knowledge and lessons learned through out the process of data collection, pre-processing and storage on the datahub as well will improve future tasks for project partners.

Outside the project, external stakeholders will have a one stop shop to explore knowledge on previous and on-going actions on FLW undertaken in various stages of the food supply chain to mitigate FLW. The relevance and impact of these actions to external stakeholders will be measured with a 7-likert scale survey strategically placed in the first data layer of the datahub (Figure 12).

Figure 12: FLW actions database evaluation survey

On the sustainability, other previously developed tools, websites and data portals, and on-going projects in the domain of FLW are crucial. It could be interesting if they become linked by some means (such as exposure of datasets through APIs: Application Programming Interfaces) with the CHORIZO FLW datahub / "Insighter" tool. This is important to further bring datasets and knowledge together, making it easily accessible for stakeholders. For example, the European Commission's FLW Prevention Hub[3] and the DjustConnect platform[4] amongst others, are of future interest to the CHORIZO FLW datahub / "Insighter" . Within the project, discussions on the future governance, exploitable components of the datahub and "Insighter" have begun. The experiences gained from the DjustConnect platform which is already operating and led by EV-ILVO, will provide a starting point for the governance model. Figure 13 and 14 below gives insights to initial thoughts on governance and exploitable components respectively.

---

[3] European Food Loss and Waste Prevention Hub - Home (europa.eu)
[4] DjustConnect
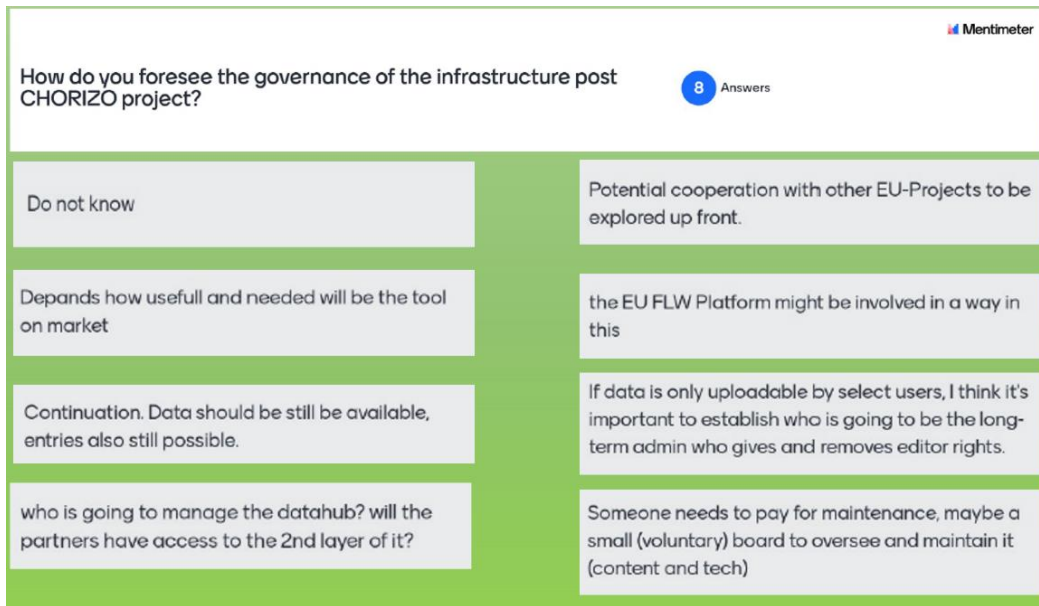
Figure 13: Governance ideas for the datahub / FLW "Insighter"
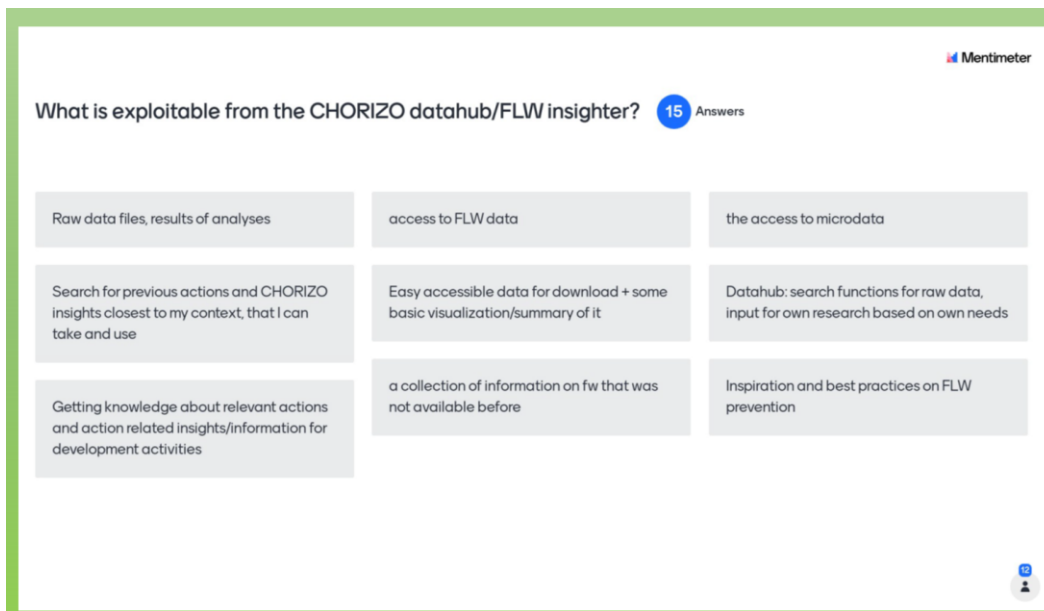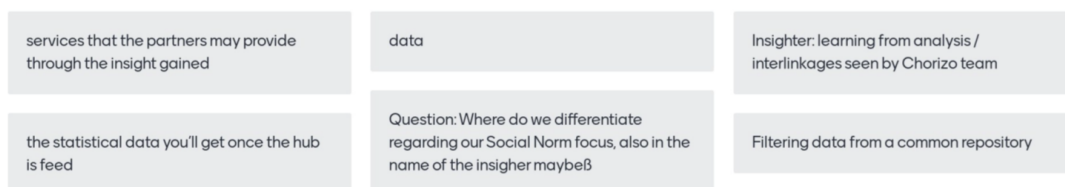

Figure 14: Exploitable components of the datahub / FLW "Insighter"

4. **Conclusion**

D2.2- CHORIZO FLW "Insighter" in this document mostly referred to as CHORIZO FLW datahub and "Insighter" is a deliverable from T2.2: CHORIZO FLW "Insighter" datahub development and validation. The development phase of the infrastructure ran from M4 to M12 of the project and is a result of an iterative development process with project partners. The FLW datahub and "Insighter" has been developed to store;

- The results of the evidence-based analysis of previous/ongoing FLW actions (data layer 1),
- The new empirical evidence on social norms, and consumer, organisation and charity (e.g. food banks) behaviour in relation to FLW, generated by the project case studies (data layer 2), and
- The contextual insights generated from analysis and modelling activities in the project (data layer 3).

Exploitation and access to the resources and datasets in the various data layers of the FLW datahub and "Insighter" are different. The results of the evidence-based analysis (previous/ongoing FLW actions) will be "open" to all interested actors. The new anonymised and/or aggregated evidence generated by the CSs being "'private" will be available through an appropriate stakeholder engagement and datahub governance model yet to be developed. Nonetheless, the experiences gained from the DjustConnect platform governance model will be exploited. It will ensure transparent and rules-based sharing of FLW datasets.

The immediate and future impacts of the FLW datahub and "Insighter" are enormous. Within the project, insights from generated data which will be readily available for project partners will positively impact their business processes and provide more knowledge. Outside the project, the access and use of the datahub and "Insighter" by stakeholders will be easy and intuitive. The applicability of the datahub and "Insighter" to facilitate decision making, will be validated by these external users themselves as they provide feedback for improvement.

Looking at the bigger picture, and the existence of other previously developed tools, websites and data portals, and on-going projects in the domain of FLW, the CHORIZO FLW datahub and "Insighter" stands a greater chance of "surviving" after the project lifespan. The possibilities

of integrating this infrastructure to others, moving towards a FLW dataspace at local, national and even regional levels will only go a long way to make data more accessible.

5. Appendices

5.1. Creating an account

- Step 1:- Click on "Register"

- Step 2:- Provide the required information on the page Register for an Account

- Step 3:- Click on Create Account
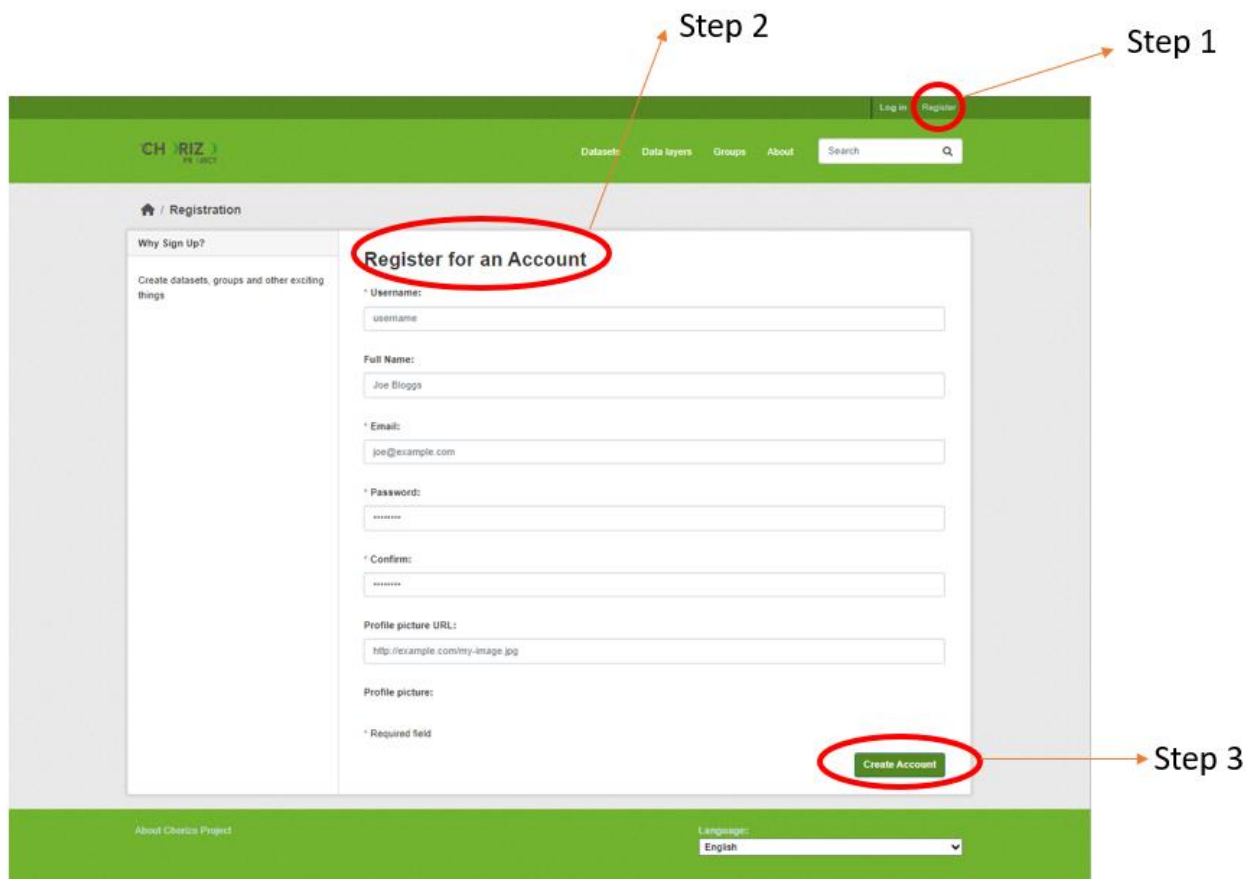
5.2. Searching resources and datasets

- Step 1:- From home page through the "Search" bar using keywords
- Step 2:- From within the "datasets" or "data layer" tabs. The layer gives results (datasets or resources) specific for the selected dataset
- Step 3:- For more defined or finer search using various filters such as data layer, groups, tags, dataset or resource format, and licenses.
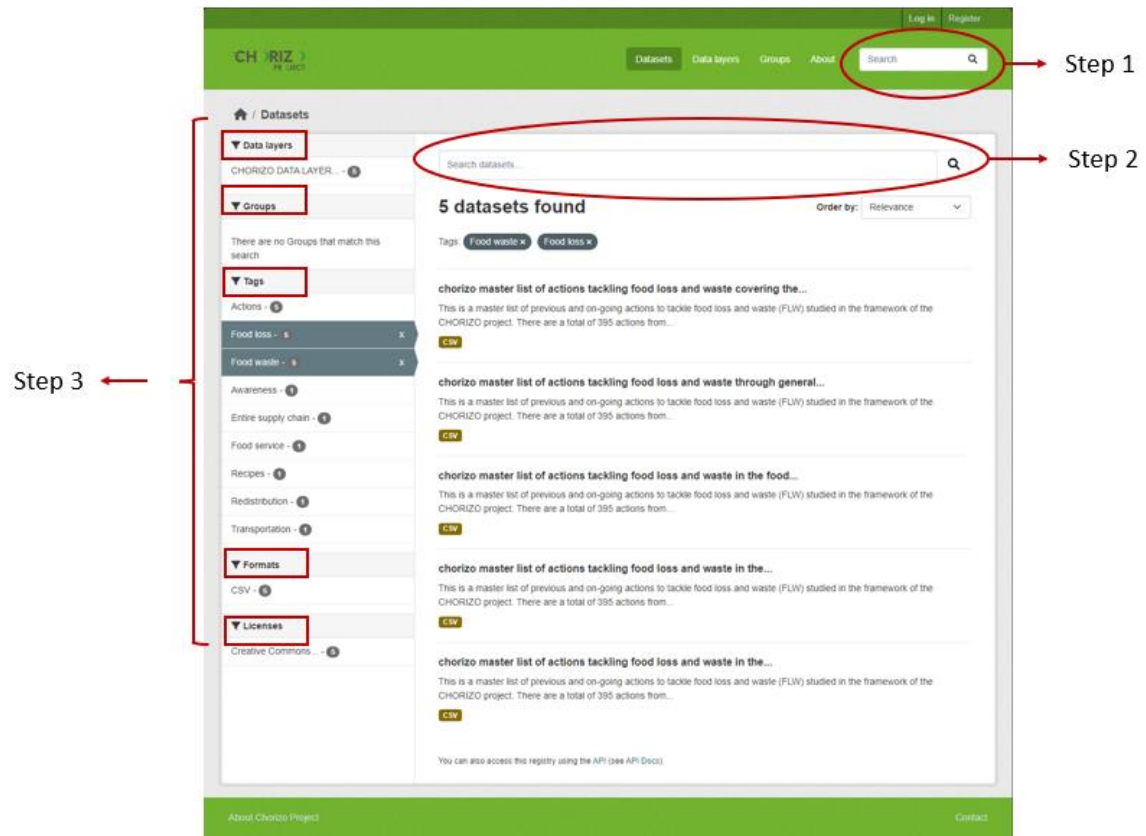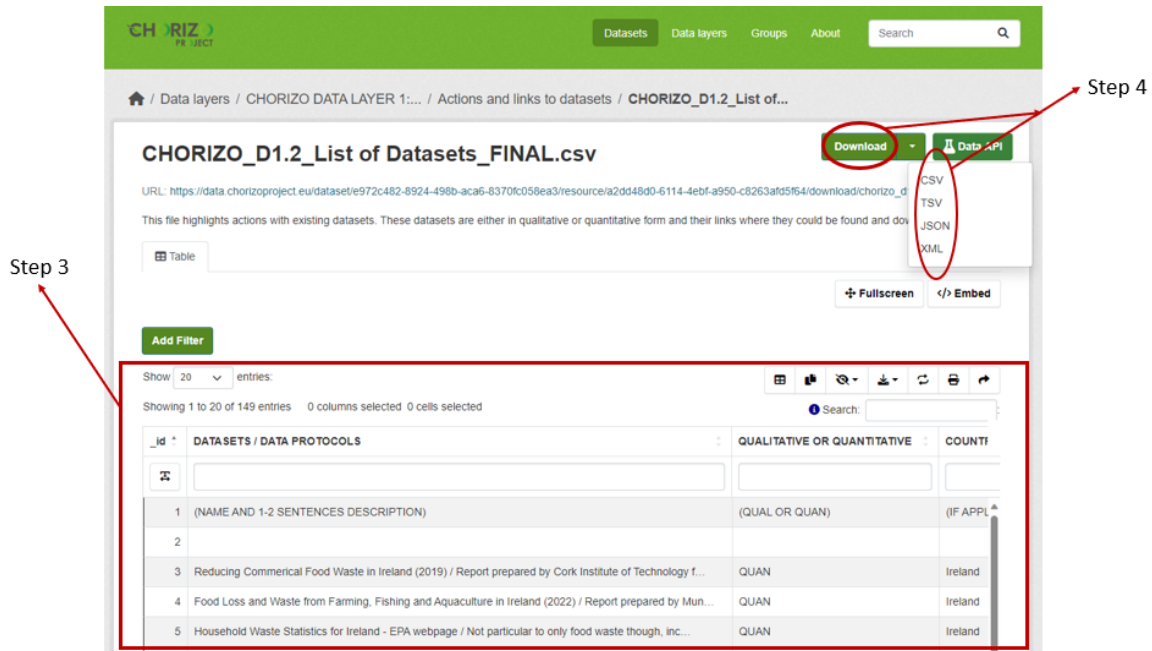
## 5.3. Accessing public datasets (previewing and downloading)

- Step 1:- Click on the desired dataset

- Step 2:- Right to the preferred dataset, click on "Explore".
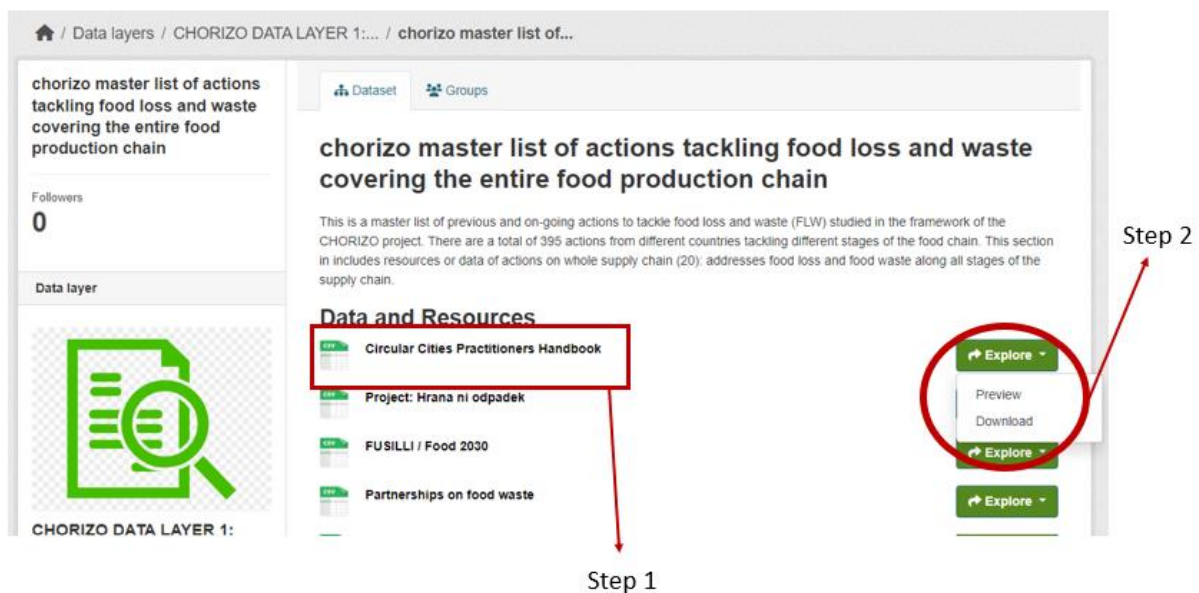
- Step 3:- Preview selected, explore the rows and column entries

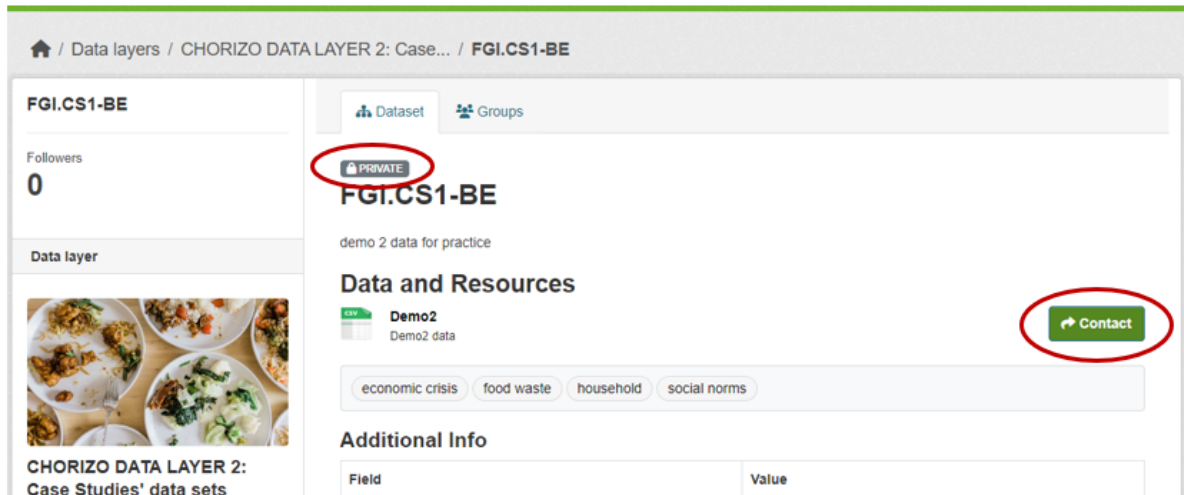- Step 4:- Download selected, select preferred format to download.

- Step 6:- Access through API access by clicking on DATA API and

- Step 7:- select preferred coding language to consume dataset or resource.
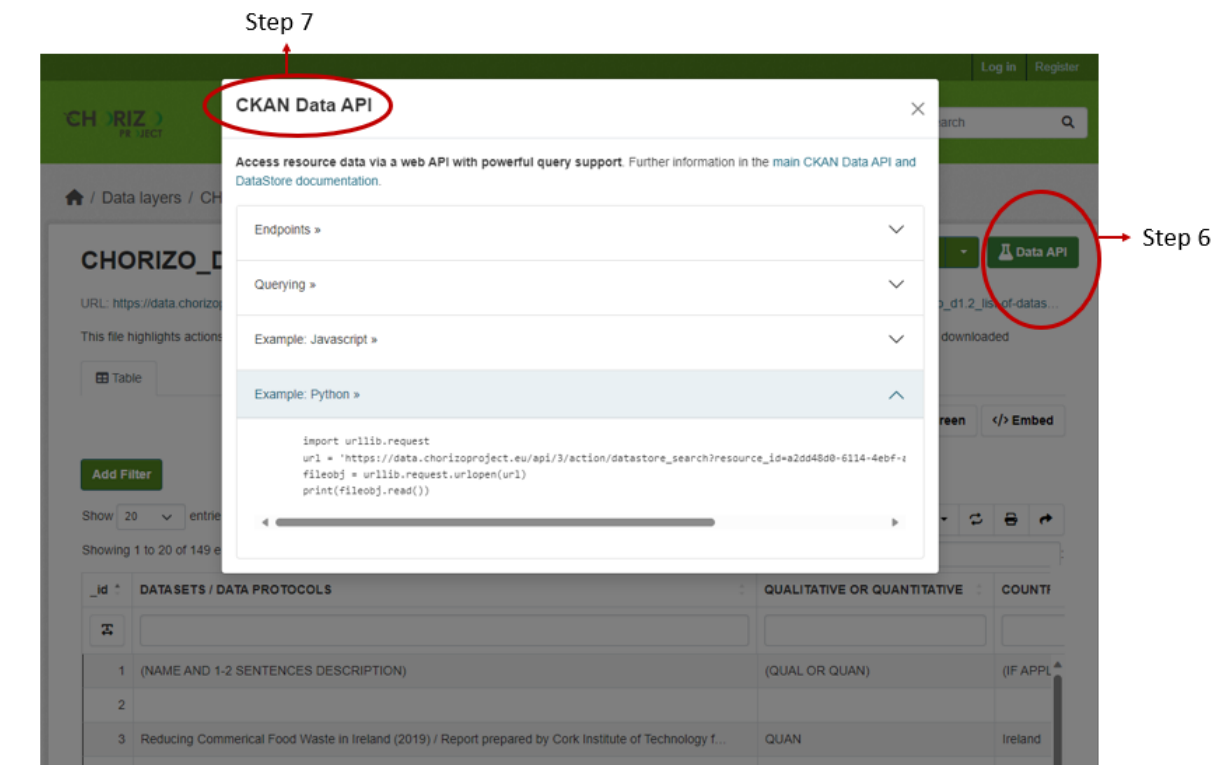
5.4.   Accessing private datasets

- Search data set

- Request dataset by clicking on "Contact"

5.5.   Uploading dataset

- *Requires an account and login

- *Data creation and manipulation can only be done within a data layer

- Administrative role roles impact this functionality.

- Step 1:- Click "Add dataset" in the data layer where you can want to create, upload or add a dataset.
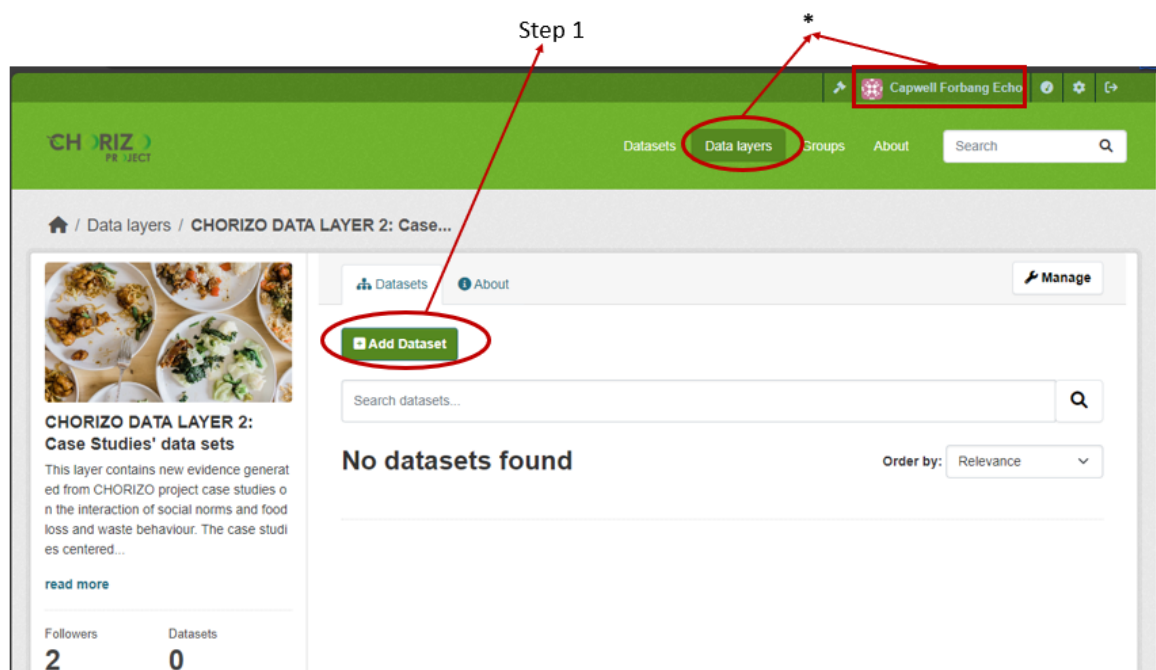
- Step 2:- Provide the required information for the dataset such as the title, description, tags, licence, assigned data layer, visibility, version, author, author email, dataset maintainer, maintainer email, and custom field (key and value to specify more meta data for the dataset). Click Next: "Add Data" at the end.

.

.

.

- Step 3:- Upload file from your local environment / PC or add link if dataset is located elsewhere on the web. A more recognizable name and description could be given to the dataset/resource. The format could be left unfilled or unselected as this is automatically detected by the system.
- Step 4:- Optional:- Click on "Save & add another", to add other closely related datasets or more information in case of qualitative data, this could be uncoded transcripts from in-depth or focus group interviews.

- Step 5:- Click "Finish" to complete the process and have the dataset uploaded and saved.
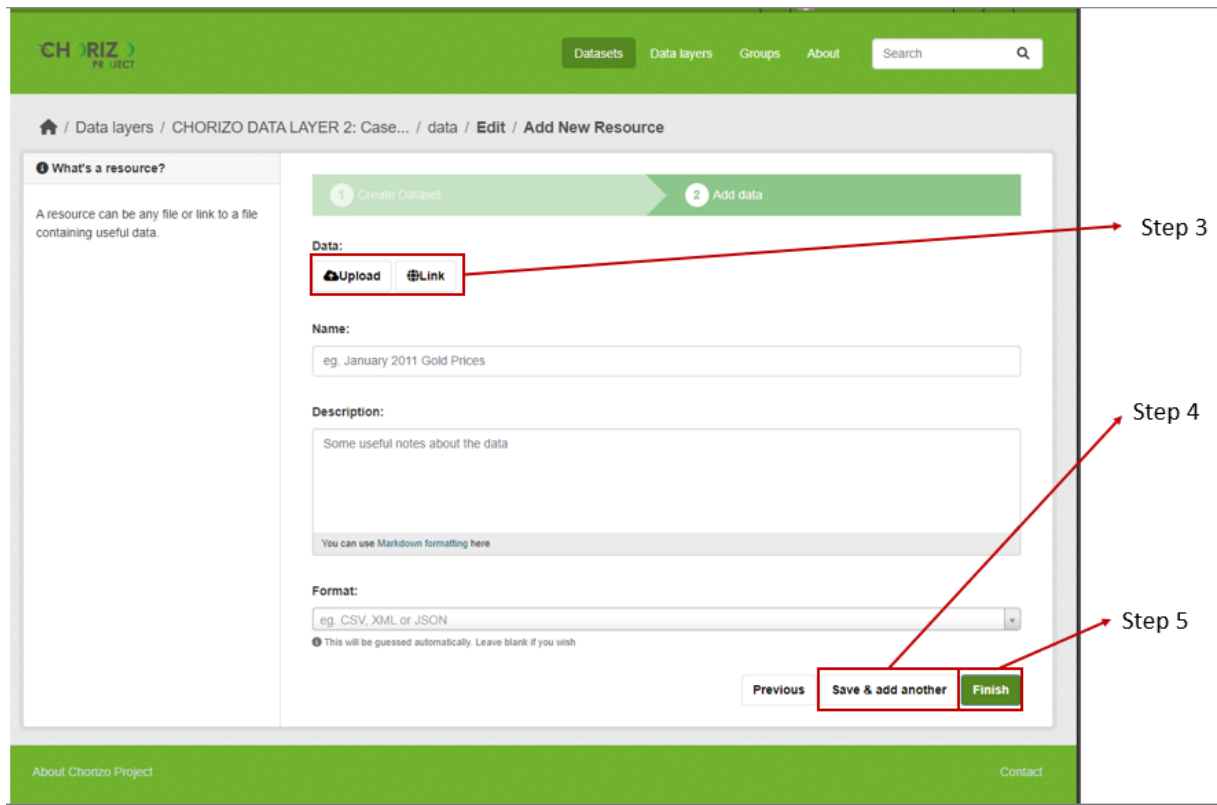
5.6. Managing datasets and resources

- *Requires an account and login
- *Data can only be done within a data layer
- Step 1:- Information of already existing datasets can be updated, version changed and more resources added. This can be achieved by clicking on "Manage" tab associated with the dataset.
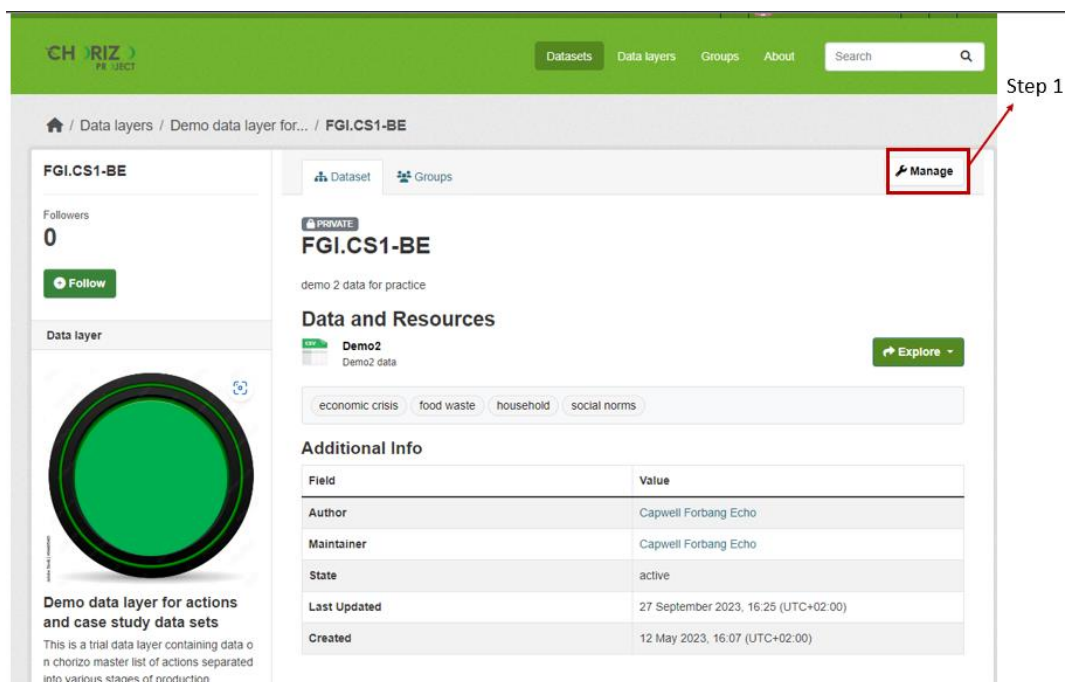
- Step 2:- The dataset can be deleted.

- Step 3:- Save desired changes made to the metadata.

.

.

.

5.7. Datahub information collection tool for partners

Information collection sheet for CHORIZO FLW Insighter

Dear CHORIZO partners,

WP2- T2.2: CHORIZO FLW 'Insighter' datahub development & validation (Lead: EV-ILVO; Participants: VLTN, ICP, CSCP) (M4-M12), would like you to provide answers to the following questions as they may fit your context. The questions are open, and we will appreciate to get your thoughts and practical possibilities as much as possible. This is very crucial to consider as we will design the datahub.

We will be collecting this information untill Friday 20th January 2023.

Thank you very much in advance.

WP2-T2.2

* Required

1. Where do you belong in the CHORIZO project? *

☐ Only Case Study (CS)

☐ Only Work Package (WP)

☐ Both CS and WP

2. Select CS number *

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ 6

☐ NA

3. Select WP number *

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ 6

☐ 7

☐ 8

☐ NA

4. In what capacity will you use the CHORIZO FLW datahub/Insighter? *

☐ Data provider

☐ Data consumer

☐ Both data provider and consumer

☐ Other

5. As a CS or data provider, how do you intend to expose your data or make your data available to the datahub? In order words, what are your technical possibilities of sharing data out of your organisation? *

Enter your answer

6. As a WP (all WPs interacting with the datahub can answer here) or data consumer, how do you foresee acessing data from the datahub? In order words, what are your technical possibilities of accessing data out of your organisation? *

Enter your answer

7. How do you foresee the future of the datahub after the project? *

Enter your answer

Submit

5.8. Feedback on datahub and FLW "Insighter" from partners

5. As a CS or data provider, how do you intend to expose your data or make your data available to the datahub? In order words, what are your technical possibilities of sharing...

12 Responses

ID ↑	Name	Responses
1	anonymous	NA
2	anonymous	I do not know yet
3	anonymous	Via the data repositories requested in the Data Protocol, but looking for guidance on how should submit data (i.e. format) to the project's datahub.
4	anonymous	We intend to create awareness about the fact that data are available. We see ourselves as facilitators
5	anonymous	We can/will share data in MS Office formats (Excel, Word)
6	anonymous	As we have a strong ICT department and close cooperation with the Faculty of Electrical Engineering and computer science and are currently developing data space, we can share data in numerous ways.
7	anonymous	N/A
8	anonymous	No idea yet
9	anonymous	We will mostly focus on producing high quality summary statistics from the experiment results and provide those to the data insighter. The kind of summary statistics available will depend on exactly what statistical operation will be performed, but we expect to know relatively soon. We will also provide anonymized summaries of the in depth interviews for further dissemination and use by other work packages
10	anonymous	Google Drive or microsoft teams
11	anonymous	Not a data provider
12	anonymous	- We have already shared resources in the EU hub and we would be willing to continue doing so. But this was rather qualitative so far. If the accessibility of the data hub is easy and not too time consuming, we could potentially also think about different kinds of contributions. However, perspectively not too big datasets. - We usually use manual inputs. Automations would not be possible in our case.

6. As a WP (all WPs interacting with the datahub can answer here) or data consumer, how do you foresee acessing data from the datahub? In order words, what are your technical...

12 Responses

ID ↑	Name	Responses
1	anonymous	if data is freely available, we can access it
2	anonymous	I do not know yet
3	anonymous	N/A as primarily a data provider.
4	anonymous	We foresee that data becomes available in an easy2use manner so that we and others can access Them for research ir policy purposes
5	anonymous	We don't have any specific technical possibilities of accessing data
6	anonymous	As we have a strong ICT department and close cooperation with the Faculty of Electrical Engineering and computer science and are currently developing data space, we can share data in numerous ways.
7	anonymous	N/A
8	anonymous	No idea yet
9	anonymous	As a work-package we expect to be able to access anonymized summaries of the in-depth interviews as well as general summary statistics of the various survey results. This will help with the modelling and to make sure the insights are consistent across case studies.
10	anonymous	Google Drive or microsoft teams
11	anonymous	The exact purpose and content of the datahub is still unclear to us - in any case, actors from the food chain and especially local authorities would benefit to learn from it, to use knowledge collected to inform policy making.
12	anonymous	- "Normal" IT skills and possibilities - No need for specific dataset-access - Interesting could be to learn about feasibility, reach and effectiveness of FLW activities and how to replicate, scale - Accessing the data hub on a website (e.g. a Chorizo project website) or on an open-access data hub (Zenodo) - When making the data publicly available on a Chorizo website, we could use a registration procedure similar to the Manifesto database or others, where one is asked what one will plan to use the data for and what ones profession is, when signing up. This could help us to undertstand the dissemination pattern and reach.

7. How do you foresee the future of the datahub after the project?

12 Responses

ID ↑	Name	Responses
1	anonymous	did not think about this yet
2	anonymous	I do not know yet
3	anonymous	Hope that it is clear, easily accessible, intuitive to use for all stakeholders, and that it can be incorporated into "datahubs" of the European Commission.
4	anonymous	Good question. My best answer would be that we should make provisions so that one of the food Resarch Infrastructures would take over the responsibility
5	anonymous	We don't know it exactly but we certainly see the datahub as a tool for getting knowledge for our future development activities
6	anonymous	Researchers and actors in the food supply chain can use data for their activities and projects.
7	anonymous	CHORIZO Insighter could be an essential tool for addressing the global issue of food loss and waste and a significant resource for organizations, governments, and individuals to understand and address the causes of food loss and waste as related to social norms. As technology continues to advance and data becomes more easily accessible, it will allow for i) more accurate tracking and reporting of food waste and loss, ii) the development of targeted solutions to reduce it considering behavioral changing. Additionally, the datahub could be integrated with other technologies, such as blockchain and IoT, to further improve data accuracy and transparency. Ultimately, the datahub could play a critical role in reducing food waste and loss contributing towards zero food waste. Overall, the future of the CHORIZO datahub seems promising (depending on its ability to adapt to the needs of the FLW market) as there is a raised demand in data driven decision making, thus data sharing solutions is likely to increase.
8	anonymous	No idea yet
9	anonymous	It will be useful to keep accessing it to make sure future models remain validated against the data and insights produced here.
10	anonymous	it will be usefull to compare data in similar studies
11	anonymous	Open access?
12	anonymous	- It would be important to understand early on how to create synergies with the EU platform hub, what expectations the EU also has on us integrating our work there, what "change / adaptation" rights we might get, etc. or if we are bound to their structure - The data should continue being publically available, so future