

Towards European Agriculture Data Spaces: A Proof of Concept in the Pig Sector

Roberto García

Universitat de Lleida

<https://rhizomik.net/~roberto>

What is a Data Space?

- First reference in 2005 by Franklin, Halevy and Maier [1], a **shift from central databases** to storing data at the source
 - A Data Space is a **decentralised**, open infrastructure for **sovereign** data exchange whose participants are **aware** and in **control** of the data they produce and consume as well as the involved services
 - No central data store, data exchange occurs directly between participants
- Part of the European Strategy for Data, proposal for **9 Common European Data Spaces** (CEDs):
 - health; industrial & manufacturing; **agriculture**; finance; mobility; Green Deal; energy; public administration; and skills

[1] Franklin, M., Halevy, A., and Maier, D. (2005). From databases to dataspace: a new abstraction for information management. ACM SIGMOD Record, 34(4):27–33

What do Data Spaces provide?

- Overcome **barriers** to **data sharing** across organizations by combining necessary tools and infrastructures and addressing issues of **trust**
- Data sovereignty:
 - Ensure stakeholders do not get exploited in the competition for their data
 - Protect how one's data is used and request for fair conditions
 - For states, organizations and also individuals

How to motivate data sharing ?

- Data-monetization (tokenization) as a data sharing incentive
- Requires an environment I can trust to share my data in, without losing its control
 - *"tragedy" of the digital*, so easy to copy and disseminate...
 - Example: digital music → Digital Rights Management (DRM),... streaming
- For data... Data Self-Sovereignty by Design?

Data Self-Sovereignty by Design

- Your data:
Never “leaves the room”
- You:
In control on who “enters” and
what they “take out” with them

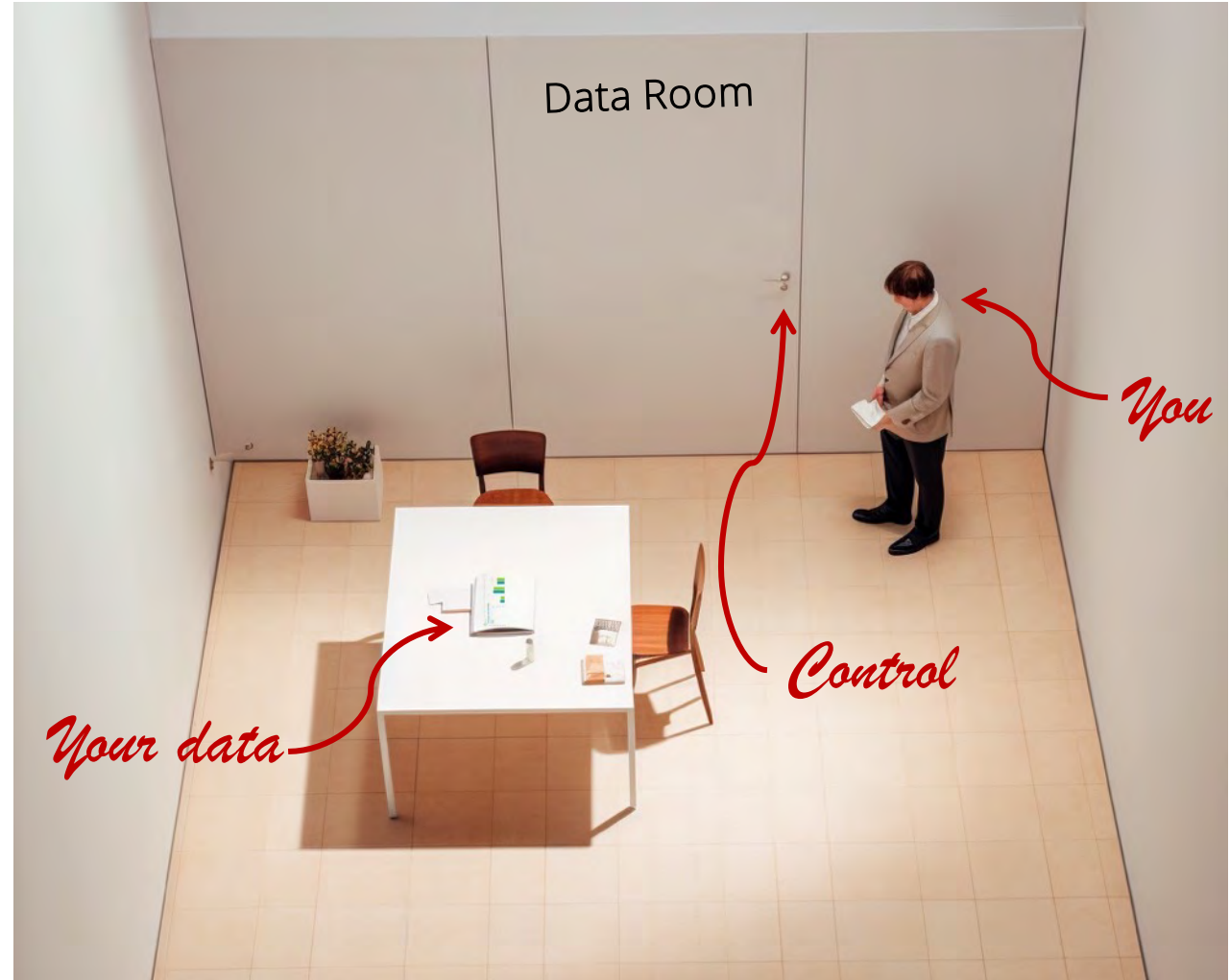
Full or partial copies

Aggregated data

AI trained models

Personal data

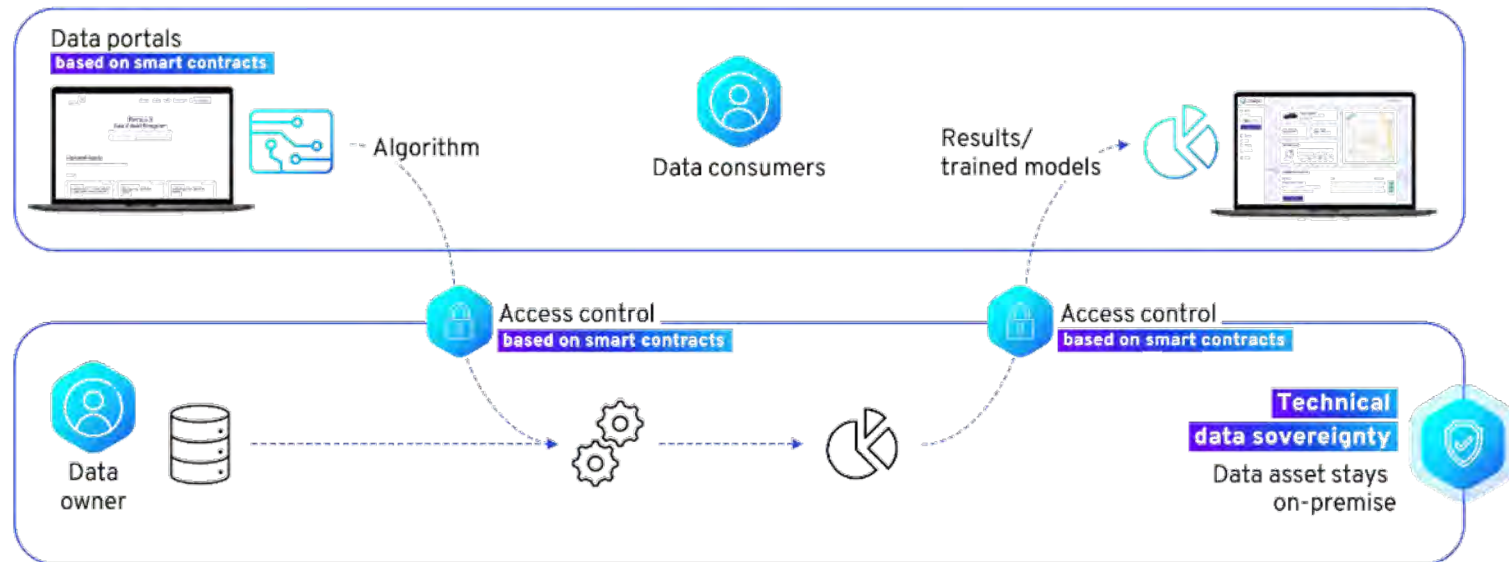
- ...



dar01995_a_bird-view_of_the_inside_of_a_whole_room_without_wind_1b2e5c91-5513-471d-9673-6c6493a76c11

Compute-to-Data

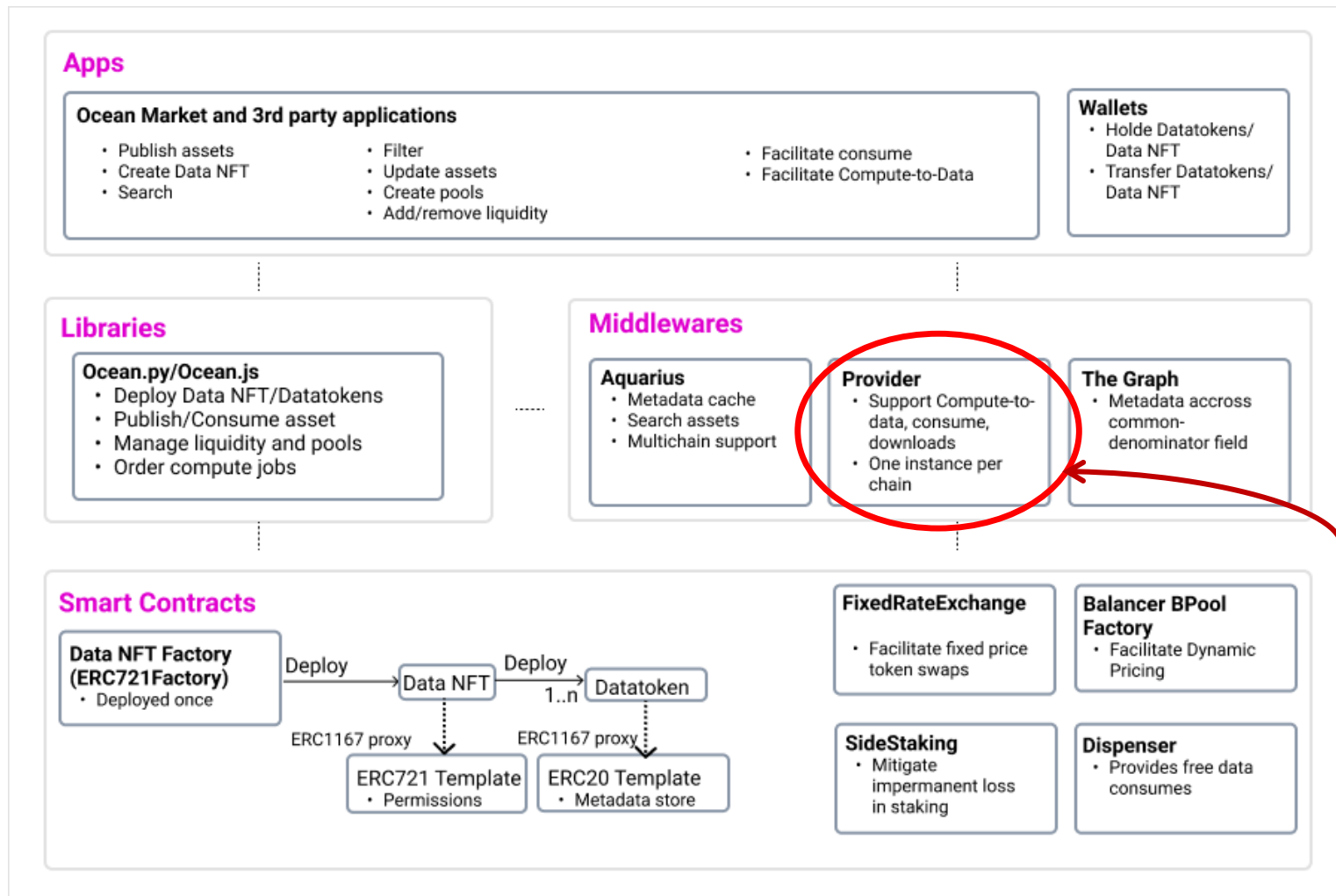
- Facilitates privacy and Data Sovereignty by design
 - Combinable with Federated Learning, Compute-to-Edge,...
- Algorithms are executed where the data is stored
 - Data originators keep **control** of data, **not copied** anywhere else



Pontus-X: Gaia-X Web3 Ecosystem



Pontus-X: built on Ocean Protocol's technology stack



Proof of Concept

- Using Pontus-X Gaia-X Web3 stack
 - Objective:
Data Sovereignty by Design + Tokenisation
 - Building blocks:
 - Compute-to-Data (CtD)
 - Trusted Execution Environments (TEE)
 - Blockchain
 - Self-Sovereign Identity (SSI)
- Control
- Ownership & Transparency

Gaia-X Web3 Stack by deltaDAO

<https://dataspace.angliru.udl.cat>



UdL Research Data Space

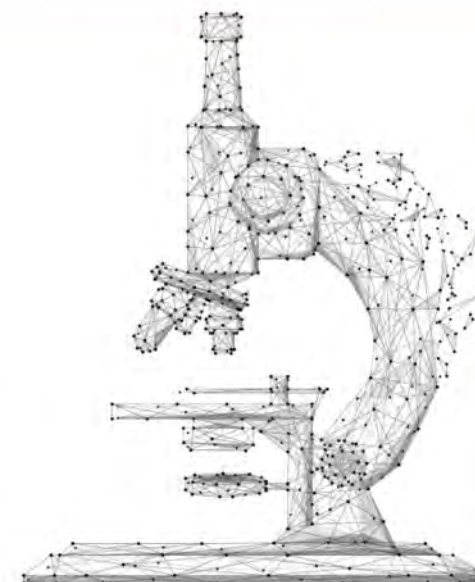
An open research platform following the FAIR guiding principles of findability, accessibility, interoperability and reusability, while guaranteeing data sovereignty.

Search for service offerings

Recently Published

CEP - Automatic Pig Feeding - 2021 S1 - 982091062894374 0x6207...6822 Automatic pig feeding data collected at the Centre for Pig Studies of Catalonia (Centre d'Estudis...) Free	CEP - Automatic Pig Feeding - 2021 S1 - 982091062894348 0x6207...6822 Automatic pig feeding data collected at the Centre for Pig Studies of Catalonia (Centre d'Estudis...) Free	CEP - Automatic Pig Feeding - 2021 S1 - 982091062894196 0x6207...6822 Automatic pig feeding data collected at the Centre for Pig Studies of Catalonia (Centre d'Estudis...) Free
RML Mapper for Daily Pig Weight CSV Data 0x6207...6822 This algorithm supports the "Pay-as-you-go" approach when sharing data through a data ... Free	CEP - Automatic Pig Feeding - 2021 - Pens 4 & 5 0x6207...6822 Automatic pig feeding data collected at the Centre for Pig Studies of Catalonia (Centre d'Estudis...) 1 EUROe	CEP - Daily Pig Weight - 2021 - Pens 4 & 5 0x6207...6822 Daily pig weight data collected at the Centre for Pig Studies of Catalonia (Centre d'Estudis...) Free

ALL DATASETS AND ALGORITHMS →



Goal

The Universitat de Lleida Gaia-X Web3 Ecosystem Portal aims to accelerate the progress of research and science, increase the quality of data, and foster collaboration between research and educational institutions. On the research portal, data scientists and researchers can find, consume and share research data sets and algorithms. Depending on the publisher's decision, the data assets can be consumed either by granting compute or download access to the data assets.

Data Audit Trail for Academia

By tokenizing data assets and using smart contracts to manage access control, each transaction, which includes publishing, sale, exchange, and consumption of data services, is logged automatically and immutable on the distributed ledger. This creates a data audit trail for academia which allows research institutions to monitor the use of their assets and gain attribution for their work.

Data Sovereignty by Design

One core concept of the UdL Research Data Space is the **Compute-to-Data (CtD)** approach. Compute-to-Data is the functionality that solves the current trade-off between the benefits of using private data and the risks of exposing it. It allows data consumers to run compute jobs on private data while the data stays on-premise with the data provider, who retains control.



AI4Pork Use Case



Precision Feeding



Temp., Hum., CO2,
Ammonia,...



Cameras



Feed and
Weight

Environment

Videos

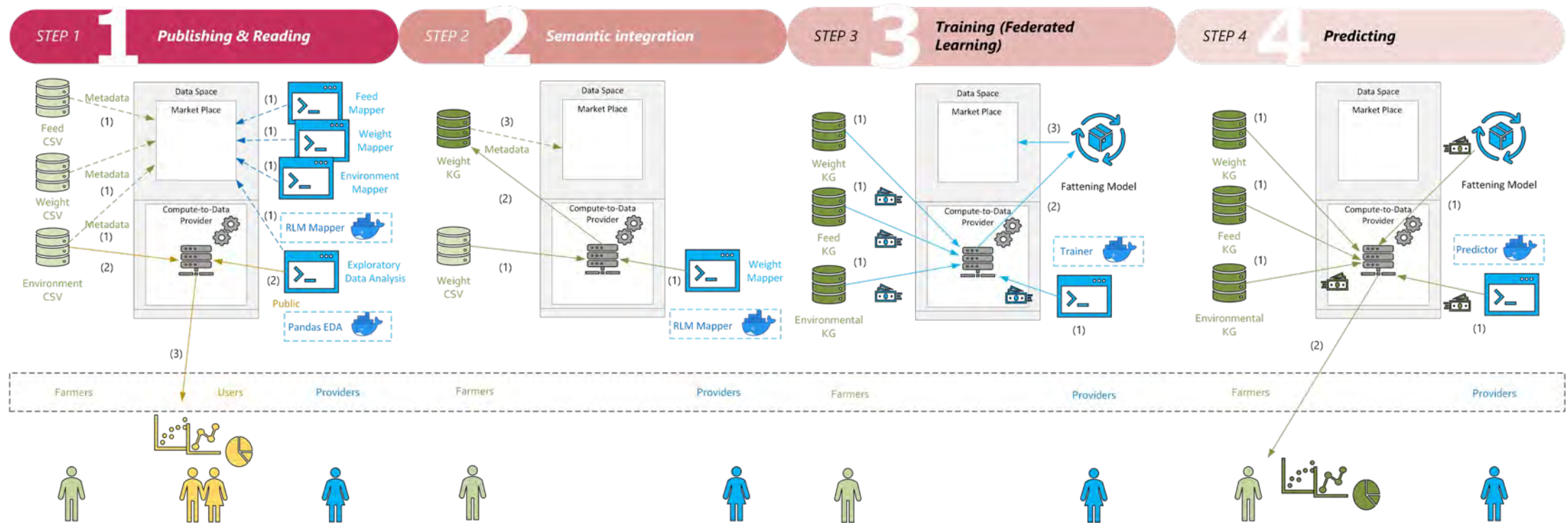


Centre of Swine
Studies of
Catalonia

Compute-to-Data Provider
at Universitat de Lleida



AI4Pork Use Case



CEP - Automatic Pig Feeding - 2021 S1 - 982091062894196

GEN-X Testnet

Owned by Ox6207...6822

Accessed with CEP-2021-S1-FEED_982091062894196

DATASET Published 4 days ago

Automatic pig feeding data collected at the Centre for Pig Studies of Catalonia (Centre d'Estudis Porcins), a consortium made up of the Diputació de Lleida, the Regional Council of La Noguera, the Torrelameu Town Hall and the Universitat de Lleida.

The feeding data includes:

- **Pen Number:** the pen the pig is located in.
- **Animal ID:** the pig identifier.
- **Date:** the date the feeding data is about.
- **Time:** the time pig feeding happened for the given date.
- **Duration (s):** the duration of the feeding event.
- **Feed (g):** the amount of food provided, measured in grams.
- **Animal Weight (g):** the pig weight, measured by the feeding machine during each feeding event.

udl pig agriculture livestock feed automatic-feeding

DATA AUTHOR

Centre for Pig Studies of Catalonia

OWNER

Ox6207...6822

DID

did:op:0d9493e7c7be9afdd6119a0099b54347ab053201d04a49ac1090deee8e73e041

Binary
56.98 kB
url

Free

Select an algorithm to start a compute job

Search by title, datatoken, or DID...

Exploratory Data Analysis

GXAT | did:op:e73ac9c332adfea204acefb4b300e...

Free

RML Mapper for Daily Pig Weight CSV Data

GXAT | did:op:f421d1f1824262a7a61e42370f65cc...

Free

You will pay

BUY COMPUTE JOB

You bought this dataset already allowing you to use it without paying again. The C2D resources required to start the job are available, no payment is required for them.
Please note that network gas fees still apply, even when using free assets.

Your Compute Jobs

SHOW

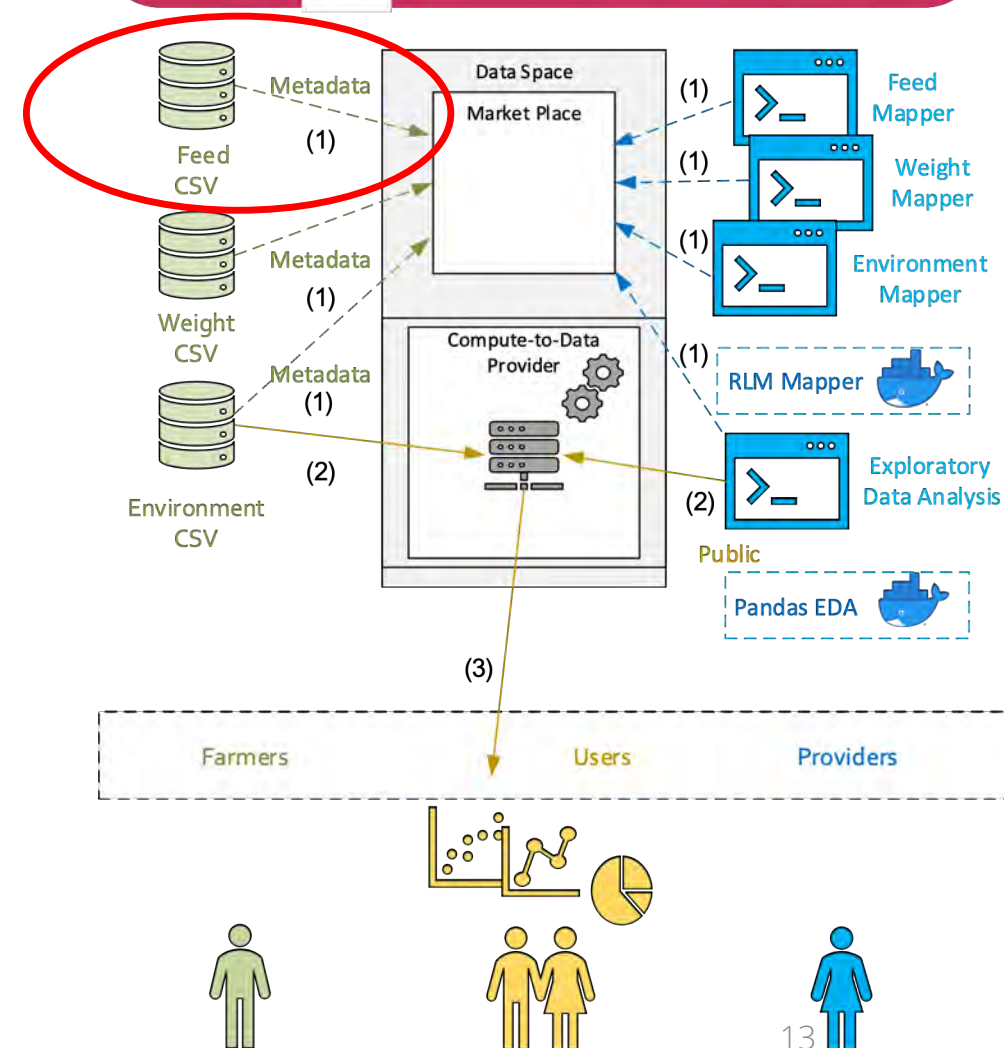
2 sales

**A14
Pork**



STEP 1

Publishing & Reading



Exploratory Data Analysis

GEN-X Testnet

Owned by 0x4cF0...2E17
Accessed with GXAT

ALGORITHM Published 19 days ago

Generate an **Exploratory Data Analysis (EDA)** report for the input tabular data.

The input data is loaded using **pandas** and then the EDA report is generated using **pandas-profiling**. The HTML version of the report can be downloaded when the algorithm finishes.

eda exploratory-data-analysis profiling pandas tabular-data
udl

DATA AUTHOR

OWNER

0x4cF0...2E17

DOCKER IMAGE

rogargon/pandas-profiling:

DID

did:op:e73ac9c332adfea204acefb4b300e0d
dde977ca9dee99017a8ed94242dc4bfa53

METADATA HISTORY

NFT metadata updated 19 days ago

DDO

Free

This algorithm has been set to private by the publisher and can't be downloaded. You can run it against any allowed datasets though!

Datasets algorithm is allowed to run on

CEP - Automatic Pig Feeding - 2021 S1 - 982091062894196	Free
CEP - Automatic Pig Feeding - 2021 - Pens 4 & 5	1
CEP - Daily Pig Weight - 2021 - Pens 4 & 5	Free

Consent

Your Compute Jobs **SHOW**

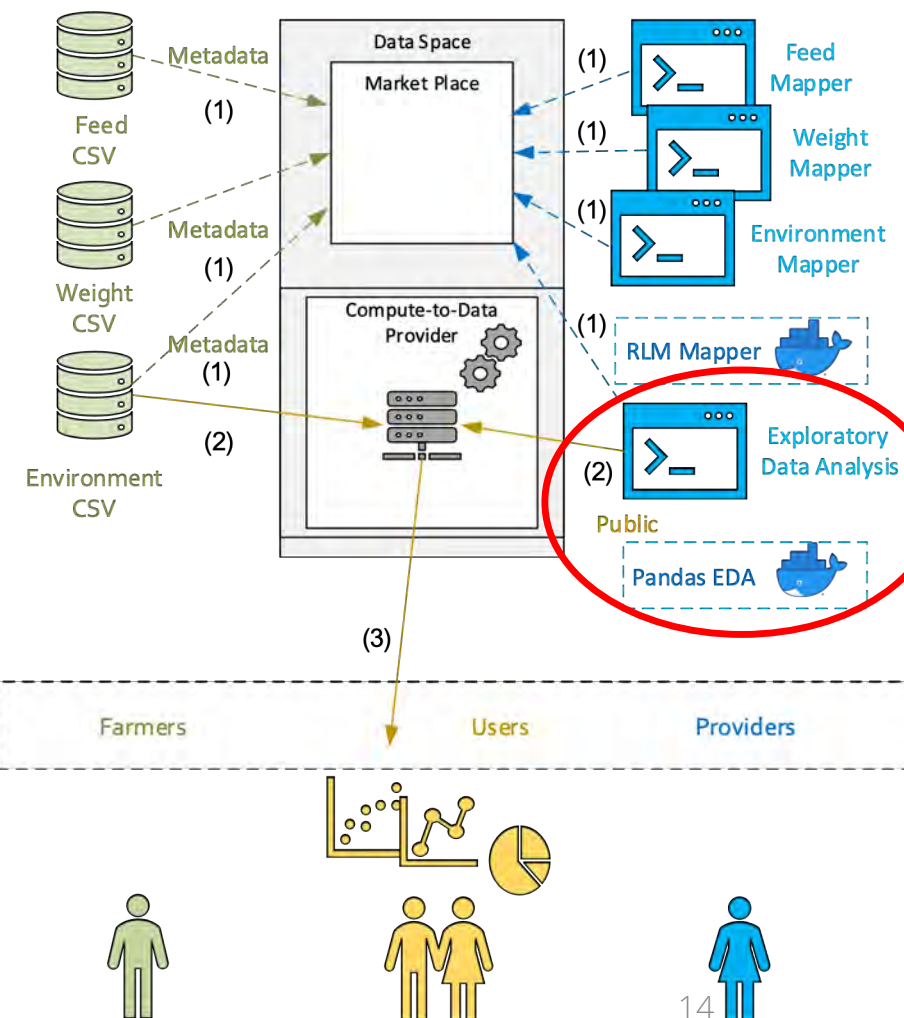
6 sales

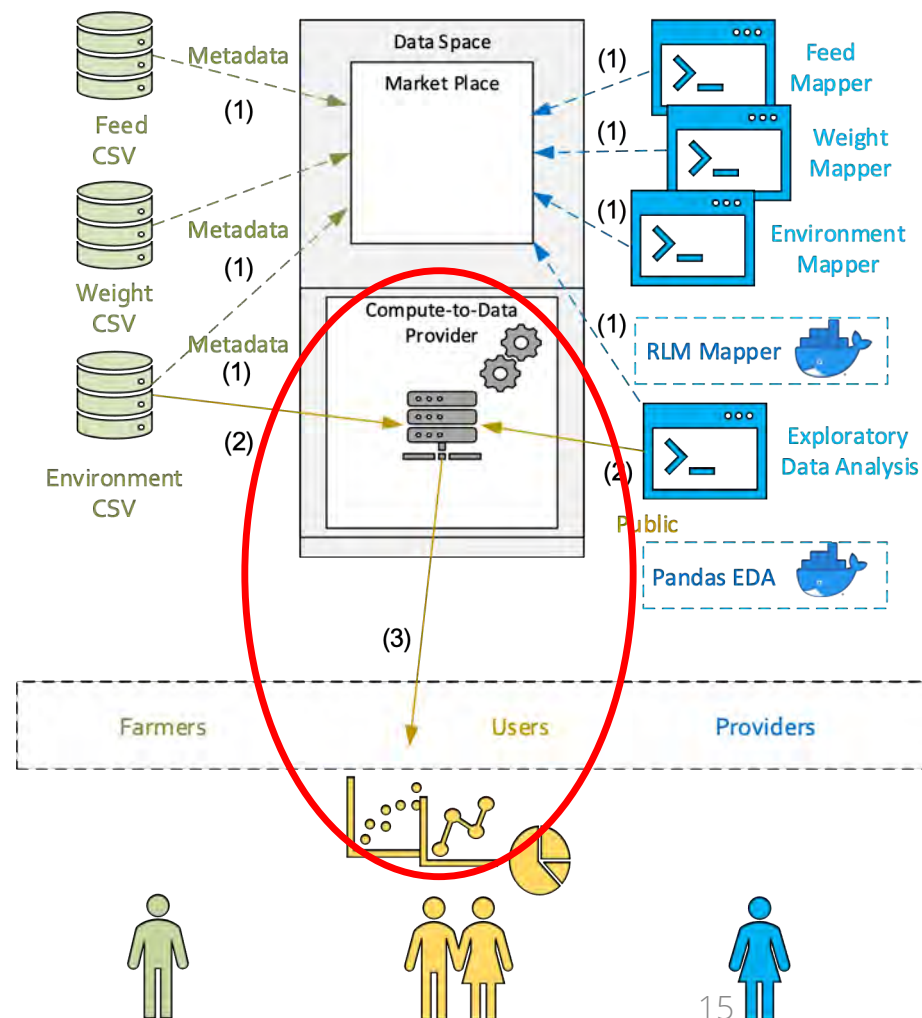
A14
Pork



STEP 1

Publishing & Reading





RML Mapper for Daily Pig Weight CSV Data

GEN-X Testnet

Owned by 0x6207...6822
Accessed with GAT 7.8

ALGORITHM Published 8 days ago

This algorithm supports the "Pay-as-you-go" approach when sharing data through a data space. Instead of requiring that the publisher integrates the data based on existing schemas, which causes a significant upfront overhead, the pay-as-you-go paradigm favors an incremental approach. This way, entry barriers are lowered and it becomes easier to share data.

In this particular case, publishers are free to share their data using their existing data schema, for daily pig-weight CSV datasets with the following columns:

- Animal ID: the pig identifier.
- Date: the date for the measured weight.
- Weight (g): animal weight in grams.

This algorithm makes it possible to follow the incremental and tiered approach proposed by the pay-as-you-go paradigm. It maps CSV data already shared through the data space to W3C RDF semantic data based on well-established vocabularies and ontologies that facilitate data integration, even across use cases and activity sectors.

The mapping implemented by this algorithm generates RDF data based on the Smart Applications REFERENCE (SAREF) ontology, a vocabulary supported by the European Telecommunications Standards Institute (ETSI) that facilitates data integration in the smart applications domain.

For instance, for the data:

Animal ID	Date	Weight (g)
982091062894196	2021-03-16	16300

The RML mapping generates the following RDF semantic data graph:



base: <https://a14pork.angliru.udi.cat/schauer/a14pork>
a14pork: <https://a14pork.angliru.udi.cat/ontology/>
saref: <https://saref.etsi.org/core/>
s4agri: <https://saref.etsi.org/saref4agri/>
s4city: <https://saref.etsi.org/saref4city/>
auto: <https://saref.etsi.org/saref4auto/>
om: <http://www.omg.org/om-2/>
time: <http://www.w3.org/2001/XMLSchema#>

Datasets algorithm is allowed to

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

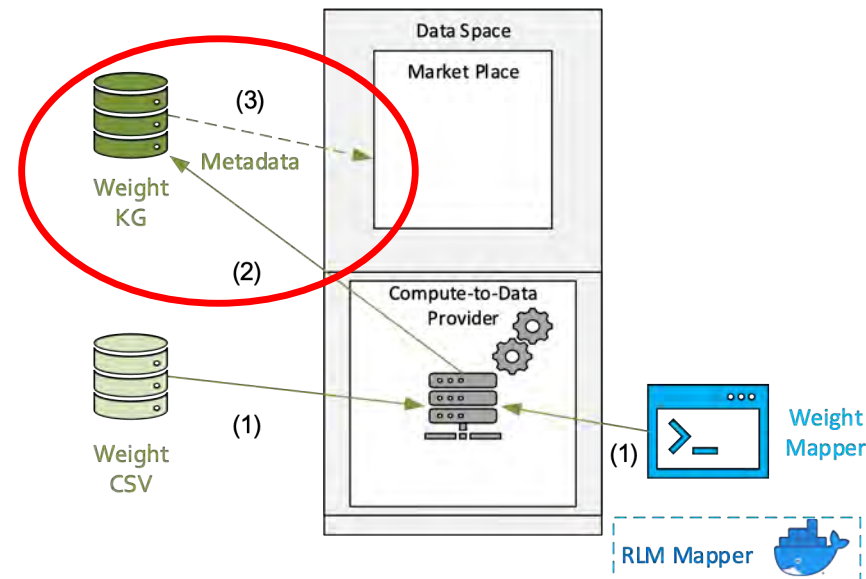
CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

CEP - Automatic Pig Feeding - 2021 51 - 982091062894196

Semantic Integration

STEP 2

Semantic Integration

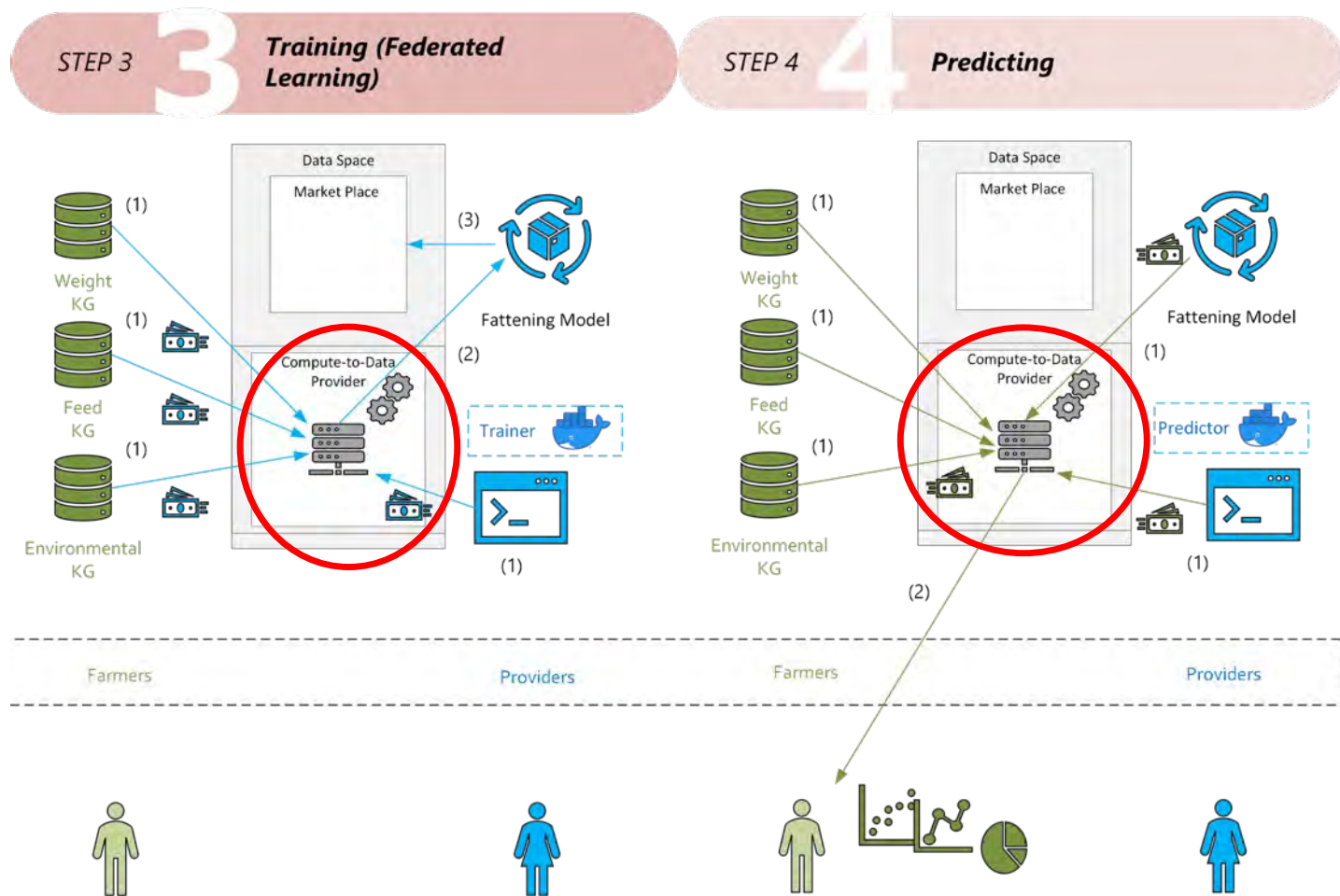


Farmers

Providers



WiP: Data & AI Models Monetization



You will pay 2.125 EUROe ⓘ

START	DATASET	1 EUROe
	+ ALGORITHM	1 EUROe
	+ C2D RESOURCES	0.125 EUROe for 898 seconds
	=	2.125 EUROe

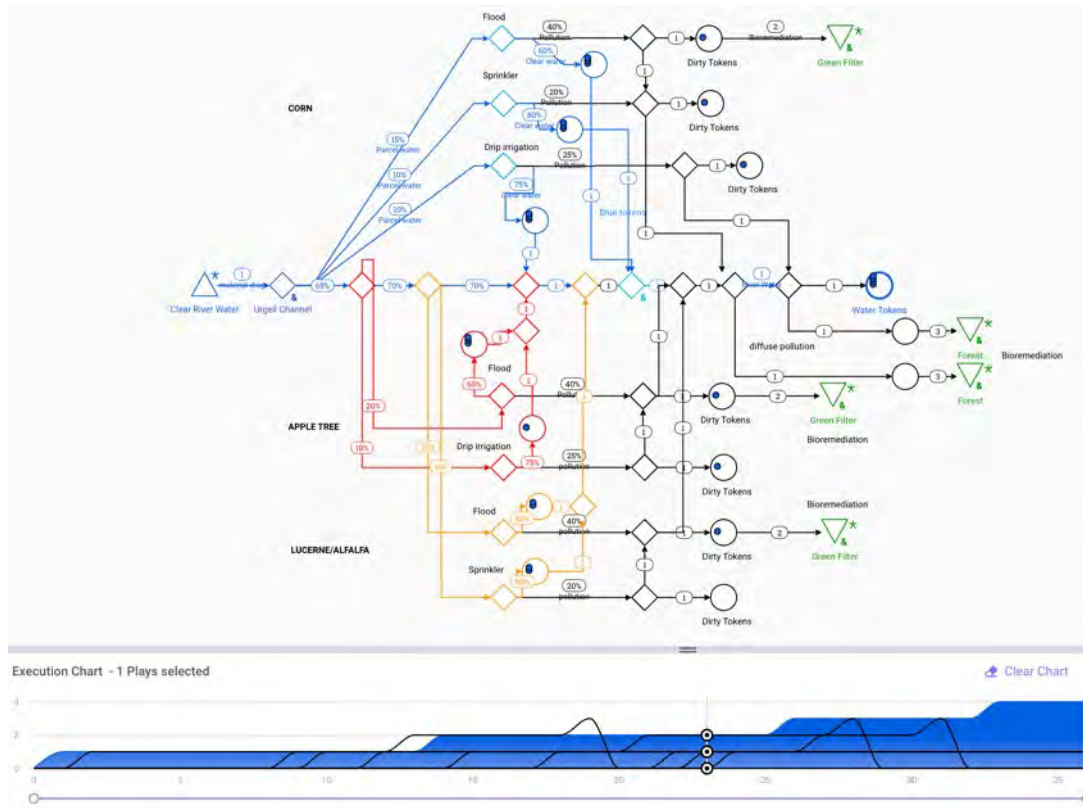
You bought this dataset already al
already bought the selected algo
again. The C2D resources required
required for them, Please note tha
free assets.

Access allowed ✓

Your Compute Jobs [SHOW ▼](#)

1 sale

Future Work: Tokenomics



<https://machinations.io/community/rgil.iranzo/water-river-tokenization-e52cd94ca89011eda2330626ff1c9bc8>

- Monetization through tokens available in Pontus-X (Gaia-X Web3 Ecosystem)
 - Payment tokens, including EUROe stablecoin
 - Aligned with EU's Markets in Crypto Assets regulation (MiCA)
 - Governance tokens
 - Representing voting power in data cooperative DAOs, where data reuse consent is delegated
 - ...

Towards European Agriculture Data Spaces: A Proof of Concept in the Pig Sector

Roberto García
Universitat de Lleida
<https://rhizomik.net/~roberto>

Acknowledgements

[AgriDataSpace](#) (Grant 101083401) funded by the European Union

[AI4Pork](#) (TED2021-130829B-100) funded by MCIN/ AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR

[ANGLIRU](#) (PID2020-117912RB-C22) funded by MCIN/AEI/10.13039/501100011033