

# Data Markets for Digital Sovereignty

26 April | 14:00 - 17:00 CET

Online



# Data Markets for Digital Sovereignty

## Welcoming Remarks

Justina Bieliauskaite, Projects Director (European DIGITAL SME Alliance)



# Data Markets for Digital Sovereignty

## Keynote Speakers



**Federico Milani**

Deputy Head of the Unit, Data Policy and Innovation  
(DG CNECT, European Commission)



**Prof. Dr. Nikolaos Laoutaris**

Director (IMDEA Madrid)







Data  
Transparency  
Group



# Understanding the Price of Data in Commercial Data Marketplaces

Santiago Andrés Azcoitia, IMDEANetworks Institute

Costas Iordanou, Cyprus University of Technology

Nikolaos Laoutaris, IMDEANetworks Institute

[Developing the  
Science of Networks]



# We checked more than 190 companies offering data products and services in order to understand how data is traded nowadays<sup>1</sup>



[1] S. Andrés Azcoitia, N. Laoutaris. "A Survey of Data Marketplaces and Their Business Models", SIGMOD Record Sept'22





**We scraped 10 data marketplaces (DMs) + 30 sellers and collected information about 215,075 data products from 2,115 sellers in total**

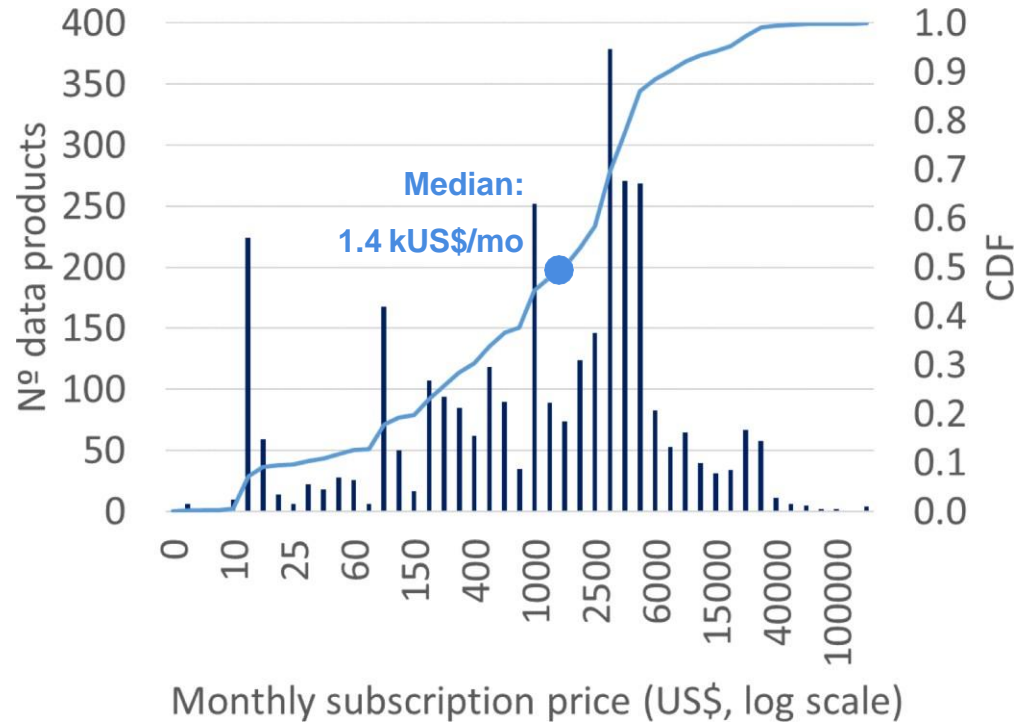
<b>Marketplace</b>	<b>#Products</b>	<b>#Paid prod.</b>	<b>#Sellers</b>
<b>Advaneo</b>	198,743	1	N/A
<b>AWS</b>	4,263	2,674	262
<b>DataRade</b>	1,592	1,592	1,262
<b>Snowflake</b>	889	889	200
<b>Knoema</b>	158	158	142
<b>DAWEX</b>	160	160	79
<b>Carto</b>	8,182	5,283	42
<b>Crunchbase</b>	9	9	15
<b>Veracity</b>	115	95	38
<b>Refinitiv</b>	187	187	76
<b>Other providers</b>	777	775	30

**Only 12k data products are “paid” and only 4k from 443 distinct sellers disclose information about their prices!**

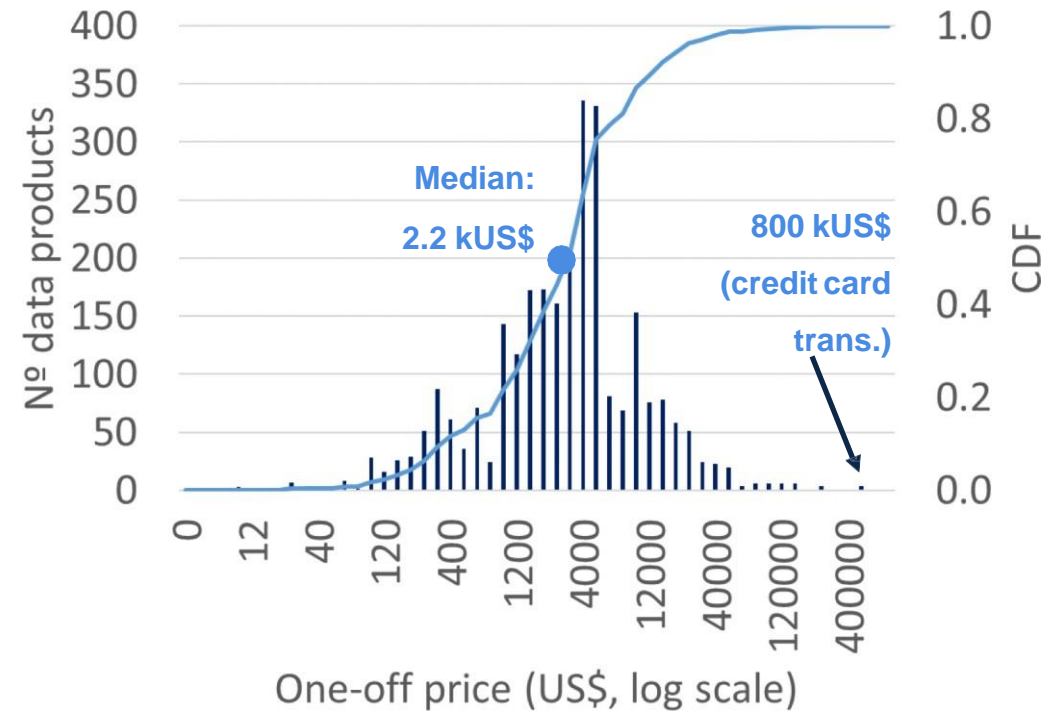


We found that data sells at an immensely wide range of prices, ...

### Subscription-based data product prices



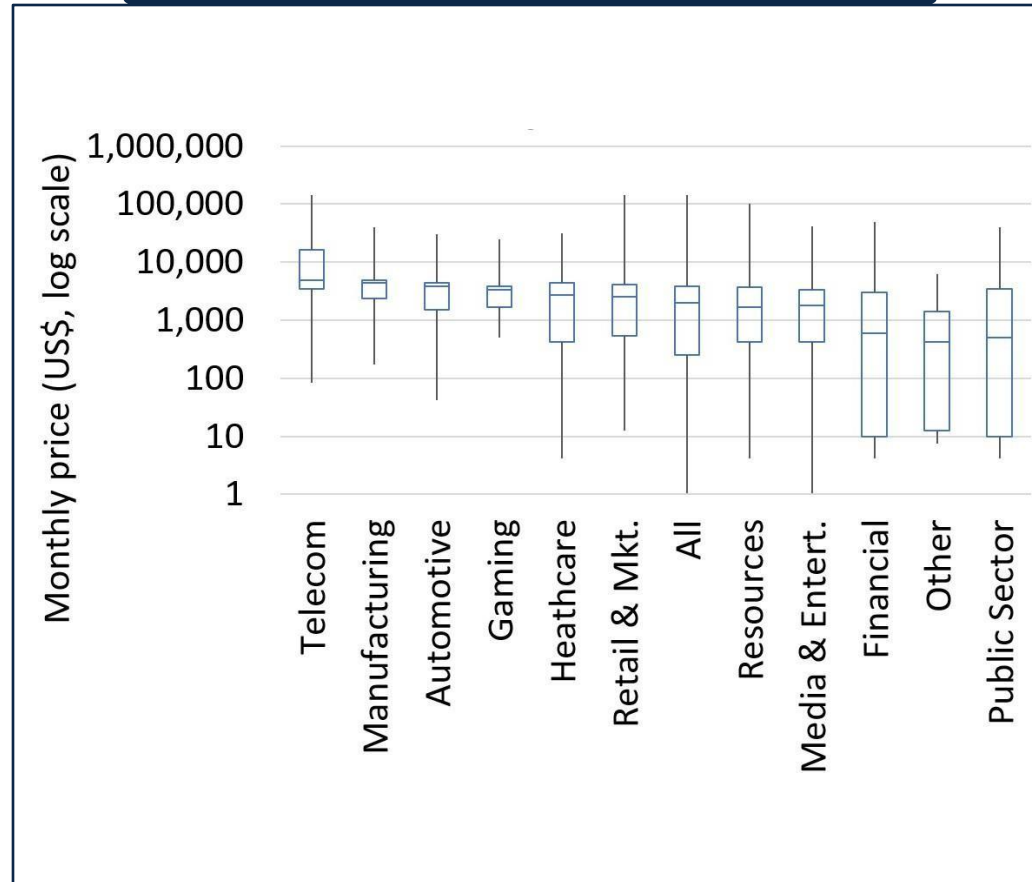
### One-off data product prices



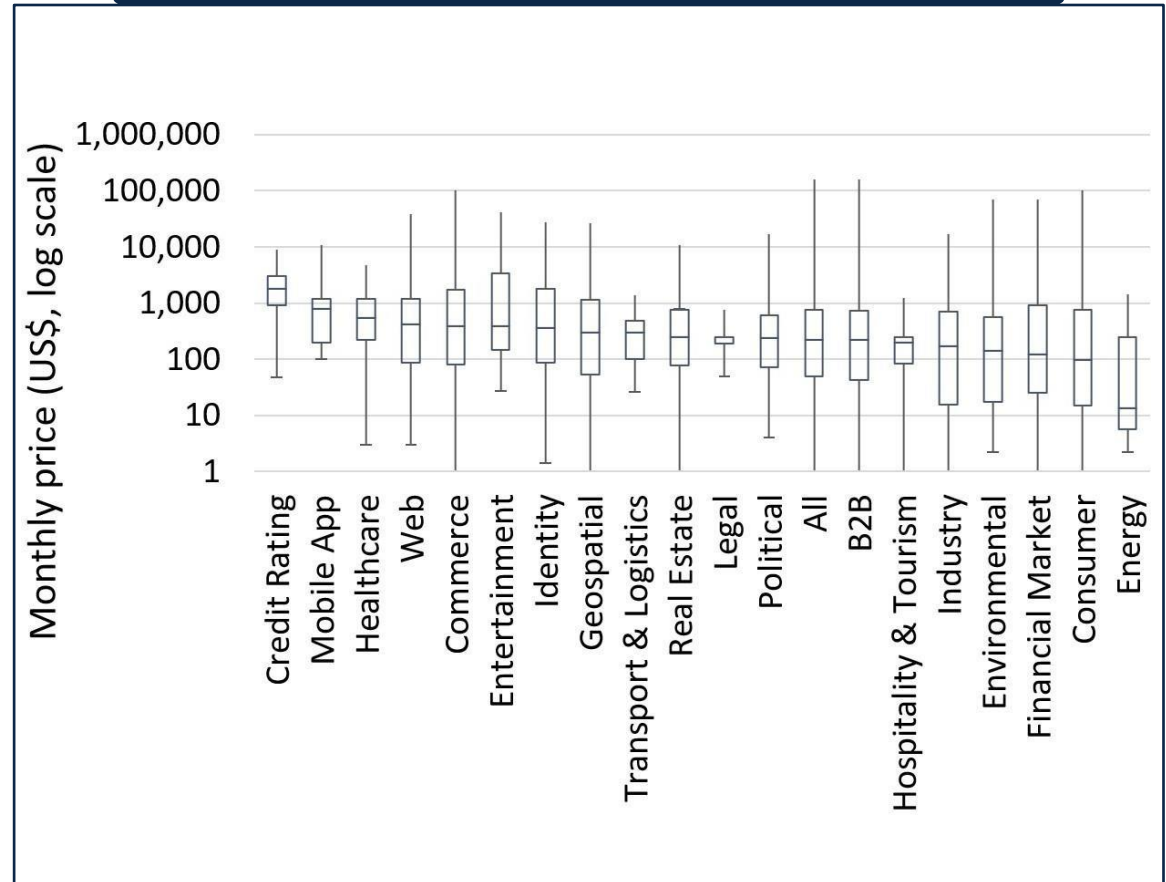


... which depend on the category of data product

Data product prices by category AWS



Data product prices by category DataRade








Cross DMs analysis is challenging, since DMs i) provide different detail, and ii) use different categorisation and criteria to assign categories to data products





Explore / Consumer Transaction Data

## Yodlee's 4M Panel (US Consumer Transaction Data, de-identified)

Starts at **\$400,000** / year

A dataset by [Envestnet](#) | [Yodlee](#)

	SECONDARY_MERCHANT_NAME	PRIMARY_MERCHANT_NAME	TRANSACTION_CATEGORY_NAME	TRANSACTION_BASE_TYPE	+4 MORE
1	Paypal	7-Eleven	Entertainment/Recreation	debit	...
2					

Request Free Data Sample →

3K Merchants

99% High precision mapping for 600 tickers

USA covered

9 years of historical data

Get a Quote

Contact Provider

"Our most granular offering providing line-by-line transactions for 4 millions US consumers."

Access to Consumer spend data of de-identified 4M users over 9 years. Clean tagged consumer transaction data on millions of merchants public and private. Suitable for all investment use cases - Fundamental, Quant, Private Equity, Venture Capital.

### Data Attributes

Attribute & Description	Example
-------------------------	---------


Envestnet | Yodlee

Powering Dynamic Innovation for Financial Services

Verified Provider

100% Response rate

Trusted by








Cross DMs analysis is challenging, since DMs i) provide different detail, and ii) use different categorisation and criteria to assign categories to data products





### Consumer transaction and payment data

Provided By: Alliant

Alliant consumer transaction and payment data, sourced from Alliant's proprietary cooperative database of billions of transactions. Examples: Credit card transactions, dollar and number broken out by block group Alliant's proprietary payment score metric

[Continue to subscribe](#)

Product offers

Overview

Usage

Support

#### Product offers

The following offers are available for this product. Choose an offer to view the pricing and access duration options for the offer. Select an offer and continue to subscribe. Your subscription begins on the date that your request is approved by the provider. Additional taxes or fees might apply.

#### Public offer

Payment schedule: Upfront payment | Offer auto-renewal: Supported

☒ \$3,500 for 1 month

☐ \$35,000 for 12 months

#### Overview

Consumer transaction and payment data, aggregated at the geographic block group level. Data is sourced from Alliant's proprietary cooperative database which aggregates hundreds of leading DTC brand's 1st party detailed transactional CRM data. Deterministic view into U.S. geographic block groups transaction and payment detail. Example data points include: -total number and dollar amount of credit card transactions by block group in last 5 years -total number and dollar amount of write offs by block group in last 5 years -Alliant's proprietary payment score metric (grouped 1-20)

Overview one sheet: [https://info.alliantinsight.com/hubfs/Downloadable%20Content%20Alliant%20AWS\\_Geo\\_Performance.pdf](https://info.alliantinsight.com/hubfs/Downloadable%20Content%20Alliant%20AWS_Geo_Performance.pdf)

Provided By

Alliant





We trained NLP NB classifiers to learn how a *source* DM labels products that belong in a certain category, and label products in a *destination* DM

Significant stems

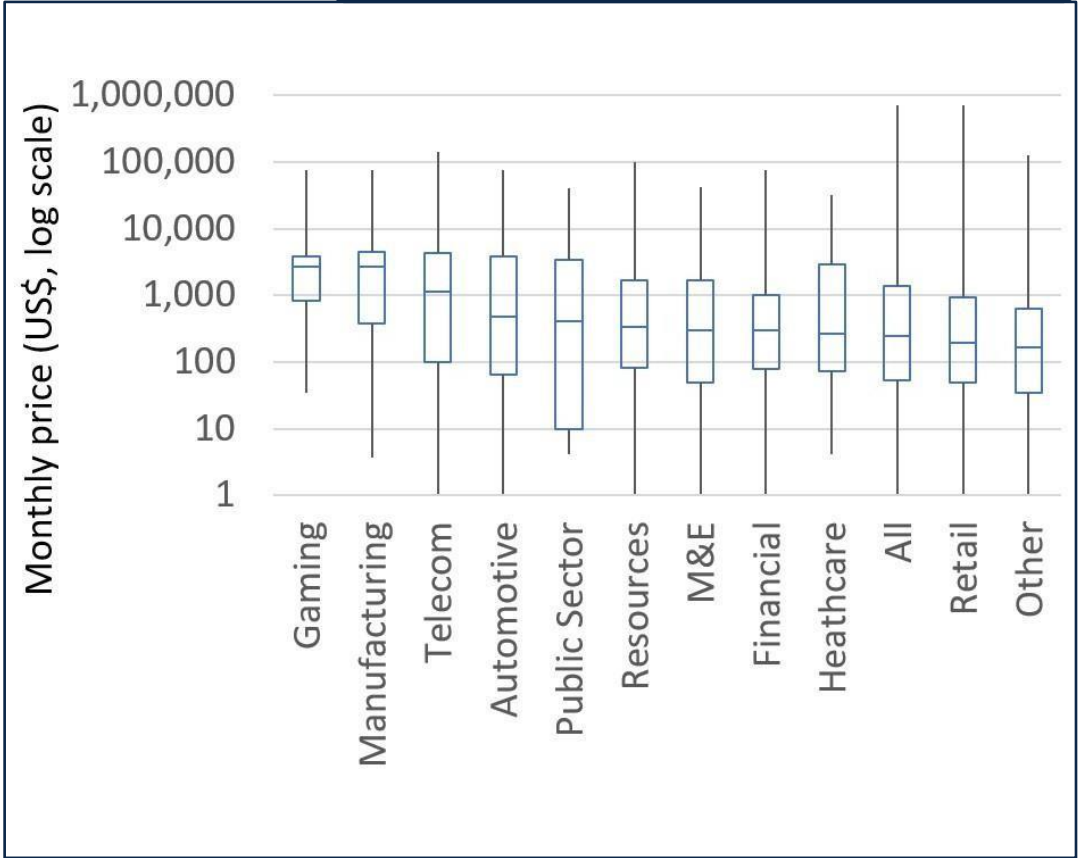
**Financial:** ‘system’, ‘sec’, ‘exchang’, ‘type’, ‘file’, ‘form’, ‘edgar’, ‘secur’, ‘act’, and ‘compani’.

**Retail, Location and Marketing:** ‘locat’, ‘topic’, ‘b2b’, ‘score’, ‘echo’, ‘trial’, ‘compani’, ‘visit’, ‘intent’, ‘consum’.

Accuracy score

	Accuracy	Precision	Recall	$F_1$ Score
Test - Financial	0.93	0.97	0.81	0.88
Test - Retail	0.95	0.96	0.88	0.91
Val. - Financial	0.89	0.72	0.88	0.79
Val. - Retail	0.78	0.81	0.68	0.74

Boxplot by AWS category (all)



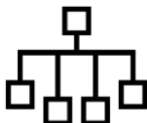




# We built a cross-DM database as a superset of metadata fields found in different DMs, and found to be driving the prices of data products



Id & Description



Category



Granularity



Time scope



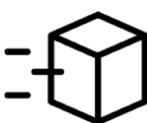
Use cases



Identifiability



Volume & units



Delivery method



Limitations



Geo scope



Update frequency



Add-ons



**So, which are the features actually driving the prices of data products?**


 We tested 9 regressors and optimized 4 of them. At least one shows  $R^2 > 0.78$  for predicting prices of financial, marketing and health-related data

TABLE IV: Accuracy achieved by regression models

Model	Financial			Marketing			Healthcare			All		
	$R^2$	MAE	MSE	$R^2$	MAE	MSE	$R^2$	MAE	MSE	$R^2$	MAE	MSE
<b>RF</b>	0.85	0.2	0.14	0.86	0.21	0.13	0.78	0.25	0.15	0.84	0.23	0.16
<b>kN</b>	0.78	0.31	0.26	0.74	0.33	0.24	0.77	0.26	0.17	0.69	0.37	0.31
<b>GB</b>	0.82	0.23	0.16	0.8	0.28	0.19	0.73	0.27	0.19	0.79	0.3	0.22
<b>DNN</b>	0.73	0.33	0.35	0.77	0.30	0.22	0.68	0.26	0.18	0.72	0.33	0.28

Note: MAE and MSE reflect the error in predicting the logarithm of data product prices

We discarded linear, Elastic-Net, Ridge, Bayesian Ridge, and Lasso regressions even though they worked well in specific cases





**We studied the most relevant individual features which sellers rely on for pricing financial, marketing and healthcare data**

Financial			Marketing			Healthcare		
RF	kNeigh	GB	RF	kNeigh	GB	RF	kNeigh	GB
units	units	units	units	units	csv	units	csv	wordlist
entities	Email	S3Bucket	entities	History	units	people	units	Del. Methods
S3Bucket	Download	wordmonthli	IdSessions	USA	yearly	wordhealth	daily	wordhospit
wordsubmit	daily	wordstock	Download	IdSessions	people	wordtrend	wordmarket	wordidentifi
Download	IdCompanies	worddeliv	REST API	N° Countries	REST API	wordmedic	wordgo	wordamerica
people	USA	people	wordcustom	Financial	wordqualiti	wordglobal	Limitations	wordhealth
txt	wordmarket	Del. Methods	USA	Others	wordaccur	csv	location data	wordreport
wordedgar	Retail	txt	yearly	people	wordidentifi	DelMethod	wordpopul	wordstudi
wordcustom	wordcontact	wordneed	monthly	wordcontact	wordwebsit	wordinsight	wordprofil	wordupdat
wordlist	realtime	wordsubmit	IdCompanies	Email	UIExport	wordreport	wordinsight	wordcontact

The table shows average scores of 5-fold executions of leave-one-out and permutation importance analysis. An average of 11 of the top 20 features by category and algorithm appear in every individual test.



Features related to data volume are present in financial and marketing data categories, but seem to be especially relevant for financial data products

Financial			Marketing			Healthcare		
RF	kNeigh	GB	RF	kNeigh	GB	RF	kNeigh	GB
units	units	units	units	units	csv	units	csv	wordlist
entities	Email	S3Bucket	entities	History	units	people	units	Del. Methods
S3Bucket	Download	wordmonthli	IdSessions	USA	yearly	wordhealth	daily	wordhospit
wordsubmit	daily	wordstock	Download	IdSessions	people	wordtrend	wordmarket	wordidentifi
Download	IdCompanies	worddeliv	REST API	N° Countries	REST API	wordmedic	wordgo	wordamerica
people	USA	people	wordcustom	Financial	wordqualiti	wordglobal	Limitations	wordhealth
txt	wordmarket	Del. Methods	USA	Others	wordaccur	csv	location data	wordreport
wordedgar	Retail	txt	yearly	people	wordidentifi	DelMethod	wordpopul	wordstudi
wordcustom	wordcontact	wordneed	monthly	wordcontact	wordwebsit	wordinsight	wordprofil	wordupdat
wordlist	realtime	wordsubmit	IdCompanies	Email	UIExport	wordreport	wordinsight	wordcontact

Due to the heterogeneity of the sample, there is no single feature other than units that relates to the price of data in every category. **The ‘what’ seems to be more important than the ‘how much’ when pricing healthcare products**



Among the rest of the features, the ones related to 'what' data is offered stand out in terms of importance

Financial			Marketing			Healthcare		
RF	kNeigh	GB	RF	kNeigh	GB	RF	kNeigh	GB
S3Bucket	Email	S3Bucket	IdSessions	History	csv	wordhealth	csv	wordlist
wordsubmit	Download	wordmonthli	Download	USA	yearly	wordtrend	daily	Del. Methods
Download	daily	wordstock	REST API	IdSessions	REST API	wordmedic	wordmarket	wordhospit
txt	IdCompanies	worddeliv	wordcustom	Nº Countries	wordqualiti	wordglobal	wordgo	wordidentifi
wordedgar	USA	Del. Methods	USA	Financial	wordaccur	csv	Limitations	wordamerica
wordcustom	wordmarket	txt	yearly	Others	wordidentifi	Del. Methods	location data	wordhealth
wordlist	Retail	wordneed	monthly	wordcontact	wordwebsit	wordinsight	wordpopul	wordreport
wordcontact	wordcontact	wordsubmit	IdCompanies	Email	UI Export	wordreport	wordprofil	wordstudi
wordsystem	real time	wordreport	wordname	UI Export	wordcover	wordregion	wordinsight	wordupdat
wordcompar	wordprice	wordcontact	location data	Download	wordfield	wordlist	Download	wordcontact





Features relating to **delivery methods** and **update rate** seem somewhat important for the prices of financial and marketing data

Financial			Marketing			Healthcare		
RF	kNeigh	GB	RF	kNeigh	GB	RF	kNeigh	GB
S3Bucket	Email	S3Bucket	IdSessions	History	csv	wordhealth	csv	wordlist
wordsubmit	Download	wordmonthli	Download	USA	yearly	wordtrend	daily	Del. Methods
Download	daily	wordstock	REST API	IdSessions	REST API	wordmedic	wordmarket	wordhospit
txt	IdCompanies	worddeliv	wordcustom	Nº Countries	wordqualiti	wordglobal	wordgo	wordidentifi
wordedgar	USA	Del. Methods	USA	Financial	wordaccur	csv	Limitations	wordamerica
wordcustom	wordmarket	txt	yearly	Others	wordidentifi	Del. Methods	location data	wordhealth
wordlist	Retail	wordneed	monthly	wordcontact	wordwebsit	wordinsight	wordpopul	wordreport
wordcontact	wordcontact	wordsubmit	IdCompanies	Email	UI Export	wordreport	wordprofil	wordstudi
wordsystem	real time	wordreport	wordname	UI Export	wordcover	wordregion	wordinsight	wordupdat
wordcompar	wordprice	wordcontact	location data	Download	wordfield	wordlist	Download	wordcontact

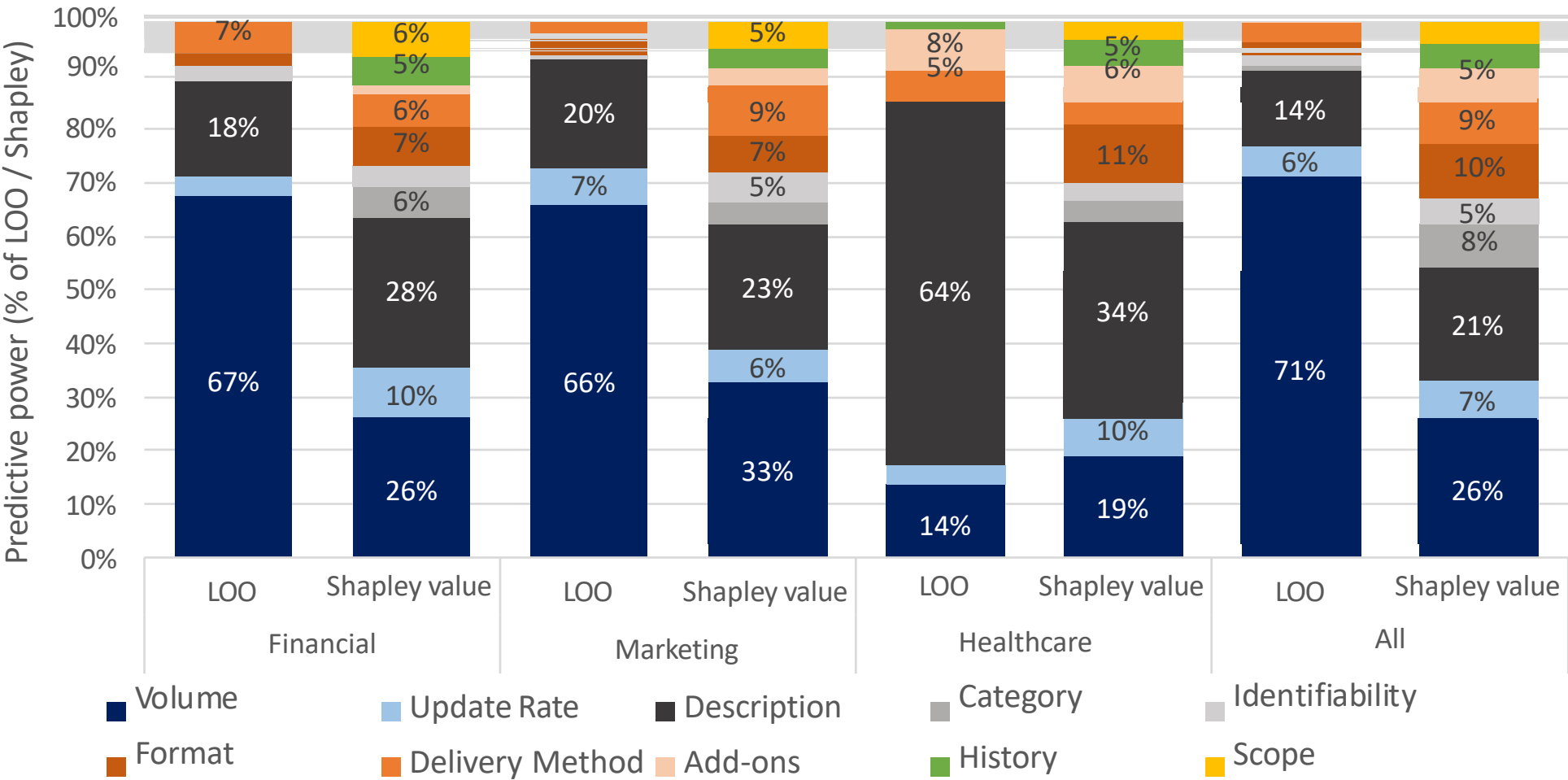


**Geo-spatial localization and scope** and the possibility of connecting data points from the same owner are relevant especially for marketing data.

Financial			Marketing			Healthcare		
RF	kNeigh	GB	RF	kNeigh	GB	RF	kNeigh	GB
S3Bucket	Email	S3Bucket	IdSessions	History	csv	wordhealth	csv	wordlist
wordsubmit	Download	wordmonthli	Download	USA	yearly	wordtrend	daily	Del. Methods
Download	daily	wordstock	REST API	IdSessions	REST API	wordmedic	wordmarket	wordhospit
txt	IdCompanies	worddeliv	wordcustom	Nº Countries	wordqualiti	wordglobal	wordgo	wordidentifi
wordedgar	USA	Del. Methods	USA	Financial	wordaccur	csv	Limitations	wordamerica
wordcustom	wordmarket	txt	yearly	Others	wordidentifi	Del. Methods	location data	wordhealth
wordlist	Retail	wordneed	monthly	wordcontact	wordwebsit	wordinsight	wordpopul	wordreport
wordcontact	wordcontact	wordsubmit	IdCompanies	Email	UI Export	wordreport	wordprofil	wordstudi
wordsystem	real time	wordreport	wordname	UI Export	wordcover	wordregion	wordinsight	wordupdat
wordcompar	wordprice	wordcontact	location data	Download	wordfield	wordlist	Download	wordcontact



# We studied the most influential feature groups, as well, resulting in notorious differences across data categories







## To probe further

Data Economy 2022

Home Call For Papers Committee Program

# CALL FOR PAPERS

Data-driven decision making powered by Machine Learning (ML) algorithms is changing how the society and the economy work and is having a profound positive impact on our daily life. With the exception of very large companies that have both the data and the skills to develop powerful ML-driven services, the large majority of provably possible ML services, from e-health, to transportation and predictive maintenance, to name just a few, still remain at the idea or prototype level for the simple reason that data, the skills to manipulate them, and the business models to bring them to market, seldom co-exist under the same roof. Data has to somehow meet with the ML and business skills that can unleash its full power for the society and economy.



## SECOND ACM DATA ECONOMY WORKSHOP

Seattle, June 18, 2023

Co-located with SIGMOD 2023 (June 18, 2023, Seattle, WA, USA)

## KEYNOTE SPEAKERS



Bruce Pon  
Founder of Ocean Protocol



Damian Boeselager  
Member of the European Parliament



Carme Artigas  
Secretary of State for Digitization and  
Artificial Intelligence, Government of  
Spain



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101070069.



### MLEDGE

Cloud and Edge Machine Learning

IMDEA Networks is the beneficiary of this project

[More info](#)

This project (REGA22e00052829516) has been funded by the Ministry of Economic Affairs and Digital Transformation and the European Union-NextGenerationEU/PRTR.





**In summary, this is the first paper measuring and understanding the price of data in commercial marketplaces, we found that:**

**1**

Data products sell at an immensely wide range of prices up to several US\$100ks per month

**2**

We homogenized heterogeneous metadata and classification labels to be able to compare data products across marketplaces

**3**

Using regression models, we managed to fit the prices of commercial products from their features with  $R^2$  above 0.84.

**4**

Features related to 'what' and 'how much' data a product contains are driving 66% of its price, and some other features (geo-scope, history, update rate) are relevant for specific categories.

**5**

We've made available code and data obtained in this study which you can find in <https://gitlab.com/sandresazcoitia1/data-pricing-tool>

Thank you!

Q&A time!

For more information please contact:



**Santiago Andrés Azcoitia**

[santiago.azcoitia@imdea.org](mailto:santiago.azcoitia@imdea.org)



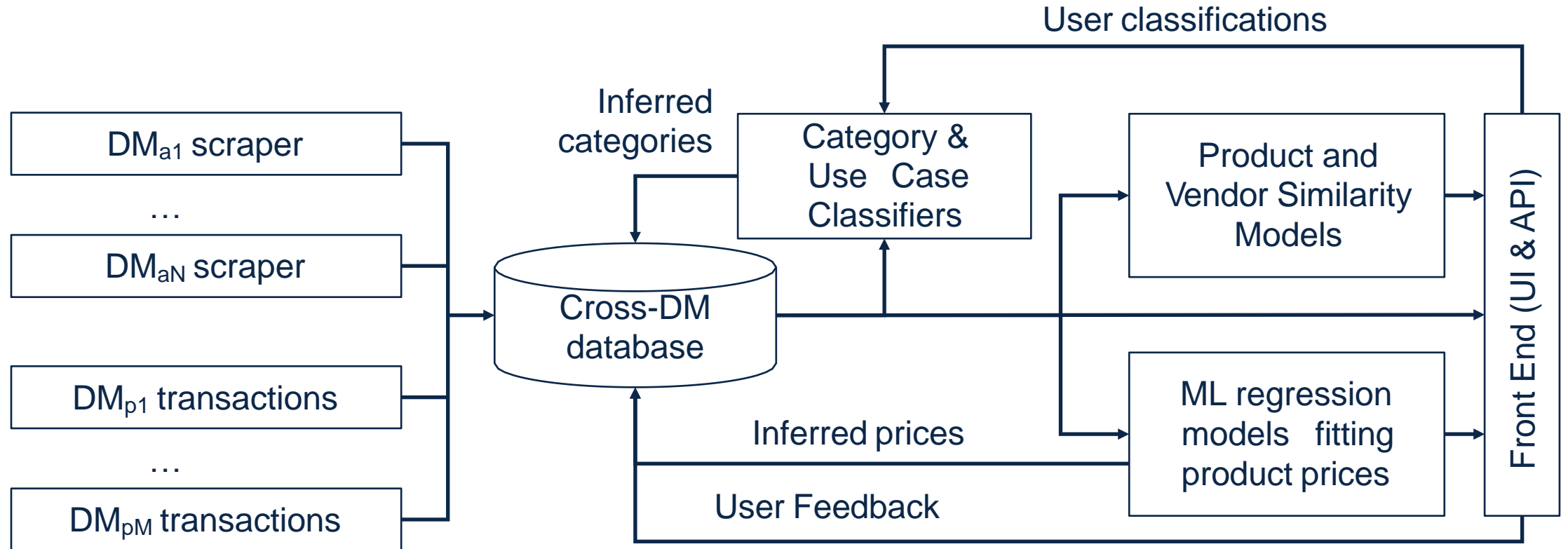
**Nikolaos Laoutaris**

[nikolaos.laoutaris@imdea.org](mailto:nikolaos.laoutaris@imdea.org)



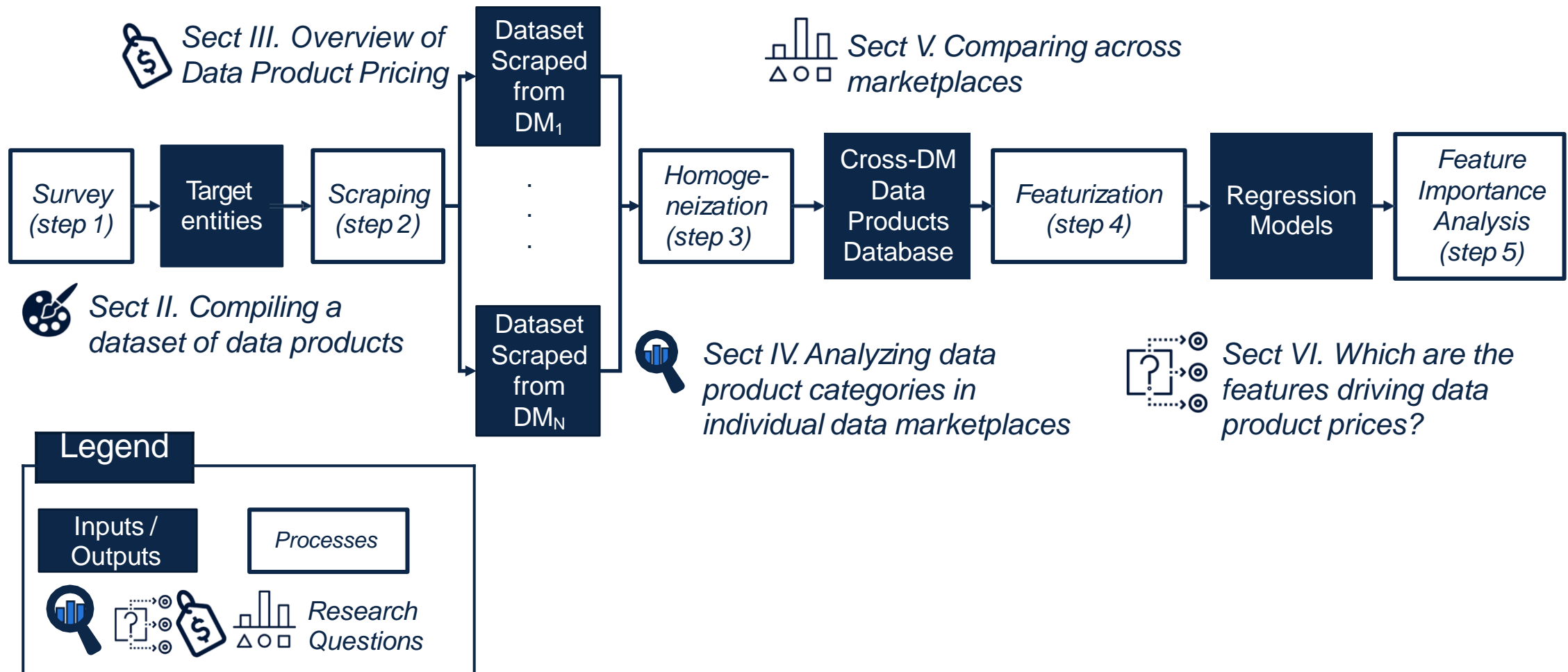


We are working on a data quotation tool<sup>2</sup> to be able to predict the prices of a data product out of its metadata based on market prices and transactions



Such a tool will have limitations, since it does not consider: i) the usability for the buyer, ii) the quality of the data, iii) the specific value for a buyer.

# So, what is the price of data in the B2B market? What are the features that are driving the prices of data products?





# Data Markets for Digital Sovereignty

## Session 1

- Pierre Gronlier, Chief Technology Officer (GAIA-X AISBL)
- Isabella De Michelis, CEO & founder ErnieApp Ltd. (Ireland)
- Martin Serrano, Project Coordinator (i3-MARKET) & Senior Research Fellow (INSIGHT Centre for Data Analytics)



# Data Markets for Digital Sovereignty

---

26<sup>th</sup> April

Pierre Gronlier - CTO, Gaia-X

# Dataspaces and Federations



- Dataspace: business point of view
  - Similar to a Domain Ownership from Data Mesh
  - Specifies and maintains governance rules (ex Gaia-X Labels)

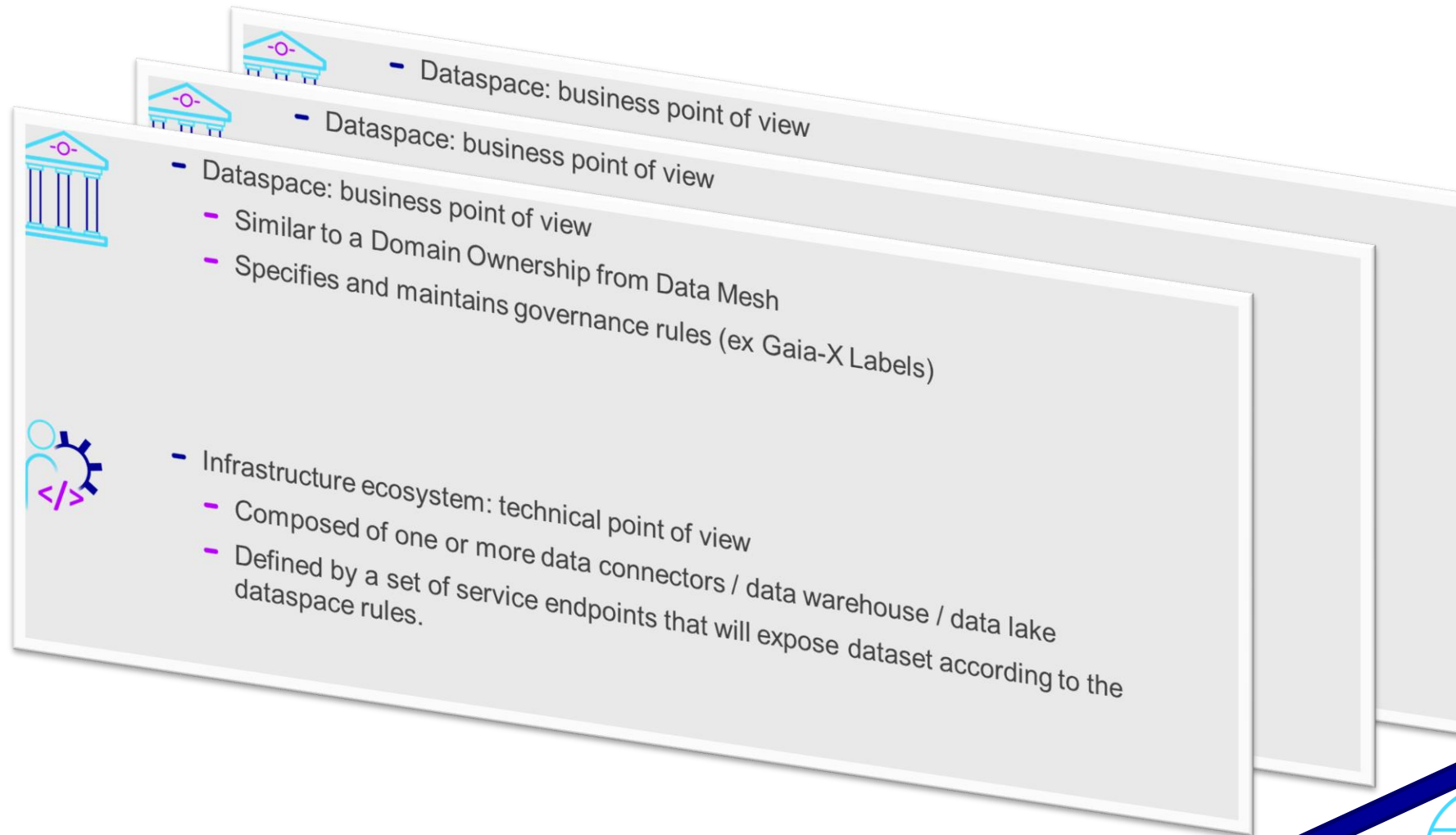


- Federations: technical point of view
  - Composed of one or more data connectors / data warehouse / data lake
  - Defined by a set of service endpoints that will expose dataset according to the dataspace rules.

**=> A dataspace can span across several federations.**

**=> A federation can be used by several dataspaces.**

# Gaia-X governance



# Our Model

## Connecting Data - Infrastructures Ecosystems



### Advanced Services

New (Cross-) Sector Innovations / Applications built by from service composition.



### Data Spaces / Federations

Interoperable & portable (Cross-) Sector data-sets and services.



### Data Exchange

Anchored contract rules for access and data usage.



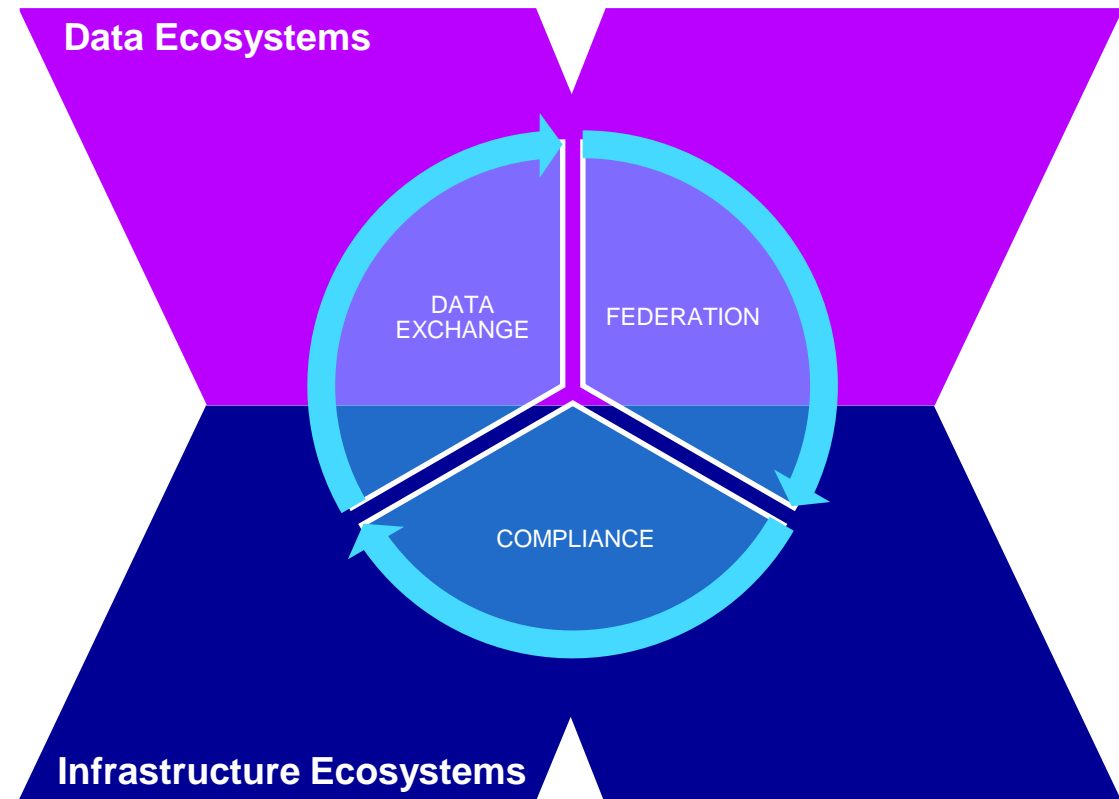
### Gaia-X Compliance

Decentralized services to enable objective and measurable trust.



### Label framework

Gaia-X and ecosystem specific Labels to ease market adoption through autonomy and self-determination.

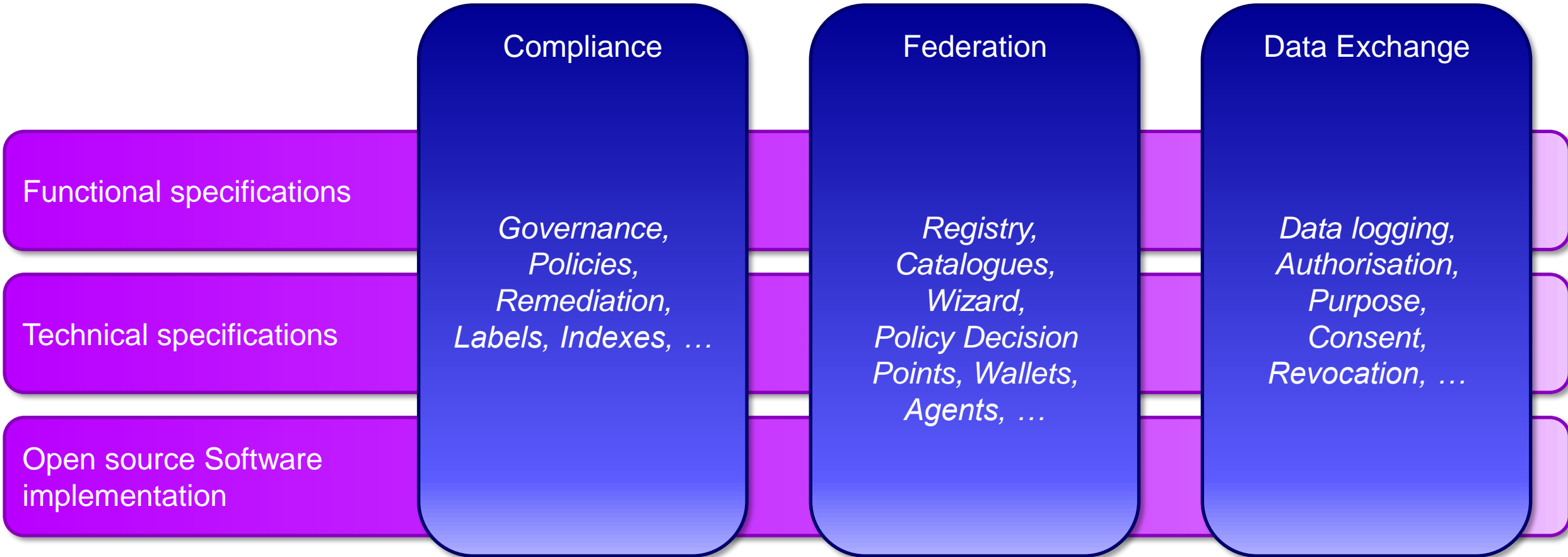




# Today's status



# Gaia-X framework



<https://gaia-x.eu/gaia-x-framework/>

# Gaia-X Compliance / Gaia-X Label

*(how to objectivize trust)*



- Based **W3C Verifiable Credentials + Linked Data (RDF)** = Gaia-X credentials
- Gaia-X credentials = building blocks of a **decentralised** machine readable **knowledge graph**
- Extensible knowledge graph by federations and dataspaces

```
{
  "@context": {"gx": "https://registry.gaia-x.eu/..."},
  "issuer": "did:example:SecNumCloudNotary",
  "credentialSubject": {
    "@type": "gx:qualification",
    "gx:schema": "https://www.ssi.gouv.fr/.../secnumcloud-referentiel-exigences-
v3.2.pdf",
    "gx:schemaOwner": {"@id": "http://example.com/anssi", "@type": "gx:Participant"},
    "gx:auditor": {"@id": "http://example.com/LSTI", "@type": "gx:Participant"},
    "gx:scopeName": "myTrustedCloud",
    "gx:scope": [
      {"@id": "http://example.com/CSP/datacenter1", "@type": "gx:PhysicalResource"},
      {"@id": "http://example.com/CSP/serviceoffering1", "@type": "gx:ServiceOffering"}
    ]
  },
  "evidence": {
    "@id": "https://www.ssi.gouv.fr/.../2022_569_np.pdf",
    "@type": "DocumentVerification"
  },
  "issuanceDate": "...",
  "expirationDate": "...",
  "proof": []
}
```

## Criterion P5.1.4

**For Label Level 3, the Provider's registered head office, headquarters and main establishment shall be established in a Member State of the EU/EEA.**

Already fulfilled with [Criterion P1.1.3](#) and the Gaia-X Compliance is **mandating**:

- the country and administrative area information for legal participant, ie including Providers, Consumers and Federators.

Example of Gaia-X Label Level 3 criterion  
(using CASCO terms and definition)

Example of Gaia-X credentials with linked data claims

# Data Exchange

(anchor enforcement of the negotiation result in the infrastructure)



- Discoverability / Data curation / Transaction traceability / ...
- Policies expressed in ODRL with recommended vocabularies, using Gaia-X credentials

## 7.4.1 Legitimate Processing of Information Related to PII

Attribute	Card.	Trust Anchor	Comment
legalBasis	1	dataController	One of the reasons as detailed in the identified Personal Data Protection Regimes, formatted as a string matching <code>&lt;protectionRegime&gt;:&lt;article number&gt;</code>
dataProtectionContactPoint	1..*	dataController	ContactPoint of the Data Protection Officer or Participant responsible for the management of personal or sensitive data
purpose[]	1..*	dataController	Purposes of the processing. It is recommended to use well known controlled vocabulary such as the Data Privacy Vocabulary: Purposes
consentWithdrawalContactPoint	1..*	dataController	ContactPoint of the Participant to whom formulate a withdrawal consent request

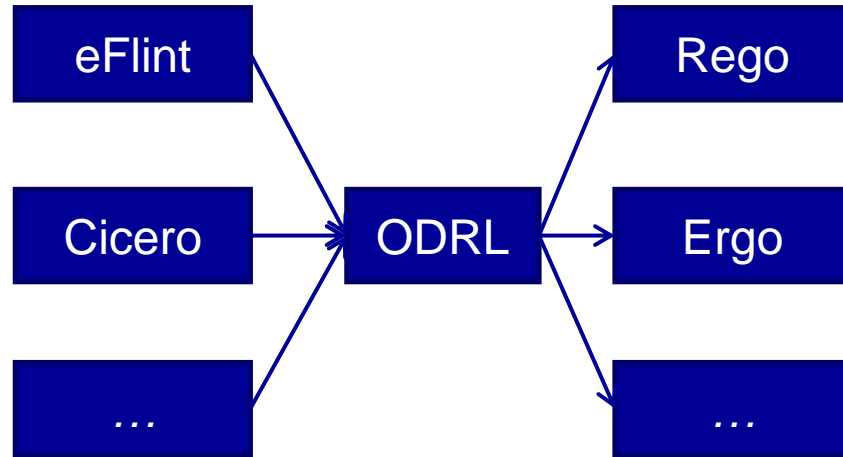
### Consistency rules

- the keypair used to sign the Data Resource claims must be traceable to the `producedBy` participant of the Data Resource.
- If the data are about data subjects as one or more Natural Persons, or sensitive data as defined in GDPR [article 9](#), then `dataController` and `consent` are mandatory.  
To avoid [data re-identification](#), this rule applies independently if the data is raw, pseudo-anonymized or anonymized. (Note: This is on purpose beyond GDPR requirements.)
- if `dataController` is specified, the keypair used to sign at least the Data Resource `consent` claims must be traceable to the `dataController`.
- Generic `authorisation` and `purpose`, not specific to PII nor consent should be expected inside the inherited `policy` attribute using ODRL [Permission](#) and ODRL [Duty](#) rules.

Example of mandatory rules

# Data Exchange

*(anchor enforcement of the negotiation result in the infrastructure)*



W3C Open Digital Rights Language  
as intermediate language

*Confidential Computing, Homomorphic encryption,  
Federated Learning, Compute to Data, secure multiparty  
computation*

Example of technical enforcements

```
@prefix odrl: <http://www.w3.org/ns/odrl/2/> .
@prefix oac: <https://w3id.org/oac#> .
@prefix dpv: <https://w3id.org/dpv#> .
@prefix cert: <http://www.w3.org/ns/auth/cert#> .
@prefix : <http://example.com> .

:app-1 a odrl:Policy ;
odrl:profile oac: ;
odrl:permission [
  a odrl:Permission ;
  odrl:assignee [
    a oac:DataController ;
    cert:key <https://example-app-1.com>
  ] ;
odrl:target oac:EmailAddress, oac:SocialNetwork ;
odrl:action oac:Use, oac:Store ;
odrl:constraint [
  odrl:leftOperand oac:Purpose ;
  odrl:operator odrl:isA ;
  odrl:rightOperand dpv:RegistrationAuthentication
]
] .
```

Example of ODRL policies in turtle



# Gaia-X Digital Clearing House / Federation

*(decentralised ecosystem of federations)*



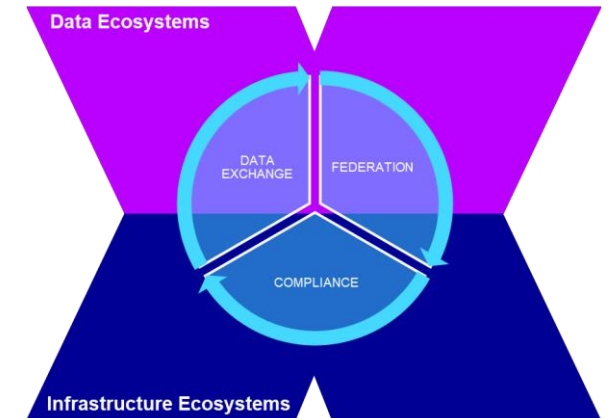
- Provide the **decentralised infrastructure** platform to be used by the **Dataspace Authorities**.
- Steer for a **no lock-in nor lock-out** effect around a handful of service providers
  - catalogues, registries, data connectors, ...
- Facilitate onboarding with reference implementation
  - Wizard, cloud wallet/agent, ...
- Provide **measurability** and service offering **comparison**
  - KPIs: Veracity / Transparency / Semantic match indexes
- **Open source** and based on existing standards:
  - RDF and shape validation (SHACL/SPARQL), **EBSI** API, W3C, ...

# Summary



## - Gaia-X

- is a **governance** for cross-dataspace exchanges
- embraces **data and infrastructure** together
- aim to **reduce** European market **fragmentation**
- working and aligning with other projects/initiatives: BDVA / DSSC / Fiware / IDSA / iShare / X-Road / ...



## - The Gaia-X **governance**

- is based on **open standards** and an **open-world assumption**
- provides **measurable and comparable indexes**
- is operationalised on a **decentralised** infrastructure
- enable users to **self-determine** their level of technical, operational and legal **autonomies**.



# i3-MARKET

H2020-ICT13-2019-2020-IA

Project Identifier: 871754

## i3-MARKET Project Overview

“Data Markets for Digital Sovereignty”

26 April 2023 14:00-17:00 Hrs. CET

D-SME Panel Participation

Presented by Dr. Martin Serrano



Marketplaces



Open Source



Security, Privacy  
and Trust



Semantics



Intelligent, Interoperable, Integrative and deployable  
open source MARKETplace backplane with trusted and secure  
software tools for incentivising the industry data economy



OLLSCOIL NA GAILLIMHE  
UNIVERSITY OF GALWAY

SIEMENS AtoS



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH



guardtime



IBM

GFT



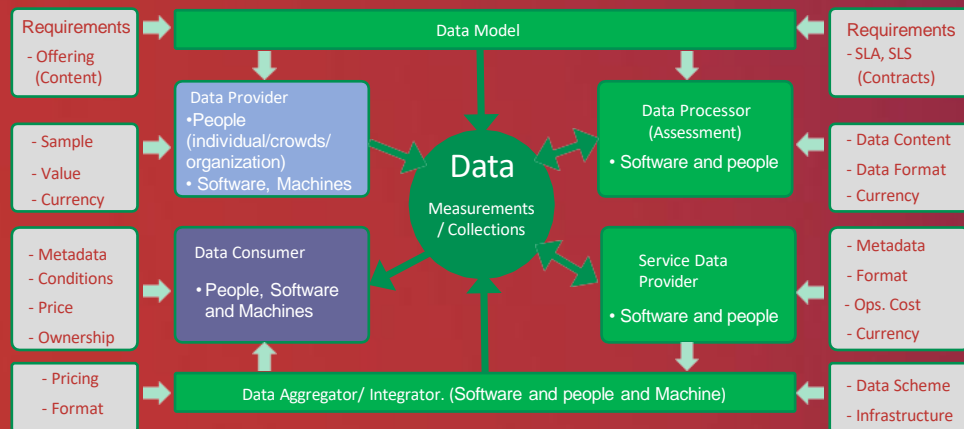
libelium

UNPARALLEL



Telesto  
Technologies  
Information Communication Technologies

# A Data-Driven Society



\*Produced by i3-MARKET Consortium 2019

# Needs a Sustainable Data Economy



# Why Data Sovereign is required now?

Data Producers



Data Consumers



Financial Benefits

≠

Better Services

## i3-MARKET Supports a New Data Economy Paradigm



# Data Marketplace Characteristics/Features



Towards →

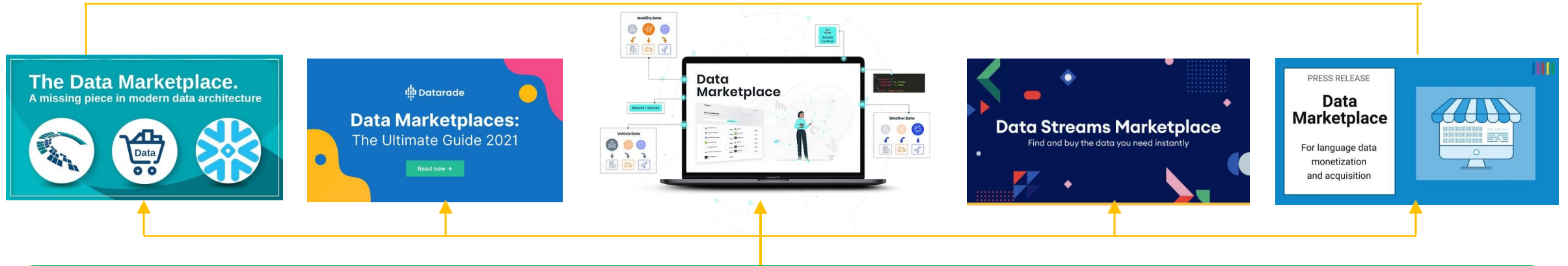


# Single Digital Market

- Exchangeable
- Adaptable
- Open
- Adaptive
- Self-Descriptive
- Scalable
- Secure
- Extensible
- Self-Contain
- Structural



## Demand Side



**Transparency, Trusted Economy, Large Ecosystem**  
(i3-MARKET Marketplace(s) Support Tools Including online Data Recommendation Pricing Tool)

## Data Bases

- Collection
- Storage
- Organised
- Structures
- Language

## Data Warehouses

- Management system
- Support business intelligence (BI)
- Analytics
- Perform Queries
- Large amounts historical data

## Data Lakes

- Information system
- Store Structured/Unstructured Data
- Data Storage as it is
- Perform Analytics
- Real-Time & Machine Learning

## Data Hubs

- Collection of data
- Multiple sources
- Organized for Distribution and Sharing mainly purpose

## Supply Side

# i3-MARKET Project Identity



# i3-MARKET Resources

from i3-MARKET idea design (2020) to industrial innovation (2023)

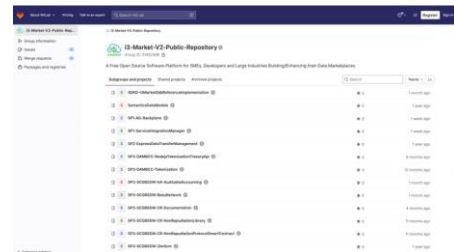
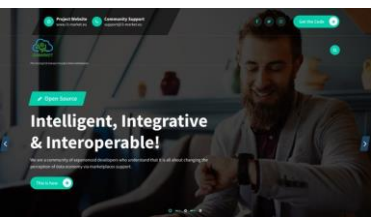
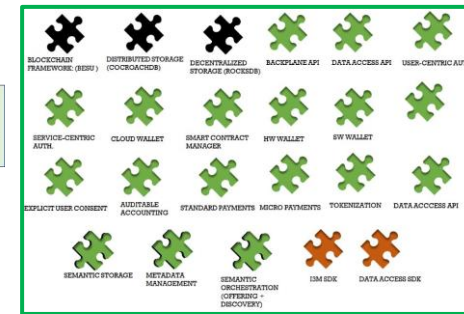
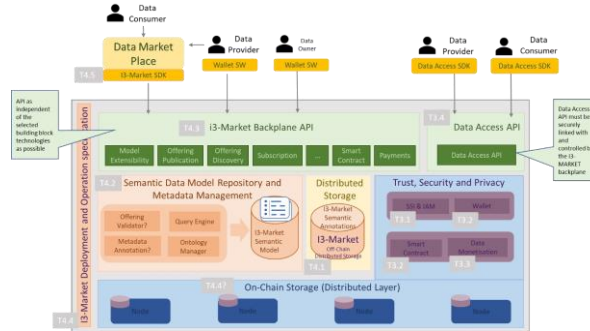
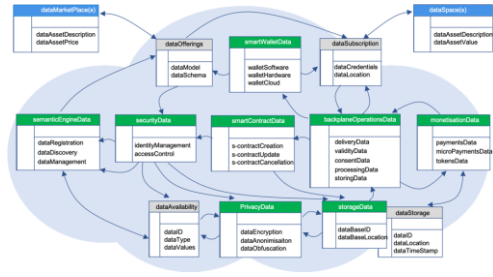
Concepts DaaS  
Design Principles

Data Flows  
Requirements  
Elicitation  
Self-Sovereign

Trustworthy  
Functional  
Architecture

Software Artefacts  
Web-RI, SDK-RI  
Docker Compose

Systems & Sub-Systems  
Industrial Applications with  
Data Marketplaces



Specifications (DaaS)

Release 1 (DaaS)

Release 2 (Tools & Components)

Release 3 (Industrial Pilots)

2020

2021

2022

2023



# i3-MARKET

H2020-ICT13-2019-2020-IA  
Project Identifier: 871754

## i3-MARKET Project Overview

“Data Markets for Digital Sovereignty”

26 April 2023 14:00-17:00 Hrs CET  
D-SME Panel Participation

*thank you!*



Marketplaces



Open Source



Security, Privacy  
and Trust



Semantics



[www.i3-market.eu](http://www.i3-market.eu)

[www.open-source.i3-market.eu](http://www.open-source.i3-market.eu)

i3-market public repository available in  GitLab and  GitHub



OLLSCOIL NA GAILLIMHE  
UNIVERSITY OF GALWAY

**SIEMENS Atos**



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH



**guardtime** 

**IBM**

**GFT** 

 **libelium**

**UNPARALLEL**





# Data Markets for Digital Sovereignty

## Session 2

- Ciarán Donohue, Legal and Policy Officer, Data Protection (DG JUST, European Commission)
- Maria Papaphilippou, Cybersecurity Officer (European Union Agency for Cybersecurity (ENISA))
- Dipl.-GeoInf Christian Linder, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Gruppenleiter Architekturprinzipien und Plattformsysteme (Germany)



# Data Markets for Digital Sovereignty

## Conclusions

**Take aways:** Sen. Prof. Antonio Nicita, Member of the Italian Parliament, former AGCOM Commissioner (Italy) & Former Chairman (European Commission Impact Assessment Board)

**What is next:** Antonio Grasso, EU Policy Director (European DIGITAL SME Alliance)



# Data Markets for Digital Sovereignty

**THANK YOU**

