Luxembourg AI Factory:

YOUR ONE-STOP SHOP FOR AI ADOPTION

The Luxembourg AI Factory is the one-stop shop **empowering businesses** of all sizes and maturity levels to harness the transformative power of artificial intelligence.

Anchored by the sovereign **MeluXina supercomputer** and a holistic public-private ecosystem, we deliver personalized guidance and AI-ready infrastructure.

From AI maturity assessments and tailored training to regulatory compliance, our end-to-end services ensure your AI journey is **secure, trustworthy, and aligned with European standards.**

Who are the partners behind the Luxembourg AI Factory

Five core partners working together to support your **AI journey**

Luxinnovation
Your single entry point: guides you through funding opportunities and connect you to the right partners, and coordinate all business support services across the national AI ecosystem.

LuxProvide
Gives you access to the MeluXina supercomputer and the high-performance computing power needed to scale your AI solutions.

LNDS
Helps you access high-quality, compliant data for your AI projects.

University of Luxembourg
Develops AI skills and talent through training and research programmes.

LIST
Brings research expertise to turn innovative ideas into practical solutions.
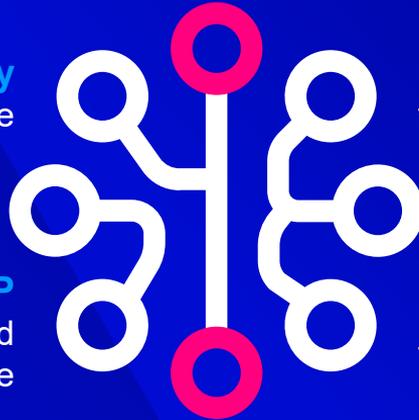
Who are the partners behind the Luxembourg AI Factory

With seven associate partners

Luxembourg House of Cybersecurity
for cybersecurity sector expertise

Technoport for startup incubation

Clarence and DEEP
for national sovereign cloud
and connectivity infrastructure

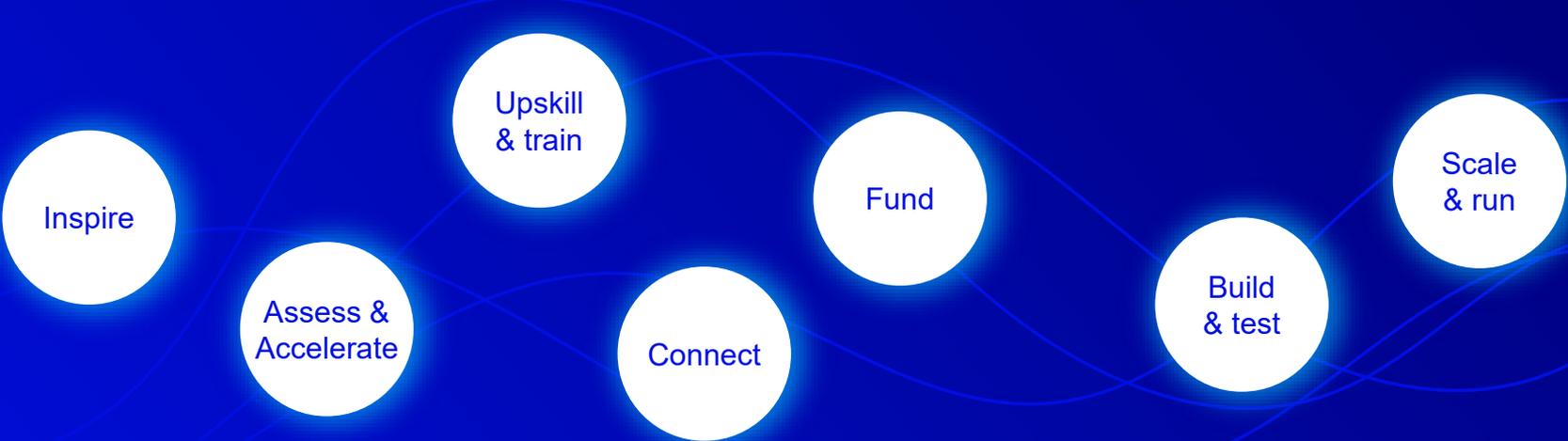**Digital Learning Hub**
for AI education and training

Luxembourg House of Financial Technologies for finance sector expertise

Luxembourg Space Agency
for space sector expertise

Service catalogue

DISCOVER & GET READY

Roadmap definition
Technical AI Scoping
Trustworthy AI Coaching
Online AI Maturity Assessment

Acceleration & Assessment

programmes
Fit4Start – Fit4Scale – Fit4AI

Data discovery
Data Governance Support
Data Access Facilitation
Data Access Request Management

Matchmaking
Enterprise & Research Center
EU AI Factories
EU Initiatives
AI Testing & Experimentation Facilities
European Digital Innovation Hubs
European Digital Infrastructure Consortia
European Common Data Spaces

Assess & Accelerate

Upskill & train

Connect

Fund

AI tech training
Programme for "AI Practitioners"
User Training

Data protection training
Compliance and legal regulations

Skills Watch
Relevant training based on skills
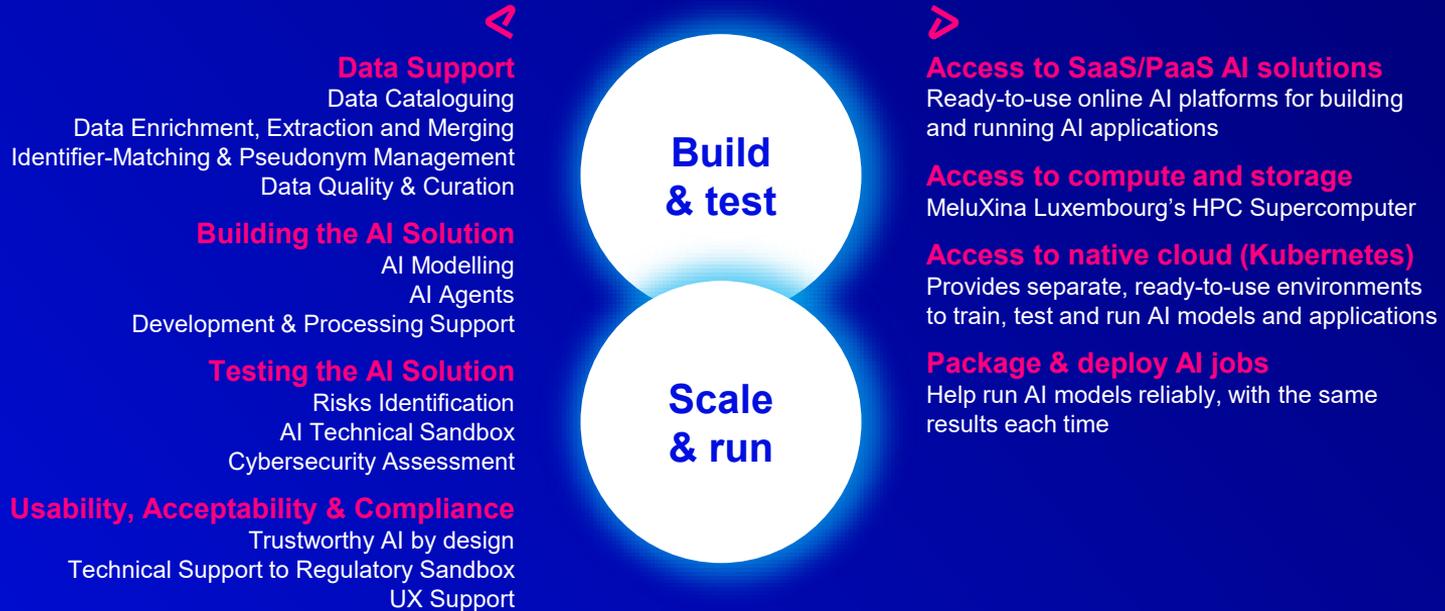
National funding
Competitive
Noncompetitive

European funding
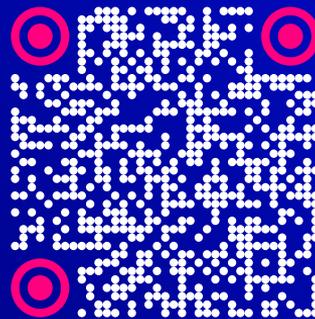Digital Europe Programme
Horizon Europe

Service catalogue

Data Support
Data Cataloguing
Data Enrichment, Extraction and Merging
Identifier-Matching & Pseudonym Management
Data Quality & Curation

Building the AI Solution
AI Modelling
AI Agents
Development & Processing Support

Testing the AI Solution
Risks Identification
AI Technical Sandbox
Cybersecurity Assessment

Usability, Acceptability & Compliance
Trustworthy AI by design
Technical Support to Regulatory Sandbox
UX Support

Build & test

Scale & run

Access to SaaS/PaaS AI solutions
Ready-to-use online AI platforms for building
and running AI applications

Access to compute and storage
MeluXina Luxembourg's HPC Supercomputer

Access to native cloud (Kubernetes)
Provides separate, ready-to-use environments
to train, test and run AI models and applications

Package & deploy AI jobs
Help run AI models reliably, with the same
results each time

Why a datasets catalogue for L-AIF?

Datasets specifically suited to AI contexts

What is "AI-ready" data?

- Dataset size: min. of several 100k records or data points

- Dataset quality: minimal quality standards including
  - conformance with domain-relevant standards
  - completeness and documentation of any pre-processing

DATA
=
VALUE

Preference towards **non-open data but all datasets are welcome!**

Use of a dataset in **successful AI model development or testing** is a bonus
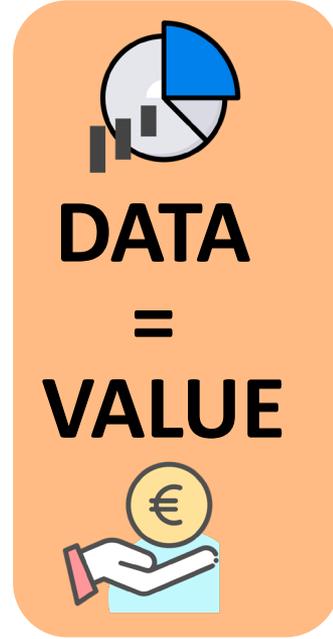
Data cataloguing ingredients

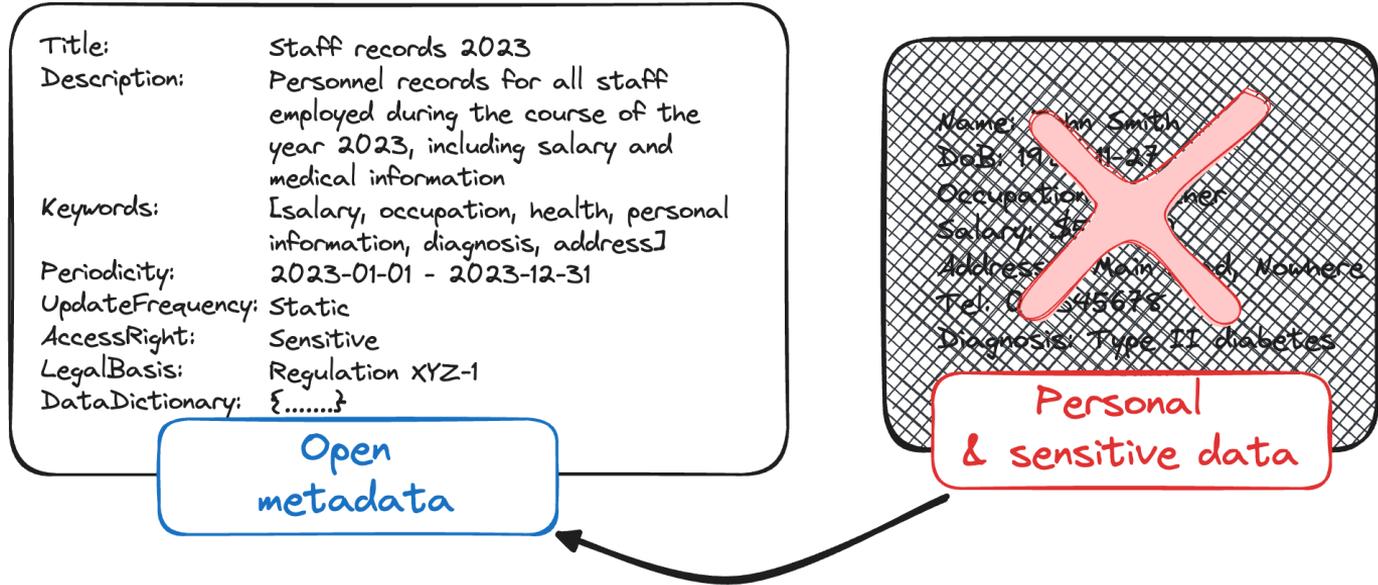|	
Metadata* model	What metadata do you need to collect?
Cataloguing infrastructure	Where do you collect the metadata?
Cataloguing processes	How and when will the metadata be collected?

* Metadata:
- is "data about data", i.e. describes the characteristics of a dataset
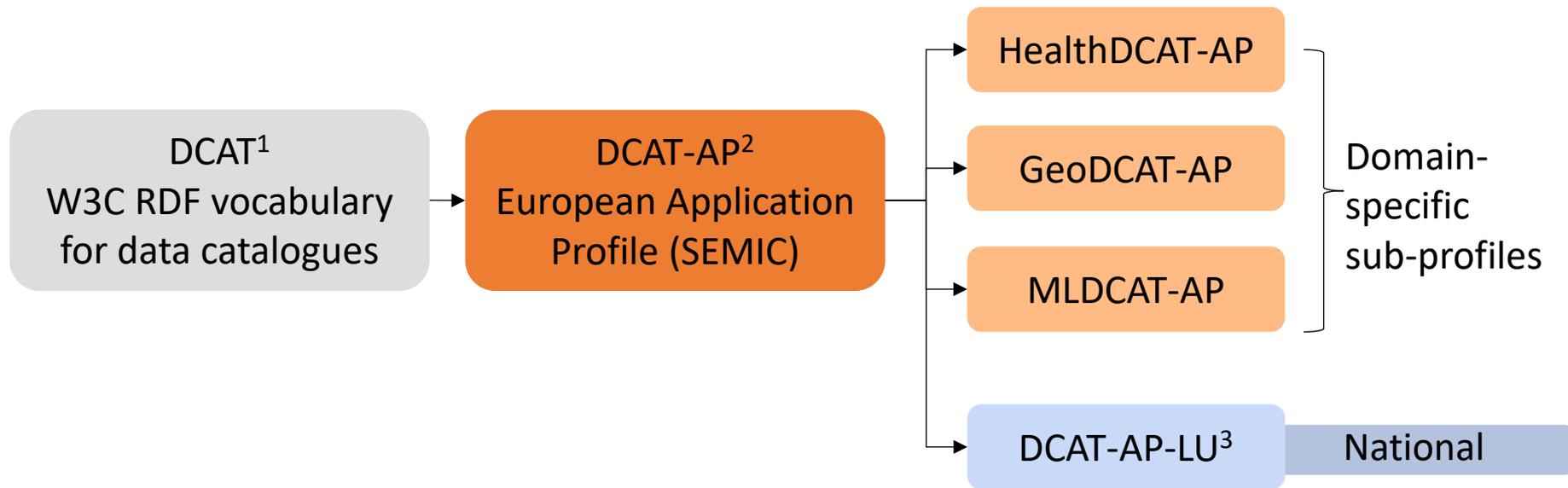- helps users find, understand and use datasets

Restricted data yet open metadata

Title: Staff records 2023
Description: Personnel records for all staff
                    employed during the course of the
                    year 2023, including salary and
                    medical information
Keywords: [salary, occupation, health, personal
                    information, diagnosis, address]
Periodicity: 2023-01-01 - 2023-12-31
UpdateFrequency: Static
AccessRight: Sensitive
LegalBasis: Regulation XYZ-1
DataDictionary: {.......}

Open metadata

Name: John Smith
DoB: 19... ...-2?
Occupation ...ner
Salary: $6...
Address: Main ...d, Nowhere
Tel. ...45675
Diagnosis: Type 11 diabetes

Personal & sensitive data

Metadata describes the data but does NOT disclose its content

The Data Catalog Vocabulary (DCAT) metadata ecosystem
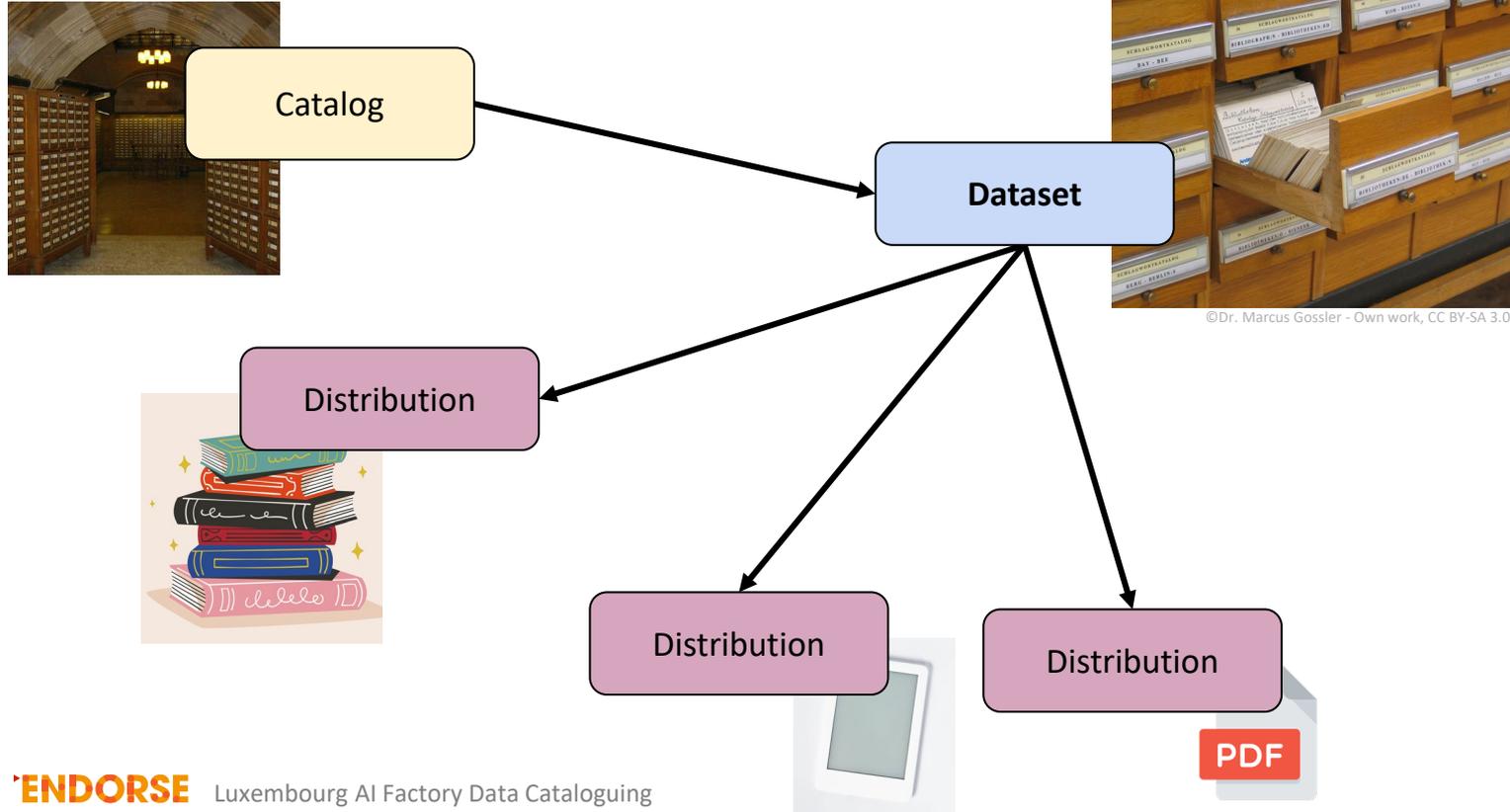
DCAT – Data Catalog Vocabulary

DCAT focuses on domain-agnostic properties to describe datasets

1 DCAT: https://www.w3.org/TR/vocab-dcat-3/
2 DCAT-AP: https://semiceu.github.io/DCAT-AP/releases/3.0.1/
3 DCAT-AP-LU: https://mindig_lu.gitlab.io/DCAT-AP-LU/, 2.0.0

DCAT - core concepts

Catalog

Dataset

©Dr. Marcus Gossler - Own work, CC BY-SA 3.0

Distribution

Distribution

Distribution

MLDCAT-AP

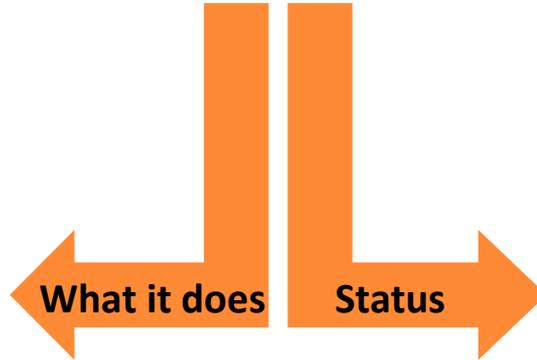DCAT-AP → **is extended by** → MLDCAT-AP → **usable in** → AI Factory

What it does | **Status**

- *Reuses* standards such as ML Schema (ML-S)
- *Adds* ML-specific entities: algorithms, ML models, benchmarks, performance metrics and risk metadata

- Under development: collaboration between SEMIC and OpenML
- Specification: https://semiceu.github.io/MLDCAT-AP/releases/3.0.0/

Luxembourg AI Factory Datasets Catalogue

https://data.aifactory.lu/

L-AI Factory Datasets Catalogue I

LUXEMBOURG AI FACTORY

Home   Datasets   Themes   Publishers   About

Luxembourg AI Factory Datasets Catalogue

Discover AI-ready data in Luxembourg

Search datasets

Science and technology
12 datasets
See datasets

Population and society
1 dataset
See datasets

Environment
6 datasets
See datasets

L-AI Factory Datasets Catalogue II

Catalogue features	
Basic	*Extended*
Data discovery via search + filter, summary metadata views, machine-readable metadata exports	"Shopping basket" for submitting data access requests (not currently available)

Metadata collection:

LNDS in-house MetaData Capture tool (MDC):
https://metadata.demo.lnds.tech/

The Legal Large Language Model (4LM) project

Legal Large Language Model (4LM) project

| What is it? | → | • Writing and review assistance for legal experts |

| Who is involved? | → | • In the scope of a wider governmental AI initiative
• LNDS & LuxProvide (LXP) collaborate on the proof of concept
    • LXP infrastructure (MeluXina)
    • LNDS (RAG pipeline) |

Status:
RAG (retrieval-augmented generation) is in place.
Now tweaking, tuning, alpha tester feedback

Get in touch

Dr. Irina Balaur

irina.balaur@lnds.lu

service@lnds.lu

aifactory.lu