aindo

AI & (Synthetic) Data:

A New Paradigm in Finance

Meet Alessandro di Pietro

Head of Business Development @ Aindo

Education

- Tor Vergata Biomedical Engineering
- Master II Procurement Management, Tor Vergata Business School
- MBA Digital Transformation, Ticinensis Pavia University

Expertise & Professional Background

- Expert in AI-driven data innovation for financial institutions
- Specialist in structured synthetic data for highly regulated markets (Finance, Healthcare, Telecoms)
- Secured €10M+ in AI innovation grants & VC funding
- Experience with banks, regulators & fintechs on AI solutions

AI & Synthetic Data are revolutionizing finance — my goal is to help you understand their role in data accessibility, privacy, and innovation

Connect with me

LinkedIn

alessandro@aindo.com

Agenda

1. **In Aindo Overview:** *Building a Deep Tech Startup*
2. **The Challenge:** *Accessibility and Quality of (Sensitive) Data*
3. **Generative AI for Structured Data:** *A Multidisciplinary Solution*
4. **Synthetic Data:** *Application*
5. **AI in Finance:** *Insurance Focus*
6. **Business Models for AI and Data in Europe**
7. **Interactive Discussion and Use Cases**

Conclusions & Next Steps

1.

Aindo Overview:

Building a Deep Tech Startup

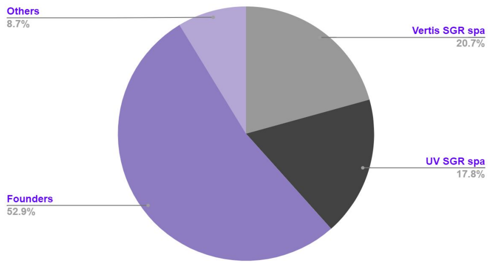
Aindo in a nutshell

SYNTHETIC DATA: THE NEW GLOBAL STANDARD FOR DATA ACCESS AND SHARING

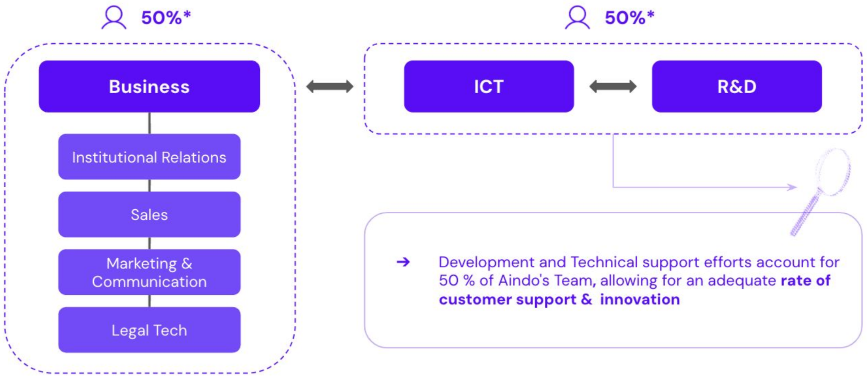
Company Info

- Aindo SpA (corporation)
- Founded in 2018 as a SISSA start-up
- 2 Co-founders
- 29 Talents (27% PhD, 65% STEM, 7 nationalities)
- Headquarter in Trieste (TS), Italy
- Powered by the support of two leading Venture-Capital (United Venture SGR SpA and Vertis SGR spa) – € 6M series A in Q4 2023

Shareholders



Organizational Structure



→ Development and Technical support efforts account for 50 % of Aindo's Team, allowing for an adequate **rate of customer support & innovation**

* % of employees within a specific division

Growth traction

2018 Q1 — **Founded in 2018**
as a SISSA startup

SISSA

2020 Q2 — **European Data Incubator Winner**
(1 out of ~500)
& €100k grant secured for SD generator development

2021 Q4 — Seed Investment

VERTIS SGR

2022 Q3 — **Awarded at the Noovle Cloud Challenge**
promoted by TIM and Google Cloud

noovle

2022 Q4 — **Attended NEURIPS**
the most important AI conference

NEURAL INFORMATION PROCESSING SYSTEMS

2023 Q2 — Patent for innovative synthetic data generation technology

Awarded by Fortune Italia

FORTUNE

2023 Q3 — **Papers publication**

Mentioned by CBInsights
(Responsible ai market map)

2023 Q4 — **Series A investment**
6.000.000 €

VERTIS SGR

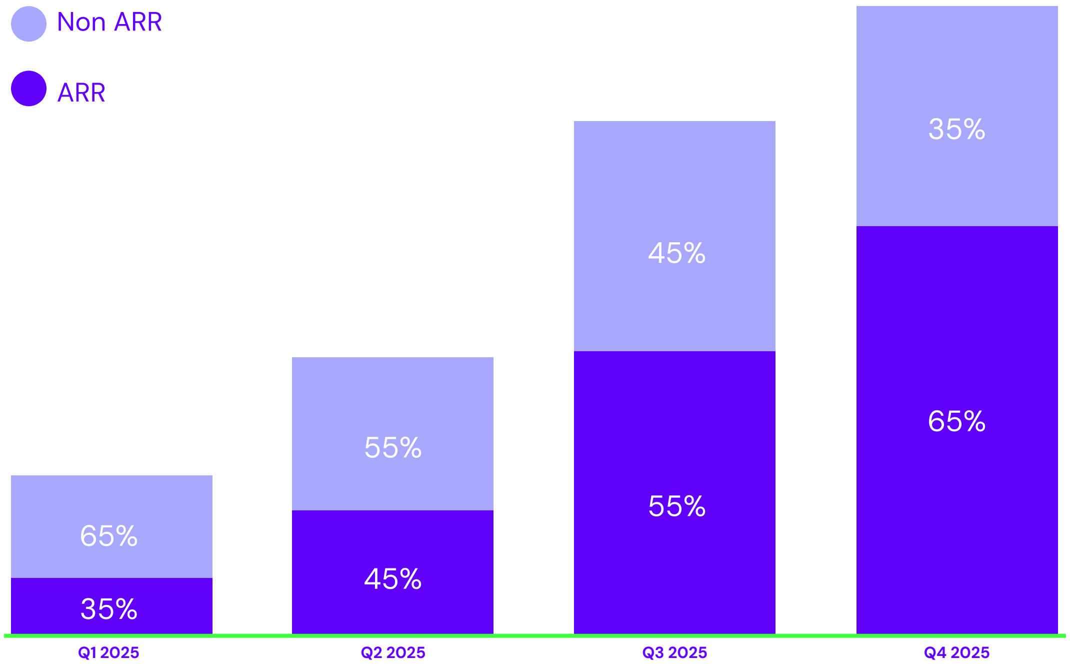
united ventures

2025 Q1 — **Selected for funding by EIC**
2.100.000 €

European Innovation Council

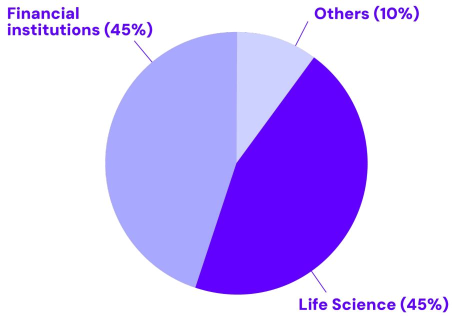
Financials & Key Metrics

Quarterly revenue target 2025



Total fiscal year 2024: 2 Million Revenues (o.w. 1 Million ARR)

Market by industry 2025



Partners & Customers

TIM ENTERPRISE · ARIA AZIENDA REGIONALE PER L'INNOVAZIONE E GLI ACQUISTI · iFAB · AREA SCIENCE PARK

bip. · FANTIX · IVASS ISTITUTO PER LA VIGILANZA SULLE ASSICURAZIONI

Televita · BE Rebel · BitBang

SISSA · National Innovation Centre Ageing · MIA CARE · insiel

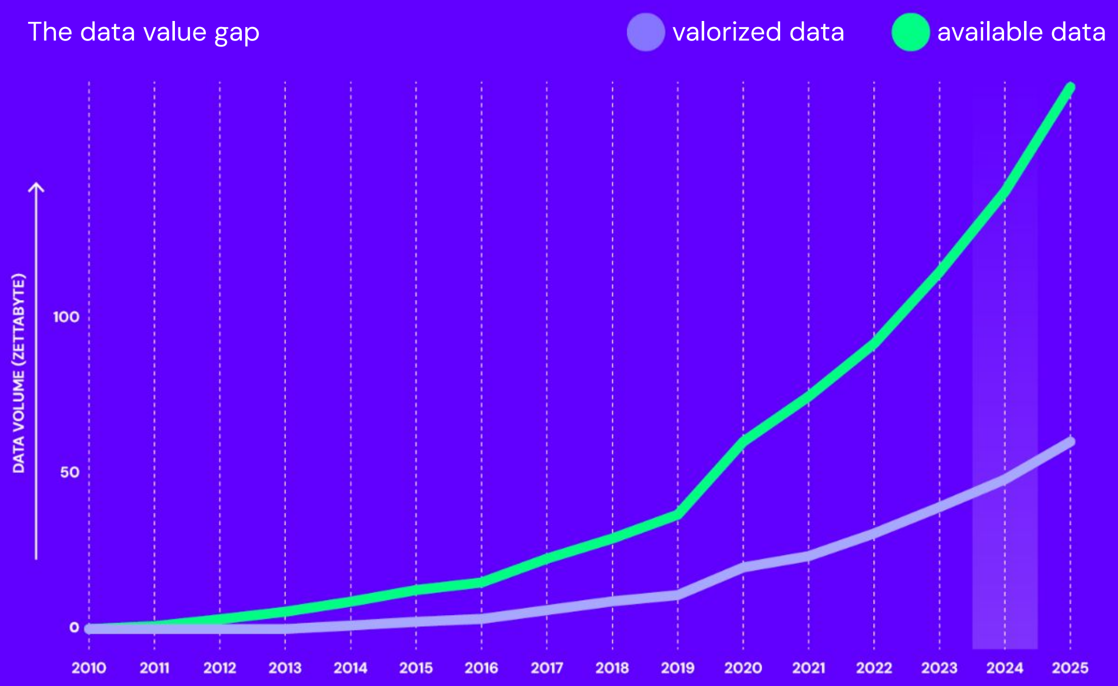
2.

The Challenge:
Accessibility and Quality of (Sensitive) Data

The Challenge

60% to 73% of all potential data value is currently not realized

The data value gap



(Source: Statista; Accenture)

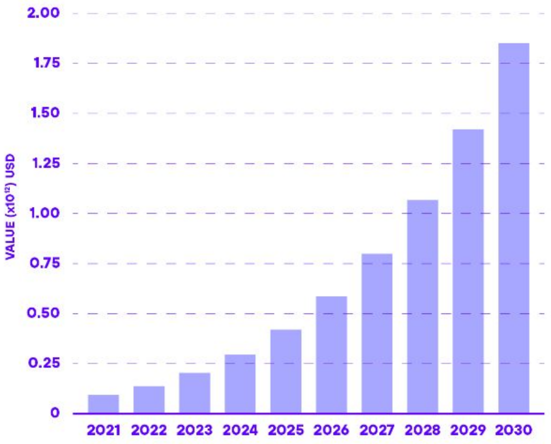
Core obstacles

- **Data distribution**
  Data is stored in different formats and systems, making it difficult to access and integrate.
- **Formatting**
  Data is often collected in unstructured, unannotated formats.
- **Privacy**
  Privacy legislation (GDPR, CCPA, DP) hinder data exchange and value extraction.
- **Incompleteness**
  Data may contain crucial gaps, leading to inaccurate analytics and models
- **Bias and unfairness**
  Datasets often underrepresent specific fragments of society, resulting in biased AI models

AI projected economic & data value by 2030

AI will add 1.5T USD to the global economy by 2030*

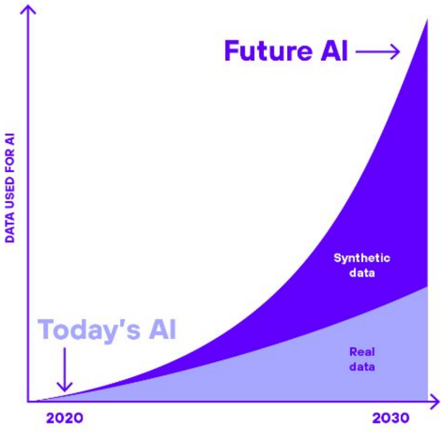
Forecast:
value of AI sector globally



*Source: Statista

By 2030, 90% of AI models will be trained on synthetic data**

Forecast:
use of synthetic data in AI training



**Source: Gartner

EU Normative landscape

EU Artificial Intelligence Act

- Synthetic data should be **preferred over pseudonymized data** or other categories of personal data
- Synthetic data **fall outside the scope of GDPR**, as they are recognized as **non-personal data**
- AI Act 'prefers' use of **synthetic data over** so-called **'sensitive' data when it comes to correcting biases** in high-risk systems

DGA Data Governance Act

- Member States should make optimal **use of** technologies like **synthetic data for privacy protection**

When is data personal?

General Data Protection Regulation (GDPR) came into effect in 2018

GDPR limits use of personal data

GDPR does not apply to **anonymous data**

EU 2016/679
"Personal data" is one of twelve attribute categories

- Name
- Identification number
- Location data
- Online identifier
- Specific physical
- Physiological
- Genetic
- Biometric
- Mental
- Economic
- Cultural
- Social identity data

Anonymized: WP29 working definition

WP29 defines "anonymized" as not susceptible to the following three risks

Singling out

Attacker singles out an individual (in an anonymized) data

Linkability

Attacker can link an individual's record in the (anonymized) dataset to another record belonging to the same individual

Linkability

Attacker can deduce the value of a parameter from the data.

Example
Singling out attack

I know that Daniel is a 35-year-old man from Milan, living in Naples.

Can we figure out which record is his Daniel's?

Name	Age	Gender	Town of origin	Residence
Hidden	34	F	Lecce	Parma
	35	M	Milan	Turin
	34	F	Parma	Parma
	35	M	Milan	Naples
	34	F	Palermo	Naples

Example
Singling out attack

I know that Daniel is a 35-year-old man from Milan, living in Naples.

Can we figure out which record is his Daniel's?

Name	Age	Gender	Town of origin	Residence
Hidden	34	F	Lecce	Parma
	35	M	Milan	Turin
	34	F	Parma	Parma
	35	M	Milan	Naples
	35	F	Palermo	Naples

Example
Singling out attack

I know that Daniel is a 35-year-old man from Milan, living in Naples.

Can we figure out which record is his Daniel’s?

Name	Age	Gender	Town of origin	Residence
Hidden	34	F	Lecce	Parma
	35	M	Milan	Turin
	34	F	Parma	Parma
	35	M	Milan	Naples
	35	F	Palermo	Naples

Example
Singling out attack

I know that Daniel is a 35-year-old man from Milan, living in Naples.

Can we figure out which record is his Daniel's?

Name	Age	Gender	Town of origin	Residence
	34	F	Lecce	Parma
	35	M	Milan	Turin
Hidden	34	F	Parma	Parma
	35	M	Milan	Naples
	35	F	Palermo	Naples

Example
Singling out attack

I know that Daniel is a 35-year-old man from Milan, living in Naples.

Can we figure out which record is his Daniel's?

<table>
<tr><th>Name</th><th>Age</th><th>Gender</th><th>Town of origin</th><th>Residence</th></tr>
<tr><td rowspan="5">Hidden</td><td>34</td><td>F</td><td>Lecce</td><td>Parma</td></tr>
<tr><td>35</td><td>M</td><td>Milan</td><td>Turin</td></tr>
<tr><td>34</td><td>F</td><td>Parma</td><td>Parma</td></tr>
<tr><td>35</td><td>M</td><td>Milan</td><td>Naples</td></tr>
<tr><td>35</td><td>F</td><td>Palermo</td><td>Naples</td></tr>
</table>

Example
Singling out attack

I know that Daniel is a 35-year-old man from Milan, living in Naples.

Can we figure out which record is his Daniel’s?

Name	Age	Gender	Town of origin	Residence
Jenny	34	F	Lecce	Parma
Sebastian	35	M	Milan	Turin
Samantha	34	F	Parma	Parma
Daniel	35	M	Milan	Naples
Elsa	35	F	Palermo	Naples

Example
Inference attack

Having singled Daniel out, can we learn any new sensitive information about him?

Name	Age	Gender	Town of origin	Residence	Has disease?
Jenny	34	F	Lecce	Parma	No
Sebastian	35	M	Milan	Turin	No
Samantha	34	F	Parma	Parma	No
Daniel	35	M	Milan	Naples	Yes
Elsa	35	F	Palermo	Naples	No

AI Act 2024: harmonised rules adoption roadmap

March 2024	ACT APPROVAL
April 2024	PLENARY SESSION & OFFICIAL APPROVAL
October 2024	PROHIBITED PRACTICES APPLICABILITY
January 2025	CODE OF CONDUCT APPLICABILITY
April 2025	GENERAL RULES & GOVERNANCE APPLICABILITY
April 2026	COMPLETE APPLICABILITY
April 2027	HIGH RISK SYSTEM OBLIGATIONS APPLICABILITY

AI Act harmonised rules also apply to already developed systems

AI Act 2024: a new perspective on Synthetic Data

EU Artificial Intelligence Act

Article 10.5

Article 59

EU legal provisions are interpreted literally in civil law, the legislator equalyze anonymous and synthetic data

Article 10.5

- De facto considers Anonymized and Synthetic Data equivalent and as interchangeable alternatives
- Set a clear distinction with pseudonymised data: these are listed separately and qualified as a different category

Article 59

- Equates Synthetic Data with Anonymized and non-personal data, making them falling out of the scope of data protection legislation (GDPR) and falling out of the definition of personal data

Aindo Synthetic data to avoid data compliance breach

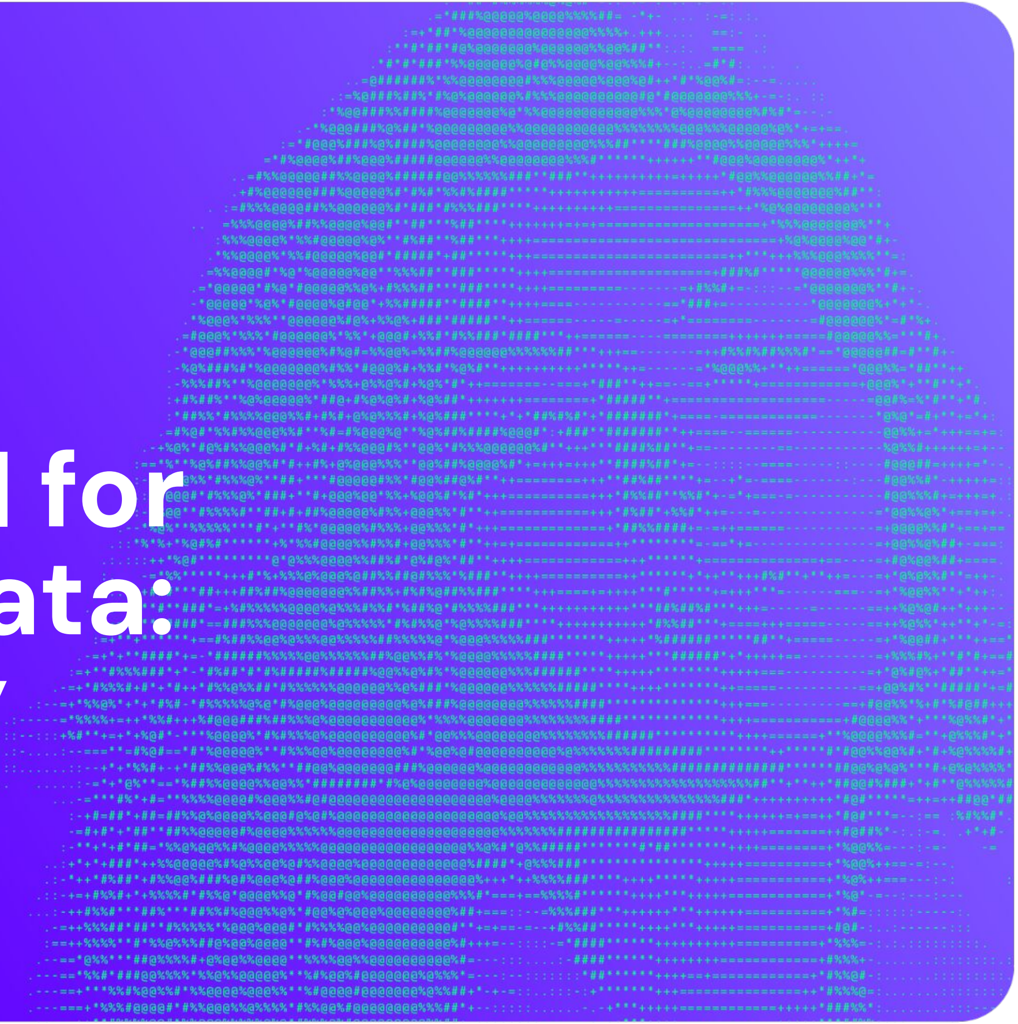
NIST National Institute of Standards and Technology U.S. Department of Commerce

Tested best-in-class by the US National Institute for Standards and Technology
Link: https://pages.nist.gov/privacy_collaborative_research_cycle/pages/archive.html

3.

Generative AI for Structured Data:

A Multidisciplinary Solution



Synthetic data

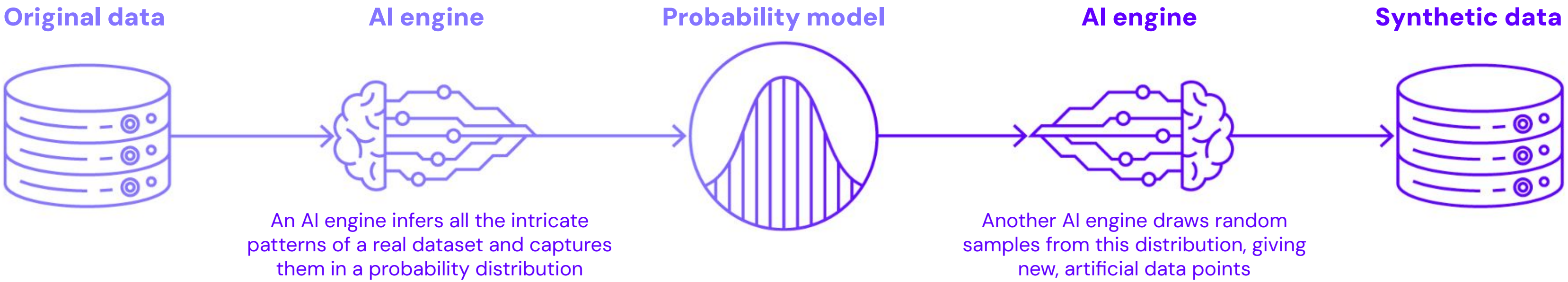
1

Synthetic data are **constructed algorithmically,** they are not collected empirically

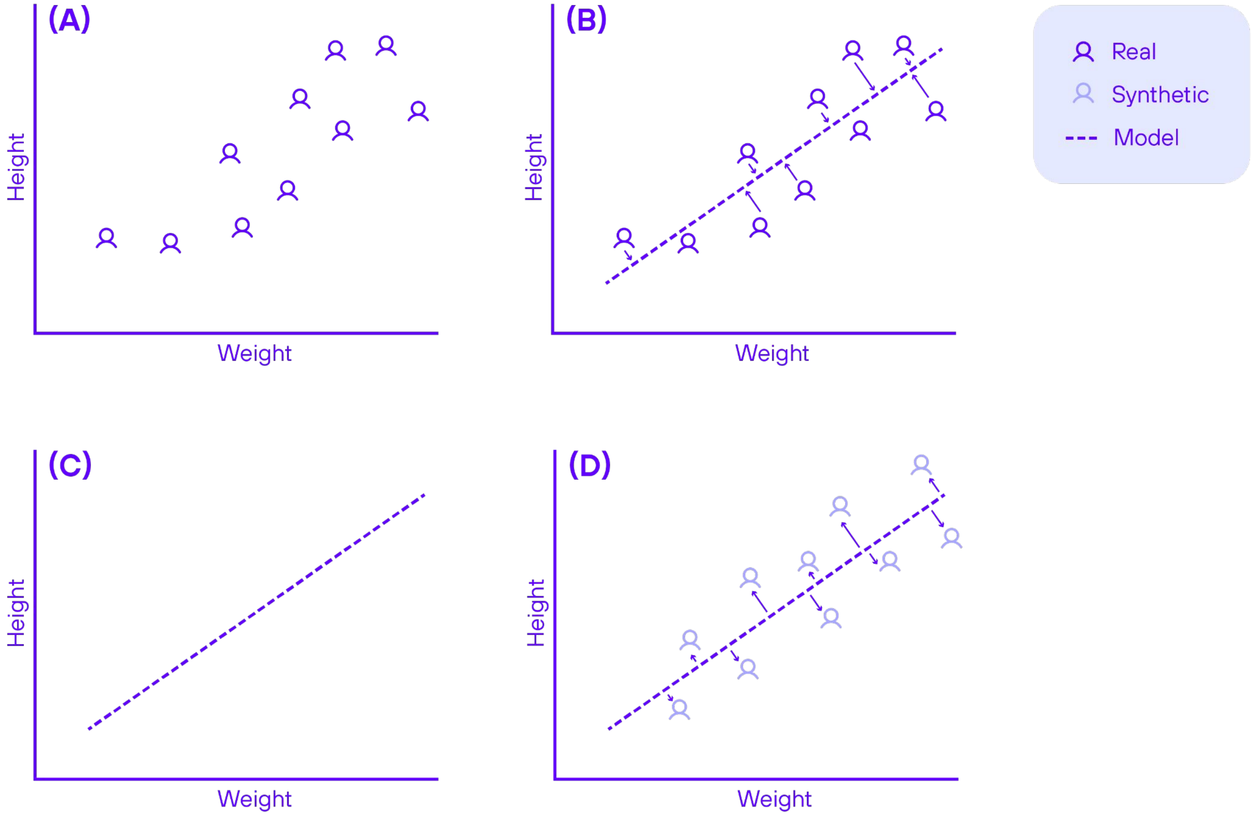
2

Through advances in AI, synthetic data are **indistinguishable from real data**, yet **void of personal information**

Under the hood



Synthetic Data: Replicating real patterns



Generational adversarial networks (GAN)

Generator
criminal

Generates counterfeit money

Counterfeit

Real

Discriminator
police

Discriminates between real and fake bills

Pass
Light mark

Fail
Dark mark

Generational adversarial networks (GAN)

Generator
criminal

Generates
counterfeit money

Counterfeit

Real

Discriminator
police

Discriminates between
real and fake bills

Pass
Light mark

Fail
Dark mark

If the discriminator (police) becomes very good at discriminating, (S)he forces the generator (criminal) to become better at counterfeiting

Generational adversarial networks (GAN)

Generator
criminal

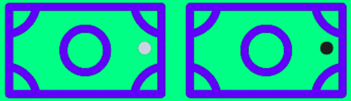
Generates counterfeit money

Counterfeit

Real

Discriminator
police

Discriminates between real and fake bills



Pass
Light mark

Fail
Dark mark

If the generator (criminal) becomes very good at counterfeiting, (S)he forces the discriminator (police) to become very good at discriminating

If the discriminator (police) becomes very good at discriminating, (S)he forces the generator (criminal) to become better at counterfeiting

Generational adversarial networks (GAN)

Generator
criminal

Generates counterfeit money

Counterfeit

Real

Discriminator
police

Discriminates between real and fake bills

Pass
Light mark

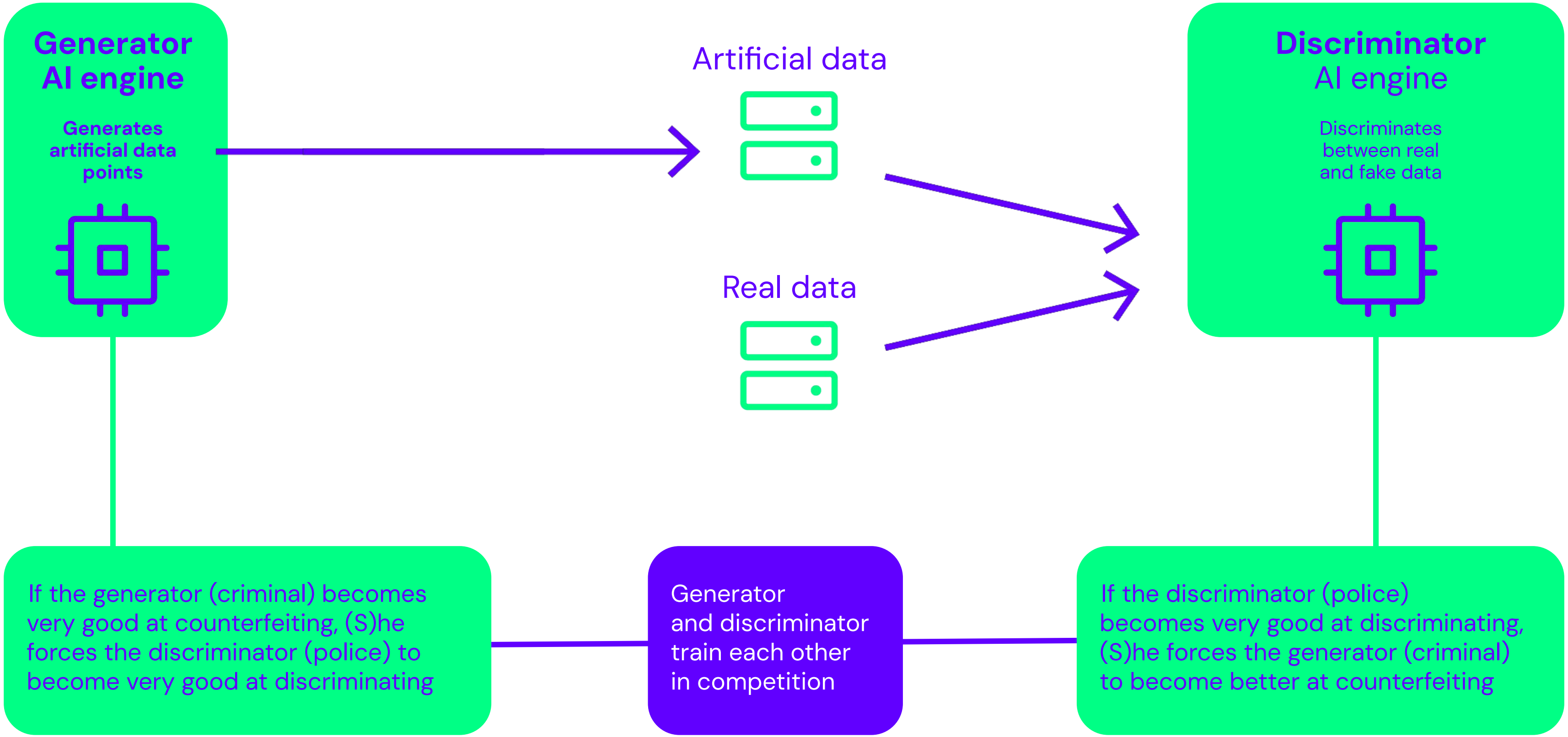
Fail
Dark mark

If the generator (criminal) becomes very good at counterfeiting, (S)he forces the discriminator (police) to become very good at discriminating

Generator and discriminator train each other in competition

If the discriminator (police) becomes very good at discriminating, (S)he forces the generator (criminal) to become better at counterfeiting

Generational adversarial networks (GAN)



Hyper-realistic synthetic data that maintains statistical accuracy and preserves privacy

1

Synthetic data are **constructed algorithmically,** they are not collected empirically

2

Through advances in AI, synthetic data are **indistinguishable from real data**

3

Synthetic data are **void of personal information.** They enhance data **quality** through augmentation by rebalancing and removal of biases.

Patented Technology

Ministero delle Imprese e del Made in Italy

Our **Gen AI patented technology** represents a key differentiator for our company.

Top Tier Technology

NIST National Institute of Standards and Technology U.S. Department of Commerce

Our technology was recognised best-in-class by the **NIST**. Accuracy and privacy outperforms competitors (link).

Europrivacy certification

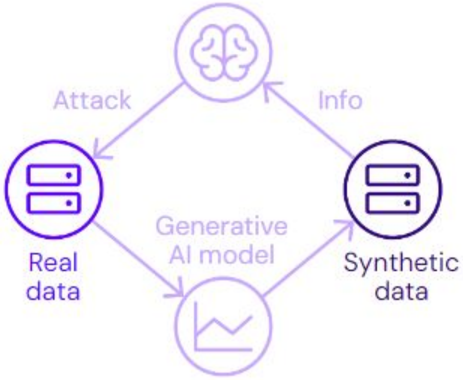
Europrivacy DNV

Our technology was certified by **DNV.** Data processing is in compliance with specification requirements v.77 (link).

Synthetic data: quality assurance

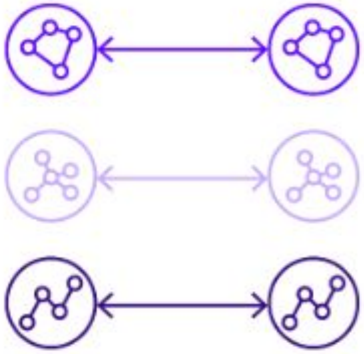
Privacy

How well does synthetic data use protect the privacy of real data subjects?



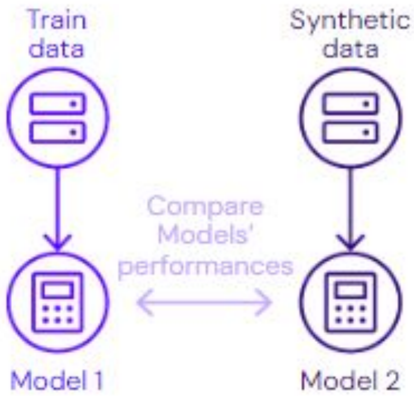
Fidelity

How well does synthetic data preserve real statistical patterns?



Utility

How useful are synthetic data in AI training?



NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

Best performing model, without privacy leak

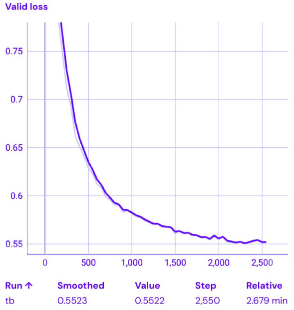
Our technology was recognised best-in-class by the **United States National Institute of Standards and Technology (NIST)** (link).

The basic privacy mechanism of our synthetic data is **generalization control**: the ability of the synthesizer to learn statistical patterns without learning individual records.

Generalization is controlled through early-stop techniques.

The user can tune the protection level setting a more aggressive early-stop.

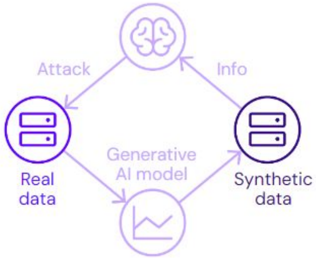
Library	Algorithm	Team	ε	Utility: SSE	Privacy Leak: UEM
aindo-synth	aindo-synth	Aindo		30.0	0.01
Anonos Data Embassy SDK	Anonos Data Embassy SDK	Anonos		30.0	0.01
Genetic SD	Genetic SD	DataEvolution	10	17.5	0.18
Genetic SD	Genetic SD	DataEvolution	1	5.56	0.04
LostInTheNoise	MWEM+PGM	LostInTheNoise	1	10.0	0.0
MostlyAI SD	MostlyAI SD	MOSTLY AI		30.0	0.01
rsynthpop	catall	rsynthpop-categorical	100	65.0	81.33
rsynthpop	ipf	CRC	10	16.67	14.29
rsynthpop	catall_NonDP	rsynthpop-categorical		50.0	63.37
ydata-sdk	YData Fabric Synthetizer	Ydata		11.85	9.7



Data Protection Mechanisms & Metrics

Privacy

How well does synthetic data use protect the privacy of real data subjects?



Mechanisms

Synthetisation

User-controlled privacy mechanism

- Differential Privacy
- Generalisation Control

Privacy Metrics

Distance to Closest Record (DCR) metrics

Deliberate Attacks-Based metrics (Anonymeter)

1st Company to be Europrivacy certified

Europrivacy™®

The certification was also issued following the positive outcome of the inspection by the "Guarantor for the Protection of Personal Data (GPDP)" that took place on May 2024

Supervisor's Inspection Findings

GPDP.Ufficio.PROTOCOLLO.U.0075685.20/06/2024

GPDP
GARANTE PER LA PROTEZIONE DEI DATI PERSONALI

DIPARTIMENTO SANITA' E RICERCA

AINDO S.P.A.
Via Pec
aindo@pec.it

CdS/CL/350777
OGGETTO: accertamento ispettivo sul trattamento di dati personali nell'ambito del processo di produzione di dati sintetici. Esito

In data 30 e 31 maggio 2024 si è svolto un accertamento ispettivo presso codesta Società, al fine di conoscere il processo di generazione di dati sintetici (strutturati e non strutturati) e di verificare tali tipologie di informazioni in relazione ai principi di protezione dei dati personali, con particolare riguardo a quello di minimizzazione. Inoltre, si è inteso verificare l'osservanza delle disposizioni in materia di protezione dei dati personali di cui al Regolamento generale sulla protezione dei dati (UE) 2016/679 (di seguito "Regolamento") con specifico riferimento ai trattamenti svolti per scopi di ricerca scientifica (ordine di servizio del 15 maggio 20224, prot. n. 58759).

L'attività ispettiva è stata verbalizzata e la Società dispone di una copia dei verbali che si intendono integralmente richiamati.

Nella prima giornata di accertamenti, è stato richiesto a codesta Società di illustrare la fase generativa dei dati sintetici a partire dall'addestramento degli algoritmi.

In via preliminare, la Società ha dichiarato che "offre ai propri clienti uno strumento di sintetizzazione che opera sui dataset di cui questi ultimi sono titolari del trattamento e mette a disposizione degli stessi una piattaforma che genera i dati sintetici sulle loro macchine"

EuroPrivacy Certification Issuance

DNV

DATA PROCESSING CERTIFICATE

Certificate No. C682246	Initial date: 22 July 2024	Valid: 22 July 2024 – 21 July 2027

This is to certify that

AINDO S.p.A.

Area Science Park - Località Padriciano, 99 - 34149 Trieste (TS) - Italy

is in compliance with relevant requirements specified by the Certification Scheme:

EUROPRIVACY™/®
Europrivacy™/® Certification Scheme General Specification Requirements v.77

Target of evaluation is in compliance for the following data processing:
Generation of synthetic data for the healthcare sector through the design, development, consulting and configuration on customer's premises of a proprietary AI Platform (Aindo Synthetic Data Platform)

A comprehensive and flexible offering for fast, secure, and reliable data-mobility

What We Offer

Aindo specializes in **relational tabular data**. Our solution offers **best-in-class performance**(1) and advanced features coupled with a user-friendly interface that makes it **accessible to both experts and non-experts**. This allows seamless integration of synthetic data into data-driven projects, even when dealing with complex data types and formats.

Core Features

- **Neural Model Synthesis**: generate complex, relational data and time-series structures using best-in-class models.
- **Broad Data Support:** handle various formats, including text and geolocation.
- **Constrained generation**: rebalance your dataset (GUI, SDK) or introduce generation rules (SDK)
- **Differentially private training**: includes mathematical guarantees of privacy

Available Through

GUI – Synthetic Data **WebApp**

REST **API**

SDK

Comprehensive Documentation(2): detailed guides for configuration and usage via GUI, REST API and SDK.

(1) https://pages.nist.gov/privacy_collaborative_research_cycle/pages/archive.html
(2) https://docs.aindo.com/

4.

AI in Finance: Insurance Focus

Data & AI in Insurance

Market Opportunity of AI Adoption in Insurance

Generative AI market value by 2032	**$1.3** trillion
AI market in insurance valuation by 2033	**$79** billion
Gen AI is vital for Risk Detection & New Business Models	**85%** insurances CEOs

AI POTENTIAL

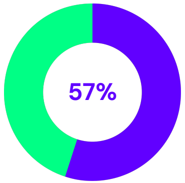
- Claims automation
- Risk assessment
- Personalization
- Combating cybercrime

SLOW ADOPTION

Adoption among insurers remains slow and siloed, due to **concerns** about the **technology's accuracy**

CALL FOR AI STRATEGY

- **C-Suites** (CTOs, CFOs, CIOs & CEOs) to craft AI strategies
- **CAIOs** to rise in Data driven organizations



AI as key technology for achieving business goals in the next 3 years

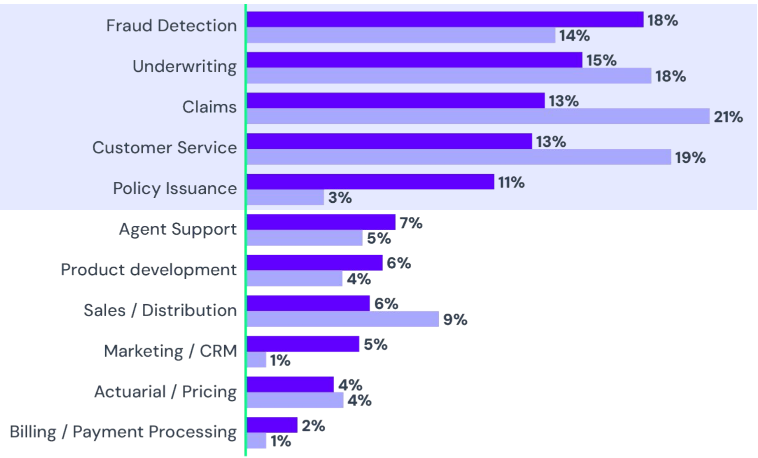
Data & AI in Insurance

Use Cases & Benefits

2023
2022

Breakdown of AI Insurance Use Case by Insurance Process

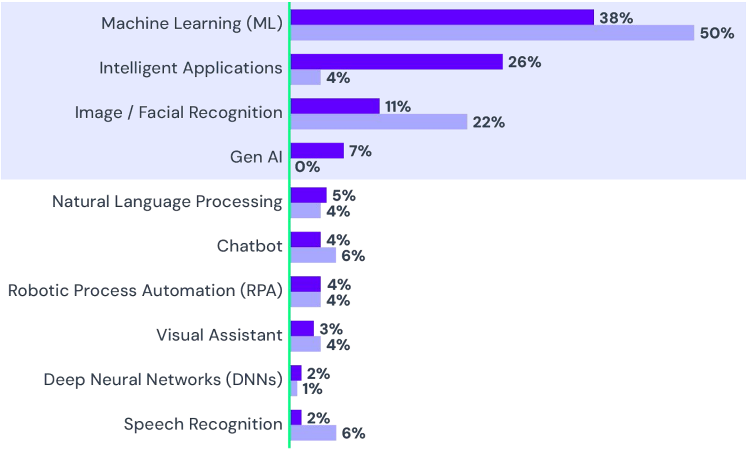
Percentage of cases



n=69 (2022), 76 (2023) use cases in the AI toolkit, analysis by type of insurance process, multiple response allowed
Source: Gartner Toolkit Artificial Intelligence Use Case for Insurance

Breakdown of AI Insurance Use Case by Technology

Percentage of cases



n=69 (2022), 76 (2023) use cases in the AI toolkit, analysis by type of insurance process, multiple response allowed
Source: Gartner Toolkit Artificial Intelligence Use Case for Insurance

Benefits

- Cost Efficiency
- Accuracy
- Scalability
- Customer Retention

Data & AI in Insurance

Challenges & Hindering factors: The central role of Data

*Often times, . . . , it's **limitations in the quality of data** platforms, master data management and data science that prevents from gaining the full value of AI*

Mike Helstrom Principal, Insurance Technology Strategy Consulting KPMG in the US

Obstacles

Data Quality & Governance

- **Poor data quality** hinders AI effectiveness & lead to **biased outcomes**
- **Legacy** Systems & **Siloed** Data

Compliance & Ethics

- **GDPR, AI Act**, & local protection standards
- **Ethical concerns:** Bias & Explainability

Cultural Resistance and Skills Gap

- **Fear of replacement**
- **Shortage** of AI professionals

The Role of Data

→ AI thrives on **high-quality, well-structured data**

→ Data integration across **silos** is critical for creating **comprehensive AI models**,

→ **Robust data governance supports AI Explainability**

5.

Synthetic Data Applications

Market & Applications

Enabling data leverage in Healthcare & Finance Industries

Finance

Healthcare

Data sharing

AI model development

Product demos

Scenario simulations

Business Analytics

Data Sharing (Intra/Inter Company)

SYNTHETIC DATA ADDED VALUE:

- ✔ Enables the sharing of authentic customer data among stakeholders, eliminating costly data anonymization processes and reduces administrative burden;
- ✔ Enhances processes within the finance industry, preserving data privacy and regulation compliance (such as GDPR, HIPAA and CCPA);
- ✔ Leverage data, hindering innovation and collaboration within the ecosystem thanks to Unified data storage and management;
- ✔ More robust and accurate predictive model, thanks to enhanced data quality through data augmentation.

AI Model Development

SYNTHETIC DATA ADDED VALUE:

- ✔ Serves as a valuable resource for training AI models, offering benefits such as:
  - ❑ Scalability;
  - ❑ Privacy preservation;
  - ❑ Augmentation of limited datasets.
- ✔ Suitables solution for enhancing model performance and generalization.

Enhancing add-on business applications: Scenario Simulation, Product Demo and Business Analytic

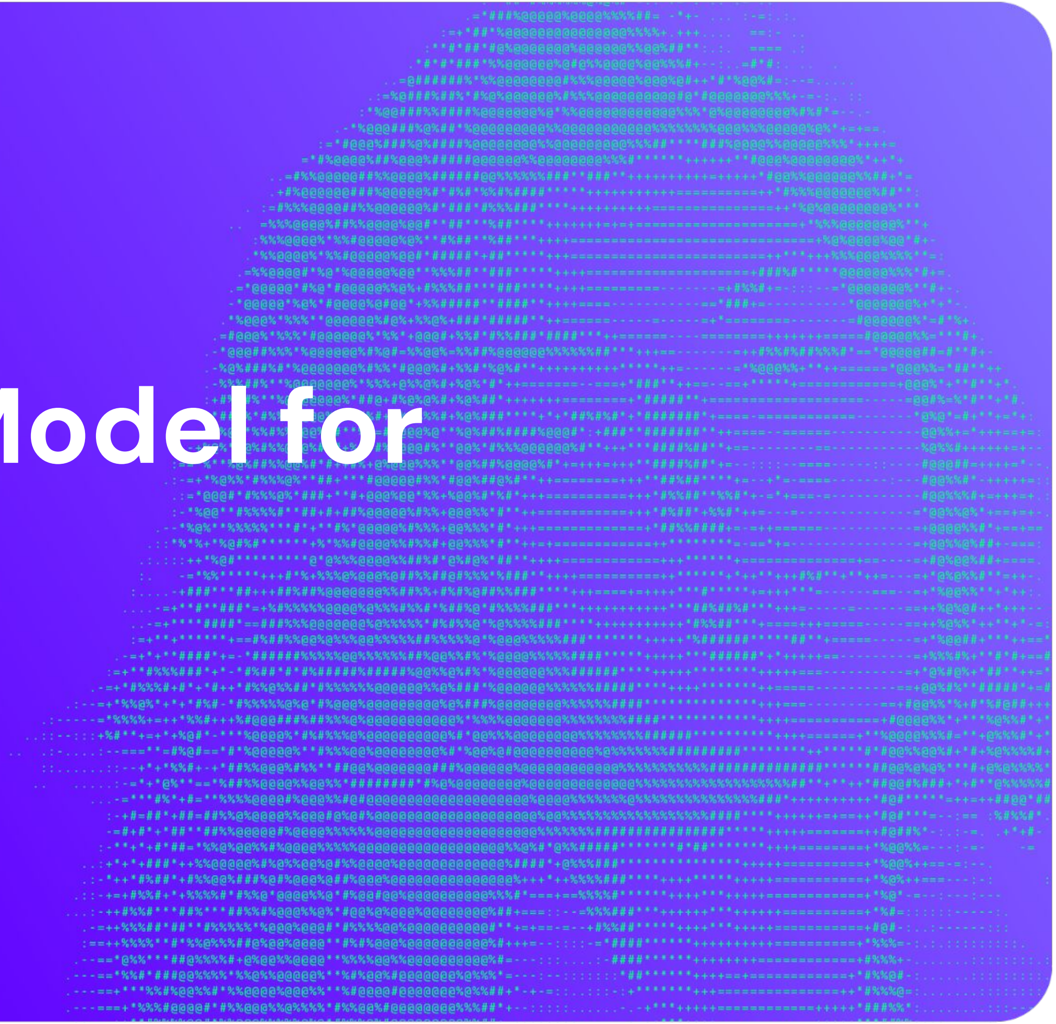
SYNTHETIC DATA ADDED VALUE:

- Enables **product seamless demonstrations**, accelerating development cycles and **allowing tests without exposing and relying on sensitive or limited data** sources.

- Enables gaining insights **equivalent to real data** and empowers **informed decision-making** and **integration into** business intelligence **tools**;

6.

A Business Model for AI and Data in Europe



Business Model

aindo

Partner ecosystem
- Independent software platform
- System integrator
- Consultancy company

B2B
- Healthcare & Life Science Companies
- Banks and Financial Institutions
- Telecommunication Companies
- Governmental institutions

Buyer personas
- CEO
- CISO
- CIO
- CTO
- Innovation manager

User personas
- Data scientist
- Data engineer
- Etl specialist
- AI/ML developper
- Data protection specialist

Subscription based

Modular by design
- Double dimension scaling
- On premises and Hybrid deployment

Pricing structure

Set up fee
Training level
Application complexity
Data types

Annual subscription
Deployment architecture
Data structured size
Model training frequency
Support level
Data types

7.

Interactive Discussion and Use Cases

AI model compatibility for IVASS Regulatory Sandbox

SANDBOX BANCA D'ITALIA EUROSISTEMA

ISTITUTO PER LA VIGILANZA SULLE ASSICURAZIONI IVASS

BE Rebel

Concept

POC

- Prove **synthetic data used for ML training** ensures adequate privacy and utility
- Verify **synthetic data's effectiveness in training two ML models:**
  - *Forecast model to predict claims frequency*
  - *Forecast model to predict Conversion/Retention Rate*

Project

STEP 1	STEP 2
MODEL COMPATIBILITY & DATA UTILITY	PREDICTIVE MODEL PERFORMANCE ENHANCEMENT

PLATFORM DEPLOYMENT

On-premises Aindo Web deployment

RECORD MANAGEMENT

- Flat Table, 20 columns / 35k rows
- Flat table, 20 columns / 180k rows

Outcomes step I

KPI ACHIEVED

- Similarity score > 90% ✓
- Privacy score > 95% ✓
- Model Compatibility (Data Utility) ✓

ADDED VALUE

- Data utility/privacy preservation
- Enhancement of AI/ML model training

VYV Group: AI model compatibility success story

GROUPE vyv

Concept

POC

Measure:

- Synthetic data **utility loss vs. real data** (from a dental consumption study already carried out)

Assess:

- Whether synthetic data provides **comparable results without personal data constraints**

Project

STEP 1	STEP 2	STEP 3
SYNTHETIC DATA GENERATION	SYNTHETIC DENTAL STUDY CREATION	SYNTHETIC vs. REAL DATA STUDY COMPARISON

PLATFORM DEPLOYMENT

SDK license integrated within the data pipeline

RECORD MANAGEMENT

Two Relational Tables:

- 5 columns / 20M rows
- 2 columns / 1M rows

Outcomes

KPI ACHIEVED

- Model Compatibility (Data Utility) ✓
- Similarity score > 90% ✓
- Privacy score > 95% ✓
- Up to < 75% for Time to Data ✓

ADDED VALUE

- Regulatory Compliance
- Operational time savings

Enabling the creation of the Data Catalog in PA-Healthcare for secure data-sharing

insiel

Business Problem

- **Develop and manage Regional Health Data Catalogs:**
- Need of strict adherence to privacy regulations, data security, governance, and ethical **standards**
- **Limitations of Legacy Anonymization Technologies:** legacy tool are impractical because they limit the utility of the data

Solution

- **Insiel developed and managed a synthetic data catalog for internal and external sharing using their technology**
- **Started with library SDK transitioning into platform for intuitive usability**

Value

- **Created market for inaccessible healthcare data to drive internal and external collaboration**
- **Reduce time-to-data for research purposes for common health benefits**
- **Avoid potential data leaks and non-compliance with privacy regulations**

Real World Evidence (RWE) Development

for an undisclosed leading Pharmaceutical company

Business Problem

- **Obtaining real-world evidence (RWE) is time-consuming:**
  - ❑ Obtaining real-world data (RWD) is an intensive process involving privacy compliance protocols, extensive permissions, and data curation.
- **RWE quality varies considerably:**
  - ❑ RWD is often siloed, hindering comprehensive insights.
  - ❑ RWD may also be static, not reflecting current unobserved variables, or lacking crucial information.
  - ❑ Traditional de-identification methods rely on information destruction (e.g. perturbation).

Solution

- **Unlock rapid, secure data access:**
  - ❑ Synthetic data is directly available, compliant, and curated. This allows for immediate RWE development opportunities.
- **Access comprehensive, current information:**
  - ❑ Integrate siloed data into one single, comprehensive synthetic data lake.
  - ❑ Create synthetic data based on the most recent RWD, accurately capturing the current research environment.
  - ❑ Avoid information destruction through traditional de-identification.

Value

- **Accelerated research**: Facilitate faster, more comprehensive RWE studies, accelerating research timelines.
- **Privacy-first approach**: Maintain compliance while enabling broad access to critical healthcare data.
- **Improved data quality**: Generate high-quality evidence based on comprehensive, up-to-date, and accurate data.

Future Roadmap

Next Steps: Expanding Synthetic Data's Role

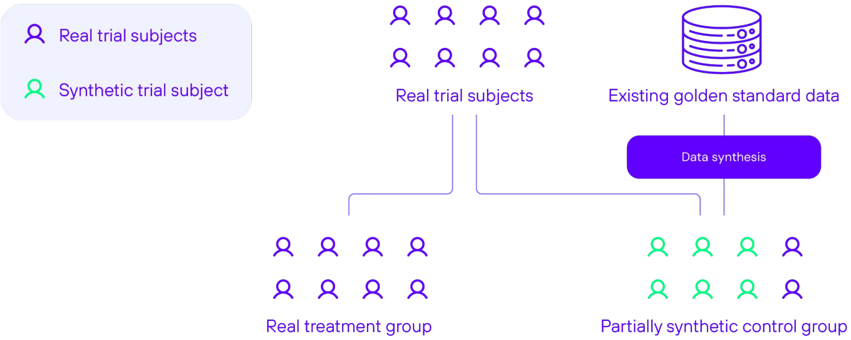
Clinical trials with fewer control subjects

Challenge

Obtaining clinical trial subjects is costly and time consuming, leading to long time-to-market for pivotal treatments.

Solution

Partially synthetic control groups can reduce the number of real trial subjects needed. This streamlines the trials, leading to efficient access to treatments.



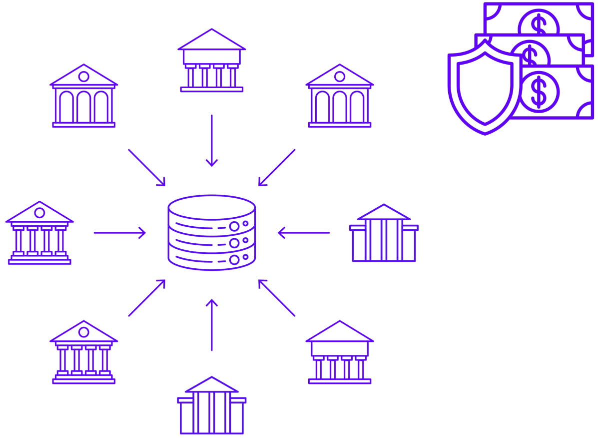
Saving & Investment Union strategy enabler

Challenge

SIU initiative is hindered by restrictions on data sharing, limiting transparency and market integration

Solution

Synthetic data enables secure, privacy-compliant data exchange, fostering a unified capital market strategy

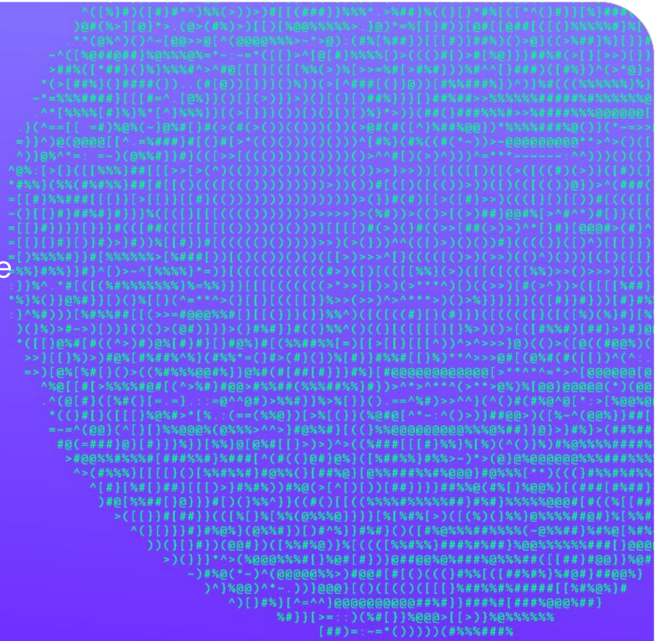


Future Roadmap

Next steps for strengthening privacy measures and industry impact.

Technology under evaluation by **CNIL**	CNIL.
Initiating engagement with the **aepd**	aepd agencia española protección datos
Active participation in **key data protection events**, including **Privacy Symposium and CPDP**	PRIVACY SYMPOSIUM
Achieving Financial KPIs & securing the next Funding Round	
Continuous Accademia collaboration	
Working with EMA for Scientific Advice to establish synthetic data as a standard for Synthetic Control Arms	EMA EUROPEAN MEDICINES AGENCY

Final Remarks

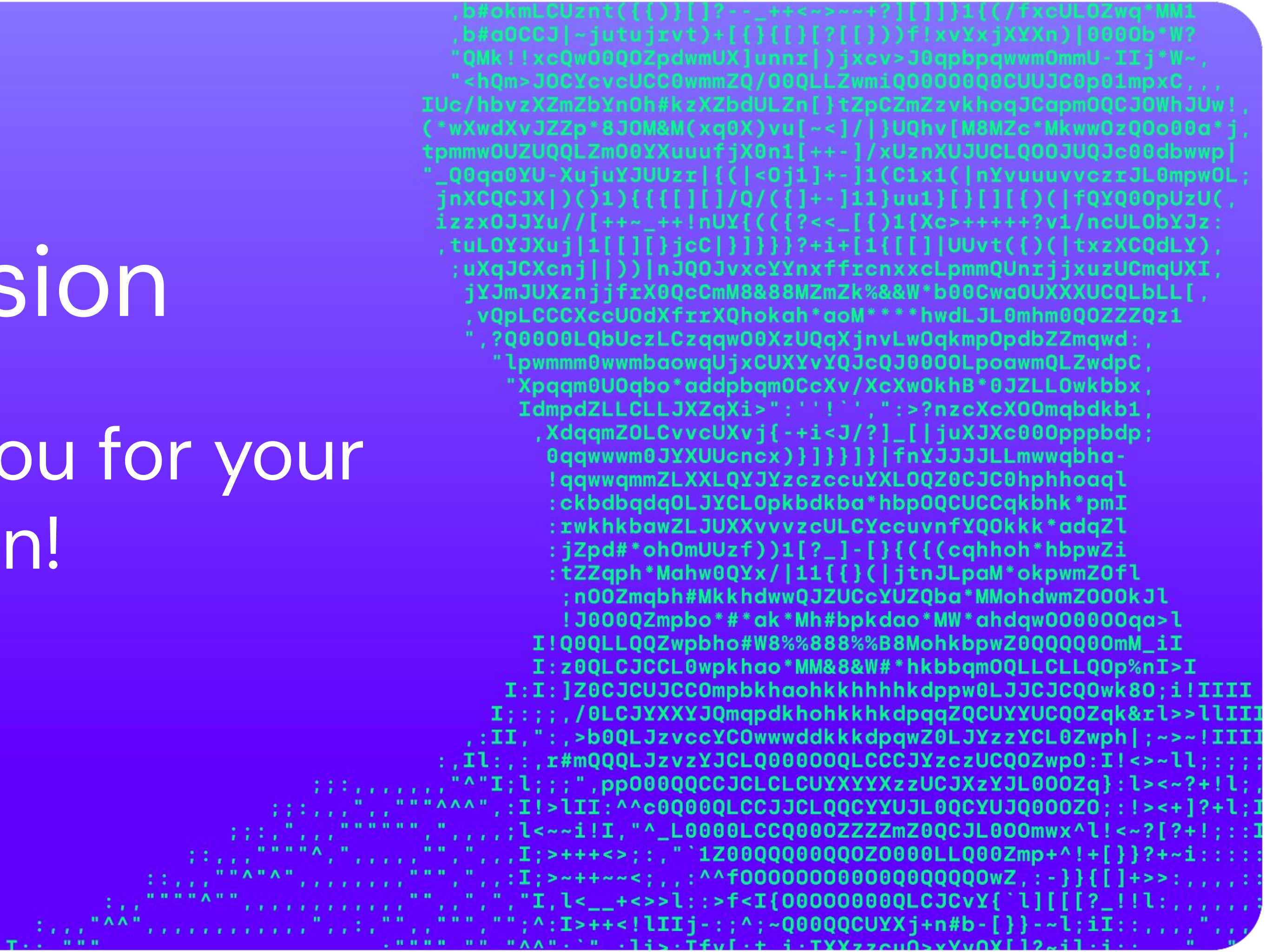


- AI development is key to Europe's competitiveness but must prioritize **safety and data privacy**, especially in **regulated sectors like healthcare and finance**.
- **Synthetic data is the most advanced solution** for protecting consumer data while enabling responsible AI use.
- Privacy evaluation must **move beyond legacy standards**, adopting more effective assessment methods.
- **European certifications and standards** are essential for ensuring **trust, accountability, and safe AI adoption**.
- **Education and training** are crucial to raising awareness and driving **data-driven innovation**.

aindo

Q&A Session

Thank you for your attention!



Alessandro di Pietro MBA,
Head of Business Development
+393290169472
alessandro@aindo.com