



Health Data Space Event

4 APRIL 2022



Chapter 2a: Breakout Use cases



#6 Enabler trusted AI

Welcome & opening

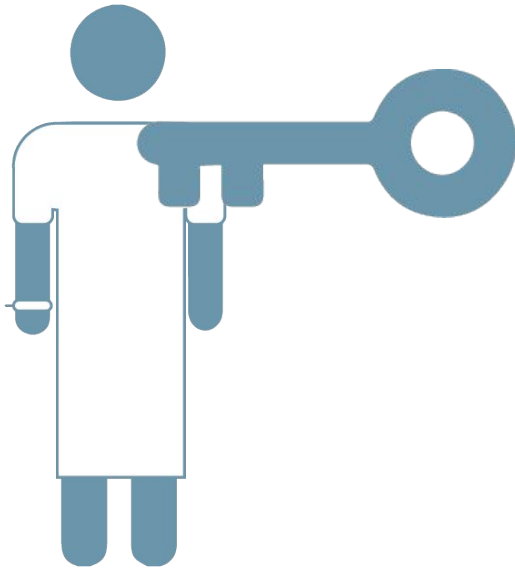


- **Prof. Ulf Nehrbass, CEO, Luxembourg Institute of Health**



AI in healthcare – the transborder Clinnova initiative

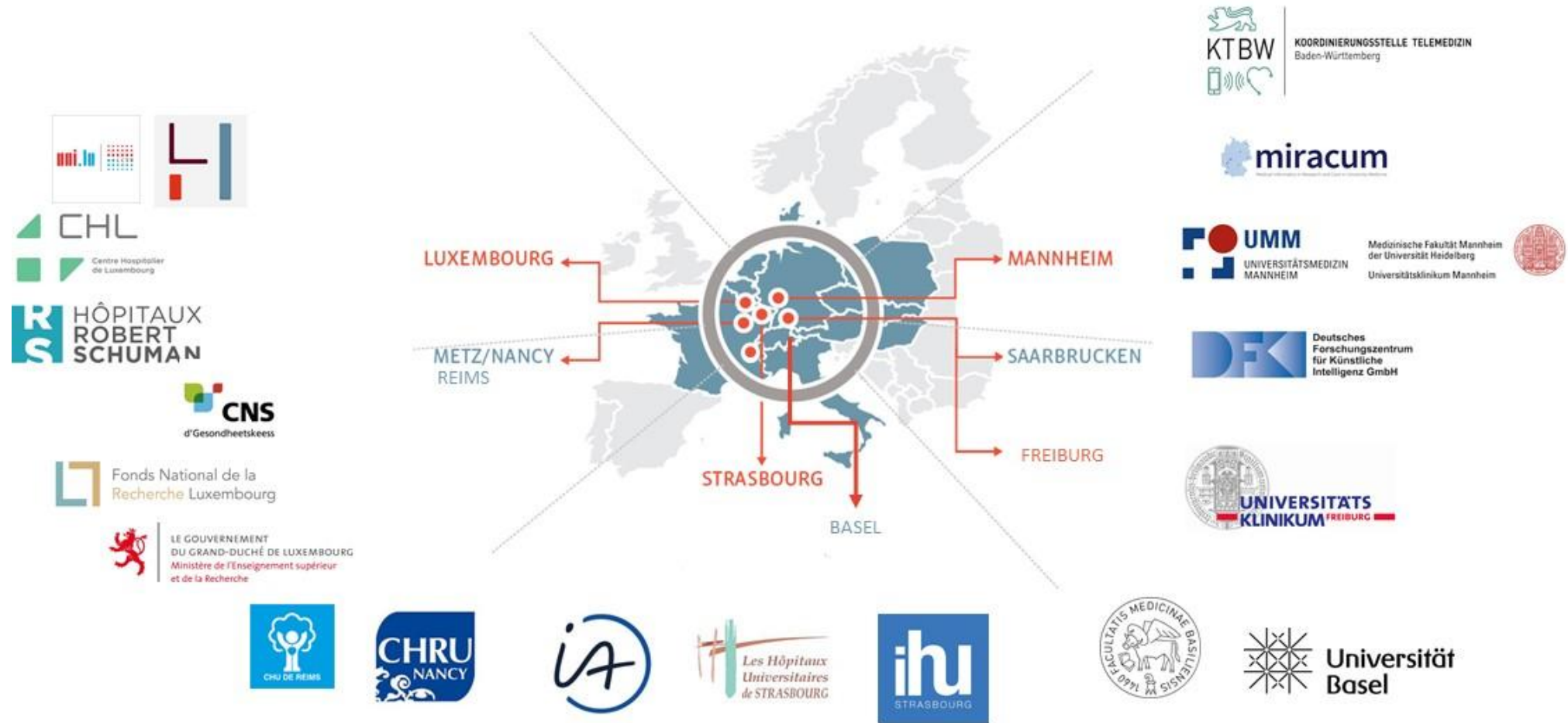
Cross-border digital health hub



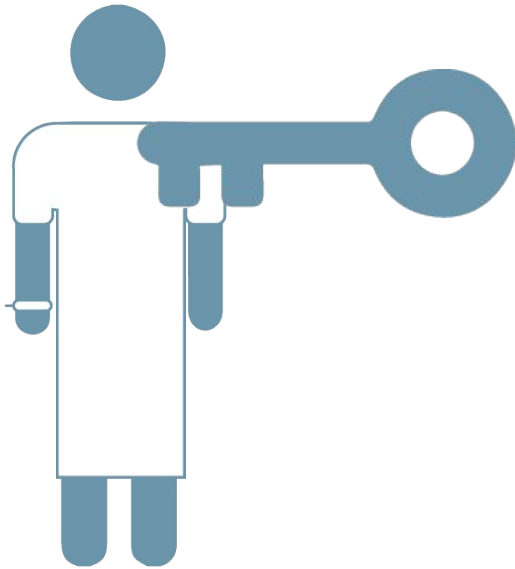
CLINNOVA'S STRATEGIC GOAL:

- Unlock the potential of data science and artificial intelligence (AI) in health care

Clinnova partners



Cross-border digital health hub

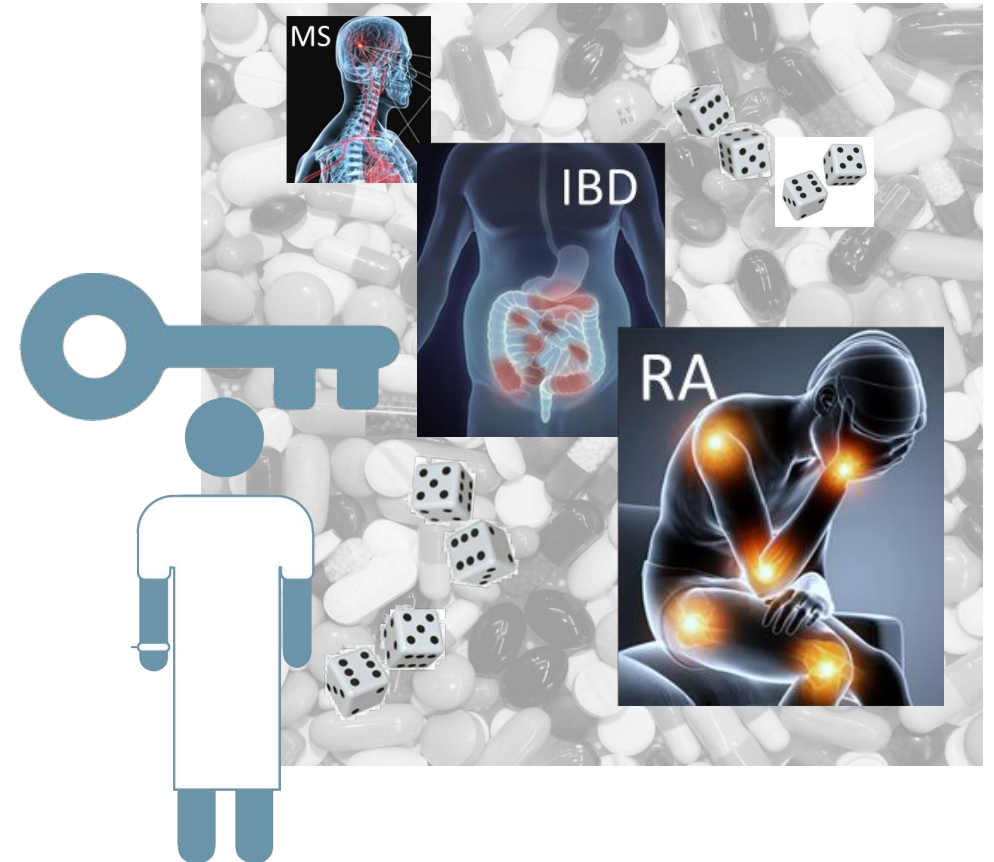


CLINNOVA CONSIDERS AI-INNOVATION IN HEALTHCARE AN OPERATIONAL AND ORGANIZATIONAL CHALLENGE

- The key innovation driver is not the AI or ML algorithm
- It rather lies in a data-enabling environment that produces standardized, quality-controlled data around relevant use cases

CLINNOVA'S PROSPECTIVE STUDIES START WITH 3 MEDICAL USE CASES

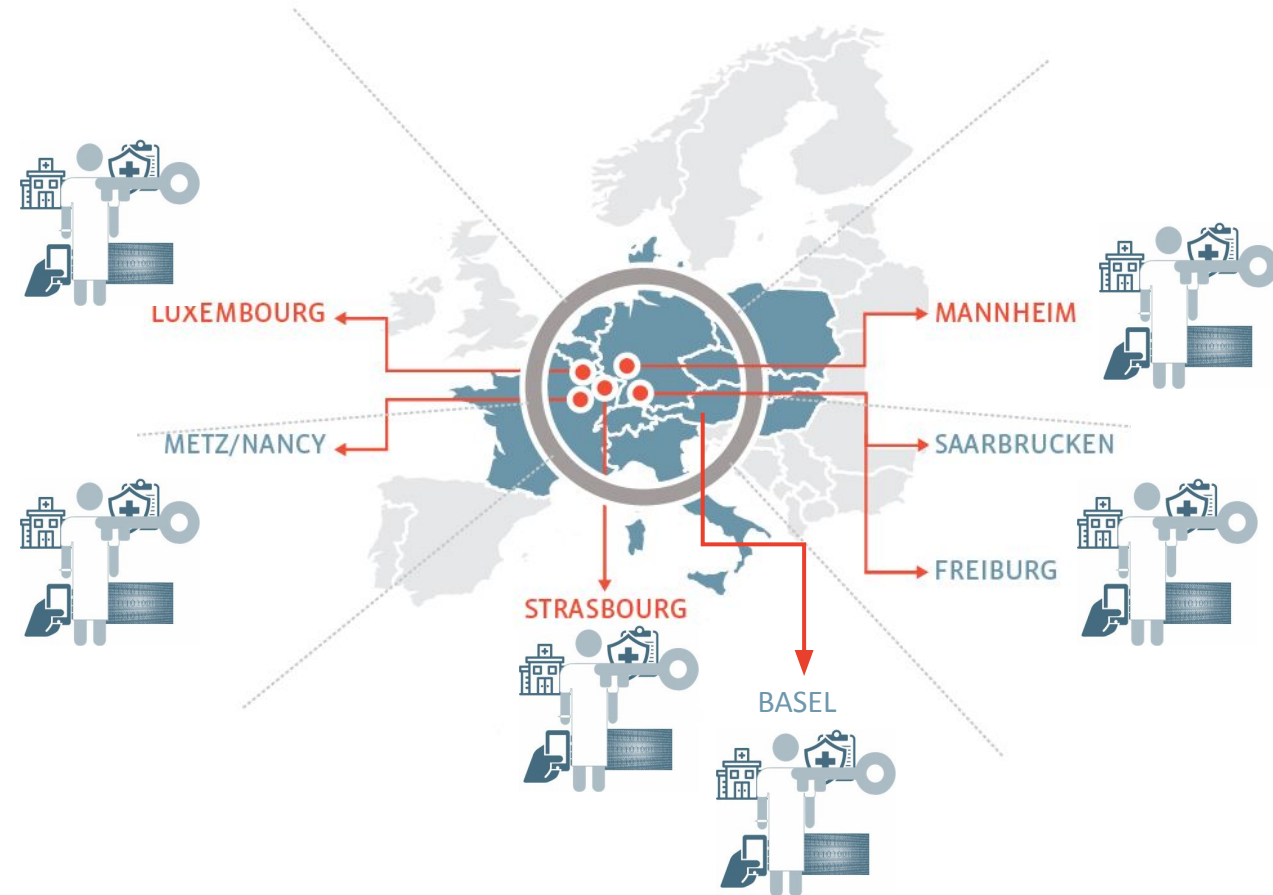
- RA, IBD and MS: 30Mio patients, €55Bn/ year market
- Standard of care drugs are abound, but it is not known which drug benefits what patient
- Clinnova's stratification approach will assign the right drug to the right patient



FACILITATING TRANS-BORDER DATA FLOWS AND COLLABORATION

- Linking up data integration platforms between clinical and research centers across borders
- Federated computing: data remain in local data integration centres and analyses are performed remotely
- Links health data context with neighbours to assure interoperability

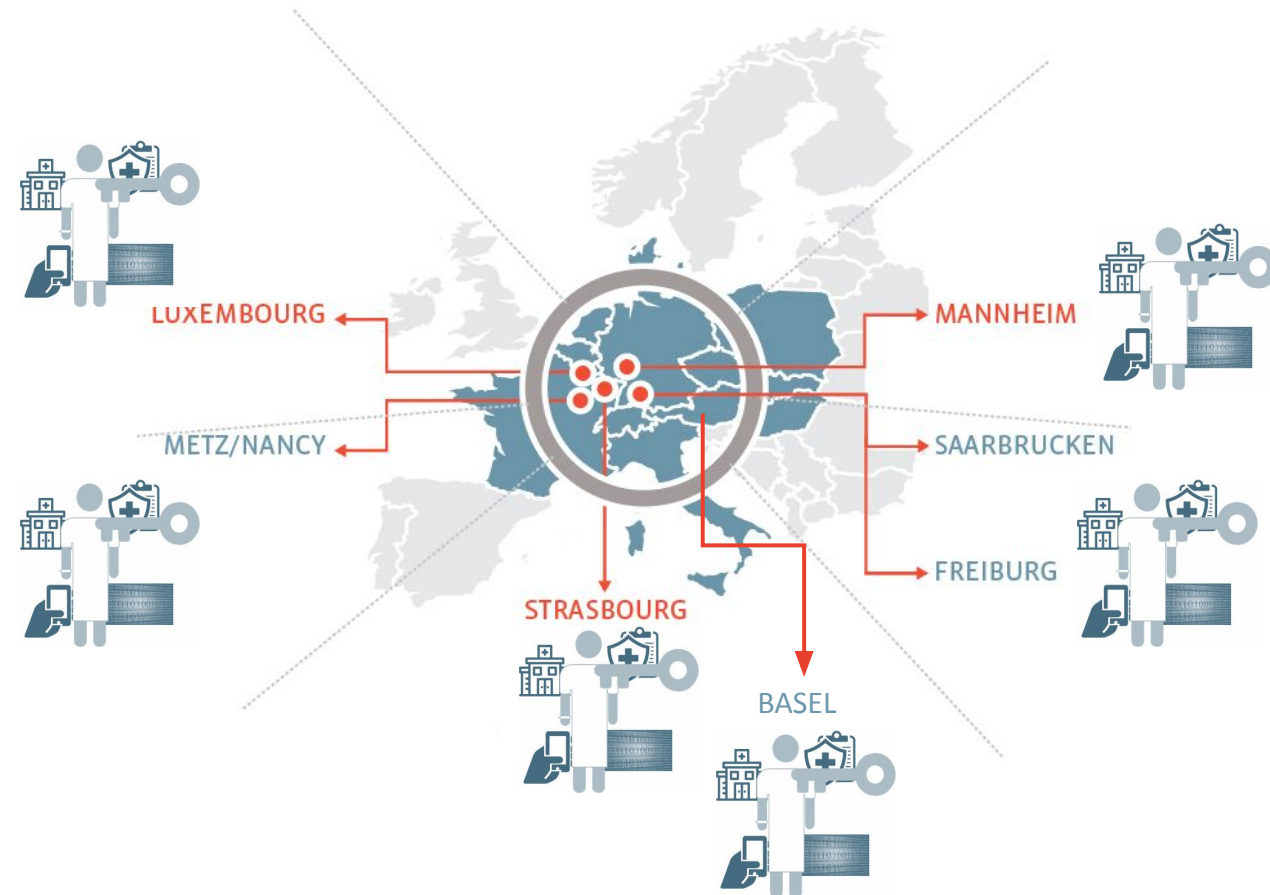
Clinnova trans-border dimension



CRITICAL QUESTIONS ON TRUSTED AI

- How can we ensure trustworthiness of AI algorithms?
- How can we achieve self-explainability of algorithms?
- How can we create a data ecosystem in which data do not need to be transferred / exchanged?

Clinnova trans-border dimension



Strategic discussion on Trusted AI



- **Prof. Philipp Slusallek**, Scientific Director, German Research Center for Artificial Intelligence (DFKI)
- **Dr Christian Schorr**, Senior Researcher, DFKI

Trusted AI



– How advanced anonymization can improve federated architectures

- **Tuomo Pentikäinen**, CEO, VEIL.AI



TRUSTED AI – HOW ADVANCED ANONYMIZATION CAN IMPROVE FEDERATED ARCHITECTURES

GDPR-free row level health data through *new next generation anonymization technology*

**Tuomo Pentikäinen, CEO
VEIL.AI**

April 4th, 2021

VEIL.AI

SHORT COMPANY INTRODUCTION:



ENABLING USE OF SENSITIVE DATA

Pseudonymized data

Anonymized data

Synthetic data



VEIL.AI HAS DEEP ROOTS IN SENSITIVE HEALTH DATA



- FIMM – Institute for Molecular Medicine Finland – with 230 employees representing tens of nationalities, is ***focusing on human genomics and precision medicine***

(Under the umbrella of the Helsinki Institute of Life Science at the University of Helsinki)



Key Data Scientists

- Our key technology experts have an ***experience of 20+ years*** related to demanding sensitive data management projects, such as:



- VEIL.AI a spin-out from FIMM, a health technology company located at Helsinki Meilahti hospital campus
- Founded in 2019
- Main customer groups: Global Pharma companies, University Hospitals, Health Regulatory bodies, MedTech companies

OUR SOLUTION:

VEIL.AI HAS DEVELOPED A GAME-CHANGING AI-BASED TECHNOLOGY THAT UNLEASHES THE POWER OF DATA

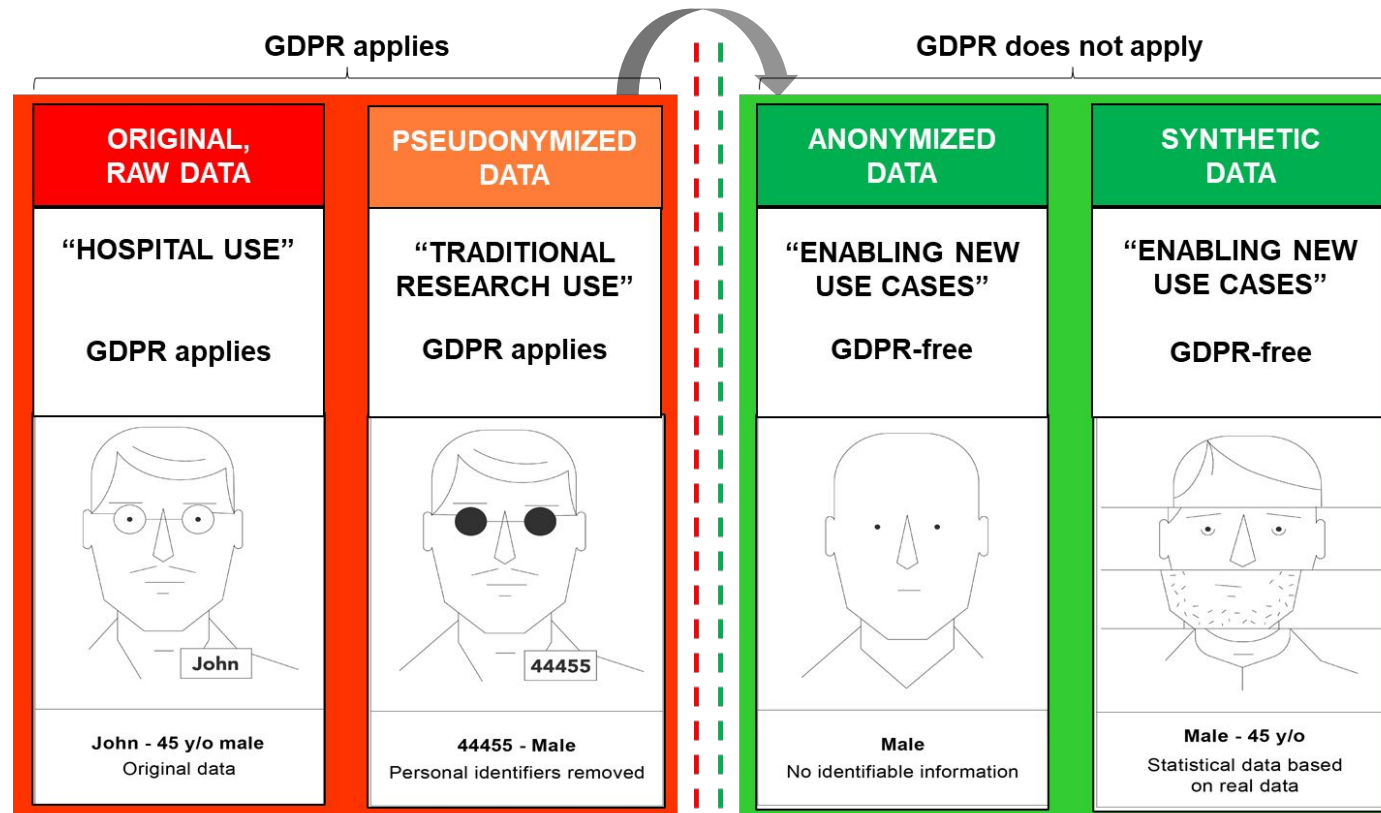


software

- The scalable *VEIL.AI Anonymization Engine* creates *extremely high-quality row level anonymized and synthetic data*, providing great new opportunities for
 - Hospitals
 - Health data hubs / Regulatory bodies
 - Pharma and Diagnostics companies
 - All organizations utilizing sensitive data

WHY TO USE ADVANCED ANONYMIZED DATA?

- One of the big advantages of anonymized data is that according to GDPR it is not considered as personal data (= it is “**GDPR-free**”). Therefore access and utility of health data improves significantly especially in transborder data collaborations.
- By utilizing also advanced anonymized data the organizations can improve the quality and quantity of the data they need (eg. clinical data).



- We can see a shift in demand from pseudonymized data to advanced anonymized data
 - Due to “GDPR-freedom”

GDPR Recital 26

...

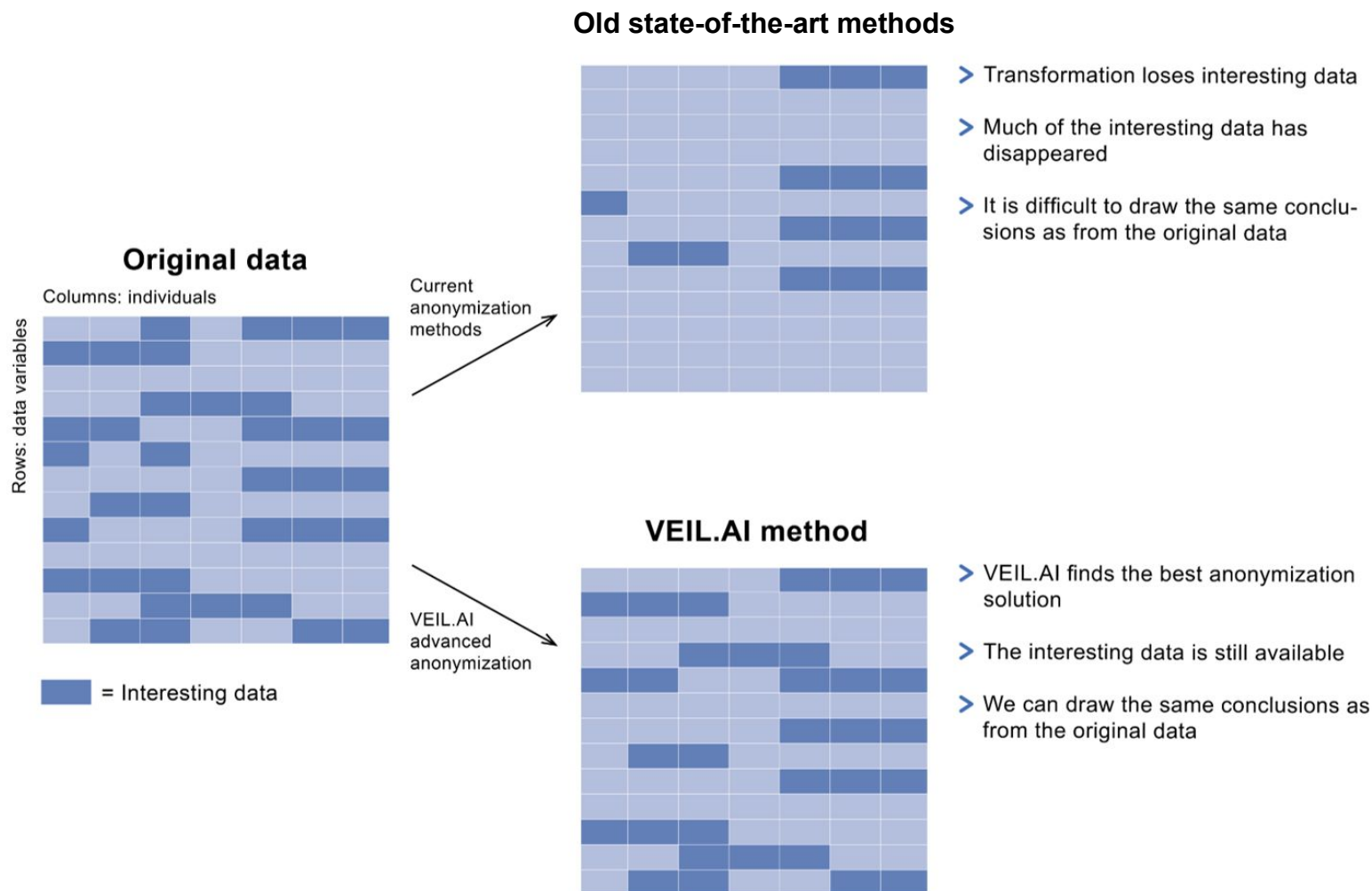
... “The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.

This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.”

COMPARISON

NEW ANONYMIZATION TECHNOLOGY vs. OLD METHODS

VEIL.AI's unique technology enables advanced anonymized datasets that can be used to draw the same conclusions as from the original data



AI based, next
generation
anonymization
technology

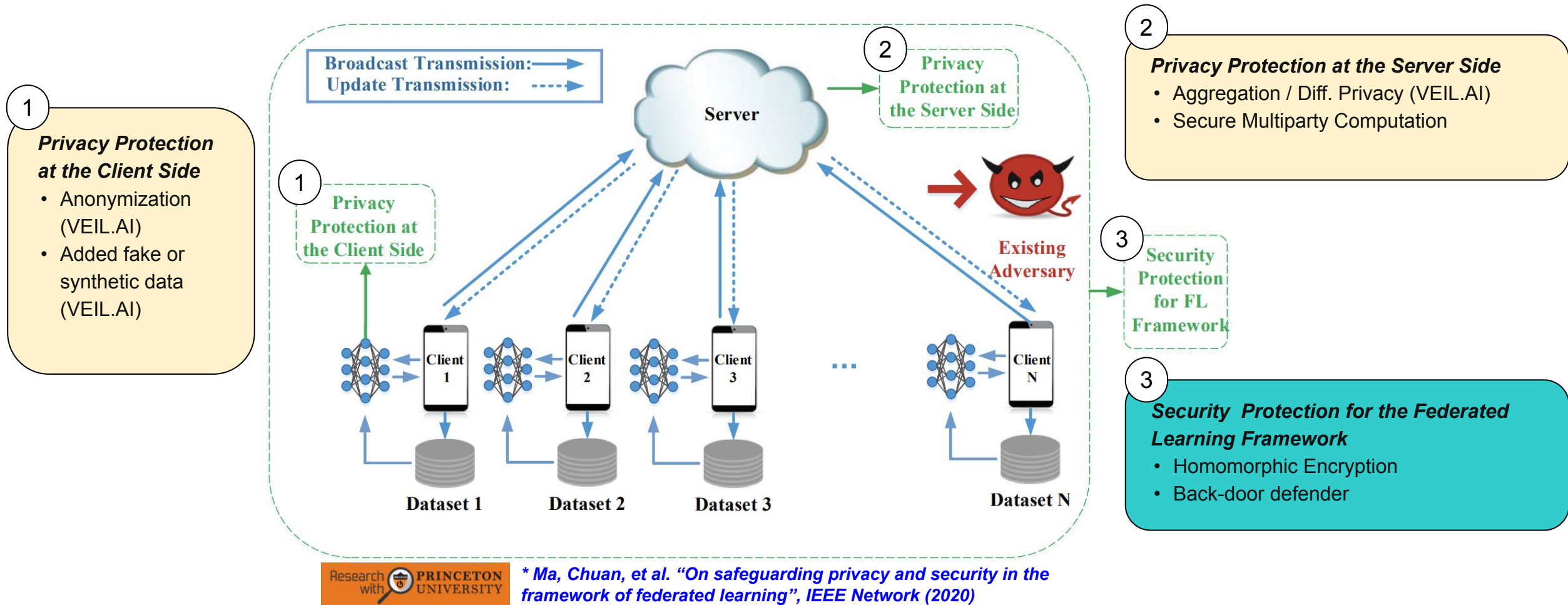
FEDERATED LEARNING ARCHITECTURE CAN BE A GOOD APPROACH

**TRANSBORDER DATA COLLABORATION INCLUDE MAJOR
DATA PRIVACY AND DATA SECURITY RISKS.**

**SOME SUGGEST THAT FEDERATED LEARNING SOLVES
ALL THESE PROBLEMS.**

UNFORTUNATELY THIS IS NOT TRUE, UNLESS...

FEDERATED LEARNING ARCHITECTURES REQUIRE ADDITIONAL PRIVACY AND INFORMATION SECURITY PROTECTIONS*



- Federation alone is not a solution to privacy problems

AFTER ANONYMIZATION – “IS THE DATA GONE NOW?”



ADVANCED ANONYMIZED ROW LEVEL DATA LOOKS THE SAME AND PERFORMS ALIKE THE ORIGINAL (RAW) DATA

Univariate analysis (Target: TenYearCHD)

Methods	Mean accuracy		Mean prediction loss	
	Orig	Anon	Orig	Anon
Logistic Regression*	0.846	0.842	0.154	0.158
Bayesian Ridge Regression	0.838	0.840	0.162	0.160
Gaussian NB	0.820	0.813	0.180	0.187

Multivariate analysis (Target: currentSmoker+TenYearCHD)

Methods	Mean accuracy		Mean prediction loss	
	Orig	Anon	Orig	Anon
Random Forest Classifier	0.836	0.838	0.071	0.073
Decision Tree Classifier**	0.732	0.745	0.150	0.147
K-Neighbors Classifier**	0.725	0.715	0.133	0.139

AutoML:

Train on Anonymized → Test on Real data
results using holdout data. (Target: TenYearCHD)

Mean accuracy		ROC_AUC (area under the curve)		Log prediction loss	
Orig	Anon	Orig	Anon	Orig	Anon
0.831	0.836	0.720	0.718	0.403	0.401

- Using automatic optimization of the anonymization process
- Summary: VEIL.AI row level anonymized data looks the same and performs alike the original (raw) data

* published model for Framingham dataset on kaggle.com

** modified model for multivariate prediction

DIFFERENTIAL PRIVACY STANDARD vs. VEIL.AI TECHNOLOGY

Multivariate analysis (Target: currentSmoker+TenYearCHD)

Methods	Mean accuracy		
	<i>Orig</i>	<i>Anon baseline</i>	<i>Anon optimized</i>
Random Forest Classifier	0.844	0.686	0.842
Decision Tree Classifier***	0.734	0.562	0.761
K-Neighbors Classifier***	0.725	0.444	0.684

Significant difference $p < 0.05$

Significant difference $p < 0.05$

The difference is not significant $p > 0.05$

* published model for Framingham dataset on kaggle.com

** a correlation coefficient value between -1 and +1 where +1 represents a perfect prediction, 0 an average random prediction and -1 an inverse prediction

*** modified model for multivariate prediction

ADVANCED ANONYMIZATION OFFERS SUPER QUALITY!

SUMMARY

1 Next generation AI based anonymization technology beats “old state-of-the-art” anonymization technologies

Feature	Unique VEIL.AI technology	"Old" state-of-the-art anonymization
Data security and privacy	✓	✓
Data with high quality and utility	✓	X
High performance and scalability	✓	X
Secure multiparty anonymization	✓	X
Multi-modal data protection	✓	X
Continuous and real-time anonymization	✓	X
AI-assisted privacy risk assessment	✓	X
Adaptive anonymization and data synthesis	✓	X

2 Advanced anonymized row data looks the same and performs alike the original (raw) data

- Advanced anonymized datasets can be used to draw the same conclusions as from the original data

Methods	Mean accuracy		
	Orig	Anon baseline	Anon optimized
Random Forest Classifier	0.844	0.686	0.842
Decision Tree Classifier***	0.734	0.562	0.761
K-Neighbors Classifier***	0.725	0.444	0.684

Significant difference $p < 0.05$

Significant difference $p < 0.05$

The difference is not significant $p > 0.05$

3 Federation is great, but it needs anonymization to solve the privacy protection problems



* Ma, Chuan, et al. “On safeguarding privacy and security in the framework of federated learning”, IEEE Network (2020)

FOR MORE INFO, PLEASE VISIT:

www.veil.ai

THANK YOU!

CONTACT INFO:



Tuomo Pentikäinen
CEO

+358 40 183 2881
tuomo.pentikainen@veil.ai

Expert.ai NLU technology

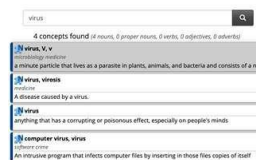


- **Gianluca Sensidoni, EMEA Account Manager, Expert.ai**
- **gsensidoni@expert.ai**
- **+39 3356595360**

What Expert.ai does

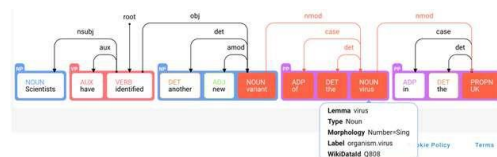
The ability to understand language and transform it into insights

Knowledge Graph



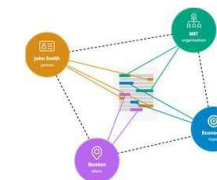
Expert.ai's knowledge graph is a representation of the real world where concepts are defined and connected to one other by semantic relationships.

Natural Language Understanding



Expert.ai distinguishes the correct **meaning of words** and expressions in context and automatically associates the attributes of more general terms that are conceptually linked to these words.

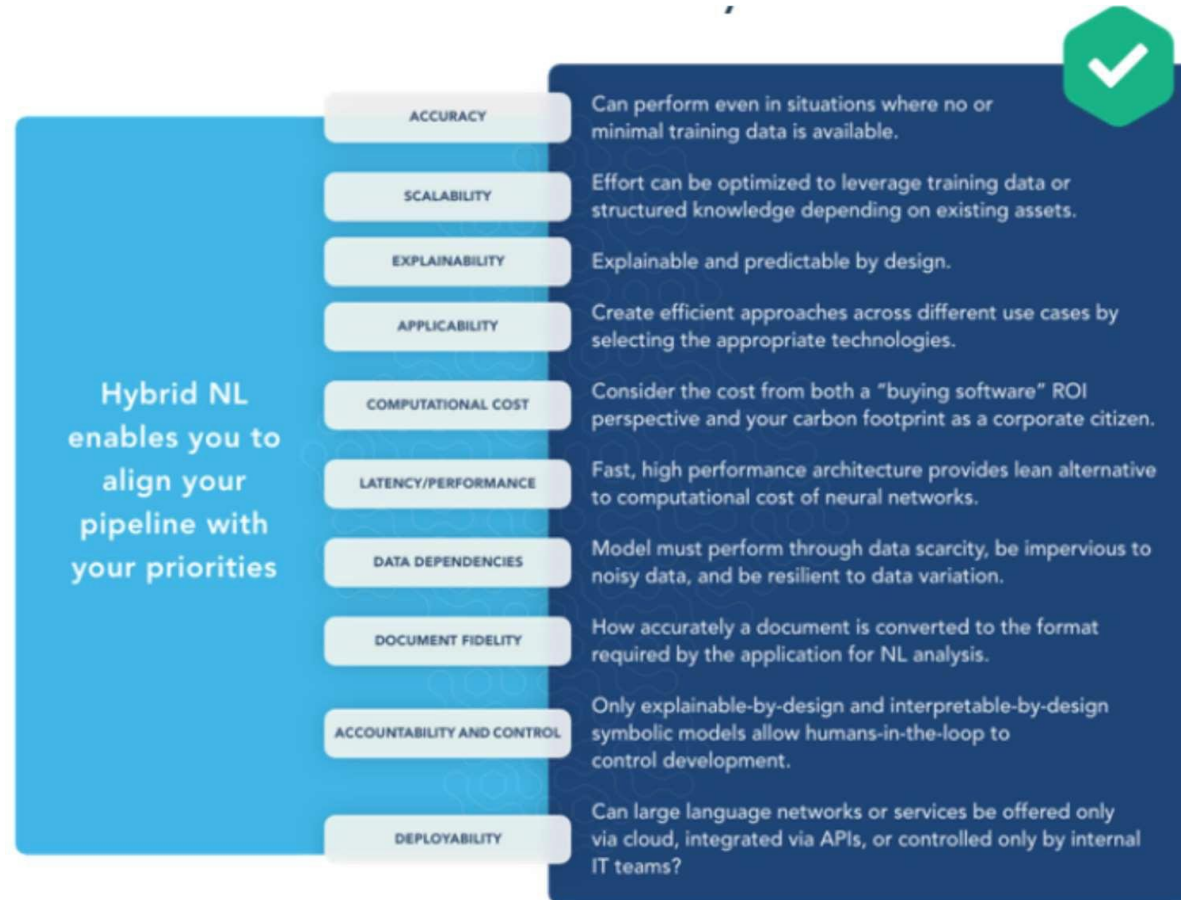
Reasoning Process



As the accurate comprehension of textual content always requires domain-specific knowledge. **Hybrid** Expert.ai technology can **mimic human** reasoning while analysing either on-line or off-line content.



The benefits of Hybrid NL



Expert.ai Technology at the end

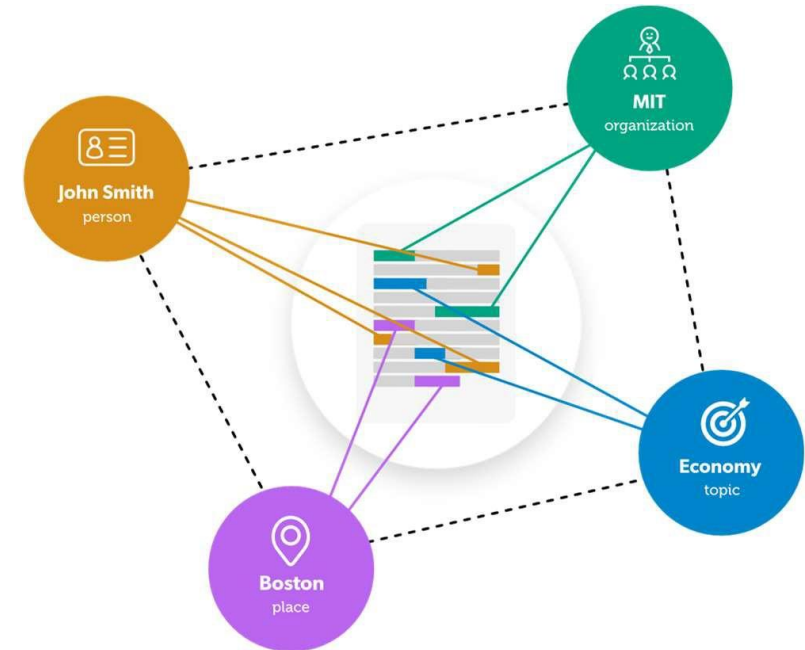
Expert.ai uses **patented AI algorithms** to mimic the human ability to read and understand any textual content

Expert.ai is based on a **pre-trained Knowledge Graph** with millions of concepts, entities, synonyms in **15 different languages**

Unlike traditional ML, it **does not require thousands or millions of examples** to be "trained from scratch"

Uses **advanced semantic rules** to classify and extract information with one of the higher accuracy levels on the market*

It's **not a black-box**: its Explainable AI concepts can always be described, documented and audited



How to use it

The target is to extract semantic data from unstructured contents in order to gain ***Information Superiority*** useful for ***Augmented DSS, Simulation & Prediction*** environment.

In the Healthcare domain f.e:

- Clinical Research
- Medical Report & Patient records analysis and cross correlation
- Disinformation vs Misinformation highlighting (Cognitive Warfare)
- Pandemic Surveillance
- Social Impact Discovering
- Digital Disease Detection
- Pharmacovigilance

INSERM in the Clinical Research



How do you detect and act on clinical insights when they are hidden across 100 million medical documents?



The Challenge

Overwhelming volumes of medical information (scientific literature, clinical studies, patents, side effect reports, and news) make it increasingly difficult for life sciences professionals to stay ahead of and develop relevant insights into their scientific and competitive landscape

How AI helps



Analyze 300+ information sources (Scientific and Healthcare)



Extract and Map connections between molecules, mechanisms of action, Pharmas & Biotechs, side effects, clinical trials, research, and networks of experts

The Impact

Days/month

**Time
Savings**

**Professionals can focus on
analysis & decisions instead of
data gathering**

About Expert.ai



€ 32m
Revenues

~270
employees

Public Company (EXSY)

- Largest global company specialized in Hybrid Natural Language Understanding technology
- Direct branches in US, CAN, UK, FR, DE, ES & IT

Dedicated to Customer Success

- Award winning Patented technology
- 30 years of R&D investments to create expert.ai technology
- Local technical & client teams committed to our customers' business & technology needs

Award Winning



2020 Best Overall
NLP Company



Magic Quadrant
for Insight Engines



AI-based Text
Analytics Platforms

Industry Recognition



FORRESTER®

AI-Based Text Analytics

People & Document-Focused Platforms

"...expert.ai's knowledge graph beats purely ML-based text analytics apps ...achieving faster time-to-value, more predictable results, and lower overall cost of ownership."

2020 Forrester Wave



Machine Learning and Artificial Intelligence Partner of the Year



Best Overall Natural Language Processing Company



KMWorld 100 Companies that matter in KM

Live Demos

Enjoy our **Free Live Demos** at:

- <https://try.expert.ai> - standard and available in 5 languages
- <https://www.intelligenceapi.com/demo/> - available in English language and dedicated to Stylometric and Emotional analysis (Behavioural algorithms for Human Factor attributes extractions)

Healthcare Live Demo with domain oriented international vocabularies such as ICD-9, ICD-10, UMLS, SNOMED, MESH, IUPHAR and ChemID*plus*.

(please ask me (gsensidoni@expert.ai) for a dedicated session)

Healthcare Live Demo (1)



UMLS TAXONOMY

- Event
- Entity

MESH TAXONOMY

SNOMEDCT TAXONOMY

ICD10CM TAXONOMY

ICD9CM TAXONOMY

- DISEASES AND INJURIES 55

DISEASE

- 13

DRUG

- Ramipril 4
- bradykinin 4
- Gabapentin 3
- Simvastatin 3

Life Sciences 1 (EN)

Model: Academic Articles

Text | XML | Recent History

and HTN on altace for 8 years
awoke from sleep around 2:30 am this morning of a
sore throat and swelling of tongue.
She came immediately to the ED b/c she was having
difficulty swallowing and some
trouble breathing due to obstruction caused by the

10171 characters

Result | Summary | Aboutness | Records | Relations | Coding

Disambiguation | Open JSON-LD

Body

Chief Complaint:
"swelling of tongue and difficulty breathing and swallowing"

History of Present Illness:
77 y o woman in NAD with a h/o CAD, DM2, asthma and HTN on altace for 8 years
awoke from sleep around 2:30 am this morning of a sore throat and swelling of tongue.

Any Concept

Abscess
DISEASE
DBpediaId : dbpedia.org/page/Abscess; ...

Adrenal Cortex Hormones
DRUG
SOURCES : LNC;MSH;NCI;SNOMEDCT_U...

Adrenergic beta-Antagonists
DRUG
SOURCES : LNC;MSH;NCI;SNOMEDCT_U...

Albuterol
DRUG
enAliases : (±)-salbutamol;albuterol;BTA2...

Allergen (brand of diphenylpyraline)
DRUG
SOURCES : NCI; VOCABS : UMLS; Display...

Allergic angioedema

Disclaimer

©2022, Expert.ai S.p.A. - build: 07 March 2022 (24 days ago)

Healthcare Live Demo (2)



disorders involving the immune mechanism (D50-D89)	Theophylline (Uniphyll) 600 mg qhs – bronchodilator by increasing cAMP used for treating asthma	Blood Urea Nitrogen
Diseases of the circulatory system (I00-I99)	Diltiazem 300 mg qhs – Ca channel blocker used to control hypertension	DRUG SOURCES : LNC;MSH;NCI;SNOMEDCT_U...
Neoplasms (C00-D49)	Simvastatin (Zocor) 20 mg qhs- HMGCo Reductase inhibitor for hypercholesterolemia	Bronchodilator Agents
Diseases of the musculoskeletal system and connective tissue (M00-M99)	Ramipril (Altace) 10 mg BID – ACEI for hypertension and diabetes for renal protective effect	DRUG SOURCES : LNC;MSH;NCI;SNOMEDCT_U...
Diseases of the respiratory system (J00-J99)	Glipizide 5 mg BID (diabetes) – sulfonylurea for treatment of diabetes	Cellulitis
Diseases of the nervous system (G00-G99)	Omeprazole (Prilosec) 20 mg daily (reflux) – PPI for treatment of ulcers	DISEASE ICD10CM : L03.90; DBpediaId : dbpedia.o...
Diseases of the digestive system (K00-K95)	Gabapentin (Neurontin) 100 mg qhs – modulates release of neurotransmitters to treat diabetic neuropathy	Chest Pain
Endocrine, nutritional and metabolic diseases (E00-E89)	Metformin 500 mg qam – biguanide used to treat diabetes	SIGNORSYMPATOM ICD9CM-Code : 786.50; enAliases : Chest...
Diseases of the skin and subcutaneous tissue (L00-L99)	Aspirin 81 mg qam - prophylaxis for MI and TIA	Chlorine
	Servant 1puff bid -	DRUG enAliases : Cl;element 17; DBpediaId : db...
	Fluticasone (Flovent) 2 puff bid - corticosteroid to treat airways in asthma/copd	Cholinergic Antagonists
	xoprex 1.25mg and Ipratropium 2.5 ml nebulized qam - anticholinergic to treat airways in COPD	DRUG SOURCES : LNC;MSH;NCI;SNOMEDCT_U...
	Review of Systems:	
	Constitutional - NAD, has been generally feeling well the last couple of weeks	
	Eyes - no changes in vision, double vision, blurry vision, wears glasses	
	ENT - No congestion. changes in hearing. does not wear hearing aids	

Disclaimer ©2022, Expert.ai S.p.A. - build: 07 March 2022 (24 days ago)

Healthcare Live Demo (3)



UMLS TAXONOMY

MESH TAXONOMY

- ▶ Anatomy
- ▶ Phenomena and Processes
- ▶ Analytical, Diagnostic and
- ▶ Therapeutic Techniques and Equipment
- ▶ Therapeutics
- ▶ Diagnosis
 - ▶ Diagnostic Techniques and Procedures
 - ▶ Medical History Taking
 - ▶ Diagnostic
 - ▶ Techniques, Respiratory System
 - ▶ Respiratory Sounds

Life Sciences 1 (EN)

Model: Academic Articles

Text | XML

Recent History

ANALYZE

CLEAR TEXT

and HTN on altace for 8 years
awoke from sleep around 2:30 am this morning of a
sore throat and swelling
She came immediately
difficulty swallowing and
trouble breathing due to

10171 characters

Result | Summary

Disambiguation | Oper

Body

Chief Complaint:
"swelling of tongue and difficulty breathing and swallowing"

History of Present Illness:
77 y o woman in NAD with a h/o CAD, DM2, asthma and HTN on altace for 8 years
awoke from sleep around 2:30 am this morning of a sore throat and swelling of tongue.

difficulty breathing (1)

Dyspnea (Freq: 1) (symptom, symptom, UMLS Sign or Symptom) NOUN

dyspnea (Freq: 1), shortness of breath , air

Score: N/A

d respiration;
irtesy of

less

bea

dbpedia.org/page/Shortness_of_breath

DisplayValue: Dyspnea

UMLSTUI: T184

WikibaseId: Q188008

Disclaimer

©2022, Expert.ai S.p.A. - build: 07 March 2022 (24 days ago)

Healthcare Live Demo (4)



▶ Named Groups		"swelling of tongue and difficulty breathing and swallowing"	Allergen (brand of diphenylpyraline) ▼ DRUG SOURCES : NCI; VOCABS : UMLS; Display...
▶ Chemicals and Drugs		History of Present Illness:	Allergic angioedema ▼ DISEASE ICD10CM-Code : T78.3; SOURCES : ICD1...
▶ Disciplines and Occupations		77 y o woman in NAD with a h/o CAD, DM2, asthma and HTN on altace for 8 years	Allergy to sulfonamide ▼ DISEASE SOURCES : CHV;SNOMEDCT_US; VOCAB...
▶ Supplementary Concept Data		awoke from sleep around 2:30 am this morning of a sore throat and swelling of tongue.	Anaphylaxis ▼ DISEASE enAliases : anaphylactic reaction; DBpedi...
▶ SNOMEDCT TAXONOMY		She came immediately to the ED b/c she was having difficulty swallowing and some	Anemia, Iron-Deficiency ▼ DISEASE ICD9CM-Code : 280;280.8; enAliases : Iro...
▶ ICD10CM TAXONOMY		trouble breathing due to obstruction caused by the swelling. She has never had a similar	Angioedema ▼ DISEASE ICD10CM : T78.3; enAliases : Edema, Ang...
▶ ICD9CM TAXONOMY		reaction ever before and she did not have any associated SOB, chest pain, itching, or	
▼ DISEASE		nausea. She has not noticed any rashes, and has been afebrile. She says that she feels like	
Angioedema	8	it is swollen down in her esophagus as well. In the ED she was given 25mg benadryl IV,	
Asthma	6	125 mg solumedrol IV and pepcid 20 mg IV. This has helped the swelling some but her	
Pulmonary Disease, Chronic		throat still hurts and it hurts to swallow. Nothing else was able to relieve the pain and	
Obstructive		nothing make it worse though she has not tried to drink any fluids because of trouble	
rash	5	swallowing. She denies any recent travel, recent exposure to unusual plants or animals or	
DM2	3	other allergens. She has not started any new medications, has not used any new lotions or	
		perfumes and has not eaten any unusual foods. Patient has not taken any of her oral	
		medications today.	
		Surgical History:	
		s/p vaginal wall operation for prolapse 2006	
		s/p Cardiac stent in 1999	

Disclaimer ©2022, Expert.ai S.p.A. - build: 07 March 2022 (24 days ago)

Healthcare Live Demo (5)



Chemicals and Drugs

Disciplines and Occupations

Supplementary Concept Data

SNOMEDCT TAXONOMY

Clinical finding

Situation with explicit context

Family history with explicit context

Family history of clinical finding

No family history of

Finding with explicit context

Procedure

Qualifier value

Observable entity

☒ Group relations (17/20)

Biguanide

DRUG

MAY_TREAT

RELATION

Diabetes

DISEASE

Metformin 500 mg qam – biguanide used to treat diabetes

Metformin

DRUG

MAY_TREAT

RELATION

Diabetes

DISEASE

Metformin 500 mg qam – biguanide used to treat diabetes

Fluticasone

DRUG

MAY_TREAT

RELATION

Pulmonary Disease, Chronic Obstructive

DISEASE

Fluticasone (Flovent) 2 puff bid - corticosteroid to treat airways in asthma/copd

Aspirin

DRUG

MAY_TREAT

RELATION

Myocardial Infarction

DISEASE

Aspirin 81 mg qam - prophylaxis for MI and TIA

Aspirin

DRUG

MAY_TREAT

RELATION

Ischemic Attack, Transient

DISEASE

SOURCES : NCI; VOCABS : UMLS; Display...

Allergic angioedema

DISEASE

ICD10CM-Code : T78.3; SOURCES : ICD1...

Allergy to sulfonamide

DISEASE

SOURCES : CHV;SNOMEDCT_US; VOCAB...

Anaphylaxis

DISEASE

enAliases : anaphylactic reaction; DBpedi...

Anemia, Iron-Deficiency

DISEASE

ICD9CM-Code : 280;280.8; enAliases : Iro...

Angioedema

DISEASE

ICD10CM : T78.3; enAliases : Edema, Ang...

Angioedemas, Hereditary

DISEASE

Disclaimer

©2022, Expert.ai S.p.A. - build: 07 March 2022 (24 days ago)

Healthcare Live Demo (6)



Investigative Techniques

▶ Diseases

▶ Named Groups

▶ Chemicals and Drugs

▶ Disciplines and Occupations

▶ Supplementary Concept Data

▼ **SNOMEDCT TAXONOMY**

▶ Clinical finding

▶ Situation with explicit context

▶ Procedure

▶ Qualifier value

▼ Observable entity

▼ Function

General metabolic function

UMLS Taxonomy

T123 Biologically Active Substance

T091 Biomedical Occupation or Discipline

T029 Body Location or Region

T023 Body Part, Organ, or Organ Component

T030 Body Space or Junction

T022 Body System

T025 Cell

T060 Diagnostic Procedure

T047 Disease or Syndrome

T196 Element, Ion, or Isotope

T033 Finding

T028 Gene or Genome

T131 Hazardous or Poisonous Substance

T125 Hormone

T197 Inorganic Chemical

T059 Laboratory Procedure

T191 Neoplastic Process

T114 Nucleic Acid, Nucleoside, or Nucleotide

T109 Organic Chemical

T040 Organism Function

DRUG

enAliases : (±)-salbutamol;albuterol;BTA2...

Allergen (brand of diphenylpyraline) ▼

DRUG

SOURCES : NCI; VOCABS : UMLS; Display...

Allergic angioedema ▼

DISEASE

ICD10CM-Code : T78.3; SOURCES : ICD1...

Allergy to sulfonamide ▼

DISEASE

SOURCES : CHV;SNOMEDCT_US; VOCAB...

Anaphylaxis ▼

DISEASE

enAliases : anaphylactic reaction; DBpedi...

Anemia, Iron-Deficiency ▼

DISEASE

ICD9CM-Code : 280;280.8; enAliases : Iro...

Angioedema ▼

Disclaimer

©2022, Expert.ai S.p.A. - build: 07 March 2022 (24 days ago)

AI-Evidence – AIR_PTE



Prof. Dr. Thomas P. Zahn
DCC Risk Analytics & DSI Data Science Institute

thomas.zahn@risikoanalytik.de

+49-30-3199829-01

AIR_PTE

AI based Risk Prediction & Treatment Effect Estimation



International Collaboration Project funded by BMWI & IRAP

- Current Partners from Germany, Canada, Denmark and Romania – open to join now

Goal: Rapid Evidence Generator (REG)

Risk Adjusted Propensity Score (RAPS) using gradient boost & deep Learning

- on historic (spatio-temporal) treatment pattern (ICD , ATC , OPS ...)
- to identify outcome risk adjusted control for multiple interventions in real world claims data

Applicable to generate RCT comparable Evidence – fast, reliable and effective i.e. for

- Innovative treatment options (DOAC for VTE treatment study) & programs (Smart CasaPlus)
- Digital Health Applications (DiGA)



Faculty of
Medicine

m_a c a d a m i a n

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages



Government
of Canada

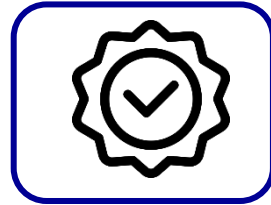
Gouvernement
du Canada



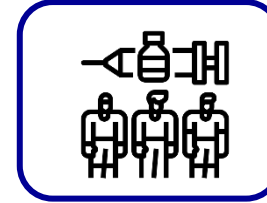
Problem



Health **innovation speed** is exponentially increasing i.e. in pharma, devices and digital health



Innovations in Health need **approval** by regulators **based on scientifically validated studies**

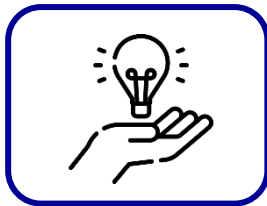


The Gold Standard - randomized control trials (RCT) – are **very time consuming and expensive**

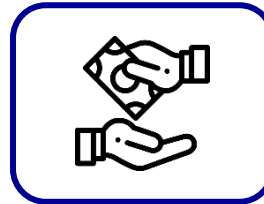


Real World Application of Health Innovations often **differ from RCT study populations**

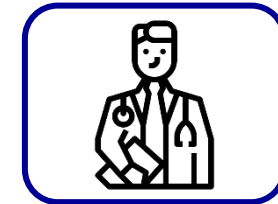
HENCE



Innovators need Real World Evidence to **continuously monitor effectiveness and efficiency** of their innovations



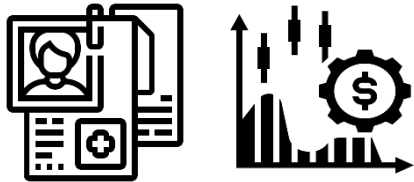
Public Healthcare Systems need Real World Evidence to **focus Innovations** to populations with highest benefits



Medical Practitioners need Real World Evidence to **support personalized treatment decisions**



AIR_PTE
AI-Evidence



Public Healthcare Systems
regularly produce **long term
health claims data**

BUT



contain highly sensitive personal information and **need
the highest level of data protection**



are obtained for the purpose of reimbursement and
not always available for research



are **highly structured and standardized** but may
lack specific information typical to RCT

Solution



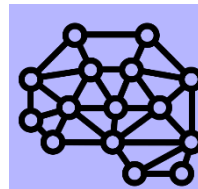
mirror the real-world effect of innovations
on the **entire population**



can be **used for the purpose of generating
Real World Evidence**



Real World Evidence from claims data is **fast,
inexpensive and continuously** obtainable



**smart AI
algorithms are
needed**

HENCE

to **ensure full data protection** when using
health claims for research

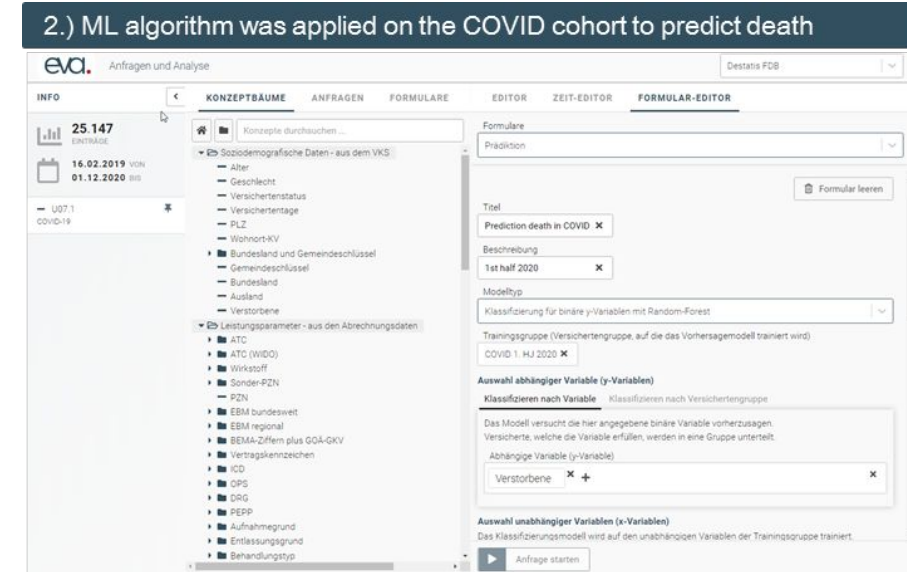
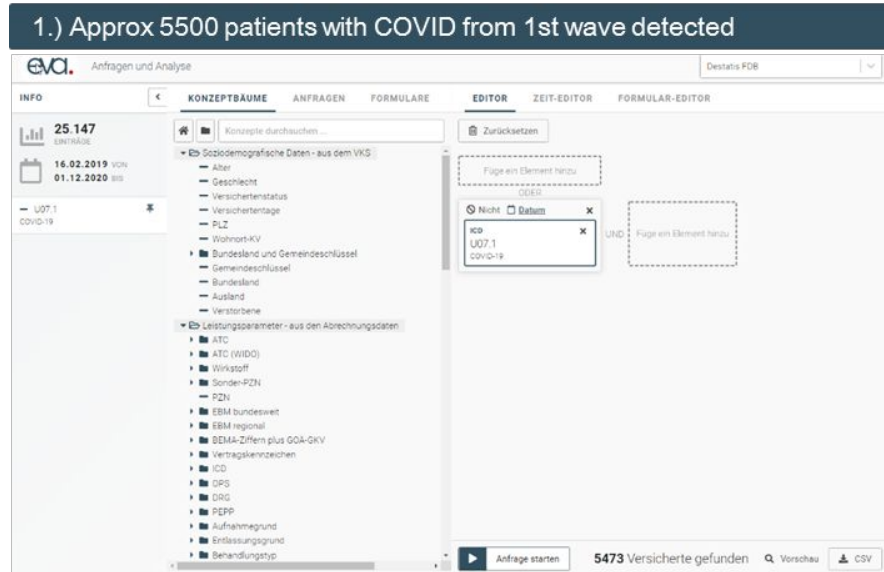
for the retrospective **identification of study
sub-populations and outcomes**

for the **identification of specific
intervention and valid control groups**

for generating **Real World Evidence for
innovative Treatment** options



we aim to adapt the REG Methods to the needs of COVID-19 Pandemic Risk predictions



WHO, Germany launch new global hub for pandemic and epidemic intelligence

Together, we will provide REG Results to Payors, Providers and Researchers via an international Rapid Evidence Repository

www.ai-evidence.de



Rapid Evidence Generators (REG)

The Rapid Evidence Generators (REG), developed in cooperation with payors, medical experts and researchers, is intended to obtain new evidence for the effectiveness of specific treatment patterns on defined outcomes, also for special patient groups. It takes advantage of the large number of individual cases and the wide range of long-term treatment courses in anonymized claims of public health insurances. Specific research questions can be submitted in the [Rapid Evidence Generator](#) section.



Rapid Evidence Repository (RER)

The Rapid Evidence Generator results will become available at the point of care by a Rapid Evidence Repository (RER) in compliance with all data protection regulations. There, physicians can enter the patient's characteristics and treatment goals after consent and rapidly obtain information on the potential effect of the different treatment options available, as well as links to the respective evidence studies.

Trusted AI for Data Interoperability



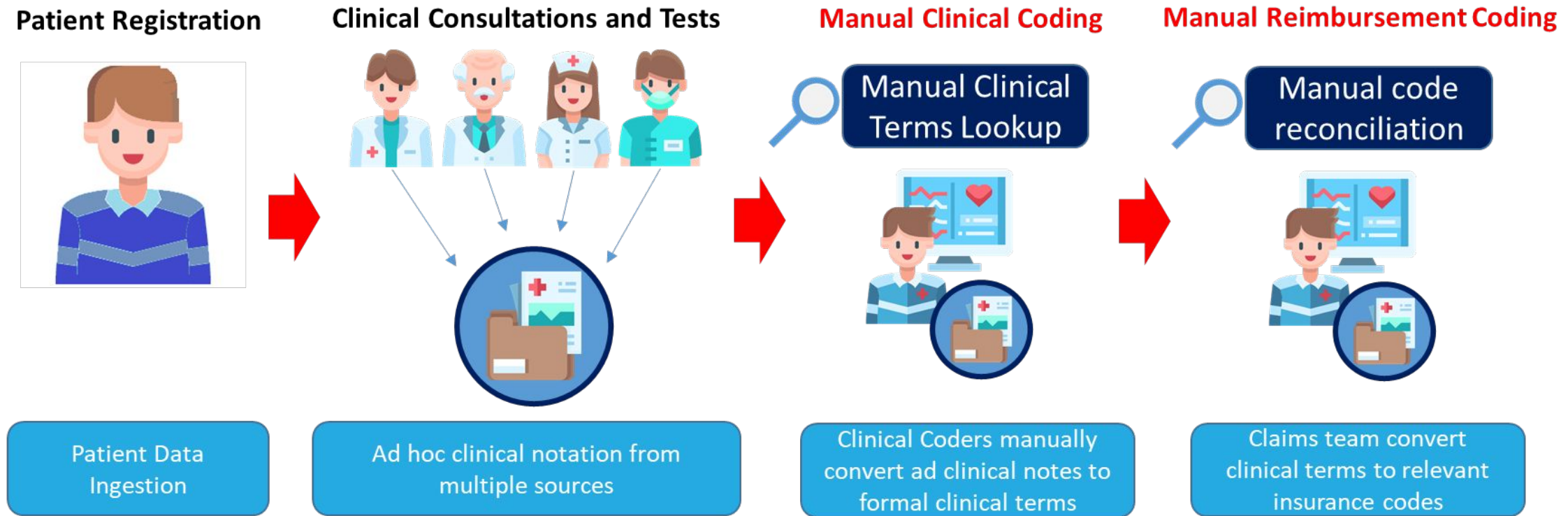
- **Dermot Doyle, CEO, Dynaccurate**

About Us...

- Dynaccurate is a spin out from Luxembourg Institute of Science and Technology (**LIST**)
- We provide **AI solutions** for health data interoperability
- We use both **Experts Systems, Machine Learning** as well as advanced search technologies
- Our mission is to provide **affordable, trustworthy AI** to the health sector to help make health data FAIR
- **The health sector actually requires AI for data management** much more than many other sectors
- Health data sector **underserved for AI solutions**

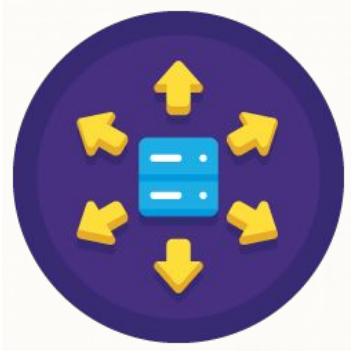


The reality of health data processing right now...



X Slow - Bottlenecked – Siloed – No Automation **X**

Why so tough??? – Financial Data vs. Health Data



In Finance

- Bank accounts and Bank identifiers rarely change **(Stable)**
- Data is largely numeric and can be worked on with formula, spreadsheets, basic mathematics **(Easier to process with applications)**
- Common on identifiers - IBANs, ISINs, CUSIPs, SEDOLs etc. - and SWIFT infrastructure **(Global regime)**
- Banks motivated to make it happen - More transaction volume, more fees, more profit – **(Common Incentives)**

In Health

- Health data identifiers like SNOMED, ICD and local terms etc. change a lot **(not stable)**
- Data is largely expert domain terminologies / ontologies and human medical knowledge **(very few applications – requires Semantic Web AI)**
- Common identifiers (i.e. SNOMED, ICD, Global Unique Identifiers) but not backwards compatible **(heavily fragmented)**
- Countries did, and do, 'their own thing' **(few incentives to change, fewer international incentives)**

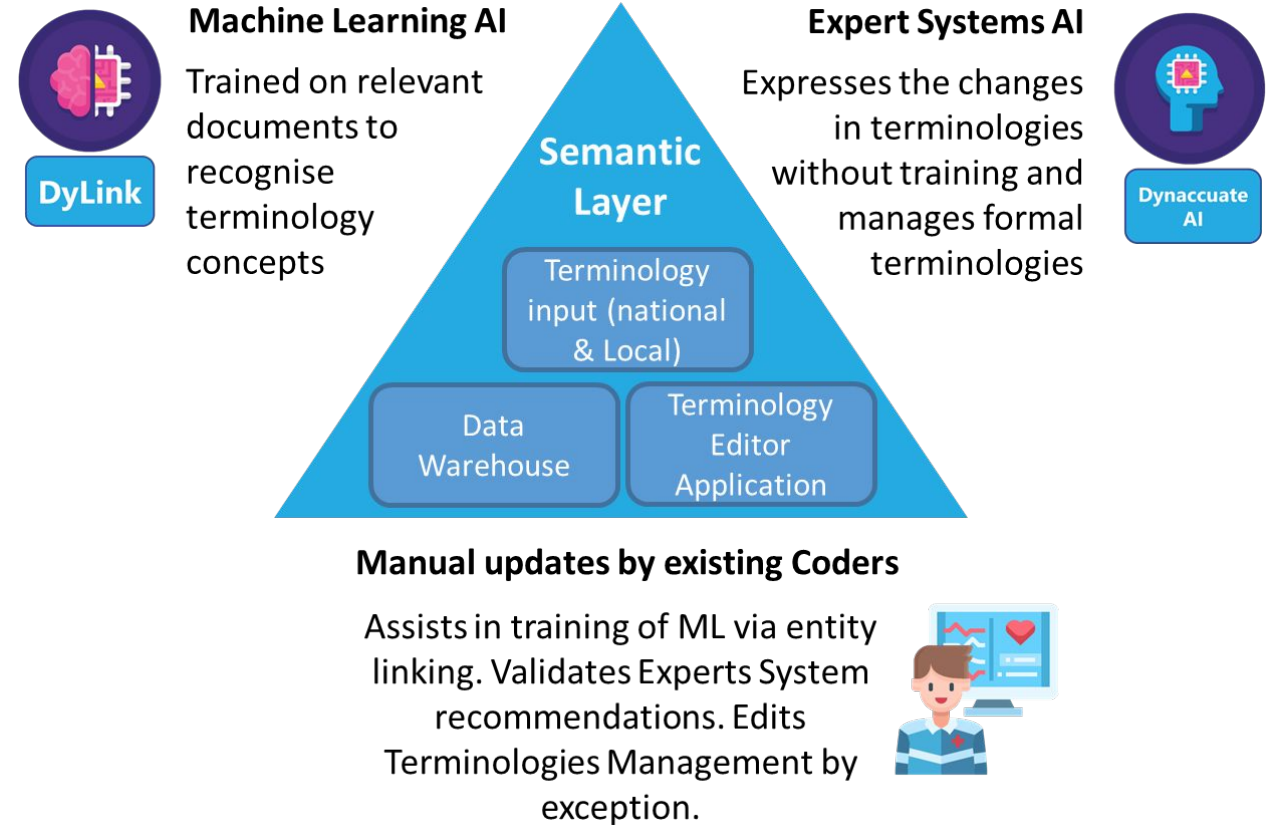


Semantic Interoperability Triad

Because of the complexity of health data, there's no easy solutions...

AI is fundamental...

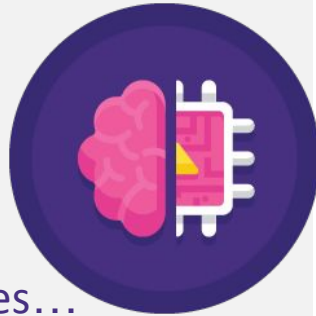
Expert Systems and Machine Learning AIs are the main solutions.



Machine Learning vs. Expert Systems

Machine Learning

- Requires data sets & human training
- Quality is determined by both data & training
- When trained can 'read' data
- Many off-the-shelf tools
- Expensive to implement
- Not transferrable between languages...



Risks/Quality factors:

- Should be trained on data sets of the type which it will later process.
- Training an AI can be quite mundane and repetitive.
- Will need ongoing training to maintain accuracy

Expert Systems

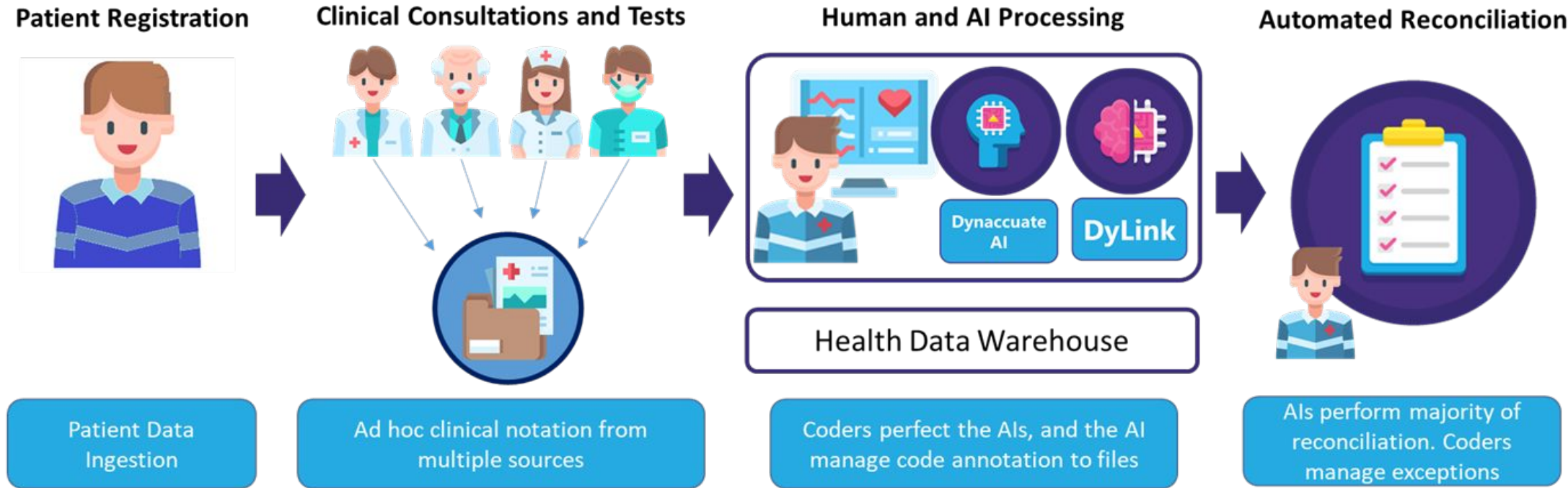
- No training – the AI simply expresses the changes
- Quality based on sophistication of the rules
- Manages the meta-data
- Much cheaper to implement
- Transferable
- Expensive to build (> €1m!!!)



Risks/Quality factors:

- 'Low hanging fruit' has all been taken
- Initial algorithms will need to be persistently tested. Better if peer reviewed
- Depending on how generic the algorithms are, they may transfer or not

Short term impact

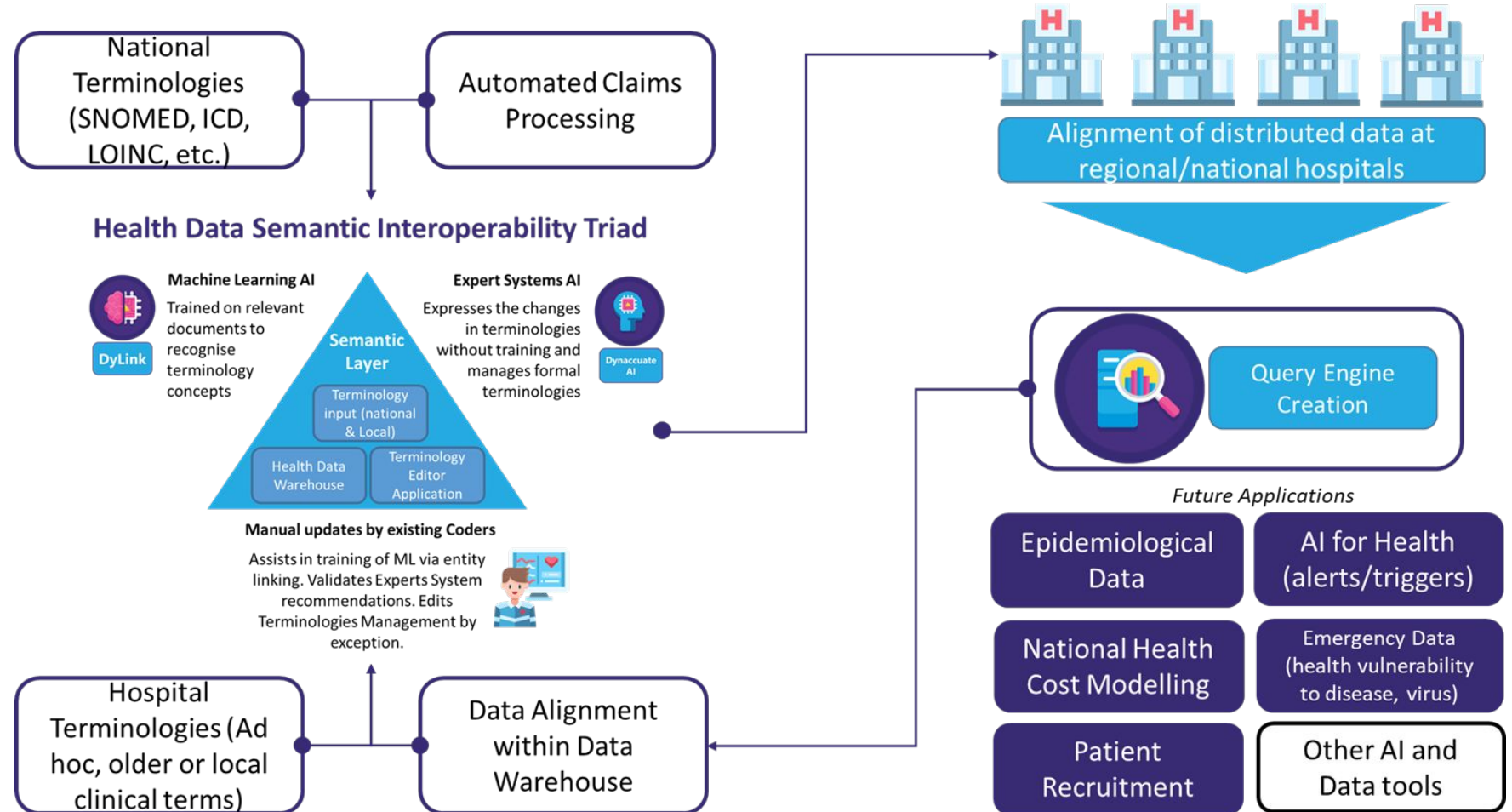


Most clinical coding becomes automated...

Long Term Impact...

As AI becomes prevalent the following will happen:

- **Manual coding disappears**
- Skilled staff work on **AI validation & perfection**
- **Capture of all health indicators** (i.e. symptoms data), not just reimbursable codes
- **'Super AI'** enabled by better health data



Reliable, granular health data enables a completely new economy

Q&A with the audience



Compile session summary

