



VEIL.AI Next-generation anonymization for safe secondary use of sensitive data

The challenges and opportunities of commercializing innovations to the public sector in Finland and abroad

**17th of April 2024
Lauri Kuronen**

VEIL.AI



ENABLING USE OF SENSITIVE DATA

**VEIL.AI is a deep tech company
specializing in data privacy enhancing technologies**

Pseudonymized data

Anonymized data

Synthetic data

Privacy Risk
Analysis

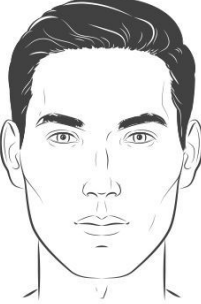

Quality
Analysis



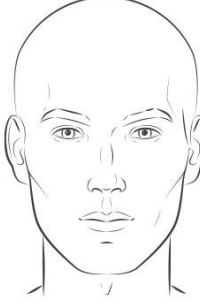

IF DATA IS GDPR COMPLIANT IT CAN BE UTILIZED IN MANY USE CASES

GDPR = European Union's General Data Protection Regulation

GDPR APPLIES

<p>Original (raw data)</p> <p>"Hospital use"</p>  <p>John - 35 y/o male Original data</p> <p>GDPR APPLIES</p>	<p>Pseudonymized data</p> <p>"Traditional research use"</p>  <p>44455 - Male Personal identifiers removed</p> <p>GDPR APPLIES</p>
---	---

GDPR DOES NOT APPLY

<p>Anonymized data</p> <p>"Enabling new use cases"</p>  <p>Male No identifiable information</p> <p>GDPR-FREE</p>	<p>Synthetic data</p> <p>"Enabling new use cases"</p>  <p>Male - 35 y/o Statistical data based on real data</p> <p>GDPR-FREE</p>
--	--

Traditionally pseudonymized data has been used in research – but GDPR applies

Anonymized data is GDPR compliant

WHY ANONYMIZED DATA OVER SYNTHETIC DATA



Next generation anonymization technology brings the quality of anonymized health data to the next level

Anonymized data

"Enabling new use cases"



Male

No identifiable information

Always from real subjects with user defined criteria.

Fulfills set anonymization criteria and there's no room for unknown bias.

Anonymity criteria apply and guarantee privacy of subjects.

Suitable for real downstream analysis, testing and system development. Scientific work can be performed on anonymized data.

High-dimensional elements can be coped with correct configuration to maintain quality and utility.

Data is generated based on some model or set of rules.

Generative model might be biased in some unknown fashion.

Can include real subjects by small chance as its generated randomly.

Suitable mainly as test data for system development and testing. Not for real downstream analysis and science.

High-dimensional elements / variables are vulnerable to Curse of Dimensionality, which can ruin quality and utility.

Synthetic data

"Enabling new use cases"



Male - 35 y/o

Statistical data based on real data

OUR SOLUTION CAN HELP IN MANY USE CASES

DO YOU WANT TO:

...IMPROVE YOUR DATA SCIENCE CAPABILITIES
(BY BRINGING ADVANCED ANONYMIZATION TO YOUR DATA ANALYSIS PLATFORM)?

...ENABLE TRANSBORDER DATA COLLABORATIONS (E.G., COMBINING EU AND US DATA)?

...RE-USE YOUR OWN EXISTING (RCT) DATA
(AND INCREASE THE VALUE OF YOUR DATA)?

... SPEED UP AND REDUCE THE COST OF CLINICAL TRIALS (BY USING ADVANCED ANONYMIZED DATA FOR VIRTUAL CONTROL ARMS)?

...DEVELOP AI ALGORITHMS WITH ADVANCED ANONYMIZED OR SYNTHETIC DATA FOR BETTER TREATMENTS?

...GET MORE RWE
(BY UTILIZING ADVANCED ANONYMIZED RWD)?

...INVESTIGATE RARE DISEASES WITH LARGER DATASETS
(THAT HAVE BEEN COMBINED GDPR-FREE)?

...ENABLE IMPROVED DIAGNOSIS AND PERSONALIZED MEDICINE (BY UTILIZING ADVANCED ANONYMIZED DATA)?

OUR SOLUTION CAN PLAY A BIG ROLE IN THESE & ALSO ADDITIONAL USE CASES

VEIL.AI

VEIL.AI ANONYMIZATION ENGINE



End-to-end software solution for sensitive health data privacy protection.

Automates health data
anonymization
synthetization
pseudonymization

Works with continuous data streams

BENEFITS

Much faster access to data (Anonymized or synthetic data is not under regulations like GDPR or HIPAA. High level of automation)

Much higher quality and utility of data than with the competing techniques.

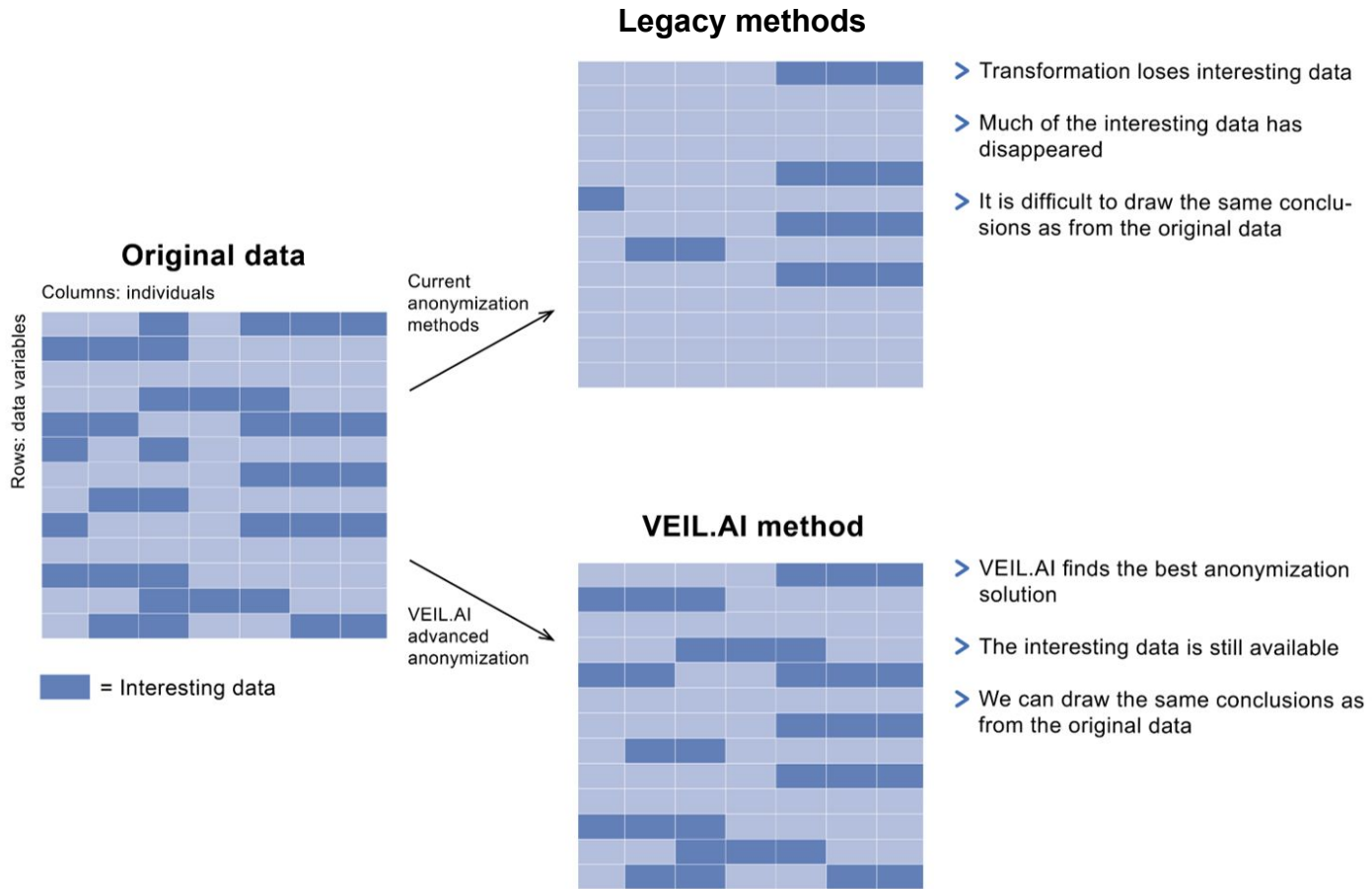
Easier and often the only way to facilitate data collaboration when data is coming from several sources.

- Patent coverage for multiparty anonymization



NEXT-GENERATION ANONYMIZATION vs. LEGACY METHODS

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



Legacy methods cannot provide good enough data quality. You cannot draw the same conclusions as from original data.

Interesting and relevant data remains after anonymization. You can draw the same conclusions as from original data.

HELSINKI UNIVERSITY HOSPITAL'S ACUTE LEUKEMIA PROJECT



#42 on the World's best hospitals (Newsweek)



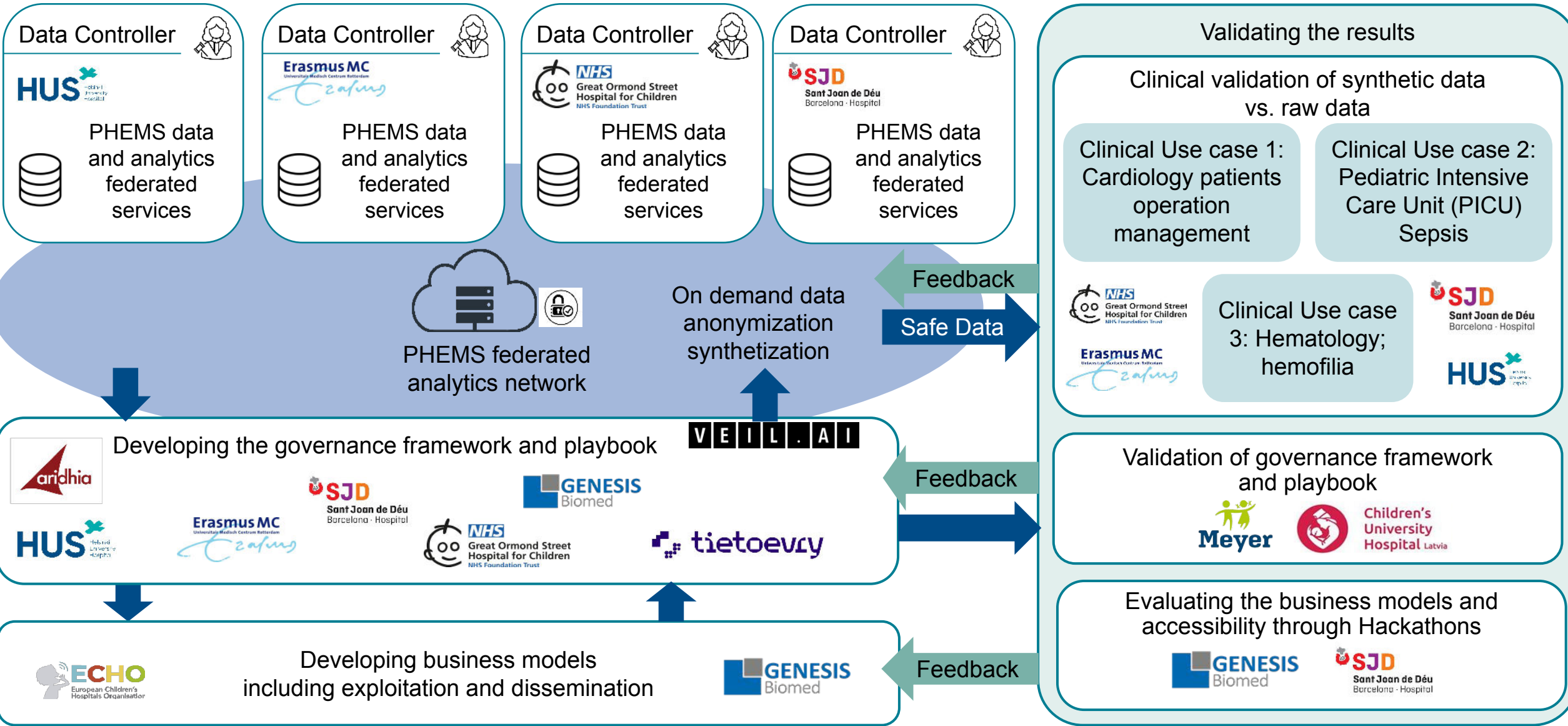
27.000 employees



870.000 surgeries

- Goal
 - Improve diagnosis and medication of acute leukemia patients using AI (pilot cohort size: 3,000, total cohort size: 310,000).
- Challenge
 - Allow data to be collaboratively worked on and shared between internal projects and with other university hospitals, academic and commercial partners.
- Result
 - VEIL.AI's technology is used to anonymize individual-level OMOP data.
 - **Anonymized OMOP data can be shared internally and with external partners.**

MULTI-PARTY HOSPITAL DATA COLLABORATION ECOSYSTEM



BAYER'S FUTURE CLINICAL TRIALS PROJECT

- The goal of the project is to improve Bayer's global drug research processes **to speed up clinical trials, reduce costs and increase patient safety.**
- Bayer benchmarked European data anonymization companies and selected VEIL.AI as their anonymization partner.
- **VEIL.AI provided the best next-generation technology for data anonymization and synthetic data generation.**
- Bayer and VEIL.AI were invited to present the "Future Clinical Trials" project in the PHUSE Data Transparency Event on June 30th, 2022



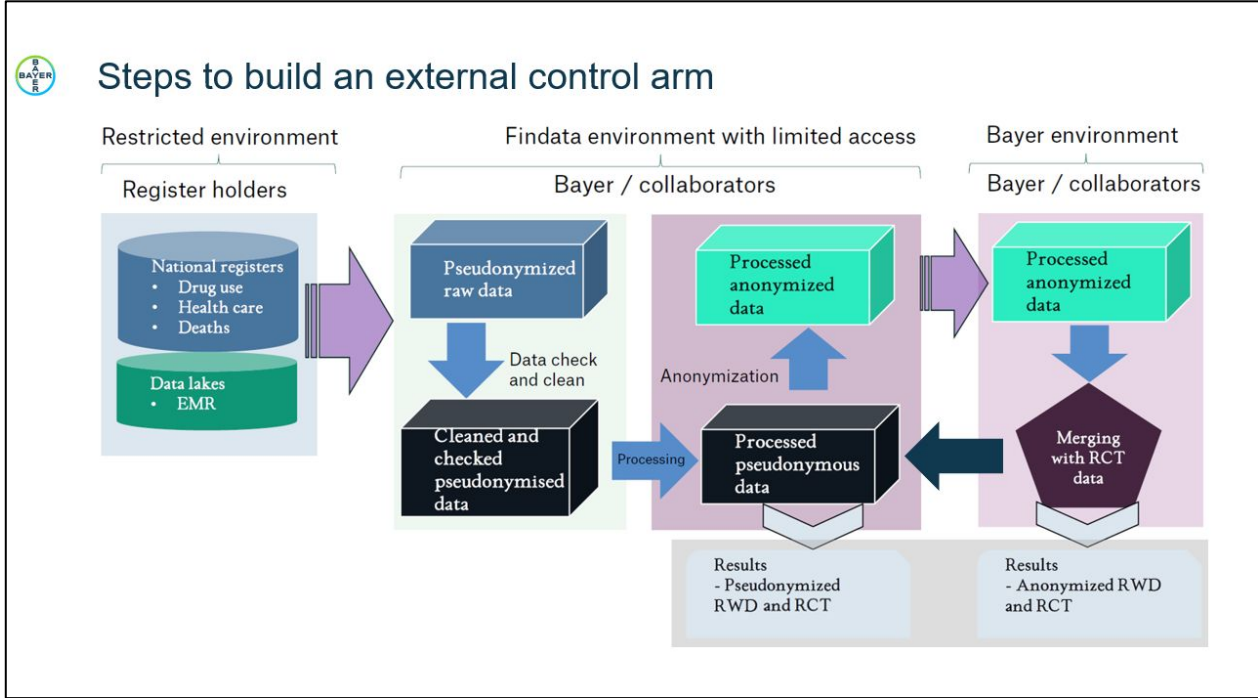
The Global Healthcare
Data
Science community



BAYER'S FUTURE CLINICAL TRIALS PROJECT

Building the external control arm for Bayer:

- 1. Data for 3300 individuals was collected from Finnish national registers and data lakes
- 2. Data was processed and pseudonymized in Findata's Kapseli® [*1] trusted environment
- 3. The pseudonymized data was anonymized by VEIL.AI in Kapseli®
- 4. Findata approved the protocol how the anonymization was verified
- 5. Bayer is free to use the anonymized data for secondary purposes



*1 The Finnish Act on the Secondary Use of Health and Social Data required that the data was processed in a trusted environment approved by the Finnish Social and Health Data Permit Authority Findata

PUBLISHED STUDY: NEXT-GENERATION ANONYMIZED DATA ENABLES SAME CONCLUSIONS AS TRADITIONAL PSEUDONYMIZED DATA

“Utilization of anonymization techniques to create an external control arm for clinical trial data”

Background:

- Individual-level real-world data (RWD) collected during daily healthcare practices are increasingly used in medical research to assess questions that cannot be addressed in the context of a randomized controlled trial (RCT)
- The aim of the study was to investigate and compare how well pseudonymized and anonymized RWD perform in the creation of an External Control Arm for RCT

Results:

- **VEIL.AI's next-generation anonymization provided excellent results**
 - Population baseline characteristics, including potential confounders were not affected
 - **Conclusions made using next-generation anonymized data are as reliable as when using the original pseudonymized data**



Jussi Leinonen
Strategic Project Lead
Bayer

“This is a significant achievement. In our study we could draw the same conclusions from anonymized data as from traditional pseudonymized, individual-level research data.”

CHALLENGES IN COMMERCIALIZING INNOVATIONS

Our experience is based on the collaboration with university hospitals, research institutes and biobanks in Finland.

- Identifying the right stakeholders and decision makers
 - o Data-driven clinicians, IT department
 - o Who does the purchase decision
- Slow processes
 - o Decision making & purchase process
 - o Care pathway
 - o Deployment widely is prolonged process
- Procurement process
 - o Country & EU specific procurement limits
- Budgeting periods
- Existing suppliers and framework agreements
- Lack of resources (company & customer point of view)
 - o Money
 - o Personnel
- Legislation
 - o Healthcare in general is heavily regulated environment
 - o MDR / IVDR / FDA (country or area specific)
- Timing
- The stigma of piloting
 - o Rarely ends in commercial cooperation

OPPORTUNITIES IN COMMERCIALIZING INNOVATIONS

Our company is a spin-off from the University of Helsinki and Institute for Molecular Medicine.

- Experience in working with public and private sector stakeholders
 - Research institutes
 - Universities
 - Hospitals
- Academic background gives you the “elephant nerves”
- Customer understanding and identifying the real needs / pains
- Public healthcare is a huge market opportunity for commercialization
- The commercial agreement
 - Usually quite long
 - Include option years
- Stable customer and payer
- Reference value with similar customers (Finland and abroad)
- Collaboration and consortiums → easier for buyer
- You need to show the value for the customer



Lauri Kuronen
CEO

VEIL.AI Denmark ApS
+358 400 244 209
lauri.kuronen@veil.ai
www.veil.ai

VEIL.AI has re-invented anonymization to solve access to data and data quality challenges. This revolutionaizes how data intensive medical and health care research and operations will be conducted.

