INNOVATION INSIGHT Unlock privacy-sensitive data with AI-Generated Synthetic Data in Healthcare



Wim Kees Janssen Co-Founder & CEO Syntho

13-14 September INTELLIGENT HEALTH 2023 Basel, Switzerland





@IntHealthAI
#IntelligentHealthAI
#SaveLivesWithAI

HEADLINE PARTNER

U NOVARTIS

PLATINUM SPONSOR INNOVATION & TECHNOLOGY

Deloitte

The Novartis

Foundation

CITY PARTNER





Synthetic data Real impact

Unlocking Privacy-Sensitive Data with AI-Generated Synthetic Data in Healthcare & Pharma

Wim Kees Janssen CEO & Founder Syntho



Wim Kees Janssen – Founder & CEO Syntho

- Visit our booth **S3**
- Contact: kees@syntho.ai





About Syntho





Recognitions







Winner Philips Innovation Award



Winner of SAS Global Hackathon in healthcare & Lifescience Shortlisted by NVIDIA as Generative Al startup to watch

Read more

Read more

Artificial Intelligence

Software

Business Intelligence





Missed data opportunities

8 | Copyright © 2021 Syntho BV. All rights reserved

Healthcare data is the most privacy-sensitive data

- Cannot simply be used
- Time-consuming
- Extensive Paperwork
- Classic 'anonymization' is no solution

Result: no data, no insights!

Turn data access into a competitive advantage

70%

An increase in industry collaborations is expected with the use of privacy tools



More revenue for organizations that earn and maintain digital trust with customers and patients 50%

Of data will be **unlocked by privacy-enhancing techniques**



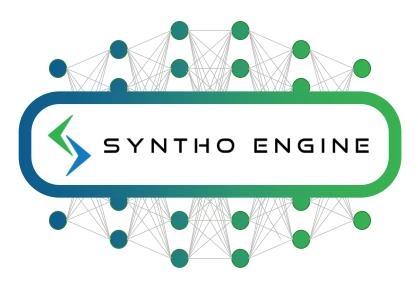
Classic 'anonymization' is no solution

Original data						Classic anonymization						
Name	Age	Gender	ltem	Price	Data		Name	Age	Gender	ltem	Price	Data
Olivia	26	Female	Shoes	€125	4 March	-	ххх	25-30	Female	Cloth	€100 - €200	March
John	75	Male	Laptop	€695	5 March	-	ххх	70-75	Male	IT	€600 - €700	March
George	41	Male	Beer	€4	7 March	-	ххх	40-45	Male	Drink	<€5	March
•••		•••			•••							
George	41	Male	Shirt	€25	9 March	-	ххх	40-45	Male	Cloth	€20 - €30	March
N=100k							N=100k				°	-

× Privacy risks

X Low accuracy

× Low scalability



Generative Al For Tabular Data

12 | Copyright © 2021 Syntho BV. All rights reserved

A synthetic data twin of your original data

Original data								
Name	Age	Gender	ltem	Price	Data			
Olivia	26	Female	Shoes	€125	4 March			
John	75	Male	Laptop	€695	5 March			
George	41	Male	Beer	€4	7 March			
		•••						
George	41	Male	Shirt	€25	9 March			

N=100k

Generate a synthetic data twin of your original data

Original data							
Name	Age	Gender	ltem	Price	Data		
Olivia	26	Female	Shoes	€125	4 March		
John	75	Male	Laptop	€695	5 March		
George	41	Male	Beer	€4	7 March		
		•••					
George	41	Male	Shirt	€25	9 March		
N=100k					1		

Synthetic Data Twin							
Name	Age	Gender	ltem	Price	Data		
NewID1	23	Male	Sofa	€790	1 March		
NewID2	23	Female	Scarf	€40	3 March		
NewID3	52	Male	Razor	€5	9 March		
•••		•••			•••		
NewIDn	35	Male	Wine	€7	7 March		
N=100k							

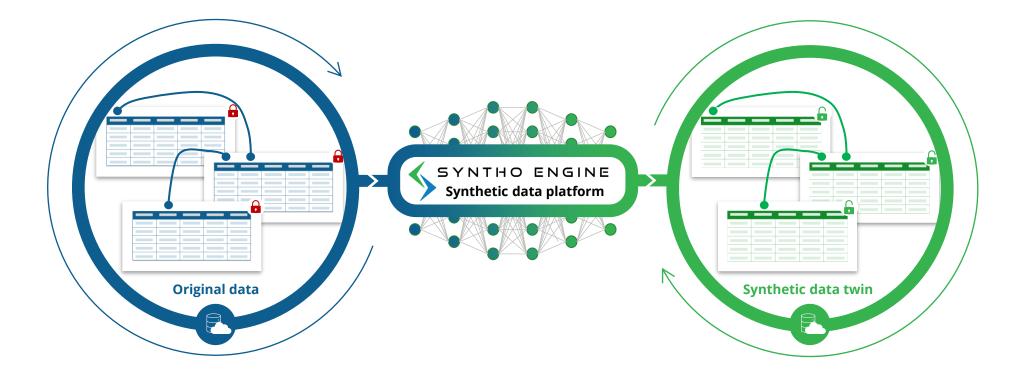
Privacy by design

✓ High accuracy

✓ High scalability

14 | Copyright © 2021 Syntho BV. All rights reserved

Al-generated synthetic twin of your original data

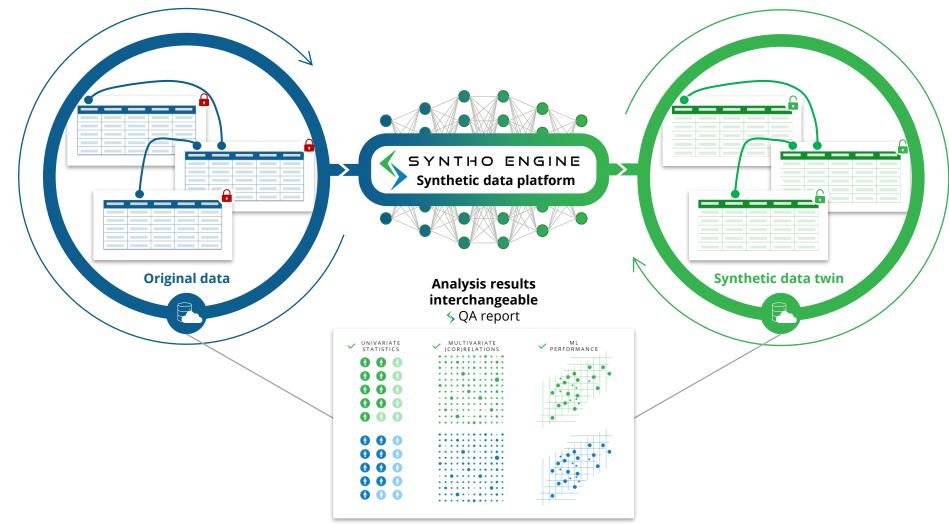


✓ High accuracy

✓ High scalability

15 | Copyright © 2021 Syntho BV. All rights reserved

Al-generated synthetic twin of your original data



What do the privacy regulators say about synthetic data?

66

"



Under the **EU General Data Protection Regulation**, data synthesis would benefit from the same reasoning as the case for **not needing legal basis for anonymization**. Under the **U.S. Health Insurance Portability and Accountability Act**, synthesis would similarly be deemed to be a form of treatment, planning and operations and therefore does **not require additional authorization by a covered entity**. <u>CC</u> iapp

EU DPAs are open **to labeling synthetic data anonymous** under data protection standards.

Article 29, EU Data Protection Work Party IAPP 1 & 2 (p.7) US Department of Health & Services HHS

A guide to the EU's unclear anonymization standards (2021)













Quality evaluation of the synthetic data

1. Syntho's Quality Assurance (QA) Report

We provide a comprehensive quality assurance report for every synthetic data run, that demonstrates the accuracy of the synthetic data compared to the original data.

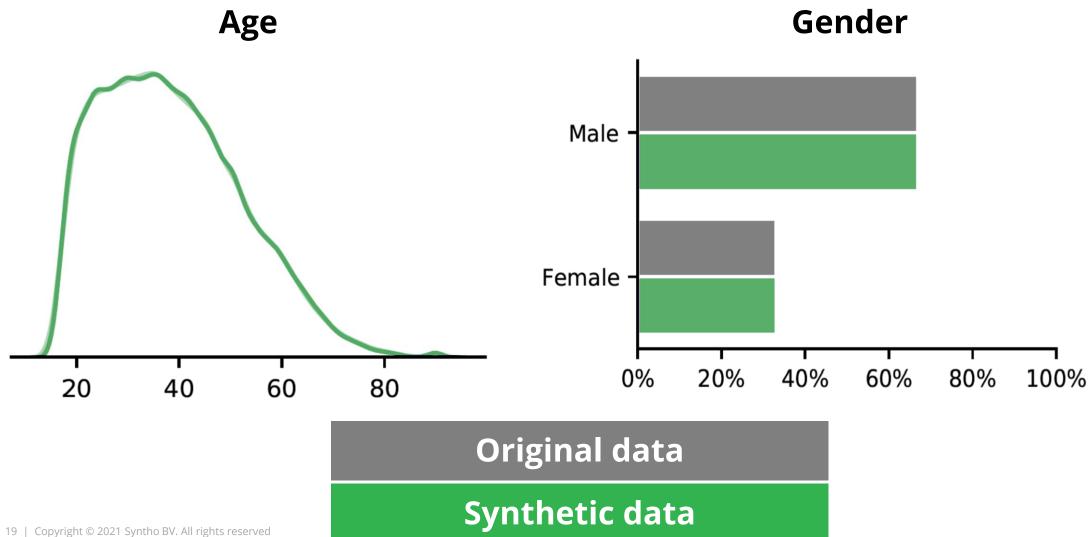
2. External assessment by the data experts of SAS

The data experts from SAS approved our Al-generated synthetic data for the use of model development as part of a real case study for a telecom customer.

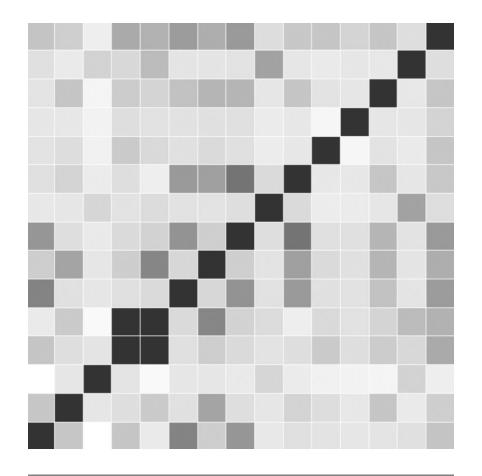
as Our synthetic data is approved by the data experts of SAS

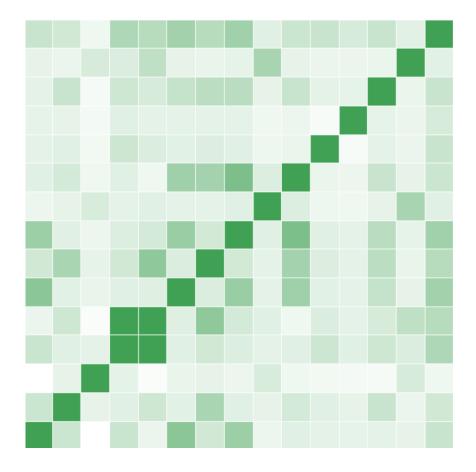


Distributions are preserved



Correlations are preserved

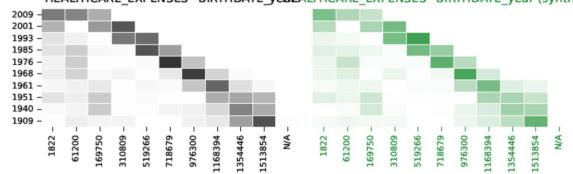




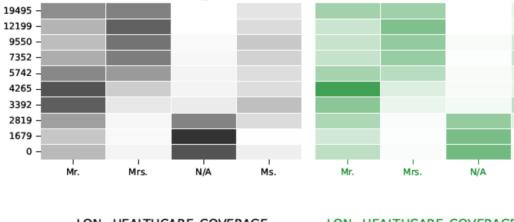
Original data

Synthetic data

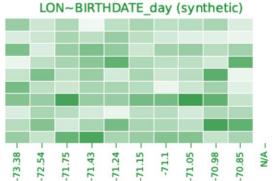
BI-variates are preserved

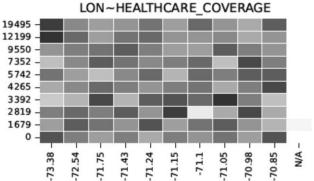


HEALTHCARE_EXPENSES~BIRTHDATE_yeldEALTHCARE_EXPENSES~BIRTHDATE_year (synthetic)

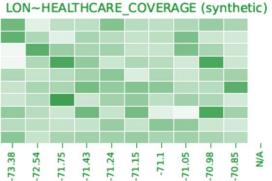


LON~BIRTHDATE_day 27 25 22 19 16 10 -70.85 -73.38 -71.75 -71.43 -71.24 -71.15 -71.1 -71.05 -70.98 54 N/A 12





PREFIX~HEALTHCARE_COVERAGE



Ms.

PREFIX~HEALTHCARE COVERAGE (synthetic

Synthetic data

Original data

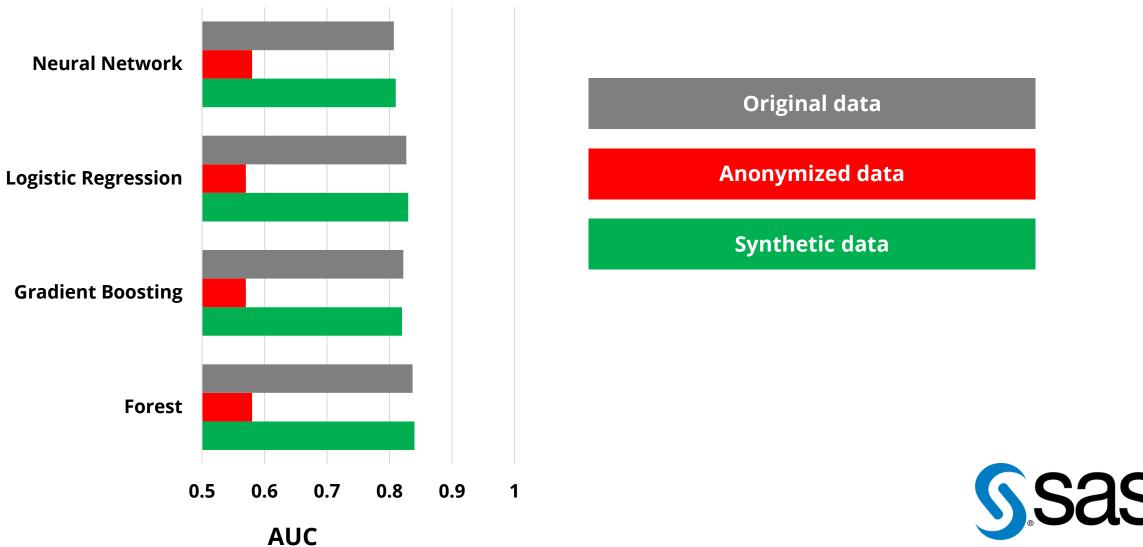
Synthetic data approved by the data experts of SAS





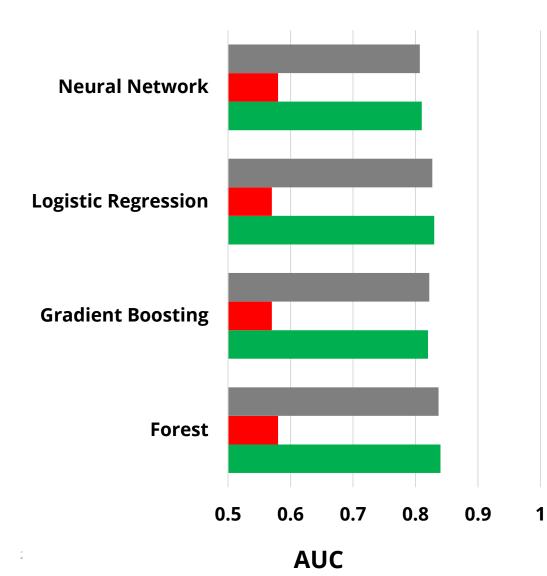
The data experts from SAS approved our Al generated synthetic data.

Synthetic data approved by the data experts of SAS



1

Synthetic data approved by the data experts of SAS



- Synthetic data is as-good-as the original data
- Anonymized data shows the worst performance
- Easy, fast, and scalable



SYNTHO ENGINE 2.0

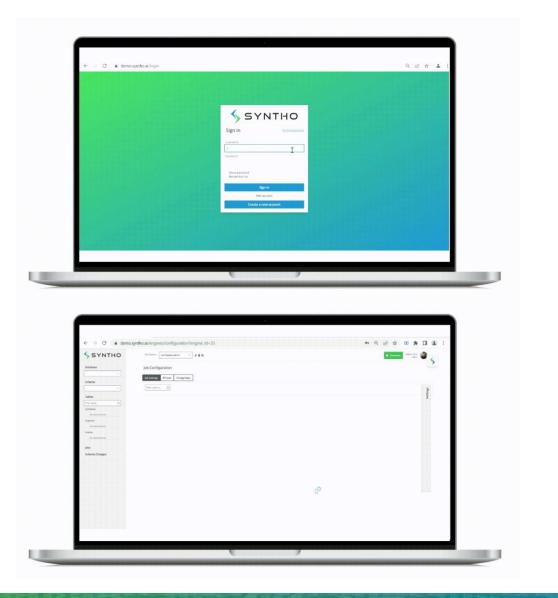
REVOLUTIONIZING AI GENERATED SYNTHETIC DATA

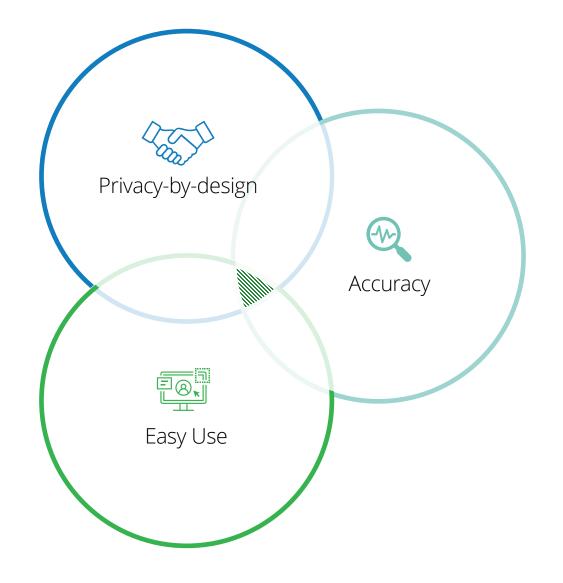
SYNTHO

SYNTHO

and the second se

3 pillars of our Syntho Engine





Why use real data when you could use synthetic data?

Now		Future
Personal		
data use	Minimize personal data	
Locked data	Unlock personal data	Synthetic data

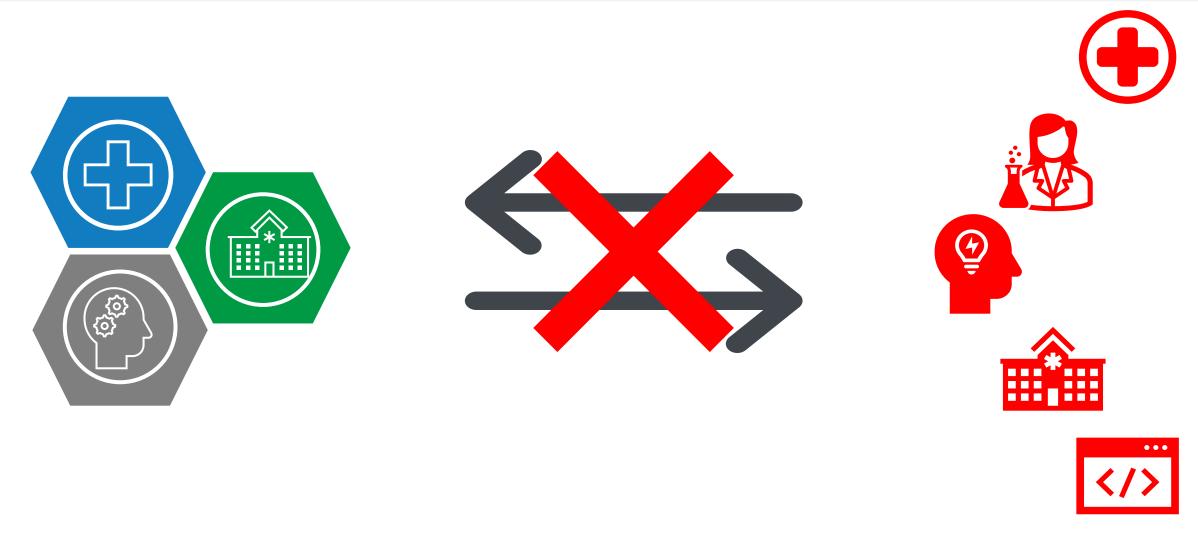
Key benefits

- 1. Less risk
- 2. More data
- 3. Faster data access

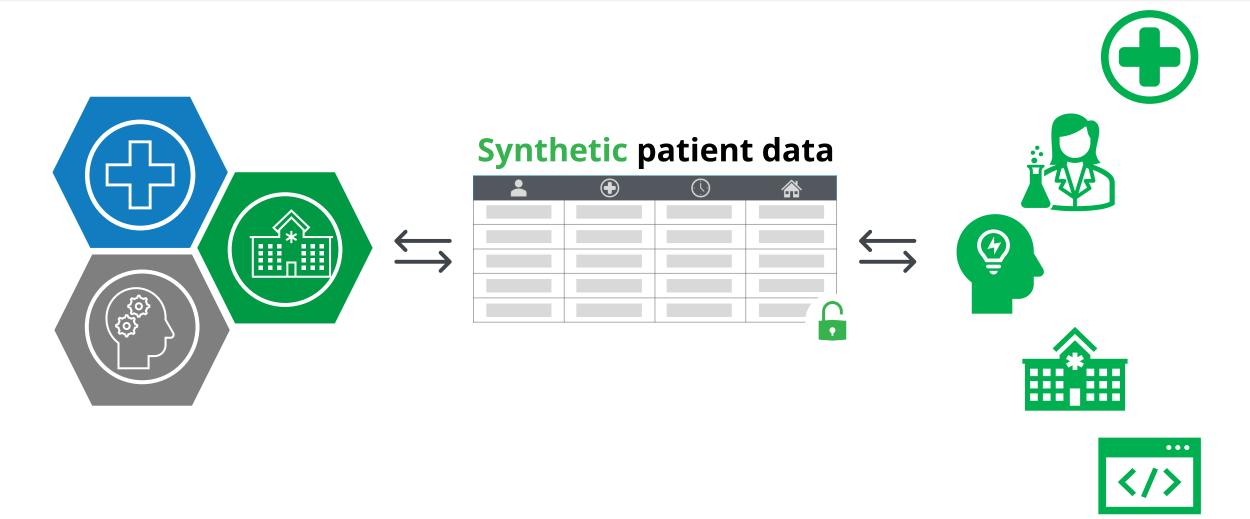


Strong data foundation

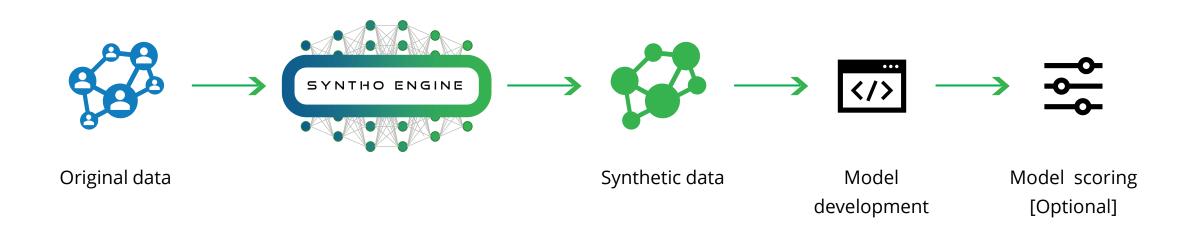
Data sharing: using and sharing data is challenging



Data sharing: freely use and share synthetic data



Analytics: Synthetic Data for Analytics



✓ Ranging from business intelligence to artificial intelligence solutions

We cooperate with leading organizations

Recognitions

Client references



31 | Copyright © 2021 Syntho BV. All rights reserved

Case Study: Synthetic healthcare data with Cedars Sinai Medical Center





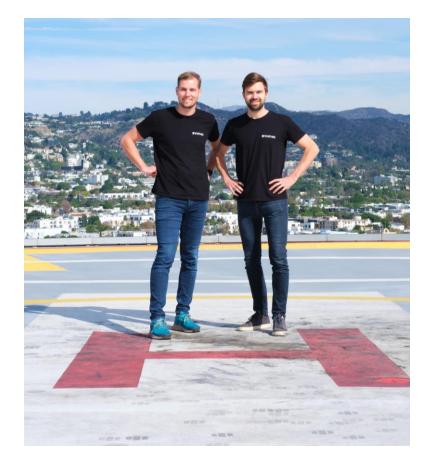


Case Study: Synthetic healthcare data with Cedars Sinai Medical Center

Examples in analytics:

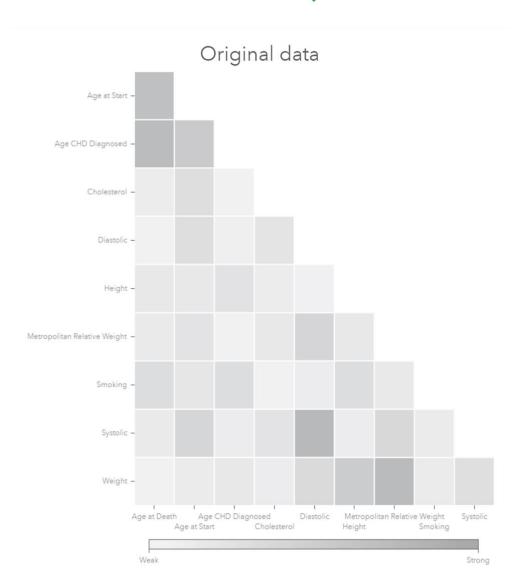
- ✓ Clinical research
- ✓ Clinical trials
- ✓ Data sharing within the industry
- ✓ Rare category diseases
- ✓ Onboarding new employees



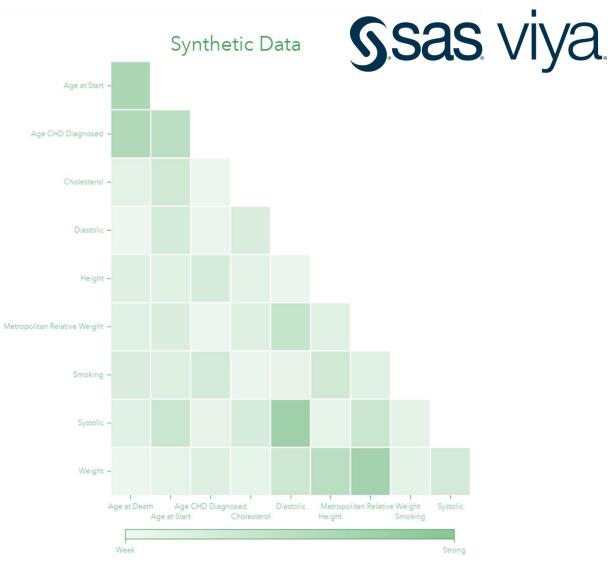


Original data

Synthetic data



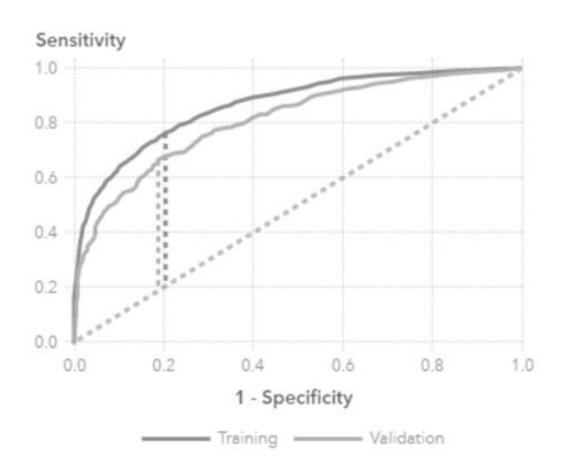
Correlations



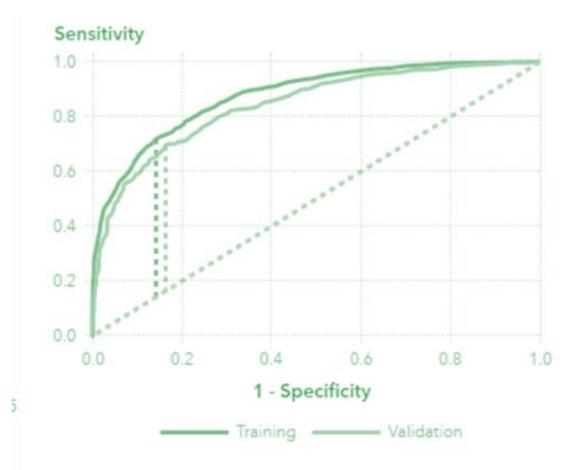
Original data

Synthetic data

Ssas viya.



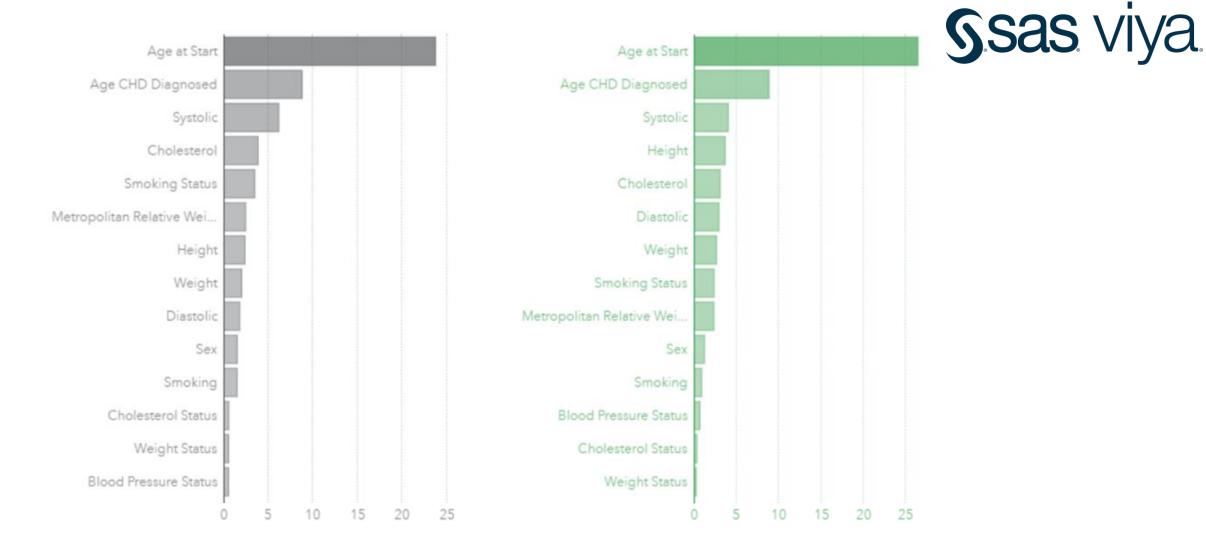
ROC Curve 🗸



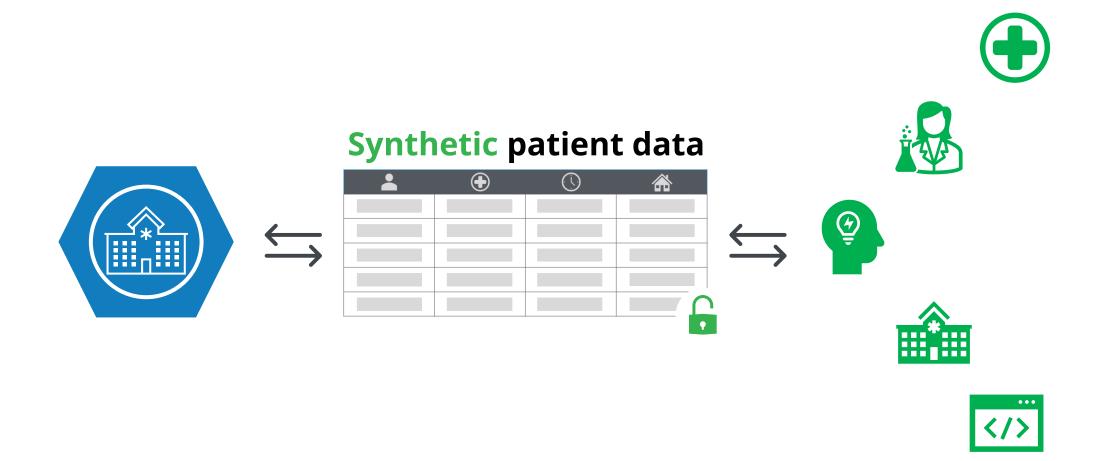
Original data

Variable importance

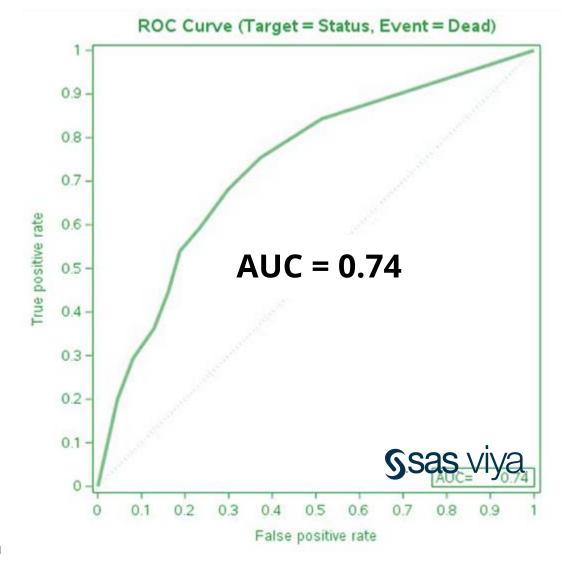
Synthetic data



Case Study: Predicting deterioration and mortality as part of cancer research on synthetic data



Case Study: Predicting deterioration and mortality as part of cancer research on synthetic data



Case Study: Predicting deterioration and mortality as part of cancer research on synthetic data

Conclusion

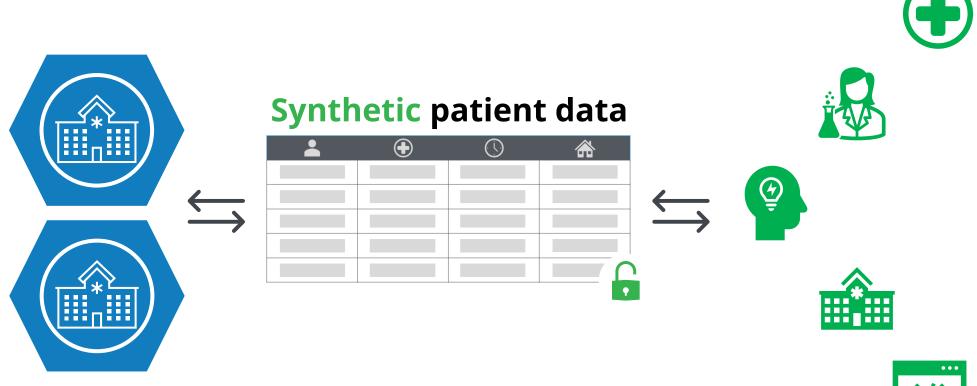
Privacy-sensitive healthcare data can be unlocked with synthetic data within the hospital to realize data-driven insights

✓ Less risk

✓ More data

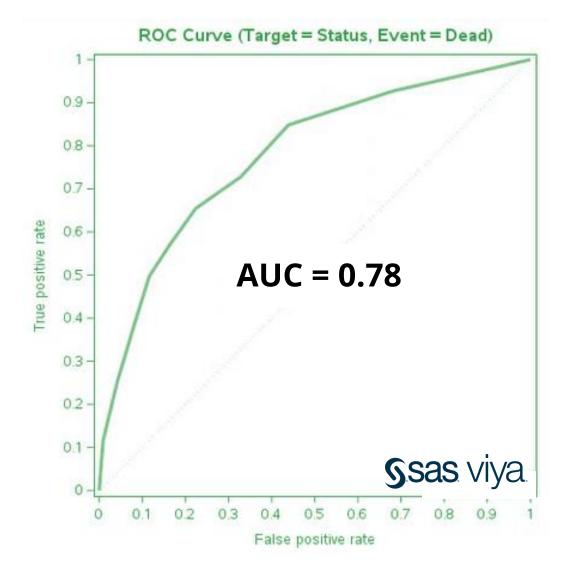
✓ Faster data access

Case Study: Combine synthetic data from *two* hospitals

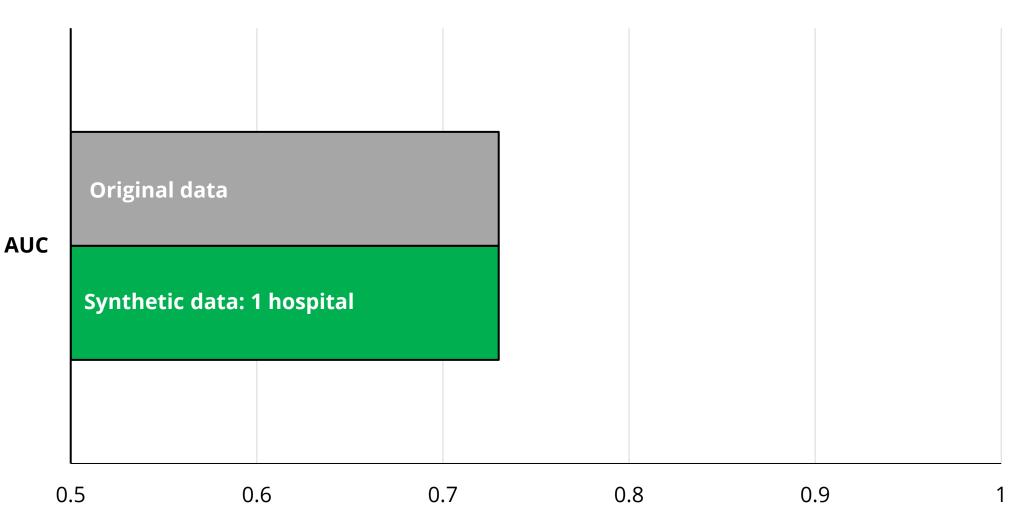




Case Study: Combine synthetic data from *two* hospitals

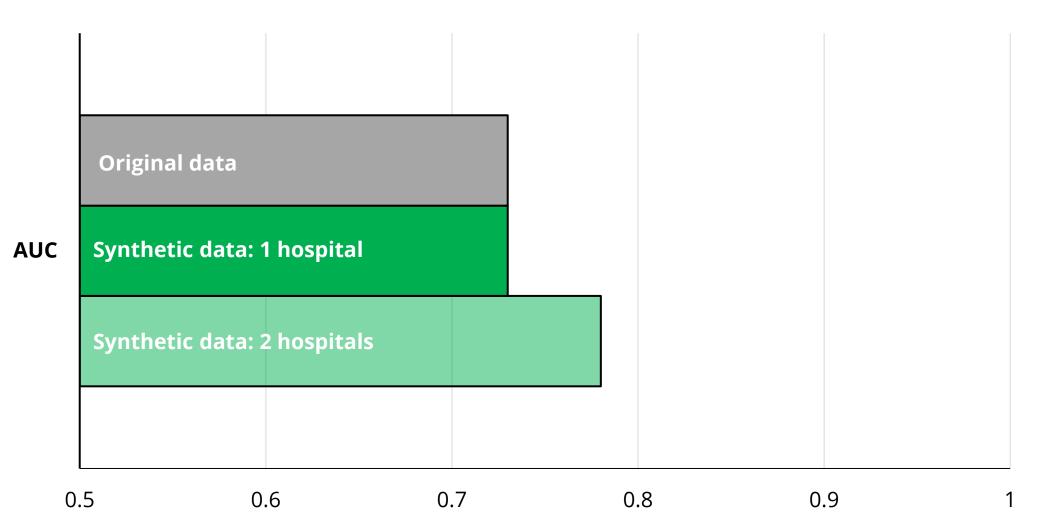


Case Study: Synthetic data is as-good-as-real and can be used for clinical research as alternative for sensitive real data



42 | Copyright © 2021 Syntho BV. All rights reserved

Case Study: Synthetic data can easily be shared. Access to more data could improve the accuracy of predictive models.



43 | Copyright © 2021 Syntho BV. All rights reserved

Special proposal to explore AI Generated Synthetic Data

Pilot	Implementation
Normal	Normal
2 Month Syntho Engine Pilot License	12 Month Syntho Engine License
Special proposal	Special proposal
✓ Deployment fee waived	✓ Deployment fee waived
\checkmark 1 additional month for free	\checkmark 2 additional months for free

Valid if interest in this proposal is indicated at our booth (S3)





Realize data opportunities

46 | Copyright © 2021 Syntho BV. All rights reserved

Turn data access into a competitive advantage

70%

An increase in industry collaborations is expected with the use of privacy tools

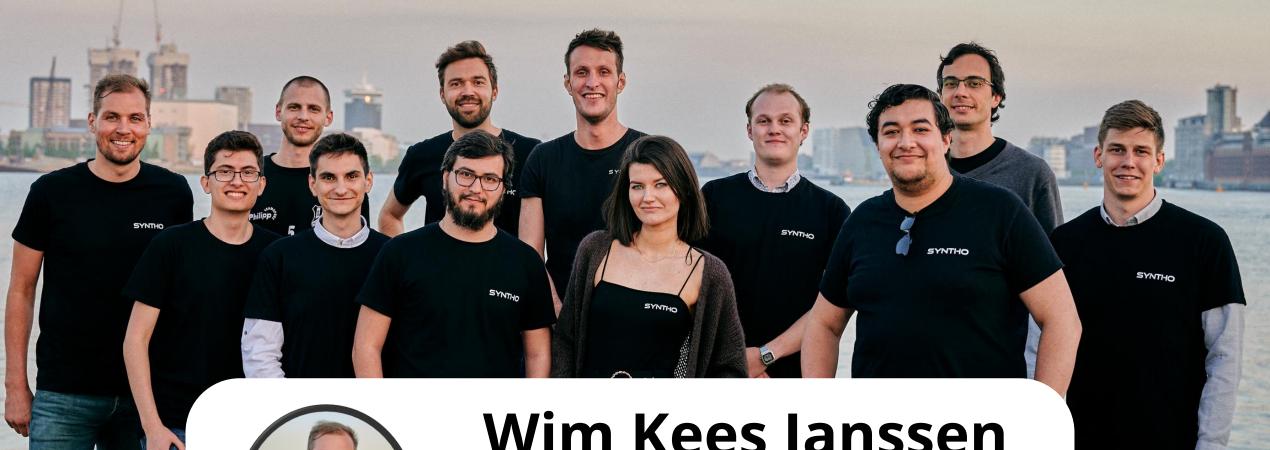


More revenue for organizations that earn and maintain digital trust with customers and patients 50%

Of data will be **unlocked by privacy-enhancing techniques**



Synthetic data – Real people!





Wim Kees Janssen

Founder & CEO Booth: **S3** Contact: kees@syntho.ai