MDClone

Liberate Data. Transform Care.

The *New Data Paradigm & The New world of Synthetic Data*

Boaz Gur-Lavie

Ziv Ofek
Founder & CEO

Founder & former Chief Innovation Officer, *dbMotion*

Acquired for $235M

dbMotion

Acquired by

Allscripts

Boaz Gur-Lavie
Founder & CFO

Former CFO STARLIMS and Abbott Labs informatics

Acquired for $123M

STARLIMS

Acquired by

Abbott Informatics

Luz Erez
Founder & CTO

Founder of ISI which became IPEX-ISI

IPEX-ISI

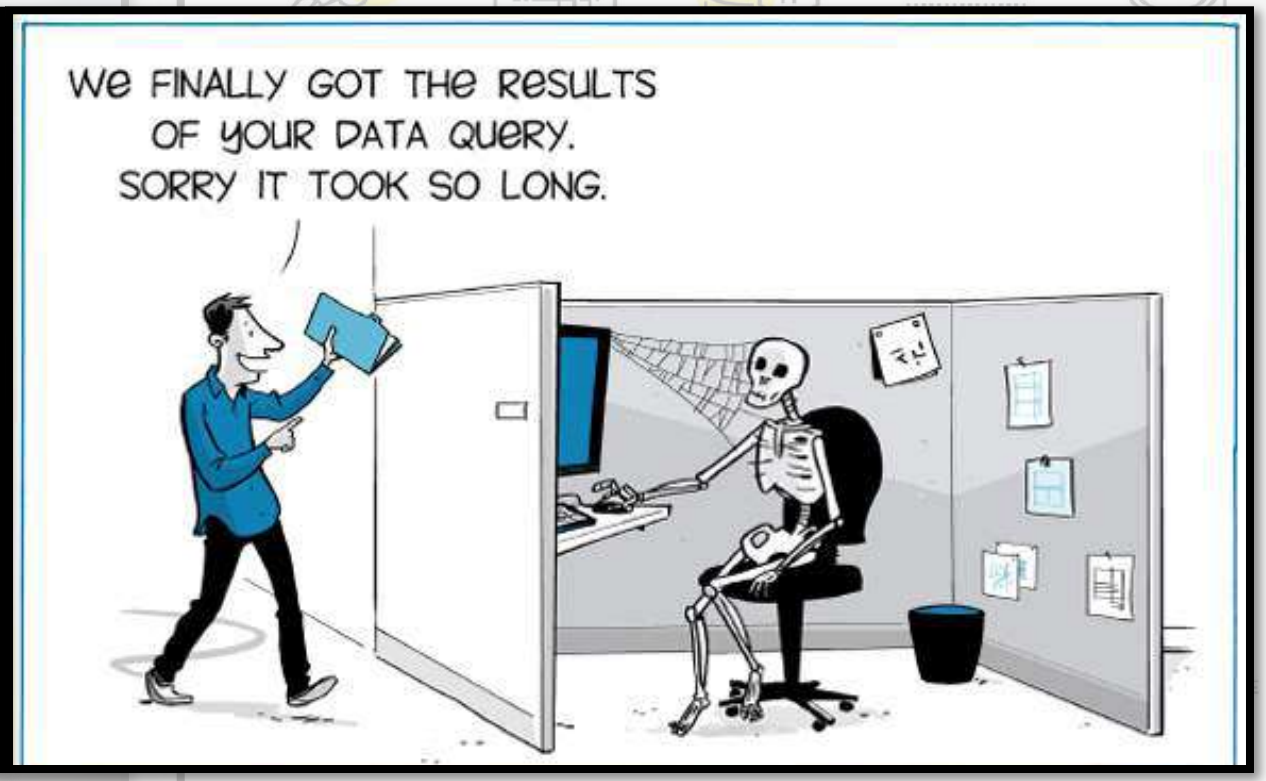
Acquired by

NESS

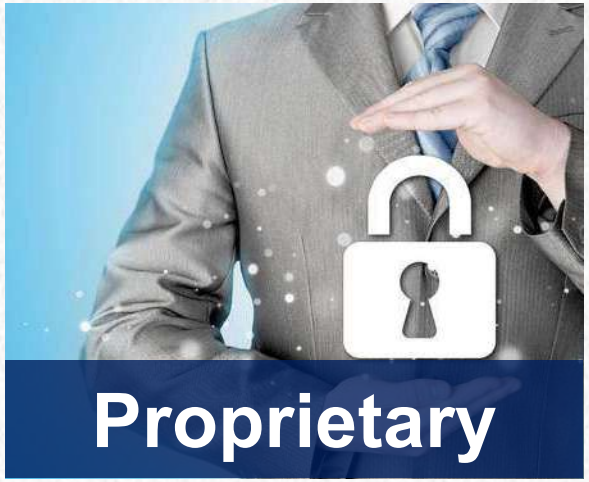
The Power of the Data… and Reality

Data

Technologies

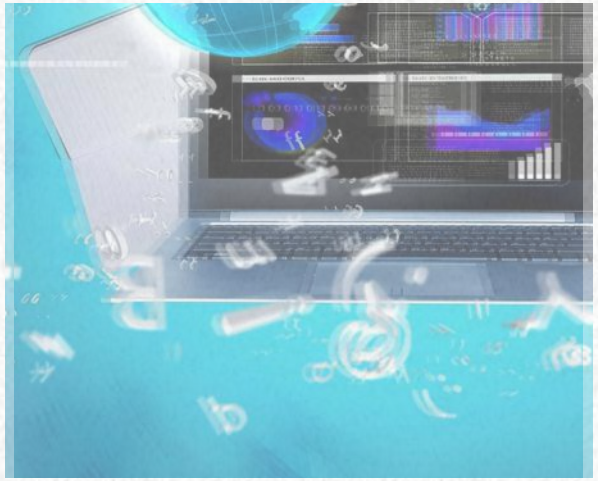


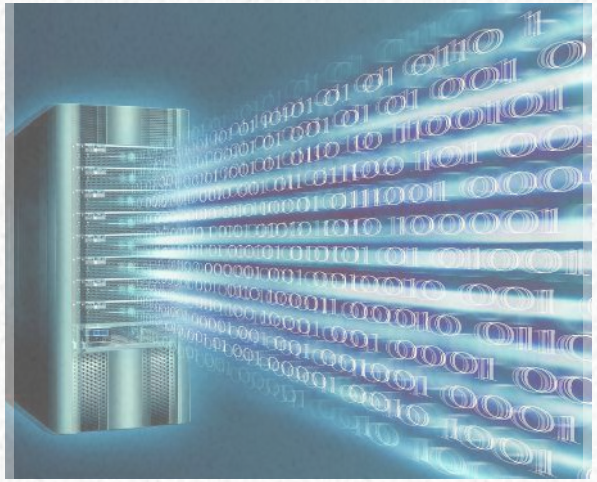
Key Challenges



MDClone

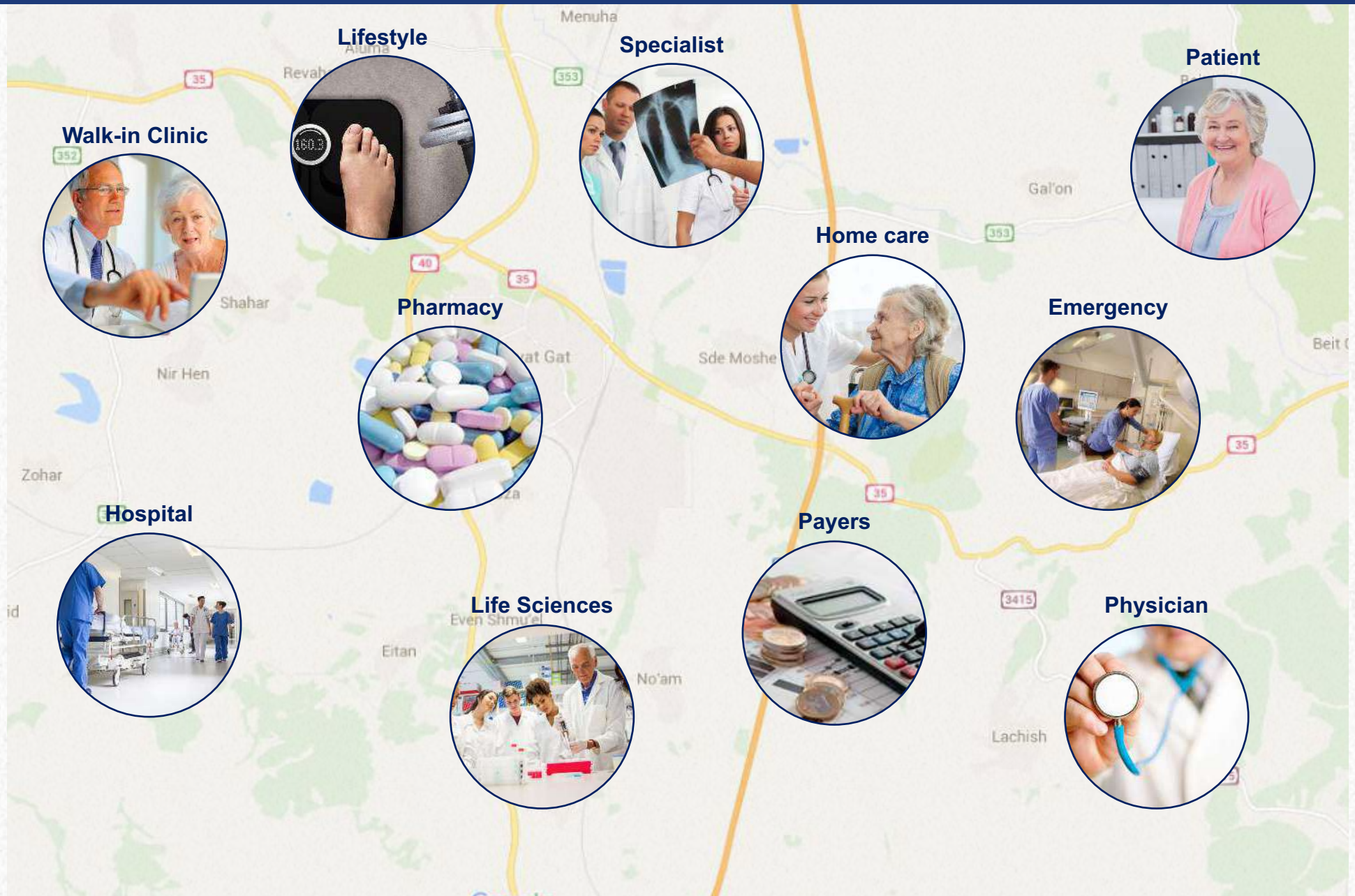
Key Challenges





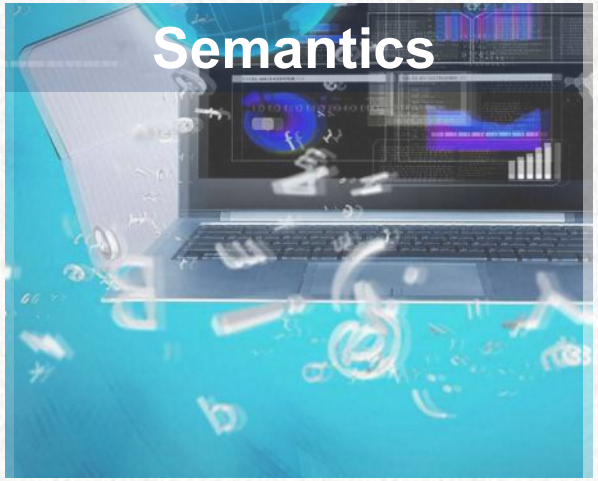
MDClone

Data Silos



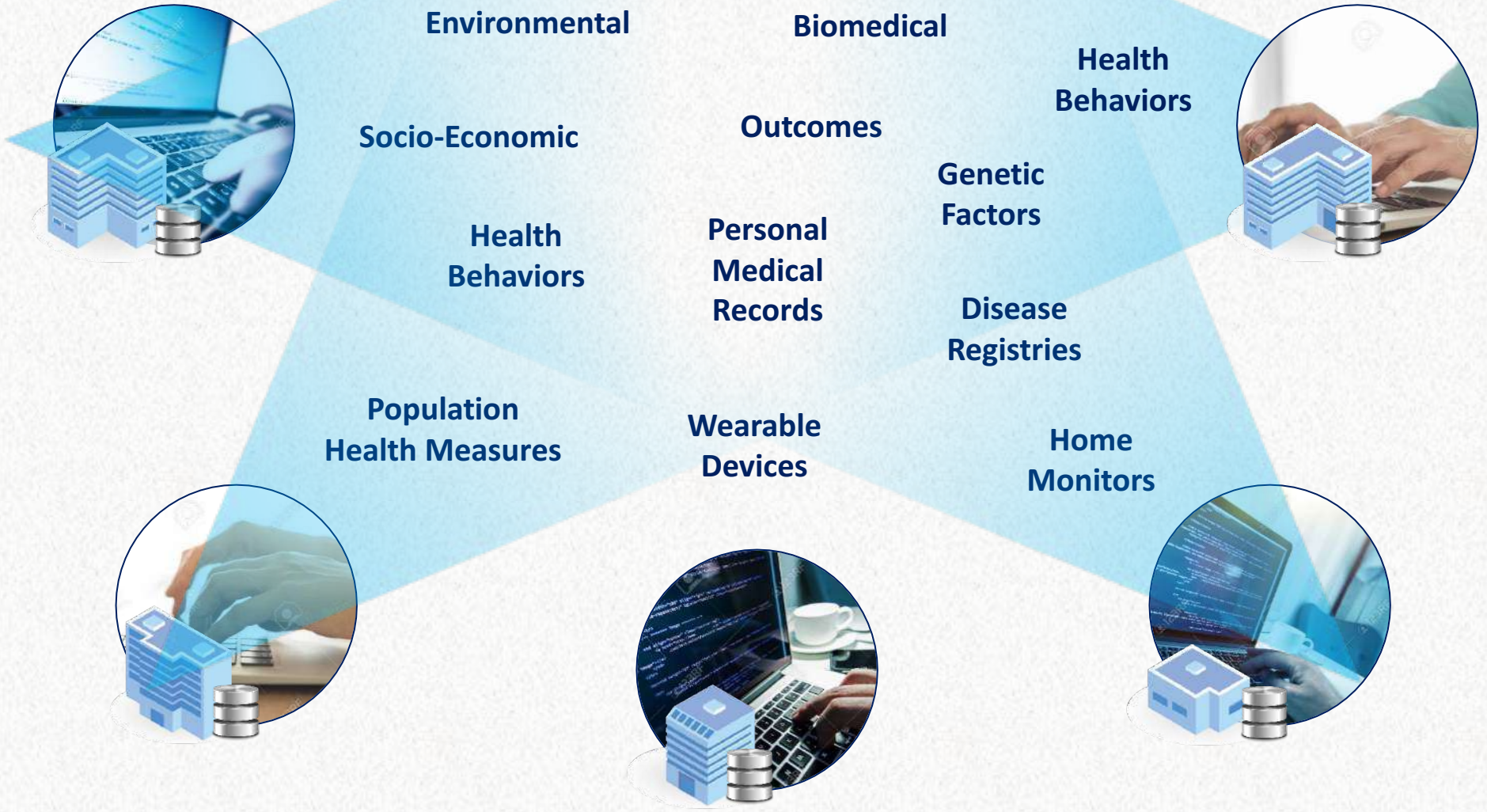
MDClone

Key Challenges



MDClone

Data Collection is Growing Exponentially

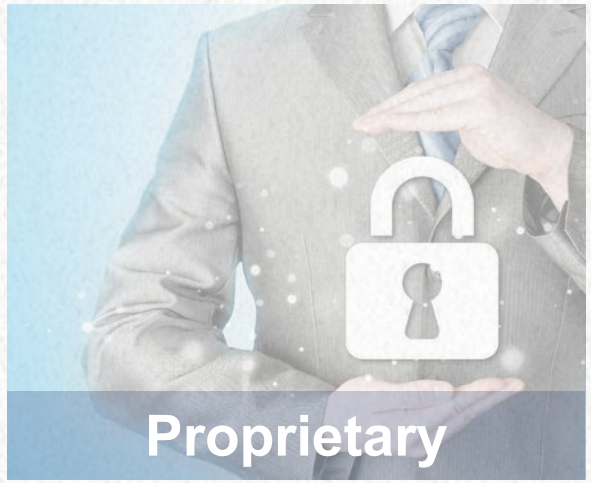


MDClone

BIG DATA

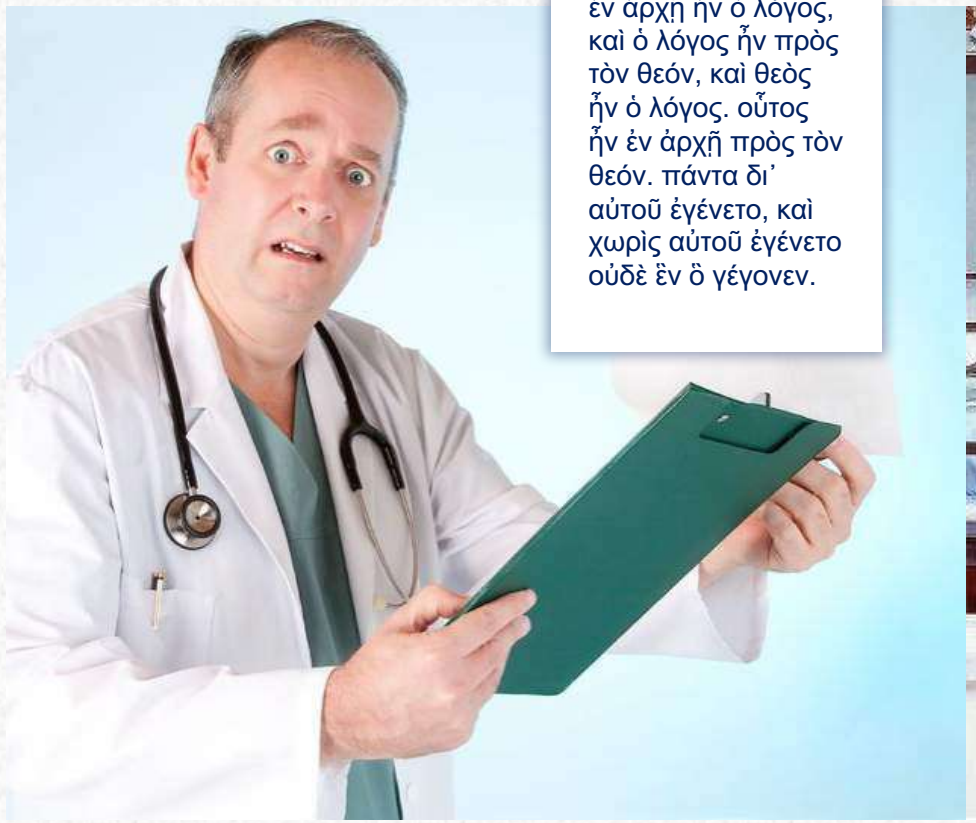
SEAST
BIG
All the rest are only technologies...
DATA
large
volume
variety
size
world
complex
analysis
technologies
different
using
now
datasets
information
definition
processing
important
problem
business
time
mean
refers
technologies
systems
analytics
new
term
storage
velocity
sources
amounts
many
tools
use
innovation
often
insights
things
means
standard
used
way
decisions
solutions
people
better
tool
changes
change
organization
behavior
gather
lot
essential
fit
making
tb
store
patterns
single
make
databases
storing
organizations
buzzword
traditional
internet
methods
memory
also
issues
products
processes
number
capturing
advances
actually
understand
right
like
technology
value
scope
made
capable
maybe
possibilities
solve
main
get
computing
meaning
need
Data
Exports
help
deal
relationships
relative
collecting
continues
enough
small
scale
evidence
collect
warehouse
derived
related
processed
approach
describes
much
system
whether
lives
social
created
challenges
set
anything
computational
reality
growth
across
dataset
services
something
otherwise
ethical
software
commerce
disparate
requires
questions
excel
aspect
later
machine
diverse
really
think
decision
Legal
Protection
particular
Mexico
Helsinki
technological
society
analytical
every
fit
lot
essential
making
store
patterns
single
make
organizations
number
buzzword
traditional
internet
methods
memory
also
issues
processes

Key Challenges

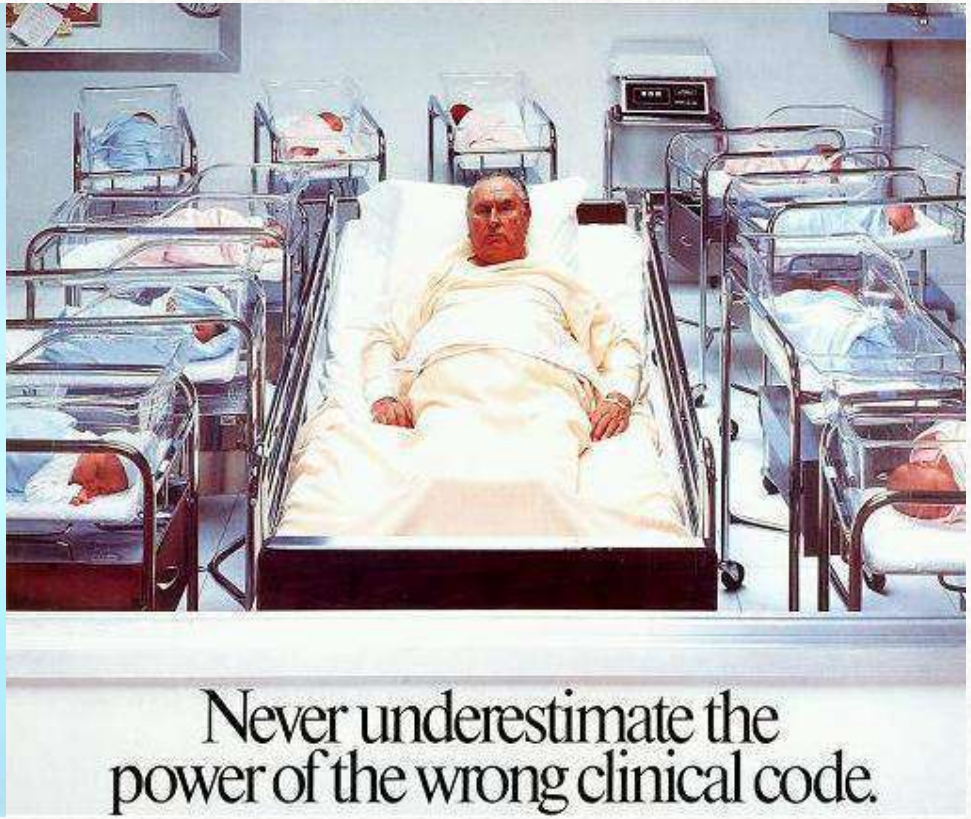


MDClone

Semantics



ἐν ἀρχῇ ἦν ὁ λόγος, καὶ ὁ λόγος ἦν πρὸς τὸν θεόν, καὶ θεὸς ἦν ὁ λόγος. οὗτος ἦν ἐν ἀρχῇ πρὸς τὸν θεόν. πάντα δι’ αὐτοῦ ἐγένετο, καὶ χωρὶς αὐτοῦ ἐγένετο οὐδὲ ἓν ὃ γέγονεν.

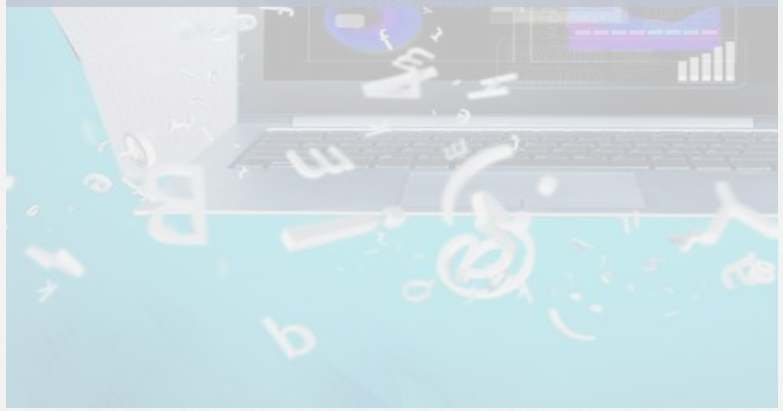


Key Challenges

Silos

Data Volume

Language



Proprietary

Data Quality

Privacy

MDClone

Access to Data Today

Risk Mitigation

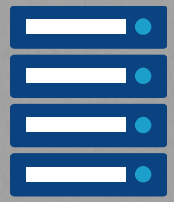
- Formal study submission
- IRB approval
- Business associate agreements
- Data sharing agreements
- Legal Reviews
- Data reviews and manual data retrieval
- De-identification process
- Limited cohort sizes
- Limited data utility

Expensive

Time Consuming

Non-Repeatable

IT/Expert Dependent



One-Off Projects

Bureaucracy

Limited Access

Medical Center – Data Owner

Restricted

Research Institutes – Insurance – IT – Pharmaceutical – Medical Device – Startups

MDClone

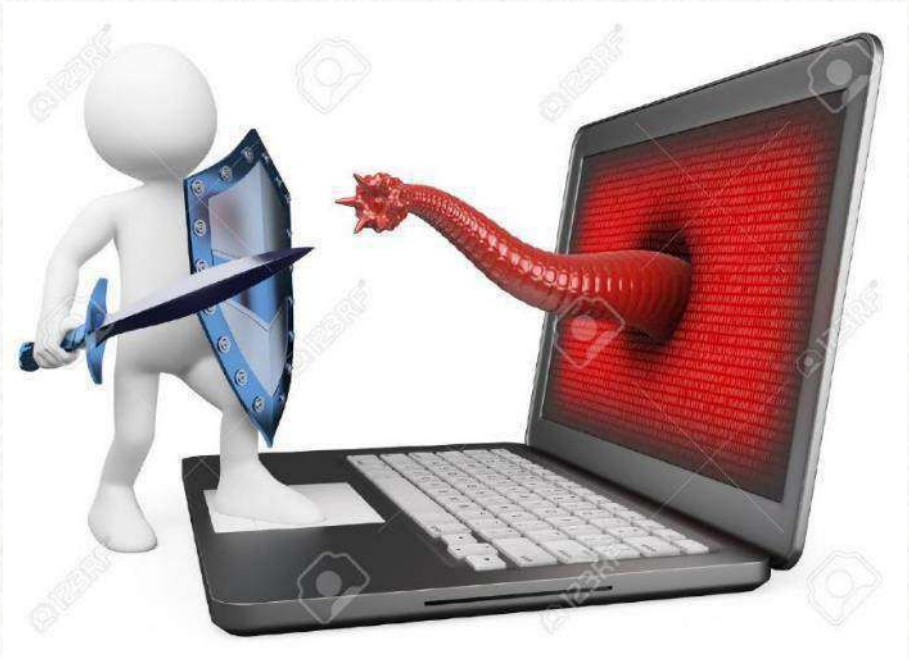
The Never Ending Battle

Anonymization

- ✓ Data Removing (Safe Harbor)
- ✓ Aggregation
- ✓ Masking & Blurring
- ✓ Noise
- ✓ Pseudonyms

De-Anonymization

- ✓ Outliers Analysis
- ✓ Ɛ Differential Privacy
- ✓ Probability Matching
- ✓ Data Mining
- ✓ Al



MDClone

DATA RE-IDENTIFICATION

Govt pulls dataset that jeopardised

Treat anonymised data as personal information: Pilgrim

By Paris Cowan
Nov 18 2015
12:17PM

SECURITY IS POWERED BY SC MAGAZINE

2 Comments

Privacy Commissioner warns of modern matching methods.

Timothy Pilgrim

Australian Privacy Commissioner Timothy Pilgrim has warned businesses that de-identifying datasets will not absolve them of the need to meet the stringent information protection demands of the Privacy Act.

Speaking to the International Association of Privacy Professionals today, Pilgrim said even anonymised datasets should be treated the same as personally identifiable information to future-proof organisations against increasingly sophisticated data matching efforts.

"The current challenge facing all organisations that handle large data sets is that data sets of

REUTERS

HOME BUSINESS MARKETS WORLD POLITICS TECH OPINION BREAKINGVIEWS MONEY LIFE PICTURES VIDEO

than your

problems

Only

Abstract
Methods
Results
Download
Authors
Citation
Data

MDClone

EUROPE

EUGDPR.org | The Regulation | The Process | More Resources | Our Partners

- Significant new fines: the greater of **20M euro** or **4% of annual global turnover**

- Consent harder to obtain:
  … **always have to be express/opt in**

WHITE & CASE People Services Our Thinking About Us Join Us

WHITE & CASE People Services Our Thinking About Us Join Us

agreed (although, as the European Parliament's press release clarified, before the GDPR becomes law a few formalities remain, which are not expected to be completed before February 2016).

Key changes and business impact

The GDPR brings significant changes to EU data protection law, many of which greatly impact businesses:

• **Significant new fines:** Maximum fines of the greater of **€20 million** or **4% of annual global turnover** *per* breach (a dramatic increase from the current typical maximum of less than €1 million).

• **Broader territorial reach:** Even businesses outside the EU will be subject to the GDPR, if they are offering goods or services to, or monitoring the behaviour of, EU residents (currently they are only caught if they operate data processing equipment in the EU). See further on this below.

• **Data breach reporting:** In most cases breaches will have to be reported within **72 hours**.

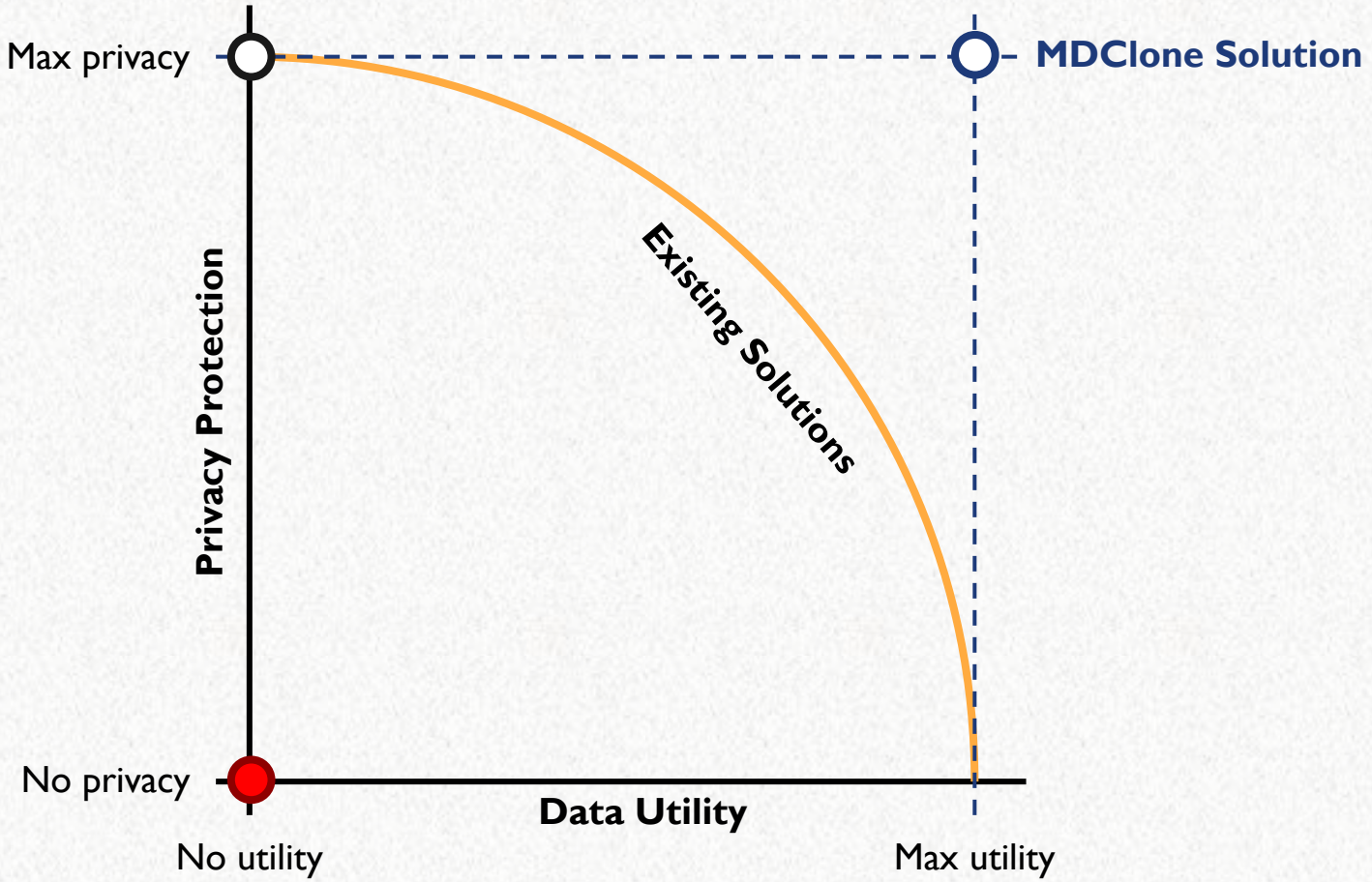
• **Stronger rights for individuals:** Individuals will have stronger rights against businesses, including the 'right to be forgotten'. Re-using existing data for new purposes will also become more difficult.

• **Consent harder to obtain:** Consent will always have to be express/opt-in (a 'clear affirmative action') whereas, under the existing regime, implied/opt-out consent is sometimes sufficient.

• **Data Protection Officers:** Most businesses that regularly monitor individuals, or regularly process sensitive personal data, will have to formally appoint an independent Data Protection Officer.

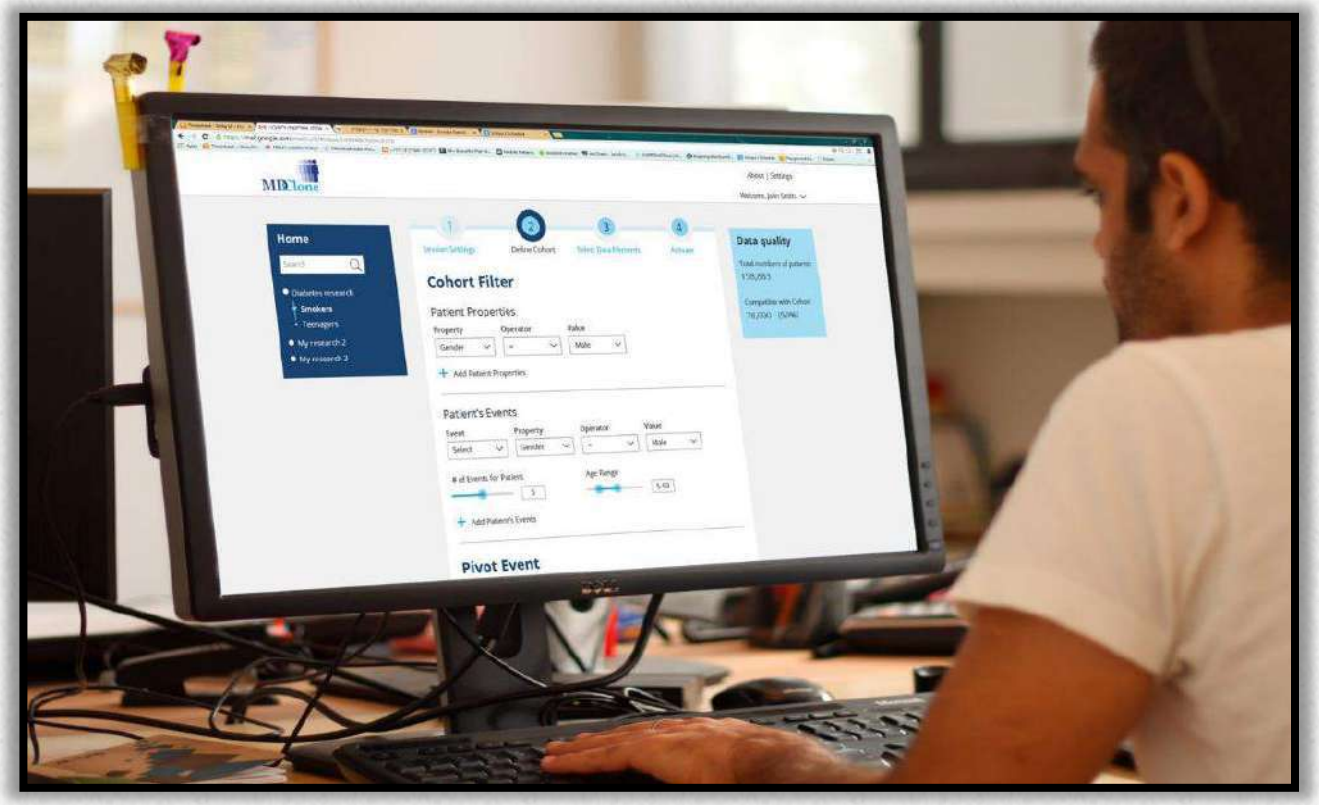
There is a need For a New Paradigm

A New Paradigm



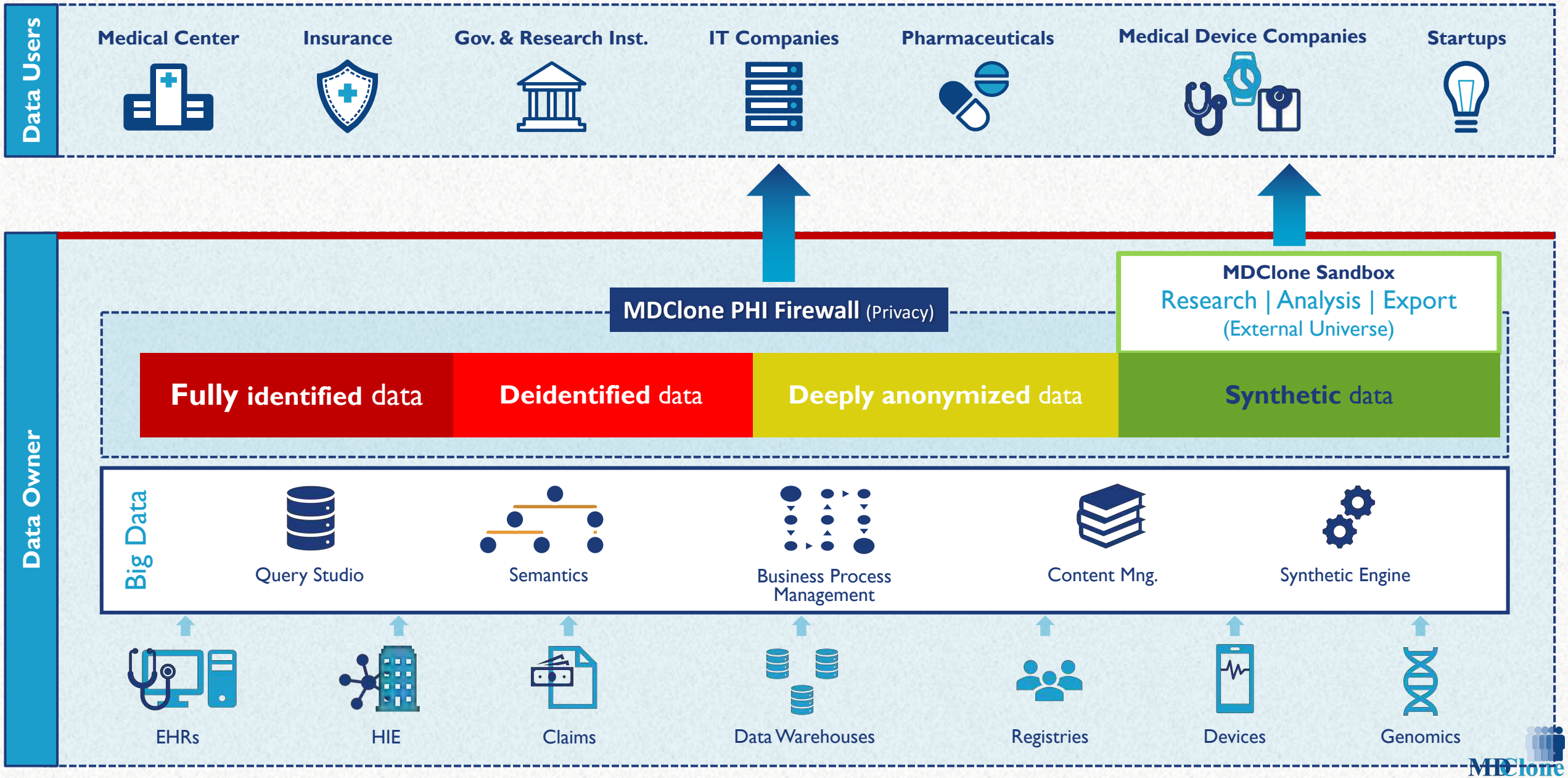
The new world of Synthetic data - MDClone Mission

Liberating healthcare data in **real time** with **zero-risk to patient privacy** to transform healthcare.



- ✓ **Any user** (no need for programming skills)
- ✓ Can ask **any question**
- ✓ With **zero time** to data & insights
- ✓ and **zero risk** to patient privacy
- ✓ With **no mediators**
- ✓ And **no bottlenecks**

MDClone Healthcare Data Platform



The Paradigm Differentiators

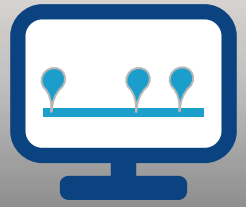
Time to Data & Insights

- No **mediators**
- **Direct** access
- No need for **technical skills**
- Familiar **terminologies**

Comprehensive Platform

- **AI-in-One** Product
- **Modern** Big-Data technology

Longitudinal Data Representation



- One **patient story**
- One **Mapping process**
- Connect data from **multiple sources**

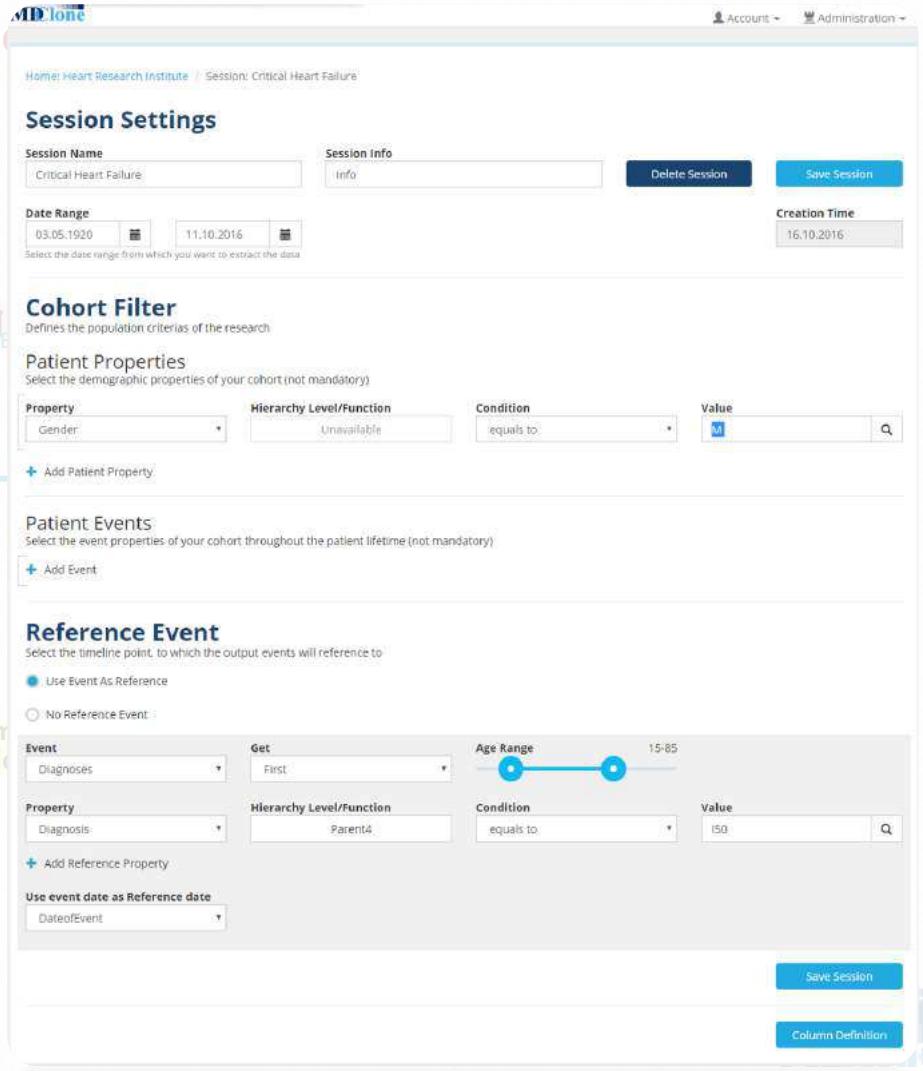
Zero-Risk to Patient Privacy

- Automatic **PHI Privacy Firewall**
- Eliminate **risk** & security **vulnerabilities**
- Maintain **data utility**

MDClone

Longitudinal Patient Representation

Outcome of **diabetes type 2** patients, treated for at least 1 year with **metformin** and **failed**, then received **oral hypoglycemic** add-on within 3 months of metformin failure.



MDClone Differentiators

Time to Data & Insights

- No **mediators**
- **Direct** access
- No need for **technical skills**
- Familiar **terminologies**

Comprehensive Platform

- **AI-in-One** Product
- **Modern** Big-Data technology

Longitudinal Data Representation

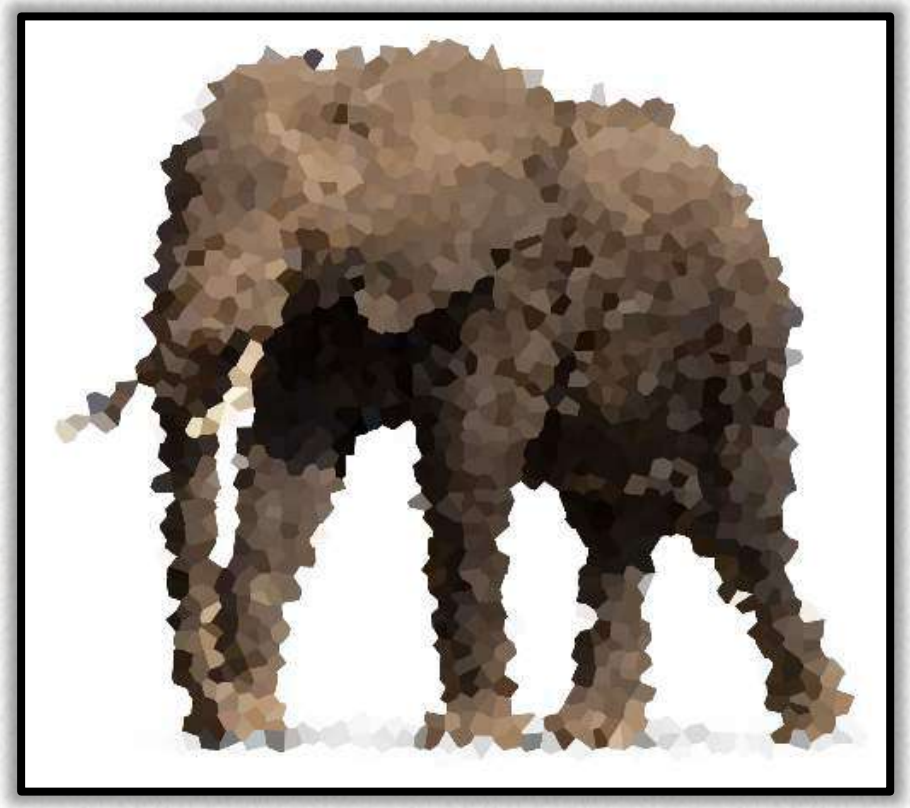
- One **patient story**
- One **Mapping process**
- Connect data from **multiple sources**

Zero-Risk to Patient Privacy

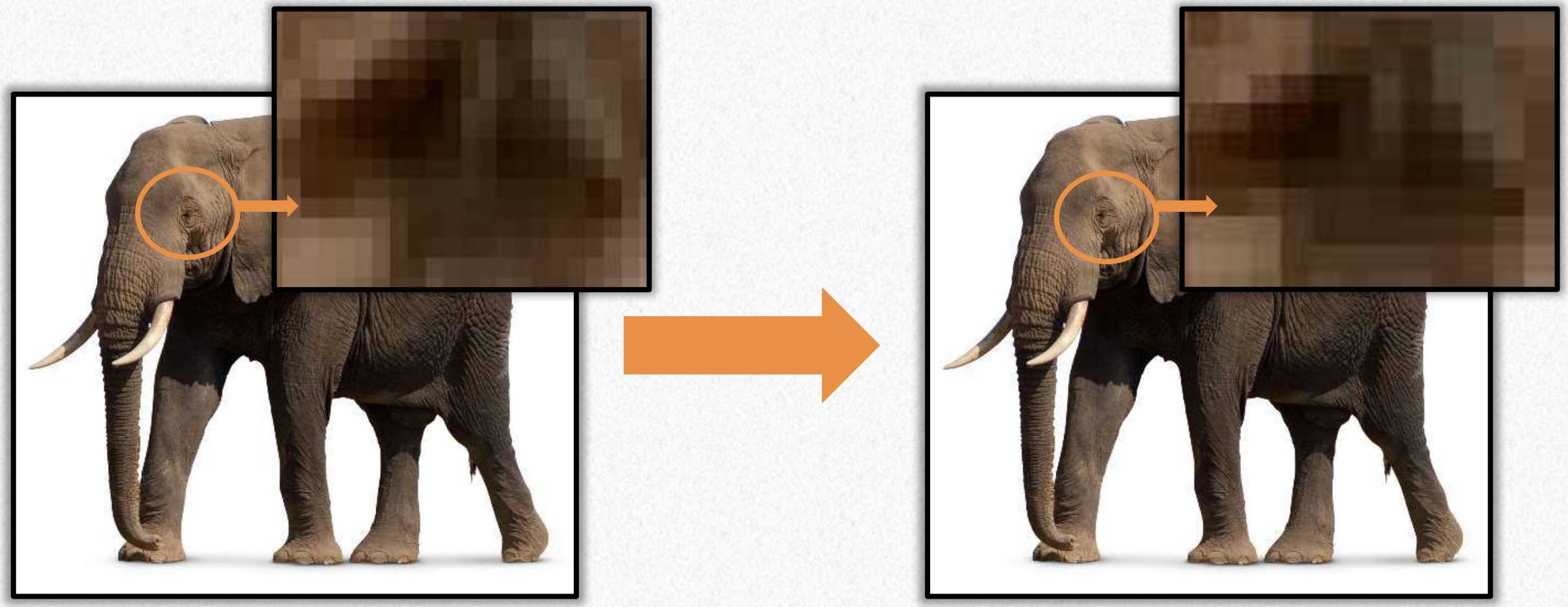
- Automatic **PHI Privacy Firewall**
- Eliminate **risk** & security **vulnerabilities**
- Maintain **data utility**

Existing Approaches to Protecting Privacy

- **Delete** information
- **Blur** numerical data
- **Hash** distinct values
- **Shift** date data
- **Block** information access
- **Aggregate** data (kill details)
- Data "**corruption**"



The MDClone Synthetic Approach



MDClone Synthetic Approach

Original Data

- Size of cohort: 950
- Average age: 37.5
- Standard deviation: 1.68
- Average weight 72.8
- …

Fully identifiable

Synthetic Data

- Size of cohort: 935
- Average age: 37.8
- Standard deviation: 1.71
- Average weight 72.3
- …

Zero-risk of patient identification

Risk factors for hypoglycemia-related hospitalization

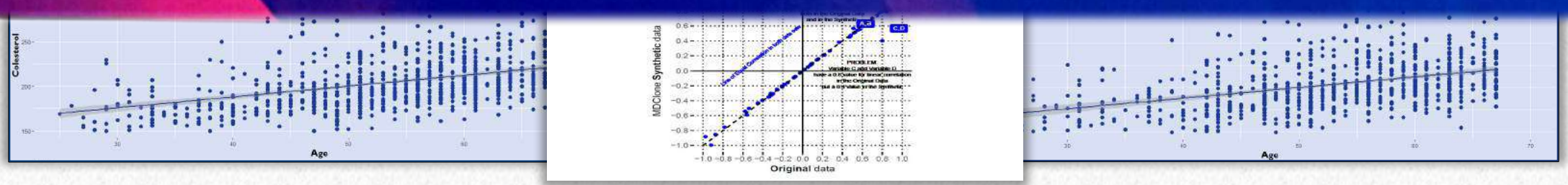
MDClone PHI Firewall (Privacy)

MDClone
Liberate Data. Transform Care.

רמב"ם
הקריה הרפואית לבריאות האדם

The Risk for Hypoglycemia Using Different Types of Basal Insulin: **Preliminary results from a retrospective study using MDClone platform**

Irit Hochberg, MD | Institute of Endocrinology, Diabetes and Metabolism, Rambam Health Care Campus, Haifa, Israel

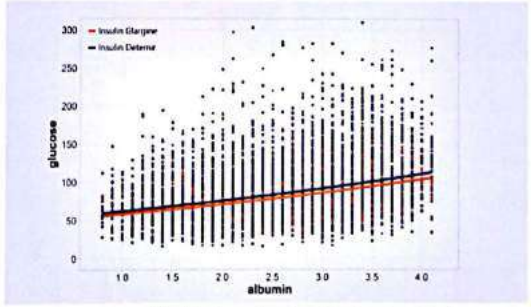


MDClone

Original Data

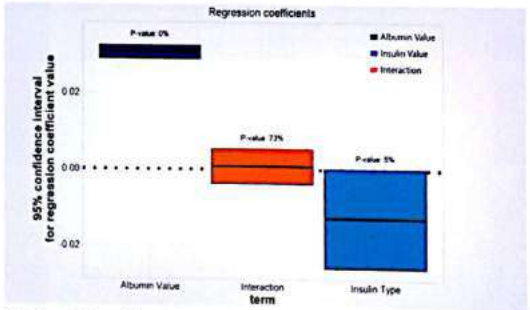
RESULTS

Insulin type influence is similar to all albumin values



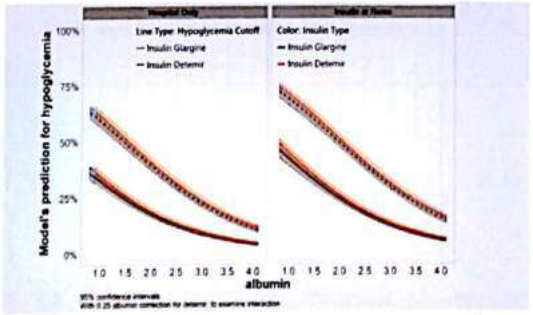
Lines are regression model with a boxcox transformation for glucose level as a function of albumin values, type of insulin and their interaction

Regression's interaction term is not significant



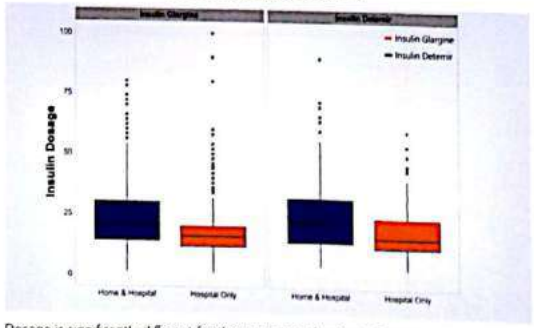
Plot shows 95% confidence for regression's coefficients. Significant coefficients should be different from 0. The interaction term is not significant, which implies that the drugs influence does not diverge in small albumin values

Logistic regression prediction for hypoglycemia



Two types of hypoglycemia definitions are used, and the model is separated for two types of patients. Insulin type was corrected to observe interaction effect.

Confounding influence of insulin dosage and insulin usage at home



Dosage is significantly different for the two types of patients. Better model is produced by correcting with home usage than by the patients dose level

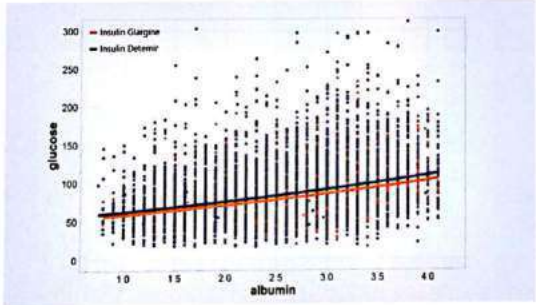
CONCLUSIONS

Hypoalbuminemia is a strong risk factor for hypoglycemia, regardless of the type of insulin used. Insulin type influence is similar to all albumin values.

Synthetic Data

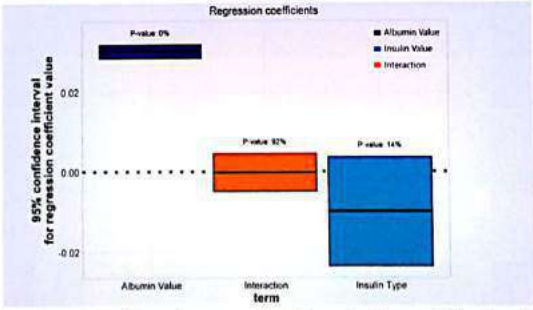
RESULTS

Insulin type influence is similar to all albumin values



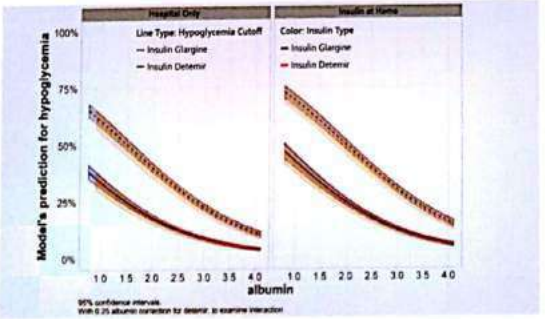
Lines are regression model with a boxcox transformation for glucose level as a function of albumin values, type of insulin and their interaction

Regression's interaction term is not significant



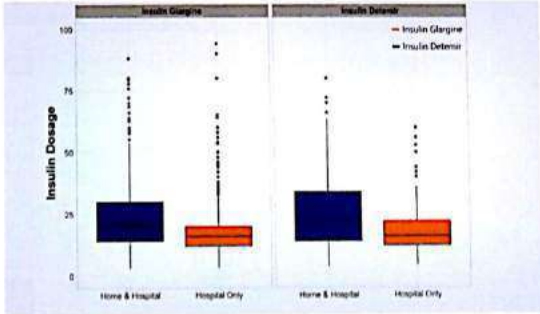
Plot shows 95% confidence for regression's coefficients. Significant coefficients should be different from 0. The interaction term is not significant, which implies that the drugs influence does not diverge in small albumin values

Logistic regression prediction for hypoglycemia



Two types of hypoglycemia definitions are used, and the model is separated for two types of patients. Insulin type was corrected to observe interaction effect.

Confounding influence of insulin dosage and insulin usage at home

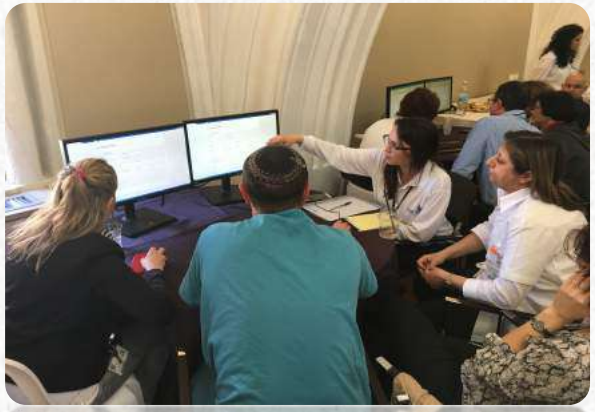


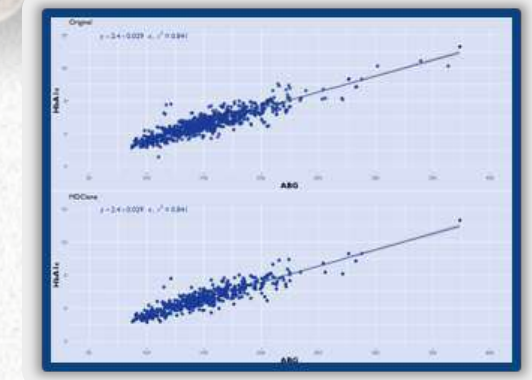
Dosage is significantly different for the two types of patients. Better model is produced by correcting with home usage than by the patients dose level

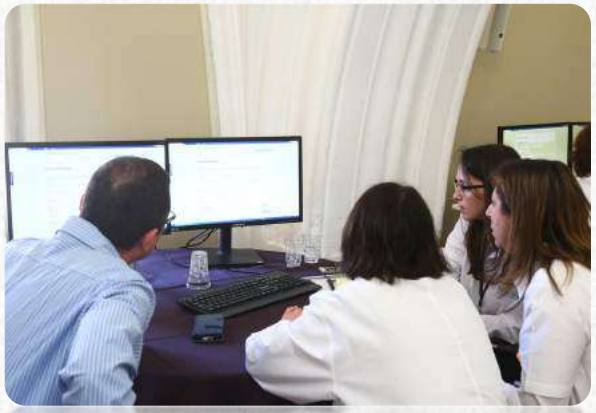
CONCLUSIONS

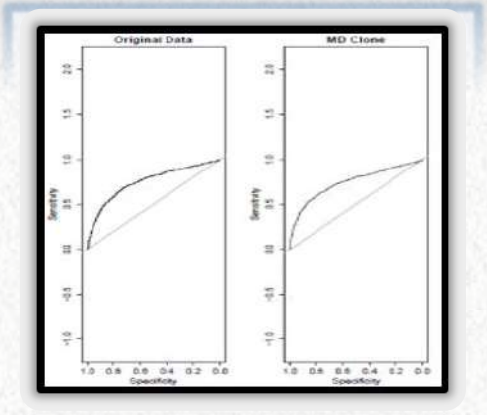
Hypoalbuminemia is a strong risk factor for hypoglycemia, regardless of the type of insulin used. Insulin type influence is similar to all albumin values.

Innovation & Field Validation







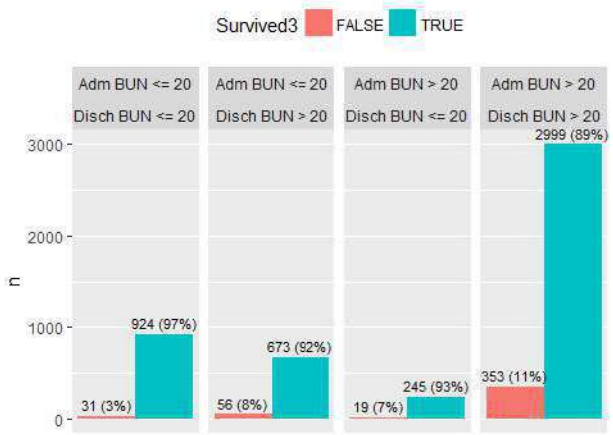


MDClone

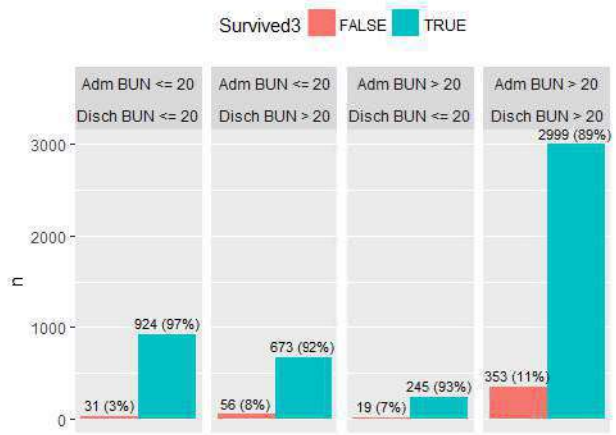
Mortality Rate and BUN Variation Between Admission and Discharge

Prof. Zaher Azam

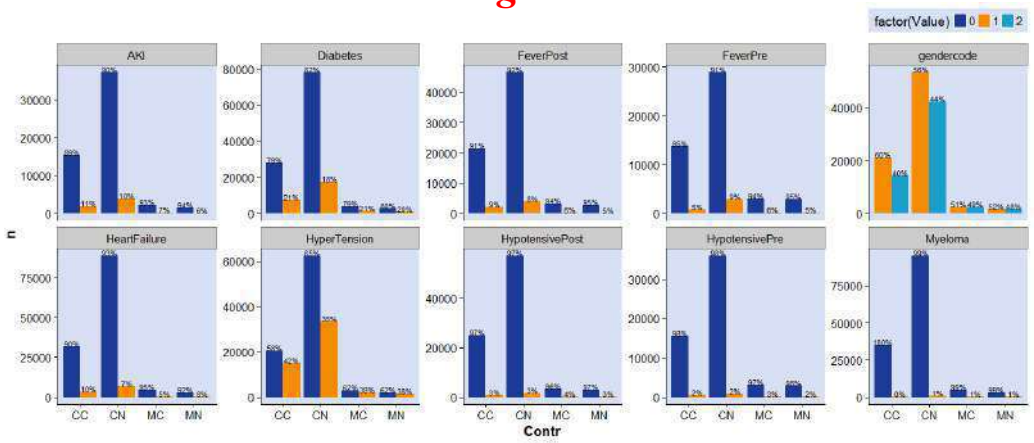
Original



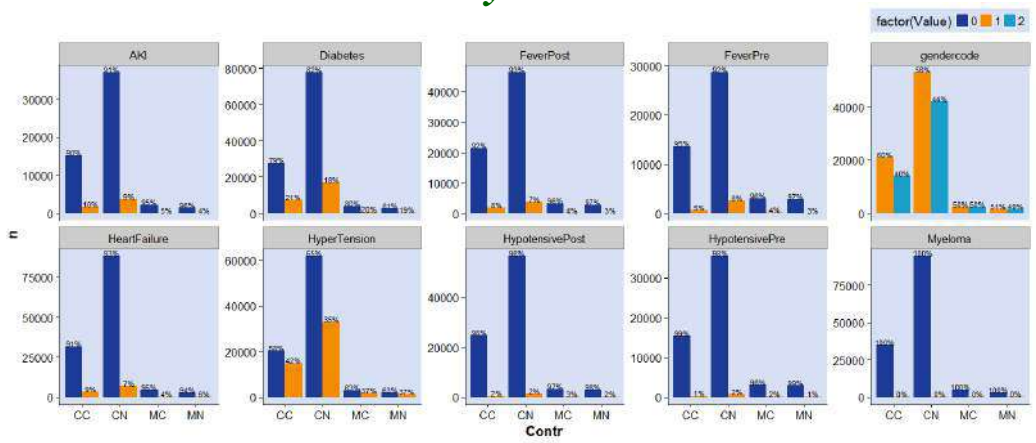
Synthetic



Original



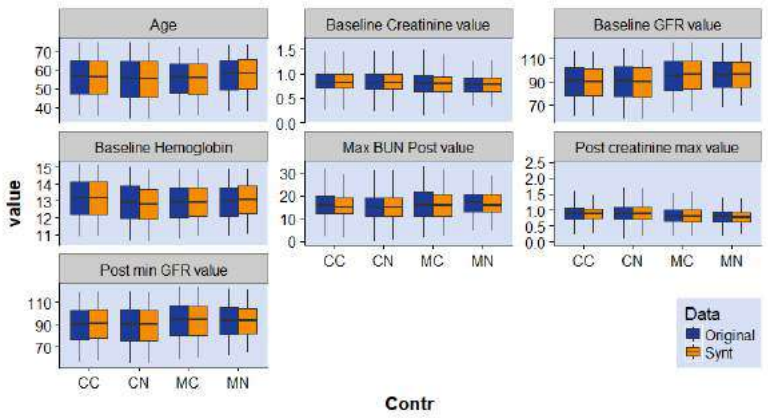
Synthetic



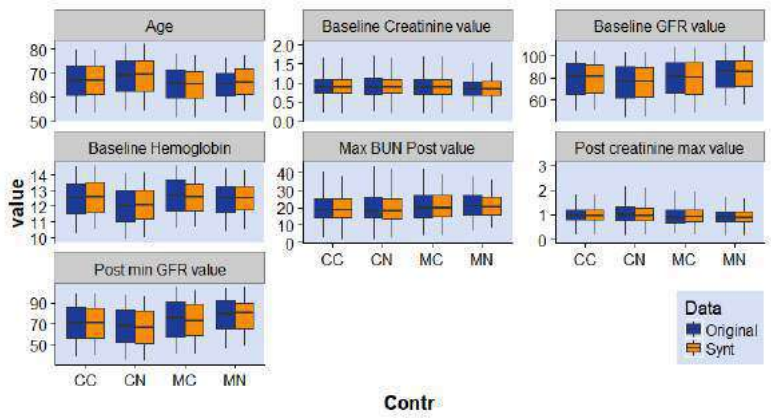
The effect of MRI contrast media on renal function

Dr. Yuri Gorelik

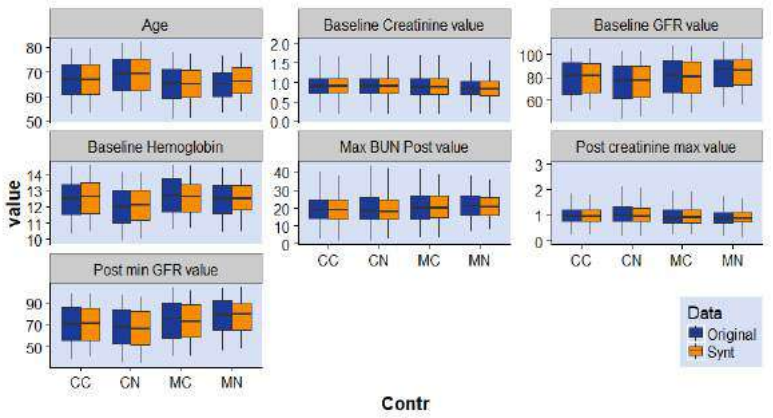
All Data



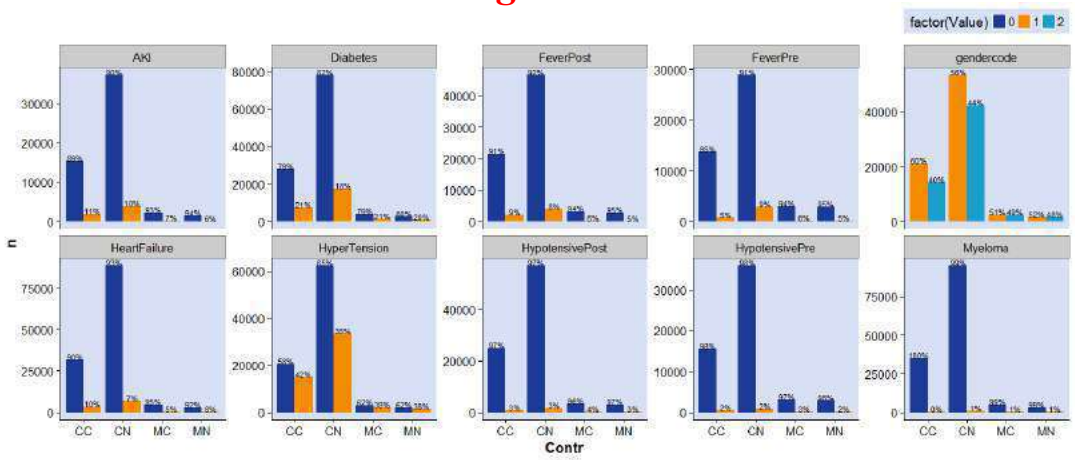
Diabetes = True



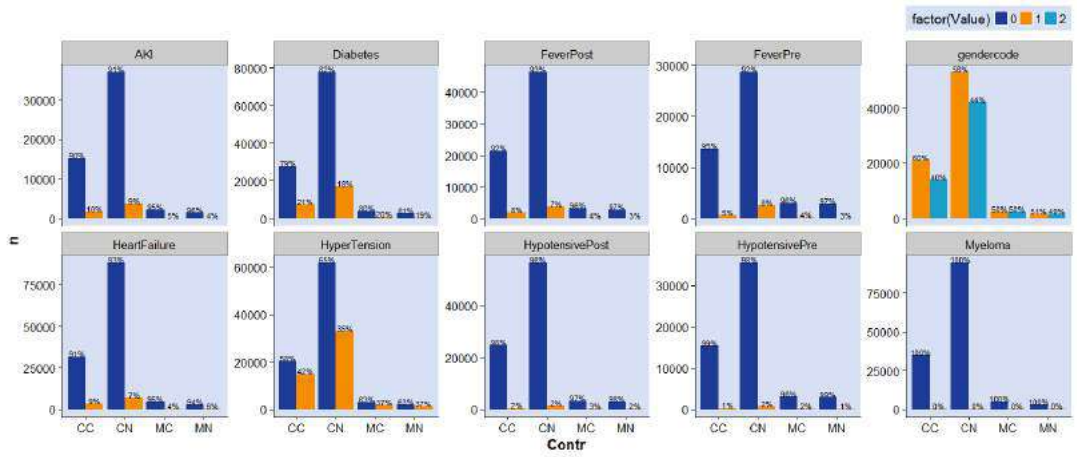
Diabetes = True



Original



Synthetic



Customer Base

- **Rambam** Health Care Campus
- **Clalit**, Israel's largest HMO and the second largest HMO in the world with 4 million patients
- **Maccabi**, Israel's second largest HMO with 2 million patients
- **Assuta**, Israel's largest and leading private medical center
- **Sheba** – the biggest hospital in the middle east

Altogether, more than 6 million Israeli lives, 80% of the Israeli market

RAMBAM Health Care Campus

HOSPITAL ASSUTA RAISING HEALTH STANDARDS

Sheba - Academic Medical Center Hospital

CLALIT over 100 years The Best For Your Family

Maccabi Healthcare Services

One of the biggest pharma's in the world

MDClone

The Power of Data Liberation

Liberate **Data**

Liberate **Research**

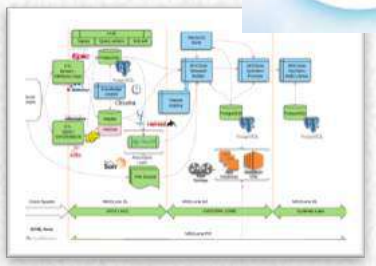
Liberate **Quality @ POC**

Liberate **Population Mng.**

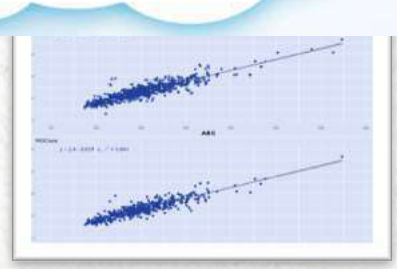
Liberate **Clinical Trials**

Liberate **Precision Medicine**

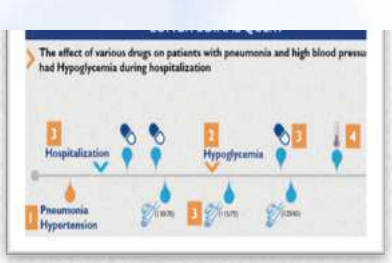
MDClone Po



Advanced Big Data Platform



Synthetic Data Privacy Zero-Risk



Longitudinal Representation

Time-2-Data

MDClone Global Grid

MDClone

Thank You!!!

Boaz Gur-Lavie

Boaz@mdclone.com

+972-54-4640119