



## ΠΑΡΟΥΣΙΑΣΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

# «ΠΡΟΓΝΩΣΤΙΚΗ ΑΝΑΛΥΣΗ ΜΑΖΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΥΓΕΙΑΣ ΤΩΝ ΕΝΟΠΛΩΝ ΔΥΝΑΜΕΩΝ, ΣΤΗΝ ΕΠΟΧΗ ΤΟΥ ΓΕΝΙΚΟΥ ΚΑΝΟΝΙΣΜΟΥ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ (ΓΚΠΔ)»



**ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:** ΚΩΝΣΤΑΝΤΙΝΟΣ ΨΑΝΝΗΣ, [kpsannis@uom.edu.gr](mailto:kpsannis@uom.edu.gr)  
**ΜΕΤΑΠΤΥΧΙΑΚΟΣ ΦΟΙΤΗΤΗΣ:** ΓΡΗΓΟΡΙΟΣ ΓΚΟΥΤΖΕΛΗΣ, [mli19011@uom.edu.gr](mailto:mli19011@uom.edu.gr)

ΦΕΒΡΟΥΑΡΙΟΣ 2021

# ΠΕΡΙΓΡΑΜΜΑ ΠΑΡΟΥΣΙΑΣΗΣ



## Σκοπός

Η εξαγωγή χρήσιμων συμπερασμάτων σχετικών με το θέμα.



## Θεωρητικό Μέρος - Τεχνολογίες

Σύγκλιση Τεχνολογιών.  
Ανάλυση Δεδομένων.  
Περιορισμοί – ΓΚΠΔ.



## Βιβλιογραφική Επισκόπηση

Εντοπίστηκε ερευνητικό κενό.  
Αναφορές σε 30 σύγχρονα άρθρα,  
χωρισμένα σε 5 κατηγορίες.



## Σενάρια Ανάλυσης

6 σενάρια Προγνωστικής Ανάλυσης με  
Ανώνυμα Δεδομένα  
(Ταξινόμηση - Παλινδρόμηση).



## Μεθοδολογία

Pearl-Growing (βιβλιογραφία)  
Ανάλυση στην πλατφόρμα ML Knowi,  
με δεδομένα data.world.



## Συμπεράσματα - Προτάσεις

Εφαρμόσιμη ή όχι η ιδέα;  
Ανοικτά θέματα για μέλλον...

## Η ΙΔΕΑ ΤΗΣ ΕΡΕΥΝΑΣ

- Η Προγνωστική Ανάλυση σε Μαζικά Δεδομένα, ασφαλώς δεν είναι μια καινούργια ιδέα. Χρησιμοποιείται εδώ και αρκετές δεκαετίες, καθώς με τη βοήθεια μεθόδων Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης μπορούν να πραγματοποιηθούν ασφαλείς προγνώσεις για καταστάσεις ή αριθμητικές τιμές, των οποίων η έγκαιρη και αξιόπιστη πρόγνωση ενδεχομένως να είναι πολύτιμη.
- Στην ιδέα αυτή ακριβώς της έγκαιρης, αξιόπιστης και πολύτιμης πρόγνωσης βασίστηκε και η παρούσα έρευνα, εστιάζοντας στα δεδομένα υγείας, προσωπικά και συνάμα ευαίσθητα, του προσωπικού των ΕΔ, αφού εντόπισε αντίστοιχο ερευνητικό κενό κατά την αναζήτηση της βιβλιογραφίας, γεγονός που καθιστά το προς διερεύνηση ζήτημα, σημαντικό και σχετικά πρωτότυπο.

## PEARL-GROWING

- Η μέθοδος «pearl-growing» χρησιμοποιείται συχνά στην αναζήτηση και χρήση άρθρων και εργασιών, σχετικών με το εκάστοτε, προς μελέτη, θέμα.
- Αναζητήθηκαν αναφορές, τόσο προς τα πίσω (backwards), δηλαδή από παραπομπές που βρέθηκαν σε ένα συγκεκριμένο άρθρο, όσο και προς τα εμπρός (forwards), δηλαδή βρίσκοντας ποιοι άλλοι ερευνητές παραπέμπουν στο προς μελέτη άρθρο.
- Νέοι όροι που εντοπίστηκαν σε κάποια αντιπροσωπευτικά άρθρα και σχετίζονταν με το κεντρικό θέμα της εργασίας, αναζητήθηκαν στο Διαδίκτυο («ScienceDirect», «ResearchGate» και «Frontiers in Medicine»), για ανεύρεση επιπλέον περισσότερο εξειδικευμένης βιβλιογραφίας.

# ΒΑΣΙΚΕΣ ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

Χρησιμοποιήθηκαν 30 άρθρα, σε 5 βασικές ομάδες και μελετήθηκαν ως προς 8 πεδία:

## ΟΜΑΔΕΣ ΑΡΘΡΩΝ

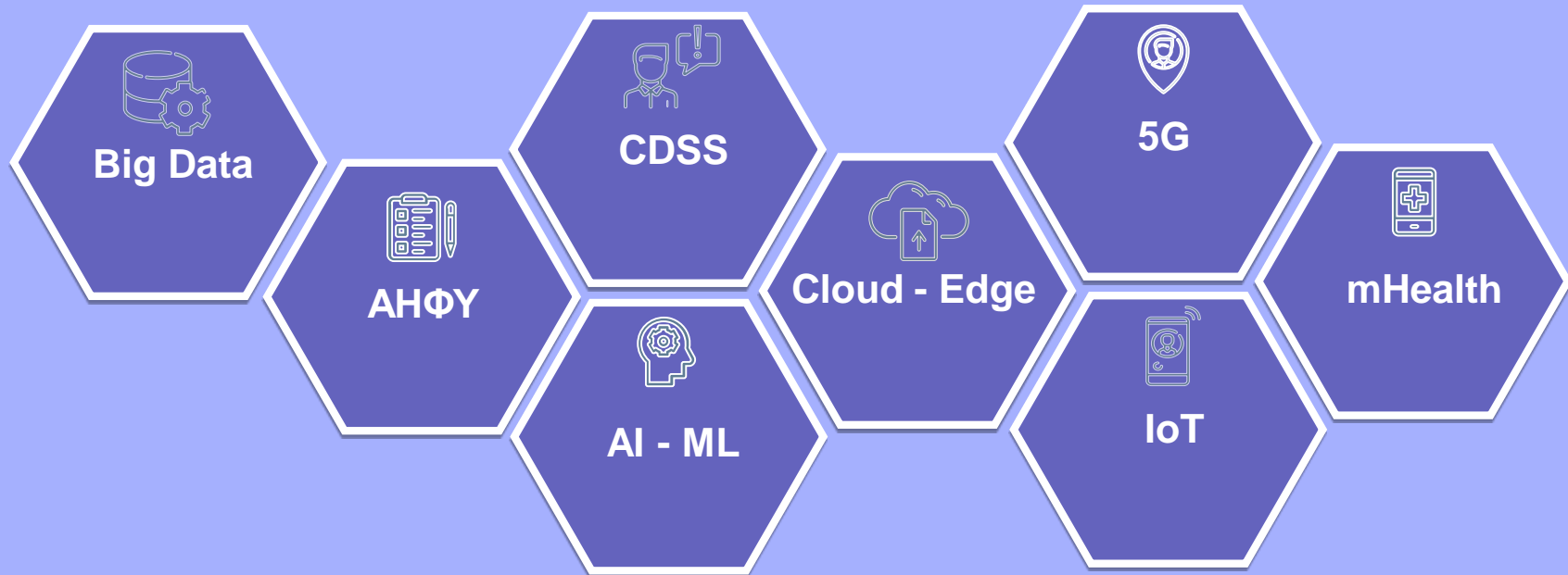
- 1. ANALYTICS
- 2. BIG DATA
- 3. PRECISION MEDICINE
- 4. IoT
- 5. MACHINE LEARNING

#1	#2	#3	#4	#5	#6	#7	#8	#9	#10

## ΠΕΔΙΑ ΜΕΛΕΤΗΣ

- 1. Συγγραφείς - Έτος
- 2. Τίτλος Άρθρου
- 3. Ερευνητικός Σκοπός
- 4. Μεθοδολογία
- 5. Κυριότερα Ευρήματα
- 6. Θέματα που Ερευνήθηκαν
- 7. Προτεινόμενες Λύσεις
- 8. Θέματα Ανοικτά για Μελλοντική Έρευνα

# ΣΥΓΚΛΙΣΗ ΤΕΧΝΟΛΟΓΙΩΝ





# TA 5Vs ΤΩΝ BIG DATA

- Volume: μεγάλος όγκος
- Variety: ποικιλομορφία
- Velocity: ταχύτητα δημιουργίας και ανάλυσης
- Variability: μεταβλητότητας (τρόπος παροχής θεραπειάς και τρόπος συλλογής των δεδομένων υγείας, διαφέρει σε τόπο και χρόνο).
- Value: αν πρόκειται να επενδύσει κανείς στην υποδομή που απαιτείται, για τη συλλογή και την ερμηνεία δεδομένων ευρείας κλίμακας, είναι σημαντικό να διασφαλίσει ότι οι πληροφορίες που δημιουργούνται βασίζονται σε ακριβή δεδομένα και οδηγούν σε μετρήσιμες βελτιώσεις.

(Anil Jain (2016), αντιπρόεδρος της IBM Watson Health)



(Πηγή: Panimalar et al., 2017)

## BIG DATA

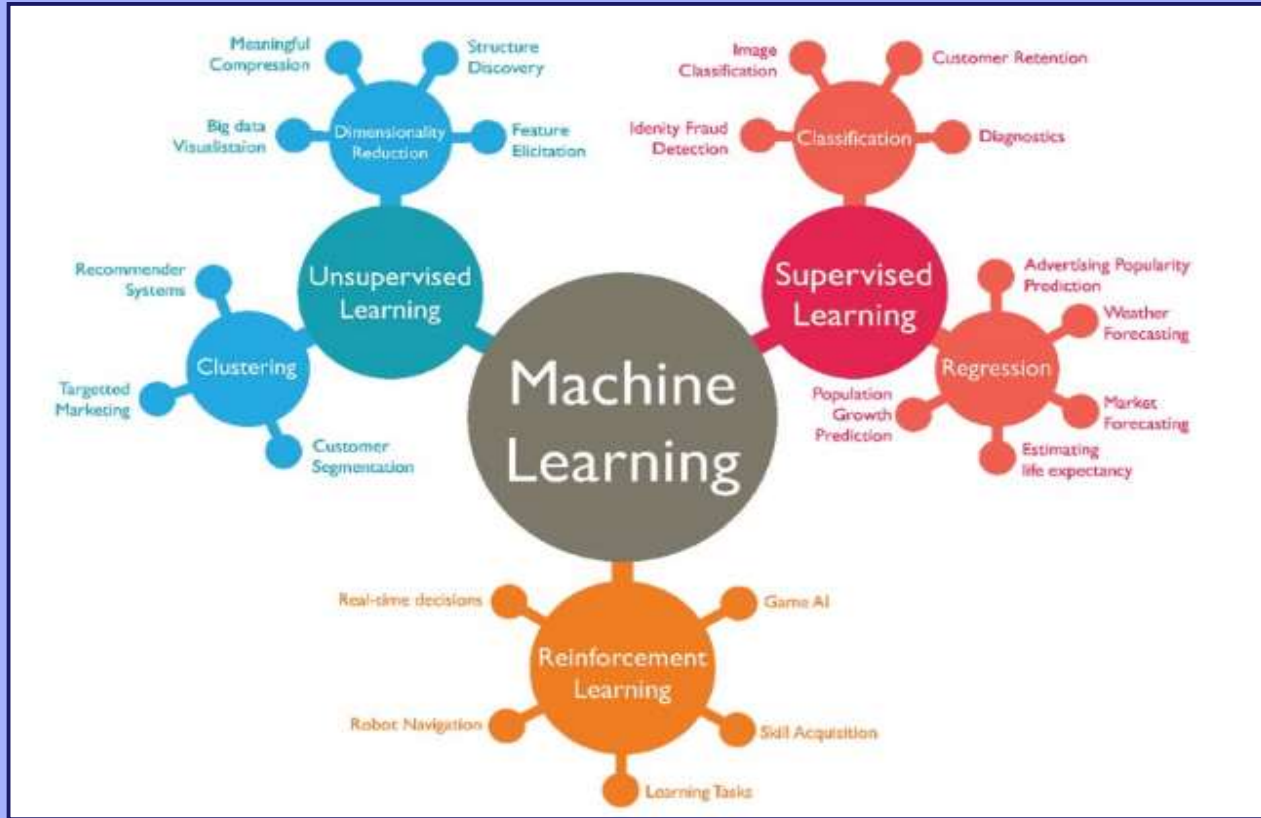
**ΜΔ: 33 zettabytes\* το 2018**  
**175 zettabytes το 2025**  
**ΜΔΥ: 2314 exabytes το 2020**

Σε 5 exabytes ανέρχεται το σύνολο όλων των λέξεων, που έχουν ειπωθεί από τον άνθρωπο, από τη στιγμή της εμφάνισής του στη γη ως σήμερα (Cottle et al., 2013).

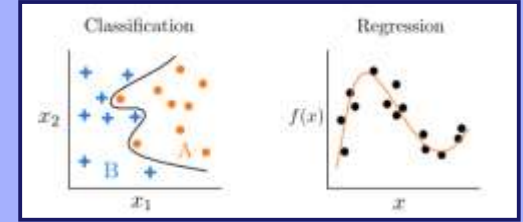
\*1 ZB =  $10^{21}$  bytes = 1024 EB



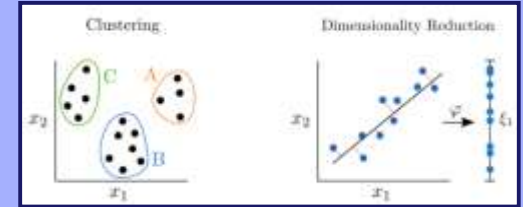
# KΑΤΗΓΟΡΙΕΣ MACHINE LEARNING



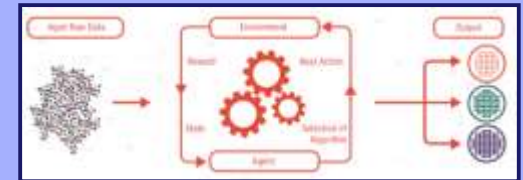
(Πηγή: nowmag.gr)



(Πηγή: Beck and Kurz, 2020)



(Πηγή: Beck and Kurz, 2020)

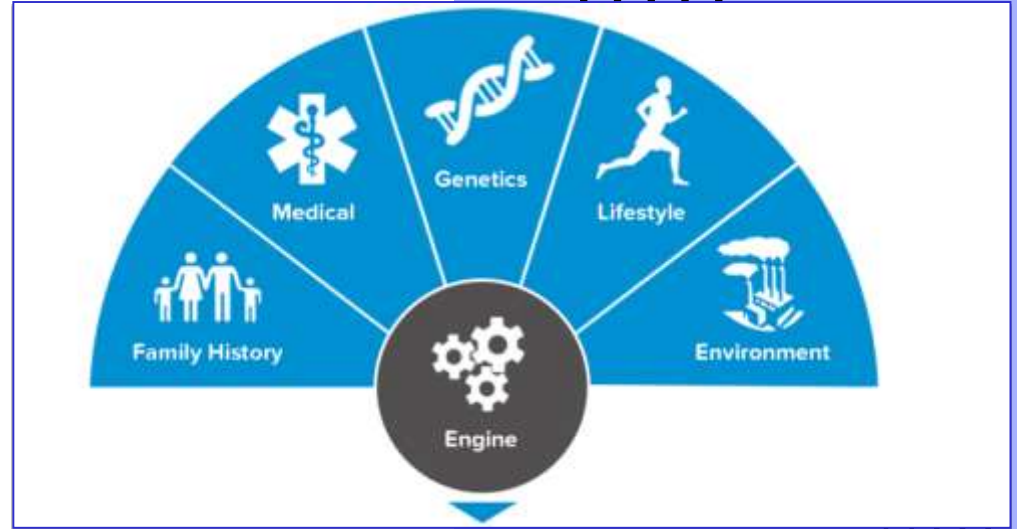


(Πηγή: Van Loon, 2020)

# BIG DATA ANALYTICS

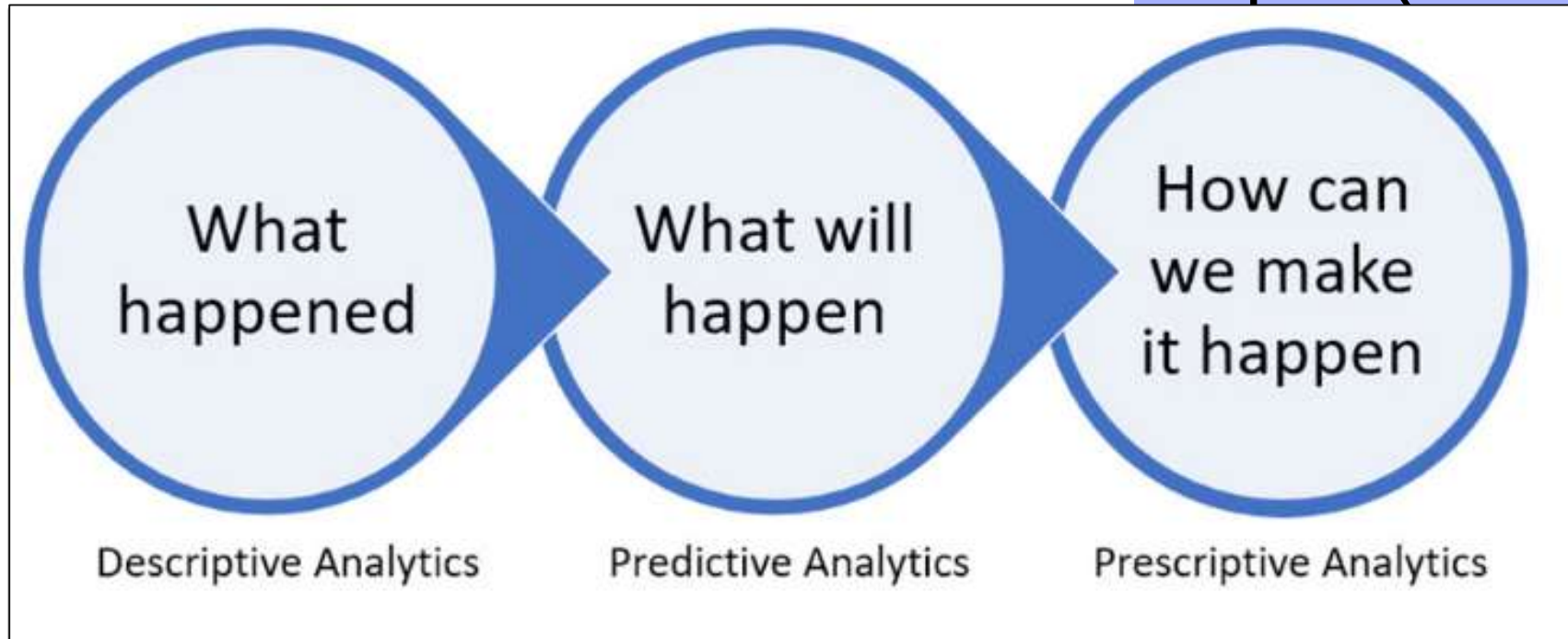
Υπάρχουν τουλάχιστον 5 τομείς πηγών δεδομένων που μπορούν να επηρεάσουν την ανθρώπινη υγεία και πρέπει να αναλυθούν, ώστε να οδηγήσουν σε χρήσιμα συμπεράσματα πρόγνωσης ή διάγνωσης της νόσου:

1. Ιατρικής ή υγειονομικής περίθαλψης
2. Γενετικής, γονιδιωματικής και ολοκληρωμένης βιολογίας
3. Συμπεριφοράς
4. Κοινωνικών παραγόντων
5. Φυσικού και κοινωνικού περιβάλλοντος.



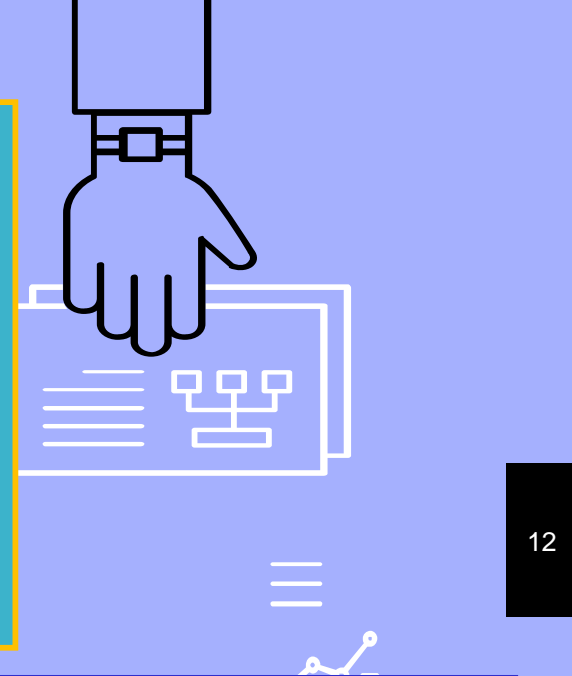
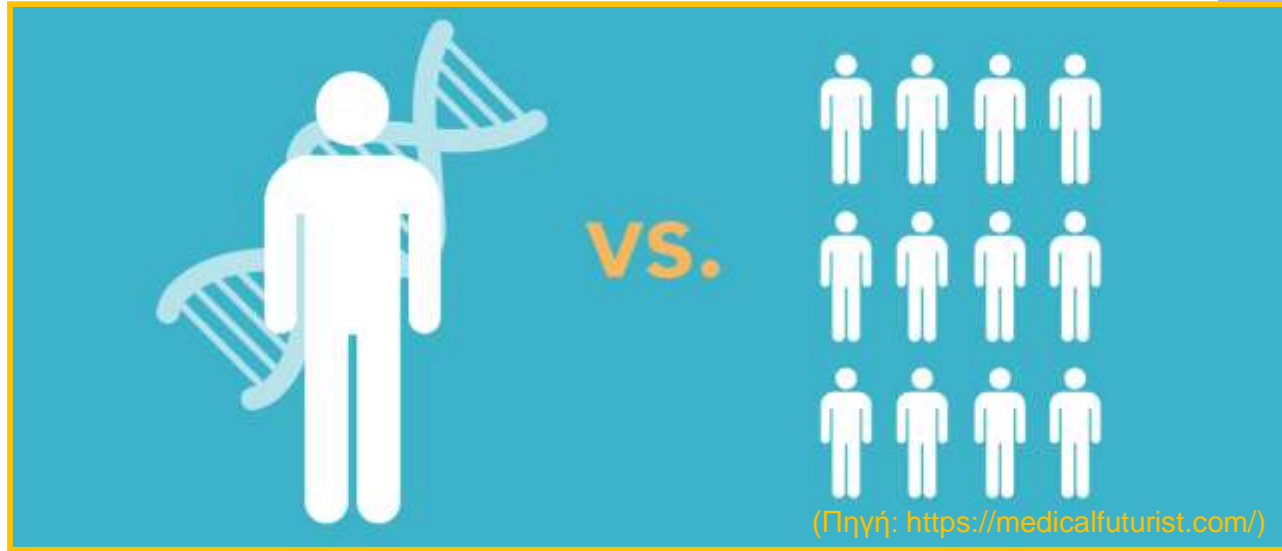
(Πηγή: <https://hitconsultant.net/>)

# ΚΑΤΗΓΟΡΙΕΣ ANALYTICS ΣΤΗΝ ΥΓΕΙΑ



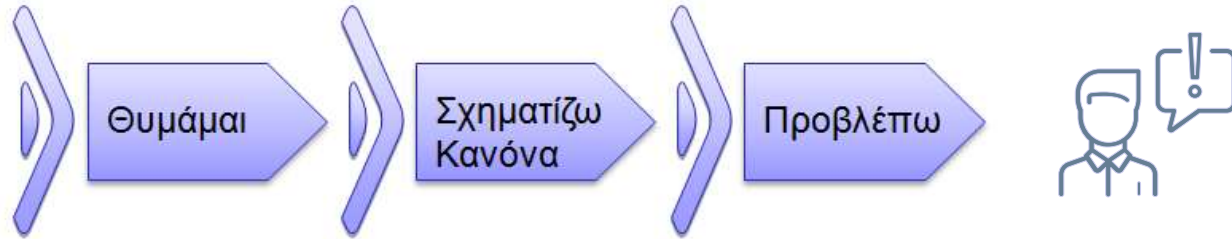
(Πηγή: vovia.com)

# PRECISION MEDICINE



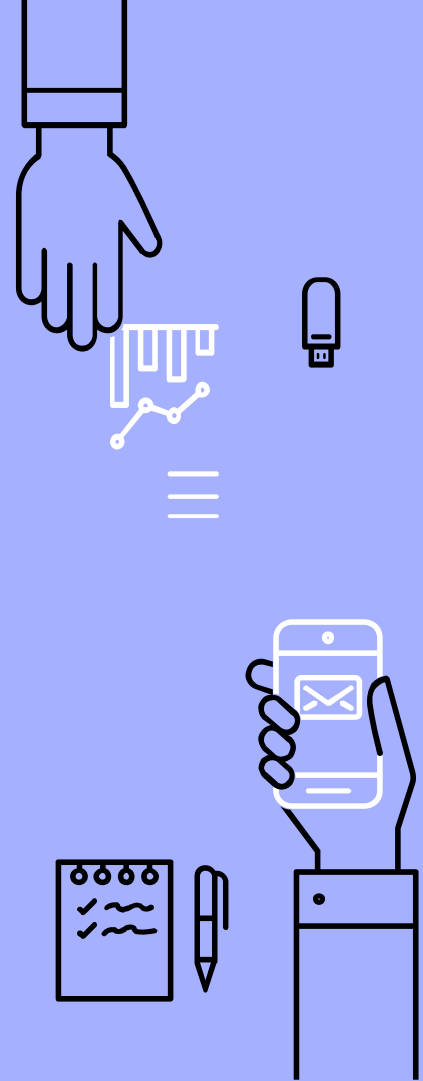
Το Εθνικό Συμβούλιο Έρευνας (National Research Council) υιοθέτησε τον ορισμό της Ιατρικής Ακριβείας το 2008, ορίζοντάς την ως την: «Προσαρμογή της ιατρικής περίθαλψης στα ιδιαίτερα χαρακτηριστικά κάθε ασθενούς, προς ταξινόμηση των ασθενών σε υποπληθυσμούς, που διαφέρουν ως προς την ευαισθησία τους σε μια συγκεκριμένη ασθένεια ή την απόκρισή τους σε μια συγκεκριμένη θεραπεία. Οι προληπτικές ή θεραπευτικές παρεμβάσεις μπορούν στη συνέχεια να επικεντρωθούν σε εκείνους που θα ωφεληθούν, εξοικονομώντας έξοδα και παρενέργειες για εκείνους που δεν θα το κάνουν».

# ΔΙΑΔΙΚΑΣΙΑ ΛΗΨΗΣ ΑΠΟΦΑΣΗΣ ML

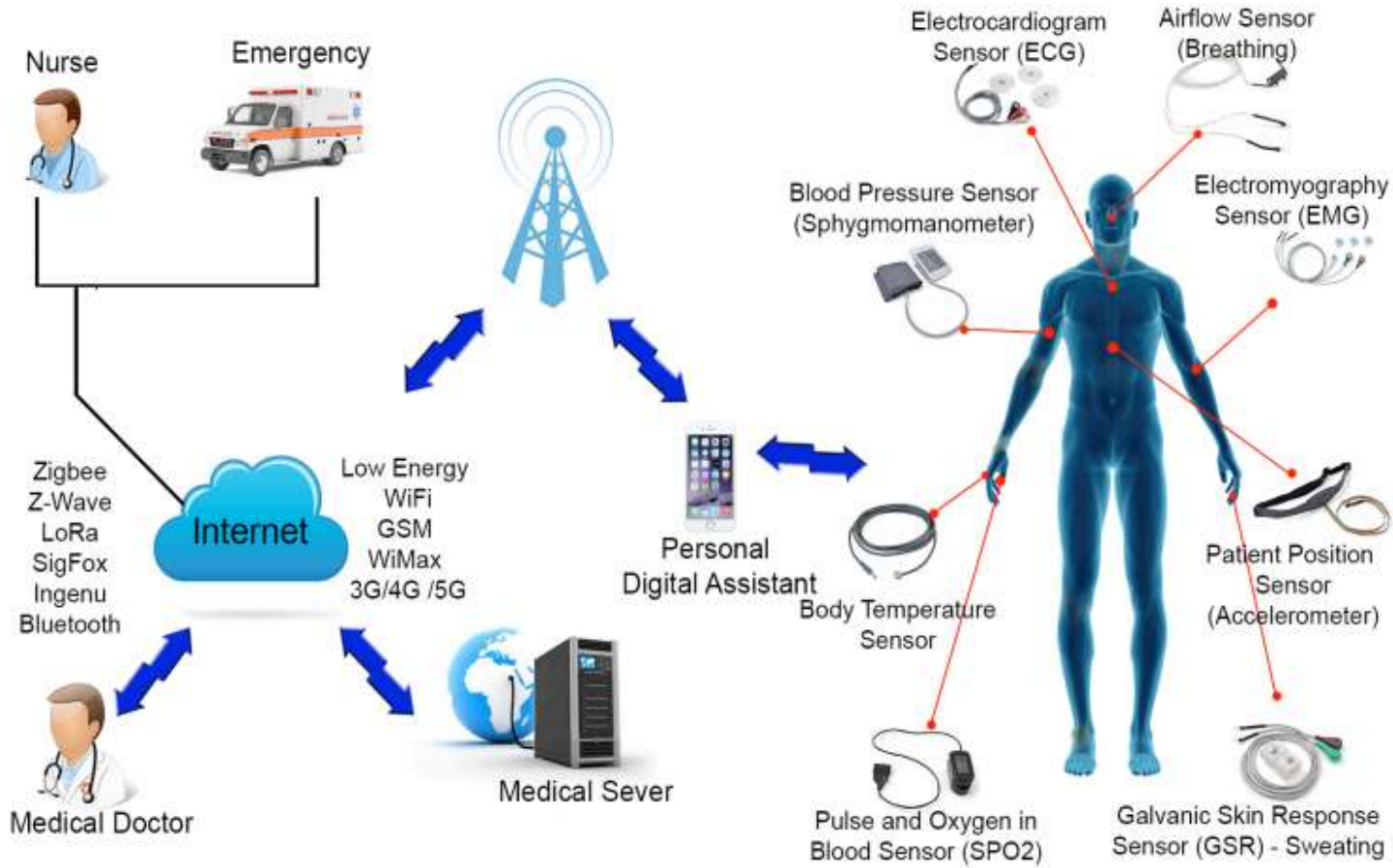


Έτσι περίπου λειτουργεί και το μοντέλο Επιτηρούμενης ML:

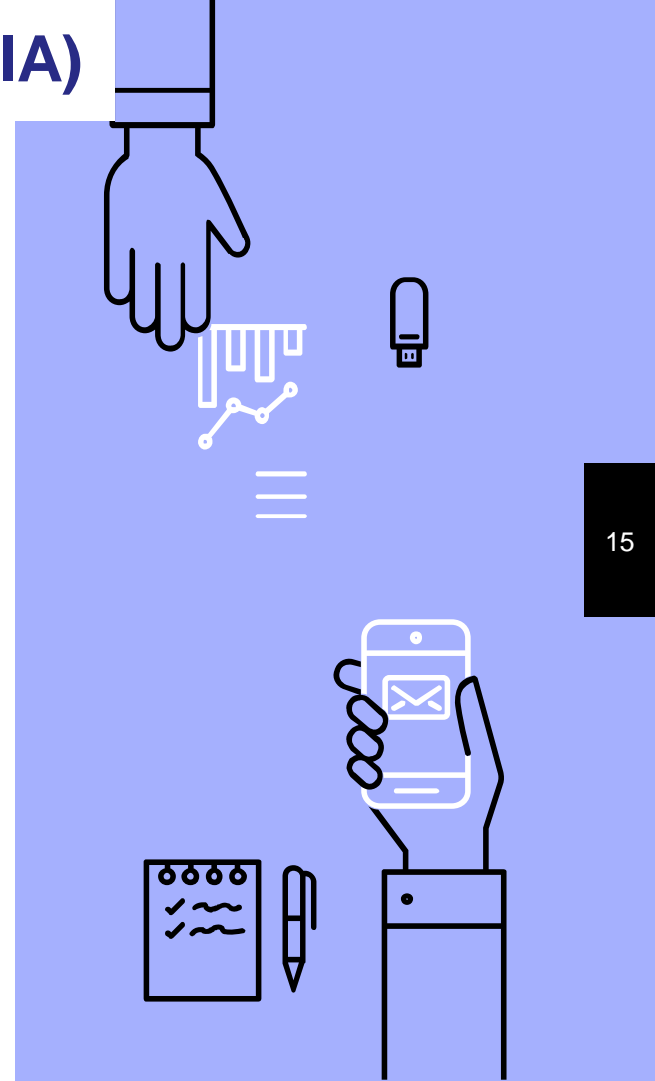
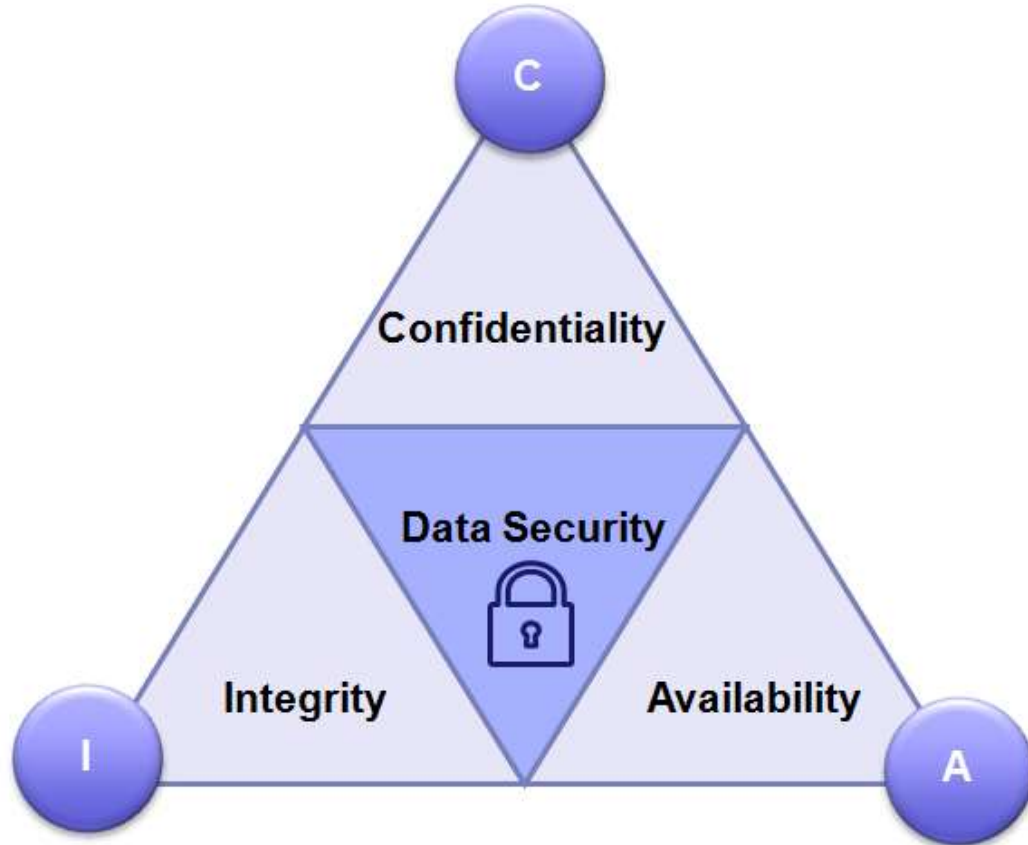
- 1.«**Θυμάται**» ένα σύνολο δεδομένων όπως συμπτώματα, ιατρικές μετρήσεις, αποτελέσματα απεικονιστικών εξετάσεων κ.λπ.,
- 2.«**Σχηματίζει κανόνα ή μοντέλο**», σύμφωνα με τον οποίο, όταν ο ασθενής παρουσιάζει τα  $y$  συμπτώματα και έχει τα  $z$  αποτελέσματα εξετάσεων,
- 3.«**Προβλέπει**» ότι (ο ασθενής) πάσχει από την  $x$  ασθένεια.



# ΤΥΠΙΚΗ ΑΠΕΙΚΟΝΙΣΗ WBAN

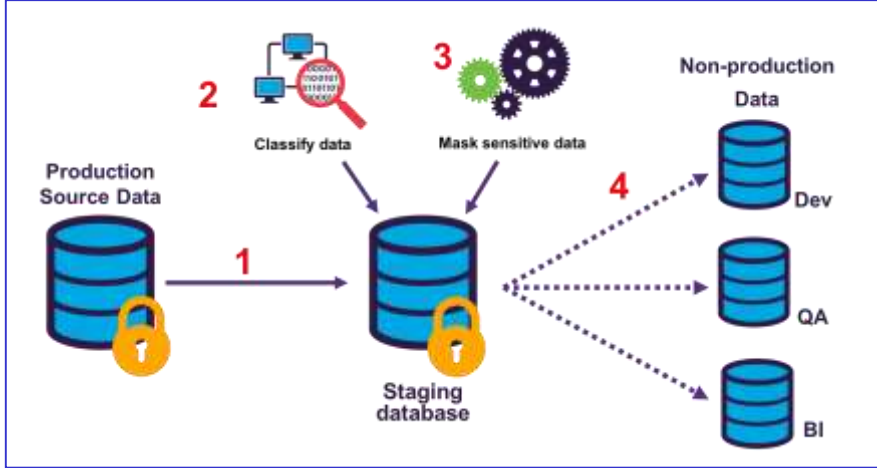


# ΤΡΙΠΤΥΧΟ ΑΣΦΑΛΕΙΑΣ ΔΕΔΟΜΕΝΩΝ (CIA)

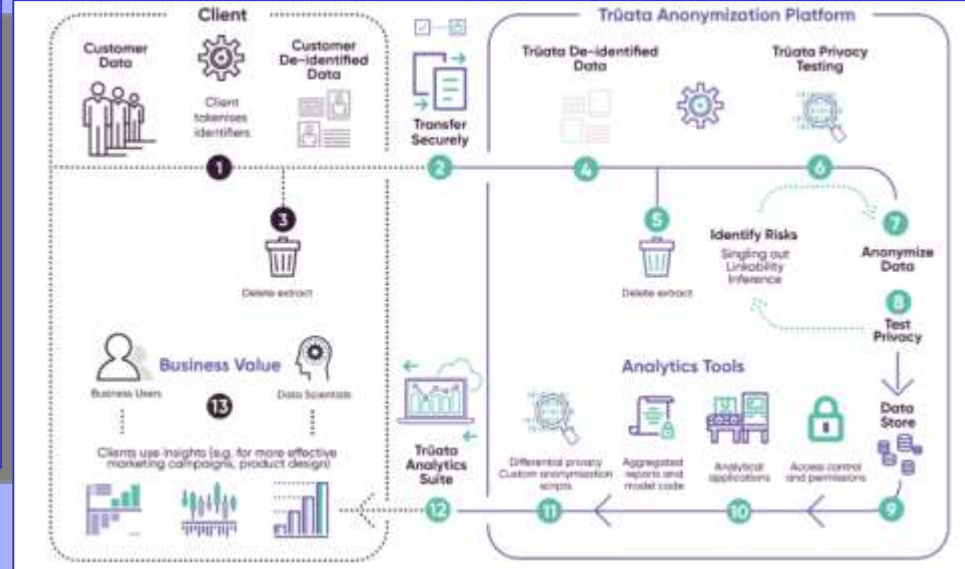




# ΑΝΩΝΥΜΟΠΟΙΗΣΗ - ΨΕΥΔΩΝΥΜΟΠΟΙΗΣΗ - ΑΠΟΚΡΥΨΗ



Λύση απόκρυψης ευαίσθητων δεδομένων Imperva Camouflage (Πηγή: srcloudsolutions.co.uk)



Πλατφόρμα ανωνυμοποίησης Truata (Πηγή: truata.com)



# ΓΕΝΙΚΕΥΣΗ ΨΕΥΔΟ-ΑΝΑΓΝΩΡΙΣΤΙΚΩΝ (QUASI IDENTIFIERS)

ID	Quasi identifiers		Sensitive attribute
	Age	Country	Political views
1	35	Greenland	Liberal
2	35	Canada	Conservative
3	38	Belize	Liberal
4	40	Belize	Liberal
5	37	Canada	Conservative
6	37	Canada	Conservative

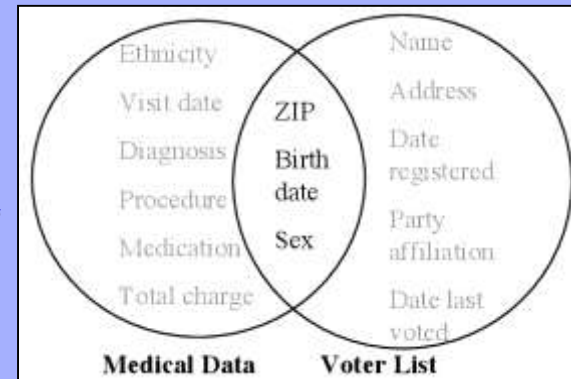
(a) Original data about the users

ID	Quasi Identifiers		Sensitive Attribute
	Age	Country	Political views
1	35-40	America	Liberal
2	35-40	America	Conservative
3	35-40	America	Liberal
4	35-40	America	Liberal
5	35-40	America	Conservative
6	35-40	America	Conservative

(b) 2-anonymous data about the users

Παράδειγμα γενίκευσης 2 ψευδο-αναγνωριστικών (2-anonymity) της Ηλικίας και της Χώρας, για να μην είναι δυνατή η ταυτοποίηση των υποκειμένων και ο προσδιορισμός των πολιτικών τους πεποιθήσεων (Πηγή: Majeed and Lee, 2020)

Παράδειγμα επίθεσης συσχετίσεων σε δύο πίνακες με κοινά πεδία ΤΚ, Ημερομηνία Γέννησης και Φύλο, για ταυτοποίηση των «ανωνυμοποιημένων» υποκειμένων (Πηγή: Sweeney, 2002)



# ΑΠΟΦΑΣΗ 20/2020 ΑΠΔΠΧ



- Η «εθνική ασφάλεια» συνιστά **αόριστη** νομική έννοια, η οποία πρέπει, με βάση τα δεδομένα που συντρέχουν σε κάθε περίπτωση, να ερμηνεύεται, προσδιορίζεται και τεκμηριώνεται ad hoc από τον υπεύθυνο επεξεργασίας.

- Συνεπώς, δεν αρκεί η αόριστη και γενική επίκληση της συνδρομής της εθνικής ασφάλειας για τον αποκλεισμό της ΑΠΔΠΧ από την άσκηση των αρμοδιοτήτων της, αλλά απαιτείται στη συγκεκριμένη περίπτωση **τεκμηρίωση** ότι η επεξεργασία διενεργείται για δραστηριότητα που αφορά την εθνική ασφάλεια της Ελλάδας, εκτός αν είναι πρόδηλη η συνδρομή της έννοιας αυτής (π.χ. αμιγώς στρατιωτικές εγκαταστάσεις με πολεμικό υλικό, μη προσβάσιμες στο κοινό).



# ΑΝΩΝΥΜΟΠΟΙΗΜΕΝΑ DATASETS

The screenshot shows the Data.world web interface. The browser address bar displays the URL: `data.world/abelvikas/diabetes-type-dataset/workspace/data-dictionary`. The page title is "Diabetes type dataset > Untitled project".

The left sidebar contains a "Project directory" section with a search bar and a list of items: "Home", "Project summary", and "Data dictionary" (which is selected). Below this are sections for "PROJECT FILES" (showing "No project files...yet."), "CONNECTED DATASETS" (showing "Diabetes type dataset (1)" and "Diabetestype.csv"), and "QUERIES" (showing "No queries...yet.>").


The main content area displays the "Data dictionary" for the file "Diabetestype.csv". It includes a description: "These dataset based on the blood sugar level, plasma glucose level and HbA1c." Below this is a table with three columns: "COLUMN NAME", "TYPE", and "DESCRIPTION".

COLUMN NAME	TYPE	DESCRIPTION
# age ⓘ	integer	This is the age of the patients
# bs_fast ⓘ	decimal	This is the Blood Sugar in fasting (Before meals), mmol/L
# bs_pp ⓘ	decimal	This is the Blood Sugar in 90 minutes after meals , mmol/L
# plasma_r ⓘ	decimal	Plasma glucose test Random taken at any time, mmol/L
# plasma_f ⓘ	decimal	Plasma glucose test Fasting usually taken in the morning because it should be taken after atleast 8 hours, mmol/L
# hba1c ⓘ	integer	mmol/mol
# type ⓘ	string	
# class ⓘ	boolean	

# ΑΝΩΝΥΜΟΠΟΙΗΜΕΝΟ DATASET (ΠΡΟΒΛΕΨΗ ΔΙΑΒΗΤΗ)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	Polyphagia	Genital thrush	visual blurring	Itching	Irritability	delayed healing	partial paresis	muscle stiffness	Alopecia	Obesity	class
2	40	Male	No	Yes	No	Yes	No	No	No	Yes	No	Yes	No	Yes	Yes	Yes	Positive
3	58	Male	No	No	No	Yes	No	No	Yes	No	No	No	Yes	No	Yes	No	Positive
4	41	Male	Yes	No	No	Yes	Yes	No	No	Yes	No	Yes	No	Yes	Yes	No	Positive
5	45	Male	No	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	No	No	No	Positive
6	60	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Positive
7	55	Male	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Positive
8	57	Male	Yes	Yes	No	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No	No	Positive
9	66	Male	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	No	No	Positive
10	67	Male	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Positive
11	70	Male	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	No	Yes	No	Positive
12	44	Male	Yes	Yes	No	Yes	No	Yes	No	No	Yes	Yes	No	Yes	Yes	No	Positive
13	38	Male	Yes	Yes	No	No	Yes	Yes	No	Yes	No	Yes	No	Yes	No	No	Positive
14	35	Male	Yes	No	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	No	Positive
15	61	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Positive
16	60	Male	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	No	No	Positive
17	58	Male	Yes	Yes	No	Yes	Yes	No	No	No	No	Yes	Yes	Yes	No	No	Positive
18	54	Male	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	No	Yes	No	No	Positive
19	67	Male	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Positive
20	66	Male	Yes	Yes	No	Yes	Yes	No	Yes	No	No	No	Yes	Yes	No	No	Positive
21	43	Male	Yes	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No	No	No	Positive
22	62	Male	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	No	No	Positive
23	54	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	Positive
24	39	Male	Yes	No	Yes	No	No	Yes	No	Yes	Yes	No	No	No	Yes	No	Positive
25	48	Male	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Positive
26	58	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes	No	Yes	Positive
27	32	Male	No	No	No	No	No	Yes	No	No	Yes	Yes	No	No	No	Yes	Positive
28	42	Male	No	No	No	Yes	Yes	No	No	No	Yes	No	No	Yes	No	No	Positive
29	52	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	No	No	Positive
30	38	Male	No	Yes	No	No	No	Yes	No	No	No	No	No	No	Yes	No	Positive
31	53	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Positive
32	57	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Positive

← → ↻ 🔒 knowi.com 🔍 🗑️ ☆ 👤 Incognito


 Platform ▾ Solutions ▾ Resources ▾ Datasources ▾ Login [Start Trial](#) [Request Demo](#)

## Unified analytics platform for ALL your data





Business Intelligence for NoSQL data, SQL data, and everything in between

[Start Trial](#) [Request Demo](#)

Free 21 day trial. No credit card required.

 **knowi**  
Unified Analytics for Modern Data

▶ 0:05 / 0:09 🔊 🗑️ ⋮

 Cloud or On-premise  Native NoSQL, SQL,  Instant Visual Insights  NLP, Machine Learning, Predictions [Chat](#)

# ΕΠΙΛΟΓΗ ΜΕΘΟΔΟΥ ML

The screenshot shows the Knowi web application interface. At the top, the browser address bar displays `knowi.com/manage#listMLs`. A notification banner states: "Your account will expire in 20 days. Contact us to convert your trial account." with a "Contact us" button. The main header is "Machine Learning" with a "NEW WORKSPACE" button on the right. A left sidebar contains navigation icons for Dashboards, Widgets, Data sources, Queries, Machine Learning, Settings, Help, and Logout. The main content area shows a table with columns for "WORKSPACES", "PUBLISHED MODEL", "WORKSPACE NAME", and "CREATED DATE". A modal dialog titled "Create Workspace" is open in the center. It features a "Workspace Name" input field containing "Predict\_Diabetes\_Classification\_1". Below the input, a message reads: "Please select type of new workspace. Type of workspace defines machine learning model which will be applied." Three options are presented in cards: "Classification" (with a decision tree icon), "Regression" (with a scatter plot icon), and "Anomaly Detection" (with a waveform icon). Each card includes a brief description and an "Example" link. At the bottom of the dialog are "Cancel" and "Create" buttons. A "Support" button is visible in the bottom right corner of the application.

← → ↻ 🔒 knowi.com/manage#listMLs 🔑 🔍 ☆ 6 ⋮

Your account will expire in 20 days. Contact us to convert your trial account. [Contact us](#)

## Machine Learning

[NEW WORKSPACE](#)

< BACK TO DASHBOARDS

WORKSPACES PUBLISHED MODEL

WORKSPACE NAME

SEARCH:

CREATED DATE


SHOWING 0 TO 0 OF 0 ENTRIES

Previous Next

### Create Workspace

Workspace Name:


Please select type of new workspace. Type of workspace defines machine learning model which will be applied.



#### Classification

The predicted output will fall into a clearly defined category or group.


Example [?](#)



#### Regression

The predicted output can have a continuous value (i.e., there are no predefined choices).

Example [?](#)



#### Anomaly Detection

The ability to identify time series events that do not conform to the expected pattern.

Example [?](#)

[Cancel](#) [Create](#)

[Support](#)

# ΑΝΕΒΑΣΜΑ ΔΕΔΟΜΕΝΩΝ ΕΚΠΑΙΔΕΥΣΗΣ

The screenshot shows the Knowi web interface at the URL `knowi.com/manage#listMLs`. A modal dialog titled "Upload Training Data" is open in the center. The dialog has a close button (X) in the top right corner. It contains two input fields: "Name" with the value "diabetes\_data\_upload" and "File" with a "Choose File" button and the filename "diabetes\_data\_upload.csv". At the bottom of the dialog are two buttons: "Cancel" (orange) and "Upload" (blue). The background interface is dimmed and shows a sidebar with navigation icons for Dashboards, Widgets, Data sources, Queries, Machine Learning, Settings, Help, and Logout. A notification at the top of the page states "Your account will expire in 20 days. Contact us to convert your trial account." with a "Contact us" button. A "Support" button is visible in the bottom right corner of the page.

# ΕΠΙΛΟΓΗ ΠΕΔΙΟΥ ΓΙΑ ΠΡΟΒΛΕΨΗ

The screenshot shows the Knowi web interface at [knowi.com/manage#listMLS](https://knowi.com/manage#listMLS). A yellow banner at the top indicates that the user's account will expire in 20 days and provides a "Contact us" button. The main content area contains the instruction: "Please select the dataset, attribute to predict and analyze the data for possible data integrity issues and recommended courses of action".

The interface is divided into three numbered steps:

- 1 Select Dataset (one item only)**: A dropdown menu is currently set to "diabetes\_data\_upload". Below it, there is a link to "Or upload new dataset".
- 2 Cloud9QL**: A text input field for entering a query.
- 3 Attribute to Predict**: A dropdown menu labeled "Select property" is open, showing a list of attributes. The attribute "class" is highlighted in red at the bottom of the list.

The attribute list includes: Age, Gender, Polyuria, Polydipsia, sudden weight loss, weakness, Polyphagia, Genital thrush, visual blurring, Itching, Irritability, delayed healing, partial paresis, muscle stiffness, Alopecia, Obesity, and class.

A vertical sidebar on the left contains navigation icons for Dashboards, Widgets, Data sources, Queries, Machine Learning, Settings, Help, and Logout. A "Support" button is located in the bottom right corner of the interface.



# ΑΝΑΛΥΣΗ ΚΑΙ ΠΡΟΕΤΟΙΜΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

The screenshot shows the Knowi ML interface. At the top, there is a navigation bar with the URL 'knowi.com/manage#listMLs' and a notification: 'Your account will expire in 20 days. Contact us to convert your trial account.' Below this, a message asks the user to 'Please select the dataset, attribute to predict and analyze the data for possible data integrity issues and recommended courses of action'. The interface is divided into two main sections: 'Select Dataset' and 'Attribute to Predict'. The 'Select Dataset' section has a dropdown menu with 'diabetes\_data\_upload' selected and a link to 'upload new dataset'. The 'Attribute to Predict' section has a dropdown menu with 'class' selected. Below these sections are two buttons: 'Analyze Data' (orange) and 'Prepare Data' (blue). At the bottom, there is a 'Preview Data' section showing a table with 17 columns and 5 rows of data. The columns are: Age, Gender, Polyuria, Polydipsia, sudden weight weakness, Polyphagia, Genital thrush, visual blurring, Itching, Irritability, delayed healin, partial paresis, muscle stiffnes, Alopecia, Obesity, and class. The rows show various patient profiles with their attributes and predicted class (Positive).

1 **Select Dataset** (one item only)

diabetes\_data\_upload

Or [upload new dataset](#)

2 **Cloud9QL**

3 **Attribute to Predict**

class

4 **Start Data Analysis**

Analyze Data

5 **Start Data Preparation**

Prepare Data

Preview Data ( Prediction candidate: class )

Age	Gender	Polyuria	Polydipsia	sudden weight weakness	Polyphagia	Genital thrush	visual blurring	Itching	Irritability	delayed healin	partial paresis	muscle stiffnes	Alopecia	Obesity	class
40	Male	No	Yes	No	Yes	No	No	No	Yes	No	Yes	No	Yes	Yes	Positive
58	Male	No	No	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	Positive
41	Male	Yes	No	No	Yes	Yes	No	No	Yes	No	Yes	No	Yes	No	Positive
45	Male	No	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	No	No	Positive
60	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Positive
55	Male	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Positive
57	Male	Yes	Yes	No	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No	Positive

520 rows / 17 columns

Support

# ΣΤΑΔΙΑ ΠΡΟΕΤΟΙΜΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

01

Data Types

02

Outliers

03

Missing  
Values

04

Rescaling

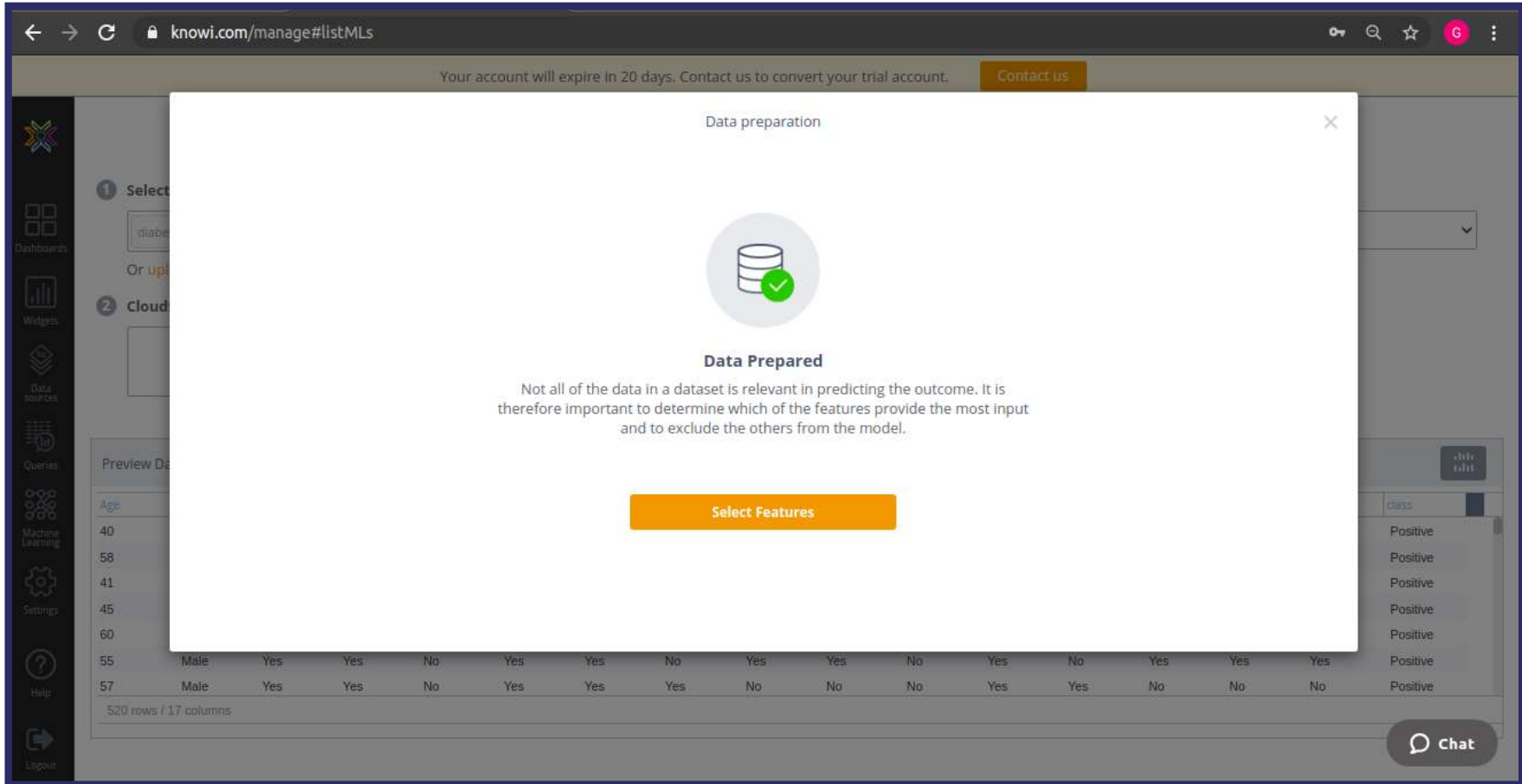
05

Discrete  
Grouping

06

Dummy  
Variables


# ΜΗΝΥΜΑ ΕΤΟΙΜΟΤΗΤΑΣ ΔΕΔΟΜΕΝΩΝ



← → ↻ 🔒 knowi.com/manage#listMLs 🔑 🔍 ☆ G ☰

Your account will expire in 20 days. Contact us to convert your trial account. [Contact us](#)

Data preparation ✕



**Data Prepared**

Not all of the data in a dataset is relevant in predicting the outcome. It is therefore important to determine which of the features provide the most input and to exclude the others from the model.

[Select Features](#)

Preview Data

Age																					class
40																					Positive
58																					Positive
41																					Positive
45																					Positive
60																					Positive
55	Male	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes				Positive
57	Male	Yes	Yes	No	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No	No	No					Positive

520 rows / 17 columns

Chat

# ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΒΛΕΨΗΣ (ΠΟΣΟΣΤΟ ΑΚΡΙΒΕΙΑΣ %)

The screenshot displays the Knowi Machine Learning dashboard. At the top, a navigation bar includes a back arrow, a refresh icon, and the URL 'knowi.com/manage#listMLs'. A yellow notification banner states, 'Your account will expire in 20 days. Contact us to convert your trial account.' with a 'Contact us' button. The main header shows 'Machine Learning' and a 'NEW WORKSPACE +' button. Below this is a breadcrumb trail '< BACK TO DASHBOARDS' and tabs for 'WORKSPACES' and 'PUBLISHED MODELS'. A table lists the workspace 'Predict\_Diabetes\_Classification\_1' with 'Classification' ML type and '0' published models. Below the table, a section titled 'Available models' allows selecting models for training. The 'Train' button is highlighted. The 'Results' section shows the accuracy of four models: Decision Tree (95.5769%), Logistic Regression (92.6923%), KNN (94.2308%), and Naive Bayes (88.2692%). A 'History' section at the bottom lists the same models and their accuracies. A 'Chat' button is visible in the bottom right corner.

knowi.com/manage#listMLs

Your account will expire in 20 days. Contact us to convert your trial account. [Contact us](#)

## Machine Learning

[NEW WORKSPACE +](#)

< BACK TO DASHBOARDS

WORKSPACES PUBLISHED MODELS

WORKSPACE NAME	ML TYPE	# OF PUBLISHED MODELS/WIDGETS	CREATED DATE
Predict_Diabetes_Classification_1	Classification	0	01/12/2021 09:33:01 PST

Available models [Select All](#) [Select None](#)

- Decision Tree [Settings](#)
- Logistic Regression [Settings](#)
- K-Nearest Neighbor [Settings](#)
- Naive Bayes [Settings](#)

[Train](#)

### Results

- DecisionTree {maxNodes=100, nodeSize=1, splitRule=GINI}: Accuracy 95.5769%
- LogisticRegression {maxNumIteration=500, regularizationFactor=0.0, tolerance=0.00001}: Accuracy 92.6923%
- KNN {neighbors=1, distance.WeightVector=, distance=Euclidean}: Accuracy 94.2308%
- NaiveBayes {model=MULTINOMIAL, smooth=1.0}: Accuracy 88.2692%

### History

01/12/2021 09:36:35 PST

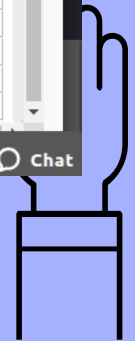
- ✘ DecisionTree {maxNodes=100, nodeSize=1, splitRule=GINI}: Accuracy 95.5769%
- ✘ LogisticRegression {maxNumIteration=500, regularizationFactor=0.0, tolerance=0.00001}: Accuracy 92.6923%
- ✘ KNN {neighbors=1, distance.WeightVector=, distance=Euclidean}: Accuracy 94.2308%
- ✘ NaiveBayes {model=MULTINOMIAL, smooth=1.0}: Accuracy 88.2692%

Chat

# ΕΜΦΑΝΙΣΗ DATASET ΜΕ ΤΟ ΠΕΔΙΟ ΠΡΟΒΛΕΨΗΣ

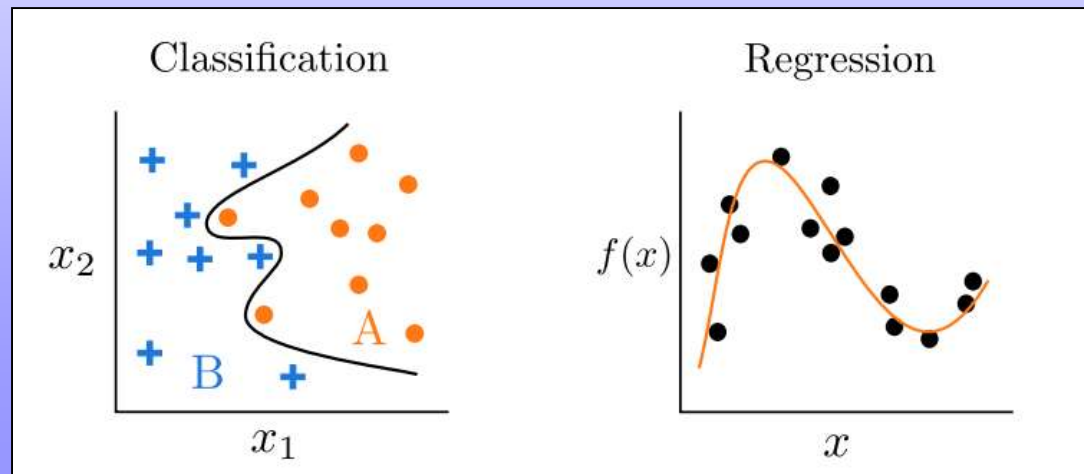


Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	Polyphagia	Genital thrush	visual blurring	itching	Irritability	delayed healing	partial paresthesia	muscle stiffness	Alopecia	Obesity	class	Predicted class
33	Male	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No	Positive	Positive
55	Male	Yes	Yes	Yes	Yes	No	Yes	No	No	Yes	No	Yes	No	No	No	Positive	Positive
48	Male	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes	No	No	No	No	Positive	Positive
56	Male	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No	No	No	Yes	No	Positive	Positive
38	Female	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Positive	Positive
28	Female	No	No	No	No	No	No	Yes	No	No	No	Yes	Yes	No	No	Positive	Positive
68	Female	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	No	No	Positive	Positive
35	Female	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Positive	Positive
45	Female	No	No	No	No	Yes	No	Yes	Yes	No	No	Yes	No	No	No	Positive	Positive
48	Female	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Positive	Positive
40	Female	Yes	Yes	Yes	Yes	No	No	Yes	No	No	Yes	Yes	Yes	No	No	Positive	Positive
57	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Positive	Positive
47	Male	No	No	No	No	No	No	No	Yes	No	No	No	No	Yes	No	Negative	Negative
45	Male	No	No	No	No	Yes	Yes	No	No	No	No	No	No	No	No	Negative	Negative
57	Male	No	No	No	No	Yes	No	Yes	No	No	No	No	Yes	No	No	Negative	Negative
72	Male	Yes	No	No	No	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Negative	Negative
30	Male	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Negative	Negative
27	Male	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Negative	Negative
38	Male	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Negative	Negative
43	Male	No	No	No	Yes	No	Yes	No	Yes	No	Yes	No	No	Yes	No	Negative	Negative
40	Male	No	No	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Negative	Negative
47	Male	No	No	No	No	No	No	No	Yes	No	No	No	No	Yes	No	Negative	Negative
45	Male	No	No	No	No	Yes	Yes	No	No	No	No	No	No	No	No	Negative	Negative



# ΣΕΝΑΡΙΑ ΠΡΟΓΝΩΣΗΣ ΠΟΥ ΔΟΚΙΜΑΣΤΗΚΑΝ

1. Κατάσταση Διαβητικός ή Όχι – Ταξινόμηση, 97.1154%
2. Τιμή Γλυκοζυλιωμένης Αιμοσφαιρίνης – Παλινδρόμηση, 20.0079
3. Κατάσταση Διαβήτη (4 Κατηγορίες) – Ταξινόμηση, 100%
4. Κατάσταση Μεταβολικό Σύνδρομο ή Όχι – Ταξινόμηση, 82.5286%
5. Κατάσταση Καρδιοπαθής ή Όχι – Ταξινόμηση, 75.9259%
6. Κατάστασης Καρκίνος Πνεύμονα ή Όχι – Ταξινόμηση, 99.9%



# PUBLISHED MONTELA ΕΡΕΥΝΑΣ

The screenshot shows the 'Machine Learning' section of the Knowi platform. At the top, there is a navigation bar with a back arrow, a refresh icon, and the URL 'knowi.com/manage#listMLs'. A yellow banner below the navigation bar indicates that the user's account will expire in 18 days and provides a 'Contact us' button. The main content area is titled 'Machine Learning' and includes a 'NEW WORKSPACE +' button. Below this, there are tabs for 'WORKSPACES' and 'PUBLISHED MODELS', with the latter being selected. A search bar shows '6 models' and a 'Quick search' input field. A table lists six published models with columns for Model Name, Workspace, Results, and Created Date. At the bottom right of the table, there is a 'Delete Multiple' button and a 'Support' button. The footer of the table area shows 'SHOWING 1 TO 6 OF 6 ENTRIES' and pagination controls for 'Previous', '1', and 'Next'.

← → ↻ 🔒 knowi.com/manage#listMLs 🔑 🔍 ☆ G ⋮

Your account will expire in 18 days. Contact us to convert your trial account. [Contact us](#)



## Machine Learning

[NEW WORKSPACE +](#)

< [BACK TO DASHBOARDS](#)

[WORKSPACES](#) **PUBLISHED MODELS**


6 models 🔍 Quick search

Model Name	Workspace	Results	Created Date	
<b>Predict_Diabetes_Classification1_DT</b>	Predict_Diabetes_Classification_1	DecisionTree {}: Accuracy 97.1154%	01/12/2021 14:57:39 PST	 
<b>Predict_HeartDis_Classification1_DT</b>	Predict_HeartDis_Classification_1	DecisionTree {}: Accuracy 75.9259%	01/12/2021 15:11:56 PST	
<b>Predict_MS_Classification_1_DT</b>	Predict_MS_Classification_1	DecisionTree {}: Accuracy 82.5286%	01/12/2021 23:28:59 PST	
<b>Predict_LungCa_Classification1_DT</b>	Predict_LungCa_Classification_1	DecisionTree {}: Accuracy 99.9%	01/14/2021 01:31:30 PST	
<b>Predict_HbA1c_Regression1_OLS</b>	Predict_HbA1c_Regression_1	OLS {}: RMSE 20.0079	01/14/2021 08:17:51 PST	
<b>Predict_DiabetesType_Classification_2_DT</b>	Predict_DiabetesType_Classification_2	DecisionTree {}: Accuracy 100.0%	01/14/2021 08:31:17 PST	

[Delete Multiple](#)

SHOWING 1 TO 6 OF 6 ENTRIES

Previous  Next

 [Support](#)

# ΣΥΜΠΕΡΑΣΜΑΤΑ

1. Αξία Προγνωστικής Ανάλυσης των ΜΔΥ για τις ΕΔ και το προσωπικό τους.
2. Εφαρμογές της έρευνας στις ΕΔ
3. Ρόλος Πληροφορικής
4. Αξία Μαζικών Δεδομένων Υγείας
5. Ζητήματα ασφάλειας Μαζικών Δεδομένων Υγείας
6. Σημεία προσοχής στα δεδομένα εκπαίδευσης αλγορίθμων ML
7. Σημασία της AI και Ανάλυσης Δεδομένων για τα ΜΔΥ
8. Αξιοποίηση των Ατομικών Ηλεκτρονικών Φακέλων Υγείας (ΑΗΦΥ)
9. Σύγχρονα Συστήματα Υποστήριξης Κλινικών Αποφάσεων (CDSS)
10. Προστασία δεδομένων υγείας: Δικαίωμα και Τροχοπέδη
11. Επιστημονική έρευνα με δεδομένα υγείας







## Αξία Προγνωστικής Ανάλυσης των ΜΔΥ για τις ΕΔ και το προσωπικό τους:

- Κατάσταση της υγείας υποκειμένων: προαπαιτούμενο της πολεμικής τους ετοιμότητας και αποφασιστικής σημασίας παράγοντας μαχητικής ισχύος.
- Υγιή στελέχη => θετικό αντίκτυπο στην εθνική μας οικονομία, Bloom et al., 2004).
- Η γνώση της κατάστασης της υγείας του προσωπικού:
  - Πολύτιμη πληροφορία για τον εκάστοτε επιχειρησιακό διοικητή.
  - Δυνατότητα στην ηγεσία ΕΔ να αξιολογεί το αξιόμαχο.
  - Οδηγός για μέγιστη αξιοποίηση του προσωπικού (προσαρμογή καθηκόντων, διαφοροποίηση δεξιοτήτων).
  - Δυνατότητα έγκαιρης συμβουλής (για τροποποίηση τρόπων εκπαίδευσης, διατροφής, άσκησης, φαρμακευτικής αγωγής).
  - Χρήση σε CDSS ιατρικών αποφάσεων (δρώντας συμβουλευτικά και συμπληρωματικά).

(συνεχίζεται)



## Αξία Προγνωστικής Ανάλυσης των ΜΔΥ για τις ΕΔ και το προσωπικό τους:

- Εστιασμένη παροχή υγειονομικών υπηρεσιών στο κάθε στέλεχος και όχι γενικά, θα επιτρέψει να εξετάζονται απογραφικά όλα τα στελέχη και να ομαδοποιούνται τα αποτελέσματά τους κατά Κλάδο, Όπλο ή Σώμα, κατά εργασία σε ακραίες συνθήκες περιβάλλοντος κ.λπ.
- Κατάλληλες τεχνικές ML μπορούν, συμπληρωματικό βοήθημα, για άτομα σε κίνδυνο, λόγω αυτοκτονικής συμπεριφοράς (Van Mens et al., 2020).
- Οφέλη από αξιοποίηση Προγνωστικής Ανάλυσης εκτιμάται ότι θα συμβάλλουν:
  - Προαγωγή της σωματικής και ψυχικής υγείας των στελεχών.
  - Καλλιέργεια αισθήματος εμπιστοσύνης στο εν λόγω σύστημα.
  - Ενίσχυση επιστημονικού έργου ιατρών και νοσηλευτών.
  - Επέκταση υφιστάμενου συστήματος Ιατρικής παρακολούθησης.
  - Εξοικονόμηση πολύτιμων υγειονομικών πόρων.



## Εφαρμογές της έρευνας στις ΕΔ:

- Τα 6 σενάρια Προγνωστικής Ανάλυσης, μπορούν να έχουν εφαρμογή στα στελέχη των ΕΔ (επιλέχθηκαν προγνώσεις, που εκτιμάται ότι αντιμετωπίζονται στην καθημερινότητα).
- Παρέμβαση, στην εντερική μικροχλωρίδα, μέσω αποτελεσματικής δίαιτας και λήψης συμπληρωμάτων διατροφής (Eetemadi et al., 2020).
- Απομακρυσμένη παρακολούθηση με συσκευές IoT (εργασία, σπίτι, συμμετοχή σε ασκήσεις ή πολεμικές επιχειρήσεις).
- Έλεγχος ζωτικών σημείων μαχόμενων στελεχών μπορεί να δώσει πολύ κρίσιμη πληροφόρηση στον επιχειρησιακό διοικητή.
- Η εφαρμογή Whoops που λειτουργεί σε περιβάλλον νέφους Azure, για ανίχνευση περιστατικών πτώσης (Mrozek et al., 2020), εντοπισμό πτώσης μαχητών στο πεδίο μάχης και έγκαιρη ενημέρωση προϊσταμένων ή ιατρικού προσωπικού.



## Ρόλος Πληροφορικής:

- Παρέχει τα κατάλληλα εργαλεία, ώστε να μετατρέπει δεδομένα σε πληροφορίες και πληροφορίες σε γνώση.
- Βοηθά στην κατανόηση των βαθέων σχέσεων μέσα από τα δεδομένα, ανακτώντας και εξάγοντας υποκείμενα μοτίβα (patterns) μέσα από αυτά.
- Ζωτική σημασία στην εκμετάλλευση της αποκτηθείσας γνώσης, για την υποστήριξη της περίθαλψης των ασθενών και τελικά την καθοδήγηση της ατομικής τους συμπεριφοράς.
- Οι σύγχρονοι υπολογιστές μπορούν να επιτύχουν ταχύτατα ομαδοποίηση μεγάλων datasets, με βάση ποικίλα χαρακτηριστικά τους γνωρίσματα, δυνατότητα χρησιμότητα στην περίπτωση της μη Επιτηρούμενης ML.



## Αξία Μαζικών Δεδομένων Υγείας:

- Δεδομένα ιατρικών φακέλων και αναλύσεις δεδομένων τους, αποτελούν πολύτιμες και αξιόπιστες πηγές συγκέντρωσης ΜΔΥ.
- Σύνθετοι αλγόριθμοι επιτρέπουν μετατροπή τους σε πολύτιμη γνώση.
- Συνδυασμός τους με datasets γονιδιωματικής ή περιβαλλοντολογικής προέλευσης, οδηγούν σε πρόγνωση σοβαρών ασθενειών (Willems et al., 2019).
- Συνδυασμός Ενισχυτικής ML και ΜΔΥ δημιουργεί μοναδικές ευκαιρίες, για βελτιστοποίηση θεραπειών, ειδικά για επαναλαμβανόμενες (Coronato et al., 2020).
- Ραγδαία αύξηση όγκου και ποιότητας εργαστηριακών δεδομένων σε συνδυασμό με τεχνικές ML, επιτρέπουν να αντληθεί βιολογική εικόνα:
  - κατανοητοί οι μηχανισμοί της νόσου, νέες θεραπευτικές στρατηγικές, βελτίωση διαγνωστικών εργαλείων για κλινική εφαρμογή (Toh et al., 2019).



## Ζητήματα ασφάλειας Μαζικών Δεδομένων Υγείας:

- Σοβαρές ανησυχίες σχετικά με ανάκτηση, αποθήκευση, ανάλυση και ασφάλεια των πληροφοριών των επαγγελματιών υγείας.
- Προκλήσεις όπως: έλλειψη τυποποίησης δεδομένων, ανησυχία για ζητήματα απορρήτου, περιορισμός ταχύτητας, δυνατότητα υποκλοπής κατά τη μεταφορά δεδομένων.
- Τεχνικές όπως: ανωνυμοποίησης, ψευδωνυμοποίησης, προσφέρουν μεγαλύτερη προστασία στα υποκείμενα.
- Τεχνικές όπως: παρακολούθηση δραστηριότητας, ομοιομορφική κρυπτογράφηση, ενσωματωμένος έλεγχος ταυτότητας και λεπτομερής έλεγχος πρόσβασης, εγγυώνται ασφάλεια των δεδομένων υγείας (Kaur et al., 2018).
- Τεχνικές απόκρυψης ευαίσθητων δεδομένων, τα αντικαθιστούν με ρεαλιστικά φανταστικά, διατηρώντας ακεραιότητα στοιχείων αναφοράς, στατιστική ακρίβεια, προστασία από κλοπή και συμμόρφωση με κανονισμούς και διεθνείς πολιτικές.

(συνεχίζεται)



## Ζητήματα ασφάλειας Μαζικών Δεδομένων Υγείας:

- Τα δεδομένα υγείας των στελεχών των ΕΔ, πολύτιμα για τον εχθρό (παρέχουν χρήσιμες πληροφορίες, περί της μαχητικής ικανότητας του στρατεύματος, κατά ηλικία, γεωγραφική περιοχή, ειδικότητα, περιόδους έντασης και αντιπαλότητας κ.λπ.)
- Αντιστρόφως, πιθανή διαρροή ή παραβίαση εχθρικών δεδομένων υγείας, μπορούν να οδηγήσουν, κατόπιν αναλύσεων, σε συμπεράσματα για την εχθρική μαχητική ισχύ, προς όφελος των ΕΔ μας.
- Στην Υπολογιστικής Νέφους, υπάρχουν κίνδυνοι παραβίασης του τρίπτυχου CIA, καθώς και της μη συμμόρφωσης των φορέων που επεξεργάζονται τα ΜΔΥ, με τις αρχές της λογοδοσίας και της διαφάνειας, που επιτάσσει ο ΓΚΠΔ. Απαιτούνται ασφαλείς αλγόριθμοι αποθήκευσης
- Στην Κινητή Υγεία (mHealth), η ιδιωτικότητα και η τήρηση αρχών ασφαλείας, είναι απαραίτητες προϋποθέσεις για την οικοδόμηση εμπιστοσύνης από πλευράς χρηστών.



## Σημεία προσοχής στα δεδομένα εκπαίδευσης αλγορίθμων ML:

- Να αποφεύγεται η μεροληπτική (biased) χρήση τους, για μείωση του κινδύνου αποτυχίας των μεθόδων ML, ως προς την γενίκευση των αποτελεσμάτων,
- Κατάρτιση μοντέλων ML από εξαιρετικά «θορυβώδεις» πηγές δεδομένων, μπορεί να οδηγήσει ακόμη και σε ανεπιθύμητα αποτελέσματα (Waring et al., 2020).
- Τα εργαλεία ML πρέπει να εκπαιδεύονται:
  - με ικανό πλήθος πραγματικών (όχι «συνθετικών») δεδομένων ασθενών και
  - με χρήση κατευθυντήριων οδηγιών (guidelines) έμπειρων ιατρών,
  - διαφορετικά πιθανόν να οδηγηθούμε σε εξαγωγή ανακριβών συμπερασμάτων (Ross and Swetlitz, 2018).





## Σημεία προσοχής στα δεδομένα εκπαίδευσης αλγορίθμων ML:

- Η ανισορροπία που οφείλεται στο μικρό μέγεθος των μελετών σειράς των ασθενών (patient cohorts), μπορεί να προκαλέσει ανεπιθύμητη «υπερπροσαρμογή» (over-fitting) όχι μόνο στα κλασικά μοντέλα AI αλλά και στα σύγχρονα μοντέλα βαθιάς μάθησης (Peng, 2015).
- Η σχέση μεταξύ της μεταβλητής στόχου, την οποία το μοντέλο προσπαθεί να προβλέψει, και των ανεξάρτητων μεταβλητών αλλάζει με το χρόνο. Δεν πρέπει να αγνοείται η έννοια της μετατόπισης (drift). Απαιτείται συνεχής παρακολούθηση, με μετρήσεις (Shendre, 2020).
- Αντί αδιάκριτης εφαρμογής τεχνικών ML, σε οποιοδήποτε σύνολο δεδομένων, η προσήλωση στο προς απάντηση βιολογικό ερώτημα, που τίθεται κάθε φορά, καθώς και στα δεδομένα που απαιτούνται για την απάντησή του, είναι πρωταρχικής σημασίας (Toh et al., 2019).
- Οι ελεγκτικές αρχές θα πρέπει να είναι σε θέση να δοκιμάζουν και να πιστοποιούν τα δεδομένα που χρησιμοποιούνται από τους αλγορίθμους (EC, White Paper on AI, 2020).



## Σημασία της AI και Ανάλυσης Δεδομένων για τα ΜΔΥ :

- Με τη σταδιακή (incremental) μάθηση, τα αποτελέσματα από ένα σύστημα AI ενσωματώνονται σε κλειστό βρόχο ανατροφοδότησης και τελειοποιούν την προγνωστική ακρίβεια του συστήματος, μέσω επαναληπτικής επανεκπαίδευσης του μοντέλου.
- Η Προδιαγραφική Ανάλυση, ως συμπλήρωμα της Προγνωστικής, συμβάλλει στην επίλυση προβλημάτων βελτιστοποίησης, με βοήθεια στατιστικής και μαθηματικών, αποτελώντας την τελική λύση υποστήριξης, για ένα πλήρως Ευφυές Σύστημα (Intelligent System).
- Η AI συνεισφέρει στη μείωση του διοικητικού φόρτου των φορέων υγειονομικής περίθαλψης, επιτρέποντας περισσότερο χρόνο για την φροντίδα των ασθενών.
- Τα σύγχρονα συστήματα υγείας ενισχύονται, με απώτερο σκοπό τη βελτίωση της υγείας και της ποιότητας των παρεχόμενων υγειονομικών υπηρεσιών στους ασθενείς.
- Τα νοσοκομεία συντομεύουν τις υπάρχουσες γραμμές αναμονής, ενώ αυξάνονται οι δυνατότητες παροχής θεραπείας, σε όσους πραγματικά τη χρειάζονται (Lopes et al., 2020).
- Για εποπτεία και διαφάνεια των συστημάτων AI είναι αναγκαία η λειτουργία κατάλληλων νομικών κανόνων, στο χώρο της υγείας (EC, White Paper on AI, 2020).

# ΣΥΜΠΕΡΑΣΜΑΤΑ (8/11)

## Αξιοποίηση των ΑΗΦΥ:

- Η επικράτηση των Μαζικών Δεδομένων και της ανάλυσής τους, προσφέρουν την ευκαιρία αξιοποίησης των ΑΗΦΥ (υποχρεωτικών πλέον στη χώρα μας με το Ν.4600/2019).
- Παρέχονται καλύτερα εργαλεία και περισσότερες πληροφορίες στον θεράποντα ιατρό, ως προς τη λήψη αποφάσεων και συνεπώς καλύτερη ποιότητα ιατρικών υπηρεσιών.
- Δύνανται να χρησιμοποιούν το **«είδωλο υγείας»**, δηλαδή την εικονική αναπαράσταση με:
  - όλες τις σχετικές με την υγεία του υποκειμένου πληροφορίες.
  - έξυπνους τρόπους διαχείρισης και πρόβλεψης της μελλοντικής του κατάστασης.



## Είδωλο Υγείας



(Πηγή: Prospero et al., 2018).

# ΣΥΜΠΕΡΑΣΜΑΤΑ (9/11)

## Σύγχρονα CDSS:

- Ισχυρότερα στους αλγόριθμους τους.
- Πιο προσαρμόσιμα σε νέα δεδομένα.
- Ενσωματωμένα πιο απρόσκοπτα στις υπάρχουσες τεχνολογίες.
- Αποτελεσματικότερα στην αντιμετώπιση ζητημάτων των ασθενών.
- Χρησιμοποιούν μοντέλα πρόβλεψης για στοχευμένη διενέργεια θεραπειών ή αποφυγή της.
- Δεν έρχονται να αντικαταστήσουν το ιατρικό και νοσηλευτικό προσωπικό, αντιθέτως έρχονται να υποστηρίξουν τη λήψη των αποφάσεων τους.



(Πηγή: <https://www.mccourier.com/>)



# ΣΥΜΠΕΡΑΣΜΑΤΑ (10/11)

## Προστασία δεδομένων υγείας, Δικαίωμα και Τροχοπέδη:

- Κάθε νοσηλευτικό ίδρυμα ενεργεί απλώς ως ο «θεματοφύλακας» των ΜΔΥ των ασθενών.
- Τα δεδομένα ανήκουν στα υποκείμενα και η πρόσβαση και χρήση τους, απαιτεί τη ρητή συγκατάθεση των ιδίων.
- Αυτό το δικαίωμα αποτελεί ταυτόχρονα τροχοπέδη, στην ταχεία εκμετάλλευση του μεγάλου όγκου δεδομένων υγείας.
- Η προοπτική απώλειας ανωνυμοποιημένων δεδομένων λόγω διαγραφής τους στο τέλος του έργου, είναι απογοητευτική για την πρόοδο της ιατρικής γνώσης (Hulsen et al.,2019).



(Πηγή: <https://weseenow.co.uk/>)



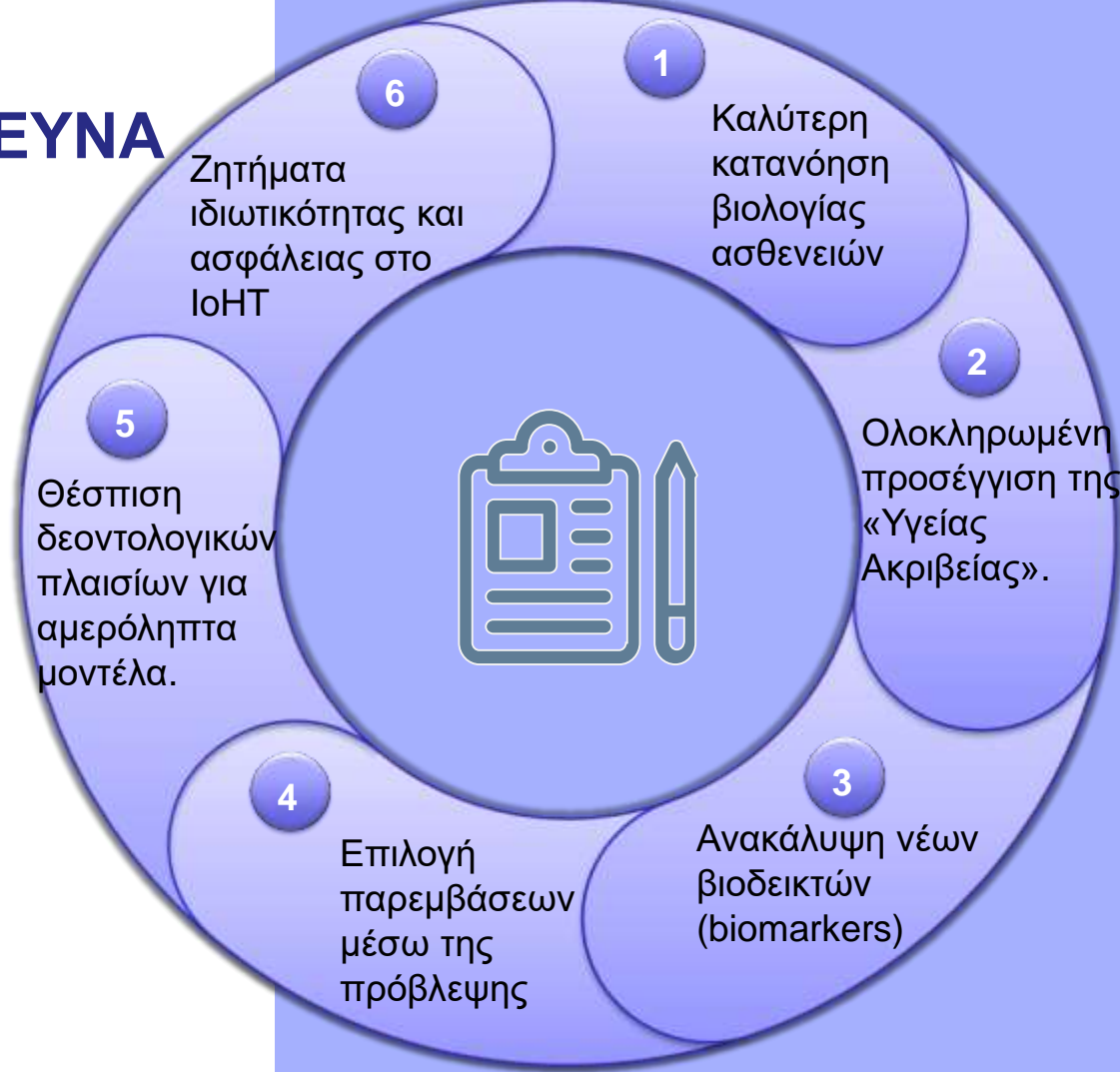
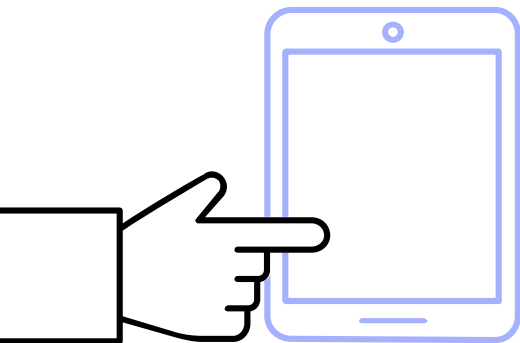
# ΣΥΜΠΕΡΑΣΜΑΤΑ (11/11)

## Επιστημονική έρευνα με δεδομένα υγείας:

- Η διενέργεια επιστημονικής έρευνας με δεδομένα υγείας ασθενών, από τα αρχεία ενός νοσηλευτικού ιδρύματος, συνιστά νόμιμο σκοπό επεξεργασίας.
- Συνιστά επεξεργασία ειδικών κατηγοριών προσωπικών δεδομένων.
- Το νοσηλευτικό ίδρυμα πρέπει να ζητήσει την άδεια της ΑΠΔΠΧ, για να επιτραπεί η πρόσβαση του ερευνητή στα εν λόγω αρχεία.
- Αν η δημοσίευση στοιχείων του ιατρικού φακέλου ασθενών, γίνεται κατά τρόπο που να μην είναι αμέσως ή εμμέσως προσδιορίσιμη η ταυτότητά τους, αυτό ΔΕΝ συνιστά επεξεργασία προσωπικών δεδομένων.



# ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ





## ΕΡΩΤΗΣΕΙΣ - ΣΥΖΗΤΗΣΗ

*«Το διοικεῖν ἐστὶ προβλέπειν»,  
Αλκιβιάδης (450-404 π.Χ.)*

PREDICTIVE  
ANALYTICS

Customers

Costs

Income

Plan