



PÁZMÁNY

Pázmány Péter Katolikus Egyetem

Információs Technológiai és Bionikai Kar



*Katolikus  
Pedagógiai  
Intézet*

# Középiskolai informatikatanárok szaktárgyi továbbképzése

---

2022. november 23.

# Adatbázis trendek

---

2022. november 23.

**Dr. Zsedrovits Tamás**  
*egyetemi docens, PPKE ITK*

# ADATBÁZIS TRENDK

---

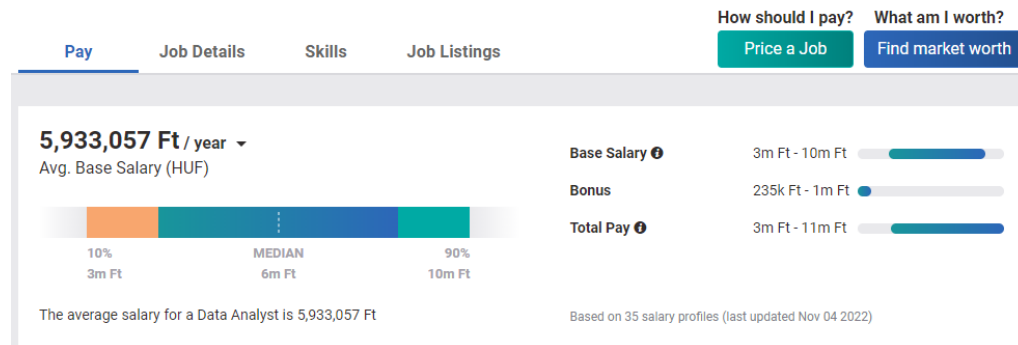
- „Adatos” állások
- Adatbázis kezelés fejlődése
- Big data
- Gépi tanulás – mesterséges intelligencia

# MUNKAERŐPIAC

---

- Forbes: [15 In-Demand Tech-Focused \(And Tech-Adjacent\) Skills And Specialties](#), 2022
  - 1. Software Development
  - 2. Artificial Intelligence
  - 5. Big Data
  - 6. Data Analytics

## Average Data Analyst Salary in Hungary



INNOVATION

# Why There Will Be No Data Science Job Titles By 2029



**Noah Gift** Forbes Councils Member  
Forbes Technology Council COUNCIL PO

POST WRITTEN BY

**Noah Gift**

Founder of [Pragmatic AI Labs](#). Adjunct professor at UC Davis Data Science program. Author of Pragmatic AI.

Fel

**Forbes**

ÜZLET PÉNZ LEGYÉL JOBB! A JÓ ÉLET FINOM INTERJÚ ZÖLD NAPI CÍMLAP



— ÜZLET

Viszlát, adattudósok! Ezek a munkakörök sehol sem lesznek 10 év múlva

[Why There Will Be No Data Science Job Titles By 2029 \(forbes.com\)](https://forbes.com)

[Viszlát, adattudósok! Ezek a munkakörök sehol sem lesznek 10 év múlva – Forbes.hu](https://forbes.hu)

# MIRE KÉSZÍTJÜK FEL A DIÁKOKAT/HALLGATÓKAT?

---

[A munkám a szenvedélyem: Vásárhelyi Orsolya adattudós | Marie Claire](#)

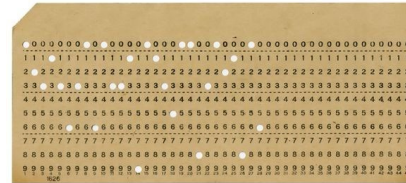
„Állítólag nem léteznek még azok a szakmák, amelyek a most iskolába járó gyerekek számára a legvonzóbbak lesznek felnőtt korukban. Amilyen gyorsan változik a világ, olyan gyorsan születnek új szakmák, és mennek ki a divatból régiek. Amikor Vásárhelyi Orsolya általános iskolába járt, Magyarországon még senki nem ismerte az adattudós kifejezést.”

[Arató Szandra - Mesterséges intelligencia munkák - Adattudós, ML engineer, Big Data Specialist - YouTube](#)

# ADATOK TÁROLÁSA ÉS KEZELÉSE SZÁMÍTÓGÉPPEL

---

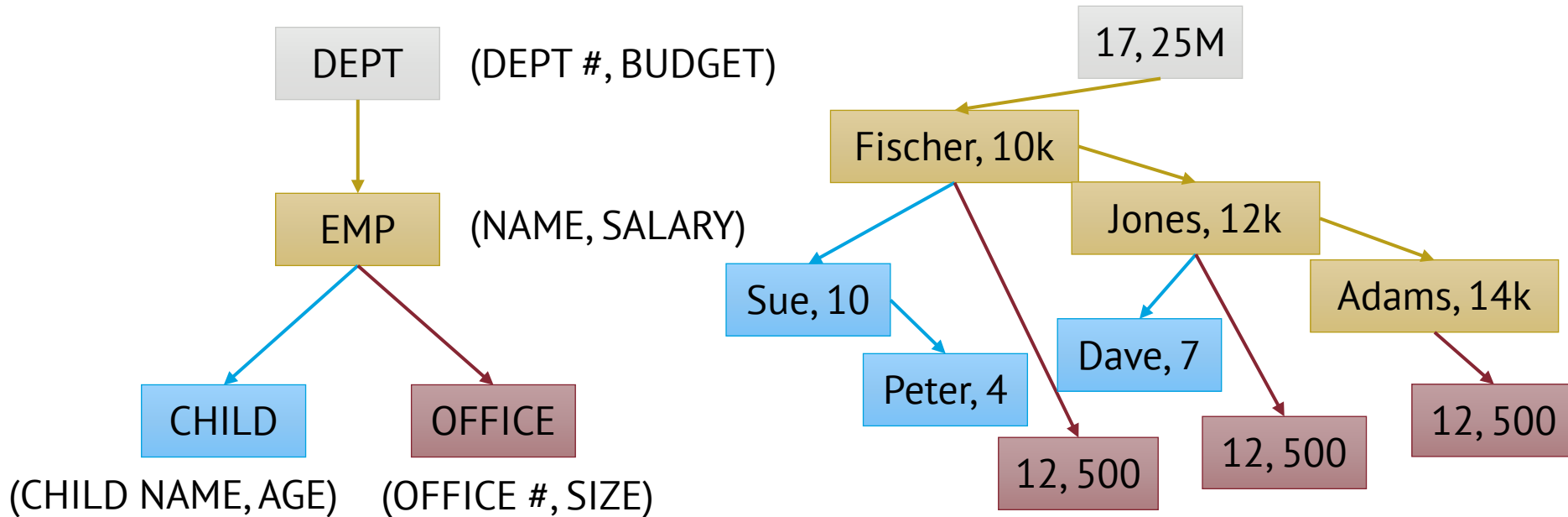
- 1950-es és korai 1960-as évek:
  - Lyukkártya
  - Mágnesszalag
- 1960-as évek vége és 1970-es évek:
  - Merevlemez – közvetlen adatelérés



IBM 3340 „Winchester”  
1974, 35/70MB



# HIERARCHIKUS MODELL

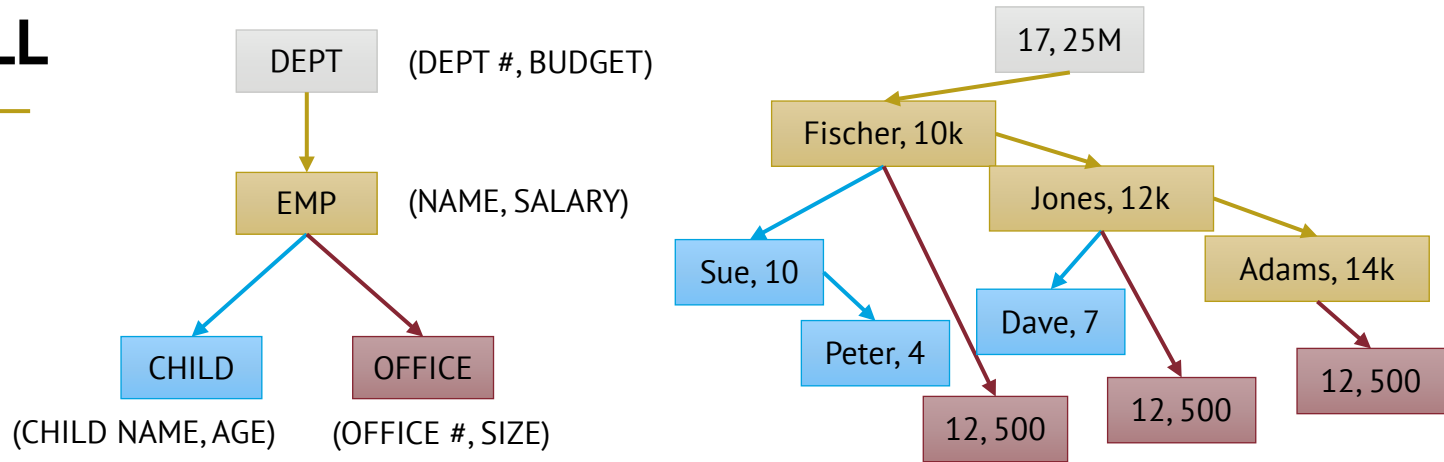


- IMS (IBM Information Management System)
  - Az Apollo programban használt űrhajók és Saturn V leltárhoz készítették
  - Még mindig használják!

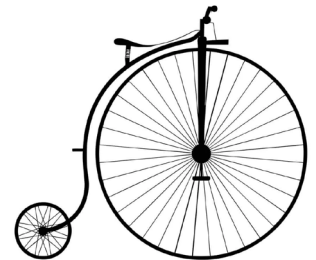




# HIERARCHIKUS MODELL



```
# find names of all employees in department 17
FIND DEPT RECORD WHERE DEPT# = 17
  if failure; return "no such department"
FIND 1st SON OF CURRENT RECORD
  if failure; return "no employee in this department"
LOOP
  save name
  FIND NEXT BROTHER OF THE CURRENT RECORD WHICH IS OF SAME TYPE
  GO TO LOOP
```



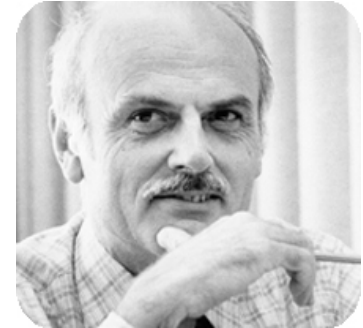
Útvonal alapú. Túlbonyolított és rugalmatlan!

A keresések nem optimalizálhatók!

(Tanuljunk a történelemből: Objektumorientált adatbázisok – rossz irány)

# MODERN ADATBÁZIS RENDSZEREK

- Ted Codd, relációs adatmodell, 1970 IBM
  - ACM Turing Díj
  - IBM Research - System R prototípus
  - UC Berkeley - Ingres prototípus
- Relációs algebra, relációs adatbázisok
  - Az adatokat „relációk” (táblák) tárolják
  - A relációkon működő relációs operátorok relációkat adnak eredményül
- Az első termék:
  - Oracle, 1979



| ID    | name       | dept_name  | salary |
|-------|------------|------------|--------|
| 22222 | Einstein   | Physics    | 95000  |
| 12121 | Wu         | Finance    | 90000  |
| 32343 | El Said    | History    | 60000  |
| 45565 | Katz       | Comp. Sci. | 75000  |
| 98345 | Kim        | Elec. Eng. | 80000  |
| 76766 | Crick      | Biology    | 72000  |
| 10101 | Srinivasan | Comp. Sci. | 65000  |
| 58583 | Califieri  | History    | 62000  |
| 83821 | Brandt     | Comp. Sci. | 92000  |
| 15151 | Mozart     | Music      | 40000  |
| 33456 | Gold       | Physics    | 87000  |
| 76543 | Singh      | Finance    | 80000  |

(a) The instructor table

| dept_name  | building | budget |
|------------|----------|--------|
| Comp. Sci. | Taylor   | 100000 |
| Biology    | Watson   | 90000  |
| Elec. Eng. | Taylor   | 85000  |
| Music      | Packard  | 80000  |
| Finance    | Painter  | 120000 |
| History    | Painter  | 50000  |
| Physics    | Watson   | 70000  |

(b) The department table



# FELHASZNÁLÁSI TERÜLETEK: ADATBÁZISOK MINDENHOL

---

- Bankolás: tranzakciók
- Vasút/Légitársaságok: menetrend, foglalások
- Telekommunikáció
- Egyetemek: regisztráció, jegyek
- Értékesítés: ügyfelek, termékek, vásárlások
- Gyártás: folyamatok, leltár, rendelések, beszerzési láncok
- HR: alkalmazottak adatai, könyvelés, adózás
- ....

# SZABVÁNYOSÍTÁS

---

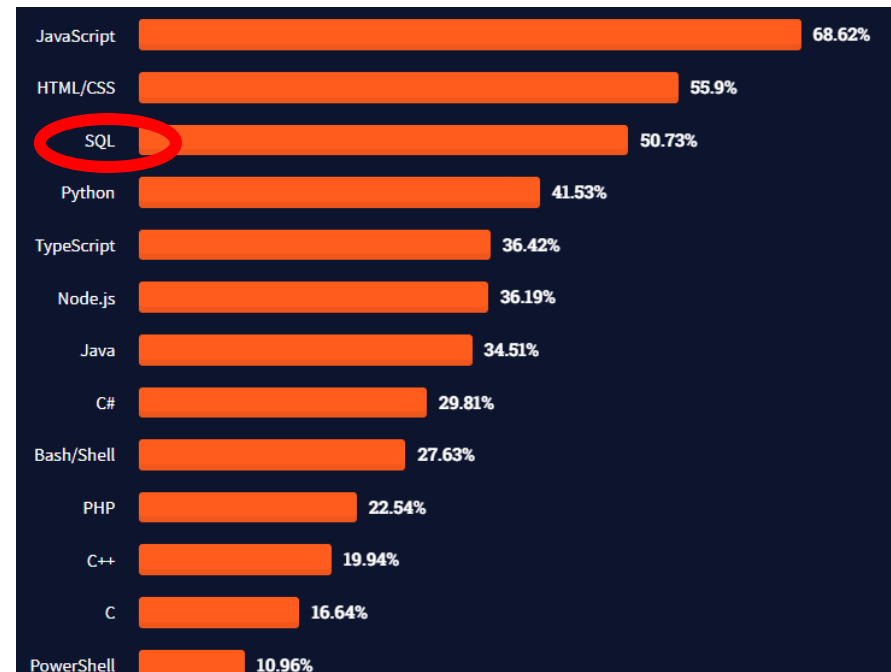


A Ford T-modellben a bal oldali kar állítja a hátsó kerék rögzítőfékjét, és üresbe állítja a sebességváltót. A jobb oldali kar vezérli a gázadást. A kormányoszlop bal oldalán található kar a gyújtás időzítését módosítja. A bal lábpedál váltja a két előremeneti fokozatot, míg a középső pedál vezérli a hátramenetet. A jobb pedál a fék.

# SQL (STRUCTURED QUERY LANGUAGE)

- Szabványok ANSI: 1986, ISO 1987
  - Akkoriban a legtöbbet használt programozási nyelvek:  
Ada, C, Pascal, Lisp, Fortran

- Stackoverflow:  
Developer Survey, 2021  
Most Popular Technologies  
<https://insights.stackoverflow.com/survey/2021#overview>
- Különböző SQL nyelvjárások  
(Oracle, MS SQL Server,  
PostgreSQL,...)



## 1990-ES ÉVEK ...

---

- Nagy több terrabájtos adattárházak, döntéstámogató és adatbányász eszközök
- Internet ...
- Webkereskedelem elterjedése
- (Objektum orientált) és objektum-relációs adatbázisok
  - Földrajzi adatok
  - (Multi-)média adatok



## 2000-ES ÉVEK



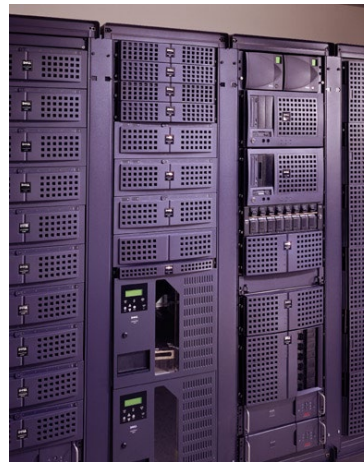
Vásárlások



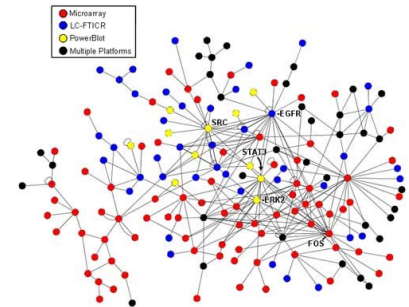
Pénzügyi tranzakciók



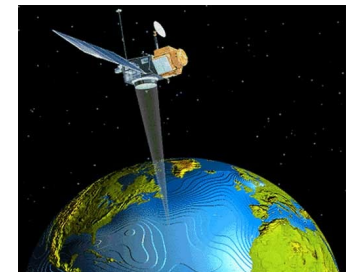
- Szöveg, kép, multimédia adatok
- Földrajzi adatok
- Gráfok
- Idősorok



Telekommunikáció  
Szenzorok



Biológiai adatok



Műholdak  
Távérzékelés



WWW  
Média  
Közösségi média

# INFORMÁCIÓ TÁROLÁS

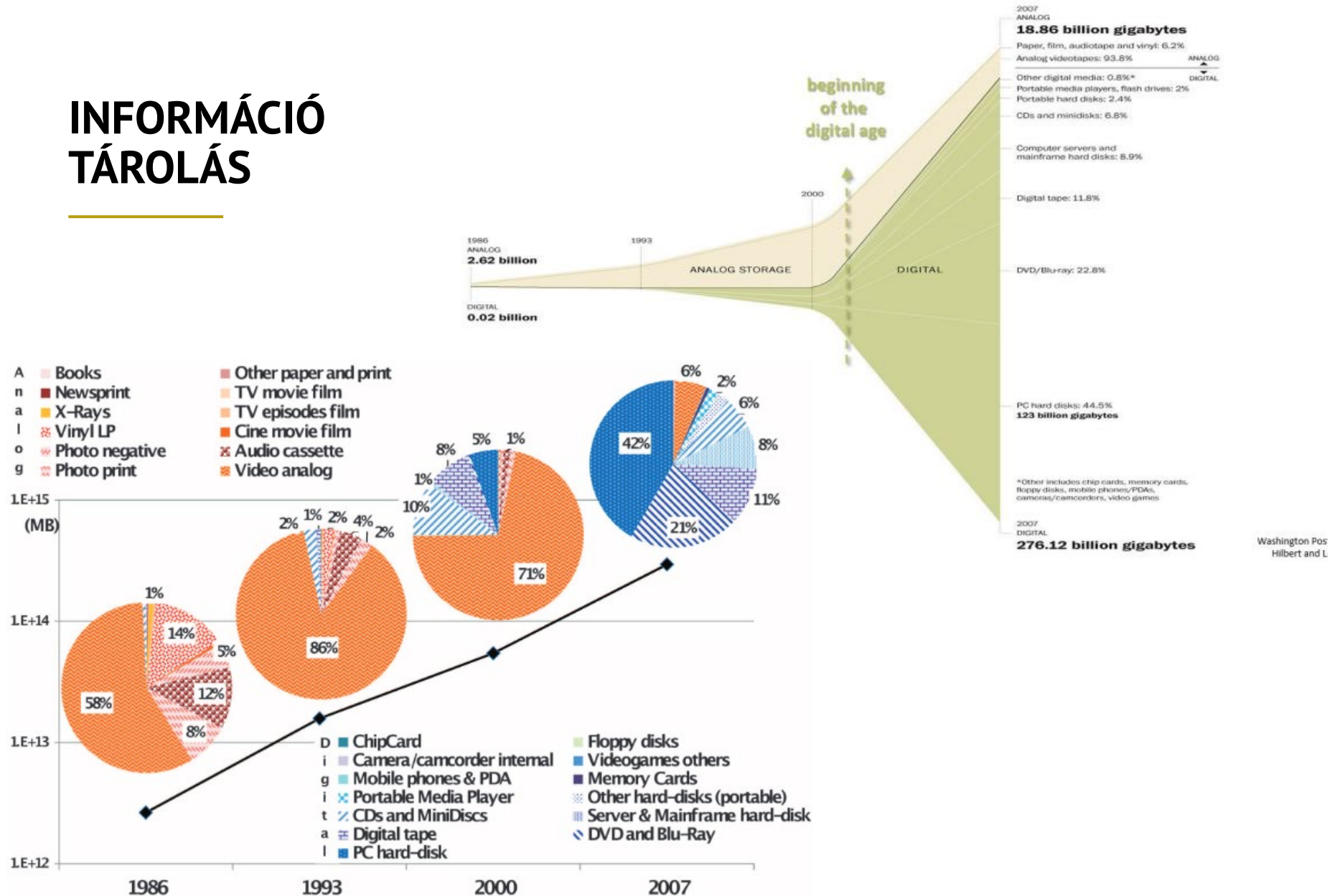
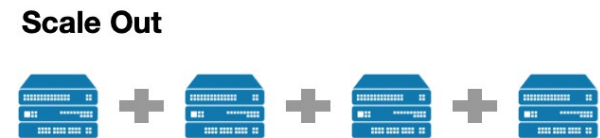
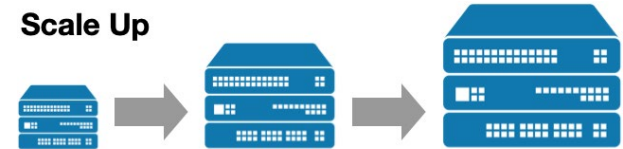


Fig. 2. World's technological installed capacity to store information (table SA1) (16).

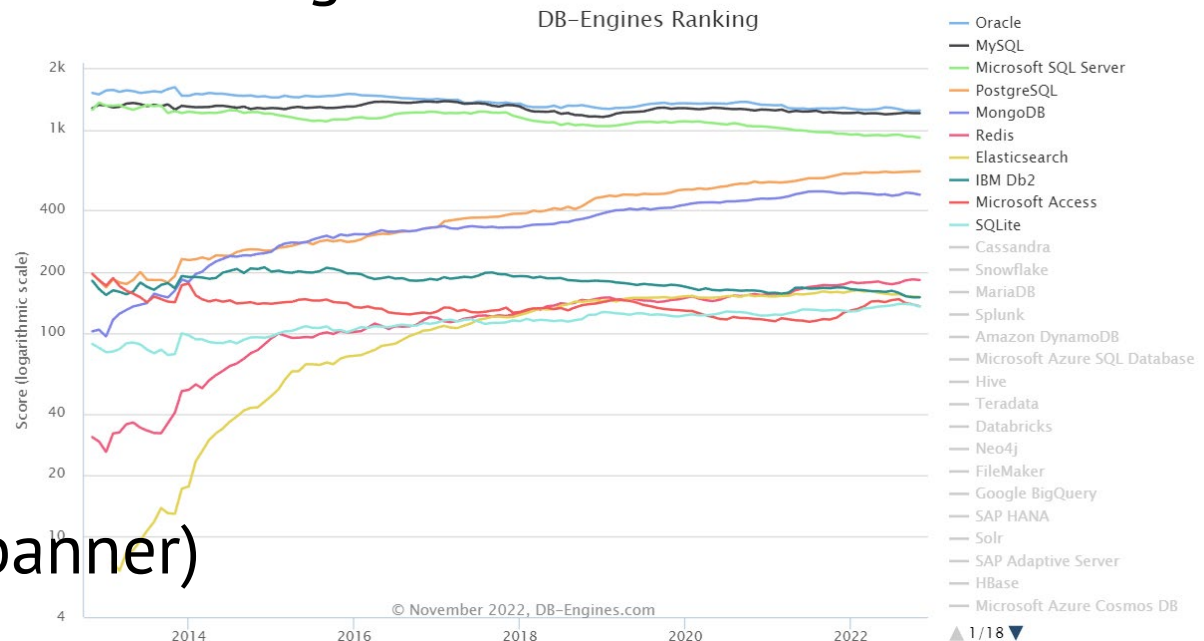


# BIG DATA -2000-ES ÉVEKTŐL

- Big data
  - Big volume, big velocity, big variety
- Scale out/horizontal scaling



- NoSQL
  - (Not using SQL)
  - (Not only SQL)
  - Not yet SQL?!
- (Google cloud spanner)

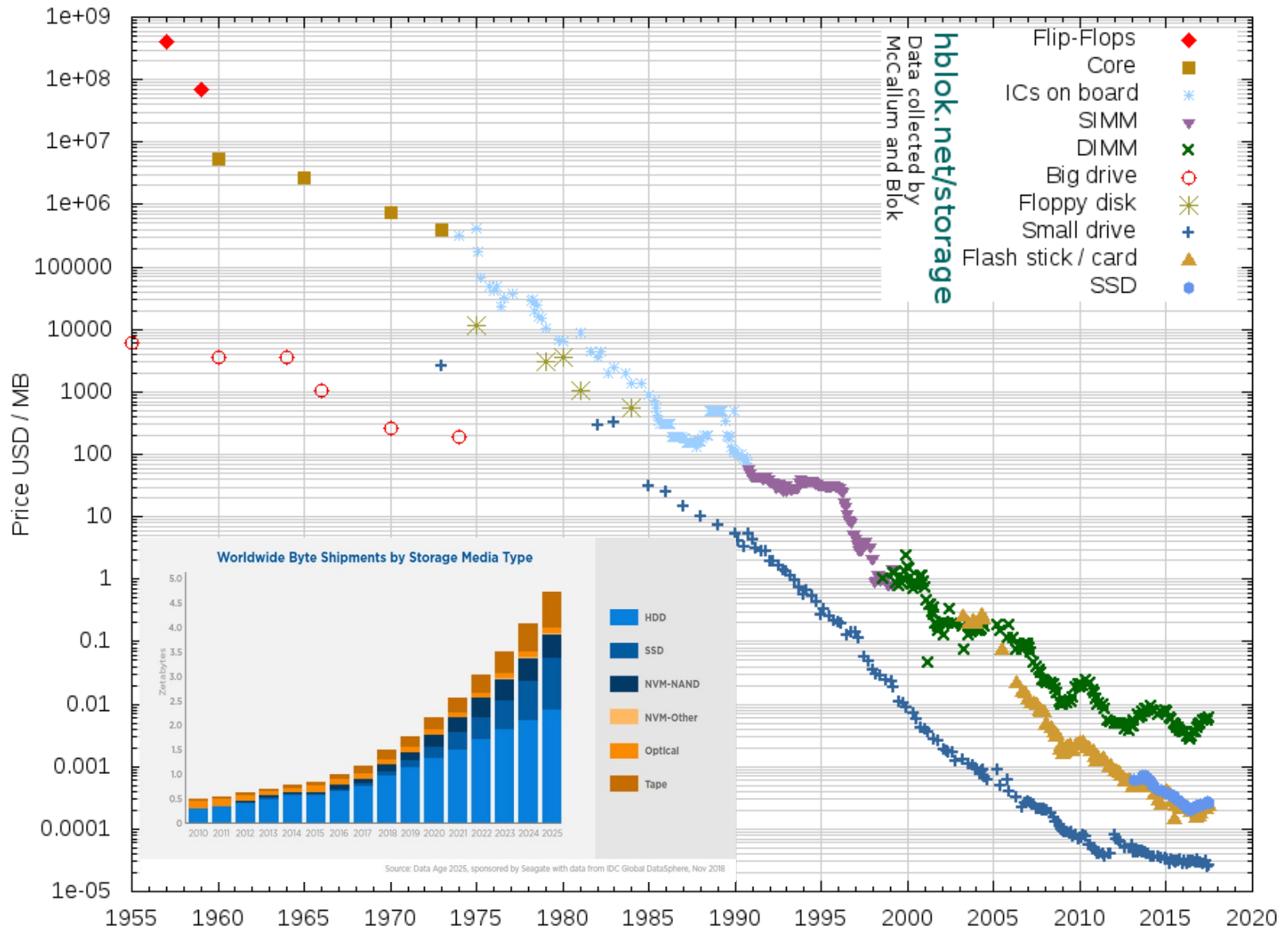


[Spanner \(database\) - Wikipedia](#)

[historical trend of the popularity ranking of database management systems \(db-engines.com\)](#)

[Scaling up vs scaling out your security segmentation | microsegment.io](#)

# Historical Cost of Computer Memory and Storage

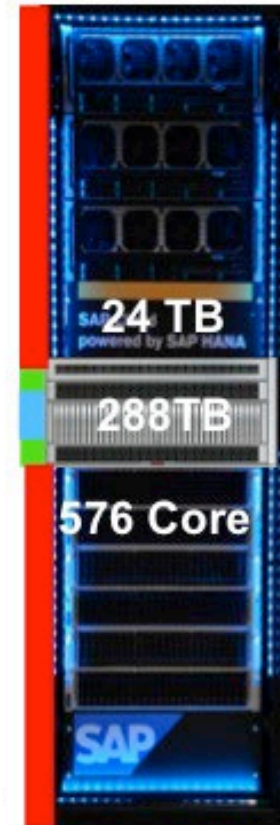


# Exadata vs HANA Cell



Oracle X5-2

|                                                                |                                                                 |
|----------------------------------------------------------------|-----------------------------------------------------------------|
| <b>Compute:</b><br>16 Sockets, 288 Cores<br>6TB DRAM           | <b>Compute:</b><br>32 Sockets, 576 cores<br>24TB DRAM           |
| <b>Network:</b><br>Per Node: 120 Gb/s<br>Bisection BW: 1.4Tb/s | <b>Network:</b><br>Per Node: 464 Gb/s<br>Bisection BW: 2.9 Tb/s |
| <b>Storage:</b><br>224 Cores<br>672 TB HDD or<br>134 TB Flash  | <b>Storage:</b><br>288 TB Shared Flash                          |

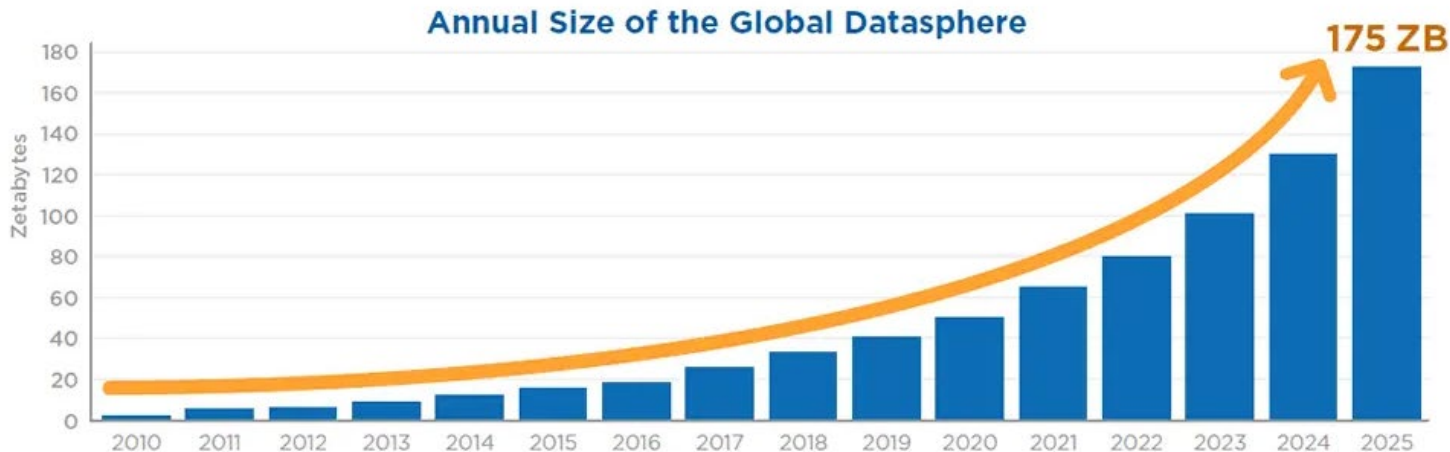


SAP (Cell 2.0)

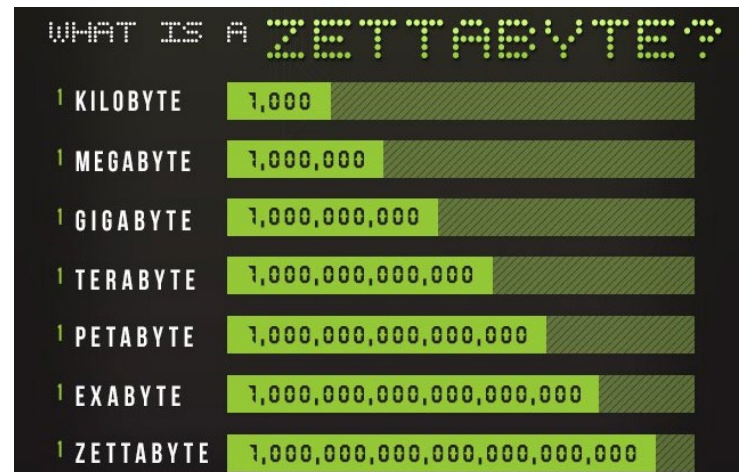


2015

# DIGITÁLIS ADATOK MENNYISÉGÉNEK NÖVEKEDÉSE



- A tárolt adatmennyiség 20 havonta duplázódik



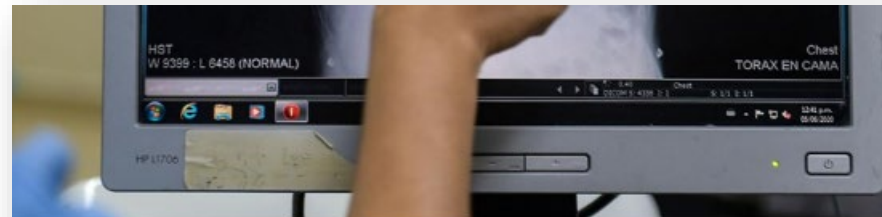
## ADATOK A TERMÉSZETBEN

---

- Az emberi testben tárolt adat mennyisége:  
150 zettabájt
  - A DNS szekvencia: 1.5GB, 10-100 billió ( $10^{12}$ , tera-)sejt
- A megfigyelhető univerzum csillagainak száma:  
 $10^{24}$  (zetta:  $10^{21}$ )



# MESTERSÉGES INTELLIGENCIA



Egy új mesterséges intelligencia 98%-os pontossággal kimutatja a tüdőrákot, tüdőgyulladást, tbc-t

prog.hu

Hírek ·

Legalább olyan jól kódol a mesterséges intelligencia, mint az emberi programozók



hws VÁLLALATI IT DEV ECO@TECH TÁVKÖZLÉS BIZTONSÁG KONZUMER

Szerző: Dömös Zsuzsanna  
2022. NOVEMBER 17. 12:41

FACEBOOK 20 KOMMENT

## A világ egyik vezető autógyártója alá kerül a magyar aiMotive



A Stellantis autógyártó vállalat vásárolja fel egyik legkiemelkedőbb hazai startupunkat. Az aiMotive a jövőben a cég alá tagozódva, leányvállalatként működik tovább.

Megszüntetheti a dugókat a mesterséges intelligencia, Amerikában már kísérleteznek vele

index VB 2022  
BELFÖLD KÜLFÖLD GAZDASÁG KULT VÉLEMÉNY TECH-TUD SPORT FOMO 24 ÓRA BLOG VIDÉO PODCAST  
Qatar 2022  
MENETREND

Munkaerő keresés robotokkal? Új korszak jön a cégek számára



hws VÁLLALATI IT DEV ECO@TECH TÁVKÖZLÉS BIZ



Szerző: Habók Lilla  
2022. NOVEMBER 21. 10:10

FACEBOOK 3 KO

## Gépi tanulás azonosított egy Renoir-művet

ORIGO  
ITTHON NAGYVILÁG PÉNZ SPORT TV FILM TUDOMÁNY TECH AUTÓ KULTÚRA TÁJÉLSPICC MÉG TÖBB

TUDOMÁNY

Több ezer kóros fehérjét talált a mesterséges intelligencia

2022. november. 19. 12:03 · TECH

Ártatlanul lett megvádolva egy 93 millió éves őslény, most a mesterséges intelligencia tisztázta



# PÉLDA: PRECÍZIÓS GAZDÁLKODÁS

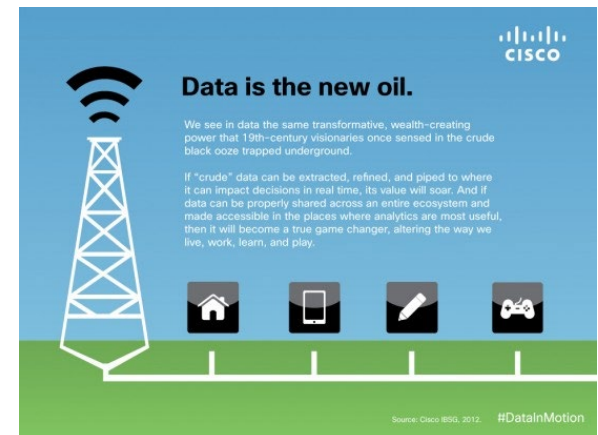
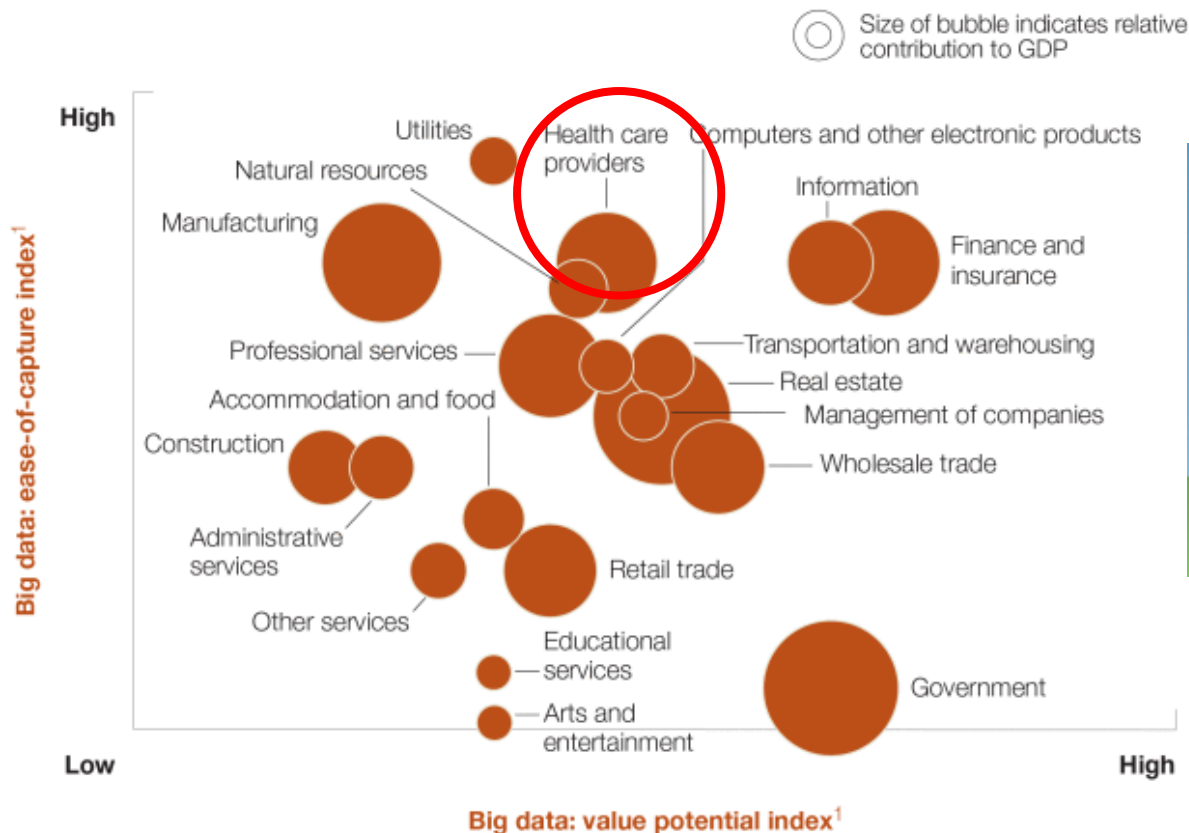
---



- A hozam maximalizálása a felhasznált erőforrások csökkentése mellett
- Mérések
  - Terményátlag, terület jellemzői, szervesanyag tartalom, páratartalom, nitrogén, pH,...
- Adatgyűjtés
  - Traktorok, kombájnok,...; drónok; műholdak
- Távérzékelés adatok
  - Hetente 2x az öntözéshez
  - Havonta 2x a növénykár felméréshez
  - 2-5m<sup>2</sup> felbontás



# BIG DATA LEHETŐSÉGEK



<sup>1</sup>For detailed explication of metrics, see appendix in McKinsey Global Institute full report *Big data: The next frontier for innovation, competition, and productivity*, available free of charge online at [mckinsey.com/mgi](http://mckinsey.com/mgi).

Source: US Bureau of Labor Statistics; McKinsey Global Institute analysis

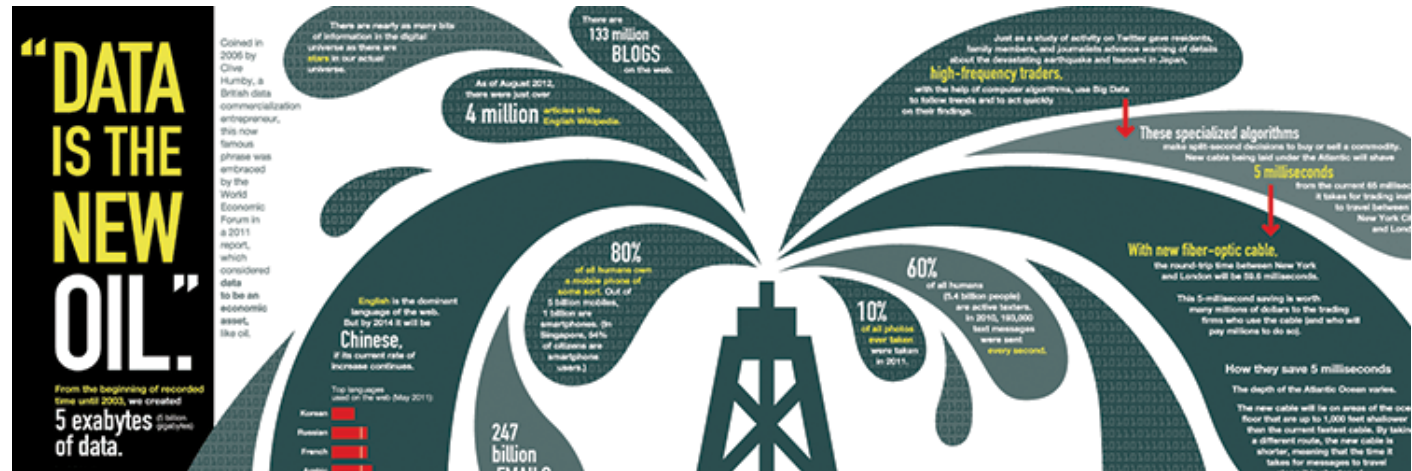
**Personalized medicine – big data**

<https://www.google.com/search?q=personalized+medicine+big+data>

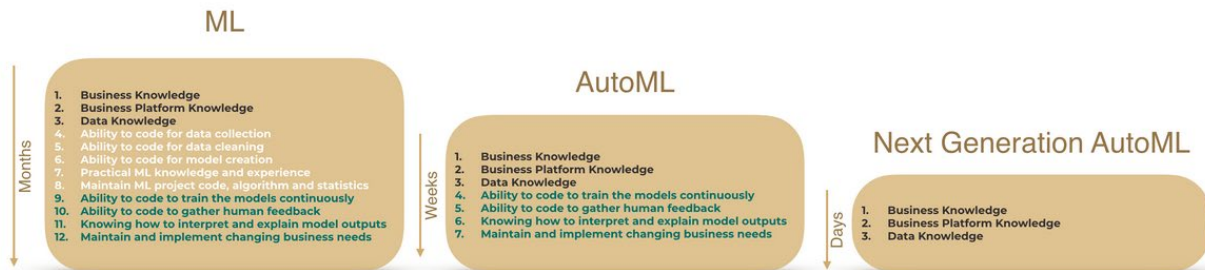
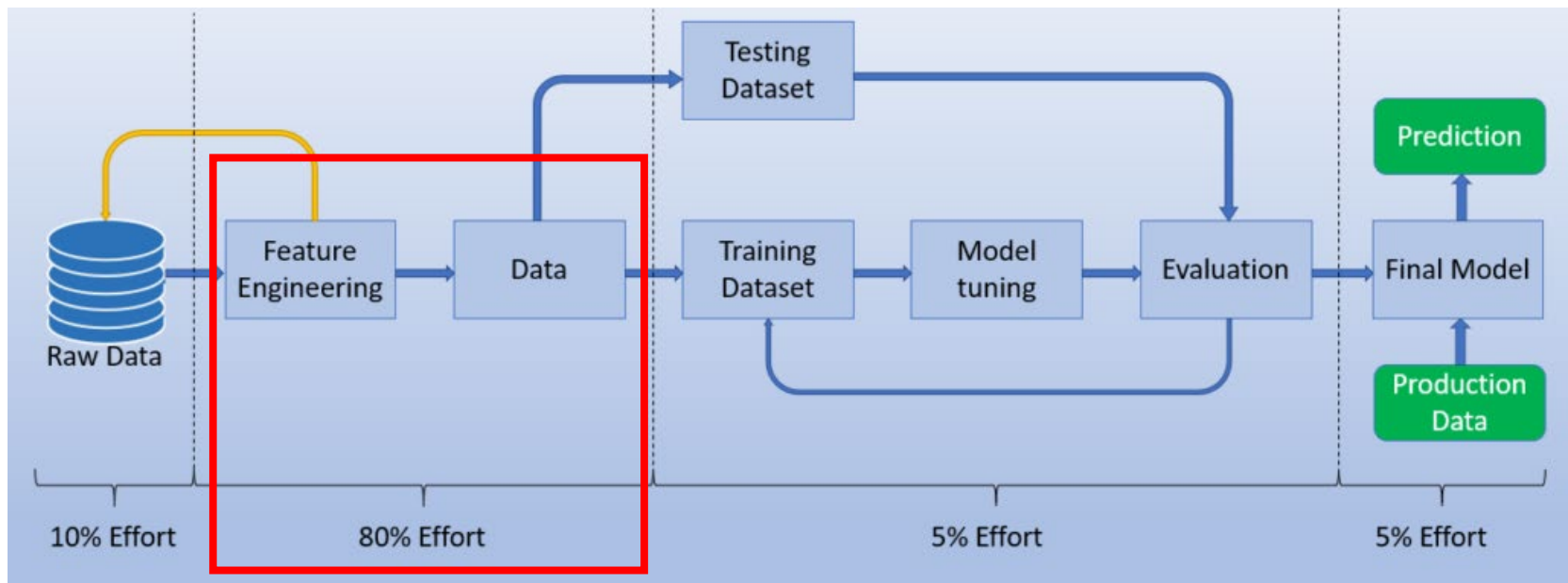
# ADAT AZ ÚJ OLAJ

- A mesterséges intelligencia csak annyit ér, amennyi/amilyen adat mögötte van  
„Artificial intelligence is only as good as the data it crunches.”

## Why Data Is The New Oil, Fortune, 11.07.2016.



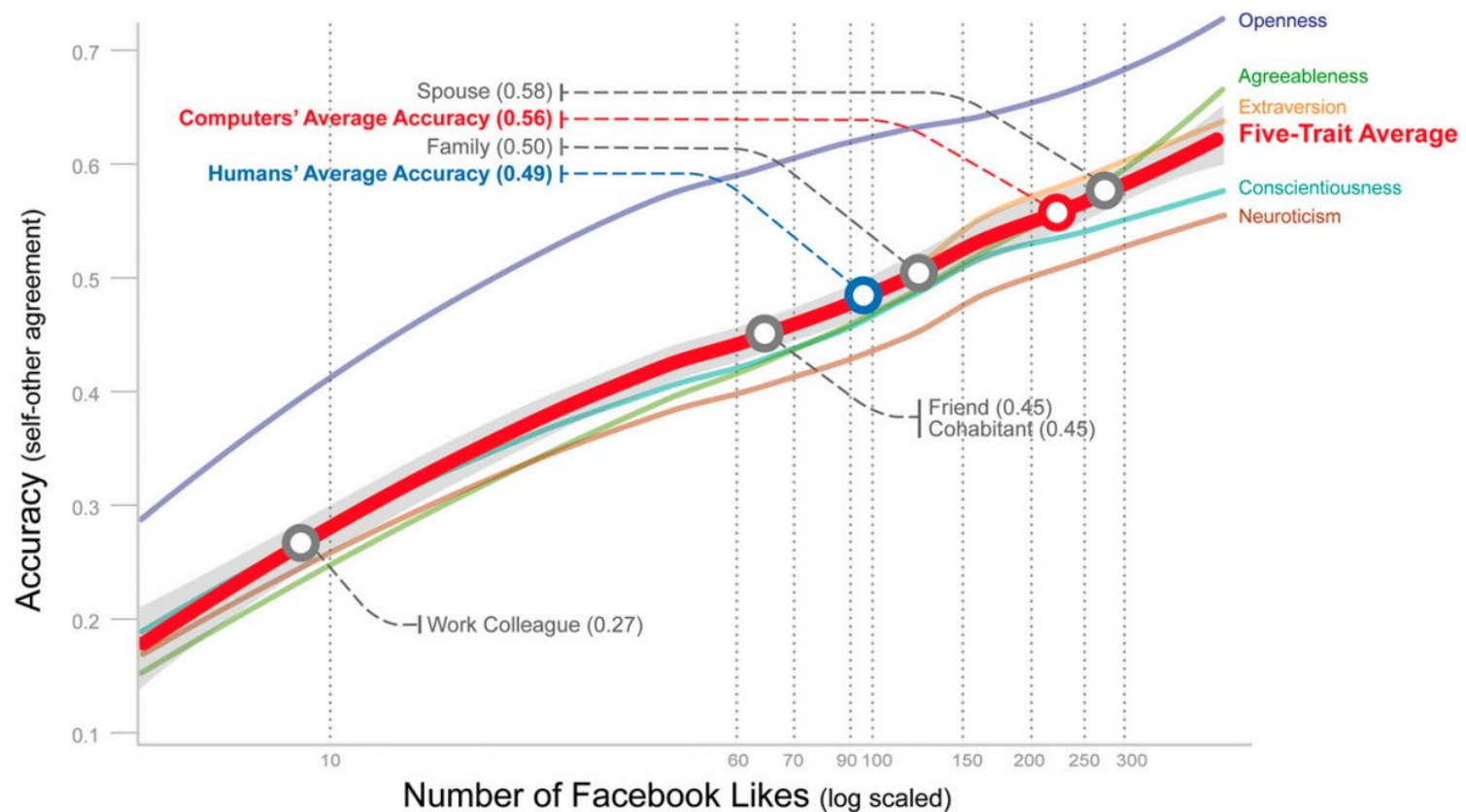
# ADATMÉRNÖKSÉG VS. GÉPI TANULÁS



Just know your business and data  
For the rest : **Next Generation AutoML**



# FIGYELMEZTETÉS: A SZÁMÍTÓGÉP ALAPÚ SZEMÉLYISÉGJEGY MEGHATÁROZÁS PONTOSABB, MINT AZ EMBERI



[Computer-based personality judgments are more accurate than those made by humans \(pnas.org\)](https://www.pnas.org)



**TED**

SIGN IN



3,246,909 views | Tristan Harris • TED2017

♡ Like (97K)

↑ Share

≡+ Add

How a handful of tech companies control billions of minds every day

[Tristan Harris: How a handful of tech companies control billions of minds every day | TED Talk](#)

## ÖSSZEFOGLALVA

---

- Az adattudományok az informatika egyik meghatározó területe
- Keresett a munkaerőpiacon
- „Állítólag nem léteznek még azok a szakmák, amelyek a most iskolába járó gyerekek számára a legvonzóbbak lesznek felnőtt korukban.”
- Adatkezelés fejlődése – lyukkártyától RDBMS-ig
- Big data – az adat az új olaj
- Gépi tanulás – jó adat kell hozzá



# TOVÁBBI HIVATKOZÁSOK

## *Munkaerőpiac*

---

- [A 20 kötelező adattudós készség az adattudományi állások megszerzéséhez \(ciksiti.com\)](#)
- [Adatelemző állás, munka - 142 állásajánlat | Profession](#)
- [\(2\) data analyst Jobs in budapest | LinkedIn](#)
- [HVG JOBLINE.HU](#)
- [Toborzás az adattudományok területén | Hays Technology](#)
- [5 Trends in the Database Job Market | Datamation](#)
- [Database Job Market Trends to Watch – ValiantCEO](#)
- [Data Analyst Salaries Around the World \(Based on Industries, Countries and Experience\) \(simplilearn.com\)](#)
- [A Complete Guide to Business Analyst Salaries in 2022 | DataCamp](#)