

A Big Data elemzéssel végső soron jobb és gyorsabb döntéshozatalt, modellezést és a jövőbeli eredmények előrejelzését, valamint továbbfejlesztett üzleti intelligenciát biztosíthat. A vállalkozások nagy mennyiségű adathoz férhetnek hozzá, és sokféle adatforrást elemezhetnek, hogy új információkhoz jussanak, és akciókat hajthatnak végre.

Az érzékelőkből, eszközökből, videókból, naplókából, tranzakciós alkalmazásokból, webről és közösségi médiából származó adatok elemzése lehetővé teszi a szervezet számára, hogy adatvezérelt legyen.

Big Data típusai

There A Big datá-nak három típusa van: a strukturált, a félig-strukturált és a strukturálatlan.

Strukturált adatok: Minden olyan adat, amely rögzített formátumban tárolható, elérhető és feldolgozható. Napjainkban az ilyen adatok mérete hatalmas mértékben növekszik, a tipikus méretek több zéta byte tartományban vannak (egymilliárd terra byte egy zéta byte alkot).

Például: Az egyik példát a relációsadatbázis-kezelő rendszerben tárolt "strukturált" adatok jelentik. Adatok a táblázatokban, pl. menutrend, fizetési táblázatok, kulcsszavak listája a kiadvány végén stb.

"Alkalmazottak" táblája szintén egy példa a strukturált adatokra.

Strukturálatlan adatok: Minden ismeretlen formájú vagy szerkezetű adat strukturálatlan adatként van besorolva. Amellett, hogy a méret hatalmas, a strukturálatlan adatok számos kihívást jelentenek a feldolgozásuk szempontjából, hogy értéket nyerjenek belőlük.

Például: Egy heterogén adat forrás text file-k, videók, képek stb. kombinációját tartalmazza

Manapság a szervezetek rengeteg strukturálatlan adattal rendelkeznek, érdekes információkkal, de nem tudják, hogyan lehet a rejtett értékeket kinyerni belőlük, mivel ezek az adatok nyers vagy strukturálatlan formában vannak jelen.

Félig – strukturált adatok: az adatok mindkét formáját tartalmazhatják.

Például: Személyes weboldalak: szövegeket, képeket, videókat, közösségi hálózatokra mutató linkeket, barátokkal való kommunikációra szolgáló linkeket stb. tartalmaznak. Egy másik példa az online boltok weboldalai reklámmal, fizetési terminálokkal.

A Big Data jellemzőit általában a **Big Data négy kötetének** nevezik.

Az **adatmennyiség** az elemzendő és feldolgozandó adatkészletek méretére utal, amelyek ma már gyakran nagyobbak, mint a terabytes¹ and petabytes². Az adatok pusztán mennyisége eltérő és eltérő feldolgozási technológiákat igényel, mint a hagyományos tárolási és feldolgozási képességek. Ez azt

¹ Terabyte: Egy bit (bináris számjegy) egyetlen bináris értéke 0 vagy 1, a számítógépes rendszer legkisebb adategysége. A terabyte az egyik legnagyobb adathordozó-egység, amelyet a piacon lévő termékek ma használnak. Vannak olyan egységek, amelyek nagyobbak, mint egy terabyte: petabyte, exabyte, zettabyte, yottabyte és brontobyte. A geopbajt 1030 bájtra utal. A gyakorlatban egy terabajt adat egyenértékű: 728.177 hajlékonylemez, 40 egyrétegű Blu-ray lemez, 85.899.345 oldal Word dokumentum, 500 órányi film, 310 000 fénykép, 17 000 órányi zene..

² Petabyte: 1 PB = 1,000 TB, pl. 1 PB több mint 1 kvadrillió bájtnak felel meg.



ITFARM

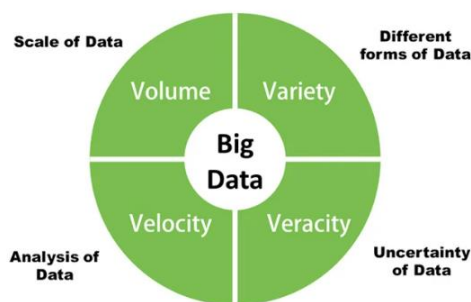
jelenti, hogy a Big Data adatkészletei túl nagyok ahhoz, hogy egy normál lappal vagy asztali processzorral feldolgozzuk őket.

A **Big Data sebessége** az adatok létrehozásának sebességére utal. A nagy sebességű adatok olyan ütemben jönnek létre, hogy különböző (elosztott) feldolgozási technikákat igényelnek. A nagy sebességgel generált adatokra példa lehet a Twitter-üzenetek vagy a Facebook-kommunikáció.

A **Big Data változatossága** teszi igazán nagygyá a Big Data-t. A Big Data sokféle forrásból származik, és általában a három típus egyike: strukturált, félig-strukturált és strukturálatlan adatok. Az adattípusok sokfélesége gyakran különböző feldolgozási képességeket és speciális algoritmusokat igényel. A nagy változatosságú adatkészletekre példa lehetnek az egy város különböző pontjain létrehozott audio- és videófájlok.

A **Big Data hitelessége** az elemzett adatok minőségére utal. A nagy hitelességű adatok számos olyan rekordot tartalmaznak, amelyek értékesek az elemzéshez, és amelyek jelentős mértékben hozzájárulnak az átfogó eredményekhez. Az alacsony hitelességű adatok viszont nagy százalékban tartalmaznak értelmetlen adatokat. Az ezekben az adatkészletekben nem értékeset zajnak nevezzük. A nagy hitelességű adatkészletre példa lehet egy orvosi kísérletből vagy vizsgálatból származó adathalmaz.

A nagy mennyiségű, nagy sebességű és változatos adatokat fejlett eszközökkel (elemzésekkel és algoritmusokkal) kell feldolgozni, hogy értelmes információkat tárjanak fel. Az adatok ezen jellemzői miatt az adatkészletek tárolásával, feldolgozásával és elemzésével foglalkozó tudástartományt Big Data-nak nevezték el.



A Big Data feldolgozásának előnyei

- A vállalkozások külső intelligenciát használhatnak a döntések meghozatala során.
- A keresőmotorok és az olyan webhelyek közösségi adataihoz való hozzáférés, mint a Facebook, a Twitter, lehetővé teszi a szervezetek számára, hogy finomra hangolják üzleti stratégiáikat.
- Továbbfejlesztett ügyfélszolgálat.
- A terméket/szolgáltatást érintő kockázatok korai azonosítása, ha van ilyen.
- Jobb működési hatékonyság.
- A Big Data technológiák adattárház létrehozására használhatók.

Összegzés

A Big Data egy olyan kifejezés, amelyet olyan adatgyűjtemény leírására használják, amely hatalmas méretű, és mégis idővel exponenciálisan növekszik. A big data lehet 1) strukturált, 2) strukturálatlan, 3) félig-strukturált. A Big Data sebessége az adatok létrehozásának sebességére utal. A nagy sebességű adatok olyan ütemben jönnek létre, hogy külön feldolgozási technikákat igényelnek. A big





ITFARM

data változatossága sokféle forrásból származik, és általában a három típus egyike: strukturált, félig-strukturált és strukturálatlan adatok. Az adattípusok sokféle gyakran különböző feldolgozási képességeket és speciális algoritmusokat igényel. A Big Data hitelessége az elemzett adatok minőségére utal: különbséget teszünk a nagy hitelességű adatok és az alacsony hitelességű adatok között. A zaj, a változatosság, a sebesség és a változékonyság négy Big Data jellemző. A Big Data néhány előnye jobb ügyfélszolgálat, a jobb működési hatékonyság, a jobb döntéshozatal.

Linkek és releváns témák:

https://en.wikipedia.org/wiki/Big_data

<https://www.guru99.com/what-is-big-data.html>

<https://www.gartner.com/en/information-technology/glossary/big-data>

Kulcs szavak

mennyiségek

karakterek

szimbólumok

Big Data

mesterséges intelligencia

mobil eszközök

közösségi média

Dolgok internete

adat vezérelt

strukturált adatok

strukturálatlan adatok

félig strukturált adatok

Big Data négy kötete

Big Data sebessége

Big Data sokfélesége

Big Data valóságtartalma



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Project: Erasmus+ KA220-ADU, Duration: since 01-01-2022 till 30 01-07-2024



**Co-funded by
the European Union**