

Oracle Big Data koncepció

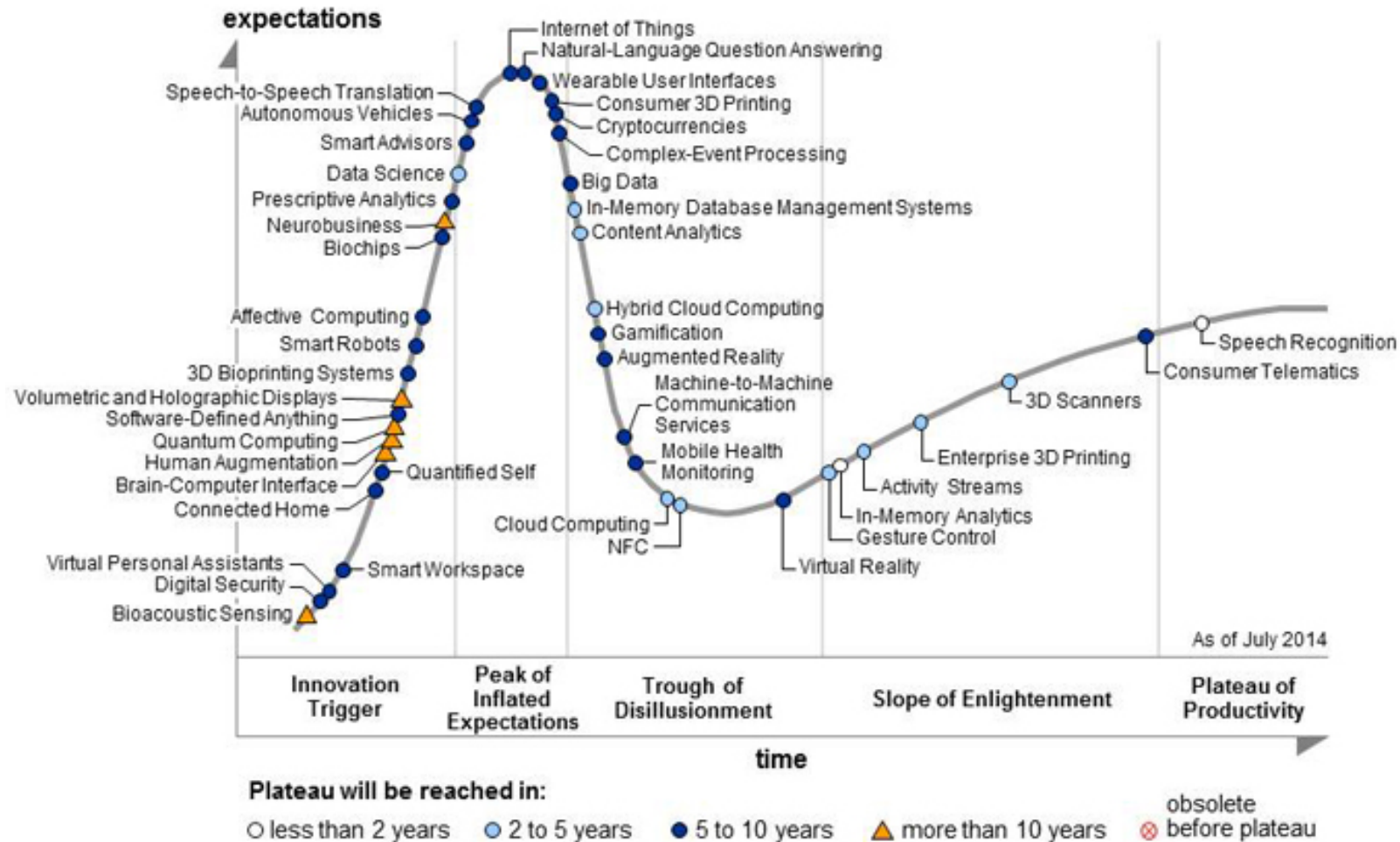
Stadler Gellért
Vezető tanácsadó
Oracle Consulting
HTE 2015 Konferencia

ORACLE®

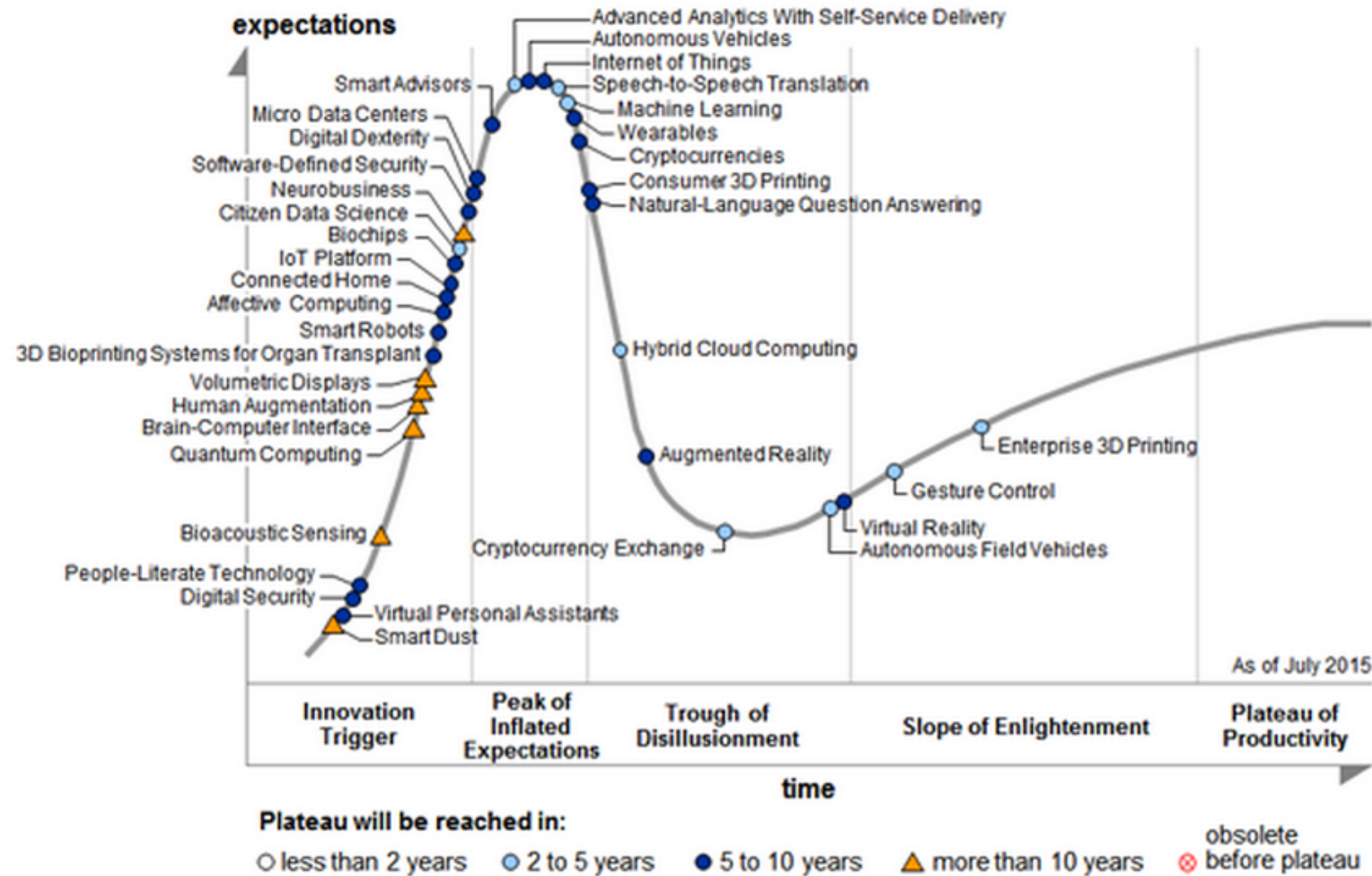
Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Gartner Hype Cycle for Emerging Technologies, 2014



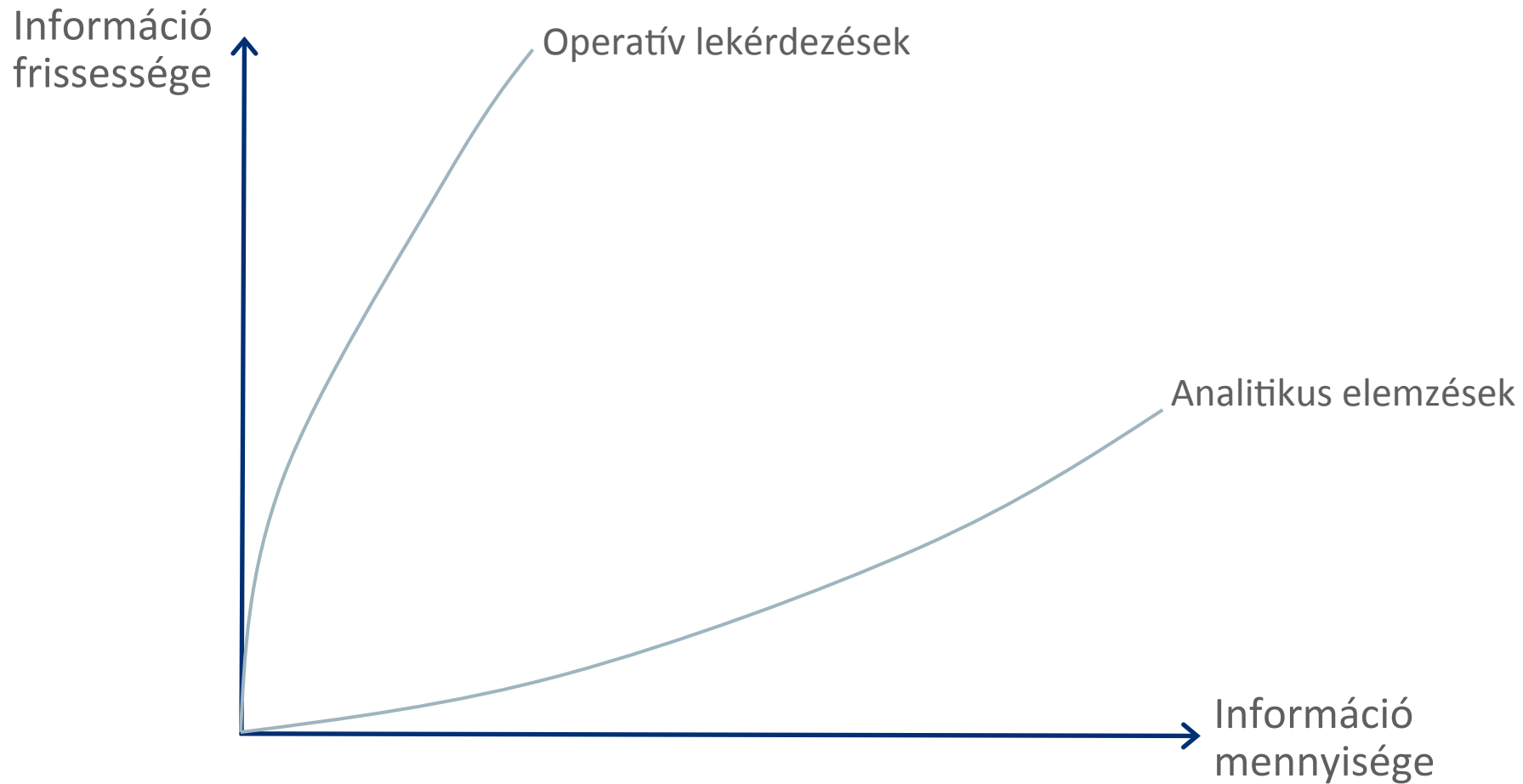
Gartner Hype Cycle for Emerging Technologies, 2015



Big Data helyett

- 1 ➤ Machine Learning
- 2 ➤ Advanced analytics With Self-Service Delivery
- 3 ➤ Smart Advisors
- 4 ➤ Internet of Things

Versenylőny az üzleti elemzésben



Hol volt a versenyelőny az üzleti riporting területén?

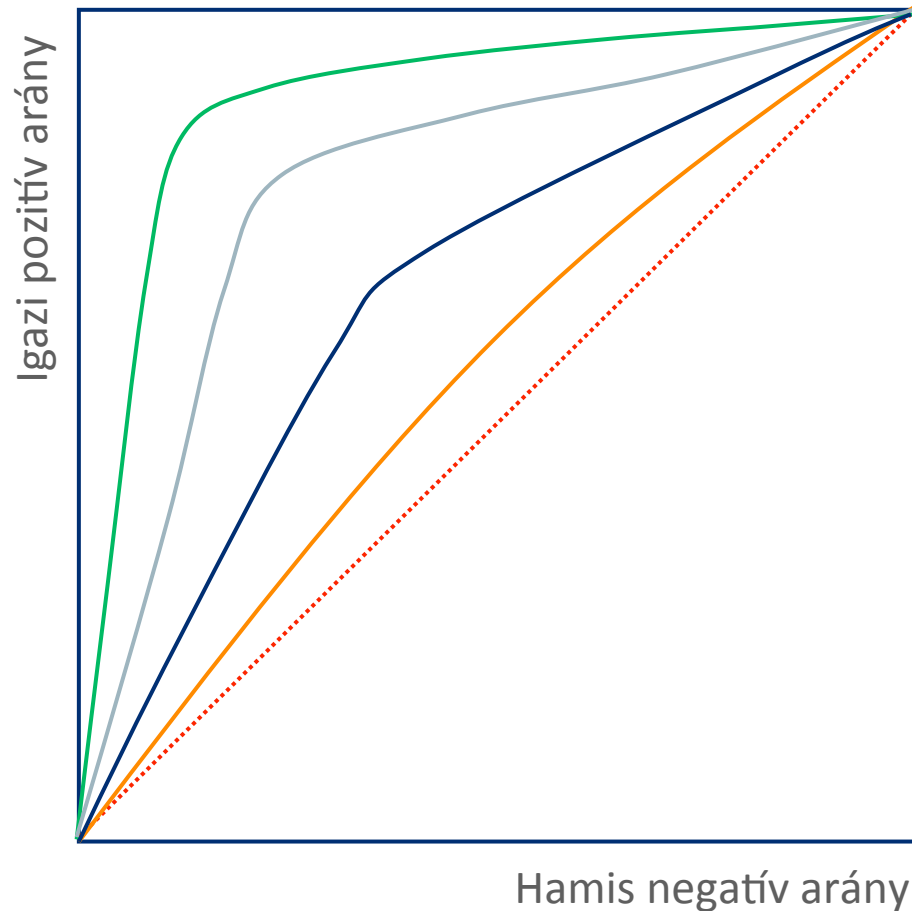
- 1 ➤ Eladási adatok, Költség adatok
- 2 ➤ Eladási adatok + Költség adatok
- 3 ➤ Egyéb adatok
- 4 ➤ Egyéb adatok bevonása
- 5 ➤ Adatok rendelkezésre állása: havi, heti, napi (T+1)

Mi történt a múltban?

Mi várható a jövőben?

- 1 ➤ Mit fog tenni az ügyfelem?
- 2 ➤ Mi érdekelhetné a jelenlegi termékein kívül?
- 3 ➤ Hogyan növelhetném az elégedettségét?

Machine learning és a Big Data

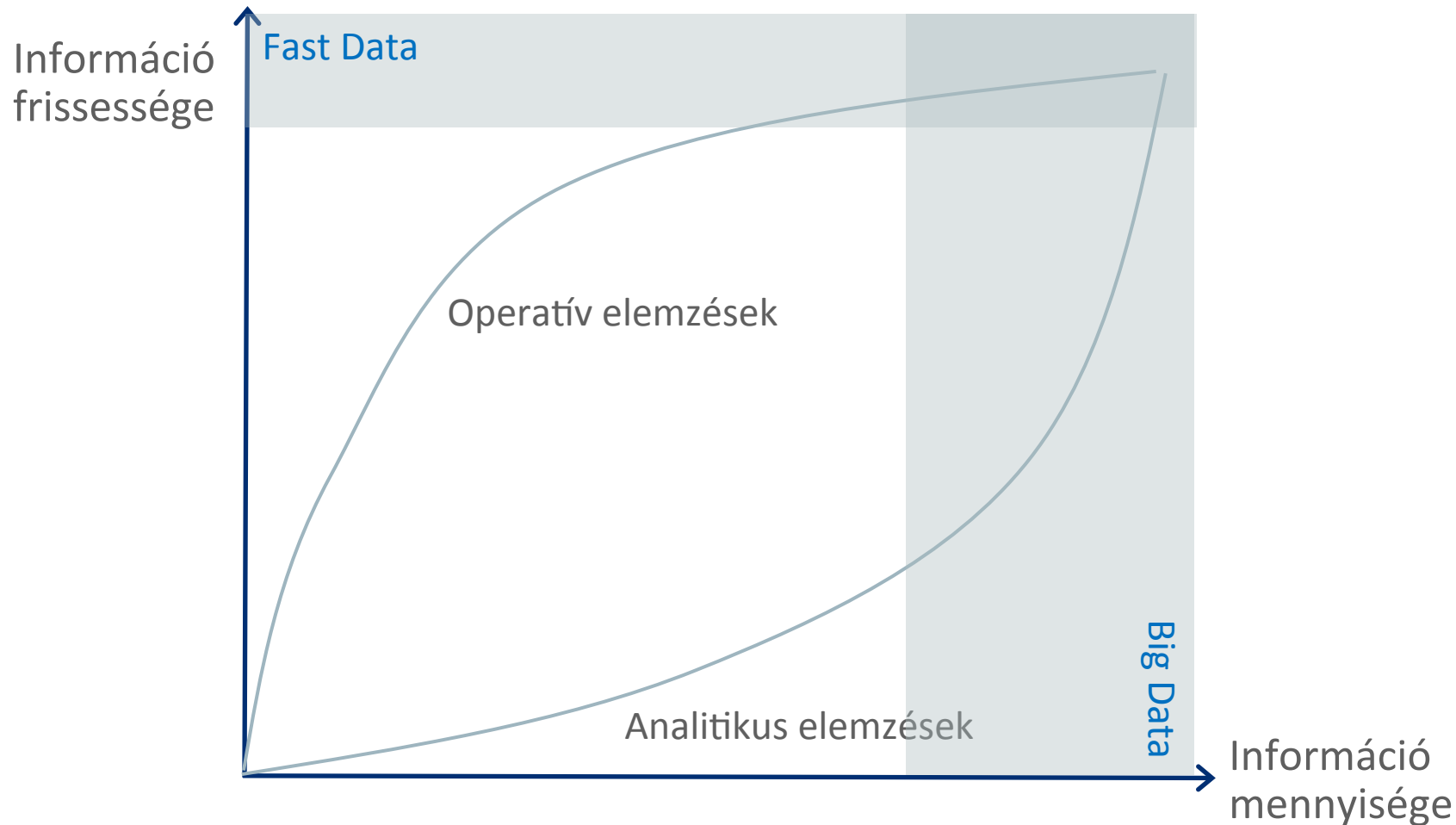


Ügyfél lemorzsolódás elemzés pontossága

- Ügyfél (törzs) adat
- Ügyféladat + pénzügyi adat
- Ügyféladat + pénzügyi adat + hívásadat
- Ügyféladat + pénzügyi adat + hívásadat + logok/szöveges adatok

Minél több adatot vonunk be az elemzésbe, annál pontosabb lesz az eredmény.

Versenylőny az üzleti elemzésben



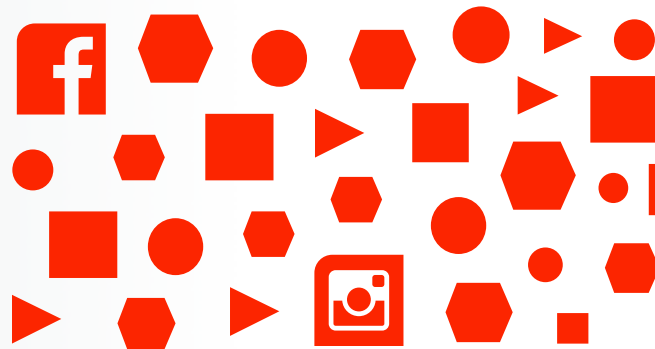
Big Data az üzleti elemzésben

	Relációs tárolás	Big Data
Adatmennyiség	<= Terabytes	>= Terabytes
Struktúra	Kötött	Szabad
Transzformációk	Jellemzően betöltés előtt/ közben	Jellemzően elemzés közben
Tipikusabb felhasználás	Hagyományos reporting	Statisztikai, adatbányászati jellegű elemzések
Tipikusabb felhasználás	Múltbéli adatok jelentése, KPI-ok kalkulációja	Jövőt kutató elemzések, mélyebb összefüggések keresése
Fejlesztési igény	Jellemzően magas	Jelentősen alacsonyabb is lehet

Big Data adatok elemzésének nehézségei

Adat bizonytalanság

- Ismeretlen, lehengetlő mennyiségű adat
- Nem nyilvánvaló a benne lévő érték
- Jelentős átalakítást igényel



Az elemzési idő 80%-a az adatok előkészítésére, értelmezésére megy el

Komplex eszkörendszer

- Szükséges háttértudás jelentős lehet
- BI eszköztámogatás limitált
- Inkább csak céleszközök léteznek, nincs egy eszközben széles funkcionalitás



Gyakran van szükség olyan szakértelmre, ami ritka vagy nehezen elérhető.

Hol van a fő nehézség?

Tárkapacitás

Elemző eszközök

Jól skálázható teljesítmény

Szakértelem

Integráció

Oracle Big Data koncepció

- **Relációs tárolás és a Big Data együttes alkalmazása**
 - Hatékonyság alapján osszuk meg az adatokat a két világ között
 - Szoros integráció (átjárhatóság) a Big Data és relációs adatok között
 - Egységes elemzői felületek biztosítása
- **Ne építsünk magunk Big Data infrastruktúrát**
 - Integráció költségét a szállító viselje
 - Egy szállítóhoz tudjunk fordulni a teljes technológiai stack-en
 - Ne töltsük a bevezetési időt infrastruktúrális feladatokkal
- **Használjunk szállítói tanácsadást a projekteknél is !**

Big Data Appliance X5-2

Sun Oracle X5-2L Servers with per server:

- 2 * **18 Core** Intel Xeon E5 Processors
- **128 GB DDR4-2133** DIMMs Memory (Upgradable to **768GB**)
- 48TB Disk space

Integrated Software (4.1):

- Oracle Linux6.x, Oracle JDK
- Oracle Big Data SQL 1.1*
- Cloudera Distribution of Apache Hadoop – EDH Edition
- Cloudera Manager
- Oracle R Distribution
- Oracle NoSQL Database CE



* Oracle Big Data SQL külön licenszelendő

Buy Your Own vs Build Your Own

ORACLE®

Integrated
Tuned
Optimized
Identical



1 Big Data Appliance
Unpack to production in days

Applications

OS

Compute &
Storage

Networking



DELL™



IBM®

Hadoop Distribution

+

Red Hat / CentOS

+

CPU, RAM, Blade, Rack

+

Cisco

=

120+ separate parts
Months from start to production

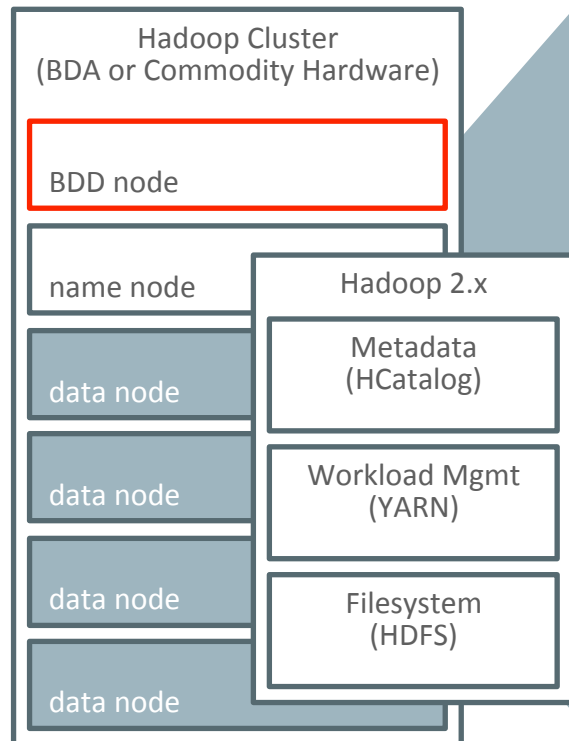
Different
Platform
Every
Time

Oracle Big Data SQL

- A teljes Oracle SQL funkcionalitás támogatott.
- Egységes jogosultság kezelés
- Smartscan funkcionalitás az egyes node-okon:
 - Where feltétel kiértékelés
 - Oszlop projekciók
 - Bloom szűrés join-okhoz
 - JSON parsing, adatbányászati funkciók



Oracle Big Data Discovery



Oracle Big Data Discovery Workloads

Studio

- **Web UI:** Find, Explore, Transform, Discover, Share

ORACLE

In-Memory Discovery Indexes

- **DGraph:** Search, Guided Navigation, Analytics

ORACLE

Data Processing, Workflow & Monitoring

- **Profiling:** catalog entry creation, data type & language detection, schema configuration
- **Sampling:** dgraph (index) file creation
- **Transforms:** >100 functions
- **Enrichments:** location (geo), text (cleanup, sentiment, entity, key-phrase, whitelist tagging)

ORACLE



Self-Service Provisioning & Data Transfer

- **Personal Data:** Upload CSV and XLS to HDFS

ORACLE

Other Hadoop Workloads

MapReduce

Spark

Hive

Pig

Oracle Big Data SQL
(BDA only)

Hogyan kezdünk hozzá- Oracle Konzultációval



Üzlet vezérelt bevezetés



*Innovation
Workshops*



*Data
Reservoir*



*Discovery
Lab*



*Data Factory
Engine*



Infrastrukturális alapozás



*Information
Management
Deep Dive*



DW Offload



Fast Data



Pilot projekt



Rapid Start Packs



*Big Data
& Analytics*

Hardware and Software

Engineered to Work Together

ORACLE®