



Adatminőség

Minden vállalat, szervezet működéséhez elengedhetetlen a folyamatait jó minőségben leíró adat előállítás és használata, ahol az adatminőség javítása többszörösen megtérül. Ehhez viszont szerkezetében és tartalmában heterogén, hiányos, hibás forrásadatok egységesítésén, javításán és feldolgozásán keresztül vezet az út.

SZTAKI megoldások

Módszereinkkel és eszközeinkkel gyorsan és költséghatékonyan tervezünk és implementálunk megbízható és robusztus adatintegráló folyamatokat. Kutatási eredményeinkre építve rendelkezünk alkalmazásokkal ügyfél-, log-, szenzor- és web adatokra, visszaélés-felderítés, IT security, okos város, trendelemzés és hálózatelemzés témakörökben.



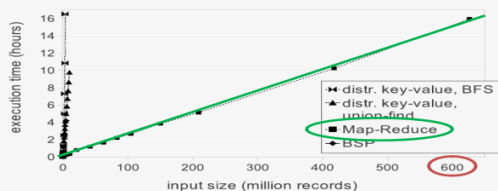
Adattisztítás, egységesítés, integráció

Adattisztító megoldásaink nyílt szabványú, újrafelhasználható transzformációs műveletekre épülnek: a valódi érték az itt leképzett üzleti, iparági tudásban rejlik. Gyorsan személyre szabható tartalommal rendelkezünk ügyféladatokra (nevek, postai címek, telefonszámok stb.) és log adatokra (applikációs és web szerver logok). Lehetőségünk van külső adatok, például címtörzsek, cégbázisok, térkép koordináták vagy akár közösségi hálókat csatolására. Transzformációs készleteink folyamatos fejlődése mellett új adatkörök feldolgozási lépései is gyorsan kialakíthatók.

Big Data

Az egyre nagyobb ütemben növekvő adatmennyiségek és a valósidejűség elvárása komoly kihívást jelentenek. Ezt szem előtt tartva megoldásaink jól tervezhető és költséghatékony módon szétoszthatók

- több processzormagra, több gépes környezetekre, felhőbe,
- Hadoop-ra köteget feldolgozáshoz,
- Strom-ra vagy in-memory adatbázisokra valós idejű feldolgozáshoz.



Azonosságfeloldás, deduplikáció

Hány ügyfelünk, termékünk, felhasználónk van valójában?

Adatbázisainkban a valós dolgok sokszor duplikálva, több lenyomattal is jelen vannak. Ahhoz, hogy valós képet alkothassunk a folyamatainkról elengedhetetlen az egyezések felderítése. Aktív kutatási tevékenységünk eredményein alapuló módszereink az elterjedt megoldásokhoz képest nagyságrendekkel több adatra alkalmazhatók olyan bonyolult esetekben is, mint például a háztartások azonosítása.