

VIKMAN LÁSZLÓ

Az üzleti célú információmenedzsment adatfeldolgozó eszközeiről*

1. Az információ szerepéről a gazdasági szférában

„Az információ a dolgok lelke. Kinyitom vele az ajtót, ami a siker kincseskamrájába vezet.” HERB COHEN¹

A XXI. századra az információ kulcsfontosságú erőforrássá vált, „amint egyre több ember munkája, kikapcsolódása és mindennapi élete kapcsolódik szorosan az adatfeldolgozáshoz és az adattovábbításhoz, az információs és kommunikációs technológiák alkalmazása a társadalom egyre nagyobb részét járja át.”² Ezek a kijelentések mára már közhelyé váltak ugyan, de vizsgáljuk meg mégis egy kicsit jobban, miért is történt ez így. A piaci verseny a globalizációnak, az egyre hatékonyabb termelésnek köszönhetően folyamatosan nő az egész világon. Alig maradt olyan termelési tényező, legyen szó természeti kincsről, olcsóbb gyártási technológiáról, képzett, avagy olcsóbb munkaerőről, ami elérhetetlen lenne a gazdasági világ számára, azaz az extenzív piaci stratégiák már nem alkalmazhatók hosszú távon. „Az információs technológiák alkalmazása minden jel szerint radikálisan átforgalmazza a szervezetek, kicsik és nagyok életét, működését. Ha valamely szervezet mindennapi működésének megszervezéséhez és fenntartásához jól alkalmazza az információs és kommunikációs technológiát, akkor képessé válik arra, hogy lényegesen jobb hatásokkal használja fel a rendelkezésre álló erőforrásokat, mint azt megelőzően.”³

1.1. Az információ - kulcsfontosságú gazdasági erőforrás és versenyképesítő

Ezek a folyamatok azt eredményezték, hogy megjelent egy komplex információ-központú

A szerző a PTE ÁJK PhD-hallgatója, az Infokommunikáció és Jog hírszerkesztője.

törekvés, melynek alapvetően három irányát jelölhetjük meg. A vállalatok célja, hogy - ügyfél- és partnerkörüket megtartsák, igényeiket, szokásaikat alapos megismerjék, - folyamatosan terjeszkedjenek a még ki nem aknázott piacok feltérképezésével, - termelési, valamint külső és belső információ folyamataikat még (költség)hatékonyabbá tegyék, kockázataikat csökkentsék.

Ezeket a célokat az informatikai fejlődés képes kiszolgálni, így világszerte terjednek az olyan szoftver- és hardvermegoldások (előbbire az ERP-rendszerek, utóbbira a hálózati technológiák jelentenek jó példát) amelyek lehetővé teszik a vállalat működése során összegyűlt adatok és információk gyűjtését, tárolását, kiértékelését és az ennek megfelelő döntések meghozatalát.

Azonban az adatok, és főként a természetes személyek adatainak ilyen tömeges méretű feldolgozása komoly kockázatot jelent az adatalanyok számára, ennek oka pedig abban keresendő, hogy sokszor még azzal sem vagyunk tisztában, hogy ki és mit tud rólunk. Fokozza az egyén, a személyiség alávetettségét, hogy az informatikai eszközök segítségével könnyen, gyorsan előállítható olyan *személyiségprofil*, amely alkalmas az egyén "átvilágítására", személyes, családi, baráti, üzleti kapcsolatainak, múltjának és jelenének az érintett tudta és beleegyezése nélkül való feltérképezésére. Következtetni lehet terveire, jövőjére, sértve ezáltal szabad akaratát, méltóságát s komoly visszaélések lehetőségét teremtve meg.⁴ Az informatikai eszközök segítségével könnyen, gyorsan előállítható olyan *személyiségprofil*, amely alkalmas az egyén „átvilágítására”, személyes, családi, baráti, üzleti kapcsolatainak, múltjának és jelenének az érintett tudta és beleegyezése nélkül való feltérképezésére; ez komoly visszaélések lehetőségét teremti meg.

Előfordulhat olyan helyzet is, hogy nem helytálló, vagy nem teljes körű ismeretek birtokában építik fel az adott alanyról készített személyiségprofil, vagy olyan adatokat kezelnek az egyénről, amelyek különösen sértik érdekeit. Ilyen különösen érzékeny adatok az egészségi állapotra, a szexuális életre, illetőleg szokásokra, a kóros szenvedélyre, faji eredetre, nemzeti, etnikai hovatartozásra, a vallási vagy más ideológiai, politikai meggyőződésre és a büntetett előéletre vonatkozó adatok.⁵ Ezek a körülmények pedig könnyen eredményezhetnek diszkriminatív vagy indokolatlanul terhes eljárásokat (pl. hitelkérelmek elbírálásánál, biztosítások megkötésénél). A jognak tehát fontos feladata, hogy kiküszöbölje az egyenlőtlen kommunikációs helyzeteket, amelyek így az érintett rovására létrejönnek.

1.2. A stratégiai tervezés kialakulása

„Annak az üzletembernek, aki adatok nélkül tervez, meg kell tanulnia profit nélkül megélni.” VIRGIL REED⁶

A stratégiai tervezés az USA-ban bukkant fel először, az 1970-es évek elején, majd ezt követően terjedt el. Azt megelőzően (50-es, 60-as évek) az össz-kereslet stabilan növekedett, jó üzleti lehetőségeket kínálva a vállalatoknak, s így a vállalatvezetés az egészen rövid távú tervezéssel is jól elboldogult. Ilyen körülmények között gondot inkább csak az operatív működés hatékonyságának és eredményességének növelése jelentett, míg a vállalat környezetével való törődés nem volt jellemző.

Az 1970-es években válság válságot követett: a közel-keleti háborúk által kiváltott olajválságok, az ebből következő és ezt kísérő alapanyag- és energia-hiány, inflációs tendenciák a fejlett országokban,

valamint gazdasági stagnálás és emelkedő munkanélküliség. Ugyanakkor megindult az olcsó és jó minőségű termékek rohamos beáramlása Japánból és más országokból az USA-ba és más fejlett országokba, mely termékek piaci részesedést hódítottak el a legfejlettebb nyugati országok vállalataitól, még olyan tradicionális iparágakban is (pl. autógyártás), ahol azok korábban kifejezetten erősek voltak. A nyugati nagyvállalatok döntő többségének tehát belföldön és külföldön is komoly versennyel kellett szembesülniük, ami nehéz és alapvetően újszerű kihívást jelentett számukra.

Ez a gyökeres piaci változás a stratégiai menedzsmentet az addigi, inkább másodlagos szerepéből az előtérbe emelte, és olyan új tervezési rendszer bevezetését tette szükségessé, mely lehetővé teszi a vállalat egészséges és előremutató működésének fenntartását akkor is, ha egyes termékek, piacok vagy üzletágak esetében zavarok keletkeznek.⁷

1.3. Az integrált vállalatirányítási és döntéstámogatási rendszerek

Az üzleti célú információ-kezelés célkitűzéseit piaci igényként jelentkeztek a rohamosan fejlődő informatikában, és olyan fejlesztéseket eredményeztek, melyek az adatfeldolgozást soha nem látott mennyiségben és minőségben végzik el. Ma már egyetlen komolyabb szoftvergyártó, informatikai megoldás-szállító sem hagyhatja figyelmen kívül ezt a szegmenst. Ez olyannyira igaz, hogy az ilyen üzleti alkalmazások fejlesztése, sikeres bevezetése, a velük szorosan összefüggő adatbázis-technológiák, hálózati megoldások, e-business termékek néhány év alatt hatalmas multinacionális cégek kiépülését hozta (például Oracle SAP).

Nézzük, melyek egy, a fentebb megjelölt stratégiai tervezési eszköz, speciális célszoftver sajátosságai. Az integrált információs rendszerben az adatfeldolgozás lépései úgy adják tovább az információt az őt követő lépéseknek, hogy közben nem változik az eszköz, a tároló-közvetítő platform (pl. nem mentik az adatokat cserélhető adathordozóra – pl. floppy lemezre – majd töltik be egy másik számítógépbe és/vagy hálózatba), és nincs többszörös adatbevitel sem, ugyanis ezekre nincs szükség. Az integrált rendszerben a funkciók, tevékenységek nem keverednek és általában nem duplikálódnak.

Az integrált információs rendszert továbbá az különbözteti meg a hagyományos, nem integrált megoldásoktól, vagyis a külön-külön működő, egy-egy részfunkciót megvalósító, egymással kapcsolatban nem lévő, elszigetelt rendszerektől, hogy az integrált információs

rendszerben a különböző alrendszerekben lezajló feldolgozások egyetlen rendszerben történő feldolgozása csak úgy lehetséges, ha az egyes alrendszerek egyrészt szorosan együttműködnek, egymásra épülnek, másrészt pedig ugyanazon, redundanciától mentesen leképzett adatokat használják. Ez a „single data depository”, az egyszeres adattárolás.

Integrált vállalatirányítási információs rendszer alatt az egy vállalaton belül lezajló valamennyi műszaki, termelési, kereskedelmi, raktározási, készletgazdálkodási, pénzügyi, vezetési, irányítási és számos egyéb folyamat egységes, integrált számítástechnikai kezelést megvalósító információs rendszert értjük. Ezen alkalmazások tehát egy adott vállalat valamennyi fontos adatfeldolgozását megvalósító, globális, átfogó információs rendszerek, amelyek az egész vállalatra kiterjedő integrációt valósítanak meg. Integrált vállalatirányítási információs rendszer alatt az egy vállalaton belül lezajló valamennyi műszaki, termelési, kereskedelmi, raktározási, készletgazdálkodási, pénzügyi, vezetési, irányítási és számos egyéb folyamat egységes, integrált számítástechnikai kezelést megvalósító információs rendszert értjük.

Az integrált információs rendszer, mint fogalom természetesen nem jelent automatikusan globális, átfogó (legszélesebb vállalati funkcionalitást megvalósító) rendszert is. Léteznek olyan integrált alkalmazások is, amelyek a vállalati folyamatok egy-egy csoportját támogatják integrált módon, így pl. integrált pénzügyi, számviteli, humánerőforrás-menedzsment, logisztikai rendszerek. Ezek az alkalmazások képesek jelentős hatékonyság-növekedést elérni az általuk támogatott folyamat-csoport területén. Ahhoz azonban, hogy vállalati szintű hatékonyság-növekedést tudjunk segítségükkel elérni, a különböző folyamat-csoportokat megvalósító rendszereket vállalati szinten is integrálni kell egymással, vagyis átfogó integrált rendszert kell kialakítani.⁸

Ilyen integrált vállalatirányítási és vezetői döntés-támogató rendszert ma már rengeteg szoftvergyártó készít, és sok ágazat-specifikus termék is készül. Ezekre a szoftverekre, rendszerekre is kialakult már egyfajta „tolvajnyelv”, rengeteg rövidítést, betűszót talál, aki a témával foglalkozik. Az átfogó rendszerek egyik jelölője és talán a legtagabb kategória az ERP (*Enterprise Resource Programming*), ami vállalati erőforrás gazdálkodást jelent. A stratégiai elemet emeli ki, így már specializáltabb az SEM (*Strategic Enterprise Management*). Az információ feldolgozására és a vezetői döntéstámogatásra koncentrálnak az EIS (*Enterprise Information System*) és az

O/DSS (*Operation/Decision Support System*). Még inkább egy külön területre koncentrálnak szoftverek az SRM (*Supplier Relationship Management*), a szállítói-kapcsolat menedzsment, valamint az SCM (*Supply Chain Management*), az ellátási-lánc menedzsment. Az IRM (*Inner Resource Management*) belsőforrás-menedzsment célja, hogy a vállalat humán erőforrásainak, illetve pénzügyi és tárgyi eszközeinek megfelelő felhasználása útján képes legyen a forgatóke csökkentésére, a ciklusidő rövidítésére és a működés általános javítására. Az IRM rövidítés másik feloldása az *Industrial Resource Management*, azaz ipari erőforrás-menedzsment. Ezen szoftverek felhasználásával a vállalat képes rögzíteni egyetlen eladás pénzügyi hatását az alapanyagok bevételezésére, a termékszállítási folyamatra a megfelelő raktározáson és szállításon keresztül⁹.

Az utolsó alkalmazástípus azért emelkedik ki a sorból – legalábbis a jelen vizsgálódás szempontjából –, mert használatának előfeltétele és elsődleges célja a vállalati ügyfelek személyes adatainak tömeges és hatékony gyűjtése és feldolgozása. Ez a CRM (*Customer Relationship Management*), az ügyfél-kapcsolat menedzsment. A CRM a holisztikus marketing eszköze, segítségével megállapítható, hogy ki a vevő, hogy viselkedik, mit akar, és mire van szüksége.¹⁰ Mivel vizsgálódásának központjában az ember áll, az adatvédelem szempontjából kiemelt jelentőséggel bír.

2. A CRM rendszerek

„Öt éven belül kétféle cég fog létezni. Azok, akik a számítógépet marketingeszközként fogják használni és azok, akik a csőd szélén állnak.” - WARREN McFARLANE¹¹

Az ügyfélkapcsolat-menedzsment rendszerek lényege hogy a vállalatok, piaci pozícióik megtartása és fejlesztése érdekében a gyártási folyamat- és termék-központúság helyett az ügyfélközpontúság felé fordulnak. Az ügyfélkapcsolat-menedzsment rendszereket is a fokozódó piaci verseny hívta életre, azonban ezek többnek tekinthetők, mint egy vagy több informatikai rendszer. Inkább egy új szemléletmódról, üzleti modellről van szó. Lényege hogy a vállalatok, piaci pozícióik megtartása és fejlesztése érdekében a gyártási folyamat- és termék-központúság helyett az ügyfélközpontúság felé fordulnak. Ennek egyik fő oka az a jelenség, hogy a piacon megjelenő termékek és szolgáltatások egyre inkább hasonlítanak egymáshoz, ami azt eredményezi, hogy kompetitív előnyt csak az ügyfél kiszolgálásának magasabb minősége

tud nyújtani. Előtérbe került az ügyfél, annak elégedettsége, az ügyfélismeret pedig felértékelődött. Ezért a vállalatoknál olyan módszer, eszköz iránt jelentkezett igény, melynek segítségével a vállalati kultúrában és a munkafolyamatokban is középpontba kerül az ügyfélkapcsolatok kezelése. A CRM modell alkalmazásának eredményeképpen a kiterjedt ügyfélkörrel rendelkező vállalat is képes egyedi ügyfélkapcsolatok kialakítására és ápolására, melynek keretében az ügyfél mindig kiemelt bánásmódban részesülhet. Ezáltal jelentősen megerősíthető az ügyfél vállalat iránti lojalitása, s ezen keresztül a vállalat szempontjából vett profitabilitása.

A CRM üzleti információk adattárházra és az Internetre épül, s célja, hogy az ügyfelekről meglévő adatok rendszerezése, elemzése révén jelentős mennyiségű és értékes információt szerezzen a vásárlói szokásokról, vevői preferenciákról, s ezáltal módot adjon annak meghatározására, hogy pl. kik a legfontosabb ügyfelek. Ezen információk birtokában, a valós vásárlói igények ismeretében megalapozottabb üzleti döntések hozhatók. További előnyt – költségmegtakarítást – jelenthet nagy ügyfél-forgalmú szervezetek esetében, hogy lényegesen olcsóbb és kényelmesebb lehet az internetes kapcsolatfelvétel az ügyféllel, mint a hagyományos ügyfélszolgálati irodákban. A CRM üzleti információk adattárházra és az Internetre épül, s célja, hogy az ügyfelekről meglévő adatok rendszerezése, elemzése révén jelentős mennyiségű és értékes információt szerezzen a vásárlói szokásokról, vevői preferenciákról, s ezáltal módot adjon annak meghatározására, hogy pl. kik a legfontosabb ügyfelek.

Általános tapasztalat, hogy a vevők egy viszonylag kis része „hozza” a nyereség legnagyobb részét, bizonyos vevők pedig veszteségesek. A fenti helyzetre tekintettel a vállalatok számára létfonosságú a nyereséget hozó ügyfelek megtartása, már csak azért is, mert általában többre kerül egy új, nyereséges ügyfelet szerezni (magas marketing-költségek, hosszú megtérülési idő, nyereség-kiesés, stb.), mint egy már meglévő megtartani. Ehhez azonban természetesen pontosan ismerni kell, kik a nyereséges ügyfelek, ügyfélcsoportok. Az ügyfelek megtartása, az „ügyfél-tőke” megőrzése szempontjából fontos, hogy azok „jól érezzék magukat” üzleti kapcsolataikban. A kiszolgálás minőségének döntő szerepe van abban, hogy marad-e a cégnél az ügyfél.

Az ügyfélkapcsolat-menedzsment megoldások elvégzik az adatok konszolidálását, strukturálását, majd azokat egy egységes adatbázisba helyezik el. Ezzel elősegítik, hogy az ügyfelekről egységes kép alakuljon ki a különböző csatornákon (kereskedők,

ügyfélszolgálati adatok, szerviz-adatok, reklamációk stb.) bejutó információkból. A közös adatbázisban elhelyezett információkat az arra illetékesek másodpercek alatt elérhetik, amikor csak szükségük van rájuk (pl. amikor az ügyféllel kapcsolatban lépnek). Miután az egyes ügyfelekről már sok információ összegyűlik, alaposan megismerik őket, s így lehetőség nyílik, pl. „testreszabott” marketingkampányokra vagy szolgáltatások nyújtására is, amely nagyban elősegítheti a nyereséget hozó vevő megtartását, illetve fogyasztást ösztönző lehet.

A vevőkről felhalmozott információk az előbbieken túlmenően úgy is hasznosíthatók, hogy korszerű eszközök, mint az *adatbányászat* (KDD) használata révén összefüggéseket tárnak fel a vevők szokásai, különböző paraméterei és jellemzői között. Az így felismert összefüggések azután jól hasznosíthatók a piaci működés során (ügyfelek elhódítása a versenytársaktól, termékek célzott bevezetése, árukapcsolási lehetőségek kihasználása és így tovább).

Az átfogó ügyfélkapcsolat-rendszer kialakítása nélkül nem lehet az ügyfelek számára testre szabott szolgáltatást kialakítani, ami pedig nagyon sok területen a vevők megtartásának, a versenyképesség megőrzésének, növelésének alapfeltétele, nem csak termelési típusú vállalkozások, hanem szolgáltatók esetében is, mint pl. közüzemi szolgáltatóknál, bankoknál, biztosító társaságoknál, vagy távközlési szolgáltatóknál. Így ezen szervezetek jelentős részénél ma stratégiai célkitűzés az átfogó ügyfélkapcsolat-kezelő rendszer kialakítása.

2.1. A CRM-alkalmazások főbb típusai és azok legfontosabb feladatai

- Az *operatív CRM* irányítja és szinkronizálja az ügyfél-interakciókat a marketing, az értékesítés és a szerviz területén. Feladata a direkt ügyfélkapcsolatok informatikai támogatása, melyhez a legteljesebb körű, egységes információ-bázisra van szüksége. Az operatív CRM által kezelt információ-bázis két forrásból táplálkozik: a közvetlen ügyfélkapcsolatok során keletkező információk, valamint a vállalati adatvagyon feldolgozása, elemzése során nyert információk.

- Az *analitikus CRM* segítséget nyújt az információforrások optimalizálásában annak érdekében, hogy elősegítse a profitábilis ügyfelek megtartását, az ügyfél- és piaci ismeretek megszerzését, a célzott piaci szegmens meghatározását, a kínálat kialakítását, az ügyfél-magatartás alakulásának elemzését támogatják információkkal, elemzésekkel.

- A *kollaboratív CRM* elősegíti az együttműködést a szállítókkal, a partnerekkel és az ügyfelekkel

a folyamatok tökéletesítése, és az ügyfelek igényeinek kiszolgálása érdekében.

Az átfogó CRM-rendszer megvalósítása általában komoly beruházást jelent. Alkalmazásával ugyanakkor a nagy ügyfélforgalmat lebonyolító, kiterjedt ügyfélkörrel rendelkező szervezetek esetében komoly költség-megtakarítás érhető el. Az átfogó rendszer megvalósulása általában egy hosszabb folyamat eredménye, melynek kiinduló pontját a jelentés-készítés, az adatbányászat, az adattárházak képezik. A CRM rendszer kiépítése, ennek megfelelően, gyakran több lépésben történik: először kiépül a széles körű adattárház, az adatbányászat, az információ-kinyerés, jelentés-készítés, majd a CRM további funkciói.¹²

2.2. A CRM alkalmazása¹³

A vevők egyedi választási környezetének megfelelő ajánlatok kidolgozása során a végső feladat az, hogy a vállalat tartós és az igényeket kielégítő kapcsolatot alakítson ki az értékesnek ítélt vevőkkel. Az ügyfélkapcsolat-menedzsment működése elképzelhetetlen a vevők vásárlási szokásait, adatait és beállítódását, illetve a vállalati kapcsolattartókra vonatkozó információkat rögzítő hírszerzési rendszerek nélkül. Az információ rendszerezése, értelmezése és elemzése útján a vállalat új ismeretekkel gazdagodhat. Az efféle tudás megszerzése és átadása érdekében a vállalatok egy része együttműködési hálózatokba tömörül. Vegyünk néhány konkrét példát:

- A *Dell Computer* a vevők preferenciáinak megismerése végett nyomon követi, hogy milyen termékeket és jellemzőket választanak. Ennek ismeretében a cég sokkal agresszívabb árképzési stratégiát és reklámozást alkalmazhat az egyes konfigurációk eladása során.

- Az *Amazon.com* figyelemmel kíséri felhasználói böngészési és vásárlási szokásait, de ez még nem minden. Az innen származó adatokat összeveti más, hasonló szokásokkal rendelkező felhasználók adataival, majd ez alapján egyedi termékajánlatokat dolgoz ki.

- A *General Motors* részlegeinek adatbázisait egyetlen központi adatbankban egyesítette. Az összesített információkat a cég olyan új autómódellek tervezéséhez használja, amelyek egymást átfedő célcsoportok számára készülnek. Adatbázisai összevonásával, valamint a telefonos ügyfélszolgálatok számának drasztikus – hatvanról háromra történő – csökkentésével a GM megbízható és egységes vevőszolgálatot épített ki.

3. A tudásfeltárás és módszerei

A tudásfeltárás (KDD - *Knowledge Discovery in Databases*) egy további olyan informa-

tikai részterület, mely a számítástechnika robbanásszerű fejlődése révén jöhetett létre. Lényege, hogy hatalmas adathalmazokból nyer ki statisztikai és matematikai módszerek segítségével olyan információkat, melyek segítik az operatív döntéshozatalt és a stratégiai tervezést. Célja tehát olyan szabályszerűségek, minták feltárása, amelyek segítségével a folyamatok könnyebben átláthatóvá és kalkulálhatóvá válnak. Az adattárház és adatbányászat a tudásfeltárási módszerek közé tartozik, mint az adatbázis-technológia, a statisztika, a szakértői rendszerek, a gépi tanulás, vagy a térinformatika. A tudásfeltárás lényege, hogy hatalmas adathalmazokból nyer ki statisztikai és matematikai módszerek segítségével olyan információkat, melyek segítik az operatív döntéshozatalt és a stratégiai tervezést; célja olyan szabályszerűségek, minták feltárása, amelyek segítségével a folyamatok könnyebben átláthatóvá és kalkulálhatóvá válnak.

E két fogalom vizsgálati tárgy és eszköz kapcsolatban áll egymással és együttes alkalmazásuk olyan üzleti döntéstámogatási módszert jelent, amely segít megtalálni és kiaknázni új üzleti lehetőségeket a nagytömegű adathalmazokban rejlő, nem ismert összefüggések feltárásával. Egyesítik az adatbázis-kezelés, a statisztika és a mesterséges intelligencia kutatások eredményeit. Az adatbányászati algoritmusoknak meglévő alkalmazásokba való beépítésével (pl. ügyfélszolgálat) intelligens alkalmazások hozhatók létre, melyek olyan komplex összefüggések feltárását biztosítják (ez a CRM egy funkciója), amelyek nem nyerhetők ki egyszerű vagy akár multidimenzionális lekérdezésekkel.

Fő céljuk – melyet változatos eszköztárral érnek el – a vizsgált múltbeli adatokban jellegzetes minták, összefüggések, trendek felfedezése, ezzel új tudás létrehozása. Használják azonban piaci szegmentálásra (fogyasztók, termékek vagy szolgáltatások vonatkozásában is), osztályozásra, előrejelzések készítésére, üzleti tervekhez adatszolgáltatásra, illetve napi rendszerességgel, vezetői döntéstámogatási rendszerekkel (Decision/Operation Support System – D/OSS) összekapcsolva rövid távú problémák megoldására is.

3.1. Adattárház (Data Warehouse)

A vezetői információs rendszerek ma már nem képzelhetők el rendszerezett, megbízható minőségű, gyorsan elérhető adatok nélkül. Ezeknek a követelménynek tesz eleget az adattárház koncepció. Olyan módszereket és eszközöket jelent, amelyekkel nagy tömegű adatot lehet hatékonyan kielemezni és az analízis eredményét jó hatásokkal

eljuttatni a felhasználóhoz, elsősorban természetesen a menedzsmenthez. Ez a technológia multidimenziós adatmodellt, multidimenziós adatbázis-kezelőt, s ezekhez kapcsolódó megjelenítő eszközöket jelent, vagyis olyan eszköztárat, amely biztosítja, hogy a vállalati menedzsment a rendszer-adminisztrátorok, rendszergazdák és fejlesztők közreműködése nélkül is képesek legyenek az őket érdeklő adatokhoz hozzájutni, azokat kielemezni, az analízis eredményét megjeleníteni, illetve felhasználni. Az adattárházakban a különböző területekről összegyűjtik, integrálják, és speciális sémában tárolják a szükséges részletezettség visszakereshető adatokat.¹⁴

Az adattárház bevezetése az addigi adatkezelési rezsimekhez képest gyökeresen megváltoztatja a vállalatok életét, munkafolyamatait. A feladat hatalmas erőforrásokat igényel, és csak hosszútávon hoz kézzelfogható eredményeket. A kockázatot jelentősen csökkenti, ha az adattárház építését lépésenként valósítják meg. Ehhez a feladathoz nyújtanak megfelelő eszközt az *adatpiacok*, melyek csak a vállalat egyes funkcionális területeinek adatait tartalmazzák. Ezek az adatpiacok az adattárházzal megegyező szerepet töltenek be az adott funkcionális területeken. A külön-külön megvalósuló adatpiacok gyorsabban, olcsóbban, kisebb kockázattal állíthatók elő, majd később könnyedén integrálhatók egyetlen adattárházba. A gyakorlat is azt mutatja, hogy a bonyolult adattárházak több adatpiac projekt együttes eredményeként jöttek, jönnek létre.

Az adattárház a vállalat birtokában lévő releváns, integrált, konzisztens, történeti adatainak gyűjteménye, mely a hagyományos, gyenge hatékonyságú információs folyamatot alakítja át gyökeresen, az alábbi elvek szerint:

- a különböző forrásokból (számlázó rendszer, integrált vállalatirányítási rendszer, stb.) származó adatok rendszeres időközönként leválogatásra és betöltésre kerülnek egy központi adatbázisba,
- az alapadatok az üzleti elemzési szempontoknak megfelelően feldolgozásra kerülnek,
- az adatok egy helyen, az üzleti gondolkodásnak megfelelő új struktúrában kerülnek eltárolásra,
- az adatok *online* elérhetővé válnak a vezetők és elemzők számára,
- az elemzéshez szükséges funkciókat standard elemző eszközök vagy egyedi alkalmazások biztosítják.

Az adatok áramlásának iránya:

1. Operatív adatbázisok:

Ilyenek az ügyfél-szerződés adatai, a számlázás, az ügyfélszolgálat nyilvántartásai,

a különböző szolgáltatások naplózása, ügyfél-kártyák, hitelkártyák felhasználási története (telefonhívás, pénzfelvétel, vásárlás, stb.).

2. Adatpiacok:

Regionális, vagy üzletági összeállítású, általában speciális céllal létrejött adatbázisok (pl. magyar *smart*-kártyák), végezhető rajtuk adatbányászat (mint minden adatbázison), de tulajdonságaik miatt még nem optimálisak.

3. Adattárházak:

Sok millió bejegyzéssel, az adatbányászat előre meghatározott, szükséges adattartalmú, speciális kritériumoknak megfelelő adatbázis.

- Időfüggettség: az adatok időponthoz kötöttek.
- Változatlanág: csak lekérdezéshez használják, sosem módosítják, csak bővítik.
- Tematikusság: nem minden adat az operatív adatbázisokból.
- Integráltság: az információkat konzisztens, egységes formában tükrözi, egy alany csak egy névvel szerepel.

4. Metaadatbázis:

Az adatbázisok nyilvántartása, kifejezi az adatbázisok relációját, kapcsolódásait. Az adatáramlás és az adatbázisok összefüggéseinek vizsgálatok jelentős szerepet játszik, térképként használható.

3.2. Adatbányászat

„A megkülönböztetés, vagyis az a képesség, hogy helyesen válasszuk meg, mire, hová és kire irányítjuk figyelmünket, az üzleti siker fontos előfeltétele, mert ez a forrása a sándéknak és az akaratnak, hogy a legfontosabb dolgot tegyük.” - MICHAEL E. GERBER¹⁵

Az adatbányászat kifejezés a 90-es években jelent meg az informatikai rendszerek területén és ma már széles körben használják. A kifejezés az angol *„data mining”* tükörfordításaként honosodott meg. Az adatbányászat az adattárházban tárolt adatok alapján történik. Célja az adatok közötti, rejtett összefüggések feltárása és a szervezet céljai érdekében történő hasznosítása.¹⁶ Ténylegesen csak egy lépése a tudásfeltárás folyamatának, azonban mint a legfontosabb és legsajátosabb rész, gyakran az egész műveletet jelöli a gyakorlati terminológiában. Az adatbányászat a gazdasági szereplők kezében kiemelten hatékony eszköz a nehezen körülhatárolható és megoldható üzleti kérdések megválaszolására. A fenti értékes információk megszerzésének útjában azonban szintén számos akadály állhat:

- A hosszú távú elképzelés, perspektíva hiánya

- Az adatállományok elavultak, nem elég frissek
- Cégen belüli ellentétek, kommunikációs zavarok
- Nem megfelelő formában közvetített adatok
- Az állományok összekapcsolásának technikai akadályai – eltérő szoftver, hardver
- Időzírási problémák, késlekedés
- A kapott eredmények értelmezési problémái
- Jogi szabályozásból fakadó akadályok

Az adatbányászati modellek:¹⁷

A, *Verification Model* – azonosító modell

A szakértők által megfogalmazott hipotézisekre és kérdésekre különböző lekérdező és analitikus eszközök segítségével keres választ, és az eredményként kapott válaszok tükrében erősíti meg, vagy veti el azokat. Ez a megközelítés megfelel a legtöbb klasszikus értelemben vett statisztika analízisnek.

B, *Discovery Model* – felfedező modell

Ebben a megközelítésben már a kiindulási alapot is automatizált úton nyert feltevések képezik. A kutatással egyidejűleg generálódnak a hipotézisek, melyeket azonnal ellenőriz is a folyamat. A közvetlen kutatásban a célváltozó adott, és a függő változókat kell megkeresni, az indirekt kutatásban előzetesen definiált változók nélkül keresnek összefüggéseket, korrelációkat.

3.3. A tudásfeltárás folyamata

1. Adatkiválasztás, tervezés

A konkrét feladat meghatározása, az ehhez szükséges szakemberek, más erőforrások és főként adatok kiválasztása.

2. Tisztítás

A humán munkához kötődő adatbázisokban gyakran előfordulnak duplikációk, elírások, hiányok, amiket a hatékony elemzéshez szükséges kijavítani. Ez a *data-cleaning* folyamata.

3. Bővítés

Amennyiben nem áll rendelkezésre megfelelő mennyiségű adat, külső erőforrások bevonásával bővíthető az adatbázis (CPEX). Ehhez igénybe vehetők adatbázis-szolgáltatók, adatgazdák cserélhetnek egymással adatokat, de végrehajtható kiegészítő adatfelvétel is.

4. Kódolás:

Célja a gépi feldolgozás és megjelenítés egyszerűsítése, módszerei:

- Pontos címek helyett területi kódok megadása.
 - Születési dátumok helyett életkor, vagy életkorosztályok.
 - Adatbázisban szereplő nagy összegek elosztása 10 hatványával.
 - Eldöntendő kérdések bináris ábrázolása (pl. van-e autója az adatalanyoknak?).
 - Pontos dátumok (pl. szerződés kötés) helyett adott időponttól eltelt hónapok száma.
- Fontos lépés, mivel gyakran adatvédelmi jogi szempontból lényeges az adatok végleges anonimizáltsága. Ezzel néhány adatkezelés esetében nem éri hátrány az adatfeldolgozó és csökkenthető a művelet jogi kockázata.

5. Adatbányászat

6. Jelentéskészítés, kiértékelés

A feladat az adatbányászat eredményeinek befogadható, üzleti döntésbe konvertálható formában való prezentálása. Ez többnyire az eredmények vizuális megjelenítésével (grafikonok, diagramok, stb.) és a megszerzett tudás rövid összefoglalásával történik. A feltárt mintával szemben támasztott elvárások:

- érthető
- érvényes és friss adatok alapján készült
- a feladattal kapcsolatos használható és releváns információ
- önálló és komplex összefüggések megjelenítése
- érdekes, fontos és jelentős.

3.4. A KDD adatbányász eszközei

Az adatbányászat nem egyetlen számítástechnikai megoldás megjelölésére használt fogalom, hanem egy alapelv, szemlélet, mely szerint az adatokban több tudás van elrejtve, mint amennyi felszínesen vagy egyenként tekintve látszik. Ezért az adatbányászatra bármilyen olyan technika (pl. adatbáziskezelési megoldás, grafikus megközelítést használó szoftvereszköz, statisztikai leválogatás) alkalmas, melynek segítségével több hozható ki az adatokból. A technikák meglehetősen heterogének, mindegyik rendelkezik valamilyen speciális előnnyel, mindegyik más célra a legalkalmasabb. A legcélravezetőbb megoldás kiválasztásában szerepet játszhat a feldolgozandó adatok mennyisége, heterogenitása, a célként megjelölt kinyerési kívánt információ tulajdonságai, a kutatástól elvárt pontosság stb. Nézzük, melyek a leginkább ismert és leggyakrabban alkalmazott eljárások, eszközök:

1. *Hagyományos lekérdező eszközök* (statisztikai módszerek, SQL, ...)

Az adathalmazok viszonylag durva elemzéséhez, előzetes feltevések tételéhez alkalmasak. Segítségükkel általános képet kaphatunk az adatokról, azok szerkezetéről. Az előzetes elemzéshez kiindulhatunk az adathalmaz egyszerű, statisztikai információiból, például megnézhetjük az általános értékeket. Bonyolult algoritmusok használata előtt mindenképp szükséges az alkalmazásuk, mivel képesek lehetünk velük bizonyos előrejelzéseket készíteni, amelyekkel a későbbi, alaposabb és specializáltabb keresések eredményeit, megbízhatósági korlátait is láthatjuk, ellenőrizhetjük.

2. Megjelenítési technikák

Szintén hasznos lehet a bonyolultabb eljárások alkalmazása előtt a különböző vizuális megjelenítési lehetőségek bevetése, melynek során az adathalmaz mintáit szabad szemmel próbáljuk felismerni. Segítségükkel – szerencsés esetben – megsejthető, merre érdemes eddig nem ismert összefüggések után kutatni. Háromdimenziós alkalmazásokkal az adatok változásai akár térbeli folyamatokként is ábrázolhatók.

3. Térinformatika – GIS

A vállalatok tevékenységének 80%-a a helyhez kötött, így adataik közvetve földrajzi tulajdonságokat is hordoznak. A térbeli összefüggések megjelenítését és elemzését a térinformatikai szoftverek támogatják. Az üzleti térinformatika a vállalati adatok értéke és az adatok térbeli elhelyezkedése közötti összefüggések elemzésére szolgáló eszköz. Az üzleti térinformatika előnye, hogy egy térképen egyszerre a földrajzi területek több tulajdonsága, változója jeleníthető meg (tematikus térkép, oszlop vagy kördiagram segítségével). Gyorsan áttekinthetőek, hamar és különösebb szaktudás nélkül értelmezhetőek, látványos elemei lehetnek a tulajdonosi beszámolóknak, valamint vezetői prezentációknak.

4. Hasonlóság és távolság

Ez a módszer is a térbeli szemléletre épít, amennyiben azt vizsgálja, hogy az egyes rekordok, illetve értékeik milyen távolságra vannak egymástól. Azok, amelyek egymáshoz közel vannak, hasonlítanak egymásra, a távoliak kevés azonos vonással rendelkeznek. Egy többdimenziós adattér pontjaiként vizsgálva a rekordokat és a már említett háromdimenziós ábrázolási eszközöket alkalmazva kaphatunk különböző „adatfelhőket”. Néha csupán ez is elég, hogy érdekes osztályokat találjunk, azonban többnyire bonyolult osztályozó (klaszterező) program szükséges az adatcsoportok meghatározására.

5. OLAP-eszközök

A közvetlen elemzésre alkalmazott OLAP technológia atyja Dr. E. F. Codd, aki a relációs adatbázis-kezelési technológiát is feltalálta. Azt meghaladva a multidimenziós elrendezés célja többek között az, hogy a különböző információ-felhasználók más-más formátumban és részletezettségben nézhessék, használhassák az adataikat, több adatdimenzióval dolgozhatnak. Az OLAP technológia ezeken az elemzési szempontokon – multidimenziós nézeteken – keresztül biztosít közvetlen, gyors adatelérést a végfelhasználók számára. Elemzési lehetőséget biztosít az egyszerű lekérdezéstől és lefűrástól a szofisztikált modellezésig és komplex tervezésig, azonban tanulásra, valamint új megoldások keresésére nem képes. Az OLAP-eszközök jellemzően adattárházakon, adatpiacokon tudnak a leghatékonyabban működni. Az adatokat speciális többdimenziós formában, úgynevezett adatkockákban tárolják. Az adattárakból nem csak adatokat importálnak, hanem eredményeiket abba vissza is töltik ezek az eszközök.

6. k-legközelebbi szomszéd

Ez a technika sem intelligens tanulási technika, hanem egyfajta keresési módszer. Előrejelzéseinek alapja az az elv, hogy minden alany várható viselkedése megegyezik a hozzá leginkább közel állóval. A módszer egyszerű, azonban nem nyújt elegendő segítséget ahhoz, hogy az adattartományhoz olyan elméletet lehessen társítani, amely segíthet az adatok szerkezetének jobb megértésében. Másik hátránya a nagy számításgépi igény, ezért csak kisebb méretű adatbázisokra, vagy adatbázisból már valamilyen szempontok szerint kiválasztott mintákra alkalmazzák.

7. Döntési fák

A szabálygeneráló tanuló algoritmusokkal végzett elemzéseknek egy olyan vizuális

reprezentálása, amikor a vizsgált kérdés egy lehetséges válaszáig vezető összefüggések rendszere pontról-pontra nyomon követhető, mintha egy fa ágain haladnánk ágról-ágra. Minden elágazás azonosítható egy kérdéssel és a továbbhaladás iránya függ a kérdésre adott választól. Például ha egy ügyfelünk nő, férjzett, és rendelkezik folyószámlával (kérdések), akkor igen (válasz), küldjünk neki a legújabb termék akciónkról direkt marketing levelet, mert valószínűleg pozitívan fog rá reagálni.

8. Társító szabályok

Olyan szabályok, amelyek egy adatbázis-tábla bizonyos attribútumaihoz tartozó értékek statisztikai kapcsolatát írják le. Az adott attribútumok alapján megjósolható, megbecsülhető az eredmény. A társító szabályokat mindig bináris attribútumokban tartározzák meg. Az ilyen szabályok felfedezése nem nehéz, problémát a sok felfedezett, de mégis érdektelen szabály kiválogatása okoz. Általában akkor használhatók eredményesen, ha már van valamilyen elképzelés arról, hogy mit is keresünk.

9. Neurális hálózatok

A tanuló algoritmusoknak egy típusa, ahol a csomópontok (neuronok) a szinapszisaikon keresztül cserélnék információt egymással. A keletkezett szabályrendszerek nem követhetők nyomon úgy, mint a döntési fák esetében, mivel a neuronhálózatok matematikai háttere nem teszi ezt lehetővé. Erre azonban nincs is mindig szükség. Egy kétértékű változósorral támaszkodó következtetés esetében például, ha nincs szükség arra, hogy melyik változó milyen értéket vett föl, a neuronhálózatok pontosabban és gyorsabban alkalmazhatók. (Az ötlet a biológiából származik, mivel a neurális hálózatok az idegrendszer szerkezetét utánozzák.)

10. Genetikus algoritmusok

A gépi tanuló algoritmusoknak olyan osztálya, amely a biológiai evolúciós elméleten alapul. Az ilyen számítások lényege, hogy a problémákat evolúcióra hasonlító módszerek alkalmazásával akarjuk megoldani, azaz a feladatként tételezett problémát legjobban kielégítő szabály lesz a vizsgálat „nyertese”. A számos alternatív megoldás kidolgozása és ellenőrzése miatt kifejezetten számításgépi technika.

3.5. Az adatbányászat sikeres alkalmazása a gyakorlatban

Befejezésül nézzünk néhány konkrét példát arra, hogyan is lehet alkalmazni ezeket az eszközöket a gyakorlatban:¹⁸

- A *Bank of America* vezetői hozzáférhetnek az egyes vevők személyes adataihoz, amelyek alapján rögtön – amíg az ügyfél még a bank épületében – bővíthet értékesítésbe kezdenek.

- Az *MBNA* bank kitüntetett figyelemben részesíti nyereségesebb ügyfeleit, illetve a bankszférában megszokott átlag felére csökkentette a vevők lemorzsolódási arányát, aminek hatására nyeresége megtizenhatszorosodott.

- A *1-800-FLOWERS.com* automatikusan figyelmezteti vásárlóit a számukra fontos dátumok (pl. születésnapok, évfordulók) közeledtére, amely jelentős forgalomerősödést jelentett.

- A *Hertz* titkos hírszerzési technológiával gyűjt versenytársairól folyamatosan információt annak érdekében, hogy képes legyen gyorsan reagálni a változó piaci feltételekre.

- A *Land's End* mindegyik vevőszegmentumnak más és más katalógust küld, és a későbbi termékajánlatok csiszolása végett gondosan nyomon követi a vevők vásárlási reakcióit.

Jegyzetek

- * A tanulmány az Adatvédelem az üzleti információ-menedzsmentben című, a XXVII. Országos Tudományos Diákköri Konferencia Állam- és Jogtudományi Szekciója „Médiajog, információs önrendelkezési jog” tagozatában I. díjat nyert dolgozat átdolgozott részlete.
- ¹ Hoffmann Istvánné: Stratégiai marketing, Aula Kiadó, 2000. p. 93.
- ² Dr. Balogh Zsolt György: Az infokommunikációs jogról, In: Infokommunikáció és jog, 2. szám, Dialóg-Campus Kiadó, p. 45.
- ³ Dr. Balogh Zsolt György: Az infokommunikációs jogról, In: Infokommunikáció és jog, 2. szám, Dialóg-Campus Kiadó, p. 45.
- ⁴ Balogh Zsolt György: Jogi informatika, Dialóg-Campus Kiadó, Budapest-Pécs, 1998. p. 169.
- ⁵ Balogh im.p. 171.
- ⁶ Hoffmann Istvánné im.p. 93.
- ⁷ Heteyi József: ERP rendszerek Magyarországon a 21. században, Computerbooks, Budapest, 2004. p. 55.
- ⁸ Heteyi im., p. 49.
- ⁹ Kotler, Philip; Jain, Dipak C.; Malcincee, Suvit: Marketinglépések, Park Kiadó, 2003. p.109.
- ¹⁰ Kotler, Jain, Malcincee im. p. 35. oldal
- ¹¹ Tapp, Alan: Direkt és adatbázismarketing, Műszaki Könyvkiadó, 1999. p. 41.
- ¹² Heteyi im.p. 76-78.
- ¹³ Kotler, Jain, Malcincee, im. p. 106.
- ¹⁴ Heteyi im.p. 66.
- ¹⁵ Hoffmann Istvánné im.p. 198.
- ¹⁶ Heteyi im.p. 66.
- ¹⁷ Barbara Seidl: Data-mining und Datenschutz, Masther Thesis, Universität Wien, 2003. p. 3.
- ¹⁸ Kotler, Jain, Malcincee:im. p. 108.