

KPI-tree modell fejlesztése predikciós eljárások alkalmazásával

Gáspár Sándor – Thalmeiner Gergő

A controlling fejlődése az elmúlt években a big data, mesterséges intelligencia, a különböző matematikai statisztikai módszertanok fejlődése és a digitalizáció által olyan modelleket hozott létre a menedzsment számára, amelyek hatékonyabb döntéshozást és tervezési folyamatot tettek lehetővé. A különböző predikciós illetve ok-okozati controlling modellek a döntéshozáshoz szükséges információtartalmat nagymértékben növelték. Számos új, különböző operatív mérési lehetőség és adatszerzési forrás jelent meg a gépészet és az informatika által létrehozott innovációk által. Ezen adatok önmagukban nagyon kevészer hordoznak információ tartalmat. Összességében egy strukturált rendszerben, illetve modellben viszont olyan információ halmazzá válhatnak ezek az adatok, amelyek akár képesek egy teljes szervezet működési modelljének ok-okozati összefüggéseit elemezni és a különböző menedzsment funkciókat a tervezéstől egészen az ellenőrzésig támogatni.

Kulcsszavak: controlling, kpi-management, kpi-tree, predikció

1. Bevezetés

Napjainkban a controlling, mint a menedzsment tudományok egyik fontos ága egy kiterjedt és széleskörű változáson megy keresztül. A különböző matematikai, informatikai megoldások lehetővé teszik a szervezetek számára azt, hogy képessé váljanak a különböző rendszereknek a holisztikus kiterjesztésére. Az információs rendszerek által felhalmozott adatokból különböző struktúrák mentén szükséges olyan információkat kinyerni amelyek relevanciájukkal elősegítik a menedzsment döntéshozását ez által pedig a termelékenyebb, hatékonyabb szervezeti működést idéznek elő.

A controlling fejlődő szerepkörével többek között lehetővé kell tegye azt is, hogy a tervezés, az ellenőrzés és az információellátás ne csak egy izolált rendszerként működjön, hanem egymásra épülve támogassák a vállalati célok elérését. Egy összehangolt controlling mechanizmus kiépítése folyamatos fejlesztési tevékenység, amely hozzásegíti a szervezetet ahhoz, hogy egy-egy fejlesztési, bevezetési ciklus eredményeként szervezetspecifikus irányítási mechanizmussal rendelkezzen (Blumné–Zéman 2014). Csak is olyan controlling rendszer lehet hatékony, amelyet a stratégiai célok és az azokat szolgáló folyamatok kontrollálására hoznak létre. Önmagában ugyan a controlling rendszerek nem lendítik fel a szervezetek működését, hiányuk azonban azt eredményezi, hogy az üzleti hatékonyság nem bontakozhat ki (Véry 2012). Jelenleg azonban a controlling szerepe változásokon megy keresztül, melyben a kihívásokat a nagyobb terjedelmű adathalmazok, kiterjedt feladatok (stratégiaalkotás, kockázatkezelés, fenntarthatóság szemlélete), és a jövőorientáltság (algoritmusok által támogatott predikciós modellek használata) jelentik.

Kutatási célunk, hogy kvalitatív módszert alkalmazva egy KPI-tree struktúra esettanulmányán, feltárásán keresztül megvizsgáljuk, milyen eljárások lehetségesek a modell prediktív továbbfejlesztésére. Tanulmányunkban a releváns szakirodalom feldolgozását követően, feltártunk egy KPI-tree esettanulmányt egy szervezet működésében, amelyet teoretikus módon prediktív módszerekkel tovább bővítettünk.

2. A digitalizáció és a Big Data jelentősége

Megfigyelhető, hogy 2012-től kezdve egyre népszerűbb lett mind az interneten mind a hagyományos sajtóban a Big Data fogalom. Pedig a története meglehetősen korábban, már 2003 környékén elkezdődött, a Google által használt GFS (Google Distributed File System) publikálásával (Sanjay et al. 2003). Ebben az időben rendelkezésre álló technológiai rendszerrel, egy 1 milliárd oldal adatait tartalmazó kereső rendszer csak nagyon magas hardver és üzemeltetési költségek mellett tudott volna működni. A költségek nagyrészt elsősorban a web kereső által előállított óriási méretű fájlok okozták, melyek kezelése az akkori fájlrendszerekben vagy adatbázisokban csak nehézkesen és hosszadalmasan volt megoldható. A hangsúly az akkori technológia skálázhatóságán volt, mivel ebben a méretarányban a folyamat már körülményesnek és költségesnek bizonyult. A GFS rendszere erre a kihívásra nyújtott megoldást, megkönnyítette az ilyen óriási fájlokkal való munkavégzést. A folyamatos fejlesztéseknek köszönhetően a létrejövő rendszerek már kinőtték a web keresés problematikáját és a 2006-os Hadoop alprojekt nevezetű fejlesztések során már függetlenedtek a web keresés szemléletétől. A projekt sikerességét tükrözi, hogy 2008-ban több cég, közöttük a Facebook vagy a New York Times is bejelentette a Hadoop technológia gyakorlatban történő alkalmazását (White 2012).

Lényegében a Big Data a nagy mennyiségben rendelkezésre álló adat felhasználásakor fellépő technológiai skálázódási problémára adott válaszként játszott fontos szerepet. Ezt követően még jelentős ideig ez volt a Big Data fő alkalmazási területe, ugyanakkor ez még nem indokolná a robbanásszerű népszerűséget. Közismertségének kulcsa abban rejlik, hogy a szervezetek többsége igényt tart a különböző elemzések, kimutatók készítésére, melyek alapja a nagy mennyiségű tárolt és feldolgozható formájú adat. Ezeket az adatokat korábban azért nem gyűjtötték, mert magas volt a tárolási költségük, csak korlátozott ideig volt a tárolásra kapacitásuk, vagy ugyan tárolták az adatokat, de nem tudták feldolgozni és transzformálni, mert a szükséges számítási algoritmusok nem álltak megfelelően rendelkezésre. Általánosan jellemző volt, hogy a kkv-k számára már közel néhány terabyte-nyi adat tárolása is komoly kihívásokat okozhatott, valamint az elemzéshez szükséges számítási metodikák, kapacitások sem voltak elérhetőek. A technológia fejlődésével, a Hadoop és a Big Data elterjedésével és elemzés szemléletű felhasználásával a szervezetek egyre több adatot kezdtek el gyűjteni, rögzíteni és feldolgozni, amely a közelmúlt elemzői számára elérhetetlen volt (Stadler 2015).

A digitalizáció hatására számtalan új információforrás keletkezésével és az egyre olcsóbb technikai eszközök kombinációjával új korszak tárul elénk. Olyan eszközök, mint például a mobiltelefonok, az online vásárlás folyamata, a szociális hálózatok, az elektronikus kommunikáció, vagy a GPS technológia működésük során

mind-mind adatforgalmat hoznak létre. Ezek az adatok jellemzően strukturálatlan formában, nem egy adatbázisban összpontosulva hatalmas mennyiségű jelként vannak jelen és csak a kiaknázásra, felhasználásra várnak. Nehéz meghatározni, hogy a világon mennyi adat áll rendelkezésünkre, ugyanakkor az IBM becslése szerint napi szinten 2,3 trillió gigabyte keletkezik. 2020-ra az adatmennyiség eléri a 40 zettabyte-ot, ami a 2005-ös évhez viszonyítva 300-szoros növekedést jelent. Az adatok keletkezéséhez, adatmennyiség növekedéséhez nagymértékben hozzájárul a digitalizáció, az információs és kommunikációs technológiák fejlődése, melyek az utóbbi években komoly változásokon mentek keresztül. A technológiák korszerűsödésével elérhetőbbé vált, hogy ez a hatalmas mennyiségben létrejövő adat összegyűjthetővé, feldolgozhatóvá, tárolhatóvá és számunkra megfelelő módon rendezhetővé vált. A technológiák közül kiemelendő a mobiltelefonok elterjedt és kiterjedtebb használata és a közösségi média befolyásoló hatása. Ezt az a tény is jelképesen szemlélteti, hogy a világ népessége, (7,7 milliárd ember) közül napjainkban megközelítőleg 6 milliárd ember rendelkezik mobiltelefonnal (IBM, 2015). Vagy példaként hozható fel a Google, amely naponta több mint 24 petabyte adatot dolgoz fel. Ez a napi szinten összegyűjtött adatmennyiség több ezerszerese az Egyesült Államok Kongresszusi Könyvtárában őrzött összes nyomtatott anyagnak. A Facebookra csak óránként több mint 10 millió új fényképes fájl kerül feltöltésre, napi 3 milliárd lájk, komment rögzül a platformon, amely során digitális nyomot hagynak maguk után a felhasználók. A cégek ezekkel az adatokkal értékes információkhoz juthatnak, beazonosíthatják és megismerhetik fogyasztóik preferenciáit (Mayer-Schönberger–Kenneth 2014).

Azon túlmenően, hogy a Big Data kifejezés leegyszerűsítve adatmennyiséget jelent, több kifejtett meghatározása is van, ezek közül az Oxford szótárak definíciója a következő: „Rendkívül nagy adatállományok, amelyek számításgényes analízisa során mintázatokat, trendeket és összefüggéseket lehet feltárni különösen az emberi viselkedés és kölcsönhatások vonatkozásában” (Oxford Dictionaries 2007). Egy másik megfogalmazásban: „Big Data nagy mennyiségű, sebességű és változatos adatok, amelyek költséghatékony módon, innovatív formában segítik a folyamatokba való jobb betekintést, a döntéshozatalt és a folyamatok automatizálását” (Gartner inc 2017). Egy 2001-es kutatási jelentésben Doug Laney a META Group (jelenleg Gartner) elemzője a Big Data-t háromdimenzióval, azaz mennyiség (volume), változatosság (variety), sebesség (velocity) szerint határozta meg, melyet az angol elnevezések kezdőbetűivel 3V-definíciónak is szoktak nevezni. A következőkben ismertetjük a 3V-t részletesebben:

- Mennyiség (volume) az összes létrehozott és rendelkezésre álló adat 90%-a az elmúlt két évben keletkezett. Az adatok mennyisége két évente megduplázódik, 2020-ra 50-szerese lesz a 2011-es adatmennyiségnek. Egy példán keresztül szemléltetve a Walmart több mint 2,5 petabytenyi (egy petabyte körülbelül 20 millió iratszekrény értékű szövegnek megfelelő adatmennyiség) adatot gyűjt minden egyes órában ügyfelei ügyleteiből.
- Változatosság (variety) a szervezetek által generált adatok 90%-a strukturálatlan (jellemzően strukturálatlan, félig strukturált, elvéve strukturált) adatnak minősül. Az adatok jellemzően olyan forrásból

származhatnak, melyek valamilyen emberi tevékenység vagy eszköz által nyomot hagy az interneten. Az adatok ma sokféle formátumban és típusban érhetők el, melyek különböző típusú elemzéseket és eszközök használatát igényli.

- Sebesség (velocity) a sebesség komoly szerepet játszik az adatok létrehozásában, tárolásában, elemzésében és vizualizálásában. Amíg régebben a számítógépek és a szerverek jelentős időt igényeltek az adatok feldolgozására és az adatbázisok frissítésére, addig a Big Data korszakban az adatok valós időben vagy közel valós időben állnak rendelkezésre és riportálható formába (Devan 2016).

Maga az adatfeldolgozás a szervezetek egyik legfontosabb feladata, ennek során két fő kategóriát különböztetnek meg: adatok kezelését és az adatok elemzését. Az adatkezelés során a hangsúly az adatok olyan kezdetleges szintű alkalmazásán (gyűjtésén, tárolásán) van, amelynek célja, hogy az esetleges jövőbeli feldolgozásukat teremtse lehetővé. Az adatok elemzésének lényege pedig az, hogy az adatok feldolgozása által a szervezetek részletesebb betekintést nyerhessenek a folyamataikba és alkalmazható tudást szerezzenek, amelyet implementálva a szervezeti folyamatok működése fejleszthetővé válik. A két cél között fellelhető különbségek mellett a nagy mennyiségű adatok feldolgozásával kapcsolatosan számos közös kihívás található. Az alábbiakban néhány fontosabb kihívási tényezőt ismertetünk.

- Algoritmusok: Elmondható, hogy minden számítógépes feldolgozás középpontjában az algoritmusok állnak. Annak ellenére, hogy a közelmúltban komoly előrelépések történtek az informatika, az adatbányászat, az elemzés és az egyéb tudományágak terén, az algoritmikus feldolgozás még mindig viszonylag korlátozott feladatokra alkalmas, ahhoz képest, mint amennyi adat létrejön. Az ismétlődő matematikai számításokon és logikai műveleteken túlmutató magasabb rendű kognitív feldolgozáshoz hasonló feladatok elvégzésére még teljes mértékben nem alkalmasak a jelenleg meglévő algoritmusok. Ezért fontos, hogy a szervezetek reális elvárásokat támasszanak afelé, hogy mit tudnak elérni jelen algoritmusokkal és az összegyűjtött adatokkal.
- Skálázhatóság. A skálázható feldolgozó rendszerek megfelelnek a növekvő adatokból származó feldolgozási igényeknek és kihívásoknak, illetve rugalmas módon bővíthetők a működés során felmerülő igényekhez. A szervezetek többségének a legmegfelelőbb bővítés lineáris, ez azt feltételezi, hogy az erőforrásigény lineárisan arányos az adatok növekedésével. Ebből következik, hogy a lineáris skálázhatóságnak tükröződnie kell mind a számítási teljesítményben, mind a pénzügyi erőforrásokban egyaránt. Példának okáért, ha az algoritmusok által eszközölt lineárisan növekvő számítási teljesítmény exponenciális költségnövekedést eredményez, akkor végül a feldolgozó rendszerek túlságosan költségesekké válnak, ez pedig azt eredményezi, hogy a szervezet számára túlságosan költséges a nagy adathalmazok feldolgozása.

- **Időbeliség.** Az időbeli követelmények felállítását követően az eredmények két fő feldolgozási módra sorolhatók: online és offline feldolgozásra. Az online mód az adatok azonnali feldolgozására utal, valós idejű rendszerek használatával az adatok azonnali elérhetővé tételével. Az offline mód az adatfeldolgozás során némi átfutási, késleltetési időt követően bocsátja rendelkezésre az adatokat. Az online és offline feldolgozási módok eltérő követelményeket támasztanak a számítást végző rendszerek és algoritmusok teljesítményére. Az azonnali feldolgozást igénylő nagy adatok lényegesen nagyobb számítási teljesítményt, technológiai fejlettséget igényelnek, mint az offline mód során feldolgozásra kerülő adatok, melyek feldolgozása során akár több nap is a rendelkezésre áll.

2.1. Változások hatása

A controlling rendszerek egyik legfontosabb feladata az előrejelzések készítése, amely az idő teltével több lépcsőben, különböző fejlettségi szinteken ment keresztül. A kezdetben még csak megérzésen alapuló előrejelzéseket (Forecasting 1.0) felváltotta az Excel alapú előrejelzések készítése (Forecasting 2.0). A módszertani fejlődés során a következő lépés az üzleti intelligencia rendszerekben történő előrejelzések (Forecasting 3.0) voltak. Az előrejelzések legújabb szintje pedig a Forecasting 4.0, ami prediktív elemzési módszerekkel, fejlett statisztikai algoritmusok támogatásával készül. Ez azt jelenti, hogy a nagy adathalmazok felhasználásával, algoritmusok által feldolgozott, statisztikai-matematikai számításokkal automatizált előrejelzések sokasága készíthető. A forecasting által felgyorsul és leegyszerűsödik az előrejelzések készítése, így rendszeresebben lehet forecast-tény eltéréseket vizsgálni. Ennek hatására az akciókényszer nő, a részletes és pontos előrejelzések hatására a szervezetek hatékonyabban tudnak reagálni a piaci hatásokra. Ebből következően a szervezeti irányítás paradigmaváltáson megy keresztül, analitikus-reaktív helyett proaktív-előre mutatóvá fejlődik (Gulyás 2017). A megalapozott előrejelzések készítéséhez az egyik legfontosabb eszköz az úgynevezett Key Performance Indicators (röviden: KPI) meghatározása. A KPI-ok alatt olyan mutatószámok meghatározását értjük, amelyek a folyamatok követésére specializálódnak és azokról szolgáltatnak információt a menedzsment számára. A mutatószámokat többek között felhasználják a stratégiaalkotás során, a következő időszak célkitűzéseinek meghatározásában, valamint azon területek jellemzésére, melyek komolyabb erőforrásokat igényelnek (Boda et al. 2005).

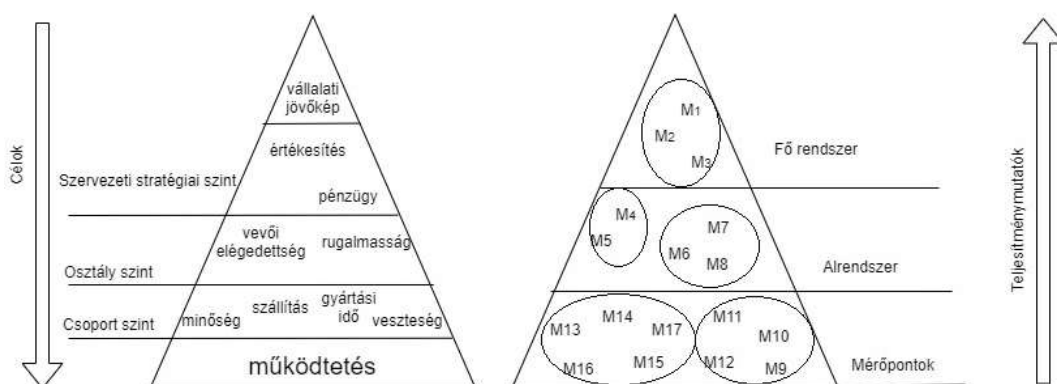
2.2. KPI-tree

A KPI-tree egy olyan controlling modell, amelyet a legtöbb esetben az iparban tevékenykedő multinacionális szervezetek alkalmaznak, illetve fejlesztenek. A tree-ben szereplő KPI mutatókat a modell adott célok és korrelációk mentén csoportokba szervezi (Schnellbach–Reinhart 2015) és egy egymásra épülő logikai felépítettség mellett rendszerezi (Ante et al. 2018). A Big Data, az adatbányászás technológiai feltételeinek fejlődése és az ipar 4.0 által generált adatok új KPI-ok megalkotására és a már meglévő KPI-ok akár percre pontos adatgyűjtésére is lehetőséget adnak (Peral

et al. 2016). A kulcsmutatók fontos szerepet képviselnek a napi szintű terv-tény elemzésekre és az ezen alapuló objektív napi szintű kimutatásokra, illetve a gazdálkodás szervezési folyamatok mérésére is (Schnellbach–Reinhart 2015).

A KPI-tree modell felépítésére a szakirodalom konkrétan nem határoz meg struktúrát, annak felépítése kreatív módon valósul meg. Az 1. ábrán szemléltetünk egy lehetséges felépítési módot, mely során kiemelendő, hogy a célok fentről lefelé, a mutatók pedig lentől felfelé épülnek egymásra, és oszlanak meg a különböző szinteken.

1. ábra Célok és mérési pontok a szervezeti rendszerekben



Forrás: saját szerkesztés Ante et al. (2018) és Schnellbach–Reinhart (2015) alapján

3. Kutatás és módszertan

Kutatásunk az Innovációs és Technológiai Minisztérium ÚNKP-19-3-I kódszámú Új Nemzeti Kiválóság Programjának szakmai támogatásával készült.

Kutatásunk során egy kiterjesztett esetelemzést végeztünk el. Esettanulmányunkként egy Pest megyei autóalkatrészgyártó multinacionális cég szolgált. Azért választottuk a kiterjesztett esetelemzés módszert, hogy felfedezzük a meglévő elméletek, illetve a gyakorlatban alkalmazott módszerek hiányosságait és a gyakorlatban alkalmazott modellt egy adott teoretikus szakirodalomban megfogalmazott modell mentén tovább fejlesszük különböző paraméterek mentén (Babbie 2013).

Célul tűztük ki a KPI-tree modell logikai felépítésének elemzését a vizsgált autóalkatrészgyártó szervezet menedzsment rendszerében. Kutatásunkban a KPI-tree kulcs mutatószámának definiálása alapvetően a menedzsment döntése. Ezen döntés sokrétű és számos funkcióra kiterjedhet. Ez az oka annak, hogy a szemléltetett ábrán „cél” KPI-oknak definiáljuk a mutatókat és a hozzá tartozó részegységeket.

4. Eredmények

Az ipari szervezetek számára a megfelelő információk alapján megvalósuló döntéselőkészítés, adatok strukturálása egy olyan kihívás, amely már évtizedek óta különböző más-más szemléletű módszereket hoz létre a piacon. A különböző információs rendszerek hatékonyan gyűjtik és egy bizonyos struktúra alapján rendszerezik, illetve időszakra vonatkozóan jelentéseket készítenek a releváns információkkal a vezetőknek.

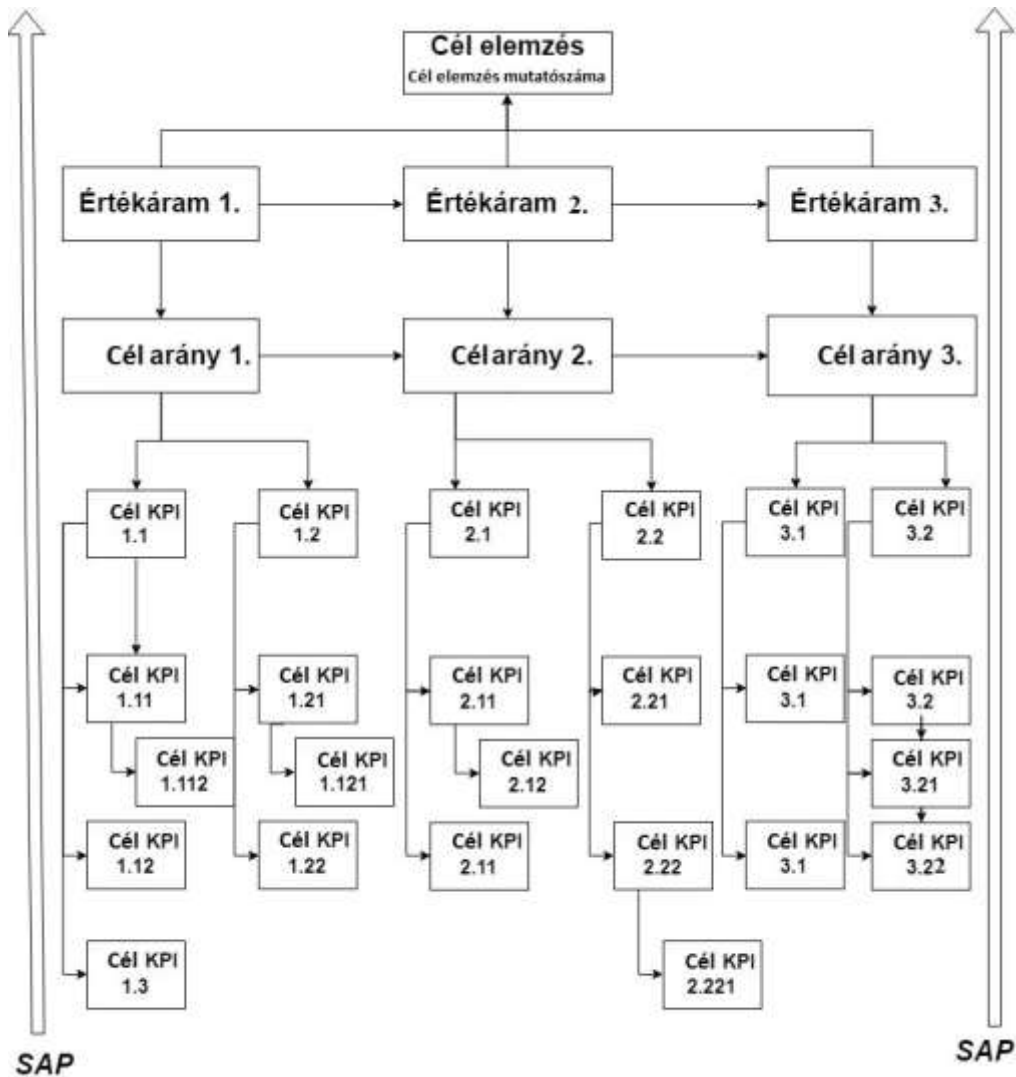
Az alábbiakban szemléltetett KPI-tree egy olyan módszer, amit a legtöbb esetben autó-autóalkatrész gyártó szervezetek alkalmaznak, illetve fejlesztenek. A vizsgált szervezetben a KPI-tree egy belső fejlesztésű rendszerként jött létre és működik. A KPI-tree kimondottan alkalmas lehet a különböző gazdálkodásszervezési folyamatok mérésére, mivel egy adott struktúra mentén képes hierarchikus sorrendbe állítani a különböző csoportba tartozó KPI-okat. A termelési folyamatokhoz kötődő lean KPI-ok számos esetben determinisztikus jellegűek. Ebből következik, hogy egy adott viszonyszám definiálásával mérhető a súlyozottság és ezáltal az ok-okozati összefüggések is kiválóan riportálhatók.

A vizsgált vállalatban értékáramok alapján történik a gyártás, ami azt jelenti, hogy a különböző módokon meghatározott értékáramok, csoportok jönnek létre. Ezekben az értékáramokban egy-egy komplex szaktudással rendelkező csoport működik egy adott gyártási folyamaton-folyamatokon, vagy egyes esetben különböző félkész termékeken. Az értékáramokhoz tartozó controll rendszerekben lévő mért mutatók egy részét felsővezetők illetve a controllerek előre definiálják. Az ezeken a definiálásokon kívüli mutatókat pedig az értékáram és az értékáram vezetője definiálja az értékáram hatékonyság érdekében. Az értékáramok összessége és az ezekhez megfogalmazott KPI-ok összessége adja meg a szervezet gyártásának teljes controll rendszerét.

A KPI-tree dinamikus rendszerként működik az értékáramok között és a hierarchiában magasabb helyet foglaló vezetők számára történő riportolás hatékonyságát segíti. Az információkat, amelyek a KPI-ok értékeléséhez szükségesek minden nap feltöltik az adott értékáramok adminisztrátorai, ezáltal a napi szintű jelentés készítés is lehetségessé válik. A vizsgált vállalat az SAP-t használja információs rendszerként. Ide töltik fel az adatokat a különböző értékáramba tartozó adatfelvitelre jogosult adminisztrátorok, amelyeket aztán a controller a KPI-tree rendszerbe integrál. A KPI-tree csúcsmutatója mindig egy olyan aggregált KPI, amelyet az összes értékáramban definiálva van.

A 2. szemléltető ábrán a különböző KPI-ok, illetve mérőpontokhoz tartozó mutatók mind pénznemben, mind a terv-tény, mind pedig pénzbeli és mennyiségi viszonyszámaként mutatható ki. A különböző százalékos arányok kimutatásával a teljeshez viszonyított arány és az adott mutató önmagához viszonyított arányaként egy ok-okozati struktúrát is meghatároz. A célhoz viszonyított arány, ami jellemzően a múltbéli adatok és egyéb controllerek által igénybe vett módszerek alapján definiált cél és adott időszak tény adataiból származó információ, az adott vizsgált részelem vagy mutató, részleg fejlődését és teljesítményét fejezi ki.

2. ábra KPI-tree modell



Forrás: saját szerkesztés Ante et al. (2018) alapján

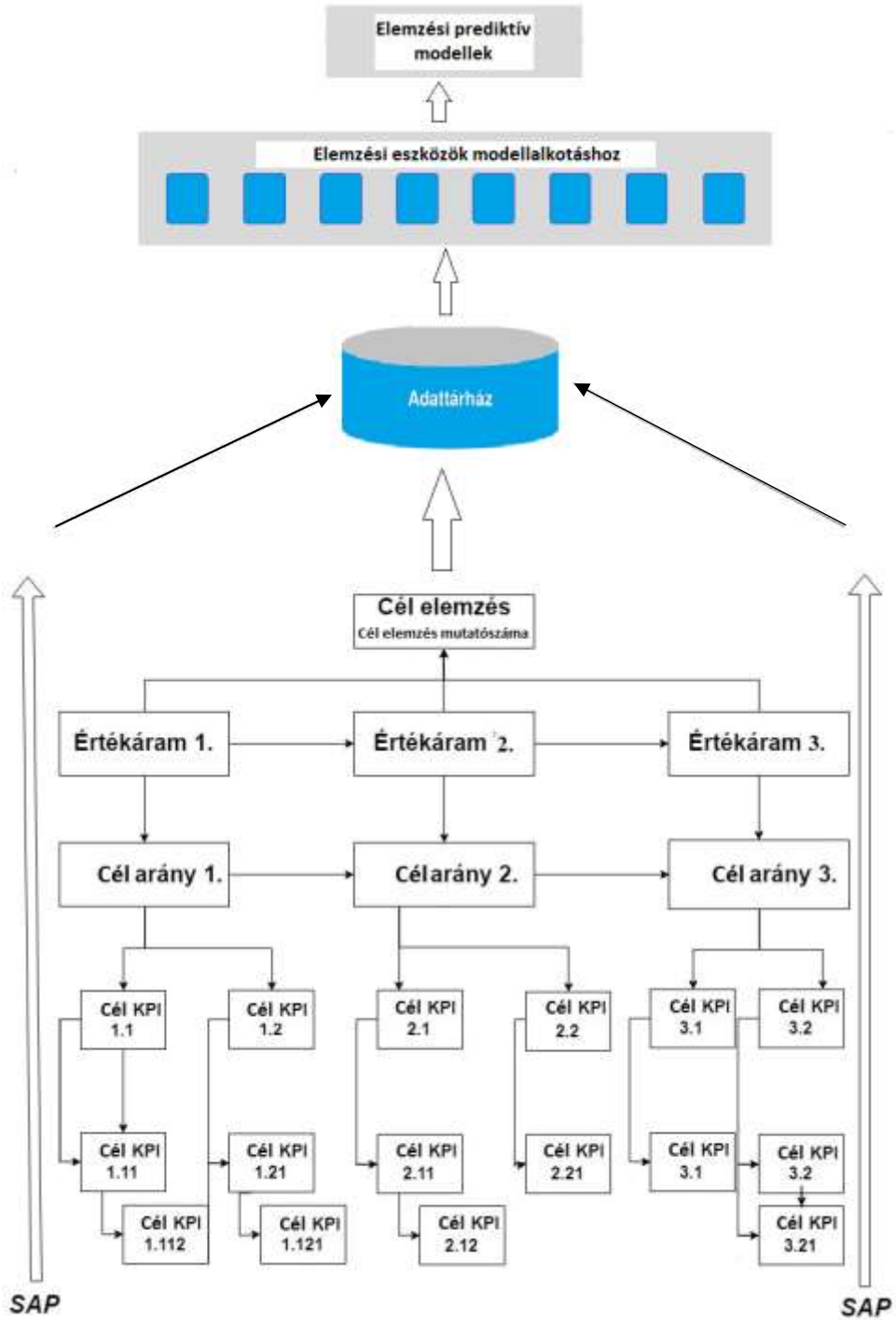
A KPI-tree-ből származó információk, mind időbeli mind a különböző funkcionális szintekről való felhalmozását követően egy olyan adattárházba kerülnek elraktározásra, amely meg kell, hogy feleljen a Big Data fejlődése során rendelkezésre álló különböző tulajdonságú adatok tárolásának. Ezen követelmények hardveres és szoftveres téren is kihívások elé állítja az adattárházak felépítettségét. Az adattárháznak biztosítania kell az elemzések készítéséhez szükséges adatok tárolását, elérhetőségét és azt, hogy az adatok az idő során ne sérüljenek, valamint továbbra is felhasználhatók, feldolgozhatók maradjanak. A működés során fontos kritérium, hogy a szervezetben alkalmazott ERP rendszer (jelen esetben az SAP rendszer)

összeköttetésben kell, hogy legyen az adattárházzal. Az adattárolásból felmerülő problémák redukálásának egyik leghatékonyabb módja lehet, ha az adattárház egy külső felhő alapú rendszer. Ezt az SAP S/4 HANA rendszer úgy biztosítja, hogy a szolgáltató már nem csak az ERP rendszerként működő SAP szoftvert biztosítja, hanem további külső szolgáltatásként az adattárházi szolgáltatást is nyújt. Ezen termék és szolgáltatás integrálásával elérhető, hogy az adatgyűjtésből és adattárolásból felmerülő kockázatok és hibák minimalizálásra kerüljenek. Az SAP S/4 HANA egy nyílt platform és a felhő megoldáshoz tetszőleges gyártók, illetve fejlesztőktől is lehet kiegészítőket és applikációkat kapcsolni, vagyis egy belső fejlesztés során létrejövő KPI-tree rendszer is összekapcsolható vele.

Az elemzési eszközök modellalkotásához az összegyűjtött releváns - kevésbé releváns külső, belső adatok és az akár nem feltétlenül transzparens adatok strukturálásával megteremtődik a lehetőség az adatok predeterminisztikus algoritmusokkal való feldolgozásához. Az adatok strukturálódását, valamint azok elemzésbe történő bekerülését nagyban meghatározza a előrejelző algoritmusok által végzett elemzések kimeneti célja. A kimeneteli célt a szervezet számos kritériumnak megfelelően határozza meg. Jelenlegi kutatásunk nem terjed ki a célok meghatározására annak összetettségének vizsgálatára. Azonban általánosan megfogalmazható, hogy a célok meghatározása legtöbb esetben top-down módszer alkalmazásával történik, de az általunk vizsgált szervezetben megfigyelhető a célok megfogalmazása esetében a top-down és bottom-up tervezési metodika is.

A KPI-tree módszer kibővítése és a módszerből származó adatok további elemzése az alábbi prediktív elemzési modellel történhet. Ezen módszerek nem feltétlenül univerzálisak. Jelenlegi kvalitatív kutatásunkban csak néhány szemléltető és KPI-tree módszer továbbfejlesztésére alkalmas prediktív elemző eljárást sorolunk fel. Nem tudunk és nem is szándékozunk a KPI-tree módszer komplex és általános prediktív modell kifejlesztésére. Ezen módszerek viszont alkalmasak a jelen vizsgálatba vont KPI-tree módszer továbbfejlesztésére.

3. ábra KPI-tree alapú prediktív elemzés IT- architektúrája



Forrás: saját szerkesztés Ante et al. (2018) és Baars (2016) alapján

4.1. Neurális háló alapú prediktív modell

A neurális háló úgy jellemezhető, mint egy a biológiai neurális rendszerek elvére felépített, hardver-szoftver jellegű, párhuzamos, osztott működésre képes információfeldolgozó eszköz. A hálók neuronokból állnak és a biológiai idegrendszerhez hasonlóan próbálja megvalósítani az információ felvételt, illetve feldolgozást. A problémák, feladatok megoldására tanulás útján jutnak el, ellentétesen a hagyományos algoritmikus eljárásoktól. A neurális hálók alapegysége a neuron. Ezek a neuronok olyan rendszert építenek fel amely „n” bemenettel és „m” kimenettel rendelkezik és amely az n-dimenziós bemeneti vektorokat m-dimenziós kimeneti vektorokká alakítja át az információfeldolgozás során (Álmos et al. 2002). Elmondható, hogy a neuronok rétegekbe szerveződve tevődnek össze, a neurális háló jellemzően három fő rétegből áll: a bemeneti, köztes és a kimeneti rétegből.

- Bemeneti rétegen belül olyan neuronok helyezkednek el, amelyek ismert információkból vagy a hálóba betáplált változókból állnak.
- Köztes rétegben az input neuronok kapcsolatban állnak a köztes réteggel, ezen kapcsolatok a bemeneti neuronok fontossága szerint vannak súlyozva, amelyek a tanulás időszak alatt folyamatosan változnak.
- Kimeneti rétegben eredményneuronok találhatóak, ezek is súlyozottan kapcsolódnak a köztes rétegben található neuronokhoz.

A neurális hálók a tanulást példákön keresztül végzik, csak úgy mint a biológiai megfelelőik (Gurney 1996). A tanulás során a bemeneti mintázatok alapján megváltoztatja a kapcsolatok súlyát, ezáltal pedig kialakul a háló végleges súlyozása. Amennyiben egy neurális hálót első alkalommal ellátunk adatokkal, a háló elkezd véletlen találgatások mentén keresni a megoldást. A súlyokat a tanulás mentén a tényleges megoldástól való eltérés függvényében fogja módosítani. Ezt a folyamatot egy olyan iteratív eljárásnak nevezzük, amely során a neurális háló által végzett leképezést egy kívánatosnak vélt leképezéshez hasonlítjuk. Azonban fontos kiemelni, hogy kis minták esetén előfordulhat a túltanulás problémája. Túltanulás során nem az általános problémát tanulja meg a neurális hálózat, hanem a megadott adatbázis sajátosságait. Ennek elkerülése végett ketté kell osztani az adatbázist tanulási és tesztelő mintákra. Kezdeként a tanuló adatbázison végezzük el a tanítást, majd futtatjuk a tesztelő mintán is. Abban az esetben, ha a tanulási mintához hasonlóan a találati pontosság kedvező, akkor a tanulási folyamat eredményesnek értékelhető. Viszont abban az esetben, ha a tesztelő minta hibázása jelentős, akkor a neurális hálózat túltanulta magát. Az eljárás lefuttatása során 0-val jelölöm a fizetéképtelen, 1-gyel pedig a fizetőképes vállalatokat.

A neurális háló súlyai jellemzően a backpropagation eljárás segítségével alakulnak ki. A neurális háló modell tanulási és tesztelő eljárása során felépíthető a sematikus összegző táblázat, melynek célja összefoglalni a minta hibáit valamint besorolási pontosságát. (Kristóf 2002)

A neurális háló és a KPI-tree módszer közös alkalmazásával a bottom-up célmeghatározási módszer lehetőségessé válik. Ezáltal pedig egy szakértői rendszert és egy objektív tervezési és elemzési rendszert biztosít a menedzsment számára. A különböző termelési, logisztikai és lean folyamatokhoz kötődő controlling rendszer

egyik meghatározó funkcióját az erőforrás allokációt is kiválóan lehetséges prediktíven meghatározni a neurális háló segítségével. Ugyanakkor meghatározó a túltanulás kockázata, ami abból adódhat, hogy a KPI-tree módszerével megadott struktúra mentén gyűjtött adatok a neurális hálóba kerülve már csak az adott struktúra mentén felépülő tanulási mintázatoknak megfelelően fogja szolgáltatni a kimeneti értékeket. Ezért fontos a tesztelési fázis futtatása és annak folyamatos kiértékelése.

Összességében elmondható, hogy a napjainkban ismert eljárások közül a neurális hálók tekinthetők a legpontosabb és legmegbízhatóbb predikciós elemzési módszernek, ezáltal pedig a legmegfelelőbb alkalmazható elemző eljárás.

4.2. Rekurzív particionáló algoritmus

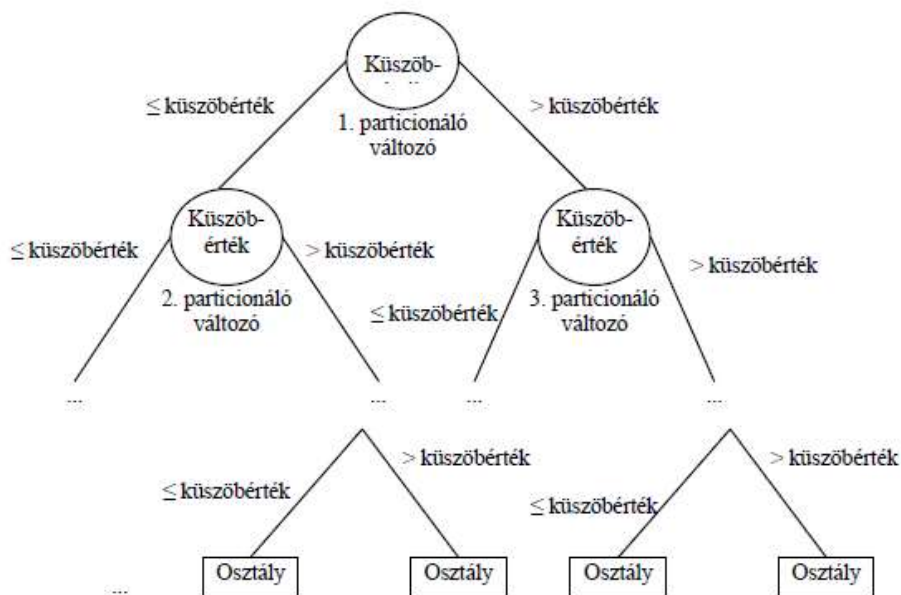
A rekurzív particionáló algoritmus egyszerű szabályok mentén döntési fákat hoz létre. Az algoritmus lépésenként kétfelé osztja az adatokat elágazásokat képezve. A mintába szereplő változókat egyesével megvizsgálva a leginkább elválasztó értékkel rendelkező változó mentén építi fel a fát. Elsődleges célok közé tartozik a döntési fa létrehozása során a homogén osztályok előállítás. Továbbá fontos szempont a variancia megléte, a megfigyeléseket úgy csoportosítja a szervezet a függő változó szempontjából, hogy a csoportokon belüli variancia minél kisebb, míg a csoportok között minél nagyobb legyen (Hámori 2001).

A rekurzív particionáló algoritmus alkalmazása abban az esetben a legegyszerűbb, amikor bináris elválasztások mentén képezünk két osztályt. Az elválasztó változó kijelölésére az algoritmus kipróbálja az összes input változót. Ahogy az összes kétfelé osztás megtörtént, az a változó kerül kijelölésre, amelyik a legkisebb hibát követi el ezáltal pedig a legjobb homogenitást eredményezi. A fa további szintjei is hasonló eljárási módban épülnek fel, így létrejön a teljes struktúra. A fa tetején a particionáló változó, legalul pedig a különböző osztályok találhatók.

Hasonlóképpen a neurális hálókhoz, ennek az algoritmusnak a lefuttatásának is fontos kritériuma a túltanulás elleni védekezés. A rendelkezésre álló adatokat tanulási és tesztelési fázisokra kell osztani. A tanulási mintára felépített döntési fát a tesztelő mintán való alkalmazás iterációi során fokozatosan „meg kell nyesni”, ameddig az osztályba sorolási hibák megfelelően közel nem esnek egymáshoz. Az így létrejött és megnyesett döntési fa képezi az előrejelzési modellt, amelyet ezután tetszőleges adatokon lehet alkalmazni.

Egy döntési fa létrehozásakor általában körökkel jelzik a változókat, valamint négyzetekkel az osztályokat. A körökbe írt számok az elágazási pontnak megfelelő értéket jelentik. Amennyiben adott változónak az értéke kisebb vagy egyenlő, mint az elágazás, akkor a bal oldali ágra kerül, különben a jobb oldalra irányul. Az ágakon megfigyelhető számok darabszámokat jelentenek, amelyek a feltételnek eleget tevő megfigyeléseket jelenti. A fa alján elhelyezkedő négyzetekben a megfelelő osztályok elnevezése szerepel. Az eljárás szimulációs kísérletezésen alapszik, továbbá a módszer fontos kritériuma a besorolási pontosság és a gyakorlati hasznosíthatóság.

4. ábra Rekurzív particionáló algoritmus döntési fája



Forrás: saját szerkesztés (Kristóf 2005) alapján

A rekurzív particionáló algoritmus az értékáramok adataiból kinyert KPI-tree mutatóit vonja be az eljárásba. A kísérletezések követően a megfelelő besorolási pontosság mellett létrejön egy olyan prediktív elemzés, amely előrevetíti a jelen elemzésbe bevont mutatókon alapuló jövőbeli teljesítményt. Az elemzés során létrejövő jövőbeli teljesítmény értékelését követően, visszafelé haladva a KPI-tree mutatóinak külön-külön való elemzésével pontosan meg lehet határozni a fejlesztendő területeket a jövőbeli összteljesítmény növelése érdekében.

5. Összefoglalás

A KPI-tree módszer egy hatékony, megfelelő és a többi komplex controlling módszerhez képest egyszerűen alkalmazható, a folyamatok és a hozzájuk tartozó költségek strukturálásában. A módszer alapvetően „mi történt” és „mi történik” kérdésekre ad választ viszont arra, hogy „mi fog történni” és „mi kellene, hogy történjen” nem ad választ. Kutatásunkban teoretikus szempontból próbáltuk ezen problémát kielemezni és megoldást találni. A neurális háló és a KPI-tree módszer közös alkalmazásával a bottom-up cél meghatározási módszer lehetségessé válik. Ezáltal pedig egy szakértői rendszert és egy objektív tervezési és elemzési rendszert biztosít a menedzsment számára. A KPI-treeből származó adatok felhasználása a rekurzív particionáló algoritmus segítségével előrevetíti a jelen elemzésbe bevont mutatókon alapuló jövőbeli teljesítményt.

A fentiekben említett két módszeren kívül számos eljárás alkalmazható KPI-tree módszerén alapuló prediktív elemzések végzésére. A KPI-tree módszer kialakítása és alkalmazása mindig szervezet specifikus, ezáltal nincs univerzális modell, ezért a prediktív módszerek kiválasztása és alkalmazása esetén figyelembe kell venni a KPI-tree módszer sajátos jellemzőit.

„Az Innovációs és Technológiai Minisztérium ÚNKP-19-3-I. kódszámú Új Nemzeti Kiválóság Programjának szakmai támogatásával készült.”



Felhasznált Irodalom

- Álmos, A. – Győri, S. – Horváth, G. – Várkonyiné Kóczy, A. (2002): *Genetikus algoritmusok*. Typotex Kiadó, Budapest.
- Ante, G. – Facchini, F. – Mossa, G. – Digiesi, S. (2018): *Developing a key performance indicators tree for lean and smart production systems*. IFAC PapersOnline, 51–11, 13–18.
- Baars, H. (2016): Predictive Analytics in der IT-basierten Entscheidungsunterstützung: methodische, architektonische und organisatorische Konsequenzen. *Controlling: Zeitschrift für erfolgsorientierte Unternehmenssteuerung*, 3, 28, 174–180.
- Babbie, E. (2013): *The practice of social research* (13th. kiad.). USA Belmont: Wadsworth, Cengage Learning.
- Blumné, B. K. – Zéman, Z. (2014): Controlling a vezetés szolgálatában. Történeti fejlődés, perspektívák. *Gazdálkodás- és szervezéstudományi folyóirat A Virtuális Intézet Közép-Európa Kutatására Közleményei*, 6, 1–2, 440–442.
- Boda, Gy. – Szilávik, P. (2005): *Kontrolling Rendszerek*, KJK-KERSZÖV Jogi és Üzleti Kiadó Kft. Budapest, 224–235.
- Devan, A. (2016): *The 7 V's of Big Data*. Forrás: Impact Radius blog. <https://www.impactradius.com/blog/7-vs-big-data/> (olvasva: 2019. 04. 15.)
- Gartner, INC. (2017): *Big Data. IT Glossary*. Forrás: <http://www.gartner.com/it-glossary/big-data/> (olvasva: 2019. 04. 11.)
- Gulyás, A. (2017): *Forecasting 4.0: Miért érdemes a jövő előrejelzési módszerét használni?*, Forrás: Menedzsment és Controlling Portál, <https://www.controllingportal.hu/forecasting-4-0/> (olvasva: 2019. 04. 02.)
- Gurney, K. (1996): *Neural nets*. Department of Human Sciences, Brunel University, Uxbridge.
- Hámori, G. (2001): A CHAID alapú döntési fák jellemzői. *Statisztikai Szemle*, 79, 8, 703–710.
- IBM (2015): *The four V's of big data* Forrás: https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg (olvasva: 2019. 04. 13.)
- Kristóf, T. (2002): *A mesterséges neurális hálók a jövő kutatás szolgálatában*. Jövőelméletek 9. BKÁE Jövőkutató Kutatóközpont, Budapest.
- Kristóf, T. (2005): A csodelőrejelzés sokváltozós statisztikai módszerei és empirikus vizsgálat. *Statisztikai szemle*, 83, 9, 841–863.

- Mayer-Schönberger, V. – Kenneth, C. (2014:) *Big Data – Forradalmi módszer, amely megváltoztatja munkánkat, gondolkodásunkat és egész életünket*. HVG Kiadó Zrt., Budapest, 11–17.
- Oxford Dictionaries (2007): *Definition of Big Data*, Forrás:
https://en.oxforddictionaries.com/definition/big_data (olvasva: 2019. 04. 11.)
- Peral, J. – Maté, A. – Marco, M. (2016): Application of data mining techniques to identify relevant key performance indicators. *Computer standards and interfaces*, 1–20.
- Sanjay, G. – Howard, G. – Shun-Tak, L. (2003): *The Google File System Proceedings of the 19th ACM Symposium on Operating Systems Principles*. ACM, Bolton Landing, New York, 20–43. (olvasva: 2019. 04. 15.)
- Schnellbach, P. – Reinhart, G. (2015): Evaluating the effects of energy productivity measures on lean production key performance indicators. *Procedia CIRP*, 492–497. doi:<https://doi.org/10.1016/j-procir.2014.07.094>
- Stadler, G. (2015): Big Data – tömeges adatelemzés gyorsan. *HTE Medianet*, 70, 1, 44–47.
- Véry, Z. (2012): Controlling – Menedzsment Fórum. Forrás:
<https://mfor.hu/cikkek/projektmenedzsment.html> (olvasva: 2019. 04. 17.)
- White, T. (2012): *Hadoop: The Definitive Guide, Third Edition*. O’Reilly Media, Inc. Sebastopol, 17–22.