

EGY FOLYÓIRAT ADATAIT TARTALMAZÓ ADATBÁZIS KONCEPCIÓS MODELLJÉNEK ELKÉSZÍTÉSE

Hampel György

Abstract: A Jelenkori társadalmi és gazdasági folyamatok című folyóiratban számos magyar és külföldi szakember színvonalas publikációja jelent meg 2006 óta. Annak érdekében, hogy az eddig megjelent, illetve a későbbiekben megjelenő számokban található információk hatékony módon kinyerhetők legyenek, egy adatbázis létrehozásának ötlete merült fel, amelynek megvalósításához az első lépés a koncepcionális modell elkészítése, amely jelen cikk tárgya. A létrehozott egyed-kapcsolat modell – amely 3 egyedhalmazt, 39 attribútumot és 3 kapcsolatot tartalmaz – alkalmas arra, hogy relációs sémává konvertálva, majd SQL-alapú adatbázis-kezelőben megvalósítva kiszolgálja a folyóirattal kapcsolatos jelenlegi információs igényeket.

Abstract: Since 2006, several Hungarian and foreign experts have published high-quality publications in the Journal of Contemporary Social and Economic Processes. To efficiently extract the information from the publications, the idea to create a database arose. The first step to achieve this is to prepare a conceptual model which is the subject of the current article. The created entity-relationship model includes 3 entity sets, 39 attributes and 3 relationships. This model can serve the current information needs related to the journal when converted to a relational schema and then implemented in an SQL-based database system.

Kulcsszavak: adatbázis, adatmodell, egyed-kapcsolat modell, Jelenkori társadalmi és gazdasági folyamatok folyóirat, Szegedi Tudományegyetem, Mérnöki Kar

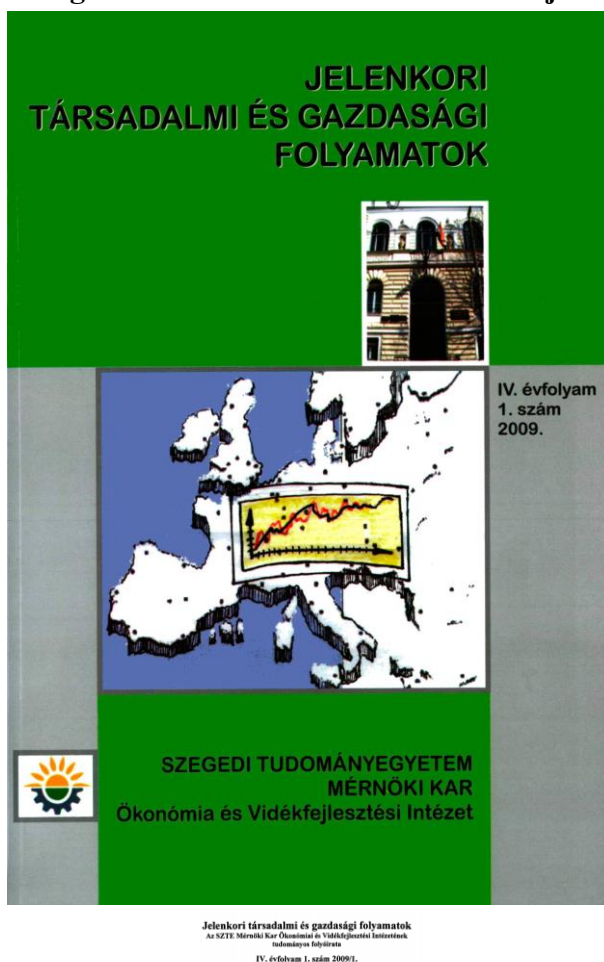
Keywords: database, data model, entity-relationship model, Journal of Contemporary Social and Economic Processes, University of Szeged, Faculty of Engineering

1. Bevezetés

A Szegedi Tudományegyetem Mérnöki Karának Ökonómiai és Vidékfejlesztési Intézete gondozásában először 2006 novemberében jelent meg az intézet oktatóinak lektorált tudományos publikációit tartalmazó nyomtatott évkönyv (ISBN: 978-963-482-799-3). Ahogy a kötet akkori főszerkesztője – Dr. Gulyás László PhD. – akkor írta, az intézet oktatói a kötetben bemutatják a kutatási témájuk egy kis darabját (Gulyás, 2006). Ezt követően 2007 decemberben jelent meg először „Jelenkori társadalmi és gazdasági folyamatok” címmel az újabb kötet, még szintén tudományos évkönyvként (ISSN: 1788-7593), és ahogy – Dr. Gulyás László PhD. és Dr. Gál József PhD. – főszerkesztők is jelzik az előszóban, ekkor már az intézet külső óraadóinak cikkei is helyet kaptak az évkönyvben (Gulyás–Gál, 2007). 2008 őszén tovább bővült a szerzők köre: négy szerzői csoportról beszélhetünk (Gulyás–Gál, 2008): az Ökonómiai és Vidékfejlesztési Intézet munkatársai, óraadói mellett megjelentek a hazai és külföldi szerzők cikkei is. Mára már gyakorlattá vált az is, hogy a kar hallgatói témavezetőikkel együtt jelentetnek meg egy-egy cikket: Panyor–Szebelédi (2018), Lendvai–Koppány (2018), Zsótér–Bagi (2020) csak néhány példa a teljesség igénye nélkül.

A 2009-ben megjelentetett 4. évfolyam 1. száma már „folyóirat”-ként nevezi meg magát (lásd: *1. ábra*) és a magyar nyelvűek mellett tartalmaz idegen nyelvű cikkeket is.

**1. ábra: A folyóirat borító címlapja 2009. (4. évf. 1. szám)
A grafikai terv Mónus Szilvia munkája.**



Forrás: Gulyás–Gál, 2008

Az azóta eltelt időben 2020. első feléig megjelent összesen 24 szám (ebbe beleszámítva a 2006. és 2007. évi évkönyvet is), a folyóirat jelenleg (2020. augusztusában) a 15. évfolyamánál tart és a 2019. évi (14. évfolyam) 3. szám már online ISSN számmal is rendelkezik (ISSN: 2676-9867).

Az eltelt évek alatt számos és szerteágazó tudományterületről származó cikk kapott helyet a folyóiratban; a 2020. 1. számot is figyelembe véve 454 publikáció jelent meg 456 különböző szerzőtől (illetve, ha figyelembe vesszük, hogy egyes szerzők többször is publikáltak és egyes cikkeknek több társszerzője is volt, a szerzők száma – az ismétlődéseket is beleszámolva – 848).

Dr. habil. Kis Krisztián felelős szerkesztő kitaró munkájának eredményeként 2020. évtől már MTA besorolással rendelkező minősített folyóiratról beszélhetünk (IX. Gazdaság- és Jogtudományok Osztály, Demográfiai Osztályközi Állandó Bizottság), amelyet az MTMT a tudományos folyóiratok között referál (Kis, 2020).

A töretlen fejlődést látva merült fel annak a gondolata, hogy létre kellene hozni egy folyóirat adatbázist, amely könnyebbé teheti a folyóirat példányaiból származó lényeges információk kinyerését.

A megvalósítandó feladat tehát egy adatbázis megtervezése magas szintű modellel (egyed-kapcsolat modellel), a logikai adatmodell elkészítése (relációs adatmodellel), végül az adatbázis tényleges megvalósítása (SQL relációs adatbázis-kezelővel).

Ebben a publikációban a folyóirathoz kapcsolódó adatbázis koncepcionális megtervezésével, vagyis az egyed-kapcsolat modell elkészítésével foglalkozom.

Az átgondolt tervezés előnyei:

- könnyen megérthető formában mutatja az adatbázis szerkezetét, az adatszoportokat és azok kapcsolatait, valamint az esetleges korlátozásokat,
- mind a felhasználóknak, mind pedig a programozóknak újabb ötleteket ad,
- mivel könnyen megérthető, segíti a felhasználó és a programozó közötti párbeszédet,
- olyan terv, leírás, amit az adatbázis-kezelő rendszer kiválasztásával, módosításával nem kell megváltoztatni.

2. Az adatbázis és logikai szintű megtervezése egyed-kapcsolat modellel

2.1 Adatbázis

Az adatbázis az adott területet jellemző adatokból, az adatok típusát és a kapcsolatokat leíró metaadatokból és az adatkezelő rendszerből áll (Tímár et al. 1997) vagy

„...véges számú egyed-előfordulásnak, azok egyenként is véges számú tulajdonságértékének és kapcsolat-előfordulásainak az adatmodell szerint szervezett együttese.” (Halassy, 1994: 38p.) vagy

„egy olyan integrált adatszerkezet, mely több különböző objektum előfordulási adatait adatmodell szerint szervezeten perzisztens módon tárolja olyan segédinformációkkal, ún. meta-adatokkal együtt, melyek a hatékonyság, integritásőrzés, adatvédelem biztosítását szolgálják.” (Kovács, 2004: 24p.)

És még lehetne sorolni jónéhány meghatározást, hiszen az adatbázisra nincs egységesen elfogadott definíció.

Egy adatbázis létrehozásának és használatának több előnye lehet, így (Kroenke, 2006; Kacsukné–Kiss, 2009):

- egységes, logikailag tiszta adatszerkezet,
- az adatok és az azokat kezelő programok függetlenek egymástól, ebből következően bármelyik módosítható a másik módosítása nélkül,
- az adatok flexibilis módon, könnyen lekérdezhetők.

2.2 Koncepcionális modell

Az adatbázist megtervezése során először koncepcionális modellt készítünk, majd az alapján készül el a logikai modell és csak ezt követi a tényleges, fizikai megvalósítás (jó esetben).

A koncepcionális adatmodell olyan eszközök gyűjteménye, amelyek célja a valóság leírása olyan módon, hogy a létrehozott modell legyen képes a valóságról feltett kérdések megválaszolására (Watson, 2006). Ez a modell nem foglalkozik az adatok szerkezetével és a fizikai tárolás problémáival sem.

A koncepcionális (magas szintű) adatmodell létrehozása elméletileg is és gyakorlatilag is jelentős szerepet tölt be az adatbázis-tervezési folyamatban (Dey et al., 1999). A koncepcionális sémák tartalmazzák az egyedek halmazát, azok jellemzőit, a kapcsolatokat és a korlátozásokat. Emellett szemléletes módon, ábrákkal mutatják be az adatbázis logikai felépítését; nem hozzáértő számára is érthető formában szemléltetik a valóság egy részét modellező adatbázist (Tímár et al., 1997). Természetesen, mint mindent, ezt is el lehet rontani; a hibás tervezéshez hozzájárul a modellező eszközök nem megfelelő alkalmazása, a felhasználók és az adatbázis-tervezők képzelőerejének korlátai, a felhasználók és az adatbázis-készítő szakemberek eltérő látásmódja és problémamegközelítése (Badia–Lemiere, 2011), valamint – bár a modell egyszerű – az lényeges, hogy a modellező tanulja meg/rendelkezzen azzal a képességgel, ami biztosítja számára a szükséges objektumok és kapcsolataik helyes azonosítását (Watson, 2006).

Akik az adatmodellezésben még nem kellően járatosak (például egyetemi hallgatók, akik adatbázis-tervezést tanulnak) Watson (2006) szerint a következő hibákat követik el leggyakrabban:

- nem ismerik fel, hogy egy attribútumnak vélt valami az valójában egyed,
- nem képesek egyedeket általánosítani és egyedhalmazhoz rendelni,
- nem ellenőrzik a kapcsolatot mindkét irányból, így a kardinalitás (összekapcsolódó egyedek számossága) megállapítása téves,
- esetleges kivételeket figyelmen kívül hagynak.

2.3 Tervezési alapelvek

Az adatbázis megtervezése során néhány hasznos alapelvet érdemes megfontolni, ezek: (Ulman–Widom, 2009)

1. Valóság-hű modellezés: Az egyedhalmazoknak, az attribútumoknak, valamint a kapcsolatoknak tükrözniük kell a valóságot (nem lehet például a CIKK egyedtípusnak tudományos fokozat attribútuma).
2. Rendundancia-mentesség (vagy legalább arra való törekvés): Törekedni kell arra, hogy minden adat lehetőleg csak egyszer szerepeljen. Például a szerző nevét – mint attribútumot – ne tároljuk el a CIKK és egyúttal a SZERZŐ egyedhalmazban is. Ezzel különböző (beviteli, tárolási, frissítési, és törlési) anomáliáknak nevezett, az adatbázis használatát és karbantartását nehezítő problémákat kerülhetünk el (Halassy, 2000). Ugyanakkor előfordulhatnak olyan esetek, amikor célszerűségi okokból mégis rákényszerülhetünk a redundáns adattárolásra – ez az adott helyzettől, igényektől, adatbáziskezelő-rendszer-től függ.
3. Egyszerűség: Ne vegyünk fel több egyedhalmazt, attribútumot és kapcsolatot annál, mint amennyi feltétlenül szükséges. Például a SZERZŐ esetén felesleges felvenni a szemszín attribútumot (bár kétségtelenül érdekes

lenne annak megvizsgálása, hogy a populációban leggyakrabban előforduló barna színhez képest van-e eltérés a szerzőknél és ha igen, az milyen mértékű).

4. Megfelelő kapcsolatok kiválasztása: Az egyedhalmazokat – kiváltképpen, ha sok van belőlük – sokféleképpen kapcsolhatjuk össze. Általában nem célszerű az összes kapcsolatot létrehozni, mert az a már korábban említett anomáliákhoz vezethet és csak megnehezíti (vagy akár el is lehetetlenítheti) az adatbázis kezelését. Az(oka)t a kapcsolato(ka)t célszerű feltüntetni, amelyek az adott feladat megoldása szempontjából szükségesek és ehhez tudnunk kell előre, hogy valójában mit is várunk el az adatbázistól. A SZERZŐ és a CIKK között például több–többhöz kapcsolatot célszerű kialakítani, hiszen egy szerző akár több cikket is írhat, illetve egy cikknek több társszerzője is lehet, de ugyanakkor a FOLYÓIRAT és a CIKK között csak egy–többhöz kapcsolat lehet, hiszen egy adott publikáció csak egyszer jelenhet meg a folyóiratban (normál esetben), másrészt egy folyóiratszám több cikket is tartalmazhat.
5. A megfelelő típusú elem megválasztása: A valóság leírása többféle módon lehetséges, így például előfordulhat, hogy a valóság egy szeletét leírhatjuk egyedhalmazzal, attribútummal vagy kapcsolattal is. Általában jellemző, hogy attribútumokat könnyebb a modellbe bevinni, mint egyedhalmazt, vagy kapcsolatot, ugyanakkor, ha minden attribútummal oldunk meg, az nem minden esetben célszerű, a kész adatbázis kezelését nehezítheti (Példaként lásd: 2. ábra).

2.4 Az egyed-kapcsolat modell

Peter Pin-Shan Chen 29 évesen korábban megjelentett „The Entity-Relationship Model – Toward a Unified View of Data” című cikkével (Chen, 1976) megalkotta az egyed-kapcsolat (Entity-Relationship, röviden ER) modell alapjait és ezzel jelentős, máig tartó hatást gyakorolt az adatmodellezésre. Az ER-modell a koncepcionális adatmodellezés legnépszerűbb eszközének tekinthető (Hartmann, 2003); népszerűsége az egyszerűség mellett az elméleti megalapozottságnak és annak is köszönhető, hogy a – mai adatbáziskezelő rendszerek által használt – relációs adatmodell előkészítésére kiválóan alkalmas. Az előbbiekből következően számos felsőoktatási intézményben tanítják az ER-modellt, ez pedig szintén hozzájárul népszerűségéhez, elterjedtségéhez (Carte et al., 2006).

A modellben az adatok szerkezetének ábrázolására ún. egyed-kapcsolat diagramot (Entity-Relationship Diagram, ERD) használunk (Ulman–Widom, 2009). Az ERD egy gráf, ahol a csúcspontok a modell három alapvető eleme (1) az egyedhalmaz(ok) (téglalappal jelöljük), (2) az attribútum(ok) (tulajdonságok) (ellipszissel adjuk meg) és (3) a kapcsolatok (rombuszsal jelöljük) (Silberschatz et al., 1996; Tímár et al., 1997; Szabó, 2013). Az egyedhalmazokat az attribútumaikkal és a kapcsolataikkal vonalak segítségével kötjük össze, ezek a gráf élei. (Ulman–Widom, 2009).

Az ER-modell tehát kifejező, egyszerű (nem szakemberek is viszonylag könnyen megérthetik), kevés fogalmat használ (így gyorsan megtanulható), szemléletes ábrákkal dolgozik, de az alkalmazott ábrázolási mód miatt a nagyon összetett adatbázisok leírása nehézkessé, az elkészült ábra pedig nehezen felfoghatóvá válhat (Dunn et al., 2005).

Az ER-modell elemkészlete:

1. Az egyed egy névvel ellátható (főnévvel megnevezhető), valamilyen típusú absztrakt objektum; a hasonló egyedek egyedhalmazt alkotnak (például: folyóirat, cikk, szerző). (Ulman–Widom, 2009).
2. Az attribútum az egyedhalmazhoz, vagy kapcsolathoz tartozik, annak tulajdonságát, tulajdonságait írja le (például: szerző esetében: név, tudományos fokozat, munkahely, neve). (Ulman–Widom, 2009) Egy egyedhalmaznak legalább egy attribútummal kell rendelkeznie, míg a kapcsolathoz nem kötelező attribútum(ka)t hozzárendelni. Az egyedtípushoz hasonlóan főnevet szokás használni az attribútum megnevezésekor (Tímár et al., 1997); fajtái:
 - egyszerű (ellipszis egyszeres vonallal),
 - összetett (több részattribútumra bontható, amelyeket az egyszeres vonallal rajzolt ellipszisben jelölünk),
 - többértékű (kettős vonallal rajzolt ellipszis) és
 - leszármaztatott (szaggatott vonallal rajzolt ellipszis; más tulajdonságok értékeiből számolható – de azt nem jelöljük a modellben, hogy hogyan) (Hampel–Heves, 2019).
3. A kapcsolat két – vagy akár több – egyedhalmazt kapcsol össze egymással és ígét használunk a leírására. Két egyedhalmaz összekapcsolása esetén bináris kapcsolatról beszélünk, ez a leggyakoribb. A modell nem zárja ki akár tetszőleges számú egyedhalmaz összekapcsolását sem. (Ulman–Widom, 2009).

Dey és munkatársai (1999) vizsgálatai szerint az egyedhalmazok és attribútumaik értelmezése többnyire nem okoz sok gondot a modell felhasználói számára, ugyanakkor a kapcsolatok értelmezésére ez már nem mondható el.

A kapcsolatok típusai lehetnek (Tímár et al., 1997; Dey et al. 1999; Ulman–Widom, 2009):

- rekurzív (visszaható, önmagára hivatkozó),
- 1:1 (egy–egy, egy–az–egyhez),
- 1:N (egy–több, egy–a–többhöz),
- N:M (több–több, több–a–többhöz),
- illetve nem bináris, több egyedhalmazt összekötő kapcsolatok esetén az egyedhalmazok számától függően beszélhetünk n-ed fokú (n: az egyedhalmazok száma) kapcsolatról.

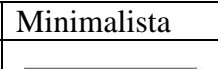



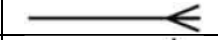
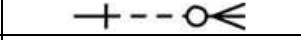


Minden felsorolt kapcsolat emellett lehet totális (teljes) és parciális (részleges, opcionális). Totális a kapcsolat akkor, ha minden egyed részt vesz a kapcsolatban és parciális kapcsolatról akkor beszélünk, ha van olyan egyed, amelynek nincs kapcsolata a másik egyedtípus valamelyik egyedéhez (Dullea et al., 2003).

A modellben fontos szerepet játszik a kulcs attribútum. Minden egyedhalmaznak kell legyen kulcsa (Ulman–Widom, 2009); amit az ERD-on az attribútum megnevezésének aláhúzásával jelölünk. Erről sokan megfedkeznek a modell elkészítésekor és ebben az adatbázis-kezelő programok is „hibásak”: engedik kulcs nélküli adatbázisok létrehozását (aminek persze oka lehet az egyszerű – egytáblás – adatbázisok létrehozásának és kezelésének egyszerűsítése is). A kulcs olyan attribútum, amely alkalmas arra, hogy egyértelműen azonosítson egy egyedet. Amennyiben több olyan attribútumunk is van, amely alkalmas a célra, akkor egyet kiválasztunk, ez lesz az elsődleges kulcs, a többi pedig alternatív kulcs. Előfordulhat, hogy egy attribútummal önmagában nem azonosítható egy egyed, hanem a minimálisan szükséges több attribútum összevonásával (egyesítésével) kaphatjuk meg az egyértelmű azonosításra alkalmas összetett kulcsot. Abban az esetben, ha egy egyedhalmaz kulcsában szereplő attribútumok között van olyan, amelyik más egyedhalmazok attribútuma, akkor gyenge egyedhalmazról beszélünk és ezt a tényt dupla vonallal jelöljük (Ulman–Widom, 2009).

A fent leírt ER-modellnek születtek bővített változatai is (Enhanced ER-model, EER-modell, kiterjesztett egyed-kapcsolat modell), amelyek további lehetőségekkel egészítik ki a modell valóságot leíró képességét, így például bevezetnek fő- és alosztályokat, az általánosítás és specializáció fogalmakat (Tímár et al., 1997; Ulman–Widom, 2009), ezeknek azonban a megvalósítani kívánt adatbázis szempontjából nincs lényeges szerepük.

Meg kell jegyezni azt is, hogy az ER-diagramok ábrázolásának többféle variánsa él egymás mellett (Hartmann, 2003), amelyek elsősorban a kapcsolatok ábrázolásában és azok milyenségében térnek el egymástól (Chua–Storey, 2011), lásd: *1. táblázat*. Watson véleménye szerint azonban nem az ábrázolás módja számít elsősorban, hanem a helyes modellezési képességet kell elsajátítani (Watson, 2006).

1. táblázat: Két adatmodellezési „dialektus”

Kapcsolat	Minimalista	DB Visual Architect
1:1		
1:1 gyenge egyedtípussal		
1:N		
1:N gyenge egyedtípussal		

Forrás: Saját szerkesztés Watson (2006) alapján

3. Az elkészült ER-modell

Az igények összegyűjtése, átgondolása során egy olyan ER-modell született, ami egyetlen egyedtípusban és a hozzá kapcsolt attribútumokban foglalja össze azokat az információkat, amelyeket az adatbázisból ki kellene nyerni. Ez a verzió folyóiratszámunként tartalmaz összesített adatokat, tehát egy-egy számon belül alapadatok (például a szerzők nevének tárolására) nem alkalmas.

Az egyedhalmaz neve: FOLYÓIRAT. Az attribútumokat, a típusukat és a hozzájuk tartozó metaadatokat az alábbi felsorolás tartalmazza a következő

formátumban: *attribútum megnevezése; attribútum típusa; metaadatok és egyéb megjegyzések*. A könnyebb áttekinthetőség érdekében az attribútumok elnevezésére több szóból álló kifejezéseket használtam. Ne felejtjük, hogy az attribútumok adatai egy-egy teljes folyóiratszámra vonatkoznak!

1. Azonosító; Összetett tulajdonság; Az egyes folyóiratszámok azonosítására szolgál.
2. Évfolyam; Egyszerű tulajdonság, az Azonosító résztulajdonsága; A folyóirat évfolyama sorszámmal.
3. Szám; Egyszerű tulajdonság, az Azonosító résztulajdonsága; A folyóirat sorszáma az évfolyamon belül.

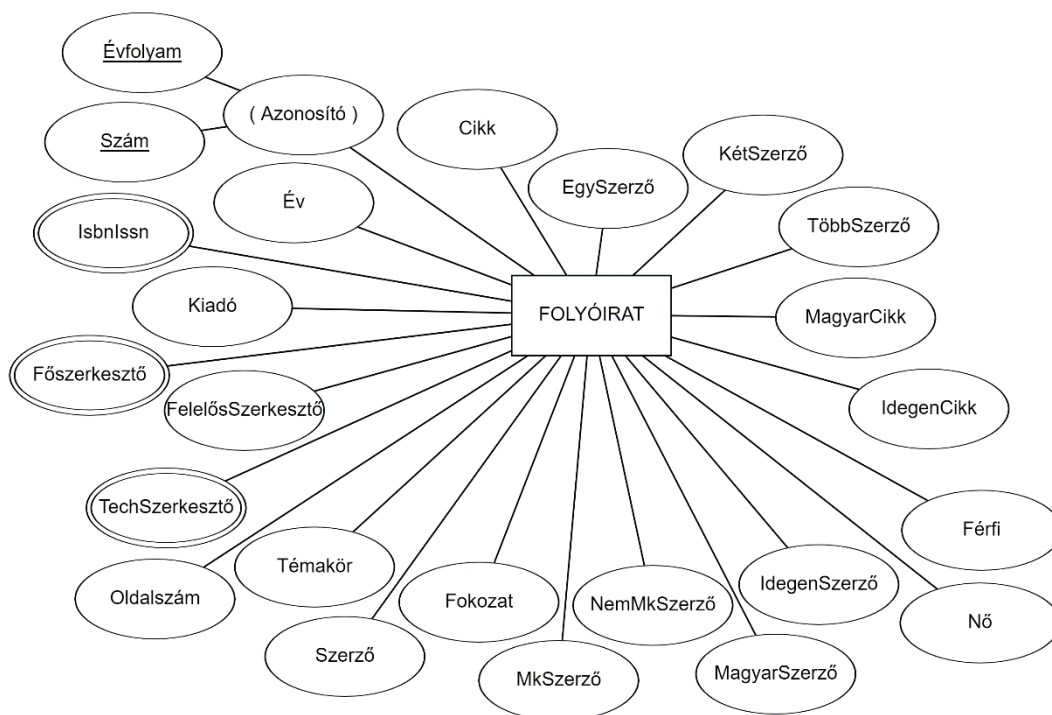
Az Évfolyam és a Szám együtt összetett kulcs (aláhúzással jelölve), ami egyértelműen azonosít egy adott folyóiratszámot.

4. Év; Egyszerű tulajdonság; A folyóirat megjelenésének éve.
5. IsbnIssn; Többértékű tulajdonság; A kiadvány ISBN, illetve ISSN számai. A nyomtatott és az online változat külön ISSN számmal rendelkezik, így egy folyóiratszám esetében egyszerre akár több értéket is felvehet.
6. Kiadó; Egyszerű tulajdonság; A felelős kiadó neve (és szükség esetén egyéb adatai). Ez az attribútum lehet akár összetett tulajdonság is (név, tudományos fokozat, beosztás stb. résztulajdonságokra felbontva), ha van erre igény ezen részadatok külön tárolására információkinyerés céljából.
7. Főszerkesztő; Többértékű tulajdonság; A főszerkesztő vagy szerkesztő pozíciót egyszerre többen is betölthetik.
8. FelelősSzerkesztő; Egyszerű tulajdonság; A felelős szerkesztő neve (és egyéb adatai, ha szükséges).
9. TechSzerkesztő; Többértékű tulajdonság; Egy számnak egyszerre több technikai szerkesztője lehet.
10. Oldalszám; Egyszerű tulajdonság; Az adott folyóiratszám teljes oldalszáma, nem csak a cikké.
11. Témakör; Egyszerű tulajdonság; A témakörök darabszáma.
12. Cikk; Egyszerű tulajdonság; A tudományos cikkek darabszáma.
13. EgySzerző; Egyszerű tulajdonság; Az egyszerű szerzős cikkek darabszáma.
14. KétSzerző; Egyszerű tulajdonság; A két társszerzővel rendelkező cikkek darabszáma.
15. TöbbSzerző; Egyszerű tulajdonság; A legalább három társszerzővel rendelkező cikkek darabszáma.
16. MagyarCikk; Egyszerű tulajdonság; A magyar nyelvű cikkek száma egy adott folyóiratszámban.
17. IdegenCikk; Egyszerű tulajdonság; Az idegen nyelvű cikkek száma.
18. Szerző; Egyszerű tulajdonság; A szerzők száma folyóiratszámanként.
19. MKSzerző; Egyszerű tulajdonság; Azoknak a szerzőknek a száma, akik az Méréskari Kar dolgozói.
20. MKSzteSzerző; Egyszerű tulajdonság; Azoknak a szerzőknek a száma, akik nem a Méréskari Kar dolgozói.

21. Fokozat; Egyszerű tulajdonság; A tudományos fokozattal rendelkező szerzők száma.
22. MagyarSzerző; Egyszerű tulajdonság; A neve alapján magyar szerzők száma (mivel állampolgársági adatok nem állnak rendelkezésre).
23. IdegenSzerző; Egyszerű tulajdonság; A nem magyar (hangzású) szerzők száma.
24. Férfi; Egyszerű tulajdonság; A nevük alapján férfi szerzők száma.
25. Nő; Egyszerű tulajdonság; A nevük alapján nő szerzők száma.

Látható, hogy ez az egy egyedhalmazból és huszonöt attribútumból álló ER-modell számos információ kinyerésére alkalmas, ugyanakkor sok előkészítést, előzetes manuális összesítést igényel, ráadásul fontos alapadatokat (például szerzők neve, cikkek címe) – amelyek lehetővé tennék további információk kinyerését – nem tartalmaz. A modell diagramja látható a 2. ábrán.

2. ábra: A folyóirat ER-diagramja (1. változat)



Forrás: Saját szerkesztés. Az ERD az erdplus.com adatmodellező weboldal eszközeivel készült.

Az ER-modell első változatát áttekintve készült a második, véglegesnek tekintett modell. A modell három egyedhalmazt (FOLYÓIRAT, CIKK, SZERZŐ) tartalmaz attribútumokkal, valamint három kapcsolatot (Ír, Publikál, Tartalmaz). A modell leírását, az attribútumokat, a típusukat és a hozzájuk tartozó metaadatokat az alábbi felsorolás tartalmazza *attribútum megnevezése; attribútum típusa; metaadatok és egyéb megjegyzések* formában.

A FOLYÓIRAT attribútumai:

1. Évfolyam; Egyszerű tulajdonság; A folyóirat évfolyama sorszámmal.
 2. Szám; Egyszerű tulajdonság; A folyóirat sorszáma az évfolyamon belül.
Az Évfolyam és a Szám együtt összetett kulcs (aláhúzással jelölve).
 3. Év; Egyszerű tulajdonság; A folyóirat megjelenésének éve.
 4. IsbnIssn; Többértékű tulajdonság; A kiadvány ISBN, illetve ISSN számai.
 5. FelelősKiadó; Egyszerű tulajdonság; A felelős kiadó neve (és szükség esetén egyéb adatai).
 6. Főszerkesztő; Többértékű tulajdonság; A főszerkesztő(k) vagy szerkesztő(k) neve (és egyéb adatai).
 7. FelelősSzerkesztő; Egyszerű tulajdonság; A felelős szerkesztő neve (és egyéb adatai).
 8. TechSzerkesztő; Többértékű tulajdonság; A technikai szerkesztő neve (és egyéb adatai).
 9. Oldalszám; Egyszerű tulajdonság; Egy adott folyóiratszám teljes oldalszáma.
 10. ÖsszesOldalDb; Leszármaztatott tulajdonság; Az folyóiratok oldalszáma összesen (az OldalSzám végösszege, a modell a számítás módját nem jelöli).
- A CIKK attribútumai:
1. CikkCím; Egyszerű tulajdonság és elsődleges kulcs; Mivel a cikk címe elsődleges kulcs, nem lehet két azonos című cikk az adatbázisban.
 2. Szekciónév; Egyszerű tulajdonság; A szekció neve (ha van), amelybe a cikk tartozik.
 3. Oldalkezd; Egyszerű tulajdonság; A cikk kezdő oldalszáma.
 4. OldalVégső; Egyszerű tulajdonság; A cikk utolsó oldalának oldalszáma.
 5. Nyelv; Egyszerű tulajdonság; A cikk nyelve.
 6. KulcsszóMagyar; Többértékű tulajdonság; A cikkhez tartozó magyar kulcsszavak.
 7. KulcsszóAngol; Többértékű tulajdonság; A cikkhez tartozó angol kulcsszavak.
 8. ForrásSzám; Egyszerű tulajdonság; A cikkhez tartozó források darabszáma.
 9. CikkOldalSzámDb; Leszármaztatott tulajdonság;
 10. CikkSzámDb; Leszármaztatott tulajdonság; A cikkek darabszáma.
 11. SzekcióDb; Leszármaztatott tulajdonság; A szekciók darabszáma.
 12. NSzerzőDb; Leszármaztatott tulajdonság; Az N-számú szerzők által írt cikkek száma.
 13. MagyarCikkDb; Leszármaztatott tulajdonság; A magyar cikkek darabszáma.
 14. IdegenCikkDb; Leszármaztatott tulajdonság; Az idegen nyelvű cikkek darabszáma.
 15. MkCikkDb; Leszármaztatott tulajdonság; A Mérnöki Karon dolgozó szerzők cikkeinek darabszáma.
 16. NemMkCikkDb; Leszármaztatott tulajdonság; A nem Mérnöki Karon dolgozó szerzők cikkeinek darabszáma.

A SZERZŐ attribútumai:

Alaposan végig kell gondolni, hogy az egyes szerzőket milyen tulajdonság(ok) alapján lehet azonosítani. A név önmagában nem elegendő, hiszen előfordulhat, hogy két vagy több különböző szerzőnek ugyanaz a neve, illetve, ha valaki nevet változtat két publikálás közben, akkor ő két névvel fog szerepelni és emiatt két különböző személynek fog számítani. A név, a fokozat (és munkahely) tulajdonságok összevonásával sem jutunk előrébb: ha egy többször publikáló szerző (új) tudományos fokozatot szerez, vagy munkahelyet vált, akkor megint csak két különböző személynek fog számítani. A probléma kezelhető egy önkényesen létrehozott egyedi azonosítóval, ami használható kulcstulajdonságként.

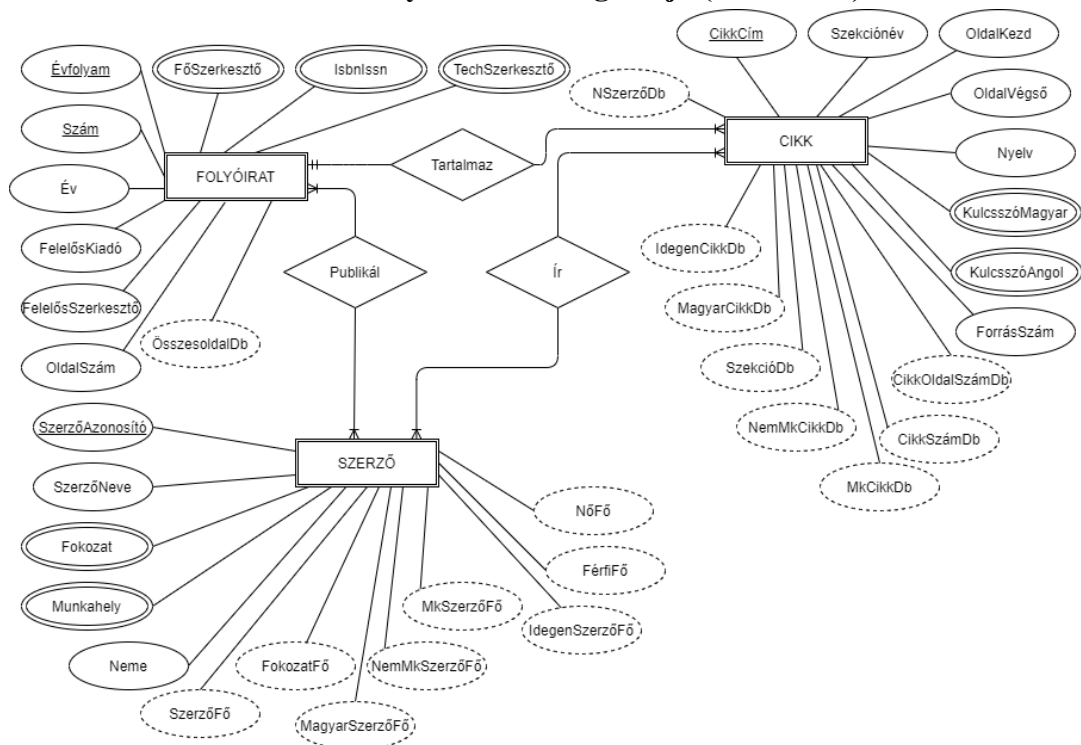
1. SzerzőAzonosító; Egyszerű tulajdonság; A szerzők egyértelmű azonosítását szolgáló kulcstulajdonság.
 2. SzerzőNeve; Egyszerű tulajdonság; Mivel elvileg előfordulhat, hogy két vagy akár több szerzőnek ugyanaz a neve, ezért – bár kényelmes megoldás lenne – önmagában nem lehet kulcstulajdonság.
 3. Fokozat; Többértékű tulajdonság; A szerző (tudományos) fokozata, amikor a cikket írta. Előfordulhat, hogy egy szerző több számban is publikál és közben új tudományos fokozatot szerez.
 4. Munkahely; Többértékű tulajdonság; A szerző munkahelyének adatai, amely lehet akár összetett tulajdonság is, hogy részeire bontható legyen (ha erre a későbbiekben lesz igény). Előfordulhat, hogy egy szerző több számban is publikál és közben munkahelyet vált.
 5. Neme; Egyszerű tulajdonság; A szerző neme a neve alapján.
 6. SzerzőFő; Leszármaztatott tulajdonság; A szerzők száma összesen.
 7. FokozatFő; Leszármaztatott tulajdonság; A tudományos fokozattal rendelkező szerzők száma.
 8. MagyarSzerzőFő; Leszármaztatott tulajdonság; A magyar nevű szerzők száma.
 9. IdegenSzerzőFő; Leszármaztatott tulajdonság; A nem magyar nevű szerzők száma.
 10. MkSzerzőFő; Leszármaztatott tulajdonság; A Mérnöki Kar szerzőinek száma.
 11. NemMkSzerzőFő; Leszármaztatott tulajdonság; Azoknak a szerzőknek a száma, akik nem a Mérnöki Karon dolgoznak.
 12. NőFő; Leszármaztatott tulajdonság; A nevük alapján nő szerzők száma.
 13. FérfiFő; Leszármaztatott tulajdonság; A nevük alapján férfi szerzők száma.
- A kapcsolat jellemzői:

- 1 számban több szerző publikálhat, 1 szerző több számban publikálhat, 1 cikk csak 1 számban jelenhet meg.
- Minden szerzőnek publikálnia kell valamelyik számban: egy szerzőnek kell legyen legalább 1 kapcsolata egy cikkhez és több cikkhez is lehet; egy szerzőnek kell legyen legalább 1 kapcsolata egy folyóiratszámhoz és több folyóiratszámhoz is lehet.

- Minden cikket publikálni kell valamelyik számban: egy cikknek 1 és csak 1 kapcsolata lehet a folyóirathoz; egy cikknek kell legyen legalább 1 kapcsolata egy szerzőhöz és több szerzőhöz is lehet.
- Minden folyóiratszámban kell legyen szerző: egy folyóiratszámnak kell legyen legalább 1 kapcsolata 1 szerzőhöz és több szerzőhöz is lehet.
- Minden folyóiratszámban kell legyen cikk: egy folyóiratszámnak kell legyen legalább 1 kapcsolata egy cikkhez.
- Az előbbiekből következően minden kapcsolat totális. A FOLYÓIRAT és a SZERZŐ között, valamint a SZERZŐ és a CIKK között több–a–többhöz (N:M) kapcsolatot, míg a CIKK és a FOLYÓIRAT között egy–a–többhöz (1:N) kapcsolatot kell létrehozni.

Az elkészült koncepciósi modell diagramját a 3. ábra tartalmazza. Az elsöre egyszerűnek (és gyorsan végrehajthatónak) tűnő feladat szisztematikus végiggondolása (mit vegyünk fel egyedhalmazként, mik legyenek az attribútumok, milyen jellegű kapcsolat szükséges az egyedhalmazok között) segített a helyes – a valóságot a megoldandó feladat szempontjából megfelelően leíró – adatbázis koncepciósi modelljének létrehozásában.

3. ábra: A folyóirat ER-diagramja (2. változat)



Forrás: saját szerkesztés. Az ERD a diagrams.net weboldal eszközeivel készült.

4. Záró gondolatok

A létrehozott ER-modell jelenlegi formájában alkalmas a megvalósításra, de ez nem zárja ki azt, hogy újabb ötletet ad a szerkesztőbizottságnak további információk

kinyerésére, ami miatt további kiegészítésre, módosításra lehet szükség még a logikai tervezést megelőzően.

A munka következő fázisa a koncepció modell átültetése logikai szintre, azaz a relációs modell létrehozása. Ebben a lépésben kell az egyedhalmazokat, az attribútumokat és a kapcsolatokat átkonvertálni táblákat, és azok szerkezetét, valamint a táblák közötti kapcsolatot leíró relációs sémává.

A befejező fázis a fizikai megvalósítás lesz, ami a relációs séma SQL-alapú megvalósítását jelenti a kiválasztott relációs adatbázis-kezelő rendszerben.

Amennyiben a későbbiekben lesz rá igény, az elkészült adatbázist akár szövegbányászatra is alkalmassá lehet fejleszteni és ebben az esetben az még eredményesebben járulhat hozzá a „Jelenkori társadalmi és gazdasági folyamatok” című folyóiratban rejlő tudás kiaknázásához.

Irodalomjegyzék

- Badia, A., Lemire D. (2011): A Call to Arms: Revisiting Database Design. *Sigmod Record*, 40 (3): 61–69. <https://doi.org/10.1145/2070736.2070750>
- Carte, T. A., Jasperson, J., Cornelius, M. E. (2006): Integrating ERD and UML Concepts When Teaching Data Modeling. *Journal of Information Systems Education*, 17 (1): 55–63.
- Chen P. P.-Sh. (1976): The Entity-Relationship Model – Toward a Unified View of Data. *ACM Transactions on Database Systems*, 1 (1): 9–36. <https://doi.org/10.1145/320434.320440>
- Chua, C. E. H., Storey, V. C. (2011): Issues and Guidelines in Modeling Decomposition of Minimum Participation in Entity-Relationship Diagrams. *Communications of the Association for Information Systems*, 29 (9): 159–184. <https://doi.org/10.17705/1CAIS.02909>
- Dey, D., Storey, V., Terence, B. (1999): Improving Database Design Through the Analysis of Relationships. *ACM Transactions on Database Systems*, 24 (4): 453–486. <https://doi.org/10.1145/331983.331984>
- Dullea, J., Song, Il-Y., Lamprou, I. (2003): An Analysis of Structural Validity in Entity-Relationship modeling. *Data & Knowledge Engineering*, 47 (2): 167–205. [https://doi.org/10.1016/S0169-023X\(03\)00049-1](https://doi.org/10.1016/S0169-023X(03)00049-1)
- Dunn, Ch. L., Gerard, G. J., Grabski, S.V. (2005): Critical Evaluation of Conceptual Data Models. *International Journal of Accounting Information Systems*, 6 (2): 83–106. <https://doi.org/10.1016/j.accinf.2004.03.002>
- Gulyás L. (2006): Előszó. *Jelenkori gazdasági és társadalmi folyamatok*, 1 (1): 1.
- Gulyás L., Gál J. (2007): Előszó. *Jelenkori gazdasági és társadalmi folyamatok*, 2 (1): 4.
- Gulyás L., Gál J. (2008): Előszó. *Jelenkori gazdasági és társadalmi folyamatok*, 3 (1): 4.
- Gulyás L., Gál J. (szerk.) (2009): Borítólapp. *Jelenkori gazdasági és társadalmi folyamatok*, 4 (1).
- Halassy B. (1994): *Az adatbázis-tervezés alapjai és titkai*. IDG Magyarország Lapkiadó Kft., Budapest.
- Halassy B. (1994): *Adatmodellezés. Elmélet és gyakorlat*. Budapest. <<https://mek.oszk.hu/11100/11144>> (2020. 08.10.)
- Hampel Gy., Heves Cs. (2019): *Informatika alapjai mérnököknek, alapszakos hallgatók számára*. Szegedi Tudományegyetem, Szeged.
- Hartmann, S. (2003): Reasoning about participation constraints and Chen's constraints. Database Technologies 2003, Proceedings of the 14th Australasian Database Conference, ADC 2003, Adelaide, South Australia, February 2003: 105–113.
- Kacsukné B. L., Kiss T. (2009): *Bevezetés az üzleti informatikába*. Akadémiai Kiadó, Budapest.
- Kis K. (szerk.) (2020): Belső címlap. *Jelenkori gazdasági és társadalmi folyamatok*, 15 (1–2): 1.
- Kovács L. (2004): *Adatbázisok tervezésének és kezelésének módszertana*. Computer Books, Budapest.

- Kroenke, D. M. (2006): Toward a Next Generation Data Modeling Facility: Neither the Entity-Relationship Model nor UML Meet the Need. *Journal of Information Systems Education*, 17 (1): 29–38.
- Lendvai E., Koppány K. (2018): Kvalitatív vizsgálat a Kígyóscsili termékeinek fogyasztói körében. *Jelenkori gazdasági és társadalmi folyamatok*, 13 (3–4): 71–77.
- Panyor Á., Szebelédi K. (2018): A termelői értékesítő szervezetek szerepe a vidék gazdaságában. *Jelenkori gazdasági és társadalmi folyamatok*, 13 (3–4): 33–42.
- Silberschatz, A., Korth, H. F., Sudarshan, S. (1996): Data Models. *ACM Computing Surveys*, 28 (1): 105–108. <https://doi.org/10.1145/234313.234360>
- Szabó B. (2013): *Adatbázis fejlesztés és üzemeltetés I.* Eszterházy Károly Főiskola, Eger.
- Tímár L., Vigh K., Tátrai J., Szigeti J., Vathy Á., Telekesi É., Vass I., Kocsis T., Priskinné R. Zs., Erdélyiné M. (19997): *Építsünk könnyen és lassan adatmodellt!* Veszprémi Egyetem & Műszertechnika-Veszprém Kft., Veszprém.
- Ullman, J., Widom, J. (2008): *Adatbázisrendszerek – Alapvetés.* Panem Kiadó, Budapest.
- Watson, R. T. (2006): The Essential Skills of Data Modeling. *Journal of Information Systems Education*, 17 (1): 39–41.
- Zsótér B., Bagi B. (2020): Gyümölcsfeldolgozó üzem létesítésének pénzügyi előkészületei. *Jelenkori gazdasági és társadalmi folyamatok*, 15(1–2): 131–134.