



KORRUPCIÓKUTATÓ  
KÖZPONT  
BUDAPEST

# Közbeszerzési eljárások dokumentumai összekapcsolhatóságának elemzése 2005-2012

*Módszertani elemzés*

*Public Procurement 4 You*

*7. Riport*

*Budapest, 2014. február*

A riportot a Korruptiókutató-központ Budapest (CRCB) kutatócsoportja készítette. A kutatás folytatását a központ pénzügyi erőforrásai, és a résztvevők önkéntes munkája tette lehetővé.

A kutatás folytatásához minden anyagi segítséget szívesen fogadunk, illetve várjuk önkéntesek jelentkezését ([info@crcb.eu](mailto:info@crcb.eu)).

## **Közbeszerzési eljárások dokumentumai összekapcsolhatóságának elemzése**

*Módszertani elemzés*

### **/ Analysis of the procurement documents interconnectivity 2005-2012**

*Technical analysis*

*Hetedik riport / 7th Report*

Együttműködő partnerek:

Erőforrástérkép projekt (MTA KRTK KTI): <http://regionaldata.org>

Átlátható Állam: <http://www.atlathatoallam.hu/>

3gteam: <http://www.3gteam.hu/>

A riportot írta:

Tóth Bence

Munkatársak:

Czibik Ágnes  
Fazekas Mihály  
Fóra Gyula  
Orbán Júlia  
Tóth Bence  
Tóth István János

közgazdász  
Ph.D. hallgató, University of Cambridge  
egyetemi hallgató  
közgazdász  
közgazdász  
Ph.D, tudományos főmunkatárs, MTA KRTK KTI

Külső szakértők:

Dr. Kelemen Zoltán  
Gyenese Jenő  
Nagy Zoltán  
Uhrin Tamás  
Dr. Volford Anita Adrienn

ügyvéd  
programtervező matematikus  
közgazdász  
programtervező matematikus  
ügyvédjelölt

Corruption Research Center Budapest  
e-mail: [info@crcb.eu](mailto:info@crcb.eu)  
internet: <http://www.crcb.eu/>

A kézirat lezárva: 2014. március 14.

## Tartalomjegyzék

|                                                                                  |    |
|----------------------------------------------------------------------------------|----|
| Tartalomjegyzék.....                                                             | 3  |
| Summary .....                                                                    | 4  |
| Összefoglaló .....                                                               | 6  |
| Rövidítések .....                                                                | 8  |
| Bevezetés .....                                                                  | 9  |
| 1. Adatok és eljárások struktúrája .....                                         | 12 |
| 2. Összekapcsolási eljárások rövid leírása .....                                 | 14 |
| 3. Empirikus elemzések a különböző hirdetmény viszonylatokban .....              | 16 |
| 3.1. Ajánlati felhívások „elsőkörös” jellemzői – 'Analitikus 2 jellemzése' ..... | 16 |
| 3.1.1. AF-ER teljes adatbázis elemzés .....                                      | 20 |
| 3.2. Korrigendumok és ajánlati felhívások összekapcsolása.....                   | 24 |
| 3.3. Korrigendumok és eredményhirdetmények .....                                 | 27 |
| 3.4. Szerződés módosítások és eredményhirdetések összekapcsolása .....           | 30 |
| 3.5. Szerződés teljesítések és eredményhirdetések összekapcsolása.....           | 33 |
| 4. Összefoglalás .....                                                           | 36 |
| 5. Melléklet.....                                                                | 37 |
| 5.1. Analitikus 1 (hosszabb).....                                                | 37 |
| 5.2. Analitikus 2 (rövidebb) .....                                               | 38 |
| 5.3. Felhasznált adatbázisok .....                                               | 38 |

## Summary

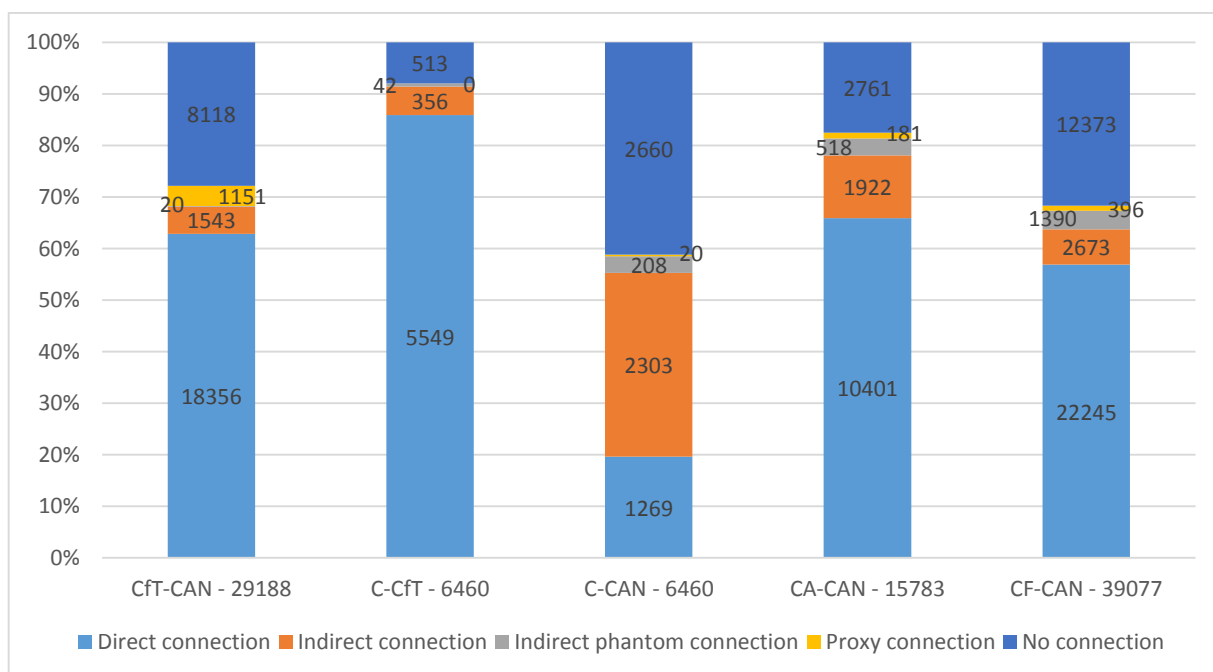
The Hungarian Public Procurement Authority (hereinafter referred to as: the Authority) does not have such a database on Hungarian public procurements, that contains the documentation of each public procurement tender in a uniform structure: namely the call for tenders, the contract award notices, the contract fulfilment, and contract amendment notices or corrigenda.

The goal of CRBC is to solve this deficiency: we would like to build a uniformly structured database based on the publicly available documents of the Authority that is suitable for statistical analysis.

This study is a by-product of this work and it analyses the connection problems of the – otherwise inconsistent – public procurement documents that are managed by the Hungarian state. The following document pairs will be discussed: call for tenders (CfT) – contract award notices (CAN); corrigendum (C) – call for tenders (CfT); corrigendum (C) – contract award notices (CAN); contract amendments (CA) – contract award notices (CAN); contract fulfilment (CF) – contract award notices (CAN).

Figure 1 shows the success of the different connection processes (direct, indirect, indirect phantom and proxy<sup>1</sup>) in case of the five document pairs. The variance of the processes' success rate are rather significant: in case of C-CAN pairs, the success rate is lower than 60%, while in case of C-CAN pairs it is above 90%. In case of the most important relations (CfT-CAN and CF-CAN), 30% of the connections are missing, meaning that in 30% of the CfTs, it cannot be known who won which procurement and what is the final contract after winning a given tender. The second conclusion is that the indirect and indirect phantom (not direct) connection processes perform poorly: in most of the cases only 10% of the missing connections can be established with these methods. Furthermore, the proxy method works only in case of CfT-CAN, C-CAN and CA-CAN pairs.<sup>2</sup>

**1. Figure: Success of different connection processes (2005-2012)**



<sup>1</sup> Indirect and indirect phantom processes are still based on registration numbers but in an alternative way, while the proxy process is based on commodity registration numbers, announcers and a text analysis procedure.

<sup>2</sup> An important note is that indirect, indirect phantom and proxy processes are exposed to type II errors: although their conditions are strict, false connections can be established.

Further implications of the statistical analysis:

- The success rate of any connection process between 2005 and 2007/8 are significantly lower compared to the years between 2009 and 2012 – in many cases it can be 10-30% or more. However, this trend is true for any kind of document pairs, hence the quality of documentation is improving.
- As neither the two kind of indirect, nor the proxy connection process have led to significantly more document connections, there is a straightforward implication that both the registry ID links and the content of the documents related to the same public procurement procedure are inconsistent.

## Összefoglaló

Magyarországon a Közbeszerzési Hatóság nem rendelkezik olyan adatbázissal a közbeszerzésekről, amelynek felhasználásával össze lehetne kapcsolni a közbeszerzési eljárások egyes lépéseinek adatait: a pályázatok meghirdetését, az eredményhirdetést, az esetleges szerződést módosítást. Ez a helyzet azt eredményezi, hogy nem lehet áttekinteni és elemezni a magyar közbeszerzési eljárások egészének folyamatát. Nem tudható, hogy az ajánlati felhívások jellemzői (például az eljárás típusa) milyen összefüggésben van azzal, hogy később, az eredményhirdetés és a szerződés megkötése után milyen gyakorisággal került sor szerződés módosításra, illetve árak emelésére. Az eljárási lépések összekapcsolhatóságának hiánya tehát nagymértékben rontja a közbeszerzések átláthatóságát, és csökkenti a közbeszerzések hatékonyságára irányuló elemzések lehetőségét.

A CRCB kutatói ezt a hiányt igyekeznek pótolni önkéntes munkában azzal, hogy a Közbeszerzési Hatóság által nyilvánosságra hozott közbeszerzési adatokat felhasználva egységes szerkezetű, a közbeszerzések matematikai-statisztikai elemzésére alkalmas közbeszerzési adatbázis (MaKAB) létrehozását tűzték ki célul.

Tudjuk, hogy ez a munka nem a mi feladatunk lenne, hanem a magyar államé.

Azt is tudjuk, hogy saját forrásainkra, a munkatársak és támogatóink önkéntes munkájára alapozott erőfeszítéseink olyan információk létrehozására irányulnak, amelyeket nemcsak mi fogunk hasznosítani, hanem mindenki más is, akit érdekel, hogy milyen hatékonysággal költi a magyar állam az adófizetők pénzét, és akiket foglalkoztat, hogy hogyan lehetne csökkenteni a korrupció, a versenykorlátozás jelenlétét a magyar közbeszerzésekben.

Jelen elemzés e munka mellékterméke és a magyar állam által gyűjtött és kezelt – egyébként inkonzisztens – adatok összekapcsolásának problémáit elemzi.

Az elemzés célja, hogy megvizsgálja a különböző közbeszerzési hirdetmények összekapcsolási eljárásainak sikerességét, továbbá javaslatokat tegyen új összekapcsolási eljárások bevezetésére. Az elemzés a következő hirdetmenypárosításokra terjed ki: ajánlati felhívások – eredményhirdetések; korrigendumok<sup>3</sup> – ajánlati felhívások; korrigendumok – eredményhirdetések; szerződésmódosítások – eredményhirdetések; szerződésteljesítések – eredményhirdetések.

Az 1. ábra azt mutatja, hogy az öt különböző hirdetmenypárosítás esetében a direkt, indirekt, indirekt fantom illetve proxy eljárás milyen sikerességgel kapcsolta össze az egyes hirdetményeket.<sup>4</sup> Látható, hogy ennek szórása viszonylag nagy: van, amikor összesen sem érik el az 70%-os összekapcsolási arányt, de pl. a korrigendum-ajánlati felhívás viszonylatban ez az arány 90% körüli.

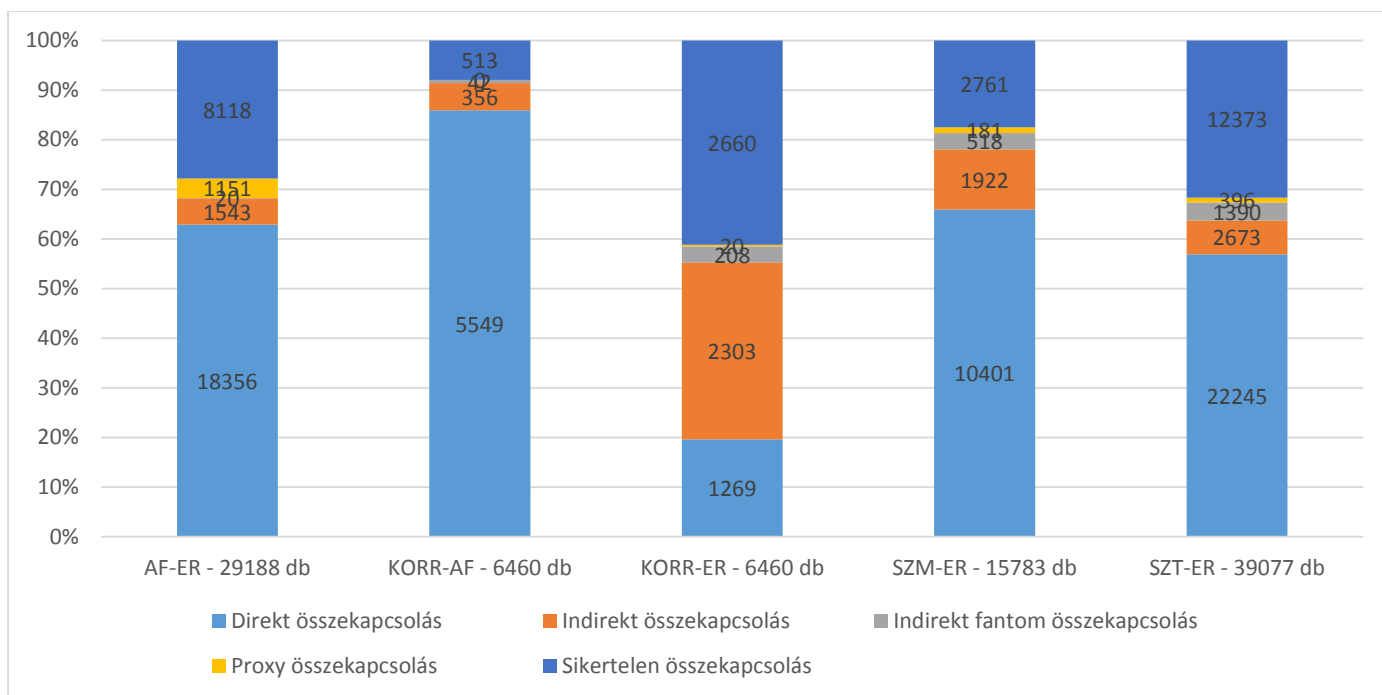
További látványos konklúzió, hogy az indirekt illetve indirekt fantom eljárások önmagukban csak mérsékelten mondhatók sikeresnek, a hiányzó kapcsolatoknak legtöbbször csak kb. 10%-át lehet ilyen módszerrel megtalálni. Ezen felül a proxy eljárás is csak a szerződésmódosítások, szerződés teljesítések eredményhirdetéssel való összekapcsolásában mondhatók sikeresnek.<sup>5</sup>

<sup>3</sup> „korrigendum”-nak hívjuk azokat a közbeszerzési hirdetményeket, amelyek korábbi közbeszerzési eljárások hirdetményeiben szereplő információkat korrigálják. Lásd [http://www.crcb.eu/wp-content/uploads/2014/01/kb\\_adatok\\_2010\\_3riport\\_120304.pdf](http://www.crcb.eu/wp-content/uploads/2014/01/kb_adatok_2010_3riport_120304.pdf). A korrigendumok megfigyelése és azonosítása nagyon fontos követelmény a magyar adatbázis tisztítása során, mivel egy eljárás adatai rajtuk keresztül duplikálva (először hibásan, aztán javítva) jelennek meg. Problémát okoz az is, hogy a vizsgált időszak alatt a Közbeszerzési Hatóság nem rendelkezett erre vonatkozó, a javítások mikéntjét szabályozó eljárásrenddel.

<sup>4</sup> A különböző eljárás típusok definícióit a második fejezetben tárgyaljuk.

<sup>5</sup> Emellett fontos megjegyezni, hogy a nem direkt eljárások esetleg másodfajú hibát tartalmazhatnak: olyan dokumentumokat kötnek össze, melyek valójában nem tartoznak össze.

**1. ábra: A különböző összekapcsolási eljárások sikeressége (2005-2012)**



**Magyarázat:**

*AF-ER: ajánlati felhívás – eredményhirdetés*

*KORR – AF: korrigendum – ajánlati felhívás*

*SZM – ER: szerződés módosítás – eredményhirdetés*

*SZT – ER: szerződés teljesítés - eredményhirdetés*

A tanulmányban szereplő leíró statisztikák további főkonklúziói:

- 2005-2007/8 közti időszak összekapcsolásai jellemzően sokkal kevésbé sikeresek, mint a 2008/9-2012-es évek összekapcsolásai, ez több 10%-nyi különbséget is jelenthet. Ez a különbség minden típusú összekapcsolási eljárás esetében megfigyelhető, tehát alapvetően egy javuló trend látható a közbeszerzések követhetőségében.
- Összességében sem az indirekt, az indirekt fantom, illetve proxy eljárás nem vezetett nagyságrendileg több hirdetmény összekapcsolásához: ez egyrészt azt jelenti, hogy az iktatószám alapú hivatkozások gyengén működnek, másrészt az egy eljáráshoz tartozó dokumentumok tartalma nem konzisztens (pl. rövid tartalmi leírások)

## Rövidítések

|      |                                       |
|------|---------------------------------------|
| AF   | ajánlati felhívás                     |
| CPV  | áruk/szolgáltatások egyedi kódjelzése |
| ER   | eredményhirdetés                      |
| EUHL | Európai Unió Hivatalos Lapja          |
| KORR | korrigendum                           |
| SZM  | szerződésmódosítás                    |
| SZT  | szerződésteljesítés                   |



## Bevezetés

Jelen tanulmány a Közbeszerzési Hatóság honlapján elérhető különböző közbeszerzési eljárásokhoz tartozó dokumentumok összekötési sikerességének eredményeit mutatja be a 2005-2012 évekre vonatkozóan. Egy-egy ilyen eljárás több stádiumból épül fel – például vannak ajánlati felhívási vagy eredményhirdetési dokumentumok –, azonban ezek összekapcsolása, tehát az egy közbeszerzési eljáráshoz tartozó összes dokumentum összekötése nem magától értetődő. A legideálisabb esetben az azonos eljáráshoz tartozó dokumentumok mindegyike tartalmaz egy olyan azonosítót, mellyel az adott közbeszerzéshez tartozó összes dokumentum megtalálható, s így ellenőrizhető.

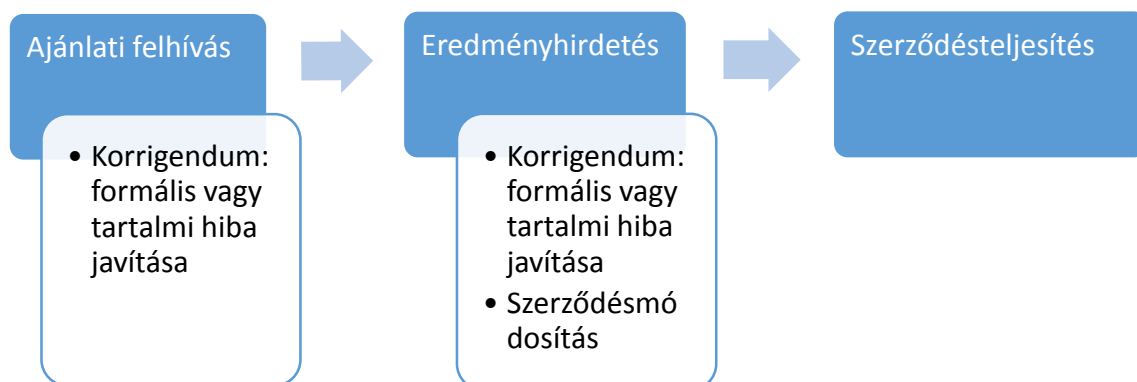
Ilyen ún. direkt linkeket azonban nem tartalmaz minden egyes dokumentum, ezért sokszor alternatív, indirekt vagy proxy módszerekkel kell vagy kellene összekötni az azonos közbeszerzési eljáráshoz tartozó dokumentumokat. Ilyen link lehet a beszerzett termék árukódja vagy az közbeszerzésben szereplő rövid (néhány szavas) tartalmi leírás egyezősége stb. Ebben a rövid tanulmányban az ilyen direkt, indirekt és proxy összekötési eljárások eredményeit foglaljuk össze.

Az 1. ábra egy teljes közbeszerzési eljárás menetét mutatja be: minden egyes stációhoz tartozik egy hirdetmény, mely a közbeszerzési eljárásokhoz tartozó dokumentumok gyűjtőneve.<sup>6</sup> Alap esetben az ajánlati felhívást (későbbiekben AF), az eredményhirdetés (későbbiekben ER), illetve a szerződésteljesítés (későbbiekben SZT) követi. „Köztes” hirdetmény lehet valamilyen korrigendum<sup>7</sup> (későbbiekben KORR), ami az AF, ER, SZT hirdetmények formai illetve tartalmi hibáit javítja ki, illetve az eredményhirdetés alapján kötött szerződéseket is módosíthatják, ez a szerződésmódosító hirdetmény (későbbiekben SZM). Optimális esetben egy adott közbeszerzési eljárásban ez a láncolat egy minden dokumentumon szereplő azonosítóval lekövethető lenne, azonban ez a valóságban sokszor nincs így. Ezért külön páronként kell összekötnünk az AF-ER, KORR-AF stb. párokat, mely alapján utólag rekonstruálni lehet egy adott eljáráshoz tartozó dokumentumcsoportot.

<sup>6</sup> Tehát mind az ajánlati felhívás, mind a szerződésteljesítés stb. egy hirdetményben van dokumentálva.

<sup>7</sup> „korrigendum”-nak hívjuk azokat a közbeszerzési hirdetményeket, amelyek korábbi közbeszerzési eljárások hirdetményeiben szereplő információkat korrigálják. Lásd [http://www.crcb.eu/wp-content/uploads/2014/01/kb\\_adatok\\_2010\\_3riport\\_120304.pdf](http://www.crcb.eu/wp-content/uploads/2014/01/kb_adatok_2010_3riport_120304.pdf). A korrigendumok megfigyelése és azonosítása nagyon fontos követelmény a magyar adatbázis tisztítása során, mivel egy eljárás adatai rajtuk keresztül duplikálva (először hibásan, aztán javítva) jelennek meg. Problémát okoz az is, hogy a vizsgált időszak alatt a Közbeszerzési Hatóság nem rendelkezett erre vonatkozó, a javítások mikéntjét szabályozó eljárásrenddel.

## 1. ábra: A közbeszerzési eljárások menete



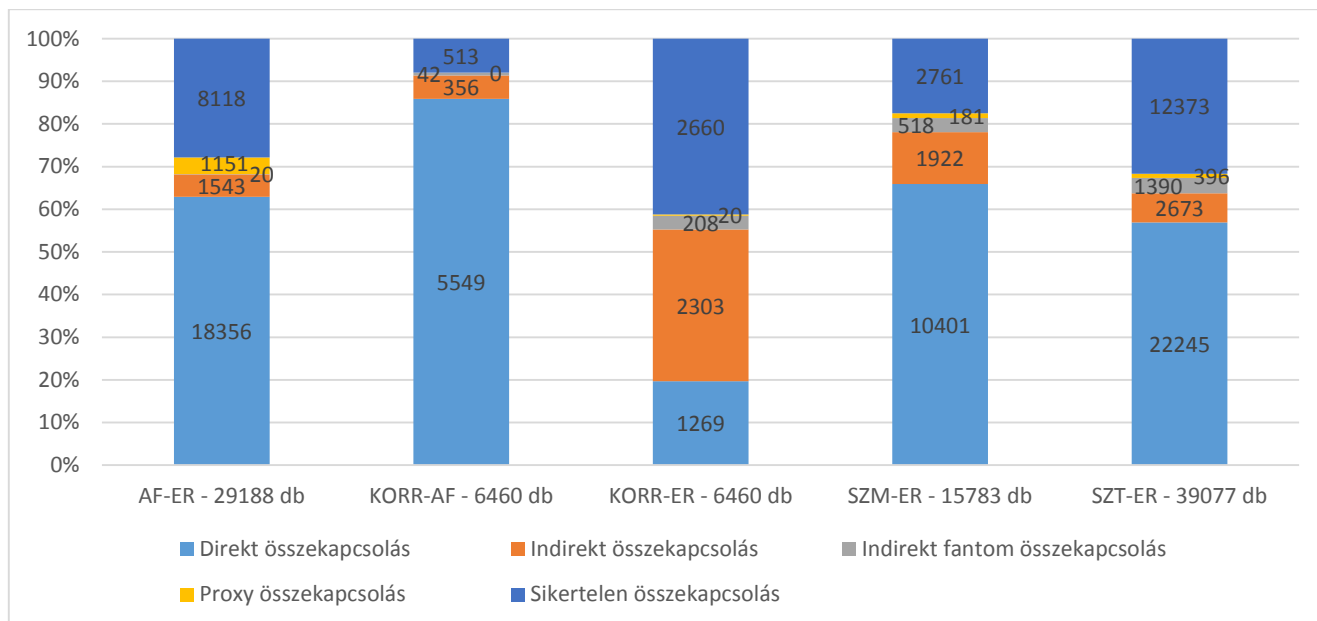
Összesen több százezer ilyen hirdetmény található a vizsgált 2005-2013-as időszakban, azonban ezek besorolása a különböző tartalmi kategóriákba (AF, ER stb.) nem feltétlen magától értetődő.

Az különböző összekötési eljárások eredményei alapján azt mondhatjuk, hogy az esetek nagyon nagy százalékában, ami a legérdekesebb relációkban kb. 10-45% között mozog, nem lehet megtalálni az összetartozó hirdetményeket közvetlen iktatószám alapú – direkt - összekötéssel. Az alternatív – indirekt és proxy – eljárásokkal ez az arány javítható, azonban az eljárások megbízhatósága nem egyértelmű. Ezekkel az összekötésekkel kiegészítve átlagosan kb. 27%-ra lehet csökkenteni a nem összepárosítható közbeszerzési dokumentumok arányát. További lényeges következtetése az elemzésnek, hogy 2005-2008-as években az összekapcsolhatóság szignifikánsan rosszabb, mint 2009-2012 között, minden egyes megvizsgált hirdetménypár relációjában.

Az 2. ábra azt mutatja, hogy az öt különböző hirdetménypárosítás esetében a direkt, indirekt, indirekt fantom illetve proxy eljárás milyen sikerességgel kapcsolta össze az egyes hirdetményeket.<sup>8</sup> Látható, hogy ennek szórása viszonylag nagy: van, amikor összesen sem éri el az 60%-os összekapcsolási arányt, de pl. a korigendum-ajánlati felhívás viszonylatban ez az arány 90% körüli.

<sup>8</sup> A különböző eljárás típusok definícióit a második fejezetben tárgyaljuk.

**2. ábra: A különböző összekapcsolási eljárások sikeressége (2005-2012)<sup>9</sup>**



**Magyarázat:**

*AF-ER: ajánlati felhívás – eredményhirdetés*

*KORR – AF: korrigendum – ajánlati felhívás*

*SZM – ER: szerződés módosítás – eredményhirdetés*

*SZT – ER: szerződés teljesítés - eredményhirdetés*

További konklúzió, hogy az indirekt, indirekt fantom és proxy eljárások önmagukban csak mérsékelttel mondhatók sikeresnek: a legtöbb esetben ezekkel a módszerekkel csak a hiányzó összekötések kb. 10%-át lehetett megtalálni. Ezzel kapcsolatban fontos megjegyeznünk, hogy ezek az alternatív módszerek potenciálisan másodfajú hibát tartalmazhatnak: lehetséges, hogy bár az összekötési eljárás logikusnak tűnő hivatkozások alapján köt össze különböző dokumentumokat, mégis olyan kapcsolatot állítunk fel, mely valójában helytelen.

Az alternatív eljárások mérsékelt sikereiből az is következik, hogy a közbeszerzési eljárások dokumentációja nem csak az iktatószámok hiánya miatt tekinthető diszfunkcionálisnak, hanem más tartalmi adatok is helytelenül vagy összepárosíthatatlanul szerepelnek ezekben a dokumentumokban. Így például ugyanannak az eljáráshoz különböző ajánlatkérővel, CPV kóddal vagy tartalmi leírással rendelkező dokumentumok tartoznak, ami az indirekt, illetve proxy változókon alapuló összekötési módszerek használatát is megnehezítik.

A tanulmány a következő struktúrát követi: a 2. fejezetben részletesen leírjuk az adatok struktúráját, illetve ábrákon szemléltetjük a lehetséges dokumentum összekapcsolási problémákat; a 3. fejezetben röviden bemutatjuk az általunk használt különböző összekötési eljárásokat; a 4. fejezetben pedig leíró statisztikákkal mutatjuk be, hogy az egyes összekapcsolási eljárások mennyire voltak sikeresek a különböző relációkban; az 5. fejezet pedig összefoglalja a főbb tanulságokat.

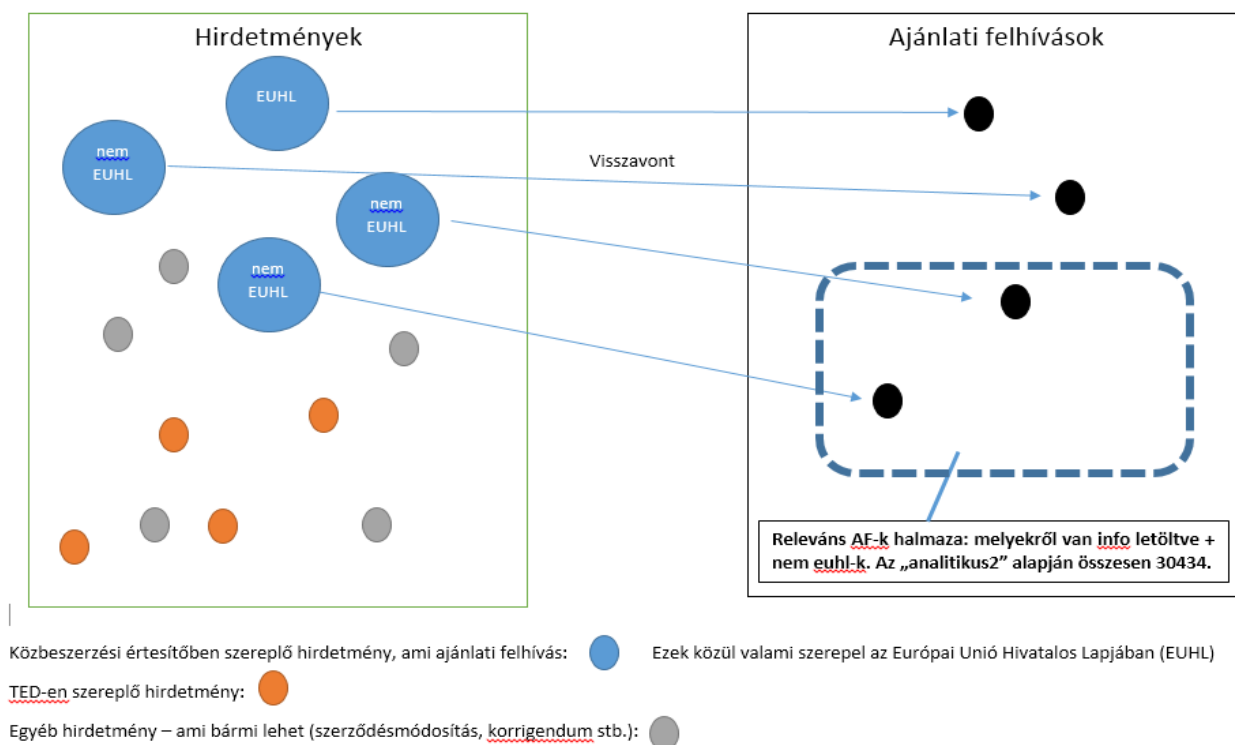
<sup>9</sup> Az egyes oszlopok kumulatíván állnak össze.

## 1. Adatok és eljárások struktúrája

Ennek a fejezetnek a célja egyrészt, hogy bemutassa a közbeszerzési eljárások dokumentációs menetét – hogy melyek azok a dokumentációs fázisok, melyeket egy láncba kell fűzni. Másrészt pedig bemutatja, hogy ezek a dokumentumok milyen formában álltak rendelkezésre, és az alapján milyen módszerekkel válogattuk szét az egyes hirdetményeket különböző tartalmi kategóriák szerint (ajánlati felhívás, eredményhirdetés stb.), továbbá vizuálisan is bemutatjuk, hogy a hiányzó kapcsolatok milyen relációkban merülhetnek fel.

Az alábbi 3. ábra azt mutatja be, hogy különböző forrásokból származó hirdetményekből hogyan szelektálódik ki az a részcsoporthoz, melyen az elemzést végezzük. Alapvetően két módszer alapján lettek besorolva a különböző hirdetmények ajánlati felhívás, szerződés stb. kategóriákba, de jelen tanulmányhoz egy olyan eljárást használtunk, mely a hirdetmények szövegében megkeresi a „felhívás”, „eredmény” és „szerződés” szavakat és ez alapján történt a tartalmi kategóriákba sorolás. Ezek után tovább szűkítettük a kört azokra a hirdetményekre, melyek egyrészt nem szerepelnek az európai unió hivatalos lapjában (EUHL)<sup>10</sup>, másrészt nem vonták őket vissza – mivel a visszavont hirdetményeket értelemszerűen nem is tudnánk összekapcsolni későbbi hirdetményekhez. Így a 3. ábra szereplő szaggatott kék vonallal bekarikázott fekete pontok jelölik a hirdetmények releváns „ajánlati felhívás” részhalmozát.

3. ábra: Hirdetmények besorolása (AF-ER példa)



A 4. ábrán egy hirdetmény példa látható, ahol a „hirdetmény típusa” sorban szerepelnek pl. az „eredmény” vagy „szerződésteljesítés” szavak, ami alapján az adott hirdetmény az ER vagy SZT kategóriákba kerül. Két ilyen dokumentum közti direkt kapcsolatot (az összekapcsolási kritériumokat már az

<sup>10</sup> Az EUHL-ben szereplő hirdetmények megvannak a magyar közbeszerzési értesítőben is, így ezzel a szűkítéssel bizonyos eljárások kétszeres figyelembevételét kerüljük el.

Összekapcsolási eljárások rövid leírása című fejezetben részleteztük) pedig a direkt „iktatószám” hivatkozás alapján képeztünk.

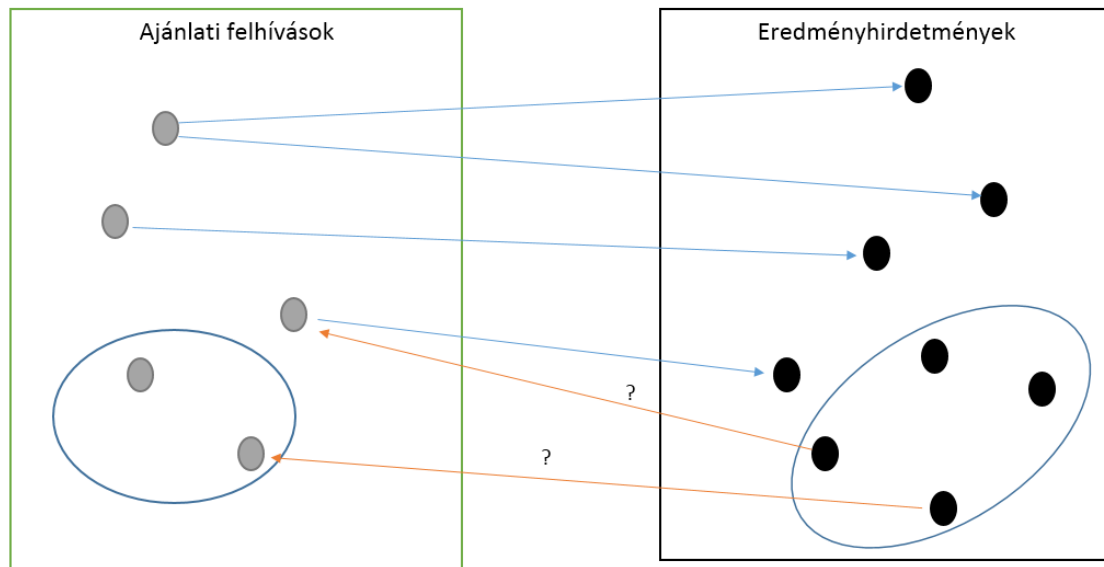
#### 4. ábra: Hirdetmény példa

| Hirdetmény                                      |  |
|-------------------------------------------------|--|
| Közbeszerzési Értesítő száma:                   |  |
| Ajánlatkérő:                                    |  |
| Nyertes ajánlattevő:                            |  |
| Teljesítés helye:                               |  |
| Beszerezés tárgya:                              |  |
| Eljárás fajtája:                                |  |
| Hirdetmény típusa:                              |  |
| CPV kód:                                        |  |
| Közzététel dátuma:                              |  |
| Iktatószám:                                     |  |
| Ajánlattételi/részvételi jelentkezési határidő: |  |
| Ajánlatkérő típusa:                             |  |
| Ajánlatkérő fő tevékenységi köre:               |  |
| Rövid tartalom:                                 |  |
| <a href="#">Tartalom letöltése PDF-ben</a>      |  |

A hirdetmények összekapcsolását a következő 5. ábra mutatja. Jelenleg az adatbázisban a kék nyíllal jelölt kapcsolatok láthatók, és a keresés fókuszában a kapcsolattal nem rendelkező ajánlati felhívások (bekarikázott szürke pontok) állnak. Az előző alfejezetben ezen bekarikázott szürke pontok továbbra is kapcsolat nélkül hagyása számítana „elsőfajú hibának”. A példaként szereplő AF-ER kapcsolatok esetében összesen 21070 AF-hez lehetett hozzákapcsolni legalább 1 darab ER-t, így 8118 olyan AF maradt (a bekarikázott szürke pontok) aminek nincs párja – a 2005-2012-es időszakban.

Azonban a „másik” irányból is meg lehet közelíteni a problémát: olyan eredményhirdetések is maradhatnak az összekötési eljárás végére, melyek valójában tartoznak egy ajánlati felhíváshoz, de nem kötöttük őket össze. **Mivel a keresési algoritmus véget ér, amint egy ajánlati felhívás legalább egy eredményhirdetést kap, nem látszódik, hogy valójában az összes hozzátartozót megtaláljuk-e vagy sem.** Ezt szemléltetik az 5. ábra piros nyilai.

## 5. ábra: Összekapcsolások problémája (AF-ER példa)



## 2. Összekapcsolási eljárások rövid leírása

Ebben a fejezetben röviden bemutatjuk, hogy pontosan milyen eljárások alapján kapcsoltuk össze az egyes hirdetményeket. Alapvetően **négyféle összekapcsolási módszert lehet megkülönböztetni: a direkt összekapcsolást, indirekt összekapcsolást, indirekt fantom összekapcsolást, illetve a proxy összekapcsolást.**

**Direkt összekapcsolás:** ebben az esetben a hirdetményekben szereplő közvetlen hivatkozások alapján történt az összekapcsolás. Minden egyes hirdetményben található egy iktatószám, amely az adott dokumentumot azonosítja. A direkt összekötés azt jelenti, hogy az egyik hirdetményben – pl. ER-ben – szerepelt egy másik hirdetmény – pl. AF – iktatószáma, ami alapján két hirdetmény azonos közbeszerzési eljáráshoz tartozik. Az alábbi példa egy ER-ből vett anonimizált idézet, amelyben XXXX/év egy direkt iktatószám hivatkozás egy AF-re, illetve szerepel az is, hogy a Közbeszerzési Értesítő mely számában és pontosan mikor jelent meg:

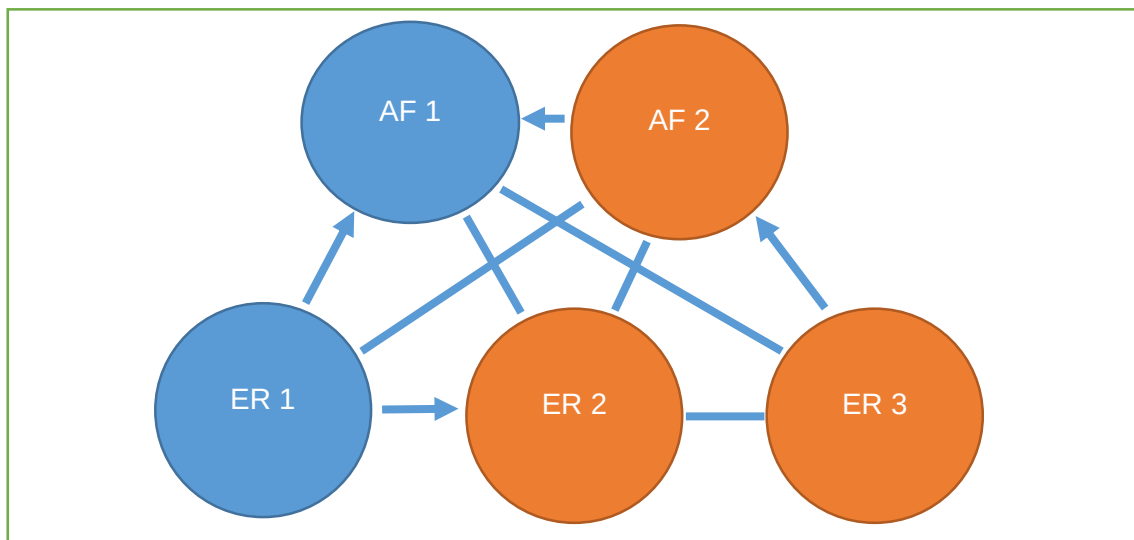
### VI.5) A SZERZŐDÉS MEGHIRDETÉSE A KÖZBESZERZÉSI ÉRTESÍTŐBEN

*XXX ajánlati felhívása nyílt eljárás megindítására (XXXX/év), Közbeszerzési Értesítő Y. évfolyam Y. szám (év. hónap nap.)*

**Indirekt összekapcsolás:** az indirekt összekapcsolás szintén az iktatószámok alapján történik, azonban itt nem tudunk közvetlen kapcsolatot felállítani két hirdetmény, például AF és ER között, hanem csak indirekt hivatkozások alapján történtek az összekötések. Az összekapcsolási módszer – direkt összekapcsolástól való különbözőségének – megértését az alábbi **Hiba! A hivatkozási forrás nem található..** ábra segíti. Az ábrán az irányított nyilak közvetlen iktatószám alapú hivatkozást jeleznek – tehát pl. az ER 1-ben szerepel az AF 1 iktatószáma (ilyen egy direkt összekapcsolás). A sima (nem irányított) egyenessel jelölt kapcsolatokat pedig az indirekt hivatkozások alapján kötöttük össze egy azonos közbeszerzési eljáráshoz. Így az ER 2 azért szerepel az AF 1-gyel kezdődő közbeszerzési eljárásban, mert az ER 1-ben explicit

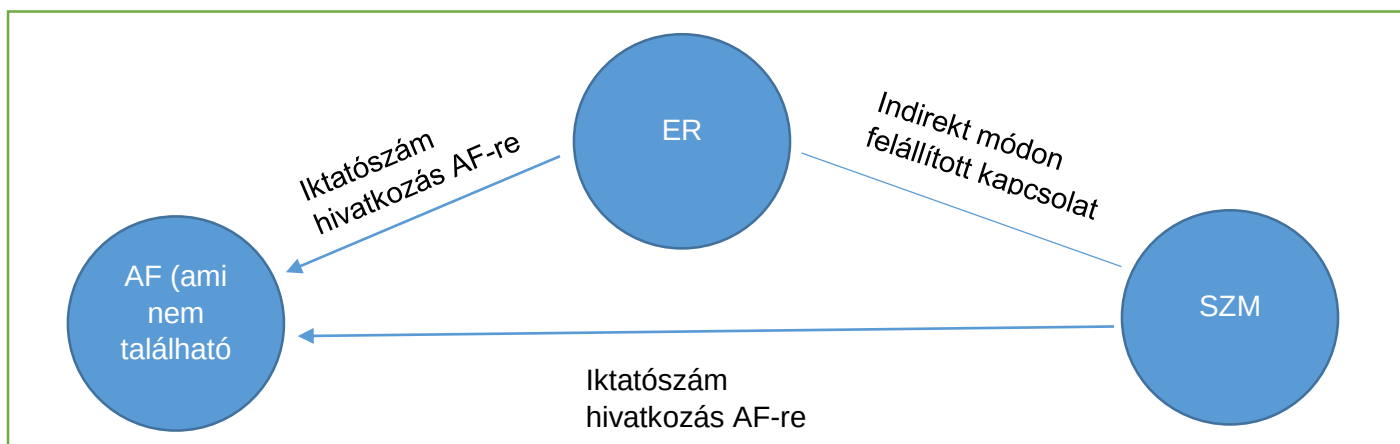
iktatószám alapú közvetlen hivatkozás van rá. Tehát az AF 1-hez közvetlenül tartozik ER 1, és ER 1-en keresztül indirekt ER 2. Egy másik nem triviális összekötés, mely ezzel a módszerrel adódhat, hogy amennyiben AF 2 iktatószámmal közvetlenül hivatkozik AF 1-re, akkor az AF 2-re iktatószámmal közvetlenül hivatkozó ER 3 indirekt hozzá köthető AF 1-hez. Ilyen módon a 6. ábra mind egyetlen közbeszerzési eljáráshoz köthető dokumentumokat tartalmaz.

6. ábra: Az indirekt összekapcsolási módszer



**Indirekt fantom összekapcsolás:** a hirdetményekben szereplő iktatószám hivatkozások alapján történt az összekapcsolás, de azokon a hirdetményeken keresztül is történt összekapcsolás, amelyek nem szerepelnek az adatbázisunkban, csak az iktatószámukat tudjuk. Ezt szemlélteti a 7. ábra.

7. ábra: A direkt fantom összekapcsolási módszer



**Proxy összekapcsolás:** ez a módszer elsősorban három kritérium teljesítésén alapul. **Egyrészt az AF-nek korábbiak kell lennie a hozzákapcsolt hirdetményénél**, de dátum szerint nem lehet köztük 60 napnál nagyobb távolság. **Másrészt mind a két hirdetményben szerepelnie kell az ajánlatkérő nevének**, melyek Damerau-Levenshtein<sup>11</sup> távolságának 5-nél kisebbnek vagy egyenlőnek kell lennie.

<sup>11</sup> A Damerau-Levenshtein távolság egy az információs- és számítástechnikában használt távolságfogalom: két szöveges jel között (végeses hosszú szimbólum sorozat) lévő távolság, melyet a két kód ekvivalenssé tételéhez szükséges legkevesebb művelet számával definiálnak. A műveletek a következők lehetnek: beillesztés, törlés, egy karakter kicserélése, vagy két szomszédos megcserélése. Ezeknek az egyszerű szerkesztési eljárások eredményeként definiálható a Damerau-Levenshtein távolság, ami megmondja, hogy hány műveletnyi távolság van két szöveg között.



**Harmadrészt, ha a hirdetmények rendelkeznek CPV kóddal, akkor legalább 2 CPV kódnek meg kell egyeznie a két hirdetményben.** Amennyiben nincs két CPV kód, akkor annyi CPV kódnek kell megegyeznie, amennyi a kevesebb CPV-t tartalmazó hirdetményben található.

A második szűrő, melyet ebben a tanulmányban csak részben használunk, hogy az előző 3 kritérium alapján összekapcsolt dokumentumokban található rövid tartalomleírások mennyire egyeznek meg. Ez alapján még egy pontosító szűrést be lehet építeni, mely ezeknek a leírásoknak a Damerau-Levenshtein távolságán alapul. Mi erre a következő változót definiáltuk:

$$rmatch = \frac{(l-r)*100}{l},$$

ahol  $l$  a hosszabb rövid tartalomleírás hossza<sup>12</sup>,  $r$  pedig a két rövid tartalomleírás Damerau-Levenshtein távolsága. **Az általunk használt szélsőérték 95, tehát amely kapcsolatokat fogadjuk el indirekt összekapcsolásnak, melyeknek az  $rmatch$  értéke 95-nél nagyobb.** Ez a választás tulajdonképpen önkényes, azonban az előzetes vizsgálódásaink azt támasztják alá, hogy a másodfajú hibák száma nagyban megugrana, amennyiben ezt a relatíve szigorú 95-ös korlátot lazítanánk.

Az összekötési eljárások minőségét két típusú hiba előfordulásával mérhetjük. Egyrészt elkövethetünk olyan elsőfajú hibát, hogy nem kötünk össze olyan dokumentumokat, melyek valójában összetartoznak.<sup>13</sup> Ezeknek a hibák a rendelkezésre álló dokumentumok minőségének hibájából következnek, mivel számos dokumentum nem tartalmaz olyan hivatkozásokat, ami alapján meg lehetne határozni a kapcsolatokat. Másrészt lehetséges, hogy a használt összekötési eljárások nem elég pontosak és ezért olyan dokumentumokat is összekötünk, melyek nem tartoznak össze. Ezt nevezhetjük másodfajú hibának. Az adatbázis építés célja, hogy minél több helyes kapcsolatot építsen fel úgy, hogy mind a két típusú hiba számát minimalizálja.

A következő fejezetben a különböző relációkban történő dokumentum-összekapcsolások eredményeit tekintjük át.

### 3. Empirikus elemzések a különböző hirdetmény viszonylatokban

#### 3.1. Ajánlati felhívások „elsőkörös” jellemzői – ’Analitikus 2 jellemzése’

Az „analitikus 2” típusú adatbázisban arról tartalmaz információkat, hogy egy adott ajánlati felhíváshoz hozzá lehetett-e kapcsolni egy eredményhirdetményt direkt, indirekt, indirekt fantom vagy proxy eljárással. Az összekapcsolások módszertanából fakadóan azonban egy-egy ajánlati felhívás többször is szerepelhet az adatbázisban, mert bizonyos összekötési eljárások az európai unió hivatalos lapjában (későbbiekben EUHL) megjelent hirdetmények alapján kapcsol össze felhívásokat és eredményhirdetményeket.

Emiatt az EUHL-ben is szereplő sorokat ki kell szűrni, hogy ezzel elkerüljük a duplikációt. Az 1. táblázat alapján látszódik, hogy összesen **46634 sort – ajánlati felhívást – tartalmaz** az adatbázis, azonban ebből **16509** pusztán technikai okokból, **az EUHL-es összekapcsolásokból kifolyólag szerepel.** Továbbá vannak olyan ajánlati felhívások, melyeket később visszavontak, így azokhoz nem is tartozhat semmilyen további hirdetmény. Ilyen ajánlati felhívásból összesen 952 darab található. Az EUHL-es és visszavont ajánlati felhívások kiszűrése után **a releváns esetek száma 29188**, így az 1. táblázat 3-5 oszlopa már csak ezeknek az eseteknek az eloszlását mutatja. A későbbi fejezetekben az adatok releváns részhalmazra szűkítését már nem részletezzük ilyen szinten.

<sup>12</sup> Megjegyzés: a rövid tartalomleírásokat módosítjuk, pl. kivesszük az „ajánlattételi felhívás” stb. szövegeket; illetve ha bármelyik rövid tartalomleírás 0, akkor a teljes „ $rmatch$ ” mutató is 0.

<sup>13</sup> Tulajdonképpen ennek az adatbázis építésnek a célja, hogy ezeket a 0%-ban összekötött dokumentumokat (tehát 100%-ban elsőfajú hibákat) megpróbálja összekapcsolni.



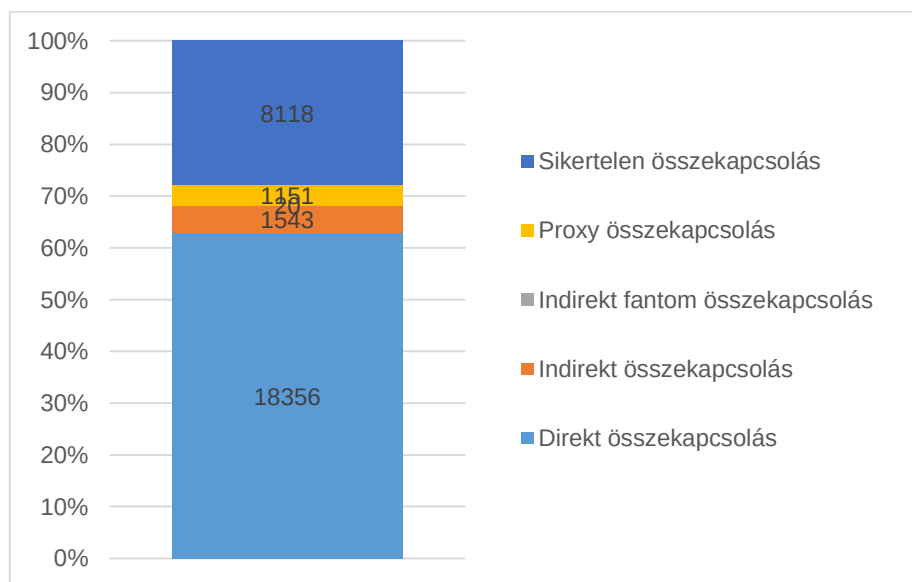
Az 1. táblázat alapján két fontos állítás tehető: egyrészt **az indirekt és az indirekt fantom eljárás sikeressége nem különbözik nagyságrendileg az egyszerű direkt eljárásétól**, másrészt **a proxy eljárással sok direkt összekötést – kb. 20%-ot – egyáltalán nem is tudtunk összekapcsolni**. Ahogy már a 0. fejezetben kifejtettük, ez az eljárás két lépcsőben köti össze a hirdetményeket: míg az első dátum, ajánlatkiíró és CPV kódok alapján közel 71%-os az összekötési arány, a rövid tartalmi leírások szövegelemzése (Damerau-Levenshtein távolsága) alapján az kapcsolatok 30%-a kiesik.

**1. táblázat: Az AF-ER összekapcsolások leíró statisztikái (2005-2012)**

|          | EUHL-e? |      | Visszavont-e? |     | Direkt eljárással összekapcsolt |      | Indirekt eljárással összekapcsolt |      | Indirekt fantom eljárással összekapcsolt |      | Proxy eljárással összekapcsolt |      |
|----------|---------|------|---------------|-----|---------------------------------|------|-----------------------------------|------|------------------------------------------|------|--------------------------------|------|
|          | db      | %    | db            | %   | db                              | %    | db                                | %    | db                                       | %    | db                             | %    |
| Nem      | 30125   | 65%  | 45682         | 93% | 10832                           | 37%  | 9289                              | 32%  | 9269                                     | 32%  | 17230                          | 59%  |
| Igen     | 16509   | 35%  | 952           | 2%  | 18356                           | 63%  | 19899                             | 68%  | 19919                                    | 68%  | 11958                          | 41%  |
| Összesen | 46634   | 100% | 46634         | 95% | 29188                           | 100% | 29188                             | 100% | 29188                                    | 100% | 29188                          | 100% |

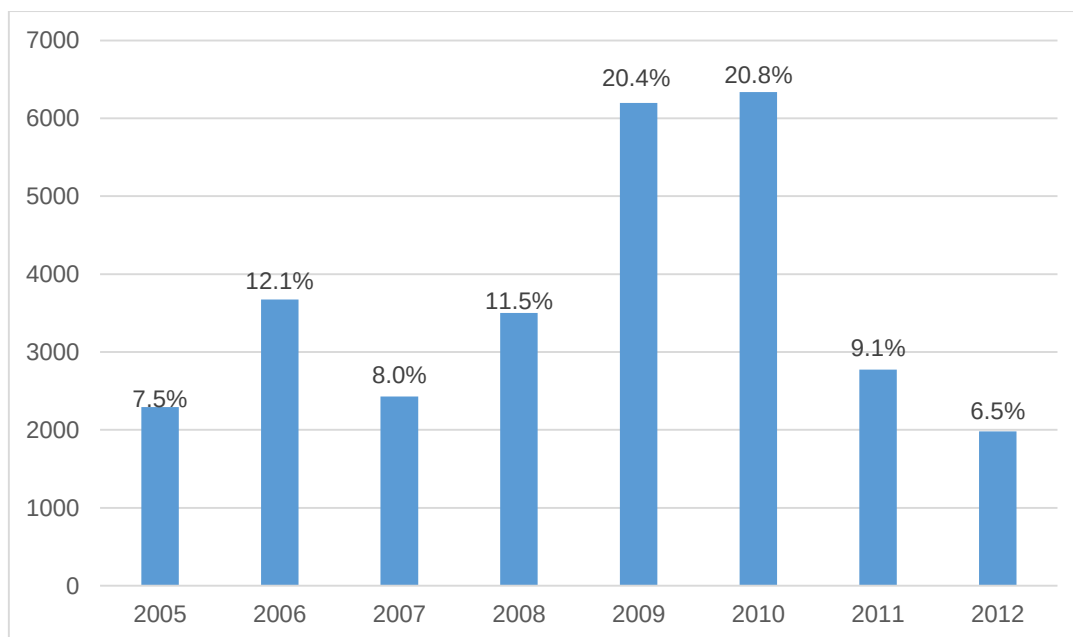
Az ajánlati felhívások között **a sikeres direkt AF-ER összekapcsolások száma 18356**, amit a sikeres **indirekt összekapcsolások száma 1534**-gyel egészít ki, továbbá még **20** kapcsolat található indirekt fantom összekötési módszerrel. **A sikeres proxy összekapcsolások száma 1151**, így **összességében 9364** olyan ajánlati felhívás van, melyet semmilyen eredményhirdetéssel nem lehetett összekapcsolni, ami az ajánlati felhívások kb. 25%-a. Ezt a 8. ábra mutatja.

**8. ábra: Az AF-ER összekapcsolások %-os eloszlása (2005-2012)**



A 9. ábra azt mutatja, hogy hány darab közbeszerzési ajánlati felhívás volt az egyes években. Látható, hogy ennek eloszlása elég változó, a 2009-10-es évekre esik a kiírások több mint 40%-a.

9. ábra: Az ajánlati felhívások évenkénti eloszlása, db és %



A 2. táblázat azt mutatja, hogy a különböző összekapcsolási eljárások mennyire sikeresek a különböző években. Látható, hogy az eljárások relatív sikeressége nagyon erősen változik: **míg a 2005-2008-as években a direkt eljárás sikeressége 23-43% között mozgott, addig a 2009-2011-es években 76-88% között alakult, illetve 2012-re is csak 60%-ra esett vissza.** Ez a trend az indirekt, indirekt fantom eljárásra is ugyanúgy igaz, illetve hasonló mintázat látható a proxy eljárások esetében is.

**2. táblázat: A különböző eljárási módszerek sikerességének évenkénti eloszlása**

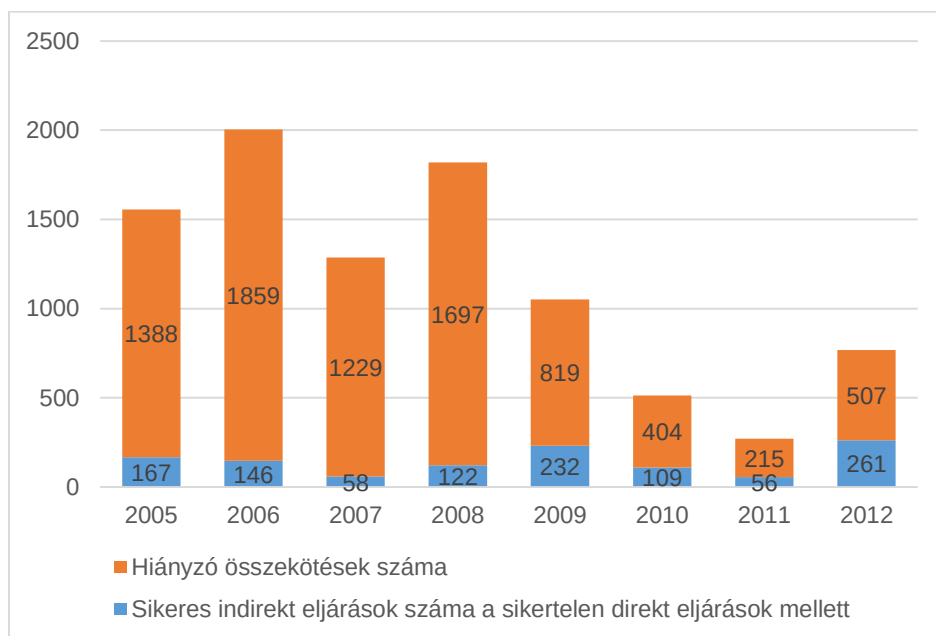
|                                              |          | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|----------------------------------------------|----------|------|------|------|------|------|------|------|------|
| Direkt eljárással összekapcsolható?          | nem      | 1757 | 2281 | 1398 | 1995 | 1465 | 770  | 363  | 803  |
|                                              | igen     | 536  | 1392 | 1031 | 1508 | 4732 | 5568 | 2412 | 1177 |
|                                              | nem (%)  | 77   | 62   | 58   | 57   | 24   | 12   | 13   | 41   |
|                                              | igen (%) | 23   | 38   | 42   | 43   | 76   | 88   | 87   | 59   |
| Indirekt eljárással összekapcsolható?        | nem      | 1560 | 2014 | 1291 | 1820 | 1052 | 513  | 271  | 768  |
|                                              | igen     | 733  | 1659 | 1138 | 1683 | 5145 | 5825 | 2504 | 1212 |
|                                              | nem (%)  | 68   | 55   | 53   | 52   | 17   | 8    | 10   | 39   |
|                                              | igen (%) | 32   | 45   | 47   | 48   | 83   | 92   | 90   | 61   |
| Indirekt fantom eljárással összekapcsolható? | nem      | 1555 | 2005 | 1287 | 1819 | 1051 | 513  | 271  | 768  |
|                                              | igen     | 738  | 1668 | 1142 | 1684 | 5146 | 5825 | 2504 | 1212 |
|                                              | nem (%)  | 68   | 55   | 53   | 52   | 17   | 8    | 10   | 39   |
|                                              | igen (%) | 32   | 45   | 47   | 48   | 83   | 92   | 90   | 61   |
| Proxy eljárással összekapcsolható?           | nem      | 1738 | 2661 | 1692 | 2418 | 3014 | 2982 | 1514 | 1211 |
|                                              | igen     | 555  | 1012 | 737  | 1085 | 3183 | 3356 | 1261 | 769  |
|                                              | nem (%)  | 76   | 72   | 70   | 69   | 49   | 47   | 55   | 61   |
|                                              | igen (%) | 24   | 28   | 30   | 31   | 51   | 53   | 45   | 39   |

A 3. táblázatból a következők láthatók: egyrészt - összhangban a korábbiakkal – jellemzően a korábbi évek (2005-2008) tartalmazzák a legtöbb sikertelen összekötést. A proxy eljárásról pedig az látható, hogy abszolút értékben nem igazán lehet kiemelni olyan éveket – legalábbis nem a korábbi törésvonalak mentén – melyekben szignifikánsan jobban szerepelne – de az utolsó 4 év összekötéseit tekintve relatíve sikeresebb. Mindemellett az AF-ER relációban számít legsikeresebbnek az indirekt eljárás. Ez mutatja be a 3. ábra.

**3. táblázat: Direkt összekötések után hiányzó összekötések, proxy eljárások sikeressége, sikertelen összekapcsolások száma**

|                                                                     | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|---------------------------------------------------------------------|------|------|------|------|------|------|------|------|
| Direkt (és direkt fantom) eljárás után hiányzó összekötések száma   | 1555 | 2005 | 1287 | 1819 | 1051 | 513  | 271  | 768  |
| Sikeres proxy eljárások száma a sikertelen direkt eljárások mellett | 167  | 146  | 58   | 122  | 232  | 109  | 56   | 261  |
| Hiányzó összekötések száma                                          | 1388 | 1859 | 1229 | 1697 | 819  | 404  | 215  | 507  |

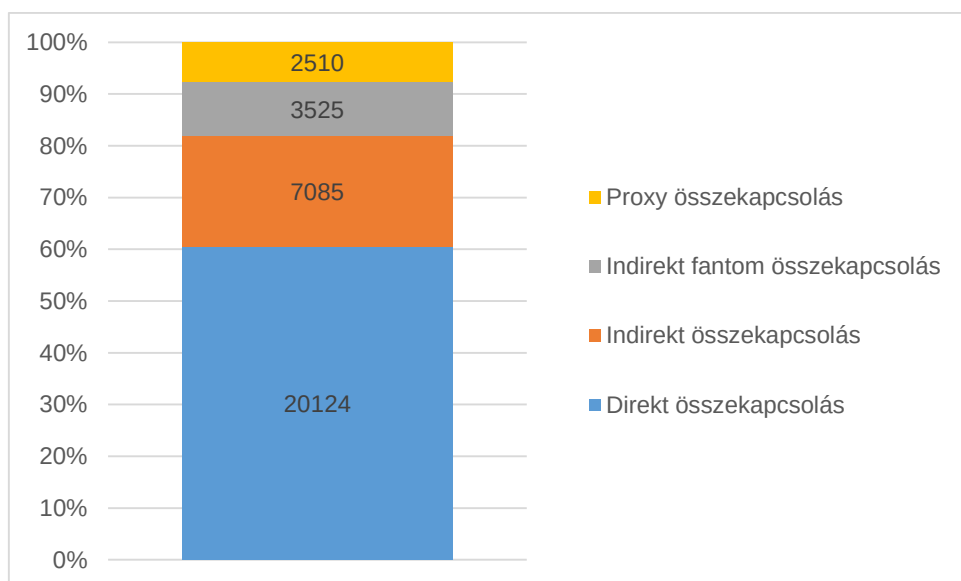
**3. ábra: Hiányzó összekötések és indirekt összekötések száma (az oszlopok teljes magassága a direkt összekötések után is egyedül álló ajánlati felhívásokat jelenti)**



### 3.1.1. AF-ER teljes adatbázis elemzés

A részletes – tehát a minden egyes összekapcsolást tartalmazó adatbázis alapján – ahogyan az 4. ábrán is látszódik – 39850 összekötést találtunk. Mivel az adatbázisban az egyedi ajánlati felhívások száma – az „analitikus2” fájljal összhangban – 29188 (a 2005-2012-es időszakban), így átlagosan kb. 1,36 eredményhirdetés jut egy ajánlati felhívásra. Ehhez képest vannak egyrészt olyan ajánlati felhívások, melyekhez egyáltalán nem lehetett hozzákapcsolni semmilyen ajánlati felhívást, másrészt olyanok, melyekhez bár hozzá lehetett kötni proxy eljárással valamilyen ER-t, azonban a tartalmi leírásaik nem egyeztek annyira, hogy helyes összekötésnek fogadjuk el.

**4. ábra: Az összes AF-ER összekapcsolások %-os eloszlása**



A 4. táblázatban az adatok struktúrája látszódik. Pl. a 3-as item\_id1, jelzi a 3-as ID-val ellátott ajánlati felhívást, amihez két eredményhirdetményt – a 97132, illetve 97017 ID-val rendelkezőt – lehetett hozzákapcsolni indirekt (item\_id2 oszlop), mind indirekt fantom (item\_id3 oszlop) eljárással, illetve a 97132-es ID-val rendelkező eredményhirdetményt proxy eljárással (item\_id4 oszlop) is meg lehetett találni. Direkt eljárással (és minden más típusú eljárással is) pedig a 2-es ID-val (item\_id1) rendelkező AF-nek lehetett megtalálni a párját, ami a 94053-as ID-val rendelkező ER. Ezt a direkt oszlop jelzi.

**4. táblázat: A teljes AF-ER adatbázis részlete**

| direkt_v<br>agy_indi<br>rekt | dir<br>ekt | indirek<br>t_fanto<br>m | proxy | item<br>_id1 | év1  | euhl1 | tipus1 | item_id2 | év2  | item_id3 | év3  | item_id4 | év4  | rmatch4 |
|------------------------------|------------|-------------------------|-------|--------------|------|-------|--------|----------|------|----------|------|----------|------|---------|
| 0                            | 0          | 0                       | 1     | 1            | 2005 | 1     | 1      |          |      |          |      | 93700    | 2006 | 100     |
| 1                            | 1          | 1                       | 1     | 2            | 2005 | 0     | 1      | 94053    | 2006 | 94053    | 2006 | 94053    | 2006 | 100     |
| 1                            | 0          | 1                       | 1     | 3            | 2005 | 0     | 1      | 97132    | 2006 | 97132    | 2006 | 97132    | 2006 | 100     |
| 1                            | 0          | 1                       | 0     | 3            | 2005 | 0     | 1      | 97017    | 2006 | 97017    | 2006 |          |      |         |
| 0                            | 0          | 0                       | 1     | 4            | 2005 | 1     | 1      |          |      |          |      | 101191   | 2006 | 34.1    |
| 0                            | 0          | 1                       | 0     | 4            | 2005 | 1     | 1      |          |      | 93640    | 2006 |          |      |         |
| 0                            | 0          | 1                       | 0     | 4            | 2005 | 1     | 1      |          |      | 280      | 2005 |          |      |         |
| 1                            | 0          | 1                       | 1     | 4            | 2005 | 1     | 1      | 92340    | 2006 | 92340    | 2006 | 92340    | 2006 | 100     |
| 1                            | 1          | 1                       | 0     | 5            | 2005 | 0     | 1      | 92709    | 2006 | 92709    | 2006 |          |      |         |
| 1                            | 1          | 1                       | 0     | 6            | 2005 | 0     | 1      | 97132    | 2006 | 97132    | 2006 |          |      |         |
| 1                            | 1          | 1                       | 1     | 6            | 2005 | 0     | 1      | 97017    | 2006 | 97017    | 2006 | 97017    | 2006 | 97.4    |

Az 5. táblázat azt mutatja, hogy az egyes ajánlati felhívásokhoz a különböző összekapcsolási eljárásokkal hány darab eredményhirdetményt lehetett kapcsolni. Látható, hogy pl. az indirekt fantom eljárás 5 AF-hez is több, mint 100 eredményhirdetményt kapcsolt, míg az indirekt fantom 11 AF esetében vezetett ilyen sok kapcsolathoz. Ezek az összekapcsolások nagyon valószínű, hogy helytelenek, látható tehát, hogy az alternatív – nem közvetlen iktatószám alapú hivatkozással összekapcsolások fals eredményekhez – másodfajú hibákhoz – vezethetnek.

**5. táblázat: A különböző eljárásokkal talált összekapcsolások száma (összes indirekt kapcsolattal)**

| Darabszám | Direkt | Indirekt | Indirekt fantom | Proxy |
|-----------|--------|----------|-----------------|-------|
| 0-10      | 29247  | 29226    | 29167           | 29188 |
| 11-25     | 9      | 15       | 25              | 0     |
| 26-50     | 3      | 7        | 60              | 0     |
| 51-100    | 0      | 6        | 6               | 0     |
| 100+      | 0      | 5        | 11              | 0     |

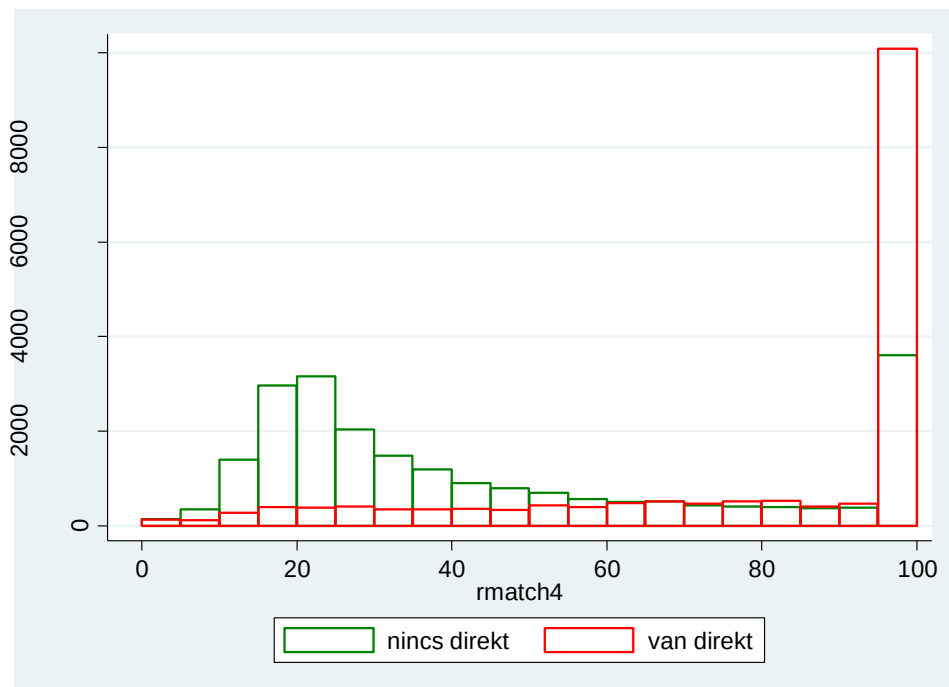
Az rmatch változó értéke – ami az indirekt kapcsolat erősségét ragadja meg egy 0-100-as skálán, ahogyan az a korábbi fejezetekben már tisztáztuk – abban az esetben, ha van sikeres összekapcsolás kb. 80, míg abban az esetben, ha nem található sikeres direkt összekapcsolás, akkor kb. 45 volt a 6. táblázat szerint.

**6. táblázat: Indirekt összekapcsolás minősége – feltételes átlagok**

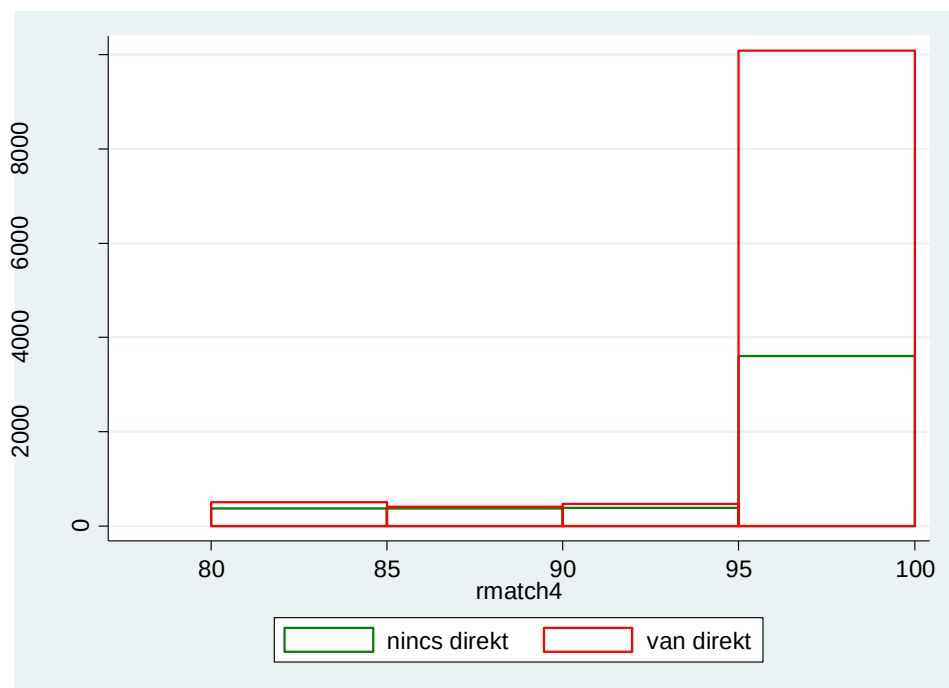
|                                    | Átlag    | Szórás    | Konfidencia intervallum |          | Megfigyelések száma |
|------------------------------------|----------|-----------|-------------------------|----------|---------------------|
| „rmatch” ha nincs direkt kapcsolat | 45,95129 | 0,2063124 | 45,5469                 | 46,35568 | 22260               |
| „rmatch” ha van direkt kapcsolat   | 80,50405 | 0,2127797 | 80,08698                | 80,92112 | 17306               |

A 5. ábra azt mutatja, hogy a „rmatch” változó – vagyis az indirekt összekapcsolás minőségi változója – hogyan alakul a sikeres és sikertelen direkt összekapcsolások esetében. Itt is látható, hogy a direkt összekapcsolások szerint szignifikáns különbség van az indirekt eljárások pontossága között: tehát az indirekt eljárással összekapcsolt hirdetmények nagy valószínűséggel valódi kapcsolatok. A 13. ábra ezt az összefüggést egy szűkített mintán (rmatch > 80) mutatja meg.

**5. ábra: Az „rmatch” változó értéke a sikeres és sikertelen direkt összekapcsolások esetén (AF-ER)**



**6. ábra: Az „rmatch” változó értéke a sikeres és sikertelen direkt összekapcsolások esetén, rmatch>80 esetén (AF-ER)**



### 3.2. Korrigendumok és ajánlati felhívások összekapcsolása

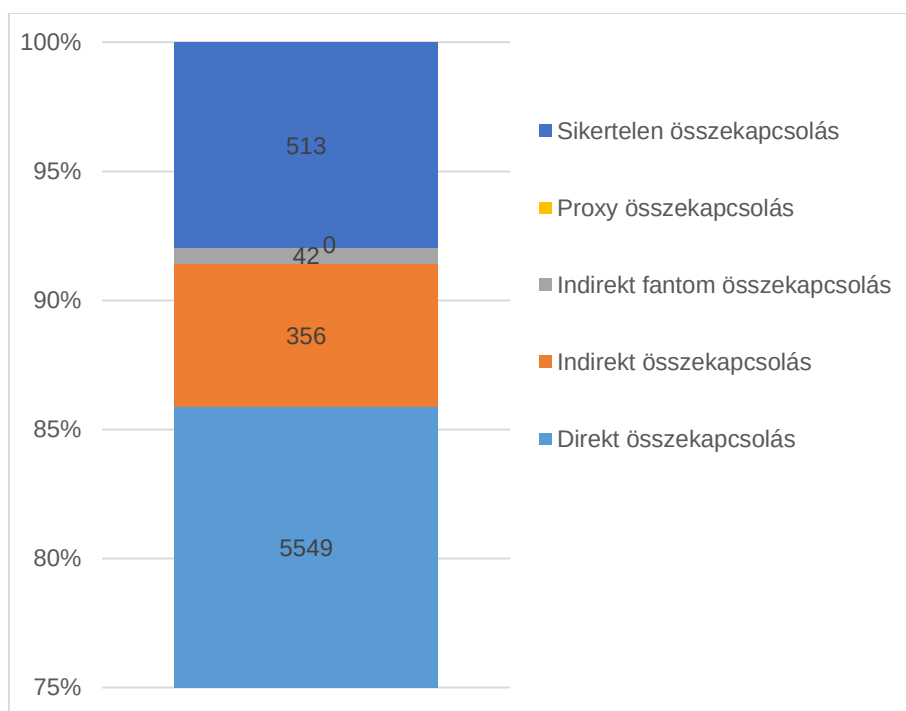
A korrigendumok és ajánlati felhívások leíró statisztikái a 7. táblázat láthatók. Ebben az esetben is, az összekapcsolási módszertanból kifolyólag, bizonyos korrigendumok kétszeresen szerepelhetnek az adatbázisban az uniós közzétételen (EUHL) keresztül, illetve vannak olyan korrigendumok is, melyeket visszavontak (23 darab, a 7. táblázat alapján). Ezen nem releváns esetek kiszűrése után **összesen 6460 darab korrigendum** maradt, melyeket össze kellene kötni egy ajánlati felhívással. A 7. táblázat alapján látható, hogy **az esetek 86%-ban ebben a relációban az összekötés direkt eljárással is sikeres**, továbbá ahogyan az ajánlati felhívások és eredményhirdetmények esetében is, a direkt fantom eljárás nem vezet szignifikánsan több összekapcsoláshoz az „egyszerű” direkt összekapcsolásokon felül. A proxy eljárással a korrigendumok 18%-hoz található ajánlati felhívás – tehát ez a módszer a direkt eljárásokhoz képest meglehetősen rosszul teljesít.

**7. táblázat: Korrigendumok és ajánlati felhívások összekapcsolásának leíró statisztikái (2005-2012)**

|          | EUHL-e? |      | Visszavont-e? |      | Direkt eljárással összekapcsolt-e? |      | Indirekt eljárással összekapcsolt |      | Inirekt fantom eljárással összekapcsolt-e? |      | Proxy eljárással összekapcsolt-e? |      |
|----------|---------|------|---------------|------|------------------------------------|------|-----------------------------------|------|--------------------------------------------|------|-----------------------------------|------|
|          | db      | %    | db            | %    | db                                 | %    | db                                | %    | db                                         | %    | db                                | %    |
| Nem      | 6482    | 61%  | 10685         | 100% | 911                                | 14%  | 555                               | 9%   | 513                                        | 8%   | 5312                              | 82%  |
| Igen     | 4226    | 39%  | 23            | 0%   | 5549                               | 86%  | 5905                              | 91%  | 5947                                       | 92%  | 1148                              | 18%  |
| Összesen | 10708   | 100% | 10708         | 100% | 6460                               | 100% | 6460                              | 100% | 6460                                       | 100% | 6460                              | 100% |

Az alábbi 7. ábra alapján látható, hogy összesen csak a korrigendumok kb. 6-7%-át nem sikerült összekapcsolni valamilyen ajánlati felhívással.

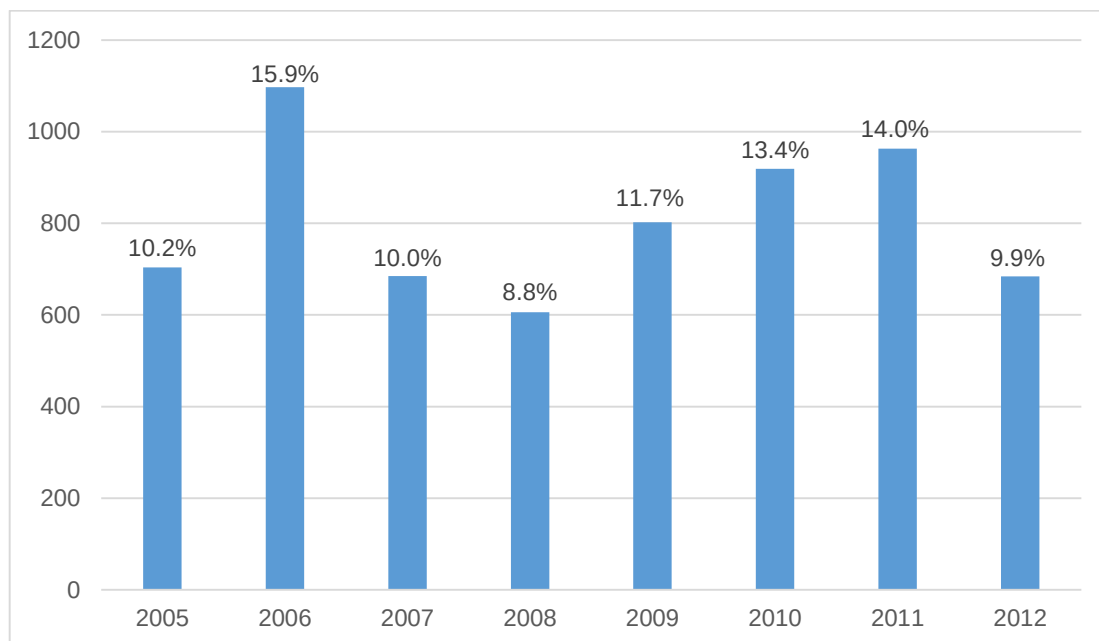
**7. ábra: A korrigendumok és ajánlati felhívások összekapcsolásának sikeressége**





Az ajánlati felhívásoktól eltérően a korigendumok évek közti eloszlása nem koncentrálódik a 2009-2010-es évekre, hanem nagyságrendileg egyenletes – bár nagy szórású eloszlást mutat – ahogyan ez a 8. ábraán látható.

**8. ábra: Az korigendumok évenkénti eloszlása**



A 8. táblázat alapján látszik, hogy a sikertelen direkt eljárások itt is jellemzően a korai 2005-2007 évekre jellemzőek, míg a sikeres indirekt eljárások pedig szintén inkább a 2008-2012-es időszakra, azonban ez a mintázat ebben a relációban sokkal gyengébb, mint más esetekben.

**8. táblázat: A különböző eljárási módszerek sikerességének évenkénti eloszlása (KOR-AF)**

|                                                    |          | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|----------------------------------------------------|----------|------|------|------|------|------|------|------|------|
| Direkt eljárással<br>összekapcsolható?             | nem      | 231  | 175  | 139  | 72   | 130  | 57   | 73   | 34   |
|                                                    | igen     | 473  | 922  | 546  | 534  | 672  | 862  | 890  | 650  |
|                                                    | nem (%)  | 33   | 16   | 20   | 12   | 16   | 6    | 8    | 5    |
|                                                    | igen (%) | 67   | 84   | 80   | 88   | 84   | 94   | 92   | 95   |
| Indirekt eljárással<br>összekapcsolható?           | nem      | 219  | 131  | 102  | 40   | 5    | 17   | 34   | 7    |
|                                                    | igen     | 485  | 966  | 583  | 566  | 797  | 902  | 929  | 677  |
|                                                    | nem (%)  | 31   | 12   | 15   | 7    | 1    | 2    | 4    | 1    |
|                                                    | igen (%) | 69   | 88   | 85   | 93   | 99   | 98   | 96   | 99   |
| Indirekt fantom<br>eljárással<br>összekapcsolható? | nem      | 214  | 111  | 86   | 39   | 5    | 17   | 34   | 7    |
|                                                    | igen     | 490  | 986  | 599  | 567  | 797  | 902  | 929  | 677  |
|                                                    | nem (%)  | 30   | 10   | 13   | 6    | 1    | 2    | 4    | 1    |
|                                                    | igen (%) | 70   | 90   | 87   | 94   | 99   | 98   | 96   | 99   |
| Proxy eljárással<br>összekapcsolható?              | nem      | 698  | 1068 | 655  | 525  | 704  | 628  | 585  | 449  |
|                                                    | igen     | 6    | 29   | 30   | 81   | 98   | 291  | 378  | 235  |
|                                                    | nem (%)  | 99   | 97   | 96   | 87   | 88   | 68   | 61   | 66   |
|                                                    | igen (%) | 1    | 3    | 4    | 13   | 12   | 32   | 39   | 34   |

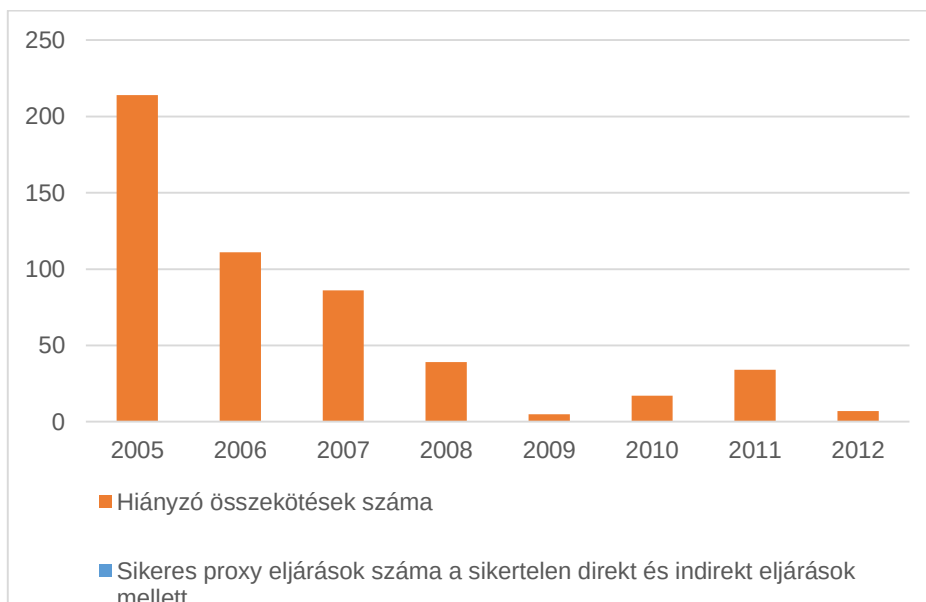
A 9. táblázat alapján látható, hogy a proxy eljárás egyáltalán nem volt sikeres a hiányzó összekötések pótlásában, továbbá a hiányzó összekötések itt is főleg a korábbi éveket jellemzik.

**9. táblázat: Direkt összekötések után hiányzó összekötések, indirekt eljárások sikeressége, sikertelen összekapcsolások száma**

|                                                                                 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|---------------------------------------------------------------------------------|------|------|------|------|------|------|------|------|
| Direkt (és direkt fantom) eljárás után hiányzó összekötések száma               | 214  | 111  | 86   | 39   | 5    | 17   | 34   | 7    |
| Sikeres proxy eljárások száma a sikertelen direkt és indirekt eljárások mellett | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    |
| Hiányzó összekötések száma                                                      | 214  | 111  | 86   | 39   | 5    | 17   | 34   | 7    |

Az alábbi 9. ábra a 9. táblázat adatait jeleníti meg.

**9. ábra: Hiányzó összekötések és sikeres indirekt összekapcsolások: az indirekt eljárás sikeressége**



### 3.3. Korrigendumok és eredményhirdetmények

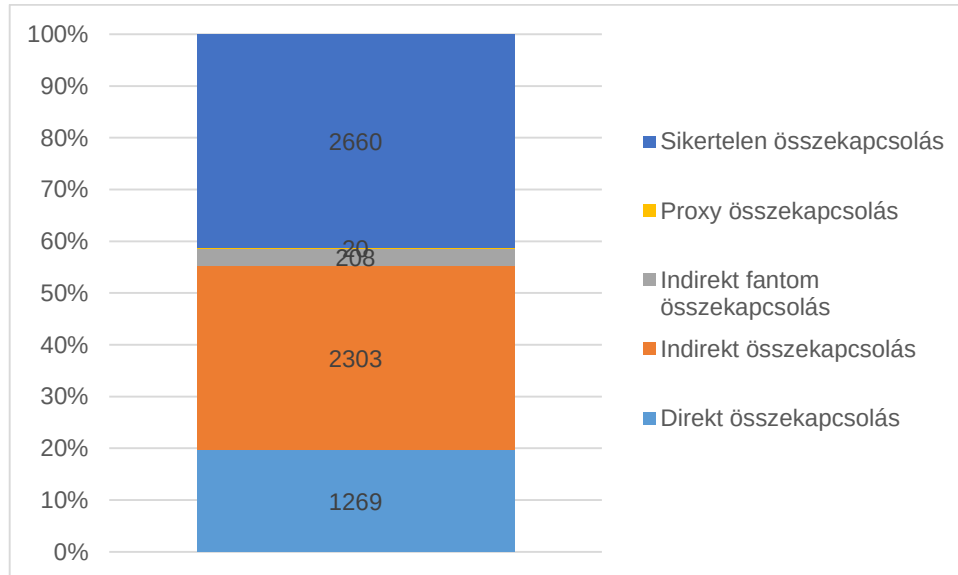
Az előző alfejezetben már tárgyalt releváns korrigendumok száma – az EUHL-s és visszavont korrigendumok kiszűrése után – definíció szerint itt is ugyanúgy 6881. Azonban a KORR-AF kapcsolatok összekapcsolásának sikerességétől jelentősen eltérően, direkt eljárással itt **a korrigendumok kb. 65%-át lehetett csak összekapcsolni valamilyen eredményhirdetménnyel**. Ehhez képest az indirekt, indirekt fantom eljárás sem vezet érdemben jobb eredményre, az proxy összekapcsolások pedig teljesen sikertelennek mondhatók, összesen 64 KORR-ER kapcsolatot lehetett ezzel a módszerrel összekötni.

10. táblázat: A KORR-ER összekapcsolások leíró statisztikái (2005-2012)

|          | EUHL-e? |      | Visszavont-e? |     | Direkt eljárással összekapcsolt-e? |      | Indirekt eljárással összekapcsolt-e? |      | Direkt fantom eljárással összekapcsolt-e? |      | Proxy eljárással összekapcsolt-e? |      |
|----------|---------|------|---------------|-----|------------------------------------|------|--------------------------------------|------|-------------------------------------------|------|-----------------------------------|------|
|          | db      | %    | db            | %   | db                                 | %    | db                                   | %    | db                                        | %    | db                                | %    |
| Nem      | 6482    | 61%  | 10685         | 92% | 1269                               | 20%  | 3572                                 | 55%  | 2680                                      | 41%  | 6396                              | 99%  |
| Igen     | 4226    | 39%  | 23            | 0%  | 5191                               | 80%  | 2888                                 | 45%  | 3780                                      | 59%  | 64                                | 1%   |
| Összesen | 10708   | 100% | 10708         | 92% | 6460                               | 100% | 6460                                 | 100% | 6460                                      | 100% | 6460                              | 100% |

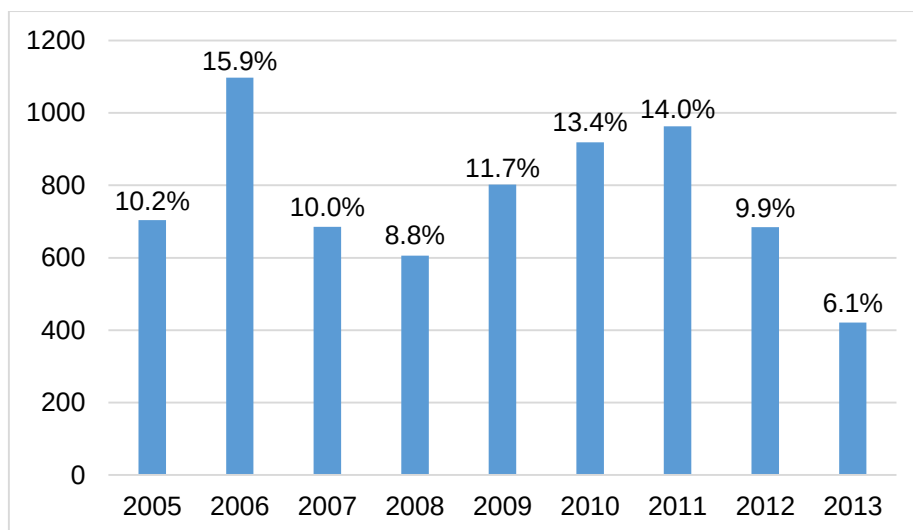
A 10. ábra részletesen is bemutatja, hogy összesen kb. a korrigendumok 30%-át semmilyen eredményhirdetményhez nem lehetett hozzákapcsolni, az indirekt fantom, illetve a proxy eljárások sikeressége is teljesen elhanyagolható.

10. ábra: A korrigendumok és eredményhirdetmények összekapcsolásának sikeressége



A 11. ábra megegyezik a 8. ábraával: a korrigendumok évenkénti eloszlását mutatja.

**11. ábra: A korrigendumok évenkénti eloszlása**



A 11. táblázat a különböző eljárások évenkénti sikerességét mutatja. A KORR-AF relációkhoz hasonlóan a 2009-2012-es évek a legsikeresebbek a direkt összekötések szempontjából, de ez a sikeresség csak igen mérsékelt: 26-33% között van a direkt összekapcsolások aránya.

**11. táblázat: A különböző eljárási módszerek sikerességének évenkénti eloszlása KORR-ER relációban**

|                                              |          | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|----------------------------------------------|----------|------|------|------|------|------|------|------|------|
| Direkt eljárással összekapcsolható?          | nem      | 657  | 1001 | 593  | 530  | 594  | 683  | 648  | 485  |
|                                              | igen     | 47   | 96   | 92   | 76   | 208  | 236  | 315  | 199  |
|                                              | nem (%)  | 93   | 91   | 87   | 87   | 74   | 74   | 67   | 71   |
|                                              | igen (%) | 7    | 9    | 13   | 13   | 26   | 26   | 33   | 29   |
| Indirekt eljárással összekapcsolható?        | nem      | 581  | 706  | 497  | 429  | 404  | 500  | 552  | 488  |
|                                              | igen     | 123  | 391  | 188  | 177  | 398  | 419  | 411  | 196  |
|                                              | nem (%)  | 83   | 64   | 73   | 71   | 50   | 54   | 57   | 71   |
|                                              | igen (%) | 17   | 36   | 27   | 29   | 50   | 46   | 43   | 29   |
| Indirekt fantom eljárással összekapcsolható? | nem      | 450  | 530  | 365  | 349  | 196  | 264  | 237  | 289  |
|                                              | igen     | 254  | 567  | 320  | 257  | 606  | 655  | 726  | 395  |
|                                              | nem (%)  | 64   | 48   | 53   | 58   | 24   | 29   | 25   | 42   |
|                                              | igen (%) | 36   | 52   | 47   | 42   | 76   | 71   | 75   | 58   |
| Proxy eljárással összekapcsolható?           | nem      | 704  | 1096 | 682  | 603  | 796  | 910  | 934  | 671  |
|                                              | igen     | 0    | 1    | 3    | 3    | 6    | 9    | 29   | 13   |
|                                              | nem (%)  | 100  | 100  | 100  | 100  | 99   | 99   | 97   | 98   |
|                                              | igen (%) | 0    | 0    | 0    | 0    | 1    | 1    | 3    | 2    |

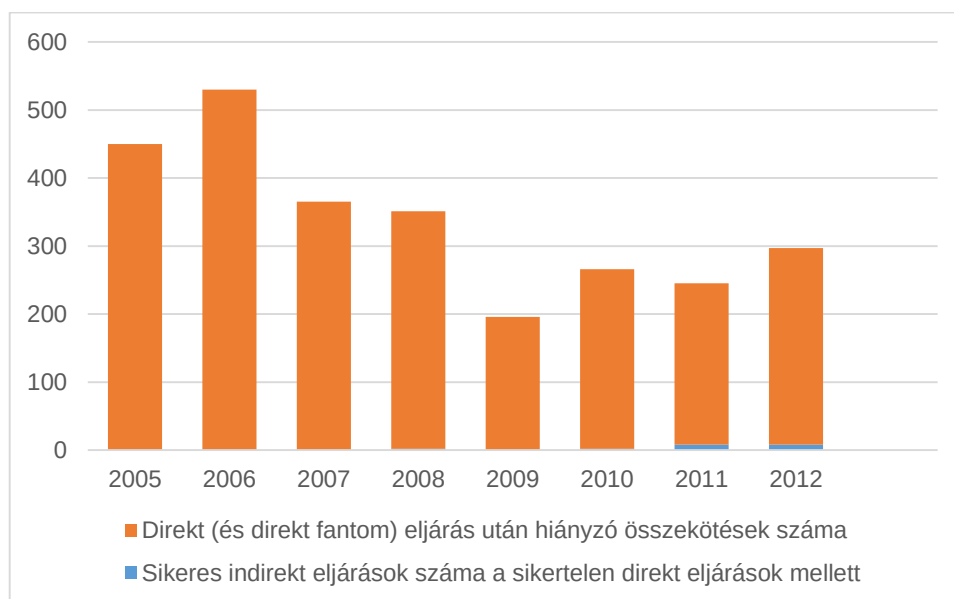
A 12. táblázat két dolgot mutat: a legtöbb egyedülálló korrigendum a 2005-2008-as években van; az indirekt eljárások az utolsó években sikeresek, de csak elhanyagolható mértékben.

**12. táblázat: Direkt összekötések után hiányzó összekötések, indirekt eljárások sikeressége, sikertelen összekapcsolások száma**

|                                                                        | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|------------------------------------------------------------------------|------|------|------|------|------|------|------|------|
| Direkt (és direkt fantom) eljárás után hiányzó összekötések száma      | 450  | 530  | 365  | 349  | 196  | 264  | 237  | 289  |
| Sikeres indirekt eljárások száma a sikertelen direkt eljárások mellett | 0    | 0    | 0    | 2    | 0    | 2    | 8    | 8    |
| Hiányzó összekötések száma                                             | 450  | 530  | 365  | 347  | 196  | 262  | 229  | 281  |

A 19. ábra a 12. táblázat eredményeit ábrázolja.

**12. ábra: Hiányzó összekötések és indirekt összekötések száma (az oszlopok teljes magassága a direkt összekötések után is egyedül álló ajánlati felhívásokat jelenti)**



### 3.4. Szerződésmódosítások és eredményhirdetések összekapcsolása

A szerződésmódosítások és eredményhirdetmények leíró statisztikái az

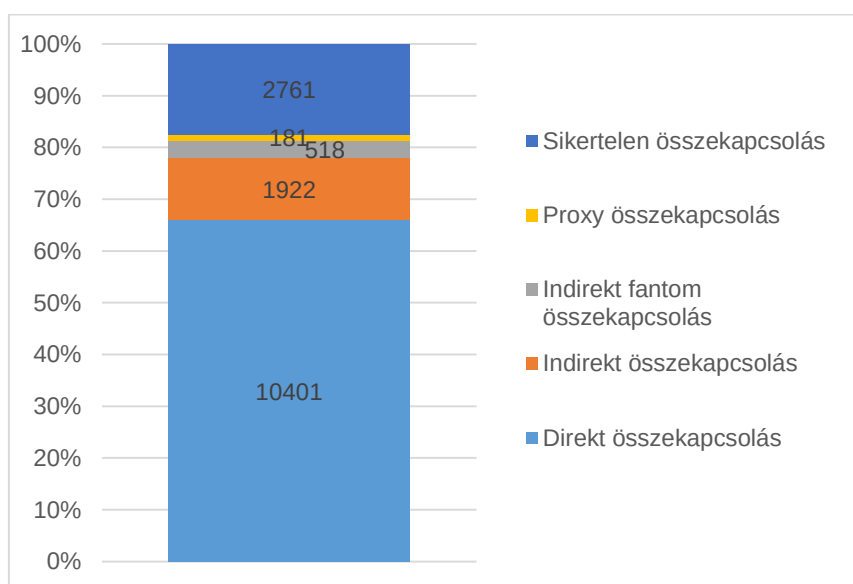
13. táblázatban láthatók. Az előző összekapcsolásokkal ellentétben itt nem okoz külön problémát az uniós hivatalos lapban való szereplés, illetve mindössze egyetlen megfigyelési sor esik ki a visszavonások miatt is. Így összesen 15783 darab szerződésmódosítás maradt, melyeket eredményhirdetményekkel kell összekötni. Ennek nagyjából 66%-át sikerült direkt összekapcsolással összekötni, illetve ebben az esetben fordult elő először, hogy az indirekt illetve indirekt fantom eljárás szignifikánsan sikeresebbnek mondható, mint az „egyszerű” direkt eljárás.

**13. táblázat: A szerződésmódosítások és eredményhirdetések összekapcsolásának leíró statisztikái**

|          | EUHL-e? |      | Visszavont-e? |     | Direkt eljárással összekapcsolt-e? |      | Indirekt eljárással összekapcsolt-e? |      | Indirekt fantom eljárással összekapcsolt-e? |      | Proxy eljárással összekapcsolt-e? |      |
|----------|---------|------|---------------|-----|------------------------------------|------|--------------------------------------|------|---------------------------------------------|------|-----------------------------------|------|
|          | db      | %    | db            | %   | db                                 | %    | db                                   | %    | db                                          | %    | db                                | %    |
| Nem      | 15784   | 100% | 15783         | 92% | 5382                               | 34%  | 3460                                 | 22%  | 2942                                        | 19%  | 12792                             | 81%  |
| Igen     | 0       | 0%   | 1             | 0%  | 10401                              | 66%  | 12323                                | 78%  | 12841                                       | 81%  | 2991                              | 19%  |
| Összesen | 15784   | 100% | 15784         | 92% | 15783                              | 100% | 15783                                | 100% | 15783                                       | 100% | 15783                             | 100% |

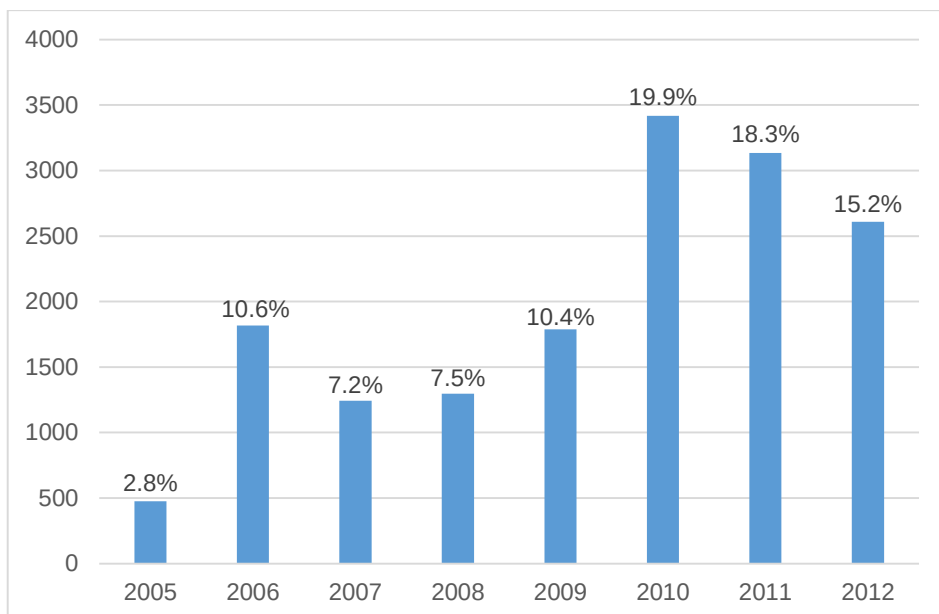
A 13. ábra azt mutatja, hogy a szerződésmódosítások közül nagyságrendileg 17%-ot nem lehetett hozzákapcsolni valamilyen eredményhirdetményhez, illetve hogy az indirekt, illetve proxy eljárásokkal kb. 16%-át lehetett megtalálni a dokumentumpárokban (melyek között azonban lehet, hogy másodfajú hibát vétünk az összekapcsolással).

**13. ábra: A szerződésmódosítások és eredményhirdetmények összekapcsolásának sikeressége**



A 14. ábra a szerződésmódosítások évenkénti eloszlását mutatja: a 2010-2011-es évekre kb. a módosítások 40%-a.

**14. ábra: A szerződésmódosítások évenkénti eloszlása**



A 14. táblázat alapján az összekötések sikeressége nagyjából a korábbiaknak megfelelően alakult: **2010-2012-es** években a legsikeresebb a **direkt eljárás**, illetve az összes többi összekötési módszer is ezekben az években működik a legjobban.

**14. táblázat: A különböző eljárási módszerek sikerességének évenkénti eloszlása**

|                                              |          | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|----------------------------------------------|----------|------|------|------|------|------|------|------|------|
| Direkt eljárással összekapcsolható?          | nem      | 377  | 1054 | 481  | 606  | 842  | 1051 | 491  | 480  |
|                                              | igen     | 99   | 764  | 762  | 690  | 945  | 2367 | 2644 | 2130 |
|                                              | nem (%)  | 79   | 58   | 39   | 47   | 47   | 31   | 16   | 18   |
|                                              | igen (%) | 21   | 42   | 61   | 53   | 53   | 69   | 84   | 82   |
| Indirekt eljárással összekapcsolható?        | nem      | 372  | 977  | 437  | 434  | 327  | 371  | 233  | 309  |
|                                              | igen     | 104  | 841  | 806  | 862  | 1460 | 3047 | 2902 | 2301 |
|                                              | nem (%)  | 78   | 54   | 35   | 33   | 18   | 11   | 7    | 12   |
|                                              | igen (%) | 22   | 46   | 65   | 67   | 82   | 89   | 93   | 88   |
| Indirekt fantom eljárással összekapcsolható? | nem      | 355  | 834  | 307  | 303  | 285  | 333  | 225  | 300  |
|                                              | igen     | 121  | 984  | 936  | 993  | 1502 | 3085 | 2910 | 2310 |
|                                              | nem (%)  | 75   | 46   | 25   | 23   | 16   | 10   | 7    | 11   |
|                                              | igen (%) | 25   | 54   | 75   | 77   | 84   | 90   | 93   | 89   |
| Proxy eljárással összekapcsolható?           | nem      | 452  | 1583 | 1085 | 1135 | 1505 | 2636 | 2325 | 2071 |
|                                              | igen     | 24   | 235  | 158  | 161  | 282  | 782  | 810  | 539  |
|                                              | nem (%)  | 95   | 87   | 87   | 88   | 84   | 77   | 74   | 79   |
|                                              | igen (%) | 5    | 13   | 13   | 12   | 16   | 23   | 26   | 21   |

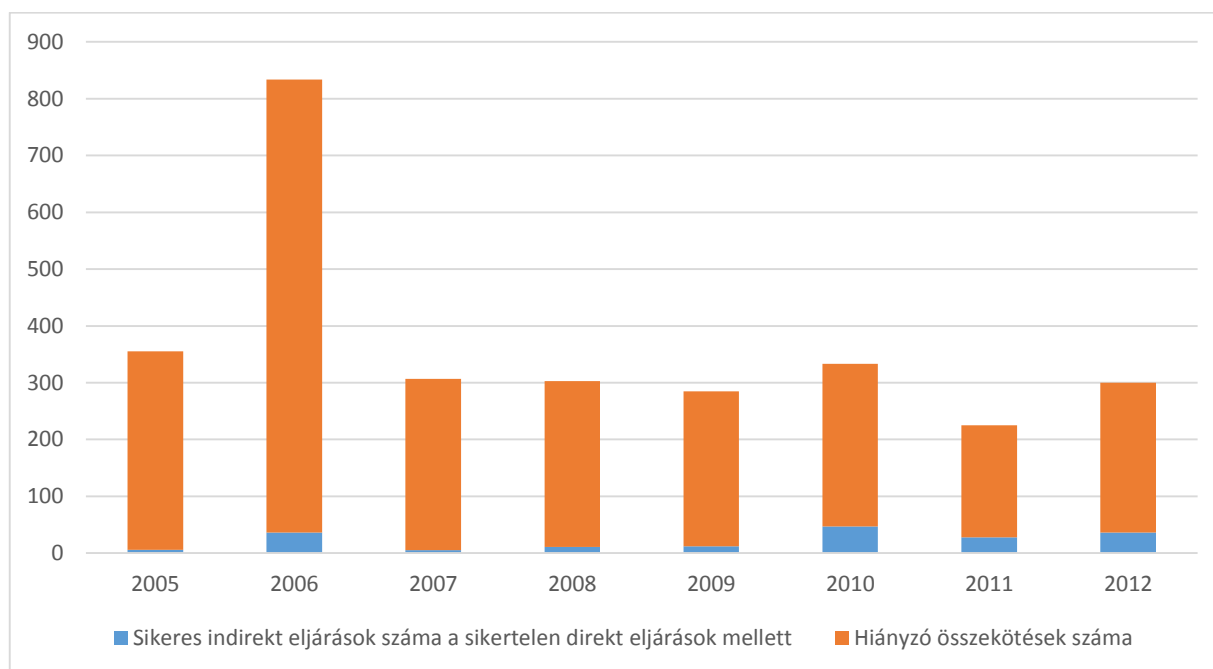
A 15. táblázat és

15. ábra alapján látható, hogy egyrészt a legtöbb hiányzó összekötés a 2005, 2006-os években van, másrészt a proxy eljárás sem ad hozzá érdemben a találatokhoz.

**15. táblázat: Direkt összekötések után hiányzó összekötések, indirekt eljárások sikeressége, sikertelen összekapcsolások száma**

|                                                                        | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|------------------------------------------------------------------------|------|------|------|------|------|------|------|------|
| Direkt (és direkt fantom) eljárás után hiányzó összekötések száma      | 355  | 834  | 307  | 303  | 285  | 333  | 225  | 300  |
| Sikeres indirekt eljárások száma a sikertelen direkt eljárások mellett | 6    | 36   | 5    | 11   | 12   | 47   | 28   | 36   |
| Hiányzó összekötések száma                                             | 349  | 798  | 302  | 292  | 273  | 286  | 197  | 264  |

**15. ábra: Hiányzó összekötések és sikeres indirekt összekapcsolások: az indirekt eljárás sikeressége**





### 3.5. Szerződés teljesítések és eredményhirdetések összekapcsolása

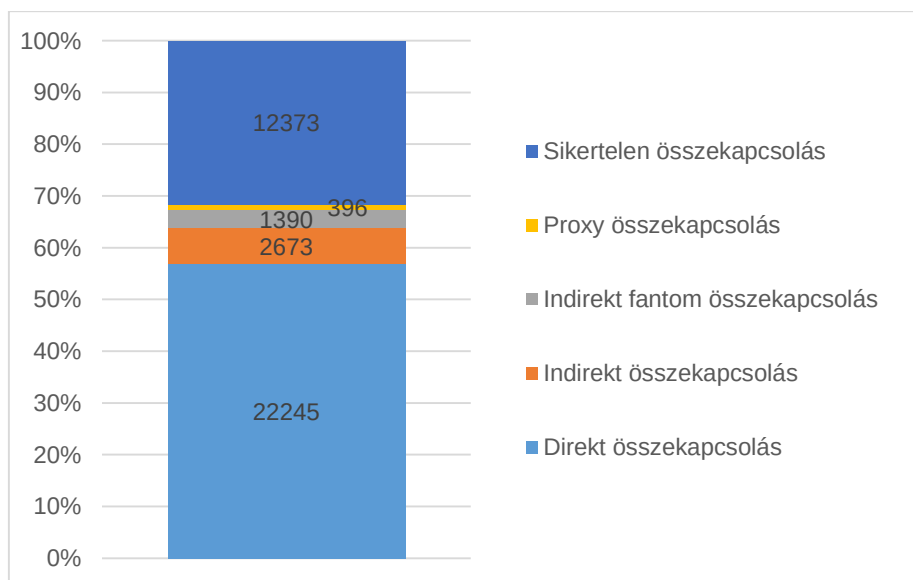
Ebben az alfejezetben a szerződésteljesítések és eredményhirdetések összekapcsolásának sikerességét mutatjuk be. A 16. táblázat alapján látható, hogy összesen 39077 releváns SZT-t tartalmaz az adatbázis a 2005-2012-es időszakban. Továbbá látható, hogy az SZT-k 57%-át lehetett összekapcsolni direkt eljárással. Az indirekt, indirekt fantom eljárás további 7%, illetve 3%-kal több kapcsolatot eredményezett. A proxy eljárással az SZT-k 16%-hoz lehetett ER-t találni.

**16. táblázat: Az SZT-ER összekapcsolások leíró statisztikái**

|          | EUHL-e? |      | Visszavont-e? |      | Direkt eljárással összekapcsolt-e? |      | Indirekt eljárással összekapcsolt-e? |      | Indirekt fantom eljárással összekapcsolt-e? |      | Proxy eljárással összekapcsolt-e? |      |
|----------|---------|------|---------------|------|------------------------------------|------|--------------------------------------|------|---------------------------------------------|------|-----------------------------------|------|
|          | db      | %    | db            | %    | db                                 | %    | db                                   | %    | db                                          | %    | db                                | %    |
| Nem      | 39077   | 100% | 39077         | 100% | 16832                              | 43%  | 14159                                | 36%  | 12769                                       | 33%  | 32867                             | 84%  |
| Igen     | 0       | 0%   | 0             | 0%   | 22245                              | 57%  | 24918                                | 64%  | 26308                                       | 67%  | 6210                              | 16%  |
| Összesen | 39077   | 100% | 39077         | 100% | 39077                              | 100% | 39077                                | 100% | 39077                                       | 100% | 39077                             | 100% |

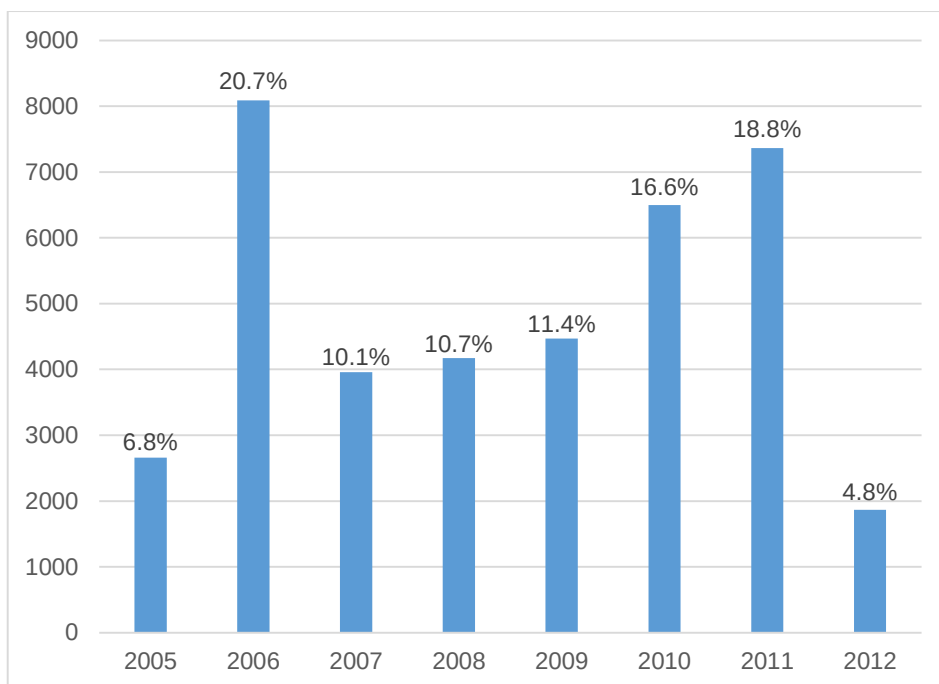
A 16. ábra alapján látható, hogy a direkt eljáráson felül mennyire sikerült újabb SZT-eket összekapcsolni ER-vel: az összesen 22245 direkt összekapcsoláson felül 2673 kapcsolatot lehetett indirekt, 1390 kapcsolatot indirekt fantom, illetve 396-ot proxy eljárás alapján megtalálni.

**16. ábra: Az SZT-ER összekapcsolások %-os eloszlása**



Az alábbi 17. ábraán az SZT-k évenkénti eloszlását láthatjuk. Ahogyan az AF-ek eloszlása alapján várható volt (ott 2009-2010-re koncentrálódik a hirdetmények nagy része), az SZT-k száma relatíve magas a 2010-2011-es években – összesen több mint 35%, illetve még a 2006-os évben is magas a teljesítések száma, azonban ennek a kiírási oldalát csak részben látjuk a vizsgált időintervallumból kifolyólag.

**17. ábra: A szerződésteljesítések évenkénti eloszlása**



Az alábbi 17. táblázatból látszódik, hogy – hasonlóan más összekapcsolási párok esetéhez – itt is ugyanaz a törés figyelhető meg az összekapcsolások sikerességében: míg 2005-2008 között az átlagos összekötési arány 10-56% között mozog (súlyozottan 37%), addig 2009-2012 között ez az arány 60-85% körül mozog (súlyozva pedig 75%).

**17. táblázat: A különböző eljárási módszerek sikerességének évenkénti eloszlása**

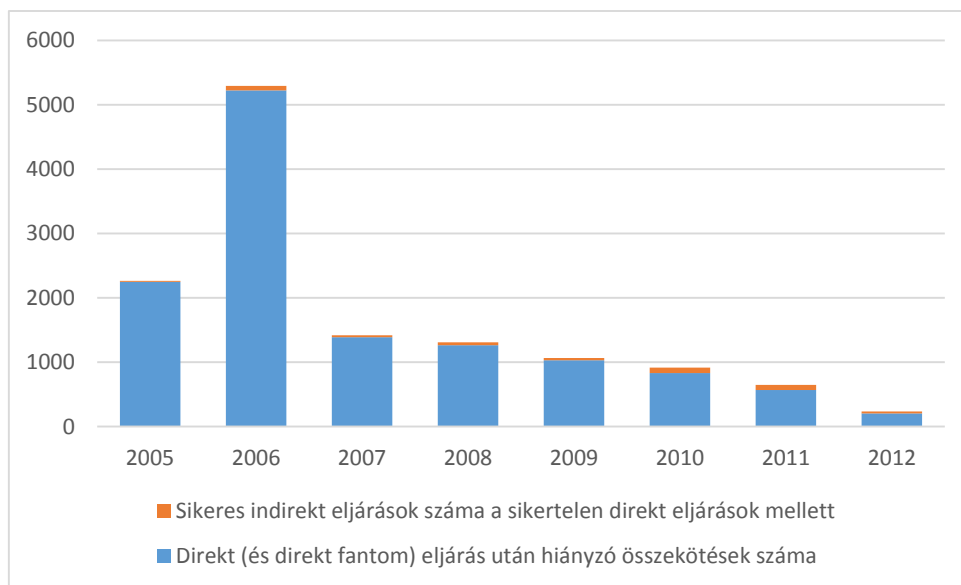
|                                              |          | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|----------------------------------------------|----------|------|------|------|------|------|------|------|------|
| Direkt eljárással összekapcsolható?          | nem      | 2326 | 5763 | 1897 | 1828 | 1779 | 1821 | 1113 | 305  |
|                                              | igen     | 333  | 2326 | 2060 | 2346 | 2690 | 4676 | 6253 | 1561 |
|                                              | nem (%)  | 87   | 71   | 48   | 44   | 40   | 28   | 15   | 16   |
|                                              | igen (%) | 13   | 29   | 52   | 56   | 60   | 72   | 85   | 84   |
| Indirekt eljárással összekapcsolható?        | nem      | 2308 | 5605 | 1740 | 1535 | 1187 | 938  | 635  | 211  |
|                                              | igen     | 351  | 2484 | 2217 | 2639 | 3282 | 5559 | 6731 | 1655 |
|                                              | nem (%)  | 87   | 69   | 44   | 37   | 27   | 14   | 9    | 11   |
|                                              | igen (%) | 13   | 31   | 56   | 63   | 73   | 86   | 91   | 89   |
| Indirekt fantom eljárással összekapcsolható? | nem      | 2251 | 5221 | 1388 | 1265 | 1034 | 832  | 570  | 208  |
|                                              | igen     | 408  | 2868 | 2569 | 2909 | 3435 | 5665 | 6796 | 1658 |
|                                              | nem (%)  | 85   | 65   | 35   | 30   | 23   | 13   | 8    | 11   |
|                                              | igen (%) | 15   | 35   | 65   | 70   | 77   | 87   | 92   | 89   |
| Proxy eljárással összekapcsolható?           | nem      | 2596 | 7601 | 3575 | 3727 | 3790 | 5036 | 5270 | 1272 |
|                                              | igen     | 63   | 488  | 382  | 447  | 679  | 1461 | 2096 | 594  |
|                                              | nem (%)  | 98   | 94   | 90   | 89   | 85   | 78   | 72   | 68   |
|                                              | igen (%) | 2    | 6    | 10   | 11   | 15   | 22   | 28   | 32   |

A 18. táblázatban az látható, hogy a direkt, indirekt (2-féle) és proxy típusú összekötési eljárások használata mellett mennyi egyedülálló SZT marad az egyes években. Látható, hogy a 2005-2006-os években van a legtöbb hiányzó összekötés, 2007-2009 között évi kb. 1000 darab, és a legutóbbi években megy csak le ez a szám a pár százaz nagyságrendre. A 18. ábra a 18. táblázatban bemutatott eredményeket vizualizálja: az proxy eljárás bár lefelezi a hiányzó összekötések számát, így is majdnem 5000 SZT-t nem lehet ER-hez kapcsolni.

**18. táblázat: Direkt összekötések után hiányzó összekötések, indirekt eljárások sikeressége, sikertelen összekapcsolások száma**

|                                                                        | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|------------------------------------------------------------------------|------|------|------|------|------|------|------|------|
| Direkt (és direkt fantom) eljárás után hiányzó összekötések száma      | 2251 | 5221 | 1388 | 1265 | 1034 | 832  | 570  | 208  |
| Sikeres indirekt eljárások száma a sikertelen direkt eljárások mellett | 13   | 73   | 34   | 46   | 34   | 87   | 80   | 29   |
| Hiányzó összekötések száma                                             | 2238 | 5148 | 1354 | 1219 | 1000 | 745  | 490  | 179  |

**18. ábra: Hiányzó összekötések és sikeres indirekt összekapcsolások: az indirekt eljárás sikeressége**



## 4. Összefoglalás

Jelen tanulmányban azt mutattuk be, hogy a Közbeszerzési Hatóság által közzétett közbeszerzési eljárásokkal kapcsolatos hirdetmények mennyire összekapcsolhatók egymással, mennyire lehet felfejteni egy adott közbeszerzéssel kapcsolatos dokumentumokat, ami az eljárások átláthatóságát biztosítja. Ehhez egyrészt bemutattuk, hogy a fellelhető dokumentumok hogyan sorolhatók be a különböző tartalmú hirdetménykategóriákba, illetve milyen direkt, indirekt és proxy változókat használó módszerek alapján alkottunk azonos közbeszerzési eljáráshoz tartozó hirdetmenypárokat. Továbbá bemutattuk, hogy az összekapcsolási hiányosságok pontosan milyen relációkban merülhetnek fel.

A tanulmány legfontosabb hozzájárulása az volt, hogy az AF-ER, KORR-AF, KORR-ER, SZM-ER, SZT-ER párok esetében külön bemutattuk, hogy az esetek hány %-ban lehetett összekötni ezeket a dokumentumokat, tehát hány olyan közbeszerzési eljárás található, mely esetén legalább egy (illetve AF-ER relációban az összes párt is megvizsgáltuk) további hirdetményt is hozzá tudunk kapcsolni az adott AF, ER, SZM, SZT vagy KORR dokumentumhoz.

A számításaink szerint ezek az összekapcsolások csak mérsékelten sikeresnek: az összekapcsolások nagyságrendileg 60-90% között mozognak a különböző relációkban. A legérdekesebb AF-ER illetve SZT-ER relációkban 68-72%-ban találtuk meg a hirdetmenypárokat. Ez azt jelenti, hogy az közbeszerzési kiírások több, mint 30%-a esetén nem lehet megtalálni a nyertes pályázót és vice versa – legalábbis a mi egyszerű módszereink alapján. Illetve ugyanígy több mint 10%-ban nem ellenőrizhető a szerződések teljesítése, hiába látható, hogy mely vállalat kapott egy adott megbízást. Fontos azonban megjegyezni, hogy ezek a számok a legoptimistább feltevésekből következnek, mivel abból indulunk ki, hogy az alternatív összekötési eljárásaink „kevés” másodfajú hibával járnak.

További fontos eredmény, hogy évenként nagyban változik, hogy milyen sikerrel megtalálhatók ezek a hirdetménykapcsolatok: míg a 2005-2008-es időszakban szignifikánsan kevésbé visszafejthetők ezek a relációk – pl. az AF-ER relációban a 2005-2008 közti átlagos direkt módszeres találati arány súlyozva kb. 37%, 2009-2012 között viszont már kb. 80%.

Ezek alapján egyrészt azt mondhatjuk, hogy szerencsére van egy nagyságrendi javulás a közbeszerzési hirdetmények összekapcsolhatóságát tekintve. Azonban még az utóbbi évek találati aránya alapján is a közbeszerzési eljárások kb. 20%-a nem transzparens, ezáltal nem ellenőrizhető, ami jelentős korrupciós kockázatokat rejt, mivel így nem ellenőrizhető, hogy pl. a szerződések valóban az közbeszerzési ajánlatban foglaltaknak mentek-e végbe, vagy nem voltak-e tendenciózus túlárázások a végrehajtáskor.

További kutatás tárgyát képezheti, hogy a jelenleg meg nem talált hirdetmenypárokat milyen más alternatív módszerekkel lehetséges összekötni. Illetve a jelenleg önállóan is vizsgálni lehet az összekapcsolható és nem összekapcsolható hirdetmények tartalmi elemeit. Így felfejthetők például földrajzi mintázatok – olyan települések, ahol jellemzően hirdetmények szintjén nem lekövethetők a közbeszerzések – vagy például azonosíthatók olyan speciális termékkategória csoportok – így iparágak –, melyek jellemzően olyan közbeszerzésekben vesznek részt, melyek nem vagy csak nehezen összeköthető dokumentációval rendelkeznek.

## 5. Melléklet

Az „analitikus2” típusú fájlokban szereplő **elj3\_min** illetve **elj3\_max** változók jelentése a következő: az adott hirdetményhez indirekt eljárással kapcsolt ER-ek közül a legrosszabb/legjobb "jósága" százalékban mérve, ahol 0 a legrosszabb, 100 pedig a legjobb. Ez tehát nem más, mint a legalacsonyabb és legmagasabb rmatch értékű indirekt módon hozzákapcsolt hirdetmény rmatch értéke. Jelenlegi az indirekt eljárások esetében az elfogadási kritérium 95 pont.

### 5.1. Analitikus 1 (hosszabb)

id: egyedi ID

item\_id: AF, SZT, SZM hirdetmény ID

iktatoszam: AF, SZT, SZM iktatószám

kozzetetel\_datuma: AF, SZT, SZM közzététel dátuma

url: AF, SZT, SZM URL

ajanlatkero: AF, SZT, SZM ajánlatkérő neve

rovid\_tartalom: AF, SZT, SZM rövid tartalom

rovid\_tartalom2: AF, SZT, SZM rövid tartalom tisztítva (pl.: ajánlattételi felhívás szöveg kivéve)

er\_item\_id: hozzákapcsolt ER hirdetmény ID

er\_iktatoszam: hozzákapcsolt ER iktatószám

er\_kozzetetel\_datuma: hozzákapcsolt ER közzététel dátuma

er\_url er\_ajanlatkero: hozzákapcsolt ER ajánlatkérő neve

er\_rovid\_tartalom: hozzákapcsolt ER rövid tartalom

er\_rovid\_tartalom2: hozzákapcsolt ER rövid tartalom tisztítva (pl.: ajánlattételi felhívás szöveg kivéve)

n1: a rovid\_tartalom2-ben szereplő legalább 5 hosszú szavak közül hány illeszthető az er\_rovid\_tartalom2-re

n2: az er\_rovid\_tartalom2-ben szereplő legalább 5 hosszú szavak közül hány illeszthető az rovid\_tartalom2-re

rmatch: rovid\_tartalom2 és er\_rovid\_tartalom2 távolsága a hosszabbik szöveg hosszával normálva (0-100)

len: rovid\_tartalom2 és er\_rovid\_tartalom2 közül a hosszabbik hossza

m1: item\_id előfordulásainak száma a teljes listában

m2: er\_item\_id előfordulásainak száma a teljes listában

knum: a direkt összekapcsolásokkal hány nem AF, SZT, SZM hirdetményt sikerült hozzákapcsolni az adott AF, SZT, SZM-hez

knum\_fantom: a direkt fantom összekapcsolásokkal hány nem AF, SZT, SZM hirdetményt sikerült hozzákapcsolni az adott AF, SZT, SZM-hez

voltmar: ez az összekapcsolás szerepelt-e a direkt összekapcsolások között (0 - nem, 1 - igen)

voltmar\_fantom: ez az összekapcsolás szerepelt-e a direkt fantom összekapcsolások között (0 - nem, 1 - igen)

## 5.2. Analitikus 2 (rövidebb)

**sorszam:** egyedi sorszám az excel fájlban

**tipus:** a hirdetmény típusa – ajánlat felhívás, eredmény

**ev:** a hirdetmény közzétételének éve

**item\_id:** a hirdetmény egyedi azonosítója

**url:** a hirdetmény URL címe

**euhl:** EUHL-es a hirdetmény? (0 - nem, 1 - igen)

**megszunt:** megszüntették? (0 - nem, 1 - igen)

**elj1:** a direkt eljárással sikerült-e AF-et (ajánlati felhívást) hozzákapcsolni a hirdetményhez

**elj2:** a direkt fantom eljárással sikerült-e AF-et (ajánlati felhívást) hozzákapcsolni a hirdetményhez

**elj3\_min:** az adott hirdetményhez indirekt eljárással kapcsolt AF-ek közül a legrosszabb "jósága" százalékban mérve (0 - legrosszabb, 100 - legjobb)

**elj3\_max:** az adott hirdetményhez indirekt eljárással kapcsolt AF-ek közül a legjobb "jósága" százalékban mérve (0 - legrosszabb, 100 - legjobb)

**elj3\_db:** az indirekt eljárással hány darab AF-et (ajánlati felhívást) sikerült hozzákapcsolni a hirdetményhez

## 5.3. Felhasznált adatbázisok

af\_er\_minta\_20140113

analitikus2\_af\_er\_20140121\_UT

analitikus2\_kor\_af\_20140121\_UT

analitikus2\_kor\_er\_20140121\_UT

analitikus2\_szm\_er\_20140121\_UT

analitikus2\_szt\_er\_20140121\_UT