

Kruzslicz Ferenc

Pécsi Tudományegyetem

Az online kérdőíves szoftverek és szolgáltatások megjelenésével szinte az elektronikus képeslapküldés egyszerűségével készíthet bárki felméréseket. A minőségi kérdőív készítésének már széleskörű irodalma van, de a biztonságos kérdőívvel kapcsolatban még nem jelent meg publikáció.

Ezt a hiányt töltjük be jelen munkánkkal, ahol az informatikai rendszerfejlesztés analógiájára alapozva és annak módszertanait felhasználva teszünk javaslatot adatvédelmi, adatbiztonsági követelményeknek is megfelelő, hatékony adatgyűjtési folyamat kialakítására. Az elméleti eredmények mellett egy valós gyakorlati példán keresztül be is mutatjuk az általunk javasolt módszer lépéseit, illetve az egyes fázisokhoz tartozó legfontosabb döntési pontokat.

Kulcsszavak: online kérdőív fejlesztés, tervezési és biztonsági kérdések, adatminőség javulás

BEVEZETÉS

A számítógépek elterjedésével szinte egy időben jelentek meg azok az elektronikus kérdőívezési módszerek, amelyekkel a papír alapú adatgyűjtést lehetett kiváltani. Első formájukban a nyomtatott verzió előállításához használt szövegszerkesztői állományokat juttatták el a felmérésben résztvevőkhöz, majd ezek visszagyűjtésével indulhatott meg az adatfeldolgozás. Az elektronikus kérdőívek terjesztése a hálózatokat megelőző időkben floppy lemezekon történt. A módszer népszerűsége abban rejlett, hogy az adatgyűjtés során az adatokat már kitöltéskor lehetett ellenőrizni, illetve begyűjtéskor azok elemzése azonnal, digitalizálás nélkül kezdődhetett meg. A kérdőív fájlok használatának tömeges elterjedését a számítógépes hálózatok megjelenésének köszönhetjük. Még az Internet előtti, lokális számítógép hálózatok korszakának nagy újítása volt, hogy a fájlokat elektronikus levél formájában vagy azok mellékleteként küldték szét. Ezzel végleg sikerült kiiktatni a kérdező és megkérdezett közötti személyes, materiális kapcsolatot. Arról, hogy ez mennyire változtatta meg a kérdőívezés hatékonyságát, Kiesler és Sproull (1986) úttörő munkájából tájékozódhatunk. A szerzők azt vizsgálták meg, hogy a tisztán elektronikus kapcsolatfelvétel hogyan hat a válaszadási hajlandóságra. Azóta is számos olyan publikáció látott napvilágot, amely az egyre újabb fajta elektronikus lekérdézési módszerek hatékonyságát hasonlítja össze a papír alapú kérdőívezéssel. Az Internet megjelenésével szinte egy időben látta meg Kehoe és Pitkow (1996) a webben rejlő új lehetőséget, és alkottak meg egy olyan CGI (Common Gateway Interface) alapú kérdőív alkalmazást, amelyik HTML (Hypertext Markup Language) oldalakon keresztül tette lehetővé az adatgyűjtést. A módszer olcsósága és gyorsasága miatt olyannyira népszerű lett, hogy a szakirodalmat ellepték a „hogyan készítsünk professzionális online kérdőívet” jellegű cikkek és szakkönyvek. Mi sem jelzi ezt jobban, mint Manfreda et al. (2006) cikke, melyben az addig

megjelent szakirodalom meta-analízisét végezték el. A közölt kutatási eredmények összevetésével azt kapták, hogy a webes kérdőívek átlagosan 11%-kal rosszabb válaszadási hatékonysággal bírnak, mint a korábbi elektronikus módokat. Nem csoda tehát, hogy 2008-tól kezdve egyre több olyan kutatással találkozunk, amelyek ennek okát, illetve e kedvezőtlen mutató javításának módszereit tanulmányozzák. Ahogyan a levelező rendszereket felváltotta a mindenhol elérhető webmail szolgáltatás, úgy épültek ki szinte azonnal a böngészőn keresztül menedzselhető webes lekérdezést nyújtó szolgáltatások. A felhő szolgáltatások (Cloud Computing) elterjedésével pedig már az adatokat sem kell letölteni az elemzéshez, ugyanis az ilyen célú szoftvereknek közvetlen hozzáférése lehet a begyűjtött adatokhoz. A vonatkozó szakirodalom központi gyűjtőhelyén, a Web Survey Methodology (www.WebSM.org) oldalon jelenleg 7 056 szakcikket találhatunk, amelyek kizárólag ezzel a szűk szakterülettel foglalkoznak.

ONLINE KÉRDŐÍVEK HATÉKONYSÁGÁNAK JAVÍTÁSA

A webes lekérdezések alacsonyabb válaszadási hajlandóságát bőven kompenzálja az, hogy az Interneten keresztül eleve jóval nagyobb az a válaszadói kör, akit el lehet érni. A különféle kapcsolat felvételi, kérdőív tervezési, motivációs módszer mögött az húzódik meg, hogy a kérdőívezés nemcsak a lekérdezőnek, hanem a kitöltőnek is költségbe kerül. Nemcsak az idejét kell rászánnia, hanem fel kell készülnie arra is, hogy bonyolultabb vagy érzékenyebb kérdések kellemetlen helyzetbe hozhatják, bizalmas információkat adhat ki magáról. De bekerülhet egy közkézen forgó, kérdőívkitöltési-hajlandóságot jelző listára is, és nem utolsó sorban saját eszközeit és erőforrásait kell felhasználnia a kitöltéshez. Mindezzel szemben a lekérdezett számára egy kutató szinte elenyésző hasznot tud felmutatni. A megkérdezett haszna, hogy a kitöltés során részben megismerhet

egy adott témát, jó érzéssel töltheti el a tudomány fejlődéséhez való hozzájárulás, illetve úgy érezheti, ezzel másoknak segít, esetleg valami anyagi díjazáshoz vagy kedvezményhez is juthat, végül pedig okulhat a felmérés alapján készült eredményekből. Fontos tehát, hogy a felmérés során semmiféle technikai hiba ne léphessen fel, és a válaszadónak ne kerüljön extra költségbe a hiba jelzése és kijavítása.

Először is fontos tudni, hogy mely réteg hajlamos online kérdőívezésben részt venni, és melyik kevésbé. Emailés, telefonos, webes, illetve ezek kombinált alkalmazására nyitott csoportok azonosításával próbálkozott meg Smyth et al. (2014). Azt tapasztalták, hogy ha mindenkit a számára megfelelő módon kérdezzük meg, akkor azt a torzítást is kiküszöbölhetjük, hogy például webes kérdőívvel eleve csak azokat érjük el, akik Internet eléréssel rendelkeznek. Márpedig a kutatás egyik megállapítása éppen az volt, hogy a kitöltés legfőbb előrejelző tényezője a felhasznált médiacsatorna megfelelősége. A webes kitöltés általában a fiatalok és magasabb iskolai végzettségűek preferálják, és ez a hajlandóság az internetfüggők között kimondottan magas. A nőket inkább az emailés úton, az idősebbeket pedig inkább személyes megkereséssel lehet elérni. A biztonságot illetően azt tapasztalták, hogy a vírusoktól és adathalásztól való félelem hatása nem kimutatható, és furcsa módon a félelmek inkább az interjúzásra hatnak negatívan.

Másodsorban a kérdőív tartalmi és formai vonzerejének fokozása lehet a keresett megoldás. Mivel a probléma e része közelebb áll a kutatókhoz, így számos „hogyan kérdőívezzünk online” kézikönyv is napvilágot látott, melyek közül Couper (2008), Bethlehem és Biffignandi (2011) valamint Tourangeau et al. (2013) munkáit kell megemlíteni. Az első még kizárólag csak a tartalmi és formai elemekre koncentrált. A másodikban már a tervezési, mintavételezési és súlyozási tudnivalók mellett a hibákkal foglalkozó fejezetben

1. táblázat: Online kérdőív tervezési javaslatok a válaszadási arány javítására

Design elem	Definíció / opciók	Hatás a válaszadási hajlandóságra és minőségre	Javaslatok a kérdőív tervezéshez
Statikus Testreszabás	<i>Meghívólevélben megszólítás:</i> általános, teljes név, csak keresztnév	Pozitív hatás a kérdőív elkezdésére, ha név szerint szólítjuk meg, még inkább, ha csak a keresztnévén.	Ajánlott, ha nem ütközik a sokaság hierarchiájával vagy kulturális háttérével.
Nyeremény	<i>Díjazás felajánlása a kérdőív kitöltéséért:</i> nincs, sokan kapnak keveset, kevesen kapnak sokat	Az egy fő által elérhető nyeremény nagysága pozitív hatással van a kitöltésre. Nincs negatív hatása a válaszok minőségére.	Ajánlott a jutalmakat utólag kisorsolni és kiosztani. Adott költségkeret esetén inkább kevesebb embert díjazunk jobban.
Hét napja	<i>A hét azon napja, amikor az érintett megkapja a felkérést:</i> hétfőtől vasárnapig minden nap külön-külön	A nap megválasztásának nincs hatása a válaszadási hajlamra, de a kitöltés megkezdésének idejére igen. Főleg igaz ez a vasárnap esetére.	A hétköznap kiválasztásának nincs jelentősége. A hétvége a családosok esetén nem hatékony. A később megkezdett kitöltések válaszai jelentősen eltérhetnek az azonnaliaktól.
Napszak	<i>A nap mely időszakában küldjük ki a felkérést:</i> a megkérdezett helyi ideje szerint 6 és 9 óra közötti sávok	Az időszáv megválasztásának nincs hatása a válaszadási hajlamra, de a kitöltés megkezdésének idejére igen. Az este kapott felkéréseket inkább csak másnap teljesítik.	Mivel az időszáv megválasztásának nincs szerepe, ezért tömeges felkéréseket érdemes időben elosztva kiküldeni. Ezzel elkerülhető kitöltéskor a kérdőív szerver túlterhelése, és tömeges kézbesítés miatt a levelek spam-nek minősítése.
Dinamikus Emlékeztetők	<i>A felkérő levél után hányszor küldünk emlékeztető levelet:</i> 1-szer, 2-szer, 3-szor	Minden emlékeztető jelentősen növeli a beérkező válaszok részarányát.	Érdemes rendszeres időközönként emlékeztetőket küldeni, de csak azoknak, akik még nem végeztek a kitöltéssel.
Türelmi idő	<i>A felkérő és emlékeztető levelek közötti idő:</i> 1, 2 vagy 3 hét	Nincs hatással a válaszadási arányra.	Tetszőleges időköz választható, de a rövidebb várakozás felgyorsítja az adatgyűjtést, és csökkenti a korai és késői kitöltésekből fakadó eltérések valószínűségét.
Szóhasználat változtatása	<i>A felkérő és emlékeztető levelek átfogalmazása:</i> új információ közlés nélkül	Jelentősen növeli a válaszadási arányt. Elsősorban a kérdőív kitöltésének megkezdését ösztönzi.	A szöveg átfogalmazása fenntartja a megkérdezett érdeklődését, és hangsúlyozza a témakör fontosságát és jogosultságát.
Értesítési idő változtatása	<i>A felkérő és emlékeztető levelek kiküldési időpontjának módosítása:</i> időszávok a napszakokon belül	Nincs hatása a válaszadási arányra.	Annak ellenére, hogy az emlékeztetők időpontjainak megváltoztatása nincs jelentős hatása az egyes emberekre, a variációjukkal heterogén közönség nagyobb eséllyel érhető el.

Forrás: Sauermaann és Roach, 2013 5. táblázatának tömörítése

megjelennek adatbiztonsági kérdések is. A harmadik könyv mutatja be a témakört leginkább átfogó jelleggel, és a válaszadási arány kérdéskörén túl már nemcsak az online kérdőívek hibáinak, hanem a design helyességének mérésével is foglalkozik. A korábbi kutatások összefoglalójaként itt jelenik meg először az online kérdőívek vizuális és interaktív aspektusa is. Az esztétikai elemek válaszadási hatékonyságra és minőségre gyakorolt hatását elemzi a Casey és Poropat (2014) tanulmány is, melyben arra a következtetésre jutottak, hogy az ergonomikus felület nemcsak növeli a kitöltési hajlandóságot, hanem hozzájárul az adatok minőségéhez is, ami végső soron a megkérdezettek bizalmát növeli. A lekérdezés helyes és biztonságos működéséhez alapvető követelmény a bizalom elnyerése, ami kulcsfontosságú tényezője a válaszadási hajlamnak.

Az eddig bemutatott, és az irodalomban fellelhető eredményekkel azonban óvatosan kell bánni! A kísérletek alanyait ugyanis a legtöbbször a kutatók számára könnyen és olcsón elérhető egyetemi hallgatóság képezi, ráadásul mindez csak az USA felsőoktatására leszűkítve. Az például, hogy ezek a hallgatók a hét melyik napjának melyik időszakában hajlandóak leginkább webes kérdőíveket kitölteni, vajmi keveset árul el arról, hogy egy hazai kisvállalkozás vezetőjéhez milyen időpontban érdemes a kérdőívet eljuttatni. Mindezek ellenére hasznos, ha ismerjük az online kérdőívek hatékonyságának azon kulcstényezőit, melyeket kimondottan *innovációs* témájú felmérések tapasztalatai alapján állítottak össze Sauermann és Roach (2013). Már az 1. táblázatban is találkozhatunk olyan biztonságtechnikai elemekkel, amelyek a szerverek túlterhelésére, illetve a meghívó levelek kéretlen reklámszemétnek minősítésére hívják fel a figyelmet. A táblázatban található javaslatokat a TÁMOP-4.2.2-A versenyképességi kutatás során is figyelembe vettük. Az adatgyűjtés egy jó részét eredetileg önkitöltős online kérdőívekkel tervezték megvalósítani, amiről hamar

kiderült, hogy mérete és komplexitása miatt sem lehet működőképes. A problémát kérézőbiztosok segítségével lehetett áthidalni, mert a pályázat költségvetése nem tette lehetővé a megfelelő mértékű nyereségyel való kitűzését.

Az 1. táblázat segítségével megpróbálhatjuk javítani a válaszadási arányt, de ezt a kis növekedést szinte romba dönti az úgynevezett „kérdőív spammelés” jelensége Rogelberg és Stanton (2007). A módszer alacsony költsége miatt egyre több hivatalos szervezet el ezzel az adatgyűjtési formával (pl. elektronikus népszámlálás). A webes lekérdezést nyújtó szolgáltatásokon keresztül pedig boldog-boldogtalan képes kérdőívvel bombázni a világot. Nemcsak az egyetemi szakdolgozatok egyre nagyobb hányada alkalmazza ezt a technikát, de a jelen folyóirat marketingkutatással foglalkozó cikkei közül több is ilyen eszközök felhasználásával készült. Ami pedig az informatikában tömegessé válik, az potenciális célpontja a visszaélési lehetőségeknek. Nem ritka az adathalászati célú, látszat-kérdőívek alkalmazása sem. Ilyen veszélyekkel a kérdőíves kutatások során kevesen számolnak. A TÁMOP-4.2.2-A adatgyűjtés tervezésekor a kutatók külön figyelmet szenteltek a kérdőív minimalizálásának, de a kérézőbiztosoknak igencsak meggyűlt a bajuk azzal, hogy korábban más kutatások már ugyanezen célközönséget „lebombázták” kérdőíveikkel. A TÁMOP-4.2.2-A versenyképességi felmérés abban is különbözik a szakirodalomban vizsgáltaktól, hogy a megkérdezettek nem egyének, hanem vállalkozások voltak. Egy komolyabb cég informatikai infrastruktúrája pedig általában el van látva tűzfal és spam-szűrő védelemmel, melyeken könnyedén felakadhat a meghívó levél vagy elérhetetlennek bizonyul a külső kérdőív weboldala.

ONLINE KÉRDŐÍVKÉSZÍTÉS, MINT ALKALMAZÁSFEJLESZTÉS

Az online kérdőívvezetés egy speciális szoftverfejlesztési műfaj. A kutatónak

lényegében a felmérés adatait tartalmazó adatbázist és annak felhasználói felületét kell informatikai eszközökkel megterveznie. Ha a lekérdező rendszer üzemeltetését a szolgáltató magára is vállalja, az esetleges lekérdezés közbeni változtatásokat a kutatónak magának kell megoldania. Nem túlzás tehát, ha azt állítjuk, hogy egy online lekérdezés összeállítása, futtatása és karbantartása nem sokban tér el egy szoftveres alkalmazás tervezésétől, fejlesztésétől és üzemeltetésétől. Legfeljebb a funkcionalitása szűkebb területet ölel fel, és ezért némiképp egyszerűbb. Azonban a kérdőív-alkalmazásban jártas emberek sem mindig ismerik a szoftverek fejlesztésére vonatkozó alapvető szabványokat és ajánlásokat. Sokszor még a webes lekérdezések nyújtotta lehetőségekkel sincsenek tisztában, nem-hogy azok kockázatát fel tudják mérni.

Jó példája ennek az a jelentéktelennek látszó eset, amikor a KSH 2011-es népszámlálási kérdőív sűgőjába vélhetően az egyik fejlesztő csempészte bele a „*Végtelen szerelemmel Audrey-nak :) (és a családomnak)*” szövegű névjegyét, amit szoftverek esetében „húsvéti tojásnak” (Easter Egg) neveznek. Az üzenetet az *enepszamlalas.hu* angol verziójából kimaradt, csak a magyar változatban volt fellelhető, ha a „Milyen fogyatékosága van Önnek?” kérdés válaszopciói alatt elhelyezett láthatatlan szövegrész kijelölésre került.

Az online kérdőív- és az alkalmazásfejlesztés közötti párhuzamot tovább erősíti, ha olyan adatgyűjtéssel állunk szemben, amit vagy más közönség felé és/vagy eltérő időben is meg kell ismételni. Ezt a feltételt már a tervezés kezdetén figyelembe kell venni. A kérdőívet úgy kell megtervezni, hogy annak kérdései és mérési skálái időtállóak legyenek, a különböző adatfelvételi panelek illeszthetőek legyenek egymáshoz. Az újrafelhasználhatóság szempontjából különösen problémás a kérdőív struktúrájának megváltoztatása, illetve az egyes adatsorokhoz tartozó kitöltők adatainak tárolása. A kérdőív változatok menedzselésekor nemcsak a felület külön-

böző verzióinak kezelésére kell gondolni, hanem a felülethez tartozó adattípusokra, adatszerkezetre és azok kódolásaira is. Egy jelentéktelennek tűnő felületi elem megváltoztatása (pl. egy kérdés megválaszolásának kötelezővé tétele) is okozhat komoly inkompatibilitási problémákat az adatgyűjtések összesítésekor. A professzionális kérdőívrendszerek esetében a fejlesztői munkát olyan technológia is támogatja, amelyik attól függően engedélyez vagy korlátoz bizonyos változtatásokat, hogy a kérdőív fejlesztési, tesztelési vagy éles üzemban működik. Éles rendszerek esetében például egy kérdés törlése már tiltott művelet, amit legfeljebb a kérdés láthatatlanná tételével lehet megkerülni. Kisebb (például helyesírás) hibákat éles rendszerben is kell tudni javítani. De ha ezzel visszaélve jelentősen megváltoztatjuk (például felcseréljük) a válaszopciók címkéit, akkor olyan inkompatibilitást idézhetünk elő, ami sajnos akár az elemzést is lehetetlenné és értelmetlenné teszi. Mivel erre a problémára tökéletes szoftveres megoldás nem készíthető, ezért nagyobb méretű kérdőívek fejlesztése során is érdemes betartani a szoftverfejlesztés alapvető (követelményelemzés – tervezés – megvalósítás – ellenőrzés – ölesítés – üzemeltetés) lépéseit, ahol az egyes fázisok pontos dokumentálása és mérföldköszertű

„Nem túlzás tehát, ha azt állítjuk, hogy egy online lekérdezés összeállítása, futtatása és karbantartása nem sokban tér el egy szoftveres alkalmazás tervezésétől, fejlesztésétől és üzemeltetésétől. Legfeljebb a funkcionalitása szűkebb területet ölel fel, és ezért némiképp egyszerűbb. Azonban a kérdőív-alkalmazásban jártas emberek sem mindig ismerik a szoftverek fejlesztésére vonatkozó alapvető szabványokat és ajánlásokat. Sokszor még a webes lekérdezések nyújtotta lehetőségekkel sincsenek tisztában, nem-hogy azok kockázatát fel tudják mérni.”

lezárása biztosítja azt, hogy az éles rendszer már lehetőleg hibamentes legyen.

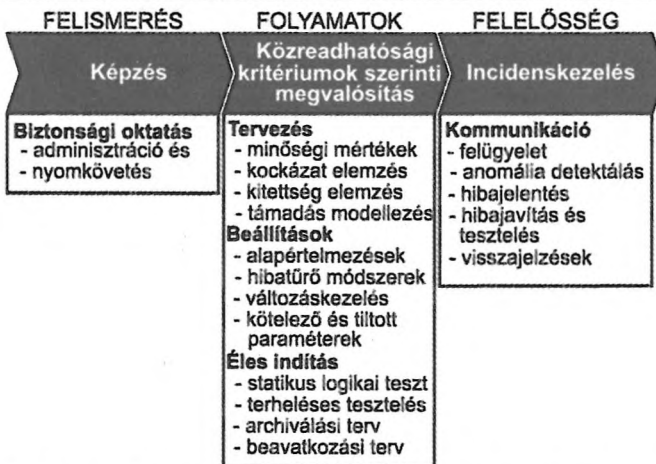
AZ ONLINE KÉRDŐÍVFEJLESZTÉS ÉLETCIKLUSA

Mivel kérdőíves adatgyűjtés biztonsági kérdéseivel foglalkozó publikáció még nem jelent meg, ezért kiindulási pontként érdemes valamilyen, a szoftverfejlesztésben már elfogadott és ismert szabványt követni. Ennek megfelelően elkészítettük az online kérdőíves folyamatának azt a biztonsági modelljét, amelyik a Microsoft által kifejlesztett SD3+C elnevezésű módszertant követi. Az SD3+C rövidítés a biztonságos tervezés, biztonságos alapértelmezett konfiguráció, biztonságos fejlesztés és kommunikáció angol (Secure by Design, Secure by Default, Secure in Deployment, and Communications) kifejezés kezdőbetűiből áll össze. Egy online kérdőív biztonsági szintjét elsősorban a tervezéskor és az elkészítésekor elkövethető hibák határozzák meg. A hibák kiküszöbölése és megelőzése mellett azonban ügyelni kell a fenyegetettségekre és fel kell mérni a támadási lehetőségeket is. Minden apró részletre azonban túl költséges és lehetetlen is lenne ügyelni, másrészt az adatgyűjtés során az informatikai architektúra jelentős része

nem is a kutatók kontrollja alatt áll. Annak érdekében, hogy az időközben előforduló, nem várt esetek minél kevesebb gondot okozzanak, a kivételek és a ritkán használt funkciók alapértelmezett beállításait úgy érdemes megválasztani, hogy azok áttekinthetőek legyenek, és ne adjanak módot a visszaélésekre. Az adatgyűjtés során a fő hangsúlyt a rendszer támogatása és felügyelete, illetve a kitöltők segítése és kérdéseik megválaszolása kapja. Az éles kérdőíven végzett módosítások veszélyeiről már volt szó. Végül a kérdőív fejlesztőinek képesnek kell lenniük a támadások felismerésére, a károk minimalizálására, de egyúttal helyesen kell kommunikálniuk a kitöltőkkel a hibák beazonosítása, reprodukálhatósága és kijavítása során. A jó online lekérdezés tehát mindig csapatmunka, ahol a szakterület kutatóinak nemcsak informatikusokkal, formatervezőkkel, adatelemzőkkel, módszertani és jogi szakemberekkel kell együtt dolgozniuk, hanem a kitöltőkkel is. A kitöltő ugyanis nem más, mint az elkészült „kérdőív-felület” kulcsfigurája: a végfelhasználó.

Mielőtt az SD3+C modell fázisain végighaladva bemutatjuk annak tipikus online kérdőív fejlesztési elemeit, fel kell hívunk a figyelmet a tervezés teljességére. Egy

1. ábra: Egyszerűsített SD3+C rendszerfejlesztési modell elemei



Forrás: Microsoft SD3+C modelljének implementálása Lipner (2004) alapján

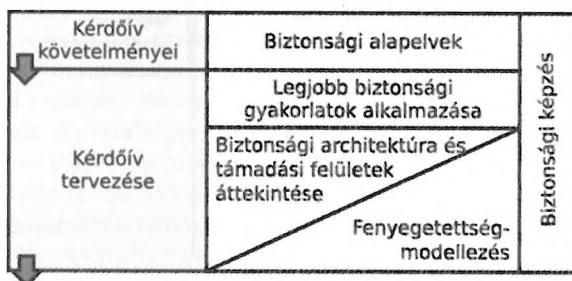
adatgyűjtést ugyanis akkor lehet élesben elindítani, ha már tényleg csak az adataink hiányoznak. Rendelkezniük kell tehát kutatási feltételezésekkel, megfelelő adat-elemzési módszertannal és az eredmények ábrázolási technikájával is. E nélkül ugyanis az adatgyűjtés gyakran túl széles spektrumú lesz (valamelyik adat majd csak jó lesz valamire alapon), és ez a kitöltési hajlandóságra is negatív hatással van. A jó kérdőívnek szakértelmet kell tükröznie, ráadásul mind-azt a kitöltő számára érthető és érzékelhető formában! Adatbiztonsági megfontolásból az is ajánlott, hogy már a kezdetektől foglalkozzunk a begyűjtött adatok minőségének mérhetőségével. Már tervezéskor meg kell határozni, hogyan fogjuk megkülönböztetni a valós és fals kitöltéseket, valamint miként ismerhetőek fel és kezelendők a válaszopciók közötti inkonzisztenciák. Ehhez manapság már nemcsak magukat a statikus adatokat tudjuk felhasználni, hanem a kitöltési dinamika (például kezdési és felhasználási idő, kérdésenkénti várakozás stb.) is a segítségünkre lehet.

KÉRDŐÍV KÖVETELMÉNYEK ELEMZÉSE

Egy online kérdőív követelményeit a kutatási célok, az adatgyűjtésben érintettek, és az informatikai, módszertani és jogi szabványok határozzák meg. Először a kutatási cél eléréséhez szükséges adatok körét, annak kezelési és feldolgozási módját kell rögzíteni. Az adatvédelmi törvény értelmében háromféle adatkört kell elkülöníteni: különleges (pl. vallási, pártállási, faji

eredet, egészségi állapot, szexuális élet stb.) adatok, személyes (pl. név, cím, azonosító stb.) adatok és egyéb adatok. Különleges adatok csak akkor gyűjthetőek, ha annak kezeléséhez az érintettek írásban hozzájárultak. Személyes adatok gyűjtéséhez elegendő az érintett hozzájárulása, vagy ha az adatgyűjtést törvény rendeli el. Kutatási célú online kérdőívezés nem lehet kötelező jellegű, ezért a résztvevők részéről elegendő a szóbeli, vagy ráutaló magatartással történő hozzájárulás. Azonban az így megszerzett hozzájárulás is csak akkor érvényes, ha az megfelelő tájékoztatáson (pl. adatvédelmi nyilatkozat, meghívó levél stb.) alapul. Az alkalmazott eljárástól függetlenül az adatokon végzett bármely művelet (gyűjtés, felvétel, rögzítés, rendszerezés, tárolás, megváltoztatás, felhasználás, továbbítás, összekapcsolás, zárolás, törlés, megsemmisítés, nyilvánosságra hozatal) adatkezelésnek számít. Online kérdőívezés esetén az adatkezelő minden esetben a kutató, hiszen ő határozza meg az adatkezelésének célját, és olyan döntéseket hoz, melyeket vagy önmaga hajt végre, vagy az általa megbízott adatfeldolgozóval hajtja végre. Személyes adatok gyűjtése esetén teljesülni kell a minimalitásnak és a célhoz kötöttségnek, azaz csak a kutatási cél elérésére alkalmas és ahhoz elengedhetetlen adatokat szabad begyűjteni és az adatkezelés során feldolgozni. Fontos tudni azt is, hogy fizikai, fiziológiai, gazdasági, kulturális vagy szociális azonosságára jellemző ismeret is személyes adatnak minősül mindaddig, amíg abból következtetni lehet

2. ábra: Online kérdőívkészítés követelmény elemzés fázisa



a kitöltő személyére. A törvényi szabályozások pontos ismerete különösen paneladatok felvétele esetén ajánlott.

Az adatvédelmi kérdések áttekintése után kerülhet sor az adatbiztonsági követelmények tisztázására, és az elvárások rögzítésére. Az adatok tisztességes hasznosítása és az adatközlők magánszférájának tiszteletben tartása az adatkezelő felelőssége. Ehhez meg kell ismernie és alkalmaznia azokat a technikai, technológiai, szervezési és minőségbiztosítási módszereket és eljárásokat, amelyekkel mindez megvalósítható. Elsősorban az adatok fizikai védelmét és a hálózaton való biztonságos továbbítását kell megoldani, valamint többszintű jogosultsági rendszer kialakításával kell az online kérdőívezésben résztvevők (fejlesztő, adminisztrátor, elemző kitöltő stb.) számára a hozzáférést szabályozni. A kérdőív kutatói oldalán állók számára egy jelszavas beléptetés után már kellő részletezettségű szerepköri jogosultságok oszthatóak ki, de a kitöltőknél a helyzet már nem ennyire egyszerű, különösen anonim felhasználók esetében.

A névtelenséget kétféle értelemben használjuk, így beszélhetünk technikai és tartalmi anonimitásról. Tartalmi anonimitás alatt azt értjük, hogy a begyűjtött személyes adatokból ne lehessen a kitöltő személyét kikövetkeztetni. Ilyen kérdésekkel a statisztikai adatbiztonság témaköre foglalkozik. A technikai anonimitás azt vizsgálja, hogy az adatgyűjtés folyamatából be lehet-e azonosítani a kitöltő személyét vagy sem. A két fogalom egymástól független, amire egyszerű példa az, ha egy teljesen semleges kérdőív kitöltéséhez olyan ütemezésben küldünk ki meghívó leveleket (például naponta egyet), hogy a kitöltés megkezdésének idejéből nagyjából be lehessen azonosítani ki is volt az illető. Az adatvédelmi törvény következtében a kettő közül a tartalmi anonimitás a könnyebben kezelhető. Ráadásul a törvény és a kérdőív ismeretében a kitöltőt nem éri bizalomvesztés, és szabadon döntheti el, hogy eleget tesz-e a felkérésnek vagy sem. Arról

viszont már sokkal nehezebb meggyőzni a kitöltőket, hogy a kérdőív technikailag is névtelenített. Honnan tudhatná, hogy a számára megadott linkben nincs-e kódolva valahogy, hogy az csak neki szól. Ha pedig a kitöltéshez előbb egy informatikai rendszerbe be is kell jelentkeznie (és ez lehet akár egy proxy szerver beállítás a böngészőben), akkor szinte elkerülhetetlen a bizalomvesztés. Érdemes megemlíteni, hogy a technikai anonimitás nemcsak a webes kérdőívezést érinti. A különféle módon kiküldött állományok szerkesztésekor a szoftverek általában bejegyzik az utolsó módosítás ideje mellett az utóljára módosító felhasználó szoftver-regisztrációs adatait is.

A technikailag valóban névtelen kérdőívezés megvalósítása ugyan nagyon fontos eleme a kitöltők bizalmának elnyeréséhez, és ezáltal az adatminőség javításához, de komoly buktatót is rejt magában. Ha ugyanis nem tudható, hogy egy adott adatsort ki rögzített, akkor az sem akadályozható meg, hogy az illető ne töltsön ki egynél többet. A többszörös kitöltés megakadályozásának technikájára még visszatérünk, de azt világosan látni kell, hogy a kitöltői spammelés megakadályozása és az anonimitás biztosítása csak egymás rovására biztosítható. A duplikált kitöltés pedig sokkal gyakoribb, mint ahogy azt a kutatók képzelik. Duplikált kitöltést eredményezhet, ha a kitöltést meg kell szakítani (pl. az adat előkeresése miatt), vagy a kapcsolat szakad meg vagy lejár az időkorlát. De nem ritka a kimondottan rosszindulatú, (a sokadik kérdőív felkérés után már csak bosszúból is elkövetelt) tömeges kitöltés sem. Ráadásul a szabványos online kérdőív rendszerek elterjedésével (pl. SurveyMonkey, GoogleForms stb.) könnyen készíthetőek olyan programok, amelyek a kérdőív keretrendszerének és magának a kérdőívnek az ismeretében automatikusan képesek tetszőleges mennyiségű adatot rögzíteni.

Ez a terület az emberi és a gépi intelligencia között is különbséget tenni képes eszközök alkalmazási területére vezet minket. A tömeges automatikus kitöltést

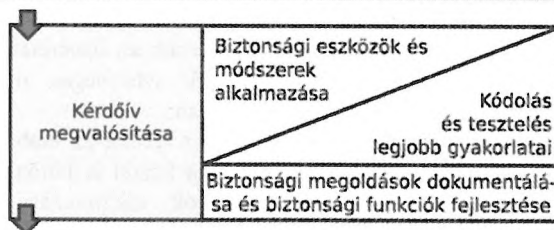
leggyakrabban úgynevezett Captcha tesztekkel lehet megakadályozni, ahol is a kitöltés megkezdése előtt olyan kérdésre kell válaszolni, amire egy átlagos intelligenciájú ember is képes, de egy gép nem. A kitöltés elé egy plusz akadály beépítése viszont sokkal több kárt okozhat, ha ettől még alacsonyabb lesz a kitöltési kedv, és arról se feledkezzünk meg, hogy az éppen megjelenített kódok újabb kiskaput nyitnak meg a technikai anonimitás megkerülésére.

A követelmény-meghatározáshoz nemcsak a kérdőív tartalmi részére vonatkozó megkötések tartoznak, hanem a lebonyolítási folyamat elemeit érintő olyan kérdések is, amelyek közül már többet a bevezetőben megismertettünk. Ide tartozik tehát a kérdőív technológiai hátterének kiválasztása, a kiértékelés rendjének és a lebonyolítás ütemezésének idő- és költségtervezése, illetve a célközönség meghatározása és annak mintavételezési módszere. Az informatikai biztonság és adatvédelem hagyományos eszköztárából pedig érdemes felhasználni egyrészt a védendő szintek elhatárolásának elvét, miszerint megkülönböztetjük az adatréteget, az (alkalmazói és rendszer-) szoftver réteget, a hardver és hálózati réteget, a fizikai környezet réteget valamint a felhasználói réteget. Másrészt az információs architektúra és a kommunikációs csatornák e rétegeinek úgynevezett CIA rendelkezésre állási (Availability), sértetlenségi (Integrity), bizalmassági (Confidentiality) tulajdonságait, hogy ennek megfelelően teljes körű, zárt, kockázattal arányos megelőző intézkedéseket tehesünk. A legjobb vezérlő elv ebben a fázisban a célhoz kötöttség szem előtt tartása.

KÉRDŐÍV TERVEZÉS ÉS MEGVALÓSÍTÁS

Egy online kérdőív tervezésekor három nagyobb folyamatot kell informatikai eszközökkel megvalósítani: a kiértékelési, a kitöltési és a nyugtázási rendet. A kiértékelési rend során biztosítani kell, hogy a kérdőív időbeli és címbeli (URL) elérhetőségéről csak azok értesüljenek, akiket a mintavétel során kiválasztottunk. Még névtelen kitöltés esetén is, ugyan korlátozottan, de van rá mód, hogy a későbbiekben követni lehessen a kiértékelések eredményességét. Ez különösen akkor fontos, ha azt valami oknál fogva meg kell ismételni, hiszen akkor már csak azoknak kell emlékeztetőt küldeni, akik még nem töltötték ki azt. A személyre szóló meghívókon kívül általában van módunk kitöltési kvóták megadására. Ez azt jelenti, hogy például egy-egy szervezethez hozzárendelünk pár kitöltési lehetőséget, de azt már nem osztjuk tovább név szerint, hanem a szervezetre bizzuk azt. Sok múlik a kiértékelés szövegének megfogalmazásán is. Egy ilyen értesítő levél kötelező elemei értelemszerűen az adatgyűjtés neve, a kérdőív elérhetőségi címe és ideje, az adatgyűjtés célja, az adatok felhasználója, illetve az erre való felhatalmazása, és végül az anonimitása. Ezen kívül ajánlott még közölni a kitöltéshez szükséges becsült időt (vagy a kérdőív terjedelmét), a kitöltéshez szükséges előkészületeket, a felmérés eredményeinek várható megjelenési helyét és idejét, egy elakadás esetén elérhető segítséget, illetve ha az adatgyűjtéshez bármiféle nyeremény tartozik, akkor annak részleteit. Tapasztalataink alapján érdemes még arra felhívni a figyelmet, hogy miként ellenőrizhetik a

3. ábra: Online kérdőív készítés tervezés és megvalósítás fázisa



kitöltésük sikerességét, vagy miként igazolhatják, hogy a kitöltési felkérésnek eleget tettek. Továbbá, ha a kérdőív túl nagy terjedelmű, akkor miként tudják felfüggeszteni és később folytatni a megkezdett kitöltésüket. Végül, mivel kommunikációs csatornáról van szó, nem hagyhatjuk figyelmen kívül annak steganográfiai felhasználását sem. A steganográfia a kommunikáció folyamatába rejtett, észrevétlen üzenetek küldésének tudománya. A bevezetőben már említettük, hogy a kérdőív URL kódjába rejtett kódokkal miként lehet tudomást szerezni még anonimnak mondott kérdőívek esetén is a kitöltő személyéről. De hasonló azonosító információkat rejthetünk el a szöveg formázásában vagy formázatlan levelek esetén a szóközök ismételt elhelyezésével. Érdeemes megjegyezni azt is, hogy léteznek különféle technikák arra nézve is, hogy névtelen adatgyűjtéshez rendeljünk hozzá nyereményjátékot, ami első hallásra egy kicsit talán furcsán hangzik, de kivitelezhető.

Nyugtázás alatt a kérdőív kitöltésének befejezésével induló folyamatot értjük. Ez a leggyakrabban egy egyszerű „köszönjük” üzenet, de lehetőséget adhatunk a kitöltés elmentésére vagy akár kinyomtatására is. Egy ilyen nyugtázó oldalon vagy levélben további információkat érdemes közölni az adott témáról vagy a kutatást végző intézményről. Ezt az utolsó oldalt érdemes olyanra tervezni, mintha jutalom vagy bónusz lenne. A legtöbb online kérdőív támogatja az összesített adatok azonnali megjelenítését is. Ez lehet egyszerűen csak annyi, hogy az illető hányadik a kitöltők sorában, de ugyanennyi erőfeszítéssel néhány vagy akár az összes kérdésre befutott gyors-statisztikákat is megjeleníthetjük rajta. Még ennél is sokkal többet mondanak azok az ábrák és jelentések, ahol a kitöltő a többi kitöltésekhez képest kap információt a válaszáiról. Mindezek természetesen csak adott kitöltésszám fölött értelmezhetőek, de önkéntes kitöltés esetén javítják az öszinte válaszok arányát. Az azonnali statisztikák alkalmazása esetén ajánlott „csak az eredmény érdekel” válaszopciókat is alkalmazni.

A kérdőívtervezést a szoftverfejlesztésből átvett analógia alapján négy részfolyamatra bonthatjuk: (1) az adat, (2) a számítás, (3) az interakció és (4) a felület tervezésére. Az adatgyűjtés feldolgozásának modellje a kérdőíven szereplő input mezőkből állítja elő az elemzéshez szükséges számítások bemenetét, és egyben meghatározza az elemzés módszertanát is. Fontos tehát, hogy először legyen ismert az elemzés modellje, és csak azután fordítsuk le azt kérdésekre, és ne fordítva. Az online kérdőívek esetén ez mindössze annyiban különbözik a hagyományos kérdőívektől, hogy a számítások adatgyűjtés közben, és akár a kitöltők számára is elérhetőek. A kérdések típusára, a szövegének helyes megfogalmazására, a válaszopciók teljességére és át nem fedésére vonatkozó ismereteinket is bátran használhatjuk, hiszen ezek mind javítják az adatminőséget. Az interakció- és a felülettervezés azonban már igazi online specialitás. A tervezés egyik alapeleme az úgynevezett kérdőív logika, ami az egyes kérdések láthatóságát, a közöttük lévő függőségi (feltételes elágazási) feltételeket, és az adatelemek adattípusainak és helyesség-ellenőrzési módszereinek, ezen belül is a kötelező kitöltés előírásának meghatározását jelenti. Adatintegritási feltételeket megadhatunk az adatelem (válaszopció) szintjén, válaszopciók között (például legalább/pontosan n opció bejelölése), de akár különböző kérdések között is (például 18 év alattiaknál nem kérünk havi jövedelmet). A kitöltési logika kialakításakor ajánlott a „minél kevesebb megkötés” elvét alkalmazni, és a megkötések nélküli állapotból kiindulva meghatározni az ideális állapotot. Kötelező válaszokat csak hibátlan kérdőíven alkalmazzunk, és ott is csak éppen annyit, amennyi feltétlenül szükséges. Ellenkező esetben könnyen rákényszeríthetjük a kitöltőt egy nem valódi válasz megadására, vagy ami még rosszabb, blokkolhatjuk a további kérdésekre adható válaszait. Ha a feltételes elágazások száma nagy, akkor egyrészt a kérdőív nehezen lesz áttekinthető, még nehezebben elemezhető,

de még ennél is nehezebben tesztelhető. Érdemes arra is felhívni a figyelmet, hogy a feltételes ugrások programozásakor a papír alapú tervezésnél általában a *válaszopció végénél* szokás meghatározni, hogy akkor ugrás erre és erre a kérdésre, míg a számítógépes logika szerint mindig azt kell megadni egy *kérdés elején*, hogy azt mikor érhető el. Bár a két leírás egyforma kifejező erejű, a lekódolt elágazások olvasása és értelmezése mind a tervezők, mind pedig a tesztelők számára gondot okozhat. A minimális megszorítások elve lehetővé teszi továbbá azt is, hogy a válaszopciókat, vagy akár a kérdéseket is véletlen sorrendben tárjuk a kitöltő elé. Nagyobb terjedelmű kérdőívek esetén ez igen hasznos eszköz ahhoz, hogy a kérdőív végére tervezett kérdések is előre kerülhessenek. Ezzel pedig elérkeztünk a navigáció kérdéséhez. Egyszerűbb esetben a kérdésblokkonkénti, csak előre haladó kitöltést alkalmazunk, de engedélyezhetjük a visszafele lapozást is. Azt is külön szabályozhatjuk, hogy a még meg nem válaszolt kérdéseket is megtekinthesse a kitöltő, vagy csak azokat, amiket már kitöltött, vagy ami éppen soron következik. Ezekkel akár kérdésenkénti kitöltés is megvalósítható. A navigáció további fontos elemei az állapotjelzés, hogy a kitöltés hány százalékával vagyunk kész, illetve a tartalomjegyzék, ami megmutatja a kérdőív blokkjainak struktúráját. Legvégül pedig egy olyan, kizárólag webes lekérdezéseknél alkalmazható funkcióra szeretnénk felhívni a figyelmet, amivel mérhetővé, sőt korlátozhatóvá tehető egyes kérdések megválaszolására felhasználható, illetve a kérdések közötti várakozási idő is. Ezen adatok felhasználása már több esetben is nagy segítségünkre volt a komolytalan kitöltések kiszűrésében.

A felület tervezés során a kérdőív megjelenésének és külalakjának elemeit határozzuk meg. A kérdőív arculatának fontosságáról már a bevezetőben Casey és Propat (2014) eredményeit bemutatva szoltunk pár szót. Ehhez képest azt fontos kiemelni, hogy az internetes technológia

miatt a tervezésekor nem tudhatjuk, hogy azt milyen eszközön (számítógép vagy mobilkészülék), milyen felbontásban, milyen operációs rendszer és böngésző alkalmazáson keresztül, illetve ezek milyen beállításait használva fogják elérni. Nem ajánlott tehát semmi olyan tervezési módszert alkalmazni, ami ezekben bármiféle kötöttséget is jelent, vagy a megtekintéséhez vagy lejátszásához speciális eszköz (szoftver vagy beépülő plug-in) szükséges. A mobil eszközök elterjedése miatt a modernebb kérdőív rendszerek már képesek a billentyűzet nélküli kitöltési üzemmódot is támogatni. Az arculati elemeket is úgy kell kialakítani, hogy azok méreteit és egymáshoz való elhelyezéseit ne abszolút értékekben, hanem relatívan határozzuk meg. Az adatelemekre vonatkozó megkötésekre érdemes vizuálisan is felhívni a figyelmet. A kötelező kitöltést például legtöbbször piros csillaggal szokás jelölni, szükség esetén megadhatunk beviteli maszkokat is. Végül arra kell felhívni a figyelmet, hogy a különböző válaszadási módszerekhez érdemes saját, lehetőleg szabványos megjelenítést rendelni (például egyszeres választást rádiógombbal, többszörös választás jelölőnégyzettel), illetve egy adott típusú kérdésnek több vizuális megjelenítése is lehetséges. Egyszeres választás például megjeleníthetünk rádiógombokkal, lenyíló listákkal vagy akár fogd és vidd technikával áthelyezhetjük a kívánt választ a megjelöltek közé. Sőt vannak olyan kérdéstípusok, mint például a sorrendezés, vagy a prioritás meghatározás, ami igazán jól csak szoftveres támogatással valósítható meg. Ezek a beviteli módszerek azonban csak akkor javítják az adatminőséget, ha nem ismeretlenek a kitöltők számára és ezáltal elősegítik a pontosabb vagy gyorsabb válaszadást.

A TÁMOP-4.2.2-A versenyképességi kutatáshoz kapcsolódó kérdőív egyik legfontosabb tervezési alapelve a komplexitás elfogadható szinten tartása volt. A tervezett index 10 összetevő pilléréhez (Humán tőke, Finanszírozás, Együttműködés, Termék, Adminisztratív rutinok, Verseny és stratégia, Technológia és IKT,

Marketing, Nemzetköziesedés és Online jelenlét) már csak témakörönként pár kérdéssel is könnyen elérhető az a limit, ami már az adatgyűjtés minőségének és mennyiségének a rovására mehet. A kérdőív bonyolultságának mérésekor is megkülönböztetünk technikai, illetve tartalmi összetevőket. A technikai összetevők közül a kérdések és válaszopciók számát kell megvizsgálni. A versenyképességi kérdőív 15 kérdésblokkjába összesen 236 kérdés került megfogalmazásra, amelyek válaszopcióihoz 1202 beviteli mezőhöz tartozott. A kérdőív terjedelméből fakadó komplexitást a kutatás tervezői többféleképpen is megkísérelték csökkenteni. A technikai eszközök közül az elágazási feltételek beiktatásával lehetett az egyes felhasználók számára szűkíteni a megjelenítendő tartalmat, bár az elhelyezett 112 ugrási feltétel nem könnyítette meg a későbbi adatelemzést. Ilyen méretű kérdőív esetén a kötelezően kitöltendő elemek számát és elhelyezkedését is jól meg kell fontolni. A kérdőívbe mindössze 6 kötelezően kitöltendő kérdés került, és ezek is mind a legelejére, mintegy szűrőként alkalmazva, hogy aki ezen a részen már túlesett, azok minden további adata relevánsnak tekinthető a kutatás céljából. A komplexitás tartalmi csökkentéséhez Schaeffer és Dykema (2011) összefoglaló munkája szolgált útmutatóul, ahol a szerzők a kérdések karakterisztikájával foglalkozó részben több megelőző kérdőív komplexitási vizsgálat eredményét foglalja össze. Szerintük a tartalmi komplexitás úgy csökkenthető, ha az alábbi mutatókat egyidejűleg próbáljuk meg optimalizálni:

- kérdéshossz: összesen szavakban mérve, illetve kérdésenként átlagosan,
- kérdésenkénti beviteli mezők átlaga: átlagos válaszopciók száma,
- kérdések szövegének összetettsége: tagadás, egymásba ágyazott mondatok, utalások,
- olvasásértési szint: milyen ismeretek, szakmai végzettség kell a szöveg megértéhez,
- kiegészítő magyarázatok száma: értelmességi pontosság, segítség,

- egyértelmű megfogalmazás: félreérthető, pontatlan kifejezések száma,
- töltelék kifejezések száma: bevezető, összekötő szövege, stop szavak,
- felidézési pontosság: tény-vélemény kérdések aránya,
- szabad és kötött válaszadási módok aránya.

A kérdőív szövegének további tömörítési lehetősége, ha kevés fajta kérdéstípust alkalmazunk, és azon belül is a válaszopciók feliratai (skálái) több kérdésnél is felhasználásra kerülnek. Ahhoz, hogy a versenyképességi kérdőív összes fenti komplexitási mutatóját meghatározzuk, sok helyen szövegbányászati módszerek felhasználásával is csak közelítő értékeket kapnánk. A tervezési folyamat eredményének hatékonyságát a 2. táblázatban foglaltuk össze.

A kérdőív ismertetőjébe az került bele, hogy a kitöltésére kb. 45 percet elegendő fordítani. Elágazási feltételek, és az adatbeviteli idő figyelembe vétele nélkül ez 190,64 szó/perc olvasási sebességet feltételez a kitöltők részéről. Bár ez a 45 perces érték a kérdőív tesztelése során került meghatározásra, ha figyelembe vesszük, hogy az átlagos olvasási sebesség 220 szó/perc, illetve az elágazási feltételek miatt nem is kell mindig mindenkinek mindent kitölteni és elolvasni sem, reális időtartambecslésnek bizonyult. Igaz csak akkor, ha valaki valóban a felkérésnek megfelelően előre felkészült, és az adatokat nem kitöltés közben kellett előkeresnie. A kitöltési idők átlaga a kérdőív lezárásakor 52 perc volt, de az előzőek miatt igen nagy, 33 perces szórással, az irreálisan rövid illetve extrém hosszú idők figyelmen kívül hagyása mellett is.

KÉRDŐÍV ELLENŐRZÉS ÉS ÉLESÍTÉS

A kérdőívek szokásos tesztelése elsősorban a kérdések nyelvtani helyességének, érthetőségének, teljességének és ellentmondás-mentességének ellenőrzését foglalja magában. Noha ide tartozik az is, hogy a tesztelés során begyűjtött adatokra alkalmazzuk a kutatás adatelemzési módszereit is, ez sajnos a legtöbb esetben

el szokott maradni. A versenyképességi kérdőív esetén is csak az adatok export/importját valamint az adattranszformációs és aggregációs algoritmusokat tesztelték. Informatikai rendszerek esetében a megfelelő minőségi és biztonsági szintet elérendő többféle tesztelési módszert alkalmaztunk. Statikus elemzés során átvizsgáljuk a kérdőív kódját, amit megtehetünk fejlesztői csoporton belül, intézményen belül, külső technikai szakértővel, vagy szoftveres helyességbizonyítás alkalmazásával. A versenyképességi kérdőív esetében először a kérdőív logikai diagramját felhasználva végeztünk szoftveres helyesség-ellenőrzést, majd intézményen belül, hallgatók és mintacégek bevonásával került sor az így nem detektálható, tartalmi hibák felderítésére

és javítására. A teljes elemzési modell futtatása a kevés hallgatói tesztelés miatt nem volt kivitelezhető. Ezt a problémát a tesztrekordok automatikus, véletlenszerű generálásával lehet áthidalni, ami szintén már része a modernebb online kérdőív rendszereknek.

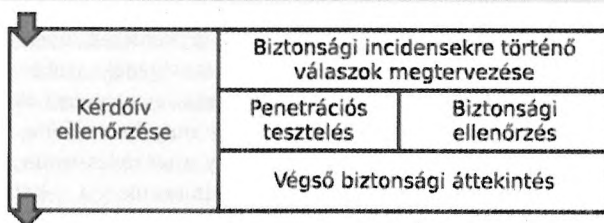
Dinamikus tesztelési módszerek esetén az eredeti rendszer egy úgynevezett „homokozó” változatához férnek hozzá a felhasználók, ahol kipróbálhatják annak teljes funkcionalitását. Az online kérdőív rendszerek esetében ez a lekérdezési időszakok szeparálásával valósítható meg a legkönnyebben úgy, hogy a kérdőívet korlátozott hozzáféréssel egy próba időszakra nyitjuk csak meg. A tesztelők munkája során talált hibákat és azok javítását nagyon

2. táblázat: A versenyképességi kérdőív komplexitási mutatói

Mutató	Érték	Megjegyzés
Kérdések száma:	236 darab	(15 blokkba rendezve)
Beviteli mezők száma:	1 202 darab	(minimális adatbeviteli módszerek kiválasztása után)
Elágazási feltételek száma:	112 darab	(az egyéb mezők kitöltésére vonatkozó feltételeket is beleértve)
Megjegyzések száma:	16 darab	
Kérdőív hossza:	8 579 szó	(bevezető és nyugtázó szöveg, valamint kérdésblokk leírások nélkül)
Ezen belül a kérdőív egyes részeinek megfogalmazásához felhasznált szavak száma:		
Kérdések szövege:	6 396 szó	
Egy kérdés átlagos hossza:	27,10 szó	
Válaszopciók szövege:	2 006 szó	
Egy válaszopció átlagos hossza:	1,67 szó	
Megjegyzések szövege:	177 szó	
Egy megjegyzés átlagos hossza:	11,06 szó	
Egy kérdés válaszopcióinak átlagos száma:	5,09 darab	
Egy kérdés teljes átlagos hossza:	36,35 szó	(válaszopciókkal és megjegyzésekkel együtt)

Forrás: Sauermaann és Roach, 2013 5. táblázatának tömörítése

4. ábra: Online kérdőív ellenőrzési és tesztelési fázisa



fontos dokumentálni. Ehhez lehetőleg ugyanazt a verziókezelő rendszert érdemes használni, amelyiket a megvalósítás során is használtunk. Csoportmunka esetén mindenképpen el kell kerülnünk, hogy a kérdőíven végzett különböző változtatásoknak nyoma legyen, hiszen szinte minden módosításnak következményei vannak az elemzési modellre nézve is. Dinamikus tesztelés során is többféle módszert alkalmazhatunk. Az integrációs tesztelés (Integration Level Testing) során a teljes rendszert átvizsgáljuk az elsőtől az utolsó csavarig, az adatgyűjtéstől kezdve az elemzésekig. Ez drága és hosszú folyamat, amit az éles indítás előtt legalább egyszer mindenképpen el kell végezni. A rendszer szintű tesztelés (System Level Test) az egyes funkcionális területek ellenőrzését jelenti, ahol a valós környezet legközelebbi szimulációját kell megvalósítani. Az ilyen tesztelés során egyrészt meg kell győződni arról, hogy a rendszer hibamentes, másrészt megfelel a felhasználók és megrendelők igényeinek. A szoftverekre kidolgozott rendszerszintű tesztelés szinte valamennyi eleme (szolgáltatás, mennyiségi, terheléses, használhatósági, biztonsági, teljesítmény, konfiguráció, megbízhatóság és dokumentáció tesztelése) alkalmazható az online kérdőívek esetére is. Ezek közül is mindenképpen ki kell emelni a biztonsági és konfigurációs beállítások tesztelését. Online kérdőívek esetén itt kell megvizsgálni, hogy szükséges-e kriptográfiailag is védett (HTTPS) csatornát használni a weblapok eléréséhez, hogyan gátoljuk meg a duplikált kitöltéseket, hogyan különböztetjük meg a kézi és gépi kitöltéseket, miként naplózunk az egyes eseményeket, és végül milyen hálózatechnikai akadályokat (például tűzfalak, proxy közvetítők) kell leküzdeni. Az ismételt, automatikus és fals kitöltések kizárása elsősorban anonim kitöltéseknél igényel komolyabb előkészületeket, hiszen nevesített esetben hozzáférési tokeneket, és elérési jelszavakat tudunk kiosztani. Névtelen kérdőívek esetén a kitöltőt csak hozzávetőlegesen azonosíthatjuk a hálózati

kapcsolatának IP címe, a kitöltés ideje, a böngésző adatbázisában elhelyezett süti (cookie) információ segítségével vagy azon link alapján, amire kattintva érkezett a kérdőív oldalára. A döntést befolyásolhatja például, hogy ha a kitöltők vállalati tűzfal mögött helyezkednek el, hiszen ekkor mindenképpen ugyanaz az IP szám tartozhat, vagy ha ugyanazt a számítógépet vagy böngészőt használják többen is. A kliens oldali azonosítás esetén pedig a képzetesebb és biztonság tudatos felhasználók könnyen törölhetik a süti bejegyzéseinket vagy tiltják le egyéb módszereinket.

Kritikus kérdés szokott lenni, hogy az online kérdőívhez kapcsolódástól számított aktív időt (session) mekkorára állítjuk be. Ha ez túl rövid, akkor a kitöltés közben lejárt miatt az adatokat már nem lehet leadni, és elvesznek. Ha pedig túl hosszú, akkor a rendszerhez egy időtartam alatt engedélyezett hozzáférések száma csökken, ami egy mennyiség fölött már kizárja az újonnan érkező kitöltőket. Ráadásul minél hosszabb a párbeszédre engedélyezett idő, annál nagyobb a veszélye az oldal elleni túlterheléses (Denial Of Service) támadásoknak. Hogy egyidejűleg mennyi felhasználót tud kiszolgálni a rendszert, azt terheléses (penetrációs) teszteléssel lehet megvizsgálni. A versenyképességi kérdőív esetében a párbeszéd időkorlátja 2 napban lett meghatározva, figyelembe véve annak komplexitását és várhatóan kicsi támadási fenyegetettségét.

Adatbeviteli funkcióval ellátott weblapok esetén „injection” típusú visszaélésekre is számítani kell. Ezek lényege, hogy a beviteli mezőkbe, vagy a kitöltési adatokat tartalmazó válasz oldal paramétereibe olyan adatokat adunk meg, amelyek nem megfelelő ellenőrzés következtében akár futóképes, ártalmas kóddá válhatnak. A tesztelés során tehát különös figyelmet kell fordítani a kitöltés szempontjából nem értelmes válaszoknak is! A tervezéskor pedig nem szabad kihagyni az adatbeviteli mezők elfogadási szabályainak megadását.

Ebben a fázisban ellenőrizzük, hogy a biztonsági másolatok működnek és az ese-

ménynaplók is elegendően részletezettek. Az élesítésre csak akkor kerülhet sor, ha a felhasználói tesztelés (User Acceptance Test) során a kritikus elemek újra ellenőrzésével ismét meggyőződünk arról, hogy az elkészült online kérdőív és annak minden beállítása megfelel a megrendelő követelményeinek. Eddigre már rendelkezünk kell az éles üzemeltetéskor előforduló esetek kezeléséhez szükséges terveknek, és az esetleges módosítások után kötelezően elvégzendő gyorsesztek listájával.

KÉRDŐÍV ÜZEMELTETÉS ÉS LEÁLLÍTÁS

Az online kérdőív élesítése és a felkérő levelek kiküldése után állandó 24/7 felügyeletet és támogatást kell adni. A folyamatosan beérkező válaszok mennyiségéről és minőségéről rendszeres jelentést kell adni a megrendelő számára, hogy az emlékeztetők ütemezését optimálisan tudja kialakítani. Kezeleni kell a kitöltőktől érkező kérdéseket, a felfedezett hibák javítását. Mivel nem mindenki vállalja azt a költséget, ami a hibák jelzésével jár, ezért a kitöltések minőségét érdemes menet közben is folyamatosan monitorozni. Ezt megkönnyítendő tervezéskor ajánlott beépíteni egy olyan extra, szabadon kitölthető, szöveges mezőt, ahova a kitöltők beírhatják észrevételeiket. Kritikus esetekben döntést kell hozni a kérdőív ideiglenes leállításáról, annak időzítéséről és az újraindításáról. Ha az adatgyűjtésnek nincs konkrét lezárási határideje, akkor ki kell dolgozni az egyes időszakok között beérkezett adatok növekményes átadásának és felhasználásának módját is. A kérdőív lezárása után még teljesíteni kell a nyugtázási folyamat tevékenységeit, és az adatokat, a kérdőív kódját, valamint az egész adatgyűjtési folyamat dokumen-

tációját át kell adni a megrendelő részére. Amennyiben pedig az online kérdőívet újra fel szeretnék használni, akkor élesítés előtt legalább a kérdőív konfigurációs beállításait és a szövegezés aktualitását felül kell vizsgálni. Még akkor is, ha azok tökéletesen működtek előzőleg. Nagy segítség, ha ezen elemekről szintén lista készül a megvalósítási fázisban, de a felhasználói tesztelést ekkor sem lehet elhagyni. Időközben ugyanis megváltozhat a hálózat konfigurációja, elemek cserélődhetnek benne, de magán az online kérdőív szoftveren is történhetnek módosítások éppúgy, mint azokon az adatbázis és webszerver háttér szolgáltatásokon, amelyeket szintén érintenek legalább biztonsági frissítések.

ÖSSZEFOGLALÁS

Cikkünkben bemutattunk, hogy ha egy internetes kérdőívre alapuló adatgyűjtést akarunk végezni, akkor érdemes megismerkedni az online kérdőíves rendszerek mögött meghúzódó technológiai ismeretekkel is. Ezzel nemcsak biztonságosabb felhasználói felületet tudunk létrehozni, hanem számos olyan lehetőséget is megismerhetünk, ami az online kérdőívezés műfajának különlegessége. Ha az online kérdőívre úgy tekintünk, mint egy speciális felhasználói felület és üzleti logika kialakítására, akkor érthető, hogy nem hagyhatjuk figyelmen kívül az iparág vonatkozó szabványait. Az internetes (azon belül is különösen az anonim) lekérdezések fokozottan ki vannak téve támadásoknak, és itt nemcsak a technikai, hanem a tartalmi (nem valós válaszközléses) visszaélésekre is fel kell készülnünk. Bár a hatékony kérdőívezés és az informatikai biztonság területe külön-külön önmagában már alaposan kutatott

5. ábra: Online kérdőív üzemeltetési fázisa



terület, a jelen munka úttörő a tekintetben, hogy ezt a két területet köti össze.

A TÁMOP-4.2.2.A-11/1/KONV-2012-0058 számú versenyképességi kutatáshoz készített kérdőíves lekérdés folyamatán keresztül röviden bemutatjuk azokat a legjobb gyakorlatokat, amelyek segítségével a több éves tapasztalatainkkal hozzá tudunk járulni a biztonságos és minőségi adatgyűjtéshez. Mindennek a nemzetközileg is elfogadott SD3+C Microsoft módszertan felhasználásával és testreszabásával adtuk keretet. Bízunk benne, hogy munkánkkal hozzájárultunk ahhoz is, hogy a rendelkezésre álló webes kérdőív szoftverek és szolgáltatások közül minden kutatáshoz sikerüljön az egyedi elvárásoknak leginkább megfelelőt kiválasztani.

Köszönetet mondunk a TÁMOP kutatócsoport vezetőjének és tagjainak támogató együttműködésért, illetve bizalmukat, hogy a teljes folyamat lebonyolítása során számítottak ránk és figyelembe vették tanácsainkat.

JEGYZET

- 1 A tanulmánya TÁMOP-4.2.2.A-11/1/KONV-2012-0058 sz. „Energiatermelési, energiafelhasználási és hulladékgazdálkodási technológiák vállalati versenyképességi, városi és regionális hatásainak komplex vizsgálata és modellezése” című pályázat támogatásával készült.

HIVATKOZÁSOK

- Bethlehem, J., Biffignandi, S., (2011), *Handbook of Web Surveys*, New Jersey, Wiley, 480 pp.
- Casey, T. W., Poropat, A., (2014), „Beauty is more than screen deep: Improving the web survey respondent experience through socially-present and

aesthetically-pleasing user interfaces”, *Computers in Human Behavior*, 30 153-163 pp.

Couper, M. P., (2008), *Designing Effective Web Surveys*, Cambridge, Cambridge University Press, 416 pp.

Kehoe, C., and Pitkow, J., (1996), „Surveying the territory.” *The World Wide Web Journal*. 1 (3), 77–84 pp

Kiesler S. and Sproull L.S., (1986) „Response effects in the electronic survey.” *Public Opinion Quarterly*. 50 3, 402-413 pp.

Lipner, S., (2004), „The Trustworthy Computing Security Development Lifecycle”, *Proceedings of the 20th Annual Computer Security Applications Conference*, 2-13 pp.

Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., & Vehovar, V. (2006), „Web surveys versus other survey modes: A meta-analysis comparing response rates.” *International Journal of Marketing Research*, 50 (1), 79-104 pp.

Rogelberg, S., Stanton, J., (2007). „Understanding and dealing with organizational survey nonresponse.” *Organizational Research Methods* 10, 195-209 pp.

Saueremann, H. and Roach, M. (2013), „Increasing web survey response rates in innovation research: An experimental study of static and dynamic contact design features”, *Research Policy*, 42 (1), 273-286 pp.

Schaeffer N. C., Dykema J. (2011): *Questions for Surveys – Current Trends and Future Directions*, *Public Opinion Quarterly*, 75 (5) 909–961 pp.

Smyth, J.D., Olson, K., Millar, M.M., (2014), „Identifying Predictors of Survey Mode Preference”, *Social Science Research* (2014) in press, accepted

Tourangeau, R., Conrad, F. G., M. P. Couper, (2013), *The Science of Web Surveys*, Oxford, Oxford University Press, 208 p.

Kruzsliz Ferenc, egyetemi docens

kruzsliz.ferenc@tkk.pte.hu

Pécsi Tudományegyetem,
Közgazdaságtudományi Kar,
Gazdálkodástudományi Intézet

Security Aspects of Online Surveys

After online survey softwares and services had appeared on the Internet data collection became as easy for anyone as sending electronic postcards. Though wide range of publications is available about creating professional quality surveys, security aspects of this kind of questionnaires has not been investigated yet. To stop this gap in our current work we suggest a framework for developing efficient data collection processes using standards from information security and privacy, based on the analogy between software and online survey development, in order to produce services meeting the researchers' requirements. Beside the theoretical results of our methodology we present its steps and usability on a real world practical example where the main decision points are discussed of the proposed phases.

Ferenc Kruzsliz