

## AZ INFORMATIKA ALAPVETŐ PROBLÉMÁJA, KULCSFOGALMA: A RELEVANCIA

Györe Pál

Kohó- és Gépipari Tudományos Műszaki Tájékoztató Intézet

Az informatikának, a tudományos információ elméletének és gyakorlatának egyik középponti problémája, kulcsfogalma a relevancia. Az informatikai szakirodalomban a relevanciát sokan, sokféleképpen magyarázták, értelmezték, vitatták. Mind ez ideig hiányzott azonban olyan összefoglaló áttekintés, amely a relevanciára vonatkozó nézőpontokat, interpretációkat elemzően, egymáshoz viszonyítva, rendszerbe foglalva vizsgálta volna. Ezt a feladatot végezte el Tefko SARACEVIC [1] a múlt év végén megjelent nagyszabású, széles alapvetésű kritikai szemletanulmányában. A következőkben ennek alapján adunk rövid ismertetést a relevanciát érintő főbb problémákról.

SARACEVIC tanulmányának ismertetése nehéz feladat. Egyrészt azért, mert szemletanulmány, s így (bár mennyire érdekes is) minden részletére nem lehet kitérni, terjedelmi korlátok miatt sem. Másrészt azért, mert stílusa helyenként nagyon „kompakt”, néha homályosan tömör, ami abból is adódik, hogy sok helyről, sokféle stíusból származó mondanivalót összegez. – Az ismertetésből tehát egyes részletek kimaradtak, s hogy az ismertetés készítője mit hagyott ki – az tulajdonképpen szintén a relevanciát (közelebbről a pertinenciát) érintő dolog! Az ismertetés tehát nem követi pontosan a szerző gondolatmenetét, helyenként gondolati átcsoportosításokat, „áthidalásokat”, összevonásokat alkalmaz, ezáltal némileg az ismertetés írójának véleményét is tükrözi a relevancia leglényegesebb problémáit illetően.

### A relevancia és a tudományos kommunikáció

A relevanciát, éppen úgy mint a tudományos információ problémáit általában, legcélszerűbb a tudományos ismeretek kommunikációjának folyamatába illesztve tárgyalni. Ezt indokolja a relevancia fogalmának történeti kialakulása, felmerülése is. Hogyan került előtérbe ez a fogalom?

A tudományos kommunikáció egyik alapvető mechanizmusa, a XVII. századtól mind a mai napig a tudományos munka részeredményeinek (kutatási témái-

nak) rendszeres, válogató közreadása folyóiratokban vagy más módon. Ez a mechanizmus tette lehetővé a tudomány, elsősorban a természettudományok gyors, hatékony növekedését, s ez vezetett el a relevancia problémájához is. A fejlődés során a tudomány eredményes műveléséhez nem egyszerűen csak információkra, hanem releváns információkra volt szükség, a tudományos kommunikáció folyamatán belül.

A nem releváns információk ugyanis mind nagyobb mértékben veszélyeztették a tudományos kommunikációt, s ezáltal a tudományos munka hatékonyságát. A tudományos ismeretek nagy méretű növekedése, specializálódása, és egyéb tényezők következtében időnként zavarok támadtak a releváns és nem releváns információk egyensúlyában. Ezekben a válságos periódusokban figyelhető meg a tudományos kommunikáció, ezen belül az információs tevékenység, korszakos fejlődési irányainak kibontakozása.

Ilyen volt a századforduló táján a dokumentáció előretörése, amikor az egyensúlyt a tudomány szerveződésének megfelelő dokumentációs feltárás (osztályozás, indexelés, kivonatolás stb.) útján próbálták megvalósítani. Ilyen volt a II. világháború után az információkereső rendszerek kialakulása, ami arra irányult, hogy technológiai úton biztosítsa az egyensúly fenntartását, a szakemberek releváns információkkal való ellátását.

Vannak, akik úgy vélik, hogy a dokumentáció azért jött létre, mert a könyvtárak kudarcot vallottak a tudományos kommunikáció problémáinak felismerésében, megoldásában. Ebből következően azt is gondolhatnánk, hogy az informatika azért keletkezett, mert a dokumentáció is csődöt mondott a tudományos kommunikáció valódi feladatainak megragadásában, megoldásában.

Valamely szakma (szakterület) azonban általában nem azért fejlődik ki, mert egy másik szakma csődöt mond, hanem azért, mert egy egészen sajátos problémakört ragad meg, határol el, s megfelelő eszmék, módszerek, eszközök alkalmazásával kísérletet tesz a jelentkező, illetőleg felismert problémák megoldására. Minden azon múlik, hogyan ragadja meg, határolja körül, célozza meg

a problematikát. *Az informatika sikerét* (bármilyen lesz is ez) *az teszi lehetővé, hogy felismerte, középpontba állította a relevanciát, s az informatika kudarcát* (bármilyen lesz is ez) *az fogja jelenteni, ha nem megfelelően, nem helyes viszonylatokban nézi, vizsgálja, ragadja meg a relevancia problémáit.*

A relevancia tehát alapvetően a tudományos kommunikáció keretébe tartozó fogalom. Ebben az összefüggésben úgy határozható meg, mint az a mérték, amely az információ forrása és rendeltetése (felhasználója) közötti érintkezés (kapcsolat) hatékonyságát fejezi ki a kommunikáció folyamatában. S mivel *minden mérték viszony, a relevancia is viszonyt kifejező fogalom.*

Mind a forrás, mind a rendeltetési hely (felhasználó) a tudományos ismeretek kommunikációjának alkotó eleme, olyan file (tár), amelyben a szakismeretek és/vagy ezek reprezentációi szervezett formában vannak tárolva. (Ilyen file pl. az emberi és gépi memória, a könyvtári gyűjtemény, a katalógus, az adatbank.) A tudományos ismeretek kommunikációja akkor és abban az esetben hatékony, ha az információ átvitele az egyik file-ből változásokat eredményez a másik file-ban. A relevancia nem más, mint ezeknek a változásoknak a mértéke.

Az információs rendszerek is be vannak ágyazódva a tudományos kommunikáció folyamatába, ezt segítik, erősítik, tartják fenn, vagy bővítik. Ennek megfelelően *a relevancia szorosan kapcsolódik az információs rendszerek működéséhez.* Ezen kívül természetesen más szempontokat is figyelembe kell venni a relevancia fogalmának megítélésében, értelmezésében.

### A relevancia fogalma – különféle nézőpontokból

A tudományos kommunikáció s a hozzá kapcsolódó információs rendszerek mindig valamely társadalmi környezetben (ökológiai szisztémában) működnek, ezért mélyen benne gyökereznek az adott társadalmi realitásokban, s az emberi tényezőkben is. Ebből következően *a relevanciát – a kommunikációs folyamaton belül – sokféle szempontból lehet és kell vizsgálni.* Ezeket a szempontokat az alábbi kategóriákba sorolhatjuk:

- szaktudomány* (szakterület, szakismeret, szaktudás),
- szakirodalom* (dokumentumok, publikációk összessége, típusai),
- egyéb nyelvi vagy szimbolikus (tartalmi) ábrázolások,*
- információs forrás* (különösen a file, vagy file-ok),
- az információ rendeltetési helye, címzettje, felhasználója* (különösen a file vagy file-ok),
- információs rendszerek,*
- környezeti tényezők,* társadalmi realitások, funkciók, stb.,
- értékrendszerek* (társadalmi, kulturális, erkölcsi stb. értékek).

A relevanciát érintő *vizsgálódások többsége a következő két problémakörre keresett választ:*

milyen tényezők, elemek, szempontok vannak jelen a relevancia fogalmában?

milyen sajátos viszonyra (relációra) vonatkozik a relevancia?

Az eddigi vizsgálatok ellentmondásai, hiányosságai jórészt abból erednek, hogy az előzőekben felsorolt nyolc szempontból csak némelyiket veszik figyelembe, ezt azonosítják a relevanciával s más tényezőket nem ismernek fel. Ha az informatikában kidolgozásra kerül majd a relevanciának teljességre igényt tartó elmélete, ez valószínűleg a jelzett kategória-rendszerben jön létre, s magába foglalja (legalább!) a felsorolt szempontokat.

Az informatikában az eddigiek során a következő szempontok – *elméleti és gyakorlati megfontolások* – merültek fel a relevancia fogalmának megítélésében (a különféle szempontok felsorolása hozzávetőlegesen azt a sorrendet követi, amely a kommunikáció egymásra következő folyamataira utal).

1. A **szaktudományi** (szakterületi) nézőpont úgy tekinti a relevanciát, mint azt a viszonyt, ami az adott szakterület ismeretkészlete és egy, az adott szakterületre vonatkozó téma (kérdés) között áll fenn.

Ez alapvető nézőpont, mert a szakterületnek (szaktárgynak, szaktudásnak stb.) rendkívül fontos szerepe van a tudományos kommunikációban. Ez a szemlélet egyébként rokonságban van azzal a megközelítési móddal, amelyet a filozófus SCHUTZ *tárgyi (topical) relevanciának* nevezett, s úgy értelmezte, hogy ez tulajdonképpen valami problematikusnak az észrevételét jelenti a problémát nem jelentő ismerős jelenségek közepette. Ez a probléma-felismerés vezet valami nem ismert dologra vonatkozó kérdés feltevéséhez, és ebből keletkezik tulajdonképpen az információs igény, amely egyúttal már pszichológiai, szociológiai „jelenség” is.

2. A **szakirodalmi** nézet szorosan kapcsolódik a szakterületihez s a relevanciát úgy közelíti meg mint olyan relációt, amely a szakterület (szakismeret) s ennek megjelenési formája (reprezentációja): a szakirodalom, vagy a szakirodalom és a szakterületre vonatkozó adott téma (kérdés) között jön létre.

Az információs rendszereket tekintve, a kommunikációs folyamat tulajdonképpen a szakirodalommal kezdődik; *az információs rendszerek célja, hogy megteremtseik és hatékonyra tegyék a szakirodalom és a felhasználók közötti kapcsolatot,* ezért a szakirodalmi relevancia minden más szempontot befolyásol.

S lényegükben véve a relevanciával vannak összefüggésben mindazok a *bibliometriai kutatások* is, amelyek a szakirodalom struktúrájára, használatára vonatkoznak; így pl. BRADFORD, LOTKA, ZIPF, MANDELBROT statisztikai eloszlási (szóródási) vizsgálatai, PRICE elemzései a szakirodalmi hivatkozások hálózatáról,

URQUHART kvantitatív megállapításai a nagy tudományos könyvtárak folyóiratállományának használatáról, stb.

A relevancia ilyen szemléletének további elmélyítése érdekében döntő volna annak a mechanizmusnak kutatása, amely egy időpontban megszabja a szakirodalom struktúrájának alakulását.

3. A logikai nézőpont a relevanciával kapcsolatban azt veti fel, hogy milyen természetű összefüggés van egy témára (szakkérdésre) vonatkozó előfeltevések (premisszák) és az adott szakterületből, vagy szakirodalomból levonható következtetések (konklúziók) között.

Ebben a tekintetben kétféle megközelítési mód merült fel: a) a *deduktív szemléletmód*, amely a premisszák és a konklúziók viszonyát a logikai következtetés alapján nézi, s b) a *valószínűségi következtetés szemléletmódja*, amely vizsgálja a premisszák közötti viszonyt, az információt mint evidenciát és a konklúziókat a megerősítés, vagy a valószínűség fokozatának alapján.

A relevanciának ez a szemlélete elsősorban CARNAP műveire támaszkodik, amelyek a valószínűség logikai alapjait tárgyalják.

4. Az **információs rendszer** nézőpontjából a relevancia azt jelenti, hogy milyen a viszony egy adott rendszer file-jainak tartalma és/vagy folyamatai, valamint a szakterület vagy a szakirodalom, továbbá egy adott téma (szakkérdés), egy felhasználó, vagy több felhasználó között.

Az információkereső rendszerek kialakulása idején *feltételezték, hogy amit a rendszer kikeres, az releváns, amit nem keres ki, az nem releváns információ*. De hamar felismerték, hogy nem minden kikeresett információ releváns s ennek (vagyis a „zajnak”) okát a rendszer belső felépítésében és működésében, elsősorban az osztályozási (indexelési), kódolási, nyelvfeldolgozási, file-szervezési, valamint a kérdéselemzési s a kereső stratégiai problémákban (fogymeghatározásokban) keresték.

Nem vették pl. figyelembe, hogy *a relevanciának milyen fontos tényezője az információk beválogatása a rendszerbe*. Ezt a szempontot később az információelemző központok kezdték tudatosan és rendszeresen alkalmazni. A relevanciának ez a tényezője azonban mind ez ideig nem került megfelelő mértékben vizsgálatra. Kísérletek történtek az információkeresés relevanciáját jellemző mérőszámok („recall”, „precision”) megállapítására, s kiterjedt vizsgálatokat folytattak a relevancia ítéletek meghatározására (CLARENDON/CRANFIELD stb.). S ezek már átvezettek ahhoz a szemlélethez, hogy a relevancia szoros összefüggésben van az információ címzettjének, felhasználójának véleményével, megítélésével.

5. A **felhasználói nézet** a relevanciát abból az aspektusból vizsgálja, hogy milyen szerepe van az emberi (szubjektív) értékelésnek (megítélésnek) az információt

hordozó dokumentumok és egy adott téma (szakkérdés) viszonyában.

Megállapították, hogy milyen tényezők befolyásolják a relevancia-ítéleteket. Ezek a következő csoportokba sorolhatók:

dokumentumok s ezek reprezentációi, szakkérdések (az információs szükségletre vonatkozó megállapítások), az ítéletalkotás során szerepet játszó helyzetek, körülmények, kifejezési módok (pl. rangsorolási skála), az ítéletet hozó személyek.

Látható, hogy a relevanciának ebben a megközelítési módjában *nagy szerepe van az emberi (főként lélektani, szociológiai) tényezőknek*. Éppen ezért meglepő, hogy nem vizsgálták kellőképpen az emberi memóriának a relevancia-ítéletekben megnyilvánuló szerepét. A memória, s ennek korlátai – amint az információszolgáltatás gyakorlatából tudjuk – komoly mértékben befolyásolják az információs/kommunikációs folyamatokat; ebből következően az ilyen irányú kutatások az informatika legfontosabb, legtöbbször ígérő feladatai közé tartoznak.

6. A **pertinencia** nézőpontja szorosan kapcsolódik a relevancia felhasználói szemléletéhez, tulajdonképpen annak változata. Ez arra a viszonyra vonatkozik, ami egy adott személy (megismerő) saját tudáskészlete és valamely szakterület vagy szakirodalom között áll fenn.

Ebben az esetben az alapvető tényezők a felhasználó szellemi apparátusában adva levő ismeretkészlet, ismeret-tartomány jellege, struktúrája, kiterjedése, az ismeretkészlet képződésének, az ismereteket feldolgozó intellektus működésének (szelektív) folyamatai stb.

A pertinencia-elmélet talán jobban megalapozható, mint az egyszerűbb felhasználói szemlélet, mert azokra a filozófiai és pszichológiai kutatásokra is támaszkodhat, amelyek annak megértésére vonatkoznak, hogy miképpen tudjuk azt, amit tudunk. Ezekhez képest az emberi ítéletekre vonatkozó kutatások kevésbé fejlettek, ennek folytán a relevancia felhasználóra orientált megközelítéséhez kevesebb támpontot nyújtanak.

7. A **pragmatikus nézet** úgy fogja fel a relevanciát, mint a felhasználónak egy sürgősen jelentkező problémája és az erre adott információ közötti relációt, mindenekelőtt az információ hasznossága és kedvező feltételeket teremtő hatása szempontjából.

Ez a szemlélet a pragmatikus-utilitárius gondolkodásmódból táplálkozik, s kialakulását nem kis mértékben idézte elő a költség/haszon tényezők elsőrendű figyelembevétele. A pragmatikus érvelés így hangzik: szép dolog, ha az információkereső rendszerek releváns információt szolgáltatnak, de *az igazi feladat az, hogy hasznosítható információt adjanak*, – olyat, amely közvetlenül hat a

cselekvésre, és tökéletesen kielégíti egy adott (döntési) helyzet igényeit.

COOPER úgy különböztette meg az információ relevanciáját és hasznosságát, hogy az előbbi főleg arra vonatkozik, hogy miről, milyen tárgyról van szó, az utóbbi pedig arra, hogy milyen a szakmai információ minősége, újdonsága, jelentősége, megbízhatósága stb. Olyan dokumentumokat, információkat kell tehát kikeresni, amelyek nem egyszerűen csak relevánsak, hanem egyúttal hasznosak is!

WILSON a logikai és szituációs relevancia fogalmával, KOCHEN a dokumentumokat jellemző hasznossági szám bevezetésével járult hozzá a pragmatikus szemlélet elmélyítéséhez, magyarázatához.

A pragmatikus megközelítés természetesen nem lehet kielégítő a relevancia végső értelmezésére, s különösen veszélyes a költség/haszon-tényező mindent megelőző szerepének hangoztatása. Ha ez a szemlélet egyetemessé válna a tudományos kommunikációban, – a legbiztosabb útja volna annak, hogy hatékonyságát teljesen tönkre tegye.

#### A relevancia-rendszerek együttese

A relevancia fogalmának megközelítésében, megértésében a felsorolt szempontok mind fontos szerepet játszanak, és nem függetlenek egymástól (amint erre már utaltunk). Úgy látszik, hogy a felmerült szempontok ciklikusan összekapcsolódó, együtt játszó rendszereket alkotnak. S ezek közül néhány úgy fogható fel, mint az általánosabb rendszer speciális esete, alrendszere. Nincs, és nem is lehet egyetlen specifikus szempontú megközelítése a relevanciának, mert a kommunikációban nem található egy ilyen rendszer.

A relevancia különféle rendszerei más-más tényezőket vesznek figyelembe, de ezek annyira összefonódnak, hogy más rendszerek nélkül nem mérlegelhetők. Így pl. a pragmatikus szemlélet nem választható el a pertinencia – vagy a felhasználói (rendeltetési) szemlélettől, és egyiket sem lehet elkülöníteni a szakterületi nézőponttól.

Az információs rendszerek igen sok gyakorlati problémájának (nehézségének), s a felhasználók olykori elégedetlenségének sok esetben az a magyarázata, hogy a relevanciának különböző rendszerei, megítélési módjai

alakultak ki. A relevanciát jellemző különféle relációknak különféle neveket adhatunk (pl. pertinencia, logikai relevancia), de állandóan szem előtt kell tartanunk, hogy amikor relevanciáról van szó, akkor tulajdonképpen mindig relevancia-rendszerek összefüggő csoportjára kell gondolnunk.

A relevanciára vonatkozó kutatásokban, gondolkodásban az lesz a legjelentősebb előrehaladás, amikor fény derül arra, hogy a relevancia-rendszerek között milyen kölcsönhatás, összjáték nyilvánul meg.

Bár a gyakorlatot egyelőre csak közvetve érintik, perspektivikusan igen fontosak a relevancia fogalmához kapcsolódó elméleti kutatások. Így pl.: MARON és KUHNS vizsgálatai a relevancia-számról, amelyek arra irányulnak, hogy valószínűségi mértéket állapítsanak meg a dokumentumok rangsorolására relevanciájuk szerint; GOFFMAN elmélete a relevanciáról mint mértékről, s ugyancsak GOFFMAN epidemiológiai elmélete, amely a járványok terjedésének analógiáját, matematikai modelljét használja fel a kommunikáció (az ismeretek terjedésének „fertőző” hatásának) vizsgálatában; HILLMAN és COOPER logikai kutatásai, különösen COOPER vizsgálatai, amelyek a tárolt s keresett információkat mint nyelvészeti entitásokat (mondatokat) elemzik, a logikai relevancia szempontjából.

Amint az előzőekből kitűnik, az informatika igen sokféle gyakorlati és elméleti szempontból vizsgálta a relevancia fogalmát. Az eddigi eredményeket áttekintve, talán az a leginkább feltűnő hiányosság, hogy milyen kevésbé kutatták a tudományos kommunikációnak s benne a relevanciának azokat a jelenségeit, amelyek a társadalmi környezettel, s az emberi értékek világával vannak összefüggésben. Remélhetőleg a jövőbeli kutatások a relevanciának ezekre a fontos aspektusaira is ki fognak terjedni.

#### Hivatkozás

- [1] SARACEVIC, T.: Relevance: a review of and a framework for the thinking on the notion in information science. = Journal of the ASIS, 26. köt. 6. sz. 1975. p. 321–343. – A tanulmány fontos kiegészítő része a függelék, amely az informatikában a relevanciára vonatkozó eddig végzett kísérleti vizsgálatok eredményeit összegezi, szintetizálja, – a szerzőnek egyik korábbi tanulmánya alapján.



*GYÖRE P.: Az informatika alapvető problémája, kulcsfogalma: a relevancia*

A cikk SARACEVIC kritikai szemletanulmánya alapján, ennek leglényegesebb megállapításait kifejtve ismerteti a relevanciát érintő fontosabb problémaköröket a következő csoportosításban: a relevancia felmerülése, szerepe, jelentősége a tudományos kommunikációban, és az informatikában; a relevancia fogalmának értelmezése különféle nézőpontokból, az eddigiekben végzett elméleti és gyakorlati vizsgálatok rendszerbe foglalása, értékelése; a különféle relevancia-rendszerek összefüggéseinek, együttes hatásának vizsgálata; a relevancia elméletére vonatkozó eddigi és további kutatások jelentősége.

\* \* \*

*GYÖRE P.: Relevance, the basic problem and key notion of information science*

Based on the critical review of SARACEVIC, and expounding its contents, the article provides a survey of main problems referring to relevance: the emergence, role and significance of the notion of relevance in scientific communication and information science; the definition of relevance from different points of view, the systematization and evaluation of theoretical and practical investigations into relevance; the correlation of various systems of relevances and of their conjugate effects; the significance of past and future experiments related to the theory of relevance.

*ДБЕРЕ, П.: Основная проблема и ключевое понятие информатики: релевантность*

На основе основных установлений обзора Саракевича статья представляет некоторые важные проблемы релевантности: роль и значительность релевантности в научной коммуникации и в информатике; объяснение понятия релевантности в разных расфасовках; систематизация и оценка теоретических и практических исследований по релевантности; изучение взаимосвязей и взаимодействия разных систем релевантности; значительность бывших и будущих исследований по теории релевантности.

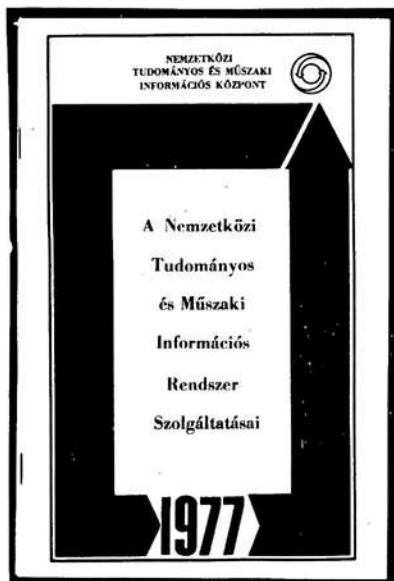
\* \* \*

*GYÖRE, P.: Das grundlegende Problem, der Schlüsselbegriff der Informatik: die Relevanz*

Aufgrund der kritischen Studie von SARACEVIC und dessen wichtigste Feststellungen erörternd behandelt der Autor einige wesentliche Probleme der Relevanz: die Erscheinung und die Rolle der Relevanz in der wissenschaftlichen Kommunikation und in der Informatik; Deutung des Begriffs der Relevanz aus verschiedenen Gesichtspunkten; Systematisierung und Auswertung der theoretischen und praktischen Untersuchungen über Relevanz; Untersuchung der Zusammenhänge und der gemeinsamen Wirkung der verschiedenen Relevanzsysteme; die Bedeutung der bisherigen und der weiteren Forschungsarbeiten über die Theorie der Relevanz.



**M  
E  
G  
J  
E  
L  
E  
N  
T**



A moszkvai Nemzetközi Tudományos és Műszaki Információs Központ (NTMIK) kiadta A Nemzetközi Tudományos és Műszaki Információs Rendszer Szolgáltatásai 1977 c. kiadványt.

A katalógus első része az NTMIR egészét érintő kiadványokról, a második rész a Nemzetközi Speciális Információs Rendszerek (NSIR), a harmadik rész a Nemzetközi Ágazati Tudományos és Műszaki Információs Rendszerek (NÁTMIR) kiadványairól és szolgáltatásairól ad tájékoztatást. A kiadványok és szolgáltatások orosz nyelvűek.

Az előfizetés vagy megrendelés módját és az árat az adott kiadványnál ill. szolgáltatásnál ismerteti a katalógus.

A katalógus térítésmentesen áll az érdeklődők rendelkezésére. Igényelhető az Országos Műszaki Könyvtár és Dokumentációs Központ Terjesztési Osztályától, 1428 Budapest, Pf. 12.