

## SZÁMÍTÓGÉPEK ÉS A TÖRTÉNETTUDOMÁNY

Budapest 1972. október 25.

(a Neumann János Számítógéptudományi Társaság, a Magyar Történelmi Társulat és a Magyar Agrártudományi Egyesület Agrártörténeti Szakosztályának kerekasztal-konferenciája)

A történészek és a számítógépes szakemberek első magyarországi szakmai tanácskozására a számítógépes kultúra fejlesztését célzó országos tervhez kapcsolódóan, a Neumann János Számítógéptudományi Társaság hathatós támogatásával került sor. Létrejöttében nagy része volt annak, hogy a legutóbbi nemzetközi történészkongresszus — Leningrád—Moszkva 1970 — behatóan foglalkozott a számítógép történettudományi alkalmazásának eddigi eredményeivel. E kongresszus tanulságaira építve, módjuk volt rá a magyar történészeknek, hogy a tanácskozásban figyelembe vegyék a nemzetközi problematikát és a fő vitakérdéseket, de a napirendben elsősorban a magyarországi történetírás konkrét kérdéseit tárgyalják. Azokat az aktuális kutatási feladatokat vették számba, ahol elsősorban indokolt a nagyfokú automatizálás, ill. ahol érdemleges feldolgozást nem is lehet nélkül tervezni. Ehhez kapcsolódóan megvitatásra kerültek azok a gyakorlati problémák és tennivalók is, melyek a számítógépes módszerek szélesebb körű bevezetésével jelentkeznek a történettudományban és a történész-képzésben.

A vitát *Tarján Rezső*, a Neumann János Számítógéptudományi Társaság elnöke vezette. Résztvevői a számítógép történettudományi alkalmazása iránt érdeklődő számítógépes szakemberek, számítógépet használó vagy használni készülő történészek és az addigi történeti feldolgozásokban résztvevő matematikusok és programozók voltak.

### Tarján Rezső: Elnöki megnyitó

Egy évvel ezelőtt örömmel de nem aggodalom nélkül fogadtam a történészek megkeresését, amikor a közös munka terve felvetődött, hiszen sajátos, a számítógépeseknek idegen szakterületről van szó. Egyes tapasztalatok mégis arra biztattak, hogy a Neumann János Társaság csatlakozzék az előkészítő kerekasztal-konferenciához.

Az egyik ilyen tapasztalatot a nyelvtudomány szolgáltatta, a gépi fordítás problémájával kapcsolatosan. Itt a kísérletek közben kiderült, hogy a nyelvészek nem ismerték eléggé a grammatika szabályait. Egyszerű mondatoknál, mint például *az ég kék* és *a fa ég* — ahol ugyanaz a két betű az egyik mondatban főnevet, a másik mondatban igét jelent — a gépnek differenciálnia kell. Hogyan lehet ezt megoldani? Kiderült, hogy a hagyományos grammatikai szabályok, úgy ahogyan a nyelv-

vészek az utolsó száz évben vagy még régebben kidolgozták, a gépi fordításra nem alkalmasak, nem kielégítőek. Így fejlődött ki a matematikai nyelvészet, melynek most már Magyarországon is neves képviselői vannak.

A számítógépesek hozzá vannak szokva ahhoz, hogy különböző természeti jelenségeket, majd később nemcsak természeti jelenségek, hanem egész szervezetek — emberi szervezetek, vagy például nagy vállalatok — viselkedését egy matematikai modellben foglalják össze. A matematikai modellt lefuttatják a gépen és a megfelelő külső hatásokat szimuláló gerjesztésekre a gép kimutatja, hogy az adott szervezet egy adott ingerre hogyan fog reagálni. Ha a történelem folyamán a különböző osztályok, társadalmak mozgásáról van szó, ha történelmi mozgásformákról beszélünk: akkor ha ma talán még nem is, de előbb vagy utóbb eljön az ideje, hogy a számítógép nemcsak egyszerű kiszolgálója, műszaki segédeszköze lesz történészeknek — olyan értelemben, hogy segít statisztikázni, különböző számia sokat elvégezni, vagy segít ismeretlen feliratokat megfejteni. Egyszer talán majd alkalmas lesz a számítástechnika — és ez a történészek nélkül nem megy — arra is, hogy egész társadalmi rétegek, osztályok mozgásformáit is vizsgálni lehessen segítségével.

Ettől persze ma még messze vagyunk. Az a néhány érdeklődő, aki itt megjelent, valószínűleg úgy fog járni, mint Mózes az ígéret földjével: elindítunk itt valamit s a végét nem mi fogjuk élni, hanem a mi esontjainkon fog felépülni az az új generáció, amely ennek a mai kezdetnek a hasznát élvezzi majd.

(A napirend ismertetése során az Elnök sajnálattal jelentette be, hogy *Malhakai László: Modellproblémák a történettudományban* című referátuma a szerző hivatalos távolléte miatt nem kerülhet megvitatásra; egyúttal bejelentette egy erről a témáról rendezendő külön vitaülés tervét.)

## TÁJÉKOZTATÓ REFERÁTUM

### Granasztói György: A számítógépek a történettudományban

(Megjelent a Történelmi Szemle XIV. évf. 1—2. számában.)

## PROBLÉMAFELVETŐ REFERÁTUMOK

### Fügedi Erik: Magyarország története és gépi adatfeldolgozás

#### 1.

Magyarországon a XVIII. század derekán kezdtek statisztikai vagy statisztikailag kiértekelhető jelentőséket és adatgyűjtéseket készíteni, de ezek csak II. József uralkodása alatt érték el a pontosságukat a mértékét, amely kvantitatív feldolgozásra érdemesíti őket (pl. lélekösszeírás 1772—1782. első népszámlálás 1784—1787).

A legfontosabbak: első népszámlálás, kataszter; a gabona- és élelmiszerárak, a külkereskedelmi forgalom, a gabonavetés és termés, a nemesfémhányászat, a népiskolák statisztikája és az úrbérrendezés.

Ezek egy része csak töredékesen maradt fenn, mégis alkalmasak arra, hogy

a) a XVIII. század végi Magyarország statisztikáját elkészítsük és ezzel

b) alapot teremtsünk egyrészt a korábbi korszakokról készült feldolgozások ellenőrzéséhez, másrészt a későbbiekkel való összehasonlításához.

## 2.

Magyarország történetének minden időpontjában hihetetlenül differenciált terület volt a földrajzi viszonyok, a település, a gazdasági életforma, a gazdálkodás színvonala szempontjából és nemzetiségileg. Ehhez 1526 után vallási különbség járult, amely a kulturális életet is különbözőképpen befolyásolta.

Elfogadva azt a tételt, hogy

a) a mai statisztika elsősorban mintavétellel dolgozik,

b) célunk az országos jelenségek felderítése, nem pedig az egyes vidékek közötti különbségé,

c) szem előtt tartva azt a tényt, hogy forrásaink egy része töredékes,

le kell vonnunk azt a következtetést, hogy első feladatunk olyan rétegezett minta kidolgozása, amely forrásaink statisztikai feldolgozását lehetővé teszi (a legkönnyebbnek tűnik a demográfiai).

## 3.

A történettudomány csak akkor válhat partnerré egy számítógépes együttműködésben, ha a) nem a forrásokat próbálja „mindentre kiterjedően” feldolgozni, hanem előre megfogalmazott kérdésekre keres a forrásokban választ (pl. stacionér volt-e az ország népessége a XVIII. század végén);

b) felfeketi az 1770—1780 közötti statisztikai források kataszterét;

c) kidolgozza dokumentációjukat;

d) tisztázza fogalmi körüket és összefüggéseiket;

e) lemond azokról az előítéletekről, amelyeket részben a múlt században szedett fel, részben a számítógépekbe vetett messianisztikus hitből fakadnak.

\*

Kevés hozzátennivalóm van azokhoz a tézisekhez, amelyeket mindnyájan megkaptak és remélem, hogy a jelenlegi értekezlet sűrűsödési időpontban is át tudnak futni. Tulajdonképpen két gondolatot szerettem volna kifejezésre juttatni, és ehhez szeretnék néhány megjegyzést fűzni.

Az első gondolat az, amit itt finoman úgy igyekeztem kifejezni, hogy „tisztázza fogalmi körüket és összefüggésüket”. Ez alatt a tisztázás alatt nemcsak fogalmi tisztázást kell érteni, hanem végig kell csinálni a hagyományos történettudománynak azt a forráskritikai munkáját, amit minden forrással kapcsolatban mindig elvégeznek. Engedjék meg, hogy erre egy példát mondjak. A mi csoportunk XVIII. század végi—XIX. század eleji népiskolai statisztikák feldolgozását vette tervbe. Úgy látszott első pillanatban, hogy ezekről pontos statisztikákat készített a Helytartótanács, amiket meg is lehetett találni. És akkor az első kérdés az volt, hogy vajon megvannak-e minden egyes évből, és amikor egy félévi munkával az egész helytartótanácsi anyagot munkatársunk megnézte, kiderült, hogy a sor lukas. Szerencsénk volt, mert azért sikerült olyan egymásra következő éveket találni — mégpedig ritmikus eloszlásban —, hogy azokat föl lehet dolgozni. Még mindig nem nyugodtam bele, a látszat túl szép volt, és azt kérdeztem, hogy vajon minden iskola benne van-e, és egyből kiderült, hogy csak a katolikus iskolák vannak benne; a protestánsokat tehát másutt, a protestáns levéltárakban kellett megkeresni. És ebben a pillanatban tartunk ott, hogy most már tudjuk, milyen anyag van a kezünkben, de tisztázni kell azokat a fogalmakat, amiket a Helytartótanács, illetőleg a korabeli protestáns statisztikák használtak. Köztudomású, — elnézést kérek a történészeketől —, hogy a fogalmi tisztázás során két nehézséggel kell megküzdenünk. Az egyik általános, a másik egy tipikusan magyar nehézség. Az általános nehézség az, hogy az emberek az új fogalmakat is, legalább is még a XVIII. századig, legtöbbször régi szavakkal fejezték ki, tebát ugyanazokat a szavakat találjuk, amelyek már régebben előfordultak, csak éppen a tartalmuk, az értelmük változott meg. A másik, a magyar nehéz-

ség az, hogy igen korán, már a XIII. században nálunk jogi formalizmus uralkodott el, és minden vagy majdnem minden gazdasági, főleg a jobbágytelekkel, a paraszbtirtokkal kapcsolatos dolog a jogi fogalmak kamuflázsa alatt jelenik meg, ugyanazok alatt a jogi fogalmak alatt nyilván különböző értelemben. Ezeknek a fogalmi köröknek és az összefüggéseknek tisztázására általában nem elég egy történész, le kell mondanunk azokról az előítéletekről, amelyeket részben a múlt században szedtünk fel. Ezek az előítéletek individuális munkára serkentenek mindenkit; és a múlt században óriási munkák készültek el egy ember erőfeszítése során, minden európai történettudományban. Ez ma már lehetetlen. Nemcsak azért, mert túlságosan hosszú ideig tart, hanem lehetetlen azért, mert egy ember nem érthet még a történettudományon belül sem mindenhez, jogtörténettől kezdve gazdaságtörténetig, és ezért valami együttműködési formát kell találni. Elnézést kérek, hogy ismét azt mondom, amit már többször elmondtam, hogy ha természetesnek tartjuk, hogy egy nagy épületet nem egy ember, hanem egy team épít fel, illetve tervez meg, amiben a vezető tervezőn kívül még statikus, épületgépész és isten tudja még ki mindenki van együtt, akkor ugyanezt valahogy meg kell alkotnunk a történettudomány keretein belül is. Külföldön is ezt az utat követik, és igen szép, természetesen az anyagtól függően különböző formáit valósították meg. A történészek a számítógépekkel szemben ma általában kétféle álláspontot foglalnak el, boldogan mondhatom, hogy mind a két álláspont szélsőséges. Az egyik elutasítja, mert egy ilyen számítógépes feldolgozás nem fér bele az individuális elképzeléseibe; a másik pedig azt hiszi, hogy a számítógép mindazokat a problémákat önmagától meg fogja oldani, amikre ő nem tudott választ kapni. Azt hiszem egyik sem vezet semmiféle eredményre. Tudomásul kell vennünk, hogy nagyobb lehetőségeink vannak a számítógépekkel, és ezzel együtt sokkal pontosabb munkát kell végeznünk, mint amilyent mi eddig általában éppen fogalmi tisztázás terén végeztünk.

*A másik gondolat*, amit szerettem volna hangsúlyozni, az, hogy a XVIII. század végén, pontosabban II. József uralkodásától kezdve, rendelkezésünkre állnak olyan eredetileg is számszerűen készített és éppen ezért feldolgozásra alkalmas anyagok, amelyek alapot nyújtanak arra, hogy a korábbi, akárcsak a XVIII. század elejének meglehetősen zűrzavaros viszonyait, és a későbbi, a XIX. sz. eleji fejlődést, az azokat reprezentáló számokat is erre a bázisra alapozzuk és ennek alapján dolgozzuk fel.

Azt hiszem, ez az a két gondolat, amivel érdemes lenne foglalkozni és nyilván az értekezlet még számtalan részletében is foglalkozni fog.

Örömmel ragadom meg a szerkesztőség szíves ajánlatát, hogy a tézisben foglaltakhoz még néhány kiegészítő megjegyzést fűzzek.

I. Mária Terézia és II. József 1769-től kezdve nagy erőfeszítéseket tettek, hogy Magyarország állapotáról felmérésekkel jobb és megbízhatóbb adatokat nyerjenek és ismereteiket rendszeresen megismétlődő jelentések segítségével módosítsák és ellenőrizzék. A rendszert, úgy látszik, Franciaországból vették át, érdeklődésük elsősorban a népesség állapotának és szaporodásának, a mezőgazdasági termelésnek és a legfontosabb mezőgazdasági termékek ára alakulásának szolt. Az ország közigazgatásának megfelelően a jelentéseket a törvényhatóságoktól, tehát a megyéktől és városoktól kérték. Ezeket a Helytartótanácsnak kellett összegyűjtenie és összesített formában a kancelláriához továbbítani; az eredeti jelentések a Helytartótanácsnál maradtak, azokat a számvetőség dolgozta fel és őrizte meg.

A nyugat-európai gondolkodást elmaradt magyar törvényhatóságok azonban az uralkodó elképzeléseit a legjobb esetben meg-nem-értéssel vagy közönyt, rosszabb esetben azzal a gyanúval fogadták, hogy mögötte az ország (ill. a nemesség) ősi szabadsága elleni akció áll. Ez a magatartás a sz. kir. városok részéről kevésbé nyilvánult meg, a megyék pedig a megelőző évszázadok alatti fejlődésük szerint voltak jobb- vagy rosszabbindulatúak, pontosabbak vagy hanyagabbak. Eddigi tapasztalataink szerint a törvényhatóságok vagy (1) csak igen nagy késéssel

küldtek jelentést; (2) rendszertelenül vagy (3) egyáltalán nem voltak hajlandók a felvételeket, ill. jelentéseket elkészíteni. A Helytartótanács ilyen körülmények között csak kivételesen tudott országos összesítést készíteni, még kivételesebben olyat, amely valóban valamennyi törvényhatóság választát tartalmazta.

Csak mellékesen jegyzem meg, hogy a hiba nem egyedül a magyar törvényhatóságokban keresendő. mert egyrészt az ország közigazgatási szervezete nem volt alkalmas nagyobb felvételek és jelentések elkészítésére; másrészt a bécsi kormány olyan fogalmak alkalmazását erőszakolta, amelyek Magyarországon nem voltak meghonosodva.

Minderre szép példa az 1772—82 között (tehát kerekén 10 éven át) bekért lélekösszeírások esete, amelyekből csupán négy országos összesítés maradt fenn (1777, 1778, 1780 és 1782), de egyetlen egy sem volt teljes, a legteljesebb még az 1778. évi, mert abból mindössze egy megye (a „non coronat” mondással jellemzett Ugocsa) maradt ki.

Mindent egybevéve Mária Terézia erőfeszítései megtörték a magyar rendek közönyén vagy ellenállásán és komoly eredményt a bécsi kormány csak II. József alatt ért el, amikor néhány felvételt a magyar törvényhatóságok kikapcsolásával a katonasággal készítettett el (népszámlálás és kataszter).

A bajoknak ezzel még nincs vége, mert a múlt században a Helytartótanács számvevősé-  
gének irattárát kiselejtezték, tehát a ténylegesen felterjesztett jelentésekből is csak az maradt meg, ami a megyei levéltárakban ma rendelkezésre áll.

Végeredményben a történész ma gyakorlatilag töredékekből kénytelen dolgozni, ezekkel kell a nagy felvételeket ellenőrizni, a fejlődést rekonstruálni. A hagyományos történeti módszertan szabályai szerint a feladat a hiányok pótlása lenne más forrásokból vagy más forrásokra átvitt becslések útján, hiszen egy emberöltő előtt még az volt az uralkodó felfogás, hogy azonos célú megyei és városi vizsgálatok eredményének összege országos eredményt ad. Egy emberöltővel ezelőtt a statisztika is országos, tehát teljes felvételekre törekedett, és mert ezek már akkor is költségesek voltak, rendszeres időközökben azokat megismételni nem tudta — kivéve a népszámlálásokat. Ez a felfogás ma már a statisztikában és a történettudományban egyaránt meghaladott.

A fejlődés során a statisztika tett nagyobb és gyorsabb lépéseket. Kifejlődött és kifinomult a mintavétel módszere, amely meghatározott matematikai alapon tud kiválasztott részekből bizonyos valószínűséggel az egészre következtetni. Ilyen módszerrel készült „mikrocenzust” ma már évenként vesznek fel, nem kell a következő évtized népszámlálását megvárni.

Történeti kérdésekben is mintavétellel kell dolgozni. Ha egy emberöltővel ezelőtt azt vallották, hogy a részegységek (megyék stb.) hiánytalan összege adja meg az országos képet, akkor most azt kell mondani, hogy a minta összege ilyen meg ilyen valószínűséggel érvényes az egész országra. Ez az egyetlen mód, hogy belátható időn belül választ kapjunk bizonyos kérdésekre. Úgy hiszem pl., hogy a józsefi népszámlálás majdnem teljes anyaga alkalmat ad ilyen minta kidolgozására, ha nincsenek is illúzióim a nehézségekkel szemben. Nemcsak a ma szokásos szempontokat (pl. a helységek nagyságának rendjét) kell majd ui. tekintetbe venni, hanem a helységek korabeli jogállását, a lakosság nemzetiségét és vallását, a település gazdasági jellegét, a lakosság jogi helyzetét stb., azaz igen bonyolult és többszörösen rétegzett mintát kell majd venni. A mintavétel jelentőségét itt egyetlen — és ismét demográfiai — példával szeretném megvilágítani. Közismert a L. Henry által kezdeményezett családrekonstrukciós módszer, amely minden hátránya mellett ma a legjobb és legmegbízhatóbb alap bizonyos népmozgalmi mutatók kidolgozására. Ha ma tíz vagy húsz helységet vennénk vizsgálat alá, akkor nem tudjuk számításaink eredményét országos szempontból értékelni, másszóval nem tudjuk megmondani, hogy mennyire szignifikáns (történeti nyelven: mennyire tipikus) eredményünk a még ki nem számított országosan érvényes mutatók szempontjából. Ha viszont előre meghatározott minta szerint választjuk ki a vizsgálandó helységeket, akkor azok összessége — előre meghatározott valószínűséggel — egybeesik a keresett országos mutatóval.

A mintavétel során figyelembe kell venni a történeti statisztika adottságait. Sokszor elmondottam, hogy a statisztikát primér és szekundér statisztikára osztják fel aszerint, hogy az adatokat tisztán statisztikai célra gyűjtötték vagy más célt szolgálnak, de statisztikailag is feldolgozzák őket. Azt is sokszor hangsúlyoztam, hogy a történeti statisztika ebben a logikai koncepcióban terciér, mert primér és szekundér statisztika adatainak helyesbítésére mindig van mód, a történetiekére szinte soha nincsen. Ennek a felismerésnek a mintavétel szempontjából is megvannak a maga következményei. Ha egy teljes vagy majdnem teljes anyagból (pl. a józsefi népszámlálás) mintát veszünk, akkor távolról sem biztos, hogy a mintában szereplő mindegyik helységre lesz megfelelő forrásunk. A számítógépesnek kell megmondania, hogy a minta csökkenése milyen mértékben csökkenti a megállapítások valószínűségét vagy hogy lehet-e (és ha igen milyen mértékben) a mintában szereplő helységet más helységgel helyettesíteni. Bármilyen legyen a matematikus ítélete, a mintavétel lesz és marad a modern értelemben vett történeti statisztika alkalmazásának legfontosabb módszere és nem szorul bizonyításra, hogy ezt csak számítógéppel lehet megoldani.

2. A XVIII. század végi források helyzete nem mindenütt olyan reménytelen, mint ahogyan azt fentebb elég pesszimistán leírtam. A rendiség lényege a dualizmus volt, az uralkodó központi igazgatásának és az önkormányzatnak dualizmusa, együttműködésük vagy annak hiánya okozta az igazgatás eredményét vagy hiányát. De éppen ott, ahol erre az együttműködésre nem volt szükség, tehát az uralkodó központi igazgatásához tartozó gazdasági szakágakban a helyzet sokkal jobb volt. A kamarai igazgatás területén nemcsakbányászatrunkról, a vámok alapján a külkereskedelmi forgalomról és egy sor más fontos gazdasági tényről már Mária Terézia alatt rendszeresen ismétlődő, sokkal részletesebb és pontosabb kimutatások és jelentések készültek, mint amilyenek a megyék, ill. a Helytartótanács által küldöttek voltak. Pontosabb adminisztráció jellemezte a sz. kir. városokat és az egyházat — legalább is Batthyány primástól kezdve a katolikust. Ezekből a forrásokból igen sok jellemző vonást állapíthatunk meg.

Az ilyen természetű anyag még egy szempontból szerencsés, hogy ti. egyes kérdésekre vonatkozólag a gépi feldolgozás programja is készen rendelkezésre áll. Elsősorban egyes csoportok és társadalmi rétegek vizsgálatára gondolok, mint amilyen a városok vezetőrétege, a hivatalnokok vagy a papság közelebbi vizsgálata.

3. A számítógép használata új eljárást jelent, amelynek következményeit a történettudománynak kell viselnie. Elsősorban gyakorlati szempontokra gondolok, és ezek közül is csak néhányat ragadok ki, miután a legfontosabbat, a gondos tervezést és a több szakember bevonását jelentő kollektív munkát (team-work) már említettem.

a) A számítógépes munka fogalma magában hordja a tömegesen felhasznált adatok jellemzőjét, ez az egyik legvonzóbb vonása a történész számára, aki mindeddig lemondott a nagyvolumenű források feldolgozásáról. Igen sokszor beszélnek ezzel kapcsolatban gépi kapacitásról, s eközben nem gondolják meg, hogy az alapvető kérdés mindenképpen az anyag gépi feldolgozásra történő előkészítése. Amint 1848 előtt keletkezett forráshoz nyúlunk, a nehézségeknek se szeri se száma. Nehézséget jelent a latin, még nagyobb a német, főképpen ha lúdtollal és gót betűkkel írták. Azok, akik ma az előkészítés munkáját szakmai felkészültségüknél fogva vállalni tudják, általában nem tartoznak a legfiatalabb korosztályok közé. Ki kell alakítani azoknak a munkatársaknak körét, akik tisztában vannak a kor paleográfijával és nyelvével, és pontos munkát tudnak végezni. Enélkül kár gépi feldolgozásról beszélni.

b) A gépi feldolgozás fokozott forrásfeltárást kíván meg mind mennyiségi, mind minőségi téren. Forrásaink a XVIII. század végén — és még inkább a XIX. század elején — levéltáriak és könyvtáriak. A levéltári állagokról jó leltárak állnak rendelkezésre, de ezeknek csupán az a feladata, hogy általános tájékozódást nyújtsanak. Bármilyen gazdasági vagy társadalmi kérdés vizsgálatánál jó kiindulópontul szolgálnak, mégis elkerülhetetlen lesz, hogy az anyagról a feltett kérdés szempontjából részletes és még részletesebb jegyzék vagy katalógus készüljön. A levéltárnak sohasem lesz feladata, hogy az általa őrzött anyagról valamilyen kimerítő tárgymutatót

készítsen, ez még géppel sem megy. Hadd említsem meg itt a XVIII. századi magyar történelem művelésének egyik fájdalmas hiányosságát és hovatovább igen komoly akadályát, a bécsi levéltárak gyakorlati hozzáférhetetlenségét.

Teljesen elhanyagolt a nyomtatott anyagból a számunkra igen fontos hírlapirodalom (bővebben valamennyi időszakos kiadványt értve alatta). Még a XIX. század elején megjelenő hazai hírlapok lelőhelye sincs összegyűjtve, ami annál nagyobb kár, mert budapesti nagy könyvtáraink anyaga elképesztően foghíjas. Vannak olyan fontos hírlapjaink (pl. a Pressburger Zeitung), amelyek igazán megérdemelnének olyan repertóriumot, mint amilyent a Le Monde-ről a franciák adtak ki.

A számítógép ezen a téren jelentős segítséget nyújthat, növelheti esélyeinket; elég itt Borsa Ivánnak a középkori oklevéladók katalógusával kapcsolatos javaslatára utalnom.

## Wellmann Imre: Agrártörténeti forrásaink feldolgozása és a számítógép

### 1. A gépi adatfeldolgozás szükségessége

A történettudományon belül elsősorban a gazdaság- és társadalomtörténelem területén adott a kvantitatív analízis lehetősége és igénye. Érthető, hogy itt már hosszabb múltra tekintet vissza az a törekvés, hogy az „impresszionista történetírás” helyébe a múltnak mérhető tényekben kifejezhető rekonstrukciója lépjen. Eddig azonban ez a törekvés a hagyományos, manuális módszerekkel csak meglehetősen szűkkörű és kevésbé mélyreható eredményekre vezetett.

Különösen az agrártörténelem nem lehet meg gépi adatfeldolgozás nélkül. Nem tudományos divat kérdése ez, nem is csupán arról van szó, hogy a magunk tudományterületén a lehetőség szerint felzárkózzunk a rohamos technikai fejlődéshez. Mindenekelőtt arra kell utalni, hogy gépi adatfeldolgozásra alkalmas történeti forrásaink legnagyobbbrészt agrárfejlődésünkhöz kapcsolódnak. Következik ez abból, hogy Magyarország hosszú évszázadokon át lényegében agrárország volt (városaink élete is jelentős részben agrár alapokon nyugodott). Már ezeknek a forrásoknak a volumene és adatgazdagsága szükségessé teszi a számítógép igénybevételét mind dokumentáció, mind statisztikai jellegű elemzés céljára. Még inkább sürgeti ezt az utóbbi vonatkozásban az a tény, hogy az agrártörténelem, sajátos helyénél fogva a tudományok hierarchiájában, szorosan érintkezik szomszédos tudományokkal, mint néprajz, földrajz, agrártudomány, közgazdaságtan, szociológia — s ezért azok szempontjait figyelembe kell vennie. Ez pedig a problémák olyan komplex vizsgálatát teszi szükségessé, ami hagyományos módszerekkel nem végezhető el.

### 2. Mit vár agrártörténetírásunk a tervezett és rendszeres gépi adatfeldolgozástól?

Semmiesetre sem csupán a manuálisan elvégezhető munkafolyamatok meggyorsítását. Hanem egyfelől kutatási területének kiszélesítését eddig feltáratlan „fehér foltokra”. Másfelől vizsgálódásainak elmélyítését: az említett komplex összefüggéseknek s az egyes tényezők viszonylagos súlyának felderítését, a jelenségek és folyamatok belső struktúrájának és dinamizmusának lehető teljes feltárását. Hogy mind a két vonatkozásban reálisan megalapozott várakozásról van szó, bőven dokumentálják azoknak a témakutatásoknak már eddig is látható eredményei, melyekről külön beszámolókat fogunk hallani. Sőt ezenfelül a számítógépes elemzés a történeti forrás kritikájához, megfelelő értékeléséhez is segítséget nyújthat.

### 3. Az agrártörténész feladatai a számítógép segítségével eredményes felhasználása céljából

Nem térve ki arra az előfeltételre, hogy az agrártörténésznek meg kell ismerkednie a gépi adatfeldolgozáshoz kapcsolódó matematikai módszerekkel, szükséges, hogy a következőkkel járjon hozzá a számítógéptudománnyal való sikeres együttműködéshez.

a) Az alkalmas források kiválasztása. Kvantitatív elemzésre elsősorban olyan történeti kútfők kínálkoznak, amelyek vagy minél szélesebb területet ölelnek fel (lehetőleg az egész országot) egységes szempontok szerint, s akkor adott időpontban részletes és sokoldalú keresztmetszet megrajzolására alkalmasak; vagy pedig (s ez a ritkább eset) az idő dimenziójában lehetőleg messzeterjedő és hézagmentes sorozatokat alkotnak, ugyancsak egységes szempontok szerint, s így fejlődési trend megvonására nyújtanak alapot. Ezek a feltételek, mint ismeretes, annál kevésbé vannak adva, minél messzebb megyünk időben visszafelé.

Már ez az adottság azt javallja, hogy a munkával ott induljunk el, ahol az adatok legteljesebben és legszilárdabban megalapozottan állnak rendelkezésünkre; ez a történeti Magyarországot illetően a XIX. század vége, amikor már valóban statisztikai jellegű s homogén felvételekre építhetünk. De fel kell hívni a figyelmet arra is, hogy az agrárfejlődésben az eltérő természeti, piaci stb. adottságok igen sokféle táji, sőt helyi változatot okoztak, s ez a sokszínűség — a természeti tényező fokozott jelentősége miatt — általában növekvő mértékben érvényesült a preindusztriális korban. Ha valóban átfogó képet akarunk kapni agrárfejlődésünk összetevőiről, legcélszerűbb, ha a XIX. század végéről készítünk részletes keresztmetszetet. Ezen a szilárd talajon megvetve lábunkat, összehasonlító alapot nyerünk, s nagyobb biztonsággal haladhatunk időben vissza, hézagosanabb dokumentált korszakok felé (elsősorban az 1828-i országos összeírást, majd a XVIII. századi s még korábbi felhasználható forrásokat dolgozva fel), s ennek alapján módunk nyílik mind messzebb terjedő fejlődési trendek kimunkálására.

b) Forráskritika. A történész feladata annak beható vizsgálata, hogy a forrás egésze s egyes adatsorai mennyire megbízhatók, hézagai máshonnan vett adatokkal mennyire tölthetők ki s gépi feldolgozásra milyen vonatkozásokban és mennyiben alkalmasak.

c) A programozás előkészítése. A kvantitatív analízist kvalitatív analízisnek kell megelőznie, kutatási hipotézis felállításának, modell megalkotásának, mely a legalkalmasabb indikátorok megválasztását teszi lehetővé. Ezt követi a forrásadatok kódolásra alkalmas formában való prezentálása — s ez az a pont, hol elsősorban van szükség a matematikai módszerek említett ismeretére.

### 4. A továbbhaladás útja

Nagy feladatok állnak előttünk; csak az erők egyesítésével, tervszerűen, lépésről lépésre haladhatunk eredményesen előre.

Az eredményesség záloga mindenekelőtt az agrártörténészeknek a számítógéptudomány képviselőivel való — eddig csak elszórt egyéni próbálkozásokban megmutakozott — rendszeres együttműködésben rejlik. Ehhez egyrészt történeti problémákkal közelebbről megismerkedő számítógépes szakemberekre, másrészt a matematikai módszerekkel megbarátkozó agrártörténészekre van szükség. Így nemcsak szilárd és messzeterjedő forrásbázis kiépítése válik lehetővé, hanem — a hazai és külföldi tapasztalatok felhasználásával — az egész agrártörténeti kutató és feldolgozó munka eddig ismeretlen lendületet nyerhet.

Úgy gondolom, az agrártörténészek és közelebbről a Magyar Mezőgazdasági Múzeum részéről nem kell külön hangsúlyoznom és részleteznem mai kerekasztal konferenciánk jelentőségét.

Nagy jelentőségű már maga a tény, hogy számítógépes problémáinkat a Neumann János



Számítógéptudományi Társaság tagjai elé tárhatjuk. S ez a tény még mélyebb és konkrétabb értelmet nyer azoknak az erőfeszítéseknek a tükrében, melyeket már eddig is tettünk agrár-múltunk feltárására. Már a századforduló pozitívista kutatói nagymennyiségű nagyrészt kvantifikálható adatot hoztak napvilágra a levéltárak mélyéből. De minden terv és rendszer nélkül, eredeti összefüggéseiből kiragadva. Ugyanakkor meg sem próbálták ennek a nyers és szétszórt adattömegnek komplex vizsgálatát, összekapcsolását, a valóságos történeti összefüggésekbe való beágyazását. Agrártörténetírásunk sokáig ezt az adattömeget próbálta hasznosítani, értékelni és újraértékelni. De éppen kiszakított és töredezett voltánál fogva csak alkalmoszerűen, egy-egy részt kiragadva tudott értékesíteni belőle. Nagyobb összefüggések felderítésére, az egésznek megragadására ilyen módon természetesen nem kerülhetett sor.

Ebből a felismerésből kiindulva a Magyar Mezőgazdasági Múzeum a Magyar Tudományos Akadémia Agrártörténeti Bizottságának támogatásával új forrásanyagnak a kutatásba való bekapcsolásához fogott. Olyan átfogó jellegű kútfőkre kívánt építeni, melyek nagy volumenük-nél fogva a kutatókat addig általában visszariasztották rendszeres feldolgozásuktól. Az a véleményünk azonban, hogy csak ilyen, az egész ország területét vagy annak nagy részét átfogó forrásanyag rendszeres és módszeres értékesítésével tehetünk az agrártörténeti kutatásban komoly lépést előre. Ennek megfelelően láttunk neki nagy XIX., XVIII. és XVII—XVI. századi, agrártörténeti szempontból alapvető jelentőségű forráscsoportjaink adatainak feltárásához. S éppen a napvilágra hozott adattömeg mennyiségét, sokrétűségét, s a történeti élet más oldalai-val való sokágú kapcsolatát látva kellett arra a meggyőződésre jutnunk, hogy a számítógéptudomány segítségével nélkül komoly, nagy lépéssel nem juthatunk előrébb. Nem olyan felismerés volt ez, amely ölbetett kézzel való várakozást jelentett volna, arra, hogy külső segítség majd csak kizökkenti szekerünket a kátyúból. Magunk is próbáltunk a sűrűben utat vágni magunknak, amint kísérleteink mutatják; elég ebben a vonatkozásban Veress Éva kutatásaira utalni. Most épp azt szeretnénk igazolva látni, vajon helyes úton járunk-e, s támogatást szeretnénk kapni további erőfeszítéseinkhez.

Említettem, hogy XIX., XVIII., XVII—XVI. századi, agrártörténeti szempontból fontos nagy forráscsoportjaink kiaknázásához fogtunk hozzá. Azért használom ezt a sorrendet, mert ennek során arra a tapasztalatra kellett jutnunk, hogy minél inkább haladunk idővel visszafelé, annál nehezebb a dolgunk, annál hézagosabb, töredékesebb, annál kevesebb vonatkozásra vet világot a forrásanyag. Éppen ezért gondoljuk, hogy leghelyesebb a XIX. század végére nézve rekonstruálható állapotból kiindulnunk. De nemcsak ez a felismerés — ti. a forrásanyagoknak egyre jobban elszürkülő és hézagossá váló volta — vezet minket erre a megállapításra. Hanem egy másik, épp az agrártörténelem szempontjából igen jelentős felismerés is. Az ti., hogy az agrár-élet területén a földrajzi tényezőnek igen nagy a szerepe, a táji, a helyi változatos adottságok rendkívüli módon beleszólnak az agrárfejlődés alakulásába. Márpedig ezt a területi kivetítést, a táji és a helyi tényezők kellő értékelését csak a XIX. századból kiindulva tudjuk jól keresztülvinni. Hiszen kartográfiai vonatkozásban is ettől az időtől fogva rendelkezünk olyan megbízható alappal, melyen meg tudjuk szilárdan vetni lábunkat. Tehát az egyik, amit itt szeretnék hangsúlyozni, az időben visszafelé haladás módszerének jelentősége. Meggyőződésünk szerint ez a módszer jó alapot ad arra, hogy a korábbi, kevésbé összefüggő, tartalmilag, sokoldalúságban korlátozottabb forrásanyagot megfelelően tudjuk hasznosítani.

Itt azonban a régebbi agrártörténeti forrásanyagra vonatkozólag mindjárt egy kiegészítő megjegyzést kell tennem. Bármennyire szükszerűek és a mai statisztikus szemével tekintve primitívek ezek a korábbi források, viszont van egy pozitívumuk, amit éppen a számítógép alkalmazása szempontjából még nem tudtunk eléggé latba vetni, és ez a számadatokat kísérő szöveges rész, amely — merem állítani — sok esetben van olyan értékes, mint maga a számszerű anyag. Tovább kell tehát lépnünk azon az úton, hogy ezt a szöveges forrásanyagot a számítógép nyelvére át tudjuk tenni. Az időben visszafelé haladás módszere mellett ez a másik probléma, melyre itt külön fel kívánom hívni a figyelmet.

És harmadikként hadd utaljak vissza a területi kivetítés fontosságára az agrártörténeti kutatásban. Ebben a vonatkozásban is megpróbáltunk utat törni: úgy gondolom, a térkép-diagrammos módszer alkalmazása, melyet ugyancsak Veress Éva indított el, e részben is szélesebb távlatokat nyit előttünk.

A nagy és sokrétű problémákkal kapcsolatban tehát, melyek megoldásához a számítógéptudomány segítségét várjuk, a megközelítés útjai-módjai már kezdenek kirajzolódni előttünk. De nem árt még egyszer hangsúlyozni: az agrártörténésznek kutatómunkáját a matematikai statisztika irányában feltétlenül el kell mélyítenie ahhoz, hogy ezeken az utakon, a számítógéptudomány képviselővel szoros együttműködésben, bizton haladhasson előre.

## Katus László: A számítógép alkalmazásáról a kapitalista korszak kutatásában

### *Mi történt eddig? Mi van folyamatban?*

A felszólaló tudomása szerint eddig hazánkban e korra nézve két kutatást végeztek illetve végeznek számítógép segítségével:

1. A magyarországi gyáripar input-output táblájának összeállítása az 1898-as gyáripari statisztika adatai alapján. (KSH — OT; Simonics Györgyné: Magyarország közgazdaságának makro-ökonómiai összefüggései a XIX—XX. században. Történeti Statisztikai Évkönyv 1967—1968. Tört. Stat. Kötetek Bp. 1970.)

2. Az 1867 és 1914 közötti szántóföldi termelés statisztikai elemzése és nemzetközi összehasonlítása (MTA Történettudományi Intézet: Puskás Júlia—Lánc Margit). Mindkét kutatás továbbfejleszthető és kiterjeszthető a nemzetgazdaság egészére.

Meg kell említenünk, hogy Amerikában is végeztek egy számítógépes kutatást az 1882 és 1913 közötti magyar külkereskedelem adatai alapján (Scott M. Eddie).

### *A jövő lehetőségei*

Az 1850 és 1945 közötti korszak igen sok lehetőséget nyújt a számítógépes kutatásokra. Egy sor téma esetében a hagyományos technikával végzett kvantitatív kutatások már eljutottak addig a pontig, hogy folytatásuk és továbbfejlesztésük csak gépi eszközökkel lehetséges:

egyrészt sok olyan forrásunk illetve forrástípusunk van, amely számítógépes feldolgozást igényel;

másrészt az utóbbi évek kutatásai során egy sor olyan probléma merült fel, amelyekre a megnyugtató választ csak számítógéppel végzett statisztikai elemzéssel kaphatjuk meg.

1. *Források, forrástípusok*, amelyek nagyszámú (több ezer, vagy többtízezer) egyedi esetre vonatkozóan korlátozott számú, azonos típusú, könnyen standardizálható adatot közölnek. Ezek részben nyomtatott, részben levéltári források; címtárak (a legfontosabbak a gazdacímtárak), almanachok, cégjegyzékek, compassok, adózók jegyzékei, sorozási jegyzőkönyvek, anyakönyvek stb.; ide sorolhatók továbbá olyan terjedelmes források, mint a munkásmozgalomról szóló rendőri jelentések, a parlamentári naplók, valamint a genealógiai kézikönyvek és életrajzi lexikonok. Szükséges továbbá az egykorú hivatalos statisztikai kiadványok hatalmas adattömegének feldolgozása a modern kutatás által felvetett új ismérvek, új osztályozási szempontok szerint, valamint magasabb fokú statisztikai analízisük.

2. *Számítógépes feldolgozást igénylő problémákat* a történeti kutatás valamennyi területéről említhetünk. Csak azokra utalok, amelyek már ténylegesen napirendre kerültek, s hagyományos technikával már történtek kísérletek feldolgozásukra.

**Gazdaságtörténet:** a nemzetgazdaság növekedésének legfontosabb aggregát mutatói: a nemzeti termék ill. nemzeti jövedelem és összetevői, valamint az alakulásukra ható tényezők; ár- és bértörténeti munkák.

**Társadalomtörténet:** struktúraelemzések országos vagy helyi szinten, egész társadalmi keresztmetszetben vagy egyes osztályokra és rétegekre vonatkozóan; kollektív biográfia: egy-egy réteg vagy csoport rekonstrukciója az egyedi esetek nagy tömegének elemzése alapján; társadalmi tömegmozgalnak (pl. sztrájkok) elemzése; demográfiai kutatások: a főbb népesedési mutatók kidolgozása országos szinten.

**Politikatörténet:** a választói és a parlamenti magatartás és az azt befolyásoló tényezők vizsgálata.

**Kultúrtörténet:** országos szintű, hosszútávú oktatásügyi statisztika összeállítása; a könyvkiadás statisztikai elemzése; a közvéleményt befolyásoló ideológiai áramlatok kialakulásának, elterjedésének és hatásának elemzése a tömegkommunikációs eszközök nyújtotta adatok alapján (sajtó).

### *A kutatások jellege*

Első közelítésben kétféle típusú munka kerülhet szóba:

1. egy forrás, vagy azonos típusú forráscsoport teljes terjedelmében való leíró feldolgozása: az adatok kódolása, lyukkártyára, mágnesszalagra vitele, az alapvető osztályozási, csoportosítási munka és az alapfokú statisztikai feldolgozás (viszonyszámok, átlagok, indexek) elvégzésével (címtárak, compassok, és hasonló természetű források esetében).

Ilyen típusú források esetében egy központi computerarchívum létesítése (adatbank). A mágnesszalagon tárolt alapadatokat bármely kutató a saját problémáinak megfelelően további statisztikai feldolgozásnak vetheti alá.

2. Bizonyos problémák megválaszolásához a hagyományos megközelítés számára rejtett összefüggések és mozzanatok feltárása magasabbfokú statisztikai analízis segítségével (korrelációk, regressziók, faktoranalízis stb.).

Ebben az esetben a kérdésre vonatkozó többféle forrás kerülne elemzésre, többnyire mintavételes eljárással. A felsorolt javaslatok nem képviselnek meghatározott, konkrét munkaprogramot, hiszen valamennyi említett probléma kidolgozása még optimális gépi ellátottság esetében is több évtizedes feladatot jelent. Csupán jelezni kívántam azokat a területeket, amelyeken — külföldi kutatók konkrét tapasztalatain és eredményein lemérhetően — lehetőség kínálkozik a számítógépek alkalmazására. Az említett források, forráscsoportok zöme olyan természetű, hogy teljes feldolgozása hagyományos módszerekkel nem is lehetséges, vagy pedig a computeres feldolgozás olyan összefüggéseket hozhat napvilágra, amelyek különben rejtve maradnának. A nemzetközi tapasztalatok azt mutatják, hogy a felsorolt források és problémák esetében a számítógépek alkalmazásával járó — kétségtől mentes — többletráfordítás megtérül az eredménytöbbletben.

Írásban készített javaslatomat csak két megjegyzéssel szeretném kiegészíteni.

1. A javaslatban említett valamennyi probléma és téma olyan természetű, amelynek számítógépes feldolgozásával a külföldi szakirodalomban már találkozunk, tehát konkrét bizonyítékunk van arra, hogy ilyen munkák elvégzése lehetséges s olyan eredménnyel jár, amely érdemessé teszi a gépi feldolgozással járó ráfordításokat.

2. A javaslatban felsorolt feladatok közül a magam részéről elsődleges fontosságúnak tekinteném bizonyos források illetve forráscsoportok feldolgozását, egy központi computerarchívum létesítését, amely az illető forrásokkal dolgozó valamennyi kutatónak rendelkezésére állna.

## TÉMAREFERÁTUMOK

## Puskás Julianna—Lánc Margit: A magyarországi szántóföldi termelés kvantitatív elemzése 1867—1913

A Katus László referátumához kapcsolódóan ismertett munka célja a szántóföldi növénytermelés matematikai-statisztikai elemzése volt, és az ebből adódó gazdaságtörténeti következtetések levonása. A szerzők itt elsősorban metodikai tájékoztatást adtak.

Kiinduló forrásanyag: valamennyi szántóföldi növény évi termésének, termelő árainak és exportált mennyiségének idősorai. Ezen idősorok alapján a következő számításokat végezték:

1. Az idősorok elemzését a legegyszerűbb és legáltalánosabban alkalmazott módszerrel, *viszonyszámok* és *grafikonok* készítésével kezdték. A fejlődési tendenciák bemutatására mindenkeleltt indexsorokat alakítottak ki: *a)* bázisindexsort (mindenkor az 1909/13 évek átlaga alapján), *b)* láncindexsort. Az indexsorok nagy többsége volumen-index. A közvetlenül össze nem adható mennyiségekhez, a különböző minőségek közös nevezőre hozásához bevonták az értékindexet is. Az összes szántóföldi termelés és az egyes növénycsoportok termelésének alakulását változatlan árák felhasználásával kísérték nyomon. Az 1909/13 évek súlyozott termelési ára alapján, változatlan árákkal képzett indexsorok itt mint a különböző használati értékek közös nevezőre hozásának illetve az érték tényleges kifejezésének eszközei funkcionálnak és ugyancsak a termelés volumenének hordozói.

2. Az idősorok hosszútávú mozgását, a termelés fejlődésének alapirányzatát *trendszámításokkal* tárták fel. A legegyszerűbb trendszámításokból kiindulva, először ún. *mozgóátlagolással* számoltak trendeket. Az idősorok ingadozásainak így nyert kiegyenlítése már megkönnyítette az idősorok természetének megismerését; az így kapott tréngörbe viszonylag jól alkalmazkodott és jelezte a termelés növekedésének azon nagyobb szakaszait is, melyeket a szöveges történeti források is említene. Azt azonban nem tette még lehetővé ez a fajta trendszámítás, hogy az idősorok egészére vonatkozó, analitikusan ismert és könnyen kezelhető görbét kapjunk eredményül.

Ezért a mozgóátlagolású trendértékek idősorainak összeállítása és grafikonjaik megrajzolása után rátértek az *analitikus* trendszámításokra. Egyáltalán nem volt könnyű meghatározni, hogy az egyes idősorokhoz melyik trendtípus illeszkedik legjobban. Ehhez többféle módszert is alkalmaztak (tanulmányozták az idősorok grafikonjait, szemmértékre behúzták a középvonalat, vizsgálták a mozgóátlagolású trendgörbéket, figyelembe vették a hasonló típusú idősoroknál alkalmazott trendszámítási módszereket), s nem utolsó sorban tekintetbe vették a szöveges történelmi források információit is. Mindezen megfontolások együttesen segítettek hozzá a szerzőket azoknak a matematikai függvénytípusoknak megválasztásához, melyek az idősorok tartós irányzatát helyesen tükrözhetik. (A simulást a legkisebb négyzetek módszerével ellenőrizték. A termelés alakulását magukban foglaló idősorokra a legtöbb esetben lineáris, néhány esetben exponenciális trendek simulnak, az áraknál viszont minden esetben parabolák illesztése vált szükségessé.) Az analitikus trendszámítással nyert idősorok összeállítása után felrajzolták a trendvonalakat (fél logaritmikusk léptékben), s ez a lineáris trendeknek is egy elhajló görbe vonalát adta. Ez az ábrázolási mód nemcsak az idősoron behúzott középvonalat mutatta jól meg, hanem a nagy ingadozásokat is az eredeti és a trendértékből kialakított idősor között.

3. Annak kimutatására, hogy az eredeti idősorokat milyen mértékben kellett kiegyenlíteni, és a sorok valóságos adatai mennyiben térnek el a trendértékektől (vagy másnéven középértékektől), kiszámították az eredeti idősor adatainak deviálásait a trendértékek idősorának adatairól. Az évről-évre való eltéréseket a *deviációk* sora mutatja. (Az áttekinthetőség és összehasonlíthatóság kedvéért a trendértékektől való deviálásokat %-ban fejezték ki.)

Vizsgálatainkban a szerzők rendkívül nagy jelentőséget tulajdonítottak a deviációk kuta-

tásának, majd ezek alapján a szóródási mutatók kiszámításának, tudatában voltak ugyanis annak a matematikai absztrakciónak, a kézzelfogható valóságtól való elvonatkoztatásnak, melyet a középértékek azaz, trendértékek kiszámításával végeztek. (Tapasztalták azt is, hogy ez az absztrakció is egyik oka annak, hogy a gazdaságtörténészek idegenkednek a trendszámításoktól.) Véleményük szerint erre a matematikai absztrakeióra éppen a történelmi valóság mélyebb és pontosabb megismerése érdekében van szükség; a történészek számára a középértékek kiszámításának fő célja és indoka éppen az, hogy megvilágíthassák és történetileg analizálhassák az ettől való elhajlásokat, szóródásokat. (A modern statisztikai módszerek igazolják ugyanis, hogy a statisztikai sokaságok másképpen világítódznak meg, ha elhajlásaik, szóródásaik középértékét ismerjük.) Az erre irányuló számítások révén biztosabban tudják az idősorokban fellépő rendellenességeket vagy jellemző sajátosságait felismerni vagy bemutatni. A sorok középértékének ismerete, és az ettől való eltérések, szóródások vizsgálata révén a termelés olyan sajátosságait lehet megismerni, amelyek — ha nem járnának meg a matematikai statisztikai elemzésnek ezt az útját — rejtve maradnának.

4. A deviációk idősorának összeállítását követően a következő *szóródási mutatók* kiszámítása következett: a deviációk abszolút átlaga és a deviációk négyzetes átlaga, az ún. standard deviáció. Ez utóbbi kiküszöböli az egyes idősorok mennyiségi különbségeinek befolyását, és ennek következtében a legtisztább mutató az egyes idősorok összehasonlításához.

5. A deviációk idősora nemcsak az eredeti idősorok körüli ingadozásait jelezte, hanem kimutatta az eltérések szakaszosságát is. Ennek megragadására a deviációk idősoraiból (az eredeti idősorokéhoz hasonlóan 7 éves mozgóátlagolással) trendet számítottak, és a trendértékeket 100-nak véve azt a trendvonalat is grafikusán ábrázolták. Így még jobban meg lehetett közelíteni az idősorok természetét, mivel a hosszútávú trendvonalon belül feltárultak a közepes tartamú szakaszok, a gyorsuló és lassuló növekedési ütemű periódusok.

6. Végül az utolsó leszámaztatott illetve számított idősorok az úgynevezett *másodlagos deviációkat* tartalmazták. A deviációkból képzett trendértékeket 100-nak véve vizsgálták az évenkénti hullámzásokat, ingadozásokat az idősor rövidebb, ciklikus mozgásainak feltárása céljából.

A matematikai statisztikai számításokkal összeállított idősorok és grafikonok már széles bázist nyújtottak a fejlődés dinamizmusának a vizsgálatához is. Ennek mutatói közvetlenül az *átlagos növekedési ütemek*. Az idősorok egészére ezt kétféle módon számolták: 1. átlagos növekedési ütem trendértékekből, 2. az idősorok első és utolsó öt évének az átlagából. Az idősorok egészének átlagos növekedési üteme ugyancsak matematikai absztrakció. A valóságban a növekedési ütemek — mint már a láncindexsorok is jelezték — két évben sem voltak azonosak. A középérték, azaz az idősor egészének átlagos növekedési üteme történészek számára elsősorban azért fontos, hogy megtudjuk az ettől való eltérések mértékét, vizsgálhassuk a szóródásokat az átlagos növekedési ütemek körül.

A deviációk idősorából, főleg trendvonalából, már világossá vált, hogy a növekedési ütemek évenkénti változásai mellett bizonyos szakaszosságok is mutatkoznak, azaz az átlagos növekedési ütemek irányváltozásai, gyorsuló vagy lassuló szakaszai. Ezeket a szakaszokat kétféle megközelítésben is igyekeztek kimutatni: 1. átlagos növekedési ütemeket számoltak egymást átfedő tízéves periódusokra, valamint 2. azokra a szakaszokra, amelyeket a deviációk, mint a növekedési ütem irányváltozásait jeleztek. — Matematikai módszereket alkalmaztak a szántóföldi termelés szerkezeti változásainak megismerésére, elemzésére is. Kiszűrési módszerrel kerestek választ arra a kérdésre, hogy a növekedésre milyen mértékben hatott a három fő tényező: a vetésterület bővülése, a vetés szerkezet átalakulása és az egyes növények terméshozamainak emelkedése. Növényenként külön is elvégezték azokat a méréseket, hogy a vetésterület, illetve a terméshozam alakulása milyen mértékben befolyásolta az egyes növények termelésének az alakulását.

7. A nemzetközi összehasonlító vizsgálat, a magyarországi fejlődés pontosabb értékelése érdekében — a lehetőség határain belül — hasonló vizsgálatokat végeztek a szerzők több európai

ország szántóföldi termelésének idősorain is. Az ausztriai és magyarországi szántóföldi termelés összehasonlítása mellett a két adatsor együttes vizsgálata lehetővé tette a főbb mezőgazdasági termékekről hosszabbtávú trendek képzését, a növekedés vizsgálatát monarchia-keretekben is.

## Csernok Attila—Ehrlich Éva—Szilágyi György: Az infrastruktúráról történeti statisztikák alapján

(Nemzetközi összehasonlító vizsgálat 1860—1968)

Az infrastruktúra kialakulásában, társadalmi-gazdasági funkciójában érvényesülő nemzetközi-történeti tendenciák megismerése érdekében a következő kérdéseket vizsgáltuk:

- a különböző országok fejlődésében a gazdasági növekedés különböző szakaszaiban hogyan alakult az infrastruktúra;
- adott ország meghatározott gazdasági fejlettségi szintjéhez — a kor technikai színvonalát figyelembe véve — milyen kiépítettségi fok tartozik;
- meghatározott gazdasági fejlettségi szinteken „mennyibe kerül”, azaz mit követel a nemzetgazdaságtól az infrastruktúra kiépítése, fenntartása, korszerűsítése és bővítése.

Az infrastruktúra hosszabb történelmi időszakban történő vizsgálata nemcsak azért szükséges, mert a különböző országok infrastruktúrájának múltbani történelmi fejlődése és az ezirányú tapasztalatok csak más országokkal való összevetésük alapján értékelhetők helyesen, hanem azért is, mert az infrastruktúra

- területeinek egy része nagy tökeigényű és hosszabb megtérülésű;
- szerepe, tartalma és szerkezete történelmi időszakonként változó.

Az infrastruktúra nemzetközi-történeti vizsgálatát az elmúlt 100 éves periódusra — 1860—1968 — végeztük, amelyet 5 időszakra osztottunk:

1. az európai iparosodás kezdetének, illetve elterjedésének évei, amelyek már többé-kevésbé összefüggő adatokkal jellemezhetők: 1860—1870;
2. a századforduló körüli évek: 1885—1905;
3. az I. világháború és a világgazdasági válság közötti szakasz: 1925—1929;
4. a II. világháborút követő évek: 1950—1955;
5. a jelenlegi infrastruktúra jellemzésére szolgáló évek: 1965—1968.

*Az infrastruktúra fogalmáról:* Infrastruktúrán — elvi értelemben — a nemzeti vagyonnak azt a részét értjük, amely sem anyagi javak létrehozását, sem azok elfogyasztását közvetlenül nem szolgálja, de amely a gazdasági fejlettség adott szintjén a mindenkori technika követelményeinek megfelelően a termelés-elosztás-fogyasztás folyamatának zavartalan mozgásterét, ún. edényrendszerét hivatott biztosítani.

Az infrastruktúrába tartozó anyagi (felhalmozódott tőkejavak állománya) és szellemi (a munkaerőállomány; munkavégzőképessége, begyakorlottsága, munkakultúrája, alkotókészsége) tőkeállomány együttműködése a szélesesen értelmezett szolgáltatások nyújtásával, áramoltatásával szolgálja az anyagi javak és a munkaerő újratermelését.

*Mutatószámrendszer az infrastruktúra nemzetközi-történeti összehasonlítására:* Az infrastrukturális ellátottság nemzetközi-történeti összehasonlítására, az infrastruktúra fejlődési útvonalai felvázolásának céljaira a speciális történeti vizsgálatnak megfelelő módszert kellett kialakítanunk. Mivel az infrastruktúra érték, illetve pénzügyi mutatók segítségével nem vizsgálható reálisan, naturális mutatók alkalmazásához folyamodtunk. 5 olyan infrastrukturális terület volt, amely az első számításokban szóba jöhetett. (Közlekedés; hírközlés; lakásellátottság és felszereltség; egészségügy; oktatás és kultúra.) A területek jellemzésére naturális mutatószámok szolgálnak. Az idő előrehaladásával a mutatók száma történelmi, gazdasági, anyagi és technikai fejlődésnek megfelelően bővül.

Minden naturális mutató esetében azon ország adatát tekintettük bázisnak, ahol a kérdéses mutató értéke a legnagyobb. Az összes többi ország adatát mutatóként és területenként ehhez a bázishoz viszonyítottuk. Az így kapott % számokat országonként és területenként átlagoltuk, és ezen átlagokat, mint pontszámokat tekintettük az illető ország infrastrukturális színvonal mutatójának. A mutató lehetséges maximuma tehát 100. Értelmezése: az adott ország infrastruktúrájához viszonyítva, amely valamennyi elem tekintetében maximum értékkel rendelkezik. Az így kidolgozott összefoglaló és részterületeket jellemző infrastrukturális mutatót, mint egészet, *orientációs jelzőszámnak* tekintjük.

*Az iparosodás és az infrastrukturális fejlődés típusai. A számítási eredményekből 3 alapvető típus rajzolódott ki:*

*Az első típust* Anglia, Hollandia és egynéhány nyugateurópai ország testesíti meg. Itt az infrastruktúra kialakulása lényegében megelőzi a gyorsütemű ipari fejlődést. Ez az infrastruktúra nem az ipari forradalom, hanem az ezt megelőző gazdaságtörténeti időszak terméke. A korábbi időszakokban létrehozott infrastrukturális keretek a későbbi meginduló iparosodás számára is biztosították a fokozatos terjeszkedés, a működés kezdetleges mozgásterét.

Az Egyesült Államok gazdasága a fejlődésnek *egy másik, korszerűbb típusát* testesíti meg. Az USA az ipari forradalom idejében vált gyarmatból független állammá. Itt a kapitalizmus kialakulásával és terjedésével egyidőben, szinte előzmények nélkül, a termelő és fogyasztó ágazatok által támasztott szükségletek folyamatos kialakulásával jött létre az infrastruktúra. *Az angol modellre az időbeli megelőzés, az amerikai modellre a kapitalizmus terjedésével láncszerűen egybekapcsolódó infrastrukturális fejlődés a jellemző.*

Amikor az USA-ban az iparosodás kezdetét vette, ismeretes, hogy ez az iparosodás elsősorban a könnyűipari ágazatok (élelmiszerfeldolgozás stb.) fejlődésével kezdődött. E könnyűipari ágazatok a későbbi modern nehézipari ágazatokhoz képest viszonylag kis tőkeigényűek voltak. Az akkori iparosodást jelentő könnyűipari ágazatok fejlődése, kis tőkeigénye adott lehetőséget a nagy tőkeigényű és hosszú megtérülésű infrastruktúra fokozatos, az ipar szélesedésével és a fogyasztási szükségletek növekedésével együttjáró infrastrukturális fejlődésre.

Ehhez külföldi tőke is rendelkezésre állott. E külföldi tőke szerepe kettős volt: egyfelől megkönnyítette az infrastruktúra kiépítéséhez szükséges, néha időben egyszerűre jelentkező tőkék iránti igények kielégítését, másfelől elősegítette az akkori könnyűipari ágazatok és a mezőgazdaság által még nem elégséges mennyiségben termelődő tőke bizonyos fokú pótlását, kiegészítését. (A gazdaság egészének akkumulációs képessége ugyanis ezidőben még kicsi, a nemzeti jövedelemből a felhalmozás aránya 10% körüli).

Az iparosodás kezdeti időszakában az USA-ban az összes beruházásoknak kb. 35%-át költötték az ún. termelőágazatok fejlesztésére és mintegy 65%-át az infrastruktúrára. Ez arra mutat, hogy az ún. termelő beruházások így bizonyultak a leghatékonyabbaknak. Az iparosodás előrehaladásával a termelő ágazatok már folyamatosan kitermelték azt a tőkemennyiséget, amely képes volt biztosítani azt a mindenkori infrastrukturális kiépítettségi fokot a mindenkori technikának megfelelő színvonalon, amelyet az egész gazdaság folyamatosan szükségelt.

Az infrastrukturális területek bővülésével, modernizálásával kapcsolatos feladatok a két világháború közötti időben a nyugat-európai országok egy részében is már az USA típusához hasonlóan a láncszerű együttfejlődés folyamatában valósultak meg. Ennek keretében a nyugat-európai országokban az infrastruktúra tekintetében nagy szerkezeti változások mentek végbe. Ezek a szerkezeti változások elsősorban a nehézipar térhódításával kapcsolatosak.

A közép-keleteurópai országok egy részében a múlt század második felétől az I. világháborúig terjedő időszakban az infrastruktúra fejlődése ugyancsak megelőzte a nagyobbütemű iparosodás beindulását. Amíg azonban a nyugat-európai országokat az I. világháborúban területileg szinte változatlanul hagyta, ezekben az országokban óriási átrendeződéseket okozott. Az átrendeződések eredményeképpen bizonyos országok a korábbinál kedvezőbb infrastrukturális feltételek közé (pl.: Magyarország), mások kedvezőtlenebbé kerültek. Ezen országok nagy

része azonban lényegében kimaradt abból az infrastrukturális szerkezeti átalakulásból, amely ugyanezidőben Nyugat-Európa jelentős részében végbement. Ennek egyik jelentős oka, hogy ekkor ezekben az országokban a könnyűipar fejlődése (vagy a balkáni országokban még ez sem) volt napirenden.

A gazdaságfejlesztési típus, amelyet az európai szocialista országok a II. világháború után valósítottak meg, az infrastruktúra-kiépítés eddig említett két típusának egyikébe sem illenék bele. Erre a típusra sem az előzetes kiépítés, sem pedig a láncszerű együttfejlődés nem jellemző. Valamiféle olyan *harmadik típus*sal van itt dolgunk, amelyből az infrastruktúra-fejlesztés legalábbis egy időre mintegy elmaradt.

E sajátos harmadik típus kialakulásának okairól esupán néhány szót ejthetünk. Ellen-tétben az előző két típus képviselő országokkal a szocialista iparosodás megindulásakor ezek az országok az infrastruktúra fejlesztésére, bővítésére és modernizálására nem használtak fel külföldi tőkét. Ugyanakkor valamennyien szinte egyidőben, olyan iparfejlesztési politikát kezdtek, amely ellentétben a kapitalista országok iparosodásával nem a könnyűipart, hanem a nehézipart s általában a tőkeigényes ágazatokat helyezte előtérbe. Ilyen körülmények között nem lehetett szó arról, hogy ezek az országok egymásközötti tőkeáramlással segítsék a maguk infrastruktúrájának fejlesztését. Miután az új ipar kialakulása úgyszólván minden tőkeforrást igénybe vett, a másik nagy tőkefogyasztó, az infrastruktúra háttérbe szorult. A szocialista országokban a beruházások többségét, mintegy 60—70%-át a termelőágazatok kapták. A termelő-ágazatokba fektetett beruházások hatékonyságának közismerten alacsony volta miatt nem sikerült megtartani azt az alapbázist, amely folyamatosan lehetővé képes tenni az infrastruktúrára fordított beruházások gazdasági fejlettségnek, a gazdasági és fogyasztási struktúrájának megfelelő mennyiségét, minőségét, szerkezetét és színvonalát.

A nemzetközi történeti vizsgálat tapasztalatai szerint az infrastruktúra a szorosabb érte-lemben vett termelőtőke működésének képezi előfeltételét. Enélkül a termelő-beruházások sokkal kevésbé lennének hatékonyak, mert működésüket vagy a munkaerő, vagy a nyersanyag-ellátás, vagy a szállítás stb. keresztmetszetei gátolták.

Mialatt a másik két típusban az iparosodás első fázisában bizonyos külföldi segítséggel, a későbbiekben kizárólagosan saját belső tőkeerőből fejlesztették, bővítették és modernizálták az infrastruktúráját, a szocialista országokban az említett körülmények folytán is, a gazdaság-politikai koncepció és a termelőtőkék elégtelen hatékonysága következtében országonként változóan, általában azonban az 1960-as évek közepéig inkább a meglévő korszerűtlen infrastruktúra feléléséről beszélhetünk.

*Termelési tényezők* (munka, azaz a foglalkoztatottság) és a tőke (azaz a beruházások) szerepe az infrastruktúrában.

Az infrastruktúrában *foglalkoztatottak* összes foglalkoztatottakhoz viszonyított arány-számának alakulásában két, sok tekintetben összefüggő tényező — az idő és a gazdasági fejlettség — sajátos ötvöződése jut kifejezésre. E két tényező közül az utóbbi a döntő szerep. A köz-lekedésben pl.: az időtényező alapvetően módosítja a gazdasági fejlettség nyomán kialakult képet; a kereskedelemben esupán részleteiben befolyásolja; az egyéb ágazatban pedig az idő tényező szinte teljesen háttérbe szorul a fejlettségi tényezővel szemben.

A magas beruházás-igényesség az infrastruktúra egyik legfontosabb tulajdonsága. Az infrastrukturális *beruházások*nak az összes beruházásokhoz viszonyított aránya időtől és országtól függően széles intervallumban ingadozik, de a legritkább esetben megy 50%-ó alá, felső határa pedig 80%-ó körül van.

Az infrastruktúra beruházásainak másik jellegzetessége, hogy nagysága általában nem függ a gazdasági fejlettségtől.

Az iparosodásnak abban a szakaszában, amely a múlt század első felétől játszódott le, a két legfontosabb infrastrukturális tényezőre: a lakásépítésre és a vasútra fordított beruházások együttesen az összes beruházások 60—70, vagy 80%-ó-áig mentek fel. A századforduló táján az.



infrastrukturális beruházásokkal szemben nagyobb szerepet kaptak a „termelő ágazatok” beruházásai. A II. világháború után az infrastrukturális beruházások újra előtérbe kerültek és súlyuk megnövekedett. Arányuk a legtöbb kapitalista országban 65—75%-ra nőtt. A szocialista országokban az infrastrukturális beruházások aránya az összes beruházásokhoz képest jóval kisebb, mint a kapitalista országokban.

Az infrastruktúra fejlesztése lassú folyamat s általában rendkívüli erőfeszítéseket követel a nemzetgazdaságtól.

### Varga Dénes—Veress Éva: Feudáliskori birtokösszeírások számítógép segítségével történő feltárása

A számítógépek alkalmazásának többféle lehetőségét veti föl *Granasztói György* tanulmányában. Ő elsősorban a statisztikai számítások lehetőségeit boncolgatja, felsorakoztatva az értékelésükkel (és túlértékelésükkel) kapcsolatos véleményeket és ellenvéleményeket. A számítógépes statisztikázással kapcsolatban természetesen mindazt a jót és rosszat el lehet mondani, amit általában a statisztikával: ha az anyag nem homogén, ha a kiválasztott minta nem elég reprezentatív vagy a kvantitatív vizsgálatokban impliciten bennerejlő kvalitatív szempontok nem helyesen tükrözik a valóságot, akkor a kapott eredmények nemhogy triviálisak, hanem egyenesen torzak lehetnek. Nem véletlen azonban, hogy a matematikai statisztika nemcsak az „átlag” fogalmának általánosításából született *várható érték* fogalmával dolgozik, hanem ezzel együtt az „átlagtól való eltérés” köznap kifejezéssel jelölhető *szórás* fogalmával. A lényeges eltérések egyedi eseteinek értékelésében természetesen tág tere nyílik a kvalitatív vizsgálatoknak. Nem szükségszerű tehát, hogy a statisztika „szűrkitse” a valóságot — ha a történész jól használja föl, eszköznek és nem célnak tekinti, többet tudhat meg belőle a valóságról. Nem biztos persze, hogy mindig több jön ki, mint amit amúgy is tudunk. A számítógépes statisztikázásnak éppen az lehet a haszna, hogy a szempontok gazdagságával sok olyan összehasonlítást is lehet tenni, amire az emberi energiából nem futná, és szemléletes ábrák segítségével könnyebben kiválogathatjuk a nem triviális eredményeket a triviálisak közül. Természetesen az sem lebecsülendő eredmény, ha valamilyen, addig csak igaznak érzett állítást tényekkel alá tudunk támasztani vagy kvantitatív megállapításokkal finomítani tudunk. Addig ugyanis más ugyanúgy igaznak érezheti az állítás ellenkezőjét. Ismeretes, hogy Leibniz-et éppen az a kettősség döbbsentette meg annak idején, hogy a matematikusok vitájában mindig kiderül, hogy kinek volt igaza, a történészek, irodalmárok vitáiban azonban egymás mellett megférnek homlokegyenest ellenkező állítások.

Mi itt a számítógépek humán célú alkalmazásának egy másik területével szeretnénk foglalkozni egy konkrét munka kapcsán. Egyelőre még csak az előkészületek folynak, a munka volumene és úttörő jellege mégis indokoltá teszi, hogy már most a történész és számítógépes szakemberek elé tárjuk a feldolgozás tervét.

Dokumentációs jellegű feldolgozásról van szó, melynek az volna a célja, hogy egyébként áttekinthetetlen, egészében értékelhetetlen anyagot tegyen a történész kutatók számára hozzáférhetővé. A gépi feldolgozást nemcsak az anyag nagy volumene és elismerten gazdag tartalma indokolja, hanem az is, hogy várhatóan sok olyan tevékenységet is át lehet hárítani a gépre, ami az ember számára roppant időrabló feladat volna. Az egyedi utánanézését kívánó döntések meghozatala most is az emberre vár, a „döntéselőkészítésben” azonban ilyenkor is komoly segítséget nyújthat a gép. A gép által hozott döntések felülbírálásának joga természetesen szintén a szakemberé.

### 1. Kiinduló feltételek

Sok tízezer, különböző célokra készült birtokösszeírás maradt ránk az 1526—1848 közötti időből. (Szórványosan vannak feljegyzések a mohácsi vész előttről is, de ezek száma viszonylag elenyésző.) A Mezőgazdasági Múzeum kutató kollektívájának kezdeményezésére és részben munkájával ezekből az anyagokból tartalmi kivonatok készültek. az eredeti dokumentumok bizonyos adatainak pontos megtartásával (pl. az előforduló helynevek betűszerinti kimásolásával). A kivonatkészítés elvei érthető módon nem voltak teljesen egységesek, hiszen ezek részben a menetközben jelentkező problémák szerint módosultak, s természetesen készítőjük egyéni felfogását is tükrözik — bármennyire törekedtek is a munka megszervezői egységes szempontok érvényesítésére.

A feldolgozásnak ez az első lépése rövideen befejeződik. Az anyag azonban még így is olyan roppant nagy, hogy kézi erővel inkább csak válogatni lehet benne. Voltak kísérletek manuális dokumentációs eszközök (peremlyukkártya, kézi mutatók) készítésére, ezek a próbálkozások azonban — érthető módon — nem hozták meg a kívánt eredményt. A gépi feldolgozáshoz természetesen komoly előkészületek van szükség. Maga az anyag felvitele megfelelő adathordozóra többhónapos munkát kíván, érdemes tehát jóelőre megfontolni: mi minden hasznos dologt lehet kipróbálni az előkészített anyagból, ha egyszer „ráeresztjük” a gépet. A gép előnye éppen abban rejlik az emberi válogatással, rendezgetésekkel szemben, hogy ugyanazt az anyagot a legkülönbözőbb szempontok szerint feldolgozhatjuk, sok oldalról hozzáférhetővé tehetjük a kutatók számára.

Mindez természetesen nem zárja ki azt, hogy az anyag elsődleges szétválogatását és csoportosítását ne kézzel végezzük. Ameddig nem egymást keresztező, hanem egymásnak alárendelt szempontok érvényesítéséről van szó („egymásba skatulyázott”, hierarchikus csoportosításról, „osztályozásról” a szó klasszikus értelmében), addig ez nem okoz különösebb problémát. Az adott esetben ezt a megközelítést két okból látjuk indokoltnak. Egyrészt azért, mert az anyag egyes részhalmazaiból már előbb, a teljes anyag gépi feldolgozása (és nem utolsósorban: lelyukasztása) előtt eredményeket kaphatunk, ennek tanulságait tehát még időben hasznosítani tudjuk (maguk a kapott eredmények természetesen önmagukban is értékesek lehetnek). Másrészt ezáltal lehetővé válik bizonyos kiegészítő információk gazdaságos bevitele, mikor azok nem egy-egy egyedi dokumentumhoz, hanem a dokumentumoknak egy egész csoportjához tartoznak. Az ilyen kiegészítő információk azután a gépen belül már minden érintett dokumentumhoz egyedileg hozzáférhetőek, s az tetszőleges csoportosításnál „viheti magával” valamennyi jellemző jegyét.

Ilyen előzetes kézi osztályozás szempontjai például a következők lehetnek:

1. A felhasznált források
2. Történelmi korszakok
3. Területi egységek
4. Birtokosok

Számunkra közömbös, hogy természetüknél fogva az egész anyagra vonatkoztatva ezek a szempontok „kereszteznek egymást”. Mesterséges hierarchiát állíthatunk fel a szempontok között: a kiválasztott elsődleges szempont szerint készíttjük el az elsődleges felosztást, majd az így kapott csoportokat bontjuk tovább a másodlagos szempont szerint és így tovább. A gépen belül később összevonhatjuk a különböző csoportok azonos jellegű alcsoportjait, s egy másféle osztályozás szerint elsődleges szempont lehet az, ami másutt másodrangú vagy negyedrangú volt.

Nehezebb problémát jelent a kiválasztott szempontok megfelelő érvényesítése. A határesetek száma néhol igen nagy lehet. Pl. a területi egységek szerinti bontásnál problémát jelent egyrészt a több helyre is besorolható dokumentumokról való döntés (legcélszerűbbnek látszik xeroxmásolatot készíteni róluk és valóban több helyre besorolni), másrészt a területi szempontnak az időtől való függése: az átfogott nagy időszak alatt a megyehatárok, uradalmak kiterjedése sokat változott, tehát még a hovatarozásra vonatkozó explicit információk elfogadása is proble-

matikus. Valamit javít a helyzeten, ha csak nagyobb egységeket választunk külön (pl. Dunántúl), egyébként azonban tudomásul kell vennünk, hogy adataink a határesetekre vonatkozóan kevésbé lesznek megbízhatóak. Mint a fényképezésben: az „időre” exponált képnél a változó részletek természetes, hogy elmozdulnak.

Az anyag legstabilabb, időben leglassabban változó objektumai maguk a dokumentumokban említett városok és falvak, az ezekre vonatkozó információk azonban, a dokumentumokban előforduló helynevek, a legváltozatosabb képet mutatják. Lényegében ez jelenti a gépi feldolgozás számára a legérdekesebb problémát: hogy lehet az alakilag különböző változatokban felismerni az azonos helynevet, illetőleg a formailag azonos helynevekről megállapítani, hogy az adott kontextusban melyik fizikai objektumra vonatkoznak. Ez a probléma bizonyos mértékig az alakfelismerés nyelvi analogonja, abban a formájában, amikor az etalon (pl. a jelenleg érvényes helységnévtár formájában) előre adva van.

## 2. A feldolgozás elveinek és céljainak tisztázása

A feldolgozás során véleményünk szerint a következő alapelvek követése kívánatos:

1. Tekintsük a feldolgozás szempontjainak halmazát nyíltnak és minden olyan információt vigyünk fel az adathordozóra, ami a feldolgozás folyamán valamikor szükséges lehet. Ezt a hozzáállást az anyag nagy terjedelme indokolja: nagy pluszmunkát jelentene, ha később kiderülne, hogy az eredetileg eltervezett rendszer szűknek bizonyult.

2. Lépcsőzetes feldolgozásra és használható részeredmények elérésére törekedjünk, hogy a menetközben szerzett tapasztalatokat bele tudjuk építeni a rendszerbe. Az ember-gép együttműködésnek ez a formája mindkét „fél” számára hasznos: az emberből újabb megvalósítható ötleteket válthatnak ki az elért eredmények, a gépi programban pedig finomítani lehet az alkalmazott szabályokat, pontosabban be lehet állítani az eredmények és hibák elemzése alapján a megfelelő paramétereket.

3. A tapasztalatok gyűjtését lehetőleg automatizáljuk a programban, hogy azonos feltételek ismételt előfordulása esetén ne kelljen mindig végigjárni az azonosítás viszonylag hosszadalmas és bonyolult útját. Ugyanakkor azt is tegyük lehetővé, hogy a kutató felülbírálhassa a gép által hozott döntéseket, nehogy esetleges téves asszociációk láncá indulhasson el.

4. Készíttessünk sok szemléletes outpoutot, összesítést az elért eredményekről, megkönnyítve ezzel az eredmények gyors egybevetését és értékelését és egyúttal a hibák időben való felderítését.

Milyen konkrét célokat tűzhetünk magunk elé az anyag feldolgozásában? A teljesség igénye nélkül — inkább csak a történéssz kollégák fantáziájának megmozgatása kedvéért — felsorolunk néhány lehetőséget.

1. A kutatás megkönnyítésére különböző típusú indexeket készíthetünk a lelőhely és a dokumentumra vonatkozó pontos információk megadásával. Lehetséges szempontok: az adatok jellege, területi egységek (helységek, megyék, uradalmak), birtokosok, többnyelvűség, az előforduló nevek, dátumok stb.

2. Speciális célokra (részben az előbb felsorolt szempontok különböző kombinálása révén) különleges összeállításokat készíthetünk:

- helységnévtárt az előforduló alakváltozatok felsorolásával;
- összesítéseket, rendezéseket, számszerű összehasonlításokat különböző csoportosításban (pl. különböző történeti korszakokra vonatkozóan);
- regisztrálhatjuk a végbement változásokat (uradalmak, megyék határainak módosulása, falvak elnéptelenedése stb.).

3. Térképszerűen ábrázolhatjuk az egy-egy időszakra vonatkozó statikus képet. A szemléletes ábrázolás felszínre hozhat egyébként nehezen felfedezhető összefüggéseket és jó segédeszközt jelenthet a kutatóknak.

Ez az ábrázolási mód alkalmazható a gép által hozott döntések könnyű értékeléséhez is (hol helyezkednek el azok a helységek, amelyek a helynevek alaki hasonlósága alapján szóba jöhetnek vagy hol vannak az országban azonos nevű helységek).

### 3. Az anyag empirikus vizsgálata

A tipikus dokumentációs jellegű problémák megoldási módszereiről most nem kívánunk beszélni (tárolás, indexek készítése stb.). Ezek elvi nehézséget nem jelentenek, csak gyakorlati megfontolásokat kívánnak. Elvileg is érdekes probléma azonban a földrajzi nevek azonosításának kérdése: hogy lehet a teljes, betűszerinti azonosság vizsgálata helyett egyéb azonosítási módokat találni, amelyek olyan változatokra alkalmazva is megfelelő eredményt adnak, amelyek időtől, helytől, nyelvjárástól, többé-kevésbé következetes helyesírási elvektől is függenek, vagy a szó szoros értelmében elírások, torzulások. E változatok sokfélesége nyilván nem sorolható fel lista alakjában, nem határozható meg egyértelműen szabályok segítségével. De az sem volna elég, ha meghatároznánk a lehetséges torzulások körét. Szükség van a torzulás *mértékének* bevezetésére is, hiszen egy nagyobb torzulást feltételező azonosítást csak akkor fogunk elfogadni, ha jobb lehetőség nincs.

A torzulások típusainak megállapítása és a fokozatok értékelése céljából empirikus vizsgálatokat végeztünk az anyagból vett minta összehasonlító elemzése alapján. Intuitíve kiválasztottuk a feltehetően ekvivalens helyneveket és formái egybevetésük alapján igyekeztünk kiválasztani és osztályozni az eltérő jegyeket, s ezek alapján fokozati különbségeket állapítottunk meg az eltérések között. Az összehasonlítás főbb tanulságai a következők voltak:

1. A szókezdő hangok a legritkább esetben változtak meg. Helyesírási variánsok előfordultak (z — sz), ezek osztályai azonban pontosan megállapíthatók.

2. A szó belsejében a mássalhangzó hanganyag jellegzetes elemei megmaradtak, ezek azonban minőségben is, sorrendben is gyakran mentek át kisebb változásokon. E változások tipikus esetei:

a) zöngés—zöngétlen átcsapás:

Filesd — Filest

Nábrád — Naprad

Tarbocz — Tardocz — Tartolcz

Terep — Terép — Tereb

Tiszabereg — Tiszaberek stb.

(Formailag ehhez hasonló a zöngés—zöngétlen különbség helyesírási összemosódása:

Szopor — Zopor

Székelyhida — Zekelhida

Asszonyvására — Azzonvasara.)

b) kemény—lágy (veláris—palatális) átcsapás, illetőleg ezek helyesírásiilag nem következetes szétválasztása (itt ezt a két altípust már nem is nagyon lehet szétválasztani):

Szinér — Szinyér

Szenfalva — Szényfalva — Színfalva — Zenfalva

Nires — Nyires

Komorzan — Komorzany — Momorzán — Komorzány

Csahol — Csaholy stb.

c) egyéb, fonctikailag közelálló hangok átcsapása egymásba:

Namtő — Nantő

Mogyorós — Monyorós

Nábrád — Nábrák

Piskárkos — Piskartos

Szókezdő hangra is találtunk ilyen példát:

Megrefalva — Negrefalva.

d) hosszúság—rövidség különbségének elmosódása:

Palád — Pallád — Palad — Pallad

Iloba — Illoba

Aszó — Asszó — Asso

Szatmar — Zatmar — Zattmar

Turvékonya — Turvékonya — Turvikonnya

e) hiányzik a következőes helyesírási különbségtétel a g és gy között:

Gyöng — Gyöngy — Göngy

Lophag — Lophagy

Geres — Géres — Gyéres

Gürkefalva — Gyürkefalva stb.

3. A hangállomány bővülésének vagy szűkülésének, asszimilációjának vagy disszimilációjának tipikus esetei:

a) l vagy r betoldása ill. elhagyása o és mássalhangzó között (vö. nyolc — nyóc, volt — vót) vagy szó végén:

Csaholcz — Csahocz

Boldad — Bodad

Darocz — Darocs — Darolcz

Fernező — Fernezély — Fernezer

b) mássalhangzótorlódás bővülése vagy szűkülése:

Porcsalma — Porcsalma

Bajfalva — Bajkfalva

Dobrovicza — Dombrovicza — Dobrapicza

Bujánháza — Burjánháza — Bojánháza

Nantő — Namtő — Nantdő

c) asszimiláció — disszimiláció:

Ballafalva — Barlafalva — Barllafalva

Lippe — Lippo — Lidpo

(inkább csak helyesírási többféleség a gy esetében:

Meggyes — Medgyes).

4. A magánhangzó hanganyag még változékonyabb, kevesebb a stabil elem:

a) hangtörténeti, helyesírási és nyelvjárási okok kevert hatása következtében számtalan esetben váltakozik egymással

a és o

o és u

ő és ü

i, e és é, pl.:

Ambod — Ombod — Ombád

Mogyorós — Mogyorus

Namtő — Nantü

Namény — Nameny — Naminy

Szenyár — Szinyér stb.

A nyiltabbaválás tendenciája nem egyformán érvényesül:

Mogyorus — Magyaros

Namtü — Namtő,

de:

Filpes — Fülpes — Fülpös.

b) Sok következetlenség forrása a speciálisan magyar hangok (é, ö, ő, ú, ü, u) jelölésének megoldatlansága az akkori írásban:

Gyiwrow — Gyürö stb.

— a példák száma tetszés szerint szaporítható.

c) A magánhangzó kiesésének két fő típusa:

a szó végén

Berenche — Berencze — Berencz

Lehencze — Lehoncze — Lehencz

Borhida — Borhid

Chege — Czegh

Németi — Német

a szó közepén:

Esztero — Esztro

Rakosa — Raksa.

5. Az előtagok (Kun-, Sziget-, Dráva-, Szamos- stb.) elmaradása sok formailag megoldhatatlan homonímia forrása. Ezek elkülönítéséhez elsősorban a kontextus alapján meghatározott földrajzi index jelenthet fogdózt.

#### 4. A probléma számítógépes megközelítése

A tanulságok alapján egy többlépcsős azonosítási rendszer tervét dolgoztuk ki.

Az *első lépcsőben* betűszerinti összehasonlítás folyik egy állandóan bővülő, valamennyi eddig előfordult variánst tartalmazó helységnévtár alapján. Amennyiben a rendelkezésünkre álló információk megengedik, a teljes helységnévtár mellett kisebb egységekről (pl. megyékről) külön helységnévtárak készülnek. A határesetek kezelése érdekében ezek nem diszjunkt halmazt képeznének, hanem minden egyes megyéhez hozzá tartoznának az időben valamikor a megyéhez kapcsolt valamennyi település (a megfelelő halmazok logikai összegét vennénk tehát).

Lényeges információként használnánk fel az azonosításban a helyneveken kívül a földrajzi mikro- és makrokörnyezet fogalmát. Ez annyit jelent, hogy csak a már addig előfordult helynevek (ill. megyék) megfelelő objektumaitól „nem nagyon messze” levő objektumokat fogunk nevük alapján azonosítani egy-egy dokumentum alapján. Kivételt jelent az, ha az újonnan megjelent objektumok *egymástól* „nem nagyon messze” helyezkednek el. Feltételezzük tehát, hogy egy-egy dokumentumban a helynevek „foltokban” elhelyezkedő helységekre utalnak. (Az ettől való eltérést az emberi kontroll beavatkozása esetén természetesen megengedjük, ehhez a gép megfelelő térképes információkkal nyújt segítséget.) Egyelőre félretesszük azokat a helyneveket, amelyek betűszerinti egyeztetéssel nem azonosíthatók. Közben azonban fokozatosan bővül a mikro- (helynevekre vonatkozó) és makro- (megyére vonatkozó) környezet aktuális fogalma, a program tehát „tanul” addigi tapasztalataiból.

A *második lépcsőben* azoknak a helyneveknek a vizsgálatára kerül sor, amelyeket az első lépcsőben nem sikerült azonosítani, felhasználva ehhez az előző lépcsőben kialakított mikro- és makrokörnyezetet. A keresés most olyan különleges elrendezésű helynévtár alapján történik, amelyben a „keresőkules”

a) a kezdőbetű és ezen belül

b) a hangtörzs

alapján rendezve tartalmazza a helyneveket (és a hozzájuk tartozó földrajzi információkat). Ez annyit jelent, hogy minden helységnévhez hozzárendelünk a kezdőbetűn kívül egy egy-két byte-os „szinképet”, amely a szó belsejében előforduló hangok ill. betűk főbb jellegzetességeiről ad információkat. Ez a részletezés nem terjed ki a hanganyag olyan megkülönböztető jegyeire, amelyek — empirikus vizsgálataink tanulságai szerint — egyébként is gyakran összemósódnak. Közös egységet jelentenek például a következő hangok:

z — sz — zs — s	(dentális és alveoláris réshangok)
v — f	(labiodentális réshangok)
gy — g — k	(paleoveláris zárhangok, ill. lágy dentális zárhang)
m — n — ny	(nazális zárhangok)
r — l — ly — j	(középujt képzett réshangok)
c — cs	(zöngétlen affrikáták)
d — t	(dentális zárhangok)
h	(laringális réshang)

Összevonásaink a Trubeckoj-féle fonológiában használatos „archifonémákra” emlékeztetnek, mivel azonban az írásképpen bekövetkező összeosódásokat is figyelembe vettük, inkább „archigrafémáknak” nevezhetnénk őket.

„Archigrafémáink” számát, sorrendjét a második lépcsőben figyelmen kívül hagyjuk, csupán az előfordulásukra vagyunk tekintettel — ott viszont az „archifonémáknak” megfelelő fonémacsoportok *jellegetessége* szerinti súlyozással vesszük őket figyelembe. A kevésbé jellegzetes hangokat jelölő betűk kiesése így kevésbé zavarja rendszerünket.

A második lépcsőben tehát kiválasztjuk azokat a helyneveket a fent leírt elv szerint rendezett helynégtárból, amelyek

- a) hanganyag
- b) földrajzi információk

szerint legközelebb vannak a keresett helynévhez, illetve az előző lépcsőben kialakított földrajzi környezethez.

A *harmadik lépcsőben* az előzőekben kiválasztott helyneveket rendezzük, most már figyelembe véve „archigrafémáink” előfordulási sorrendjét is a vizsgált helynévben, legvalószínűbbnek minősítve azt a helynevet, melynek archigraféma struktúrája egyáltalán nem, vagy csak kevésbé tér el a vizsgált helynév hasonló struktúrájától, s földrajzilag sem fekszik a megfelelő település, az eddig azonosított környezettől nagyon távol. (E vizsgálatok közben természetesen bővül a földrajzi környezet aktuális fogalma is.)

Végül a *negyedik lépcsőben* a „gyanúba vett” helynevek részletes, lépésről lépésre való azonosítása megy végbe, visszatérve az „archigraféma” szintről a konkrét betűelőfordulás szintjére. „Mozaikkép”-szerűen végigpróbálhatjuk az előforduló betű vagy betűcsoport lehetséges variánsait, a dokumentum keletkezési idejének és az eddig figyelmen kívül hagyott magánhangzóstruktúrájának a figyelembevételével. (Az alkalmazandó szabályok és megszorítások természetesen még közelebbi nyelvészeti vizsgálatokat követelnek).

A vázolt rendszer alkalmas arra, hogy a Salamon—Szolimán—Szulejmán-szerű változatok azonosságát felismerje, s a már felismert helynév-variánsokat úgy építse bele a rendszerbe, hogy legközelebb már az első lépcsőben hasznosítani tudja a szerzett tapasztalatokat. Gépi idő tekintetében egyébként is nagyon takarékos a rendszer, mert fokozatosan szűkíti a szóba jöhető helynevek körét, s csak annyi változat vizsgálatára használja teljes arzenálját, amennyire feltétlenül szüksége van.

## Cseh László — Haraszti György — Veress Éva: Térképdiagrammok előállítására számítógéppel és az adatok tárolásának-feldolgozásának topográfiai módszere az agrártörténetben

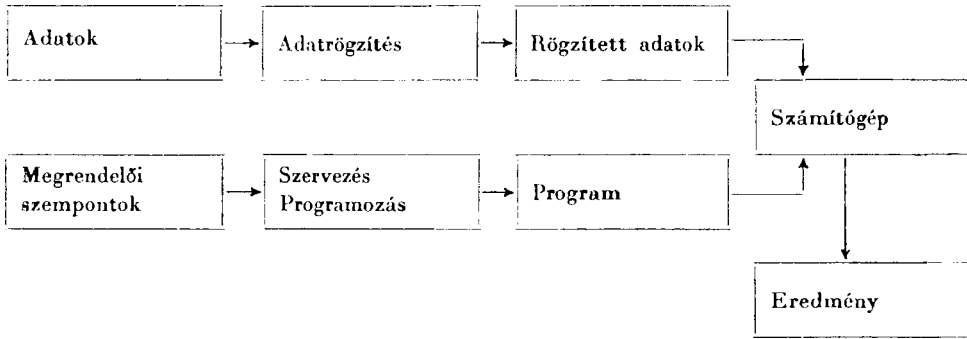
### I. Térképdiagrammok előállítása számítógép segítségével

#### 1. Elvi megfontolások

*Számítógépes feldolgozás szempontjából* (a kérdést a gép oldaláról közelítve) értelmetlen az adatok tematikus megkülönböztetése, szaktudományok szerinti osztályokba sorolása.

A számítógép csak meghatározott műveletek végrehajtására képes, a világ anyagi egységének kifejezőjeként is tekinthetjük, mindenesetre olyan dolog, amely teljesen érzéketlen a felépítmény finomabb árnyalataival szemben.

A számítógép a megrendelő által szolgáltatott adatokat a megrendelő szempontjait kielégítő program utasításai szerint feldolgozza, esetleg a kapott eredményeket kiértékeli:



A megrendelő igényeinek a gép oldaláról határt szab a gép kapacitása, az alkalmazható programnyelvek sajátosságai s végül de nem utoljára — hazai viszonyok közt ez lényegi szempont — a „gépre jutás” (felhasználói költségek, gépidő stb.).

A lényegi probléma azonban nem a számítógép, hanem a gép felhasználója részéről jelentkezik.

Egy adott szakterület buktatói a gép számára is végzetesek lehetnek, legfeljebb a gép segítségével a hamis eredmények könnyebben és gyorsabban érhetők el.

A számítógép által megoldani kívánt problémákat *előben gép nélkül* is meg kell tudni oldani. A gép szerepe — jelenlegi ismereteink szerint — „csupán” egy már korábban megalkotott logikai-matematikai lánc *konkrét* lépéseinek a hagyományos eljárásokhoz képest gyors és pontos elvégzésére korlátozódik.

A fenti megállapítások érvényesek a számítógépes technikának a történettudományban való alkalmazása esetén is.

A számítógép sem az adatok kiválasztását, sem forráskritikai vizsgálatát stb. nem végzi el a történész helyett (bár ilyen feladatokra is felhasználható, csak nem a történész helyettesítőjeként); a rendelkezésére bocsátott adatokat mintegy „végső igazságok”-nak tekintve egyaránt képes mind a leghajmeresztőbb prekonceptiók igazolására, mind — a helyes kiválasztási elvek alkalmazása esetén — a munka nagyarányú meggyorsítására, elmélyítésére, finomítására.

A számítógép gyakorlati haszna tehát a jól megválogatott adatok gyors és logikai ellentmondásoktól mentes feldolgozásában rejlik, ez teszi a nagyszámú forrással dolgozó és sokoldalú vizsgálódásokat végző történész számára nélkülözhetetlen segédeszközzé.

## 2. A topográfiai diagram előállítása

A sornyomatással készült topográfiai diagramok a tematikus kartográfiában használt pontdiagramokra hasonlítanak annyiban, hogy az ábrázolni kívánt földrajzi területen elhelyezkedő kisebb területi egységeket (pl. helységeket) pontokként jelölik. A pontokat a diagramon különféle, a számítógép karakterkészletéből vett, meghatározott szempontok szerint kiválasztott jelek képviselik.



Az eljárás lényege:

- az adatok mindig a pontokkal reprezentált helységekre vonatkozóan vannak megadva,
- mindig a helység topográfiai (térkép-) koordinátaival együtt kerülnek rögzítésre az adatkártyákon ill. mágnesszalagon,
- a számítógép központi memóriájában egy mátrixban tárolódnak, ahol az egyes adatok mátrixon belüli elhelyezkedését a térképkoordináták adják meg;
- valamennyi adat feldolgozása után a mátrix a központi memóriából kiírásra kerül.

Ez az eljárás alkalmas arra, hogy egy meghatározott léptékű térképen elhelyezhető szignifikáns pontokat (helységeket) egy számítógép sornyomatójának tábláján térképszerű elrendezésben vigyünk fel. Az adott földrajzi pontot egy-egy meghatározott történeti, természet- vagy gazdaságföldrajzi, mezőgazdasági, demográfiai stb. szempont szerinti kategorizálásban, a besorolásának megfelelő jellel ábrázoljuk. A kategória jele lehet direkt (adatokkal közvetlenül megadott) vagy indirekt (adatokból számítható).

E jelhalmaz vizuális áttekintést nyújt az egy területet jellemző tényezőkről, az ábrázolásmód pontossága azonban egy jól definiálható érték alatt marad. Ugyanis az egyes földrajzi pontok a sornyomatató táblájának (leperellojának) fix felosztású hálózatában kerülnek koordinálásra. A tábla koordináta-felosztása a „vízszintes” és a „függőleges” tengely mentén különböző. Egy sor hossza ugyanis 120 egységgel (írópozícióval) véve 30,5 cm; egy oszlop 25,4 cm, ami 60 egység (sor); így a legkisebb koordinálható mező egy 2,54 mm × 4,23 mm méretű téglalap. Ez utal egyben az egyes pontok ábrázolásánál a koordinálás pontosságára is. Az eltérés maximális értéke:

$$\text{vízszintes irányban} \quad \frac{2,54}{2 \cdot 10^6 \cdot M} \text{ (km), (Ny—K irány)}$$

$$\text{függőleges irányban} \quad \frac{4,23}{2 \cdot 10^6 \cdot M} \text{ (km), (É—D irány)}$$

ahol M a térkép lépték-aránya.

Első kísérletünk során Vas megye 350 vizsgált helységének túlnyomórészt XVIII. századi adatait egy számítógépes program mágnesszalagra vitte fel lyukkártyákról. E mágnesszalag ilyenformán 350 azonos felépítésű blokkot tartalmazott, a helységek számának megfelelően.

A topográfiai diagramokat egy másik program állította elő egy-egy teljes szalagátfutásban. Egy átfutás során a mágnesszalagról a megfelelő jelek bekerülnek a számítógép központi tárolójában kijelölt 60×120 méretű mátrixba, úgy, hogy a betöltendő mátrix-elem indexeit a program minden blokk beolvasásakor a helység térkép-koordinátáiból generálja.

Az összes helység blokkjainak beolvasása után a program a sornyomatató tábláján a mátrix artalmát soronként kiírja, megállapodás szerinti pozicionáló jelekkel és a tábla azonosítását télző aláírással együtt.

### 3. Térképdigramok előállítás

Második kísérletünk a történeti Magyarország megyei részletességű munkatérképe volt. Itt az ábrázolt földrajzi egységek nem pontok, hanem jelekkel kitöltött felületek (vármegyék, megyei jogú városok területe). Pl. a kivitelezés ennek megfelelően módosítja a topográfiai diagram készítésének alapelvét: a térképkoordináták nem a tábla egy sorának és oszlopának találkozási pontját (egy írópozíciót) adnak meg, hanem mindig az írópozíciók egy csoportját.

A program a korábban leírt mátrixsémának megfelelően működik (129×60); itt a mátrix-elemek az egyes területekre jellemző mezőgazdasági adatok egységesített jelei.

A programcsomag első kártyája a vezérkártya, ennek jelenleg, a történeti Magyarország XIX. századvégi munkatérképéhez 33 variációja működik, ez szabja meg, hogy melyik térkép készüljön el, s ennek megfelelően írónak ki a fejlé- és oldalszövegek.

A magyarázó jelek is belekerülnek a mátrixba és sor-oszlopfolytonosan azzal együtt kerülnek sornyomtatásra (printerre).

A térképdiaagramokhoz ellenőrző táblázat tartozik. Ebből a térkép jellemző adatai olvashatók le és a térképszerkesztés ezen adatok segítségével korrigálható, illetőleg a térkép értékelésekor a kimutatott hiba figyelembe vehető.

Ellenőrző eljárásunkat egybevethettük azzal, hogy a Csebisev-tétel alapján az egyes területek százalékos hibája

$$a = n^{1/4}$$

ahol  $n$  az adott terület térképelemeinek száma, az  $a$  együtttható pedig a térképre jellemző állandó, a legkisebb négyzetek módszerével számítható.

A jelenlegi feldolgozás során  $a = 0,188$ .

A relatív hiba (az összterületre vonatkoztatva):

$$a/n^{3/4}$$

A fenti egyenletek alapján számított hiba az empirikus adatok segítségével számított értékekkel nincs ellentmondásban, és a tematikus térképek rendeltetését tekintve elfogadható.

#### 4. A diaagramokhoz kapcsolódó statisztikai feldolgozás

A diaagramok használhatók vizuális segédeszközként, grafikus segédletként. Sok diaagramból álló sorozat kiértékelésénél statisztikai eljárásokra van szükség, illetőleg az adatok térképszerű ábrázolásán túl szükség lehet a térképektől független statisztikai elemzésre is. A diaagramokhoz felhasznált, a mágnesszalagok törlése után lyukkártyákon tárolt adatok erre, megfelelő számítógépes program csatlakoztatásával, tetszés szerint felhasználhatók, de ilyen program már a diaagramok elkészítésével párhuzamosan is alkalmazható. Pl. a pontdiaagramok esetében a 350 Vas megyei helység egyes adatait kontingencia táblázatokban foglaltuk össze. A történeti Magyarország munkatérképlapjaival párhuzamosan statisztikai diaagramok is készültek, itt a két program fizikailag egybeépült.

#### 5. A térképdiaagramok rendeltetése és a topográfiai elv

1. Az ismertetett diaagramok kísérleti céllal készültek. Kidolgozásukra a magyarországi parasztgazdaság üzemszervezetének a XVI. századtól a XIX. század végéig terjedő monografikus kutatása kapcsán került sor, abból a célból, hogy a földrajzi módszer megfelelően érvényesíthető legyen a vizsgálatokban.

A topográfiai diaagramokat akkor alkalmazzuk, ha a térképen feldolgozott terület sok kis egységből áll, és ezek területe ill. határvonalaiak nincsenek figyelembe véve, magától a kinyomtatott táblától nem vizuális képet várunk, hanem tanulmányi anyagként használjuk, a feldolgozott adatok földrajzi elrendezésben való kimutatásaként.

A térképdiaagramnak, melyet a történeti Magyarország 1895. évi területére, az ugyanilyen, évi mezőgazdasági összeírás és ehhez kapcsolódó kiegészítő adatok felhasználásával dolgoztunk ki, van bizonyos vizuális értéke (ez a továbbiakban fokozható a jelek karakterének jobb megválasztásával), de a hangsúly itt is a tanulmányi jellegén van. Ez a diaagram típus akkor célszerű, ha a felvett területi egységek tényleges területének és a határvonalaknak kimutatása fontos, a táblán ábrázolt terület pedig nem aprózódik fel túl sok kis egységre.

A kartográfiában alkalmazott térképa automatizálással ezek a térképdia gramok csak lát szólag vannak rokonságban: míg a kartográfia számára a kinyomtatott vagy automatikusan kirajzolt térkép mindig végcél, a történeti kutatásban a térkép nem több, mint az elemzés egyik eszköze. A történettudományban alkalmazott Bertin-Arbellet módszer áll munkánkhoz legközelebb, de technikai adottságaink elváltakoztak. (Térképdia gramjaink nem pótolják a vizuális célra, publikációk számára manuálisan rajzoltatott térképeket, viszont azzal, hogy tet szés szerinti sorozatban állíthatók elő, megkönnyítik a publikálandó térkép kiválasztását, a grafikus munkáját pedig lényegesen gyorsítják.)

Eljárásunk lényege, hogy a térkép a vizsgálat számára rögzített adatokból készül (tehát az adatoknak csak egyik felhasználási formája), az adatfeldolgozás folyamata nem szakad meg a térképn yomtatás miatt. Elvileg előállítható olyan program, mely a nyers alapadatokon végzett legkülönbözőbb műveletek végeredményeit összefoglaló térképdia gramok formájában adja ki, vagy a közbeeső eredményekből készít térképdia gramokat és ezeket újabb térképdia gramban szintetizálja, tehát egy teljes elemzési folyamatot végez el automatizáltan, anélkül, hogy a folya matot — elvileg — meg kellene szakítani. Ilyen programok előfeltétele azonban az, hogy az elemzés minden lépését meg tudjuk tervezni; ezt részint hagyományos („papír-ceruza”) módon, részint géppel végzett részkiérletekkel készítjük elő. (Alapfeladat a térképsorozatok szintetizá lásához, az automatikus zónameghatározáshoz szükséges kategória-problémák megoldása, ami kifejezetten történeti statisztikai kérdés.)

2. A földrajzi módszer agrártörténeti jelentőségéről nem szükséges szólni. A paraszti üz em történeti statisztikai vizsgálatánál azért jutott az általánosnál is kiemelkedőbb szerephez, mert ennél a témánál az agrártörténet úgyszólván teljes problematikáját figyelembe kell venni (beleértve a források kezelésével kapcsolatos problémákat is), s ennek megfelelően a problémák — a rész kérdés szemszögéből nézve — többnyire élesebben vetődnek fel, mint egy-egy ö ssze fogláló szemléletű munkánál. Magának az üz em problémának összefoglaló áttekintéséhez viszont szükség van tájékozódásra ott is, ahol a hazai kutatás még nem jutott összefoglaló eredményekre. A rendelkezésre álló adatok territoriális áttekintése kiinduló vizsgálati feltétel, melyet csak a nagy sorozatban előállítható munkatérképek segítségével lehet biztosítani. Ebből az igen gy a korlati problémából keletkezett a számítógépes térképdia gram.

3. Maga a topográfiai elv, miután menetközben jutott kiemelkedő szerephez egy számító gépes statisztikai feldolgozásban, a számítógépes megvalósítás következtében általánosabb meg gondolásokhoz is elvezet:

Az a körülmény, hogy a térképdia gramot előállító program területi egységeket (az adott két esetben helységeket ill. megyei törvényhatóságokat) vesz alapul, következik az adatok kon zekvens topográfiai rendszerének szükségessége.

Abból, hogy a térképdia gram számára a területi egységeket földrajzi koordinátákkal kell meghatározni, a területek ill. helységek koordinátákkal való azonosítása következett; ettől már csak egy lépés az, hogy az adattárolásban a földrajzi név másodlagos információvá válik, és csupán a közérthetőséget szolgálja (azt, hogy a kutató ne veszítse el közvetlen kapcsolatát az adatokkal). Az azonosítás céljára a koordináták jobban és sokoldalúbban használhatóak, mint akár a földrajzi nevek, akár a neveket pótló sorszám vagy egyéb mesterséges szimbólum.

A térképdia gramokhoz szükségessé vált, hogy a mindenkori vizsgált terület földrajzi térképét a számítógép által nyomtatott tábla beosztásával egyeztetve, a nyomtatott soroknak illetoleg oszlopoknak megfelelő koordináta rendszerrel lássuk el. (Ez a munka független az alap térkép léptékétől; előállítható azonos léptékű számítógépes térkép különböző méretű térké pekből is, ha azok léptékét ismerjük.) Az egyszer már megadott földrajzi koordináták, ha az adatok tárolását is ezek alapján végeztük, lehetővé teszik, hogy térképdia gramok nyomtatása nélkül is végezzünk földrajzi keresést az adatokban.

A topográfiai koordináták bevezetése — mint a helynév helyettesítője az adatfeldolgozás ban — némileg függetleníti a kutatást a helynevek változékonyságától, az igazgatási-uradalmi

beosztások folytonos átalakulásától, az ezekből származó technikai nehézségektől. A topografikusan felépített adattárolás és feldolgozás elnélkülözi a megyei vagy az ÁBC szerinti rendezést és keresést, az alkalmazott rendszer pedig természetes, tehát kézenfekvő és indokolt. (A tárolásnál alkalmazott sorrend, a vízszintes sorokon végighaladó, a sorokat egymás után sorravevő rend célszerűségének felismeréséhez nem kell számítógép (már Kniezsa is ezt alkalmazta), de emellett más rendezési elveknek is megvan a létjogosultságuk. Ahhoz, hogy ezt egységes topográfiai rendszerező elvként vezessük be, a számítógépes feldolgozásnál fellépő kényszerítő követelményre volt szükség.

Az 1970—1972 folyamán végzett munka részleteinek, tárgyi eredményeinek és dokumentációjának ismertetésére itt nincs hely. Az első munkafázis, a topográfiai diagram kidolgozását és alkalmazását a Mezőgazdasági Múzeumok III. Nemzetközi Kongresszusának (Bp. 1971) anyagát közlő kötetben ismertettük (M. Mezőgazd. Múz. Közl. 1971—1972).

## HOZZÁSZÓLÁSOK

*Dányi Dezső:*

A számítógépek és a történettudomány problémaköréből egy részkérdéssel kívánok foglalkozni.

A számítógép alkalmazása a történettudományban akkor válik optimálissá, ha célratorően, a ráfordítási költségeket mindig szem előtt tartva, a rendelkezésünkre álló információs halmazból a legmegfelelőbbeket válogatjuk ki és próbáljuk megszólaltatni. Ehhez természetesen szükséges, hogy megfelelő áttekintéssel rendelkezünk a számítógéppel feldolgozható információk tartalmáról, szerkezetéről és nagyságrendjéről. A számítógép, mint eszköz, eléggé összeegyeztethetetlennek látszik a történettudomány rendelkezésére álló információs halmazok jelenlegi leltáraival. Ezek a leltárak maximálisan a könyvtári katalógusok szintjét, vagy még ennél is lazábban — a fond-jegyzékek formáját és tartalmát érik el. Lényegében a tárgy megjelölésénél a hely vagy az időpont és néhány kiemelkedő személy meghatározásánál nem adnak mélyebb tájékoztatást. Gépi adatfeldolgozás esetén ezeket a katalógusformákat alaposan elemeznünk kell és fel kell derítenünk — nem kis fáradsággal — részletes — ha összeírásokról vagy számszerű anyagról van szó — a rovat mélységéig terjedő tartalmukat. Másként fogalmazva: a történettudomány — számítógéppel feldolgozható — forrásainak adatbázisa nincs úgy elrendezve, hogy jelentős munkabefektetés nélkül felhasználható legyen.

Ennek az adatbázisnak a megteremtése hosszú, előzetes munkát kíván, de meghozná eredményét, mert a számítógépes értékelés, az ismerethalmazból való kiválasztás pontosságát, komplexitását növelné.

Azt hiszem, hogy már Katus László is célzott arra, hogy a számítógép és a történettudomány problémakörében nemcsak a numerikus — statisztikai összeírás — feldolgozás jelent előnyöket, hanem a kvalitatív, minőségi változókat hordozó információké is. A minőségi információk számítógépes feldolgozásával kapcsolatban újabb archeológiai kutatásokra utalok. A számítógép tehát módot ad olyan logikai konstrukciók és modellek — magatartások és vélemények például számszerűsített formában való elemzésére, amelyeket korábban — berendezések hiányában — nem végezhattünk volna el.

Végül arra kívánom felhívni a figyelmet, hogy a számítógép alkalmazása megnyitja a torz, töredékes, nem teljes minták, információs halmazok feldolgozásának, értékelésének lehetőségét is.

Mindezek a lehetőségek elméletileg adottak. Gyakorlati megvalósításuk igen sok szervezési, tudományszervezési munkát igényel — többek között a gondolatok kicserélésének, a közös nyelv kialakításának, ilyen — kerekasztal — módozatait.

*Perjés Géza:*

Igen öröndetes, hogy végre sikerült ezt a megbeszélést összehozni, melynek remélhetően valójában hasznos gyakorlati következményei lesznek. A kezdetekről van szó, így jól meg kell fontolni minden lépésünket, és az itt felmerült javaslatokat és terveket alaposan át kell gondolni.

Természetesen nem kétséges, hogy igen nagy szükség lenne a Katus László által javasolt adatbankra, továbbá mind arra, amit Wellmann Imre és Fügedi Erik terjesztett elő. Ugyanakkor azonban azt hiszem, szükség lenne a matematikai statisztika módszereinek kikísérletezésére is, hiszen alig tudjuk, hogy a történettudományban miként alkalmazhatók, vagy legalább is nem mindegyikről tudjuk, és éppen ilyen szempontból hatalmas meglepetések érhetnek bennünket. Ezek a módszerek mégis csak elsősorban a természettudományokra álltak be, azokból eredtek, viszont könnyen lehetséges, hogy a történelemben egészen más populációkkal, egészen más eloszlásokkal, következésképp más törvényszerűségekkel találkozunk majd, s mindazok a módszerek, melyek ma a természettudományban — de akár az ökonometriában, vagy más humán jellegű tudományokban — oly sikeresen alkalmazhatók, a mi esetünkben kétes értékű eredményekhez vezetnek.

Amennyiben tehát kísérleteznünk is kell, akkor az említett nagy tömegű adatok feldolgozását kilátásba helyező tervek mellett két dolgot kellene csinálni: kisebb forrásosoportokat végigmatematizálni, azonkívül tömegesen végezni szekundér elemzéseket, azaz már kiadott, forráskritikailag vitathatatlan értékű adatokat a matematikai statisztika segítségével újra elemzés alá venni, amint ezt Veress Éva is csinálta Vörös Károly művével.

Ezenkívül azonban szó van arról is, hogy miután a matematikai statisztika megtanulása kardinális kérdés, viszont a filozófus beállítottságú agyak számára a matematikai apparátus egésze megemészthetetlen, — azaz, mi legfeljebb az egyes módszerek logikáját tudjuk megérteni, de pl. már a matematikai levezetéseket alig tudjuk követni — felmerül a kérdés: van-e királyi út számunkra, történészek számára a matematikai statisztikai módszerek megtanulásához?

Én azt hiszem van, és itt saját tapasztalataimról szeretnék beszélni. (Persze nagyon érdekes lenne, ha éppen azok, akik már dolgoztak computerrel, mint pl. Veress Éva és Puskás Júlia, beszámolnának arról, hogy ők miként sajátították el azt a matematikai tudást, ami munkájukhoz óhatatlanul kellett.) Én úgy gondolom, hogy az ember akkor tudja legkönnyebben megközelíteni a matematikai statisztikai módszerek lényegét, ha saját kutatási anyagán tanulja meg azokat. Kb. 8 éve dolgozom az 1728. évi adóösszezáron és kezdettől fogva arra törekedtem, hogy az anyagban található mennyiségi, sőt minőségi adatokkal is különböző matematikai műveleteket végezzek el. Elsősorban szóródás-, korreláció- és regressziószámításról van szó, valamint az asszociáció erősségének méréséről. E műveleteket részben könyvből, részben tanfolyamokon tanultam meg, tulajdonképpen azonban csak technikájukat, mivel a matematikai levezetések és a problémák sajátosan matematikai kifejtése jórészt idegen, beszéljünk őszintén, érthetetlen maradt számomra. Jellemző pl., hogy a matematika nyelve, a szimbolika, ami a valódi matematikus elméjé számára a legbiztosabb támasz, számomra óriási nehézségeket okozott, és sokszor könnyebben megértettem egy levezetést, ha a szimbolikumokat a konkrét számokkal helyettesítettem. A sok számolás viszont mégsem volt eredménytelen, mert részben erősen fejlesztette számérzékemet, másrészt sok mindent megértettem a matematikai eljárások logikájából, és ami a döntő: ezt a logikát sikerült — legalább is úgy érzem, hogy sikerült — azzal a logikával összeegyeztetni, mely magából az anyagból kibontakozott.

Ismétlem tehát: ha van egyáltalán királyi út a matematikai statisztika módszereinek megértéséhez, akkor az a történész saját kutatási anyagán át vezet. Itt van pl. a szóródás-számítás. Ha a történész hosszú ideig bibelődik valamilyen adattömegegél és azt esetleg még táblázatba is foglalja, észreveszi, hogy az adatok valamilyen érték körül tömörülnek, amire aztán ki is mondja, hogy az tipikusnak tekinthető. Ezután azonban már jóval kevesebb nehézséget okoz a számtani, vagy más átlag levezetésének megértése, mint ha idegen anyagon mutatják azt be

reki. Azután történész fejjel annak érzékelése és megértése sem túlságosan nehéz, hogy a tipikus annál „tipikusabb”, minél inkább tömörülnek az adatok körülötte, azaz az átlag körül. Ettől viszont már csak egy lépés választja el attól, hogy megértse a szóródás lényegét és elővéve egy kézikönyvet, fogalmat alkosson a normális eloszlásról és annak valószínűségszámítási következményeiről. Így azután tulajdonképpen már nem is túlságosan érdekes, hogy lépésről-lépésre tudja a standard-deviáció számítás levezetését, hanem sokkal fontosabb, hogy megérezze és megértse, hogy annak a törvényszerűségnek a logikája, amely az anyag tanulmányozása alapján világossá vált előtte, azonos a szóródás-számítás logikájával, és hogy öröm fogja el azért, hogy amit addig csak verbálisan és csupán megközelítően tudott megfogalmazni, azt most számszerűen és viszonylag nagy pontossággal tudja kifejezni.

További kérdés, hogy nem lehetne-e olyan kisebb kapacitású számológépekre szert tenni, melyeken nem nagy tömegű — tehát az adathordozóra való felvitel szempontjából különösebb problémát nem okozó —, viszont sok, vagy bonyolultabb számítást igénylő adatokon közvetlenül és különösebb adminisztratív procedurák nélkül el lehetne végezni bizonyos műveleteket.

A számítógépekkel sokszor triviális eredményeket hoznak ki, említette Dányi Dezső. Ez, sajnos, nagyon is igaz. Ugyanakkor viszont a triviális eredmények, éppen azért, mert triviálisak, néha nagyon is fontosak lehetnek, pl. forráskritikai szempontból. Én, amikor Vörös Károly anyagát elemzés alá vettem, egy igen triviális kérdést tettem fel: van-e összefüggés a Vörös által közölt és elemzett forrásban a szántó területe és az ígásállatok száma között? A kérdés azért triviális, mivel a válasz csak egy lehet: kell lenni összefüggésnek, hiszen több szántó megműveléséhez több ígásállat kell. Korreláció-számítással kerestem a választ, és bebizonyosodott, hogy a két tényező között igen erős — ha jól emlékszem — 0,7—0,8-as erősségű kapcsolat volt, aminek alapján el kellett fogadnom, hogy a forrás megbízható, hiszen igen csekély a valószínűsége annak, hogy rossz adatok alapján is hasonló erősségű összefüggést kaptam volna eredményül.

Végül pedig a perspektíva. Teljesen egyetértek azzal a véleménnyel, hogy a számítógép nem hoz, nem hozhat forradalmat a történetírásban. Sőt nagyon is gyanús lenne egy ilyen forradalom. Forradalomról legfeljebb olyan területeken lehet szó, melyek forrásainak adatai számszerűek vagy számszerűsíthetők. Ez viszont a források elenyészően kisebb részére áll fenn. Ezen a szűkebb területen végbemenő forradalom azonban már pszichológiai, alkotáslélektani is egyben. Arról van szó ugyanis, hogy a tudomány nem lehet meg hipotézisek nélkül, ezekhez viszont intuíció és fantázia kell. Bizonyos adatmennyiségen felül azonban a felállított hipotézisek és sejtések verifikálásához irdatlan tömegű adatscsoportosítás, kombinációs tábla és számítás kell, ami kézi úton elvégezhetetlen. A computer tehát tulajdonképpen az emberi szellem legértékesebb és legnemesebb adottságait, az intuíciót és a fantáziát szabadítja fel, eltávolítva azoknak a technikai nehézségekből származó korlátait, ugyanakkor — s ez nem mellékes dolog! — szoros ellenőrzés alá is veszi őket, mert a felállított hipotézis helyességéről pillanatok alatt kimondja az ítéletet.

Gondoljunk csak arra, hogy annak idején Pascal, Bernoulli, Laplace és mások még azt gondolták, hogy az emberiség a valószínűségelmélet birtokában olyan eszközre tett szert, mely a helyes emberi viselkedéshez és a racionális döntések meghozatalához fog biztos fogódzókat nyújtani, hiszen — mondták — az emberi döntés nem más, mint valószínűségek mérlegelése. Ez akkor persze álmom volt, de ma, a computerek korszakában a döntés- és játékelméletben a megvalósulás felé halad.

Hogy most már szűken vett szakmából, a hadtörténetből hozzak példát: annak idején Zrínyi a háború valószínűségeit még szerencsének nevezte, amiben a korabeli tudomány szintjén tulajdonképpen teljesen igaza volt. Éppen a valószínűségelméletnek XVII—XVIII. századi robbanásszerű fejlődése ösztönözte Clausewitzot, a katonai elmélet eddigi legnagyobb alakját, hogy a háborús szerencsét a valószínűséggel helyettesítse. Megállapításai semmilyen visszhangot sem keltettek a hadtudományban, és közel másfél évszázadnak kellett eltelnie ahhoz, hogy a valószínűségelmélet alapján a háborút modellizálják. Amilyen örömmel tölti el e nemes művé-

szet, — ti. a hadművészet — híveit ez a tény, ugyanolyan szomorúvá is teszi, mert ennek az egésznek napjainkban, az atomháború korszakában túlságosan sok értelme nincs, hiszen a háború eredményének most már csak egy valószínű kimenetele van: a tökéletes megsemmisülés. Úgy, hogy jobban is tennénk, ha a háborút, mint a konfliktusok elintézésének módját, félretennénk. Ami most már mondanivalóm szempontjából lényeges, az az, hogy a számítástechnika helyettesíteni, vagy legalább is erősen támogatni tudja azt az intuíción, melytől vezetettve, a hadvezérek évtizedeken át győztek vagy vesztek.

Nos, nálunk történelemszempontnál is hasonló a helyzet. A történeti folyamatok valószínűségi, stochasztikus folyamatok, melyeket a történelem megsejtethetett, s néha verbálisan meg is fogalmazhatott. A computer viszont e sejtéseket igazolhatja vagy cáfolhatja és exact módon meg is fogalmazhatja. Legalább is a história egyes területein.

*Sándor Pál:*

Olyan primitív dolgokat szeretnék szóvá tenni, hogy szerény hozzászólásom legfeljebb néhány percet vesz igénybe. Kapcsolódnék a Dányi Dezső által felvetett információ problémájához.

Ő ezt a kérdést igen erőteljesen exponálta, nézetem szerint azonban a helyzet nem olyan rossz, ahogyan ezt lefestette, mert — tudomásom szerint — a repertóriumok sora készült el, mind az Országos Levéltárban, mind a területileg illetékes levéltárakban. Évek óta folynak ezek a szisztematikus munkálatok, és a történelemszempont az első feladata az, hogy erről tudomást szerezzen. Hadd utaljak erre a tényre, főleg mert *összeírásokról* van szó. Emlékszem, hogy pl. a konkripciókról szép anyag áll rendelkezésünkre. Már most, azt nem lehet várni, hogy a kutató számára az íróasztalhoz vigyék az anyagot és a tájékoztatókat, jegyzékeket, vagy repertóriumokat oly részletességgel készítsék el, hogy azokban a számára fontos kérdések mindjárt, közvetlenül előbukkanjanak, mert ezeknek a jegyzékeknek — azt hiszem — nem is ez a szerepük. Ezek ismerete alapján olyan kutatásokat kell még elvégezni, ami nélkül történetet írni nem lehet. Mert ki gondolta volna azt (én legalább is egy ideig nem gondoltam), hogy a cs. kir. Főkezelő s a visszaállított Helytartótanács iratai között olyan a birtokrendezéssel kapcsolatos, eddig ismeretlen statisztikai kimutatások vannak, amelyek mind agrárstatisztikai, mind agrártörténeti szempontból egyaránt nagy érdeklődésre tarthatnak számot. A jegyzékek, repertóriumok az ilyen természetű adatokat *közvetlenül* nem ugrathatják ki, de azok ismerete mégis nélkülözhetetlen, sőt — szakszerű használat mellett — az információk adatok feltárását is lehetővé teszik. Ezzel azt kívánom hangsúlyozni, hogy Dányi Dezsőnek abból a szempontból teljesen igaza van, hogy információk anyagra szükség van, de talán abban a vonatkozásban nincs teljesen igaza, hogy ezen információk anyagok ne állnának már ma — bizonyos szinten — rendelkezésünkre.

T. k. az ő hozzászólásában az a lényeges, amit talán *közvetlenül* nem mondott ki, de észrevételeiben — értelem szerűen — mégis benne rejlik: ez a fajta munka, amit mi csinálunk, kifejezetten történelemjellegű munka. Ha pedig történelemjellegű munkáról van szó, akkor az első tennivaló: a források felkutatása, azután a forráskritika elvégzése. Erről különösen Wellmann Imre előterjesztésében kaptunk hangsúlyozott utalást és Fügedi Erikében is, ami ugyan annyira magától értetődő, hogy erről nem kellett volna szólnom, miután azonban ez az alapelv eddig nem került szóba, ezért azt egy pár mondat erejéig mégis hangsúlyoznom kellett.

Továbbá: a történelemszempont az is el kell végeznie — egyelőre függetlenül minden gépviteltől — (e szakaszban még szó sincs a gépről), hogy ezen forráskritika által rendezett adatok az adott kor ismeretében, abba ágyazottan, *történetileg hitelesek* legyenek. Mert ha nem azok, és csak úgy megőrülünk a nyersen kapott számoknak (van példa erre, nagy munkák készülnek így; megtaláltak egy kimutatást, nekiültek, 5 évig dolgoztak rajta, és aztán kiderült, hogy a munka azért rossz, mert a genezist, az alapvető történelemszempontot nem végezték el), kárba vesztett munkáról van szó. Ezzel csak azt kívánom röviden hangsúlyozni, hogy *mielőtt* gépre tesszük és

számítástechnikai szintre emeljük az anyagot, addig is (ezt az előterjesztések nagyon helyesen mondják, ehhez csatlakoznék én is, nekem is ez a véleményem, s ha nincsen igazam, kérem, mondják meg), a munka első fázisában a szokványos „konzervatív” módszerrel kell dolgozni, mert ez a módszer az, amely lehetővé teszi, hogy az adatok később egyáltalán értelmezhetőek legyenek. Gépre nekem akkor lesz majd szükségem — etekintetben teljesen ismeretlen területen mozgok —, ha olyan összefüggéseket keresek, amelyeket az anyag csak sejtet, de bizonyosan nem mond meg, és gépre téve valaminő többletet nyújtanak a történész számára; vagy olyan értelemben, hogy az általa felvetett összefüggéseket igazolják, illetve megcáfolják, vagy más — eddig nem gyanított — összefüggéseket mutatnak ki. *Ebben az esetben* a munka már arra a szintre emelkedik, hogy az adatok gépre vihetőek. Amíg azonban a munkálat említett első szakasza (forráskutatás, forráskritika, az adatok történetileg hiteles értelmezése) lehetővé teszi a tökéletességgel nem készült el, addig a második fázisra való átvitelt nem szabad megvalósítani, mert itt megint kárbavesztett munkától kell félnünk.

Ami végül rövid felszólalásom harmadik pontját illeti, az csak arra vonatkozik, hogy a legnagyobb mértékben egyetérték az elnöknek az itt felvetett szervezeti és tanulmányi-szakmai vonatkozású konstruktív javaslataival, mert hisz sokunknak szükséges azok valóra-váltása.

*Veress Éva:*

Tudománytörténetileg tekintve azt a folyamatot, amellyel eljutottunk a számítógép igénybevételeéhez, emlékeznünk kell a történész közvélemény jellegzetes kettéhasadására, mely nem szakmánk sajátossága volt. Szerencsére eléggé múltóban van ahhoz, hogy a benne testet öltött félreértésekkel leszámoljunk. Egyik oldalon a modernizáló és számítógéppárti közvélemény, néha erős túlzásokkal és végletekkel, másik oldalon a konzervatívnak minősített, néha valóban elzárkózó, de általánosságban inkább óvatos és mértéktartóan gondolkodó tábor. A kibontakozásra nézve sokáig komoly veszély volt a „félíg értett dolgok ígérete” (Kuczka Péter jellemezte így a sci-fi vonzóerejét). A matematikai módszerek és számítógép lényegének felszíni ismerete, megoldva a humán képzelettel. Ez az objektív lehetőségeknek nemcsak túlértékeléséhez, hanem félreismeréséhez s így eltorzításához vezetett, ami nem vált a korszerűsítő, törekvések javára. Ami viszont a hagyományos oldalt illeti, kérdés, hogy jogos-e az ún. „narratív” és a „kvantitatív” történetírásnak olyan ellentétpárként való szembeállítása, amelyben — ha teljesen nem is zárják ki egymást — de igazából csak az egyik oldalé lehetne a jövő. Én nem hiszek benne. A kvantitatív szemlélet érvényesülésének eddigi nemzetközi tapasztalatai indokoltá teszik, hogy a kvantifikációnak ne a hagyományossággal való ellentétét, hanem a történetiséghez való viszonyát, a történettudomány stabil hagyományaihoz való kötöttségét tegyük szavá. A „hagyományos” történetírás olyan erősségei, mint a forrás-filológia és a közérthető kifejezés-mód például nem avulnak el és a kvantitatív szemlélet sem nélkülözheti a múltba való beleélés intuitív képességét. (A forrástapasztalatból, korismeretből táplálkozó intuícióról van szó.) Amikor továbblépünk, a jelenkori, a matematizálódás irányába hajtó fejlődés nem páfordulást kíván tőlünk. Meggyőződésem — és más tudománysszakok példája is ezt mutatja —, hogy legalább annyiban kell arra építenünk, ami a történettudomány sajátja, mint arra, amit kívülről adaptálunk.

A kvantitatív kutatásoknál a műfaj buktatóival érdemes elsősorban törődni. Ilyen buktató például az, hogy ha a tömeges forrásfeldolgozásnál mélyreható, sokoldalú munkát akarunk végezni, a számítógép segítségével kapott eredmény gyakran komplikáltabb, mint a kiindulás: néha rejtjelessége, de leg többször áttekinthetetlen tömege miatt. (Ismeretes J. Bertin 1969-ben, az Annales-ban leírt figyelmeztető példája, de idézhetnénk a Sztanyiszlav Lem egyik gépmeséjében megfogalmazott általánosabb intelmet is.) Ha a géptől kívánt információt nem szelektáljuk és



meghatározásában túl nagy szerepet engedünk a gép fizikai kapacitásának, megfelelkezve az értékelő elme kapacitásáról, olyan túlradó információtömeget nyerhetünk, mely *menyisége arányában veszít információs értékéből*. — Az információs szervezés még bizonyára számos részletvitát fog kiváltani, ahol a kvantifikációnak ez a problémája nyilvánvalóan nem önmagában nyer majd megoldást.

A team-munkával, szükségességével láthatóan mindannyian egyetértünk. Idekapcsolódik a team-munka várható hatása túlzott, bénító specializáltságunkra, mely nem annyira tematikus, mint történeti korokra szabott. Tarthatatlan, hogy ma egy XVI. százados és egy a XIX. század második felével foglalkozó történészt például nagyobb távolság választ el egymástól, mint egy-egy, velük valamilyen problémában találkozó közgazdásztól, szociológustól, geográfustól, nyelvésztől. Átgondolt kutatás, jó téma- és szakembercsoportosítás sokat segíthetne, hogy feloldódjék szakmánknak ez az atomizáltsága. A számítógépnek ebben — mint látjuk — serkentő és elhatározó szerep jutott.

Ami magát a számítógépet illeti: hatékony segédeszközt lássunk benne, vagy ennél többet? Szerszám vagy interdisciplinális módszer jutott kezünkbe? Ezek azok a kérdések, melyeket nem kell elvileg megválaszolni, de a következőket *tudjuk*: munka közben a kutató számára egyelőre a konkrét teljesítménybővülés a fontos. A történész szemlélete szempontjából viszont az a döntő, hogy olyan technikáról van szó, melynek léte — függetlenül attól, hogy személyesen igénybe vesszük-e vagy nem — tudománytörténeti korszakot határoz meg; mint ahogyan a kultúra történetében korszakot határoz meg például a vízierő: függetlenül attól, hogy egyidejűen több fajta energiaforrást használtak, az adott univerzális energián alapuló fejlődés vált a korszak jellemzőjévé.

Varga Dénes:

Az eddigi hozzászólók — történészek lévén — természetesen történész szemmel nézték a gépesítés problémáját. Szeretném néhány szóval a másik oldalt, a számítógépes oldalát érinteni a problémáknak. Ahhoz, hogy eredményesen használjuk a számítógépeket a történelemtudományban, előbb egyet kell értenünk az alapvető dolgokban, meg kell találnunk a közös platformot.

Azzal kezdeném, hogy nemcsak a számítógép, a matematika sem az, aminek a nem-matematikuskok legtöbbször hiszik. A legtöbb nem-matematikusk még ma is abban a hitben él, hogy a matematika „a számok tudománya”. A matematikának ez a platóni koncepciója még a XIX. században is kísértett, a modern matematika kialakulásával azonban — a XX. században — már teljesen tarthatatlanná vált. Közelebb állunk az igazsághoz, ha azt mondjuk, hogy a matematika a *relációk tudománya*. A matematikai absztrakciónak egy viszonylag alacsonyabb szintjén valóban a számfogalom dominált, és a matematikusok sokszor maguk is megelégedtek a számokkal megfogható relációk kifejezésével. Sokkal helyesebb azonban általában objektumokról és az objektumok között fennálló relációkról beszélni. Maguk a relációk is néhány elemi tulajdonságukból magyarázhatók meg matematikailag, hasonlóan ahhoz, ahogyan az atomfizika is eljutott az elemi részecskék fizikájához.

A számítógépek felhasználásához nagyon fontos, hogy ne csak számokat lássunk mindenben, ne kösse meg a történészek fantáziáját az, hogy csak statisztikázást várnak a géptől, azt hiszik, a gép ennél többre nem képes.

Roppant fontos szempont természetesen az is, hogy ne csak azt lássuk, mit *lehet*, hanem azt is, mit *érdemes* géppel végeztetni. Megint visszatérve az előbbi példára: egyáltalában nem biztos, hogy egyszerű számítások kedvéért érdemes a számítógépet igénybe venni. Gyakran a milliósámra feldolgozott anyag sem ad sokkal pontosabb képet, mint ha véletlenszerűen kiragadott mintából kézzel végezzük el ugyanazt a számítást.

A fő előnye a számítógép alkalmazásának az, hogy egyidejűleg nagyon *sokféle szempontból* lehet feldolgozást végezni, hogy úgy mondjam, „több bőrt le lehet húzni” a gépre vitt anya-

gokból, míg régebben óriási energiabefektetéssel is legfeljebb odáig jutott el a kutató a nagyobb anyagok feldolgozásánál, hogy egyetlen szempontból végzett csak értékelést.

Az anyagok kiválasztásánál tehát előre jól gondoljuk meg, hogy lehetőleg olyan anyagokat vigyünk gépre, amelyeket többféle szempontból vizsgálni lehet. Itt azután nem kötelező, hogy előre ismerjük mindazokat a szempontokat vagy a szempontoknak mindazokat a kombinációit, amelyeket a vizsgálat során alkalmazni fogunk. Ez is egyik alapelvünk lehet a gépi feldolgozás előkészítésében, hogy kalkuláljuk bele munkánkba a későbbi módosítások, bővítések lehetőségét. Nem szabad túlságosan szűkkeblűen, practicista módon csak egyetlen konkrét, előre teljesen behatárolt célra irányítanunk a feldolgozást, hanem próbáljuk belefoglalni a gépre vitt anyagba mindazokat az információkat az adott adatokkal kapcsolatban, amelyekre a feldolgozás későbbi stádiumában valamikor szükségünk lehet. Ha a feldolgozás nagyobb méretű, akkor szinte kötelező elv az, hogy először mintafeldolgozással kezdjük, s az erre épített visszacsatolás alapján ellenőrizzük, tökéletesítsük az alkalmazott módszereket.

Fontos alapelv a számítógépek felhasználásában a *szemléletesség* lehetőségének kihasználása. Sok energiát takaríthatunk meg vele, ha sok-sok adat hosszadalmas értékelése helyett pl. áttekintő grafikonokból képet kapunk a vizsgált folyamatról, s a kérdéseket később már ennek megfelelően tehetjük fel. A gép és az ember közötti együttműködés az, amibe bele kell tanulnunk, hogy a gép lehetőségeit jól ki tudjuk használni. Nagyon jól felhasználható eszköze lehet például a történész kollégáknak a rajzolóberendezés, a *plotter*, nagyon jól felhasználható arra, hogy egyes összefüggésekről szemléletes képet kapjanak. A plotterek alkalmazása során egyébként általánosan elfogadott módszer az, hogy előbb durvább lépésközzel általánosabb képet rajzoltatnak a vizsgált folyamatról, s csak ennek alapján vizsgálják meg finomabban.

Önmagában a gép nem különösebben okos jószág, de nagyon butának sem kell tartanunk — az a kérdés, hogy mi magunk milyen okosan használjuk. Még azt sem állíthatjuk, hogy a gép képtelen a hibák kijavítására. Érdekes példákat lehet találni éppen abban az írásos anyagban, amit Veress Évával együtt készítettünk, arra, hogy egészen vad torzulások esetén is ki lehet javíttatni számítógéppel az ember által elkövetett hibát (pl. eltorzult falunek azonosításában).

Végül szeretném felhívni a figyelmet arra a módszerre, amely Csernok Attila, Ehrlich Éva, Szilágyi György közös munkájában is felhasználást nyert a véletlen ingadozások hibájának kikorrigálására. Az ún. *csúszó átlagolásra* gondolok, ahol az egy évre vonatkozó átlagok önmagukban nagyon sok véletlen eseménytől függhetnek, és ezért nem mondanak elég sokat; ha nagyobb időközökre vizsgáljuk a folyamatot, akkor meg túlságosan vázlatos képet kapunk. Ha viszont minden évben az érték meghatározásánál a *környezetet* is figyelembe vesszük korrigáló tényezőként, vagyis az előtte és az utána levő értékekkel együtt az átlagot vesszük, az átlag kiszámításánál figyelembe veendő értékek csúsznak évről évre tovább a skálán. Hasonló elvet követünk egyébként nemcsak a számításokban, hanem magának a tervezésnek a folyamatában is a távlati tervezés esetében. A távlati terveknek nem az a szerepe, hogy egy olyan keretet adjon, amelyet évekre lebontva kötelező érvénnyel be kell tartani. A távlati terv arra való, hogy előre lássunk — egy bizonyos pontossággal, olyannal, amilyennel egyáltalán lehet. Egy év múlva már nem teljesen az előző évben megállapított értékek érvényesek az előttről levő évekre, hiszen az azóta eltelt idő bizonyos mértékig módosíthatta elképzelésünket: van, amit igazolt, van, amit megcáfolt, mindenképpen közelebb hozta azonban a részleteket.

Ilyen és hasonló elvek bevezetésével lényegesen hatékonyabbá lehet tenni a számítógépek alkalmazását. A történészek által feldolgozandó anyagoknál is ez a feladatunk: olyan módszereket próbáljunk kidolgozni, amelyek segítségével többet ki tudunk belőlük hozni, mint amit emberi energiával megtehetnénk.

*Benda Gyula:*

Történészként szeretnék hozzászólni, mégpedig egy kicsit saját kollégáim ellen szólva. Itt állandóan a források problémájáról beszélünk, pedig szerintem itt nem a források problémájáról van szó. Ez nem újkeletű kérdés, mert amikor a Monumenta Hungarorum kötetét kiadták, akkor is meg kellett találni az okleveleket, és itt az, hogy modern információ feldolgozási módszereink vannak, előny, de továbbra is meg kell találnia a történésznek a maga forrását. Én az elméletre szeretném felhívni a figyelmet.

Minden történeti magyarázatnak, elméletnek megvan a maga forrása, tehát ha én ma fel akarok állítani egy adatbankot, amely matematikailag vagy számítógéppel feldolgozható, az holnap elavulttá válik abban az esetben, ha új módszer pl. lehetővé teszi eddig kallódó levéltárak polcain heverő iratkötegek feldolgozását is.

Azt hiszem az a vita, hogy az egyik felszólaló azt mondja, van ilyen adatrepertórium, a másik azt, nincs, az onnan fakad, hogy más kérdéseket tesznek fel az adatoknak. Azt kell tehát tisztázni, és számítógépesekkel is csak így fogjuk egymást megérteni, ha tudjuk, hogy milyen összefüggések, milyen elmélet segítségével próbáljuk rekonstruálni a valóságot. Tehát meglepszünk-e a társadalmi tagozódásban az osztályminőség leírásával, vagy pedig más ismérvek alapján mennyiségileg próbáljuk a társadalmi struktúrát megközelíteni. Nyilvánvaló, hogy az egyik történész számára nem forrás az, ami a másik kutató számára forrás, és a hagyományos, ill. modern történet közti különbséget nem ott látom, hogy az egyik csak papíron és ceruzával hajlandó számolni, hanem ott, hogy egyáltalán fel sem merül benne a számítás szükségessége, tehát nem ott van eltérés, hogy az egyik fejlődési trendet számolva azt mondja, hogy a géphez ő nem ért, ő csak papíron számol, hanem egyáltalán a gazdasági fejlődés jellemzésére nem a trendet tartja jellemzőnek. Ha már oda eljutott, hogy trendet számol, akkor már tényleg csak egy olyan számítás kell, hogy az adatok tömegessége, az időszakasz hosszúsága kifizetővé teszi-e a gépi munkát vagy nem.

Ezt a számítást olyan összefüggésben is elvégezhetjük, mint ahogyan egy francia történész mondta, hogy ők kiszámították: egy történész munkaórája belekerül egy bizonyos összegbe, a számítógép munkaórája is. Ha a történész számítógép nélkül dolgozik X órát dolgozna, géppel X—Y órát; akkor fizetődik ki a számítógép, ha megtakarítás jelentkezik. Nyilvánvaló, a rögtönzött számítások szerint a magyar történész munkaóráját kevésbé díjazták, a számítógép többbe kerül, nem szabad elfelejteni a számítógép felhasználásának ezt a problémáját sem.

Szeretnék utalni Katus Lászlónak a referátumára, ahol nagyon világosan megfogalmazza azt, hogy vannak kipróbált módszerek, amelyek bizonyos változtatásokkal alkalmazhatók, természetesen nekünk a forrásaink különböznek valamivel, tehát módosítani kell a módszereken. De az első feladat mégiscsak az lenne, hogy számítógépesink és történészeink is megismerjék ezeket a módszereket, a történész mondja meg, hogy kérem, nekünk ilyen forrásunk nincs, tehát eldobhatjuk a módszert, a matematikus pedig esetleg azt mondja, ilyen mi még nem tudunk csinálni. De mindenesetre az első kiindulásnak adott ez a lehetőség. Ne akarjunk túl nagyot lépni. Próbáljuk meg az elért eredményeket Magyarországon is bevezetni.

Egy második lépcső, erre van már példa Magyarországon, az, hogy mivel mi történetileg másképpen látjuk a valóságot, eltérően próbáljuk magyarázni, eltérő módszereket alkalmazunk. Erre jó példa, hogy a kartográfiai módszernek kialakult egy sajátos magyar változata, ugyanakkor a franciák is kidolgoztak egy másikat. Itt nyilvánvalóan más jellegű problémáról van szó, mint az, amit Katus László ismertetett.

A számítógépesek számára szerintem nem az a fontos, hogy jelenleg mi történészek el tudjuk-e dönteni azt, hogy egy latin kifejezés mit jelent, vagy nem tudjuk eldönteni, mert ez korábban is fontos volt a történész számára. A számítógépesnek, azt hiszem, az a lényeges, hogy mi most struktúrában gondolkodunk, időfolyamatban stb. Tehát pl. a Marzewski-féle kvantitatív történelem gondolati rendszer, amelyik ráépül egy matematikai rendszerre és vele szemben áll

az „histoire sérielle”, amelyik független idősorokat elemez anélkül, hogy ezeknek egymáshoz való viszonyát pontosan tisztázná, ez megint más matematikai módszereket szükséges.

A számítógép és áttételesen a matematikus is eszközt jelent a történész számára, de a végső szót mégiscsak a történész sajátos szempontjai kell, hogy kimondják. A történészek tudják feltenni a kérdéseket. Legfeljebb a matematikus azt fogja mondani, hogy ezt mégsem tudjuk megoldani, ezt meg lehet csinálni egy már meglevő programmal. De azt, hogy mit vizsgáljunk kvalitatíve, vagy kvantitatíve, azt a történésznek kell megmondani.

*Kelemen József:*

Elsősorban azt szeretném megkérdezni, hogy a társulásban vagy a munkában érdekelt szervezeteknek lenne-e kifogása az ellen, ha a Magyar Nyelvtudományi Társaság Elnöksége előtt javaslatot tennék arra, hogy társuljon a számítógépes vizsgálatokhoz. Amennyiben ennek az együttműködésnek nincs akadálya, a legközelebbi választmányi ülésen előterjeszteném ezt a javaslatot. Azt hiszem, rendkívül hasznos lenne ez a közreműködés a mi társaságunk részére, s mivel a nyelvészetnek sajátos és sok tekintetben eléggé egzakt módszerei vannak, talán bizonyos területen a történész kollégák is nyernének egy ilyen együttműködésből.

Hozzászólásom másik részében szeretnék megerősíteni egy pár olyan gondolatot a nyelvtudomány oldaláról, ami itt a történészek részéről elhangzott. Ilyen pl. a jelentől visszafelhaladás módszertani követelménye. Ezt mi a nyelvtudományban az utóbbi évtizedben megpróbáltuk előkészíteni olyan módon, hogy a Magyar Nyelv Értelmező Szótára munkálatai során elsősorban a magyar irodalmi klasszikusok és a XX. század nyelvének szótári és bizonyosfokú morfológiai elemzésével meghatároztuk azokat a sajátságokat, amelyek nyelvünkre jellemzőek. Ez nagyon nagy munkát jelentett. Kb. 12 évig dolgoztunk rajta, míg 60 000 címszónak ezeket a sajátságait kielemeztük, és 7 kötetben meg is jelentettük. Az máris látszik, hogy az ezután következő történeti vizsgálatokhoz ez a szótár rendkívül hasznos segítséget fog nyújtani; így pl. a legújabbban előkészület alatt álló új magyar nyelvtörténeti szótár jelentés-elemzése minőségileg lényegesen más és jobb lesz az előzőnél, amely a múlt század végén jelent meg. Azért, mert az értelmező szótár jelentés-elemzése nagy mértékben támogatja a történeti emlékek jelentés-elemzését, természetesen mindig figyelembe véve azt a körülményt, hogy időközben a jelentésben is és egyéb részletekben is változások vannak. A párhuzamokat kritikával kell ugyan néznünk, de e párhuzamok mindenesetre gondolatébresztők, s az eddigi vizsgálatok is azt mutatják, hogy az Értelmező Szótár hatása termékenynek bizonyul.

Egy másik gondolat az volt, — azt hiszem Fügedi Erik említette —, hogy a régi szavakat új jelentés jelölésére használjuk. Ez bizony sokszor elég nehéz feladat elé állítja a kutatót: ez a nyelvtudományban közismert. Egészen triviális példákat idézve, gondoljunk arra, mit jelentett a *ház* szó néhány ezer évvel ezelőtt, amikor csak egysejtű házak voltak; mit jelentett néhány száz évvel ezelőtt, amikor már nem egysejtű, de még főleg földszintes házak voltak, s mit jelent ma, az emeletes házak és a felhőkarcolók korában. A fogalomváltozásokat és a fogalomváltozással kapcsolatos jelentésváltozásokat tehát nem lehet figyelmen kívül hagyni. Arra a kérdésre, hogy a *toll* szóban mikor történt meg az a jelentésváltozás, melynek során a madártoll mint íróeszköz helyett a fémtoll mint íróeszköz jelenik meg tudatunkban, egyelőre aligha adhatnánk kronológiailag kielégítő választ. De bizonyos elemzésekkel támpontokat kaphatunk a fogalomváltozások és jelentésváltozások irányára és korára vonatkozólag. A fogalomváltozásra valló szövegrészletek és egyéb adatok tanulmányozását tehát termékenynek és hasznosnak tartom. Jó, ha a történelmi kutatásban is tekintettel van a kutató az ilyen szempontokra.

A történészek is felvetették a szövegelemzés fontosságának kérdését. Ezen a téren már számos tapasztalattal rendelkezünk, s az újabb szövegelméleti vizsgálatok is fontos tanulságokkal szolgálnak. Igaz, elsősorban stilisztikai szempontból végeztek szövegelemzéseket (kulcssza-

vak alapján végzett szöveg- és stílselemzések, egy-egy korszak, stílusirányzat jellemző sajátosságainak vizsgálata). Meggyőződésem, hogy nagytömegű anyagra támaszkodó minták vizsgálata alapján olyan etalonszerű felmérések is végezhetőek, amelyek a történettudomány számára is hasznos szolgáltatokat nyújthatnak.

Intézetünkben eddig elég sok nehézséggel kellett megküzdenünk a gépi statisztikai vizsgálatok előkészítése érdekében, főleg azért, mert vezető nyelvészeink idegenkednek még a gépi vizsgálatoktól. Így sem anyagiakkal, sem munkaerővel nem igen támogattak bennünket. De fokozatosan mégis sikerült elindítani három irányú gépi vizsgálatot intézetünkben. Egyik a magyar XX. századi prózai nyelv hangstatisztikai elemzése; másik egy irodalmi műnek (Az Ember Tragédiája) komplex, ritmikai, stilisztikai szempontból, a szókészlet szempontjából, a szereplők nyelvhasználata szempontjából való elemzése; a harmadik csak most van kibontakozóban: a mai magyar nyelv gyakorisági szótárának előkészítése, ezzel kapcsolatban a leggyakoribb 5000 szó meghatározása, az előfordulások számával, műfajonként és összesítve, az öt műfaj gyakorisági adatai alapján a diszperzió és ennek figyelembevételével a „használat” (úzus) relatív értékének meghatározása. Ha ezeket a munkákat elvégezzük, a szógyakoriságon kívül a szófajgyakoriság, a diszperziószámítás terén is, az egyes stilisztikai-műfaji típusok számzerű jellemzése terén is fontos matematikai-nyelvészeti eredmények várhatók.

Ezenkívül beszélt nyelvi fonológiai (fonéma- és fonémagyakorisági) vizsgálatok készültek intézetünkben, irodalmi nyelvi, köztük költői nyelvi fonémastatisztikai vizsgálatok készültek a debreceni Kossuth Lajos Tudományegyetemen.

Saját problémáink közül legfontosabbnak látszik ez idő szerint megfelelő matematikus-programozó, lyukasztási kapacitás, gépidő és a munkálatokhoz szükséges pénzügyi fedezet biztosítása. Ha ezek megoldásában az érdeklét Társaságok és Intézetek erkölcsi támogatást nyújthatnának, nagyon hálásak lennénk érte.

Felmerült az a kérdés is, hogy a gép hibázik-e vagy nem. Köztünk is volt, aki azt vitatta, hogy a gép nem hibázik, s volt aki szerint a gép nagyon is megbízhatatlan. A kérdés lényege azonban szerintem nem ez, hanem az, hogy a hibák részint gépi ellenőrzéssel, részint emberi munkával kiküszöbölhetőek.

Varga Dénes hangsúlyozta a gépi vizsgálatok sokszempontúságának kérdését. A gyakorisági szótárral kapcsolatban komoly vitáink voltak arról, mennyi adatot vigyünk rá az adathordozóra. A fő ellenvetés az volt, hogy minél több szempontot akarunk érvényesíteni, annál drágább az előkészítés. Bár tudjuk, a gépi adatfeldolgozás annál előnyösebb és kifizetődőbb, minél több szempontból dolgozzuk fel jó szervezéssel ugyanazt az anyagot, mégis tekintetbe kell vennünk az egyelőre rendelkezésünkre álló anyagi kereteket, s a munkákat ennek figyelembevételével kell megterveznünk és szerveznünk.

*Hanak Péter:*

Előjáróban hadd jelezsem, hogy a Történettudományi Intézet is részt kíván venni a munkaközösség munkájában. Ezt a részvételi szándékot annál inkább jelezhetem, mert amint látom, a Történelmi Társulat jelenlevő képviselői is túlnyomórészt az Intézet tagjai, és mert Intézetünkben már több, kvantitatív módszert alkalmazó munkát is folyt, közülük egy, a Puskás Júlia—Láncz Margit készítette mezőgazdasági termelés kvantitatív analízis számítógéppel.

Ami a mostani „alakuló” megbeszélésünket illeti, úgy gondolom, kedvező időpontban jött létre.

Először, a számítógép történettudományi alkalmazásával egy jó évtizedes késésben vagyunk a nyugati, a francia, amerikai, és jó néhány éves késésben a szovjet történészekkel szemben. Ez az év-évtizednyi késés talán ezúttal hasznos lesz: gyakorlati példákon mérhetjük

le a számítógép alkalmazásának lehetőségeit, előnyeit, és egyúttal korlátait, hibáit is, szövege mások kárán is tanulhatunk. Tehát nálunk most olyan időpontban indulhat meg a módszeres, koordinált számítógépes feldolgozó munka, amikor már nemcsak a történeti kvantifikáció csodaszerűen messianisztikus hitével, hanem már a kijózanodás egészséges kritikájával, legalábbis a módszer korlátainak és hibáinak felmérésével indulhatunk el.

Másodsor, úgy hiszem, azért is fölöttébb aktuális a munkaközösség megalakítása, mert már elég sok történész foglalkozik — kézzel, fejjel, géppel — kvantitatív munkálatokkal ahhoz, hogy legyen tényleges anyag, eredmény, amire támaszkodhatunk, és legyen elég nézetkülönbség, amiről elvitakozhatunk. Úgy is fogalmazhatnám, éppen a mai ülés némely benyomásai alapján: elég nagy már a káosz a történészek között a számítástechnika és a számítógép alkalmazása körül ahhoz, hogy valamilyen társulatot, munkaközösséget alakítsunk. A rend — ugyanis mindig a káoszból teremődik.

Harmadsor, azért is hasznosnak tartom a munkaközösség megalakítását, mert mi most itt, együtt, nem valaminő véletlenszerű vagy rétegezett mintát alkotunk, hanem *egy* szempontból válogatott mintát. Úgy látom, azok vagyunk ma itt, akik már eddig is a történeti kvantifikáció és a számítástechnika mellett foglaltunk állást, akik a történettudomány egzaktsága, a korszerű módszerek alkalmazása, a számítástechnika elsajátítása vagy legalábbis alaplemeinek ismerete, a matematikai statisztika technikájának megismerése mellett, és az „impresszionista” anyagfeltárás és feldolgozás, a benyomásokra épített kvalitatív értékelések ellen foglaltunk állást. De a történészek másik táborra, alighanem nagyobb része, akik nincsenek meggyőződve a történeti kvantifikáció és a számítástechnika alkalmazásának lehetőségéről vagy szükségességéről, vagy akik éppenséggel ez ellen vannak, akik ebben valaminő „polgári objektivizmus” modernista változatának becsapódását látják, nincsenek jelen. És úgy vélem, Granasztói barátom kissé optimista módon reménykedik abban, hogy a szovjet példákra való hivatkozással könnyen legyőzhetjük az ellenérzést és ellenállást, amely a történészek eléggé széles és magas körében is él a számítástechnika történettudományi alkalmazásával szemben. Nagyonis aktuális tehát a munkaközösség megalakítása az idegenkedők, a bizalmatlanok meggyőzésére is, a magunk gyengeségeinek kiküszöbölésére is, hogy nem kompromittáljuk gyenge, céltévesztett munkákkal a számítástechnika történettudományi alkalmazhatóságának elvét, programját.

Ezúttal nem kívánok a nagyon érdekes elméleti és módszertani kérdésekhez hozzászólni, bár hasonló kérdésekkel mindannyian, magam is találkoztunk már. Pl. azzal a metodikai kérdéssel, amely Ehrlich Éva és munkatársainak kutatásai kapcsán vetődött fel. A fő probléma a társadalmi struktúra elemzéseknél, úgy vélem, az, hogy az alapkategóriákat már eleve adottnak vesszük, az anyagfelvételt már eleve eme kategóriák szerint végezzük el, ezeket az alapkategóriákat tápláljuk be a gépbe. Nyilvánvaló, hogy a gép ezután csak ugyanazokat a kategóriákat adja vissza, és annyit tesz, hogy számszerűen, vagy százalékos eloszlásban kimutatja a mennyiségi mutatókat. Hogy ilyen volt-e pontosan a társadalom rétegződése egy bizonyos időszakban, az nem bizonyos. Ahhoz másféle módon kellene felvételezni az anyagot, és a tipizálásban is, a kategorizálásban is többet kellene rábízni a gépre, bizonyos, alaposan meggondolt jellemző kritériumok megadása mellett. Hasonlóképpen roppant érdekesek és fontosak azok a problémák, amelyeket Granasztói, Dányi, Perjés kolléga említett. Mégis, úgy vélem, jelenleg nem a problémák felsorolása, jelzése, taglalása az elsőrendű feladatunk. Ezek alkotják *majd* a munkaközösség tényleges tevékenységének tárgyát.

A magam részéről úgy értelmezem az összehívott történészek, számítástechnikai szakemberek, matematikusok feladatát, a mai értekezlet optimális célkitűzését, hogy adjunk ötleteket, témákat, adjuk elő kívánságainkat, elvárásainkat, és ezzel nyújtunk tájékoztatást a megalkuló munkaközösség vezetői számára. Én három fontos programpontra említenék.

Az első feladat a munkaközösség programjának kidolgozása lehetne a következő fél évre, vagy egy évre. Melyek azok a feladatok, amelyek ellátása, azok a témák, amelyek megvitatása erre a társulatra vár. Etekintetben az itt képviselt két szélső vélemény között állok. Nem értek

teljesen egyet azzal, amit Dányi Dezső elsődrendű feladatnak jelölt meg: egy nagy történeti adatbank sürgős felállítását. Ilyenre persze a jövőben szükségünk lesz, de kérdés, van-e elég erőnk e feladat azonnali kitűzéséhez. Ne felejtsük el, hogy a történéseknek konkrét munkáik, sürgető feladataik vannak, amelyeket mielőbb meg kell oldaniok. Mielőbb választ várnak a kutatásaik során felmerült kérdéseikre. (Pl. sürgetnek bennünket egy régebben indult kvantitatív kutatás befejezésére, amelynek témája a kialakuló magyarországi munkásság összetétele. A programozás megkezdődött, én is, más történészek is kíváncsian várjuk az eredményt.) Az említett, és még sok más folyó munkával nem várhatjuk meg, hogy a statisztikai források generális feldolgozása után, egy nagy adatbank segítségével majd erre a kérdésre is választ kaphatunk. (Mellékesen: nem is biztos, hogy az adatbank minden fontos történeti kérdésre tartalmazni fog-e adatokat.) Egyetérték Perjés barátommal abban, hogy nagyonis szükséges kis, jól körülhatárolt vagy határolható források feldolgozása, kis témák kutatása. Ezek is beletartoznak a programunkba. De azzal a végtellessel sem értek egyet, hogy a munkaközösség csupán tudomásul vegye a folyamatban levő munkákat, és nyújtson a már meglevőkhöz elvi-módszertani segítséget.

A munkaközösség programjának másik fontos részét éppen abban látom, hogy ne csupán regisztrálja, hanem valaminő formában koordinálja is a folyamatban levő, esetleges, a kutatók egyéni érdeklődése vagy bizonyos korábbi tervfeladatok által megindított kvantitatív munkálatokat. Koordinálja: ez egyúttal azt is jelenti, hogy igyekszik összhangba hozni a párhuzamos kutatásokat, átadni a hasonló témán dolgozóknak a már kész programozásokat. Koordinálni: ez azt is jelenti, hogy egymást kölcsönösen tájékoztassuk s végül azt is, hogy új témák fontosságára, új lehetőségekre felhívjuk a kutatók és a nagy intézmények figyelmét.

A harmadik feladat az önművelés, a tanulás megszervezése lehetne. Egyrészt szeretnénk tanulni a matematikusoktól, akik érdeklődnek a történelem, a történeti kvantifikáció iránt, és már eddig is sok segítséget nyújtottak nekünk. Másrészt egymástól is tanulhatunk, mindazoktól, akik már ilyen típusú munkát végeztek, kikínlódtak, tapasztalatokat szereztek.

Úgy gondolom, ha meginduláskor nem akarunk túl sokat markolni, akkor e három feladattal meg kell elégednünk, és meg is lehetünk elégedve, ha az új munkaközösség vezetősége a három feladatot meg tudja oldani.

*Granasztói György:*

Egyszerű gyakorlati dolgot szeretnék javasolni, amelynek megvilágításához elmondom saját esetemet. A számítástechnika segítségével végzett kutatásoknál a legnagyobb probléma a tudományos intézetek vertikális tagozódása, ami több diszciplína alkalmi együttműködését csak nehezen teszi lehetővé. Nálunk, a Történettudományi Intézetben én igen könnyen meg tudom értetni saját kutatási célkitűzéseimet közvetlen főnökeimmel, azonban már ők is nagy nehézségekbe ütköznek akkor, ha a többször említett együttes, ú. n. *team* munkát próbálják megszervezni. Saját, már befejezett munkámat, amelynél gépet is igénylő módszert kíséreltem meg alkalmazni, intézetünk igazgatójának személyes támogatása tette lehetővé. Mint a Közgazdasági Egyetem Gazdaságtörténeti tanszéke professzorának módja volt arra, hogy munkámhoz megszerezze a tanszék támogatását, s ezáltal kapcsolatba kerülhettem az Egyetemi Számítógépponttal, az ott dolgozó programozó és matematikus szakemberekkel. A személyes kapcsolat, a szerencse segített. Ez egészségtelen szituáció. Akadémiai intézetünkben ma gyakorlatilag nincsen kerete az ilyen komplex *team* együttműködésnek.

Azt szeretném javasolni, hogy a most szervezendő bizottság vagy valamilyen más szerv nyisson afféle házassági hirdetéshez hasonló rovatot, nyilvántartást és ily módon segítsen, hogy a téma iránt érdeklődő, különböző tudományterületekhez tartozó szakemberek egymásra találjanak, a leendő együttműködés céljából. Pillanatnyilag nem látok más kiutat, mert az megoldhatatlan, hogy egy olyan intézet, mint a történettudományi, most programozó és matematikus szakembert foglalkoztasson teljes munkaidőben. Jelenleg elegendő munkával sem lehetne ellátni.

A feladat tehát az, hogy a történeti jellegű feladatok megoldása iránt felkeltsük a matematikus és a számítástechnikai szakemberek érdeklődését.

Sajnos tüstént felvetődik a kérdés: mennyiben lukratív egy ilyen téma a számítástechnikai szakember számára? Lukratív-e anyagi, de szellemi értelemben is? Csak az utóbbi szempontnál maradva — tehát megéri-e a fáradságot, a történészek által támasztott igények kielégítése egészen primitív programok adaptációja-e, avagy szó lehet-e egyáltalán bonyolultabb, speciális esetekről, amelyek sajátos tapasztalatokkal gazdagíthatnák a *software*-ek készítőit. Elméletileg nem zárható ki az utóbbi eshetőség sem. Ma még mindenestre a kezdeteknél állunk, nehéz előre látni a jövőbe. Javaslatomat megismételve azt szeretném, hogy a történeti problémák számítógépes megoldása iránt érdeklődő történész, matematikus és programozó szakemberek egymásra találhassanak valamelyik tudományos társulat, vagy a most alakuló bizottság védnöksége alatt.

\*

*Az elnöki zárszó* a vita eredményeit elsősorban a jövő tennivalók szempontjából foglalta össze.

A résztvevőkkel egyetértésben javasolta az Elnök a társadalmi keretek között megkezdett közös munka hasonló keretek között való folytatását, hangsúlyozva ennek a szervezeti formának már eddig is bebizonyosodott és várható eredményességét. A továbbiakban a maga részéről különösen olyan vitaulések közös rendezését tartja szükségesnek, melyeken a folyó kutatások ismertetésére és behatóbb megvitatására nyílik lehetőség.

Köszönetet mondva a történészeknek a konferencia előkészítésében végzett szervező munkájukért, biztosította a résztvevőket a Neumann János Számítógéptudományi Társaság további szakmai és szervezeti támogatásáról.

## ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ И ИСТОРИЧЕСКАЯ НАУКА

### РЕЗЮМЕ

25 октября 1972 г. г. Будапешт

Конференция круглого стола, организованная Научным обществом имени Яноша Неймана, Венгерским историческим обществом и Секцией аграрной истории Венгерского научного общества аграрных наук

Первое совещание историков и специалистов по вычислительным машинам в Венгрии было проведено в соответствии с государственным планом по развитию культуры вычислительной техники в Венгрии. В подготовке этой конференции венгерские историки могли опираться на два благоприятных обстоятельства. Во-первых, на последнем международном конгрессе историков (Ленинград—Москва, 1970) был сделан обзор о применении до сего времени вычислительных машин в различных областях историографии, а также был обобщен опыт применения их и были обсуждены спорные вопросы; принимая во внимание эти факты, совещание имело возможность сосредоточить свою работу на отечественной проблематике и на конкретных задачах. С другой стороны, значительная помощь была оказана венгерским Научным обществом вычислительной техники, носящим имя Яноша Неймана, которым проводится интенсивная межпрофессиональная деятельность в интересах преодоления первоначальных трудностей, наблюдаемых в новых областях применения вычислительной техники. Его ценная помощь подчеркивалась и тем жестом, что дискуссию проводил сам председатель этого общества, Ржеже Тарьян.

Участниками совещания, проведенном в узком кругу, были специалисты, интересующиеся применением вычислительной техники в историографии, специалисты вычислительной техники, которые уже до сих пор принимали участие в исторических исследованиях, а также историки, использующие в своей работе ЭВМ (или готовящиеся к этому). На повестке дня фигурировали те актуальные вопросы венгерской историографии, при исследовании которых в первую очередь является обоснованным высокая автоматизация; среди них речь шла о задачах и методах разработки массовых источников, взаимосвязанных рядов данных, имеющихся в распоряжении о территории страны в целом или ее большей части.



## I. Рефераты

1.

**Дьёрдь Гранастон: Вычислительные машины в исторической науке**

Этот дискуссионный,двигающий ряд проблем доклад, в котором автор показывает также и свои собственные исследования, был опубликован в журнале «Тёртенельми Семле» (Историческое обозрение, 1972 год, № 1—2) поэтому здесь отпадает необходимость в его изложении.

2.

**Эрик Фюгеди: История Венгрии и компьютер**

Рассматривая источники, происходящие из периода формирования сообщения статистических данных в Венгрии, автор выдвигает проблемы их разработки выборочным методом, подчеркивая при этом необходимость знания их происхождения и критической оценки.

Во второй половине XVIII века Венгрией, (которая в то время была гораздо больше, чем в настоящее время) управляла королевская администрация, находящаяся в Вене, и бывшее в политическом отношении почти независимым комитатское дворянство, имевшее право самоуправления. Императоры, в том числе Мария Терезия (1740—1780 гг.) и Иосиф II. (1780—1790) прилагали значительные усилия в целях того, чтобы систематически и обстоятельно изучать страну. Мария Терезия, видимо, попыталась перенять систему, которая применялась в то время во Франции. Она придавала большое значение демографии, аграрному производству и ценам на основные пищевые продукты. В интересах получения необходимой информации ей было отдано распоряжение о том, чтобы комитаты и свободные королевские города из года в год представляли ей сведения и о движении народонаселения (*Conscriptio Animarum*), о посевах и сборе урожая (*Tabella procreationis frugum*) и т. д. Однако ее усилия были обречены на неудачу в результате сопротивления дворянства, которое не желало предоставлять сведений ни о чем, что было связано с дворянством (дворяне, например, были исключены из статистики по движению народонаселения). Иосиф II, которому наскучили споры, распорядился о переписи населения (в том числе и дворянства, 1784—1787 гг.), используя для этого армию, по его указу офицеры королевской армии составили земельный кадастр страны (опять же вместе с дворянскими владениями). Возмущенное дворянство после смерти Иосифа II в большинстве комитатов уничтожило данные замеров и переписи, и волна возмущений на долгое время сделала невозможным предоставление каких-либо статистических сведений.

Правда, были административные области, где начинания королевской власти оказались удачнее. Так как её представители располагали полной властью королевскими правами (регалиями), император мог получить любую информацию, какую только хотел. Венгерская палата систематически предоставляла сведения о внешнеторговом товарообороте, о производстве благородных металлов и т. д.

Небольшие исследования, проведенные в наших архивах, также сделали очевидным, что несмотря на все трудности, которые нужно было в то время преодолеть королевскому управлению, а также несмотря на варварскую отбраковку (селекцию), в ходе которой в прошлом веке были уничтожены различные статистические отчеты, мы всё же располагаем ценным материалом. Этот материал кажется достаточно надежным для того, чтобы можно было дать основанный на солидной статистике обзор о демографическом, экономическом и социальном положении Венгрии конца 18 века. Поскольку это можно констатировать на данной стадии исследований, мы имеем несколько статистических съемок, охватывающих всю страну, и целый ряд отрывочных переписей. Этот факт сам по себе предлагает воспользоваться методом выборки, что еще подтверждается тем обстоятельством, что, как нам известно, Венгрия 18-го века была раздробленной во всех отношениях. В Венгрии существовали дворяне и крепостные, помимо католической церкви имелось три протестантских и православная церковь, население Венгрии слагалось по меньшей мере из шести национальностей. С экономической точки зрения различия были еще сильнее. В стране имелись такие районы, где господствующим было монокультурное растениеводство (главным образом выращивание винограда), в других местах основным источником существования было животноводство и рыболовство. Излишне говорить о том, что в такой обстановке культурная жизнь также не могла быть повсюду одинаковой.

Автор подчеркивает, что он полностью отдает себе отчет о стоящих перед нами в связи с этим трудностях. Не существует такого выборочного метода, который был бы панацеей на все случаи; во всех областях — демографической, в сельском хозяйстве и т. д. — необходимо составлять отдельно расчлененные выборки, но, как полагает автор, это единственный способ для того, чтобы найти выход из положения, и одновременно это единственная возможность для того, чтобы мы смогли использовать также и имеющиеся в нашем распоряжении отрывочные материалы.

## 3.

### Имре Велльманн: Обработка наших аграрно-исторических источников и вычислительная машина

В историографии, в первую очередь в социально-экономической истории имеются возможности и потребности количественного анализа, вследствие чего в этом отношении уже имеет за собой значительную давность стремление к тому, чтобы на место «импрессионистской историографии» вступила реконструкция прошлого, которая может быть выражена в измеримых фактах. Аграрная история особенно не может обойтись без машинной обработки данных.

1. *Необходимость машинной обработки данных* — это не вопрос моды в науке, и даже речь идет не о том, чтобы примкнуть к техническому прогрессу. В Венгрии наши исторические источники, пригодные для машинной обработки, в своей основной массе связаны с нашим аграрным развитием по той причине, что Венгрия в течение многих столетий по сути дела была аграрной страной (жизнь наших городов в значительной части также базировалась на сельском хозяйстве). Уже объем этих источников и обилие, многообразие данных делают необходимым применение вычислительных машин как в целях документации, так и статистического анализа, но еще более настоятельно проявляется тот факт, что аграрная история — исходя из своего особого места в иерархии наук — тесно соприкасается с другими смежными науками (как этнография, география, аграрные науки, экономика, социология), и поэтому необходимо принять во внимание также позиции этих наук. Это в свою очередь делает необходимым такой комплексный анализ проблем, который не может быть проведен обычными методами.

2. *Что ожидает аграрная историография от планомерной и систематической машинной обработки данных?* Ни в коем случае лишь ускорения рабочих процессов, которые могут быть выполнены мануальным способом. С одной стороны, ожидается распространения исследования на до сих пор еще не исследованные «белые пятна», с другой стороны, углубление исследований (выявление вышеупомянутых комплексных взаимосвязей и относительно веса отдельных факторов, по возможности наиболее полное раскрытие внутренней структуры и динамизма явлений). В обоих случаях речь идет о реально обоснованных ожиданиях — это документируется и подтверждается видимыми уже и в настоящее время результатами тематических исследований. Кроме того, анализ, проведенный с помощью вычислительных машин, может содействовать критике исторических источников, соответствующей их оценке.

3. *Задачи историка-аграрника в целях успешного использования вычислительной техники* выходят за пределы того, что он должен познакомиться с математическими методами, связанными с машинной обработкой данных. Необходимо еще и следующее:

а) *Отбор пригодных источников.* Для количественного анализа в первую очередь запрашиваются такие исторические первоисточники, которые охватывают либо как можно более широкую область по единым точкам зрения (по возможности всю страну) — и, таким образом, пригодны для начертания подробного, многостороннего разреза в данный период; или же (и это более редкий случай) в хронологическом аспекте образуют по возможности обширные и не имеющие пробелов ряды данных, точно так же по единым точкам зрения — и таким образом предоставляют основу для начертания тренда развития.

Эти условия, как известно, наличествуют тем менее, чем к более раннему историческому периоду мы обращаемся.

Уже и эти предпосылки побуждают к тому, чтобы начать работу там, где в нашем распоряжении имеются наиболее полные и наиболее обоснованные данные. Что касается истории чешской Венгрии, это конец 19-го века, когда мы можем опираться уже на данные действительно статистического характера и на однородные данные. Но необходимо обратить внимание и на то, что в аграрном развитии различные природные усло-

вия, условия рынка и т. д. явились причиной возникновения многообразных региональных, более того, местных вариантов, и эта многогранность — вследствие усиленного воздействия природного фактора — обычно в нарастающей мере имела место в преиндустриальную эпоху. *Если мы хотим получить действительно полную, всеохватывающую картину о компонентах нашего аграрного развития*, наиболее целесообразным будет, если мы сделаем реальный разрез о конце 19-го столетия. Находим твердую почву на этой прочной основе, мы получим сопоставимую базу и с большей уверенностью сможем продвигаться назад во времени к эпохам, документированным с большими пробелами, и на основе этого мы получим возможность для разработки трендов развития на все более отдаленные периоды.

б) *Критика источников*: задачей историка является тщательное исследование того, насколько надежен источник в целом и его отдельные данные или группы данных, насколько можно пополнить пробелы данными, взятыми из других источников, и в каком отношении пригодны они для машинной обработки.

в) *Подготовка программы*: Квантитативному анализу должен предшествовать качественный анализ, построение гипотезы исследования, создание модели, которая делает возможным выбор наиболее пригодных индикаторов. За этим следует презентирование данных источников в форме, пригодной для кодирования — и это тот пункт, где в первую очередь необходимо вышеупомянутое знание математических методов.

4. *Путь дальнейшего движения вперед*. Перед нами стоят большие задачи, и успешно шаг за шагом планомерно продвигаться вперед мы сможем лишь объединенными силами. Залог успеха прежде всего кроется в систематическом сотрудничестве историков-аграрников с деятелями вычислительной техники, которое до сих пор проявлялось лишь в разрозненных индивидуальных попытках. Для этого, с одной стороны, необходимы специалисты вычислительной техники, ближе знакомые с проблемами исторических наук (чему имеются примеры как в международном масштабе, так и у нас в стране), с другой стороны, нужны историки-аграрники, которые не чуждаются математических методов. Благодаря этому будет возможным не только построение прочной и широкой базы источников, но и — с использованием отечественного и зарубежного опыта — исследование истории аграрного дела в целом и разработка её смогут получить невиданный до сих пор подъем.

#### 4.

### Ласло Катуш: О применении вычислительной техники в исследовании эпохи капитализма

1. *Что сделано до сего времени и что находится в стадии разработки?* Как известно автору, до сего времени в Венгрии по теме этой эпохи было проведено, вернее проводятся два исследования с помощью вычислительной машины:

а) составление таблицы выпуска — затрат фабрично-заводской промышленности Венгрии на базе фабрично-заводской статистики 1898 года.

б) статистический анализ выращивания полевых культур в период с 1867 по 1914 годов и сопоставление в международном масштабе.

Работы *Шимонич Дьердьне* и *Юлиш Пушкаш*. В Америке Эбби М. Скотт производил исследование с помощью вычислительной машины на базе данных венгерской внешней торговли за период 1882—1913 гг.

2. *Возможности и требования будущего*. Период с 1850 по 1945 год в Венгрии дает весьма большие возможности для проведения исследований с помощью вычислительных машин. В случае целого ряда тем квантитативные исследования, проведенные с помощью традиционной техники, уже подошли к тому, что их продолжение и дальнейшее развитие возможно лишь с помощью машинных средств: с одной стороны, мы имеем много таких источников или типов источников, которые требуют обработки с помощью вычислительной техники, с другой стороны, в ходе исследований, проведенных за последние годы, возник целый ряд проблем, успешный ответ на которые мы сможем получить лишь путем статистического анализа, проведенного с помощью вычислительных машин.

а) *Источники, типы источников*, в которых сообщается ограниченное количество данных идентичного типа, которые могут быть легко стандартизированы касательно большого числа (порядка тысяч или десятков тысяч) единичных случаев. Это отчасти печатные, отчасти архивные источники: адресные книги (наиболее важные адресные книги

землевладельцев), альманахи, реестры торговых фирм, компасы, регистры налогоплательщиков рекрутские списки, метрические книги и т. д. Далее сюда могут быть отнесены такие обширные источники, как донесения полиции о рабочем движении, журналы заседаний парламента, а также генеалогические справочники и биографические лексиконы. Далее, необходимо проводить обработку огромной массы данных, содержащихся в прежних официальных статистических изданиях, согласно новым критериям, выдвинутым современным исследованием, согласно новым методам классификации, а также проводить их статистический анализ на более высоком уровне.

б) *Проблемы*, требующие разработки с помощью вычислительных машин, существуют во всех областях исследования. Здесь говорится только о тех проблемах, которые уже действительно поставлены на повестку дня, и уже были совершены попытки их обработки традиционной техникой.

*Экономическая история*: наиболее важные агрегатные показатели роста национальной экономики (компоненты национального продукта и национального дохода, а также факторы, действующие на их формирование), работы по истории цен и заработной платы. — *Социальная история*: структурные анализы в общегосударственном или местном масштабе, в разрезе всего общества или же относительно отдельных классов и слоев; коллективная биография (реконструкция одного слоя или группы на базе анализа единичных случаев или большой массы), анализ общественных массовых движений (например, стачек), демографические исследования разработка наиболее важных демографических показателей в масштабе страны. — *История политики*: исследование поведения избирателей и депутатов парламента, а также исследование влияющих на это факторов. — *История культуры*: составление статистики по делу просвещения на длительный период в масштабе страны; статистический анализ книгоиздания; анализ формирования идеологических течений, оказывающих влияние на общественное мнение, на базе данных, предоставленных средствами массовой коммуникации (печати).

3. *Характер исследований*. В первом приближении речь может идти о работе двух типов:

а) Описательная разработка источника или группы источников идентичного типа в полном объеме; кодирование данных, нанесение на перфокарту или магнитную ленту, фундаментальная работа по классификации и группировке с выполнением первичной статистической обработки;

в случае источников такого типа создание центрального «компьютерного архива» (банка данных), основные данные которого, хранимые на магнитной ленте, любой исследователь может использовать для дальнейшей статистической обработки в соответствии со своими проблемами.

б) Выявление связей и моментов, скрытых в случае традиционных методов исследования, для решения определенных проблем с помощью более сложного статистического анализа,

в этом случае анализу подвергается несколько источников, касающихся этого вопроса, в большинстве случаев путем выборочного метода.

4. *Излагая задачи исследований, решаемых с помощью вычислительных машин*, автор хотел указать, какой путь открывает вычислительная техника перед исследованием, поскольку вышеупомянутое обобщение даже при оптимальной оснащенности ЭВМ представляло бы собой многодесятилетнюю программу исследования. Автор подчеркивает, что решение аналогичных проблем известно в зарубежной литературе. Опыт зарубежных исследований показывает, что более крупные расходы, связанные с применением вычислительной техники, в таких случаях окупаются большими результатами. Автор вносит ряд предложений, первостепенным из которых он считает *создание центрального компьютерного архива* — с обработкой определенных источников или группы источников.

## 5.

### Юлианна Пушкаш—Маргит Ланц: Квантитативный анализ производства полевых культур в Венгрии в 1867—1913 гг.

Целью рассматриваемой работы, подключающейся к реферату Ласло Катуш, было проведение математико-статистического анализа производства полевых культур и подведение вытекающих из этого итогов по экономической. Авторы здесь в первую очередь дают методическую информацию.

Исходный материал источников: динамические временные ряды годового производ-

ства всех полевых растительных культур, цены производства и экспортируемого количества. На базе этих рядов динамики были выполнены следующие расчеты:

1. Анализ динамических рядов начали с самого простого и наиболее широко применяемого метода — с составления *чисел-индексов* и *графиков*. Для показа тенденций развития прежде всего были образованы ряды индексов: а) ряд базисных индексов (всегда на основании средней величины 1909/13 годов), б) ряд цепных индексов. Подавляющее большинство рядов индексов — индексы объемов. К количествам, которые не поддаются непосредственному суммированию, для приведения к общему знаменателю различных качеств, привлекли также и индекс стоимости. Формирование валового производства всех полевых растительных культур и производства отдельных их групп прослеживали с использованием неизменных цен. Ряды индексов, образованные на базе взвешенной цены производства 1909/13 годов с неизменными ценами, здесь функционируют как средства приведения различных потребительных стоимостей к общему знаменателю, и как средства фактического выражения стоимости, и в то же время они также являются носителями объема производства.

2. Движение динамических рядов в длительной перспективе времени, основная тенденция развития производства было вскрыто *трендовыми расчетами*. Исходя из самых простых расчетов тренда, в первую очередь рассчитали тренды с помощью исчисления т. н. подвижной средней. Полученное таким образом сглаживание колебаний динамических рядов уже облегчило ознакомление с природой временных рядов; полученная таким образом кривая тренда относительно хорошо приспособлялась и обозначала также те наиболее крупные отрезки (этапы) роста производства, о которых упоминается также и в текстах исторических источников. Однако такого рода расчеты тренда еще не позволили получить результатом кривую, касающуюся временных рядов в целом, которая аналитически известна, и которой легко пользоваться в работе.

Поэтому после составления временных рядов трендовых стоимостей с подвижным средним и графического их изображения перешли к *аналитическим* трендовым расчетам. Вовсе нелегко было определить, какой из типов тренда лучше всего коррелирует с отдельными динамическими рядами. Для этого использовалось также несколько различных методов (изучались графики временных рядов, на глазок проводили медиану, исследовали трендовые кривые с подвижной средней, принимали во внимание методы трендовых расчетов, примененные с аналогичного типа временными рядами, и не в последнюю очередь, учли также и информации текстов исторических источников. Все эти соображения в совокупности содействовали авторам в выборе тех типов математических функций, которые могли бы правильно отражать устойчивую тенденцию временных рядов. (Корреляция проверялась методом наименьших квадратов. С временными рядами, включающими в себя складывание производства, в большинстве случаев коррелировали линейные, а в некоторых случаях — экспоненциальные тренды, а для цен во всех случаях стало необходимым сопряжение (совмещение) парабол.) После составления временных рядов, полученных аналитическим расчетом тренда, были начерчены линии тренда (полулогаритмическим масштабом), и это дало изогнутую кривую и для линейных трендов. Этот метод изображения не только хорошо показал медиану, проведенную в ряду динамики, но показал также и большие отклонения между исходными (первоначальными) динамическими рядами и динамическими рядами, созданными из трендовых величин.

3. Для того, чтобы выявить, в какой степени необходимо было сгладить исходные (первоначальные) временные ряды, и также показать, насколько отличаются фактические данные рядов от трендовых величин (или, по другой терминологии, от средних величин), были исчислены отклонения данных исходного динамического ряда от данных временного ряда трендовых величин. Отклонения, получившиеся из года в год, показаны рядом *девиаций*. (Для большей наглядности и сопоставимости отклонения от трендовых величин были выражены в процентах.)

В своих исследованиях авторы придавали чрезвычайно большое значение исследованию девиаций, затем на основании их — исчислению показателей дисперсии; так как они сознавали ту математическую абстракцию, то абстрагирование от реальной действительности, к чему они прибегали при вычислении средних, т. е. трендовых величин. (Или наблюдалось также и то, что эта абстракция является одной из причин того, почему историко-экономикой чуждаются трендовых расчетов). По их мнению, эта математическая абстракция необходима именно в интересах более глубокого и точного познания исторической действительности; главной целью и доводом для расчетов средних величин для историков является именно то, чтобы они могли осветить и исторически проанализировать отклонения от этих средних величин, дисперсию. Ибо современные статистические методы подтверждают, что статистические совокупности освещаются по-другому, если нам известны их отклонения и дисперсии. Благодаря расчетам, направленным на это, можно более уве-

ренно опознать или показать нарушения порядка, проявляющиеся во временных рядах или характеризующие их особенности. В результате знания средних величин рядов и отклонений от них, в итоге исследования дисперсии можно опознать такие особенности производства, которые могли бы остаться нераскрытыми, если бы мы не воспользовались математическим статистическим анализом.

4. После составления временного ряда девиаций последовало исчисление следующих *показателей дисперсии*: абсолютная средняя величина девиаций, и квадратичная средняя величина девиаций, т. н. стандартная девиация. Эта последняя устраняет влияние количественных различий между отдельными рядами, и в результате этого является самым чистым показателем для сопоставления отдельных временных рядов.

5. Динамический ряд девиаций означал не только колебания исходного временного ряда по отношению к медиане, но выявил также и разрывность, цикличность отклонений. Для фиксирования этого, из временных рядов девиаций (аналогично исходным динамическим рядам с исчислением 7-летней подвижной средней) был рассчитан тренд, и, приняв величину тренда за 100, линия тренда также была изображена графически. Таким образом можно было еще лучше выяснять природу временных рядов, так как в составе линии перспективного тренда раскрылись периоды средней продолжительности, периоды ускоряющегося и замедляющегося темпа роста.

6. И в заключение последние производные или расчетные временные ряды содержат так называемые *вторичные девиации*. Приняв величины образованных из девиаций трендов за 100, исследовали ежегодную флуктуацию, колебания временного ряда в целях раскрытия его более коротких циклических движений.

Составленные с помощью математических статистических расчетов временные ряды и графики уже предоставили широкую базу и для исследования динамики развития. Показателями его являются непосредственно *средние темпы роста*. Для временных рядов в целом это рассчитывалось следующим образом: 1) из величины тренда среднего темпа роста, 2. из средней величины первого и последнего пятилетия временных рядов. Средний темп роста всех динамических рядов в целом также является математической абстракцией. В действительности темпы роста — как уже было показано рядами ценных индексов — не были аналогичными даже за два года. Средняя величина, то-есть средний темп роста всего временного ряда, важен для историков в первую очередь потому, что мы сможем узнать величину отклонения от него, сможем исследовать дисперсию относительно средних темпов роста.

Из временного ряда девиаций, главным образом из линии тренда уже становилось ясным, что наряду с ежегодными изменениями темпов роста проявляются также и определенные цикличности, то есть изменения направлений средних темпов роста, ускоряющиеся или замедляющиеся периоды. Эти циклы старались выявить также двумя различными путями: 1) проводили расчет средних темпов роста на десятилетние периоды, покрывающие друг друга, а также 2) на те этапы, которые девиации характеризовали как изменения направления темпа роста. — Применялись математические методы для определения и анализа изменений структуры производства полевых культур. Методом элиминирования изыскивали ответ на вопрос о том, в какой мере влияли на рост три основных фактора: увеличение посевной площади, преобразование структуры посевов и повышение урожайности отдельных культур. Для каждой растительной культуры порознь были проведены также те измерения, с помощью которых можно было выяснить, в какой мере формирование посевной площади или урожайности оказывало влияние на складывание производства отдельных растительных культур.

7. В интересах проведения сравнительного анализа в международном масштабе и более точной оценки развития Венгрии — в пределах возможности — аналогичные анализы проводили авторы также и с временными рядами производства полевых культур во многих других европейских странах. Наряду с сопоставлением производства полевых культур в Австрии и в Венгрии был проведен совместный анализ двух рядов данных, который позволил построить перспективные тренды для основных сельскохозяйственных продуктов, а также исследование роста в масштабах всей австро-венгерской монархии.

6.

### **Атила Чернок—Ева Эрлих—Дьёрдь Силадьи: Об инфраструктуре на базе Международное сравнительное исследование исторических статистик (1860 1968 гг.)**

Исследование, проведенное коллективом экономистов, репрезентировало на конференции проблему приближения экономической истории и современного экономического исследования. Само исследование в первую очередь касалось следующих вопросов:

1) Как формировалась инфраструктура на различных этапах экономического развития различных стран.

2) Какой уровень устроенности инфраструктуры — принимая во внимание технический уровень эпохи — соответствует определенному уровню экономического развития данной страны.

3) На определенных уровнях экономического развития «чего стоит» или чего требует от народных хозяйств создание инфраструктуры, её поддержание, модернизация и расширение.

Международное историческое исследование инфраструктуры охватывает период в 100 лет (1860—1968 гг.).

О понятии инфраструктуры: в принципиальном смысле под инфраструктурой мы понимаем ту часть национального богатства, которая не служит непосредственно ни созданию материальных благ, ни их потреблению, а которая на данном уровне экономического развития, в соответствии с всегдашними требованиями техники, призвана обеспечить свободный простор, так называемую сосудистую систему, для процесса производства — распределения — потребления.

Для исторического сопоставления инфраструктуры в международном масштабе служит специальная система показателей, созданная для целей исследования.

Типы индустриализации и инфраструктурного развития: из результатов исчислений вырисовывается три основных типа. Представителями первого типа являются Англия, Голландия и несколько западноевропейских государств. Здесь формирование инфраструктуры по существу предшествует быстрому промышленному развитию. Инфраструктурные рамки, созданные в ранние периоды, обеспечили возможность последовательного распространения, начальный простор и для позднее начавшейся индустриализации. Экономика Соединенных Штатов Америки репрезентирует второй, более современный тип развития. Здесь одновременно с развитием и распространением капитализма, почти без предварительных предпосылок, складывалась инфраструктура с постепенным образованием потребностей, предъявляемых, производящими и потребляющими отраслями. Вначале этому содействовал также и иностранный капитал. Для английской модели характерно опережение по времени, а для американской модели характерно развитие инфраструктуры, которое неразрывной цепью сопрягается с распространением капитализма. Часть стран Центральной и Восточной Европы, особенно в период после первой мировой войны, по существу не прошли путь того структурного преобразования инфраструктуры, которое в то же самое время происходило в значительной части стран Западной Европы. Тот тип экономического развития, который был осуществлен европейскими социалистическими странами после второй мировой войны, не сходится ни с одним из двух вышеуказанных типов создания инфраструктуры. Для этого типа не характерны ни предварительное создание экономических условий, ни «сцепленное» совместное развитие. Здесь мы имеем дело с таким третьим типом, в котором развитие инфраструктуры как бы отсутствует некоторое время. В этих странах до периода середины 1960-х годов мы можем говорить скорее об утилизации наличествовавшей несвоевременной инфраструктуры.

Что же касается роли факторов производства в инфраструктуре, то можно установить, в складывании относительно соотношения занятых в инфраструктуре с общей численностью занятых находит выражение своеобразный сплав фактора времени и фактора экономического развития. Из этих двух факторов решающую роль играет последний.

Одним из наиболее важных свойств инфраструктуры является высокая капиталоемкость, потребность в инвестициях. Относительное соотношение инфраструктурных капиталовложений с общими капиталовложениями колеблется в широком интервале в зависимости от времени и от государства, но в самых редких случаях снижается до уровня менее 50%, а его верхний предел достигает примерно 80%. Второй характерной чертой инфраструктурных капиталовложений является то, что их величина обычно не зависит от экономического развития. — В социалистических странах доля инфраструктурных капиталовложений гораздо меньше по сравнению с капиталовложениями в целом, чем в капиталистических странах.

Развитие инфраструктуры является медленным процессом, и, как правило, требует чрезвычайных усилий от народных хозяйств.

#### 7—8.

В связи с рефератом по аграрной истории Имре Вельманна публикуются две специальные работы по применению вычислительных машин. В этих работах с привлечением коллективов специалистов в составе, соответствующей решаемой задаче, рассматриваются и подготавливаются общие методы, которые могут быть использованы в общих работах по разработке источников и обработке данных по аграрной истории.

## Денеш Варга Ева Вереш: Разработка переписей земельных владений феодальной эпохи с помощью вычислительных машин

С помощью математиков-лингвистов составляются индексы источников по группе переписей земельных владений, сохранившихся от периода 1526—1848 гг. Число таких кадастров составляет порядка десятков тысяч. При выполненной до сего времени с помощью вычислительных машин систематизации резюме примерно 15.000 источников используется элементарная классификация, дающая возможность быстрого обзора источников (в период конференции проводилась её непосредственная подготовка). Для этого участвующие в работе математики-лингвисты и историки разработали такую систему, которая принципиально в дальнейшем может быть расширена произвольно по желанию, вплоть до подробной тематической разработки. Это может стать необходимым при разработке отдельных, менее крупных разделов, частей материала, избранных с помощью вычислительных машин.

Разработке с помощью вычислительных машин предшествовали эксперименты по применению мануальных документационных средств (традиционная система каталогов, перфокарты). Эти работы понятным образом не принесли желаемого результата, и трудности, связанные с подготовкой, содержанием в порядке и использованием такого количества материала, стали очевидными. И всё же авторы считают весьма целесообразной предварительную подготовку машинной обработки материалов, осуществляющуюся с помощью мануальных документационных средств; обладая предварительными информацией, полученными о материале, можно более экономично организовать машинную обработку данных, можно приобрести немало полезного опыта и не в последнюю очередь: большинство проблем систематизации, определения понятия выдвигается уже на этом этапе, в силу чего документационная система, приспособляющаяся к материалу, попадает в машину уже в зрелой форме. (В 1-ой части доклада рассматривается целесообразная комбинация мануальной и машинной систематизации.)

При составлении индексов переписей земельных владений необходимо решить специальную историческую задачу. В этом отношении эта работа принимает совершенно другой характер, чем доклада (например, библиотечных и т. д.). Прежде всего необходимо было обеспечить осуществление точек зрения, связанных с критикой источников. (Для этого основой явились данные критики исторических источников, содержащиеся в резюме, составленных на источники; по ним можно было составить аттестации относительно достоверности документов и данных.)

С исторической точки зрения одной из наиболее важных частей составляющегося регистра является географический указатель, составление которого является, пожалуй, самой интересной частью работы с точки зрения применения вычислительной техники. Вследствие непостоянной орфографии географических названий и частого изменения названий поселений, а также неоднократного переустройства административных единиц кажется нерезальным составление географического индекса из массы данных,<sup>1</sup> охватывающих более четырех столетий и несколько сотен тысяч географических названий. В машинной обработке данных задачей здесь является не сличение (сверка, согласование) географических названий по буквам, так как таким путем может быть упорядочена лишь часть этого материала. Здесь необходимо решить проблему, автоматического отождествления географических названий в различных по форме вариантах, и, наоборот, по одинаковым по форме географическим названиям нужно установить, к какому физическому объекту они относятся в данном контексте. Эта проблема до некоторой степени является языковой аналогией распознавания образов, при условии когда эталон дается заранее (например, в форме существующих в настоящее время географических словарей (топографических каталогов). — В 3 и 4 части очерка рассматриваются результаты, полученные из эмпирического исследования, а также метод, разработанный для подхода к решению этой проблемы с помощью вычислительных машин.

## Ласло Чех — Дьёрдь Харасты Ева Верешш: Топографический метод хранения и обработки данных в аграрной истории (Составление картографических диаграмм с помощью вычислительных машин)

В статье описывается два вида картографических диаграмм, начерченных с помощью построено печатающего устройства (принтера) ЭВЦМ:

1. Точечные диаграммы, на которых одна географическая единица (например, поселение) изображается в виде точки; точка изображается каким-либо письменным зна-



ком, который соответственно теме выбирается из набора знаков принтера. У одной тематической карты определенная серия письменных знаков соответствует серии качественных категорий (например, серия цифровых знаков может обозначать категории количественного порядка). Эти диаграммы служат не для визуального просмотра, визуальность в них осуществляется лишь постольку, поскольку необходимо избирать письменные знаки, хорошо различимые друг от друга; их целью является просто фиксирование данных или полученных в ходе обработки материалов результатов в географическом порядке. Это рабочие карты, которые исследователю более легко проанализировать, чем сводки в виде таблиц. В них осуществляется известная редукция, выполненная согласно заранее заданным аспектам. Они могут быть дешево изготовлены в больших количествах и с ними легко обращаться со всех точек зрения.

2. По существу по аналогичной технологии изготавливаются картографические диаграммы, пригодные также и для визуального просмотра, и более всего приближающиеся к графическому выполнению. Отдельные территориальные единицы внутри изображаемой географической площади (например, комитаты внутри территории страны) составляют поверхности, заполненные письменными знаками, соответствующими их аттестации. Путем выбора знаков можно достичь тех визуальных контрастов, которые достигаются штриховкой различной силы при вычерчивании карт. Напечатанная таким образом карта дает уже проанализированную обзорную картину об окончательных результатах обработки данных.

Эти диаграммы в виде карт близки к тематической картографии, но их нельзя смешивать с автоматизированной картографией. Их составители придерживаются принципов картографии, насколько этого позволяет техника, но в окончательном итоге эти диаграммы — всего лишь простые технические пособия, и как таковые они означают лишь один из этапов процесса обработки данных. Существенной чертой применяемой техники является то, что картографическая диаграмма органически связана с процессом обработки в целом, обработка данных, поставленных уже на носитель данных вычислительной машины, может быть завершена без перерыва, и полученные результаты по желанию могут быть занесены на картографическую диаграмму, таблицу или в любой список.

Осуществление географической точки зрения в такой большой мере в процессе обработки данных дало возможность (а также и содействовало) тому, чтобы хранение данных более крупного количества источников также было осуществлено по единым географическим признакам, была создана система обработки данных, пригодная для региональных исследований таким образом, чтобы её можно было легко использовать для обработки данных источников различных эпох и различного состава. Создание этой системы требует значительных предварительных работ, но от нее можно ожидать и значительных результатов. Первые подготовительные работы были начаты на основании опыта, приобретенного в связи с описанными здесь первыми картографическими диаграммами.

## ДИСКУССИЯ

В ходе оживленной дискуссии, завязавшейся в несвязанной регламентом части совещания, участники — наряду с выдвижением отдельных конкретных, немаловажных деталей вопросов и замечаний — по сути дела единодушно высказались за необходимость планомерного распространения количественных исследований, углубление тематики и организационной сплоченности. Было подчеркнуто требование обоснованности количественных исследований, необходимость соединения их с комплексными исследованиями. К этой точке зрения присоединились также и присутствовавшие на конференции представители родственных отраслей наук — в первую очередь лингвисты, предложив при этом свое сотрудничество.

В других областях применения вычислительных машин при исследовании истории наук (например, при анализе текстов, анализе содержаний и т. д.) в Венгрии до сего времени не было или почти не было инициативы. Исследователи учитывают имеющиеся в этих областях возможности.

Совещания совместно с представителями-специалистами по вычислительной технике желательно проводить и в будущем, в отношении подготовки и усовершенствования специалистов на них также можно возлагать максимальные надежды. В итоге конференции участниками были вынесены конкретные практические предложения в сторону научных органов-организаторов конференции о продолжении сотрудничества между различными областями наук.

## ORDINATEURS ET L'HISTORIOGRAPHIE

(Budapest, le 25 octobre)

Conférence de Table Ronde de l'Association János Neumann des Sciences Cybernétiques, de la Société Hongroise d'histoire, et de la Section Historique Agraire de l'Association Hongroise des Sciences Agraires

La première conférence interprofessionnelle des experts des ordinateurs et des historiens fut organisée dans le cadre du plan national visant le développement de la culture des ordinateurs en Hongrie. Lors des travaux de préparation, les historiens hongrois furent appuyés par deux facteurs. D'une part, la dernière conférence internationale des historiens (Léningrad—Moscou, 1970) — tout en mettant au jour tant les expériences que les problèmes — donna une vue d'ensemble sur l'application jusqu'à présent des ordinateurs dans plusieurs domaines de l'historiographie; prenant en considération ces conclusions, les délibérations pouvaient être concentrées sur les problèmes du pays et sur les tâches concrètes. D'autre part fut considérable l'appui reçu de l'Association Hongroise des sciences cybernétiques portant le nom de John von Neumann — qui développe une activité intense interprofessionnelle pour pouvoir surmonter les premières difficultés se manifestant dans les nouveaux champs d'application.

Les participants du colloque tenu dans un cercle restreint furent des experts des ordinateurs s'intéressant à l'utilisation historiographique des ordinateurs ou ayant déjà participé à des expériences, ainsi que des historiens qui s'étaient déjà servis — ou bien voulaient le faire — des ordinateurs. A l'ordre du jour figurèrent les problèmes actuels de l'historiographie hongroise, dans l'étude desquels l'automatisation est avant tout indiquée; surtout il fut question des méthodes et des tâches du dépouillement des sources massives (séries continues de documents couvrant une partie du pays ou le pays entier).

### Rapports

#### 1.

#### György Granasztói: Ordinateurs dans l'historiographie

(Ce rapport initial, qui posa les problèmes et présenta les recherches de l'auteur, fut publié dans le numéro 1—2. XV. de *Történelmi Szemle*, c'est pourquoi il n'est pas analysé ici.)

#### 2.

#### Erik Fügedi: Histoire de la Hongrie et l'ordinateur

Passant en revue les sources remontant à l'époque où commencent les relevés statistiques en Hongrie, l'auteur posa les problèmes du dépouillement de ceux-ci par l'échantillonnage, tout en soulignant la nécessité de connaître leur genèse et de les soumettre à une analyse critique.

Dans la deuxième moitié du XVIII<sup>e</sup> siècle, la Hongrie d'alors (qui était beaucoup plus grande que celle de nos jours) fut gouvernée par des rois siégeant à Vienne et par la noblesse des comitats presque indépendante au point de vue politique et disposant d'une autonomie. Les souverains, notamment Marie-Thérèse (1740—1780), Joseph II (1780—1790) faisaient de grands efforts pour connaître le pays d'une manière approfondie, systématique. Marie-Thérèse essayait — paraît-il — la transplantation du système appliqué dans la France contemporaine. Elle attribua une grande importance à la démographie, à la production agricole, aux prix des provisions de base. Pour se procurer les informations nécessaires elle ordonnait aux comitats et aux libres villes royales de fournir d'année en année des renseignements sur la démographie (*Conscriptio Animarum*), sur la moisson et les semailles (*Tabella procreationis frugum*) etc. Ces efforts échouèrent à cause de la résistance de la noblesse qui se refusa à fournir des données touchant ses affaires. (Les nobles furent omis par exemple dans les statistiques démographiques). Joseph II, las des discussions, se servit de l'armée, organisa un recensement de la population (y compris les nobles) (1784—1787), et le cadastrage du pays (y compris les propriétés nobiliaires), à l'aide de ses officiers. Après sa mort la noblesse indignée annula le cadastrage dans la plupart des comitats et la vague d'indignation rendit pour longtemps impossible de procéder à n'importe quel relevé statistique.

Il y avait des domaines administratifs dans lesquelles le gouvernement royal avait plus de succès. Étant donné que le roi disposait des droits royaux (*regale*), il pouvait se procurer toute

information qu'il voulait. La Chambre royale présentait régulièrement des rapports sur les échanges commerciaux avec les pays étrangers, sur la production des métaux précieux etc.

Même les recherches de courte durée dans les archives hongroises montrent clairement que malgré les difficultés que l'administration royale avait dû surmonter et malgré le triage barbare par lesquels les relevés avaient été détruits au siècle passé, nous possédons pourtant des matières de grande valeur. Ces matières semblent bien se prêter à un travail par lequel nous puissions donner un aperçu solide, basé sur des statistiques, des conditions démographiques, économiques et sociales dans la Hongrie de la fin du XVIII<sup>e</sup> siècle. Autant qu'il est possible d'en juger à l'état actuel des recherches, nous possédons quelques échantillonnages statistiques embrassant tout le pays et aussi quelques-uns fragmentaires. Déjà ce fait motive le procédé d'échantillonnage qui nous semblera encore plus indiqué si nous prenons en considération le fait que la Hongrie de XVIII<sup>e</sup> siècle fut divisée à tous les points de vue. Il y avait des nobles et des serfs, en dehors de l'Église catholique il y avait trois Églises protestantes et une orthodoxe, la population appartenait à plus de six nationalités. Sous l'aspect économique les différences étaient encore plus grandes. Dans certaines régions la monoculture (surtout la viticulture) régnait, dans d'autres la pêche et l'élevage étaient les éléments de subsistance les plus importants. Il va sans dire que dans ces conditions la vie culturelle ne pouvait être homogène non plus.

L'auteur souligne qu'il connaissait bien les difficultés à cet égard. Il n'y a pas d'échantillonnage qui puisse servir de remède universelle, il faudrait faire des échantillonnages stratifiés dans tous les domaines (démographique, agraire etc.), c'est le seul moyen de trouver la solution et la seule possibilité de profiter des échantillonnages fragmentaires étant à notre disposition.

### 3.

## Imre Wellmann: Dépouillement des sources d'histoire agraire et l'ordinateur

Vu que, dans l'historiographie, c'est surtout l'histoire économique et sociale qui permet et même impose l'analyse quantitative, et que par conséquent, dans ces domaines-là — depuis assez longtemps — les chercheurs, au lieu de reconstruire le passé à l'aide de «l'historiographie impressionniste», s'efforcent à recourir aux données mesurables, l'histoire agraire ne peut plus exister sans dépouillement mécanique des données.

1. La nécessité du dépouillement automatique ne relève pas d'une mode scientifique, il ne s'agit pas non plus de simplement rattraper le développement technique. En Hongrie les sources historiques se prêtant au dépouillement automatique concernent surtout l'évolution agraire, car la Hongrie fut un pays agraire pendant des siècles. (La vie des villes était basée en grande partie aussi sur l'agriculture.) D'une part déjà la volume et la richesse de ces sources nécessitent l'utilisation de l'ordinateur (tant dans la documentation que dans l'analyse statistique), d'autre part l'histoire agraire — vu sa place spéciale occupée dans l'hierarchie des disciplines — a des rapports très étroits avec d'autres disciplines voisines (ethnologie, géographie, agronomie, économie politique, sociologie), c'est pourquoi il est inévitable de prendre en considération les aspects spéciaux de ces dernières. Tout cela exige l'analyse complexe des problèmes et ce n'est plus exécutable avec des moyens traditionnels.

2. Que peut apporter à l'historiographie agraire le dépouillement systématique automatisé des données?

Non seulement l'accélération des procès de travail exécutables à la main, mais d'une part l'agrandissement du domaine des recherches en l'étendant à des «taches blanches» et d'autre part l'approfondissement de l'analyse (l'analyse complexe de la structure intérieure et du dynamisme des phénomènes ainsi que l'éclaircissement de l'importance relative de différents facteurs et l'exploration des rapports complexes susmentionnés). Les résultats atteints jusqu'à présent dans les recherches prouvent qu'il s'agit d'un espoir fondé réellement. En plus, le dépouillement automatisé aura prêté aide à l'appréciation correcte et à la critique des sources historiques.

3. Les tâches de l'historien agraire dans l'utilisation efficace des ordinateurs dépassent la simple connaissance des méthodes mathématiques concernant le dépouillement automatique. Les autres tâches sont les suivantes:

a) Le choix des sources adéquates. La possibilité de l'analyse quantitative est donnée surtout dans le cas de sources historiques qui couvrent — sous des aspects homogènes — un territoire assez vaste (si possible, le pays entier) et qui permettent donc de traiter le profil détaillé, richement nuancé, valable à l'époque en question;

ou bien (c'est plus rare) si elles forment, classées également selon les principes homogènes, des cycles complets s'étendant aussi loin que possible dans le temps et peuvent ainsi servir de base pour dresser des trends d'évolution. Plus on recule dans le temps, moins ces conditions sont données.

Tout cela indique qu'il nous faut commencer le travail dans un domaine et à une époque où nous disposons de données complètes et justifiées. Concernant la Hongrie historique, cette époque est celle de la fin du XIX<sup>e</sup> siècle, quand nous possédons déjà réellement des relevés homogènes, ayant un caractère statistique. Mais aussi faut-il attirer l'attention sur le fait que dans le développement agricole les différences dans les conditions naturelles et dans celles du marché etc. firent surgir bien des variétés régionales, voire locales, et que cette diversité — vu la prédominance du facteur de la nature — entra en jeu dans une mesure croissante à l'époque préindustrielle. Si l'on veut avoir une vue d'ensemble des facteurs du développement agricole en Hongrie il est recommandé d'élaborer un profil de la fin du XIX<sup>e</sup> siècle. C'est une base solide de comparaison qui nous permet de remonter avec plus d'assurance dans le temps vers les époques moins documentées et sur cette base il nous sera possible d'élaborer des trends plus longs de l'évolution.

*b)* Critique des sources

L'historien doit examiner soigneusement dans quelle mesure l'ensemble de la source ou ses séries de données sont crédibles, dans quelle mesure ses lacunes peuvent être complétées par des données prises à d'autres sources et sous quels rapports elles sont propres au dépouillement automatique.

*c)* Préparation du programme

Avant l'analyse quantitative il faut procéder à une analyse qualitative, dresser l'hypothèse de recherche, établir le modèle — qui permettra le choix des indicateurs les plus adéquats. A tout cela succède la présentation de la documentation sous une forme se prêtant au codage et c'est le point à partir duquel la connaissance des méthodes mathématiques est indispensable.

4. Voie de l'évolution

Nous avons encore beaucoup de problèmes à résoudre et nous ne pouvons marquer un progrès appréciable que de pas à pas, selon un plan, unissant les forces systématiquement. La condition primordiale du succès est avant tout la collaboration des experts des ordinateurs et des spécialistes de l'histoire agricole. Jusqu'à présent cette collaboration ne se manifesta que dans des tentatives individuelles. Pour pouvoir réaliser la collaboration, nous avons besoin d'une part d'experts des ordinateurs connaissant bien les problèmes historiques (à l'échelle internationale et nationale il y en a des cas), d'autre part d'historiens se familiarisant avec les méthodes mathématiques. Ainsi il sera possible non seulement de créer une base de sources large et sûre mais aussi de donner un élan sans précédent aux travaux de recherches et dépouillement — en s'appuyant sur les expériences acquises à l'étranger et dans notre pays.

4.

### László Katus: Sur l'utilisation des ordinateurs dans les recherches sur l'époque capitaliste

1. Ce qui est déjà accompli et ce qui se fait actuellement.

A la connaissance de l'auteur, deux recherches furent effectuées ou sont en train en Hongrie — concernant cette période — à l'aide d'ordinateurs:

*a)* le tableau input — output de l'industrie manufacturière en Hongrie dressé sur la base des statistiques industrielles de l'année 1898;

*b)* l'analyse de la production des cultures agricoles et une comparaison internationale entre 1867—1914 — études de Mmes György Simonics et Julia Puskás. Aux États-Unis c'est Eddie M. Scott qui fit des recherches à l'aide d'ordinateurs sur la base des données du commerce extérieur hongrois des années 1882—1913.

2. Possibilités et besoins futurs en Hongrie

L'époque entre 1850—1945 offre plusieurs possibilités pour les recherches poursuivies à l'aide d'ordinateurs. Dans le cas de beaucoup de thèmes les recherches effectuées par la méthode traditionnelle sont arrivées à un point où leur développement n'est plus possible qu'avec des moyens automatiques: d'une part nous avons beaucoup de sources ou des types de sources qui demandent le dépouillement automatique, d'autre part au cours des recherches des années dernières, plusieurs problèmes se posèrent auxquels nous ne pouvons répondre d'une façon satisfaisante qu'à l'aide d'analyse statistique effectuée par des ordinateurs.

*a)* Des sources, des types de sources qui renferment des données en nombre limité, du même type, faciles à standardiser et concernant plusieurs milliers ou dizaines de milliers de cas individuels. Ce sont d'une part des sources imprimées, d'autre part des sources d'archives: les livres d'adresses (ceux des cultivateurs sont les plus importants), des almanachs, des listes de firmes, des compas, des rôles des contribuables, des procès-verbaux des conseils de revision, des registres d'état civil etc. Les documents volumineux, comme rapports policiers sur le mouvement

ouvrier, journaux du Parlement ainsi que les manuels généalogiques et les encyclopédies biographiques peuvent être y classés aussi. En plus, il est nécessaire de procéder au dépouillement des données des éditions statistiques officielles de l'époque conformément aux nouveaux principes de classement, et aux critères nouveaux introduits par les recherches modernes, ainsi que les soumettre à une analyse statistique d'un niveau plus élevé.

b) Dans tous les domaines des recherches historiques nous pourrions citer des problèmes qui exigent le dépouillement automatique, mais nous ne nous occupons ici que de ceux qui furent déjà mis à l'ordre du jour effectivement et dont le dépouillement fut déjà essayé par des méthodes traditionnelles.

#### *Histoire économique*

Les indices agrégats les plus importants de la croissance de l'économie nationale (les composants du revenu national et ceux du produit national ainsi que les facteurs ayant une influence sur leur évolution), les travaux écrits sur l'histoire des salaires et des prix.

#### *Histoire sociale*

L'analyse des structures sur le plan national ou local, concernant certaines couches et classes ou bien l'ensemble de la société; biographie collective (reconstruction d'une couche ou d'un groupe sur la base d'un grand nombre des cas individuels), l'analyse des mouvements de masse (p. ex. des grèves), les recherches démographiques (l'élaboration des indices les plus importants du mouvement démographique à l'échelle nationale.)

#### *Histoire politique*

L'analyse du comportement des parlementaires et des électeurs et celle des facteurs qui les influencent.

#### *Histoire de la civilisation*

L'établissement de statistiques du système d'enseignement à l'échelle nationale pour une longue période, l'analyse statistique de l'édition, l'analyse de la formation des courants idéologiques influençant l'opinion publique — fondée sur les données offertes par les mass-media (presse p. ex.).

### 3. Caractère des recherches

Comme premier abord, il peut être question d'un travail de deux types:

a) le dépouillement descriptif in extenso d'un groupe de documentation de même type ou d'une source; le codage des données, leur enregistrement sur bande magnétique ou carte perforée, en procédant au travail fondamental de classement, de groupement et du dépouillement statistique de base;

dans le cas de sources de tel type il est recommandé de fonder une «archive d'ordinateur central» (banque de données) dans laquelle les données fondamentales enregistrées sur bandes magnétiques pourraient être dépouillées par les chercheurs au point de vue statistique conformément à leurs propres problèmes.

b) L'exploration des rapports et des phases restés cachés pour le travail traditionnel, à l'aide d'une analyse statistique d'un niveau plus élevé, pour pouvoir répondre à certains problèmes,

dans ce cas-là différentes sources concernant le problème en question seraient analysées à l'aide d'un échantillonnage.

4. Prenant en considération les tâches des recherches poursuivies à l'aide d'ordinateurs, l'auteur désire signaler — le résumé susmentionné constituant un programme de recherches de plusieurs décennies, même en cas d'approvisionnement optimal en machines — quelle est la voie qui s'ouvre devant les recherches par l'introduction des ordinateurs. Il souligna que la solution des problèmes de ce genre est connue dans la littérature internationale et que les expériences internationales montrent que le surcroît de frais dû à l'utilisation des ordinateurs est compensé par les résultats. Parmi ses propositions l'auteur estime de toute première importance celle de la fondation d'une archive centrale d'ordinateur pour le dépouillement de sources ou de groupes de sources déterminés.

### 5.

## **Julianna Puskás—Margit Lánç: Analyse quantitative de la production agricole des labours en Hongrie de 1867 à 1913**

L'étude s'attachant au rapport de L. Katus avait pour but l'analyse statistique-mathématique de la production agricole des labours pour en tirer ensuite les conclusions économico-historiques. Les auteurs donnèrent surtout un renseignement méthodologique dans ce qui suit.

Les sources de base: les séries des quantités exportées, des prix de production, et celles de la récolte annuelle de chaque culture. Les auteurs effectuèrent les calculs suivants fondés sur ces séries:

1. L'analyse des séries fut commencée par la méthode la plus simple et la plus générale: celle en dressant des diagrammes et des chiffres corrélatifs. Des séries d'indicateurs furent formées pour présenter les tendances de développement, notamment des indices de base, toujours sur la base de la moyenne des années 1909—1913 et des indices de chaîne. La plupart des séries d'indicateurs sont les indices de volume. Pour pouvoir réduire au même dénominateur les quantités impropres à l'addition et les différentes qualités, les auteurs introduisirent l'indice de valeur. Elles suivirent le changement dans la production de certains groupes de cultures et dans la production agricole totale en utilisant des prix invariables. Les séries d'indicateurs formées sur la base du prix de production pondéré de la période 1909—1913 et de prix invariables, fonctionnent comme moyens de la réduction au même dénominateur des valeurs d'usage différentes et comme expression effective de la valeur, et elles représentent le volume de la production.

2. Le changement de longue durée des séries et la tendance fondamentale du développement de la production furent explorés par des calculs de trends. Partant des calculs les plus simples, les auteurs calculèrent d'abord les trends par des moyennes mobiles. Le rajustement ainsi atteint des flottements des séries facilita déjà la connaissance de la nature de ces séries, et la courbe de trends s'adapta relativement bien et signala les phases plus grandes de l'accroissement de la production qui étaient mentionnées dans les sources historiques avec texte. Ce système de calcul n'a pas pourtant permis d'obtenir une courbe facile à manier d'une façon analytique et concernant les séries entières.

C'est pourquoi les auteurs recoururent aux calculs de trends analytiques, après avoir établi les séries des valeurs de trends calculées par le système des moyennes mobiles et après en avoir tracé les diagrammes. Il ne fut point facile de déterminer, quels sont les types de trends le mieux adaptés à certaines séries chronologiques. Les auteurs se servirent de plusieurs méthodes (elles étudièrent les diagrammes des séries chronologiques, marquèrent au jugé la médiane, examinèrent les graphiques des trends (de calcul) à moyennes mobiles, prirent en considération les méthodes de calcul de trends appliquées dans le cas des séries de type pareil) et prirent en considération aussi les informations offertes par les sources historiques. Tirant toutes ces conclusions, les auteurs choisirent les types de fonctions mathématiques qui pouvaient refléter d'une manière juste la tendance continue des séries. (Les auteurs élaborèrent une méthode de contrôle: celle des moindres carrés. Aux séries chronologiques, englobant la formation de la production, s'adaptèrent dans la plupart des cas des trends linéaires, en quelques cas des trends exponentiels, mais quant aux prix, il fut toujours nécessaire d'appliquer des paraboles.) Après avoir établi les séries de calculs de trends, elles tracèrent les diagrammes de trends (à une graduation ini-logarithmique) et tout cela eut pour résultat une courbe divergente des trends linéaires. Cette solution graphique montre bien non seulement la médiane, mais aussi les divergences entre la série originale et celle calculée d'après des valeurs de trends.

3. Pour démontrer dans quelle mesure il fallait égaliser les séries originelles et dans quelle mesure les données réelles des séries divergent des valeurs de trends (autrement dit: des médianes), les auteurs calculèrent les déviations des données de la série originelle partant des données de la série des médianes. La différence d'année en année est démontrée par la série des déviations. (Les déviations des trends, pour pouvoir être résumées et comparées sont exprimées en pourcentage.)

Les auteurs attribuèrent une grande importance à l'exploration des déviations et, en partant de celles-ci, au calcul des indices de dispersion, elles furent conscientes en effet de l'abstraction mathématique obtenue par le calcul des médianes. (Elles ont remarqué que c'est cette abstraction qui est une des raisons pour lesquelles les historiens économistes sont réticents aux calculs de trends.) Selon les auteurs on a besoin de cette abstraction mathématique précisément pour pouvoir connaître plus profondément et plus exactement la réalité historique. Pour les historiens le but primordial des calculs des médianes consiste précisément à pouvoir démontrer et analyser de point de vue historique les diversions et les dispersions. (Les méthodes statistiques modernes prouvent que les multitudes statistiques se présentent d'une façon différente si l'on connaît la médiane de leurs diversions et de leurs dispersions.) Par les calculs visant ces dernières, on a la possibilité de reconnaître ou de présenter les caractéristiques de la production lesquelles seraient cachées sans cette analyse statistique mathématique.

4. Puis les auteurs ont calculé les indices de dispersions suivants: la moyenne absolue des déviations, la moyenne quadratique des déviations et la déviation dite standard. Cette dernière est l'indice le plus pur, car il élimine l'influence des différences quantitatives des séries.

5. La série des déviations signala non seulement le flottement de la série originelle autour de la médiane, mais démontra aussi la périodicité des différences. Pour la saisir, les auteurs calculèrent des trends par la méthode des moyennes mobiles et représentèrent graphiquement cette ligne de trend — considérant la médiane comme 100 pourcent (avec une moyenne mobile de 7 ans comme pour les séries chronologiques). Par cette méthode elles pouvaient approcher plus facilement la nature des séries, car à l'intérieur de la ligne de trends à long terme se manifestèrent les périodes moyennes et les périodes de rythme accélérés et retardés.

6. Les séries calculées contiennent les déviations secondaires. Pour démontrer le mouvement cyclique et les flottements, les auteurs examinèrent la périodicité annuelle — considérant la médiane comme 100 pourcent.

Les séries et les diagrammes calculées par des méthodes statistiques, mathématiques servirent de base solide pour l'examen du dynamisme du progrès. Les indices de ce dernier sont *des rythmes de croissance moyens*. Les auteurs calculèrent ces indices concernant les séries entières de deux façons: 1. à partir des médianes du rythme de croissance moyens, 2. à partir de la moyenne des premières et dernières cinq années des séries. Le rythme de croissance moyens des séries entières est également une abstraction mathématique. Dans la réalité, les rythmes de croissance — comme le signalèrent déjà les indices de chaîne — ne furent pas les mêmes pendant deux ans non plus. La médiane, c'est à dire le rythme de croissance moyen des séries entières est important pour les historiens afin qu'ils puissent connaître les déviations et puissent analyser les dispersions autour des rythmes de croissance moyens.

D'après l'analyse des séries des déviations et surtout de leurs lignes de trends il est devenu clair qu'au delà du changement annuel des rythmes de croissance certaines périodicités se manifestaient, à savoir les périodes accélérées ou retardées des rythmes de croissance. Les auteurs s'efforcèrent de démontrer ces périodes par deux méthodes: 1. par le calcul de rythmes de croissance moyens superposés pour des périodes de 10 années, 2. pour les périodes que les déviations signalèrent comme changements du rythme de croissance. Les auteurs se servirent des méthodes mathématiques pour l'analyse du changement structural de la production des cultures de champs. Elles cherchèrent — par une méthode de filtrage — la réponse à la question de savoir dans quelle mesure les trois éléments essentiels influencèrent la croissance. Ces trois éléments sont: l'aggrandissement de l'emblavement, la transformation structurale des semailles et l'accroissement du rendement des récoltes. (Les auteurs analysèrent le rendement par cultures aussi.)

7. Pour l'analyse comparative internationale et pour l'appréciation plus exacte du développement hongrois, les auteurs soumièrent dans la mesure possible à un examen analogue les séries de la production des cultures de champs dans plusieurs pays européens. En dehors de la comparaison de la production en Hongrie et en Autriche, l'examen concerté des deux séries permit d'établir la formation des trends à long terme concernant les principaux produits agricoles et d'examiner la croissance au niveau de la Monarchie Austro-Hongroise aussi.

### Attila Csernok—Éva Erlich—György Szilágyi: Sur l'infrastructure sur la base des statistiques historiques (Examen comparatif international 1860—1968)

L'examen exécuté par le collectif d'économistes illustra à la conférence les problèmes que pose le rapprochement des recherches en économie et en histoire économique. Les recherches se poursuivirent surtout dans les domaines suivants:

1. Comment évolua l'infrastructure dans les différentes phases de la croissance économique des différents pays? 2. Quel est le degré de développement — prenant en considération le niveau de la technique de l'époque — appartenant au niveau économique du pays donné; 3. «Combien coûtent» — à des niveaux définis du développement économique — l'établissement, l'entretien, la modernisation et l'extension de l'infrastructure, combien ils exigent des économies nationales?

L'examen international historique de l'infrastructure embrassa une période de 100 ans (1860—1968).

Sur la notion de l'infrastructure: Les auteurs entendent par cette notion la partie de la fortune nationale qui ne sert directement ni à la création ni à la consommation des biens matériels, la partie qui doit assurer la marge pure du processus de «production — distribution — consommation», conformément aux exigences de la technique, à un niveau donné du développement économique.

Un système d'indices spécial sert à la comparaison internationale historique de l'infrastructure sur le plan international.

Les types du développement infrastructural et de l'industrialisation:

Trois types fondamentaux se dessinèrent en résultat des calculs: Le premier est incarné par l'Angleterre, la Hollande et quelques pays de l'Europe occidentale, dans ces pays la formation de l'infrastructure précède le développement industriel accéléré. Les cadres infrastructuraux créés antérieurement garantirent la marge primitive pour le fonctionnement et l'extension graduelle de l'industrialisation commençant plus tard. L'autre type du développement, plus moderne, est incarné par l'économie des États-Unis. L'infrastructure s'est formée ici presque

sans antécédents, parallèlement à la formation du capitalisme et à son extension, en même temps que se présentèrent les besoins provoqués par la production et la consommation. Au début, le capital étranger appuya aussi ce procès. Le modèle anglais est caractérisé par l'antériorité de l'infrastructure, celui des États-Unis par le développement infrastructurel accompagnant l'extension du capitalisme. Plusieurs pays de l'Europe centrale et orientale restèrent à l'écart — surtout après la I<sup>re</sup> guerre mondiale — de la transformation infrastructurelle qui s'effectua à la même époque en Europe occidentale. Le type de développement économique réalisé par les pays socialistes européennes après la II<sup>e</sup> guerre mondiale diffère des deux autres types susmentionnés. Ce type n'est caractérisé ni par l'antériorité ni par le codéveloppement. C'est un troisième type où le développement infrastructurel reste absent, du moins pour un certain temps. Dans ces pays-là — jusqu'au milieu des années soixante — il s'agit plutôt de l'usure de l'infrastructure existante, périmée.

Deux recherches spéciales avec ordinateur furent exposées suivant le rapport de I. Wellmann concernant l'histoire agraire. Toutes les deux préparent des méthodes applicables généralement dans l'exploration des sources et analyse des données de l'histoire agraire. Elles se poursuivent avec la collaboration des collectifs, composés de divers spécialistes, conformément au but.

En ce qui concerne le rôle que jouent les éléments de production dans l'infrastructure, on peut constater que l'évolution du rapport entre le nombre de ceux travaillant dans l'infrastructure et la population active montre la fusion particulière des facteurs du temps et du développement économique. Entre les deux facteurs c'est le deuxième qui est décisif.

Une des caractéristiques les plus importantes de l'infrastructure est le fait que cette dernière exige de grands investissements. La proportion des investissements infrastructureux comparée à la proportion des investissements totaux montre de larges variations, en fonction du temps et du pays. Cette proportion ne baisse presque jamais sous 50 pour-cent et sa limite supérieure s'élève à 80 pour-cent. L'autre caractéristique des investissements de l'infrastructure est le fait que leur volume ne dépend pas en général du niveau de développement économique. Dans les pays socialistes la proportion des investissements infrastructureux comparée aux investissements totaux est moins grande que dans les pays capitalistes.

Le développement de l'infrastructure est un procès long et exige de grands efforts des économies nationales.

7—8.

### Dénes Varga—Éva Veress: L'exploration à l'aide d'ordinateurs des recensements des terres de l'époque féodale

A l'aide de linguistes-mathématiciens nous préparons des index de sources sur plusieurs dizaines de milliers de recensements des terres remontant aux années 1526—1848. En systématisant à l'aide d'ordinateurs les 15 000 extraits de documents, nous exécutons d'abord le classement élémentaire qui peut faciliter l'aperçu général. (Au moment de la conférence c'est ce travail qui se prépara.) Les linguistes et les mathématiciens participant à ce travail élaboreront une méthode qui pourrait être développée ultérieurement jusqu'au dépouillement thématique détaillé qui peut s'avérer nécessaire dans le cas de certaines parties de la matière choisie à l'aide de l'ordinateur.

Antérieurement on a fait des expériences concernant l'utilisation des moyens de documentation manuels (système de catalogue traditionnel, cartes perforées). Tout cela n'a pas eu de résultat satisfaisant, les difficultés de leur préparation, entretien et utilisation étaient évidentes, vu la quantité des matières. Les auteurs estiment quand-même opportune de préparer le dépouillement automatique par des moyens de documentation manuels; possédant les informations antérieurement obtenues sur la matière, on peut organiser plus économiquement le dépouillement automatique, on peut se procurer des expériences utiles et en plus, la plupart des problèmes de classification et de définition se posent dès ce travail et le système de documentation adapté à la matière arrive à l'ordinateur sous forme mûre. (C'est la première partie de l'exposé qui traite le problème de la combinaison avantageuse de la mise au point manuelle et automatique.)

Dans le cas des indices des registres des propriétés terriennes il faut que le travail diffère d'une façon fondamentale des autres systèmes de documentation spécial (p. ex. la documentation des bibliothèques.) Avant tout, il fallait garantir la prépondérance des aspects de critique de sources (pour cela servaient à base les données critiques des résumés de source, c'est d'après ces données qu'on peut former les qualifications concernant certaines données et les documents.)

Sous l'aspect historique, la partie la plus importante du registre en préparation est l'indice géographique et du point de vue de l'utilisation des ordinateurs — c'est le travail le plus intéressant. À cause de l'orthographe variée des noms de lieu, à cause du changement fré-



quent des noms des habitants et de la réorganisation successive des unités administratives, le dressement de l'indice géographique semble être irréal pour plus que quatre siècles à la base des données de plusieurs cent mille datations de noms de lieu. Lors du dépouillement automatique, ce n'est pas l'accord littéraire qui est le devoir primordial parce qu'à cette méthode on ne peut classer qu'une partie de la matière. Il faut résoudre le problème; comment on peut reconnaître le même nom de lieu d'après les variantes, c'est à dire, comment on peut définir d'après les noms de lieu ayant des formes uniques, quel est l'objet physique concerné dans un contexte donné. Dans une certaine mesure ce problème est l'analogon linguistique de «pattern recognition» dans le cas où l'étalon (sous forme de registre p. ex.) est donné d'avance.

Les troisième et quatrième parties de l'exposé passent en revue les résultats obtenus grâce à l'examen empirique de la toponymie et à la méthode élaborée pour aborder le problème à l'aide de l'ordinateur.

### László Cseh—György Haraszti—Éva Veress: Méthode topographique du dépouillement et du stockage des données dans l'histoire agraire (établissement de diagrammes géographiques à l'aide de l'ordinateur)

L'étude expose deux sortes de diagrammes géographiques dressés à l'aide de printer:

1. Les diagrammes de points: des unités géographiques (p. ex. une agglomération) se présentent ici sous forme de points, le point est représenté par un signe de ponctuation choisi du jeu de signe du printer — conformément au thème. Sur une carte thématique une série définie de signes de ponctuation répond à la série des catégories qualificatives. (P. ex. la série des chiffres peut signifier des catégories d'échelle.) Ces diagrammes ne servent pas à la netteté visuelle, le facteur visuel joue un certain rôle lorsqu'il faut choisir des signes de ponctuation qui puissent être distingués nettement, leur but est d'enregistrer dans une disposition géographique les données et les résultats obtenus lors du dépouillement. Ces diagrammes sont des cartes de travail que le chercheur peut dépouiller plus facilement que les tableaux récapitulatifs. Une réduction effectuée selon certaines considérations fixées à l'avance y joue son rôle. Ils peuvent être exécutés à bas prix même dans une quantité importante et ils sont facilement maniables de tous les points de vue.

2. Des diagrammes de cartes pouvant être mesurés visuellement aussi, se préparent avec une méthode pareille. Des unités territoriales à l'intérieur du pays (p. ex. les comitats) forment une surface couverte des signes de ponctuation conformes à leur qualification. Par le choix des signes de ponctuation on peut atteindre les contrastes visuels qui sont atteints lors du tracé des cartes par les hachures de différente densité. La carte imprimée de cette façon offre sur le résultat définitif du dépouillement une vue d'ensemble comprenant l'interprétation aussi.

Ces diagrammes ressemblant aux cartes sont très proches de la cartographie thématique, mais ils ne doivent pas être confondus avec la cartographie automatisée. Lors de leur dressement on prend en considération les aspects cartographiques dans la mesure où la technique le permet, mais en dernière analyse, ils sont de simples instruments de travail et en cette qualité ils représentent une phase du processus de dépouillement. Une caractéristique importante de la méthode utilisée est le fait que les diagrammes de carte ont des rapports très étroits avec l'ensemble du processus du dépouillement. Le dépouillement des données portées à l'ordinateur peut être achevé sans interruption et les résultats peuvent être représentés soit sur les diagrammes de cartes soit sur des tableaux et peuvent être transmis au listing.

La prépondérance de l'aspect géographique dans le procès de dépouillement des données permet et facilita aussi de résoudre le problème du stockage selon un principe géographique homogène des données provenant de sources volumineuses. Il devient possible d'élaborer un système de dépouillement approprié aux examens régionaux, de façon à ce que celui-ci soit facilement applicable aussi à des sources de composition et d'âge différents. La mise au point de ce système exige un travail préparatoire considérable mais le résultat est aussi important. Le travail préparatoire est commencé sur la base des expériences faites avec des diagrammes de carte susmentionnés.

## DISCUSSION

Dans la deuxième partie des délibérations, au cours de la discussion, les participants — tout en posant quelques problèmes de détail non négligeables et concrets — parvinrent quant au fond à la même opinion. Ils tirent tous à l'élargissement méthodique des recherches quantitatives, à leur approfondissement thématique et à l'organisation nécessaire du travail. Ils sou-

lignèrent l'importance de la préparation des examens quantitatifs et la nécessité de les relier aux recherches complexes. Les représentants des disciplines voisines partagèrent la même opinion (surtout des linguistes) tout en offrant leur appui.

Dans d'autres domaines historiographiques de l'utilisation des ordinateurs (p. ex. analyse des textes, analyse du contenu etc.) il n'eut pas ou presque pas d'initiatives jusqu'à présent en Hongrie. Les chercheurs tiennent compte de ces possibilités.

Il est recommandé de continuer la collaboration entre les historiens et les experts des ordinateurs — elle peut être très fructueuse dans le domaine de la formation des experts. Les participants — tirant les conclusions de la conférence — firent des propositions concrètes, pratiques, aux associations scientifiques, organisatrices de la présente conférence, concernant la collaboration interprofessionnelle.