

HÍREK ITTHONRÓL

Az atomoktól a csillagokig

Előadássorozat az ELTE Természettudományi Kar Fizikai Intézetében főleg – de nemcsak – középiskolásoknak

Tudod-e, hogyan működik a mobiltelefon és a mikrohullámú sütő, hogy keletkezik a szivárvány, milyen információkat lehet leolvasni az úrfelvételekről, hogyan lehet kincset keresni a GPS-szel és exobolygókat a Magyar Automata Teleszkóppal? • *Érdekel-e*, hogy milyen fizikai háttere van az internetnek, a sejtek önszerveződésének, a plazmatévének és az alternatív energiaforrásoknak? *Látod-e*, hogy mindennapi életünkben is mindenütt jelen van a fizika, a bennünket körülvevő modern technikai eszközök működése pedig egyenesen elképzelhetetlen nélküle? • *Akarsz-e* érdekes, meglepő, sőt meghökkentő fizikai jelenségeket, kísérleteket látni, esetleg saját kezded elvégezni a kísérleteket? • *Szeretnél-e* többet tudni az új felsőoktatási rendszerről, a három egymásra épülő cikluson alapuló egyetemi képzésről?

Ha érdekelnek e kérdések és a tudomány válaszai, gyere el a 2006 szeptemberében kezdődő előadássorozatra! Az ELTE TTK Fizikai Intézetében tevékenykedő kutatók és oktatók beszélnek a fizika frontvonalába tartozó érdekességekről, újdonságokról.

Az első előadás: 2006. szeptember 28-án (csütörtökön), 17.00-kor lesz az ELTE TTK légymányosi északi tömbjében (1117 Budapest, Pázmány Péter sétány 1/A), az Eötvös teremben (földszint 0.83 terem). Az előadások ingyenesek. A részletes program megtalálható a <http://www.atomcsill.elte.hu> honlapon, ahonnan a tavalyi előadások hang- és képanyaga is letölthető.

Az előadás-sorozat programtervezete:

2006. szeptember 28. – *Tichy Géza*: Mikrohullám árnyékolása, mobiltelefon – bevezetőt mond *Lendvai János*, a Fizikai Intézet igazgatója

2006. október 12. – *Vattay Gábor*: Az internet fizikája

2006. október 26. – *Cserti József*: A szivárvány fizikája

2006. november 9. – *Pál András*: Tű a szénakazalban – Bolygókeresés a Hungarian Automated Telescope projekt keretein belül

2006. november 23. – *Rajkovits Zsuzsa*: Buborékok, képek, káprázatok – a szappanhártyák színes világa

2006. december 7. – *Honyek Gyula*: A plazmatévétől a SIM-kártyáig ... Válaszok olyan kérdésekre, amelyeket csak kevesen tesznek fel

2006. december 21. – *Jubász András*: A mindennapok fizikája

2007. január 11. – *Czirók András*: Sejtek önszerveződésének fizikája

2007. január 25. – *Derényi Imre*: Sejtstruktúrák fizikája

2007. február 8. – *Horváth Ákos*: Mit ad nekünk a Nap: alternatív energiaforrások fizikai háttere

2007. február 22. *Kovács Béla*: Kincskeresés GPS-szel: a korszerű navigáció alapjai

2007. március 8. – *Tímár Gábor*, *Kern Anikó*: Otthonunk az űrből – felvételek az ELTE műholdvevő állomásáról
Szeretettel várunk Téged, barátaidat, tanáraidat és szüleidet sorozatunkon!

ELTE TTK Fizikai Intézet

MINDENTUDÁS AZ ISKOLÁBAN

eSCIENCE

Adatözönben élünk. Műholdak térképezik fel Földünk felszínét és időjárását, távcsövek pásztázzák automatikusan a Világegyetem egyre távolabbi régióit, elektronikus berendezések rögzítik a nagyenergiás részecskeütkezésekből kijövő számtalan részecske összes adatát. Személyes adatainkat hitelfelvételi, vásárlási és utazási szokásainkról magán- és közintézmények gyűjtik, a humán genom projekt keretében pedig elkészült az első emberi géntérkép. Az interneten felgyülemlő adatmennyiség ma már csak *exabyte* nagyságrendben mérhető, és ez is kétévente duplázódik. Az összegyűjtött adatok kiértékelése általában nagyobb gondot okoz, mint maga az összegyűjtésük, mert sokszor nehéz átlátni az adatok közötti összefüggéseket, és kiszűrni belőlük a lényeges információkat: előre jelezni esetleges katasztrófális jelenségeket a természetben vagy a hálózata-

tokban, felismerni a terroristagyanús személyeket, megtalálni az öregedésért és örökletes betegségekért felelős géneket, azonosítani a potenciálisan piacképes termékeket, megbecsülni egy-egy döntés kockázatát.

Felvetődik a kérdés, hogy miként lehetne a – fentiekben példaként kiragadott esetekben meglevő – hatalmas adatözönt kezelni. A választ az *eScience* jelentheti.

Mi az eScience?

Az *eScience* kifejezést egy új, „feltörekvő” technológiára használják, melynek révén nagyléptékű, komplex tudományos tevékenység fejthető ki a modern információs technológia felhasználásával. Legfőbb jellemzője a rend-

kívül sok, gyakran különböző helyekről elérhető adattal operáló kiértékelő munka, melynek eredményes véghezviteléhez az adatok automatikus gyűjtésére, optimális adatbázisba rendezésére, rendkívül nagy számítástechnikai kapacitást igénylő feldolgozására, és a lényegret megragadó vizualizációra van szükség.

Az eScience az alkalmazott tudomány, az alap kutatás, valamint a modern információs technológia (IT) interdiszciplináris egymásra hatásából született, és egy új minőséget képvisel a kiértékelésben. Az IT kifejleszti a korszerű eszközöket és informatikai rendszereket az alap- és alkalmazott kutatások mai problémáinak megoldásához. Az így keletkezett rendszerek, illetve az általuk szolgáltatott adatok, komplexitásuk miatt, már csak az eScience segítségével vizsgálhatók. Az eScience – jelenlegi fejlettségi szintjén – elsősorban a szervezett adatgyűjtés anyagának feltárására irányul, ugyanakkor a spontán felhalmozódó adatmennyiség feldolgozásához is előkészíti a szükséges technológiát.

Miért fontos az eScience?

Az elmúlt évtized jelentősen megváltoztatta szinte minden tudomány módszertanát. Ahhoz hasonlóan, ahogy a matematika egy évszázaddal ezelőtt szétáradt szinte minden tudományban, és az elméleti modellek alapeszközzé vált, ma már szinte lehetetlen bármelyik tudományágot a modern információs technológiák használata nélkül művelni. Ezzel párhuzamosan egy másik irányú folyamat is megindult: a modern információs rendszerek annyira összetetté és bonyolulttá váltak, hogy vizsgálatukhoz, fejlesztésükhöz, valamint az általuk szolgáltatott adatok kiértékeléséhez az informatika eddigi módszerei már nem elégségesek. Szükség van a hagyományos, de bonyolult természeti jelenségek vizsgálatára kidolgozott módszerek eszköztárára is.

Sem a tudomány, sem az elektronika történetében nem ez az első forradalom. Az olcsó személyi számítógépek megjelenése megváltoztatta az informatika viszonyát a tudományhoz, az üzleti és a mindennapi élethez. Az olcsó internet forradalmasította a munkacsoportok működését, az itt közölt eredmények mindenké számára gyorsan elérhetővé váltak. Ma pedig a szemünk előtt zajlik az adattechnológia forradalma: rohamosan nő a műszerek, érzékelők, felvevő berendezések sávszélessége és mennyisége, valamint az internet-felhasználók száma, akik hála a ma már könnyen kezelhető webtechnológiának, szinte ontják a különböző értékű adatokat. Mindez oda vezetett, hogy napjainkra mind az igények, mind a már létező kapacitások kinőtték az informatika által biztosított eddigi kereteket: a már meglévő adatbázisok sokszor szétszórtan, „ömlesztve” tárolják az adatokat, az új adatgyűjtési eljárásoknak egyre nagyobb mennyiségű adatot kell automatikusan begyűjteniük. A meglévő adatokat egyre nagyobb számítástechnikai teljesítménnyel lehetne csak feldolgozni, és az adatbázist használók sokszor sötétben tapogatóznak, mert nem mindig tudják, hogy pontosan milyen adatokra van szükségük, illetve, hogy mi mindenre lehetne használni az összegyűjtött információt.

Az eScience technológiája, bár a tudomány eredményeiből gyökerezik, a mindennapi élet számos területén is alkalmazható. Mindenütt alapkövetelmény a jó trendek felismerése és kiaknázása. A megfelelő irányban tett gyors elmozdulást általában siker kíséri: példa erre a Google-portál, vagy az Amazon e-kereskedelem több – az összegyűjtött adatok alapján megtervezett és jól célzott – reklámakciója. Ugyanakkor az e-buborék kipukkanása az ezredforduló körül jelezte, hogy számos technológia még nem érett be.

Mit gondolnak mások?

Az eScience névvel jellemzett új technológia kifejlesztésének és alkalmazásának szükségessége már felvetődött az USA Nemzeti Tudományos Alapjának, a kanadai Innovációs Alapnak, valamint az angol Nemzeti eScience Központ bizottságainak ülésein, ahol megállapították, hogy az adatgyűjtés, tárolás és karbantartás kezelésére új megközelítést kell kidolgozni, mivel az eddigi eljárások rohamosan bonyolódhatnak, és egyre több pénzt fognak felémészteni. Az IBM kivonul a hardver-üzletágból, helyette adatbázisrendszerekre és a bioinformatikára helyezi a hangsúlyt, a Microsoft Research-nél hasonló irányvonal figyelhető meg. A *New Scientist* 2005. áprilisi száma a rákkutatás jelenlegi helyzetéről közöl áttekintést, melyből kitűnik, hogy ezen a területen szűk keresztmetszetet jelent olyan szakemberek hiánya, akik egyrészt birtokában vannak bizonyos biotechnológiai ismereteknek, ugyanakkor szakszerűen képesek kezelni az eScience eszköztárát.

Mi a jövő?

A számítástechnika és a kutatás-fejlesztés küszöbön álló forradalma a következő technológiákra fog épülni: az egyre nagyobb adatözmön olcsó begyűjtésére, strukturált tárolására, a különböző helyeken elszórt kapacitásokhoz való egyszerű hozzáférésre, valamint az adatok és keresések megjelenítésére. Mivel hatalmas adattömegek a tudósok műszereinél jelentek meg először, a kutatók már sok tapasztalatot szereztek ezeken a területeken, ebből adódik, hogy a tudományinformatika az eScience élén járhat olyan technológiák előállításában – illetve hazai átvételükben és továbbfejlesztésükben –, amelyek az élet egyéb területein is kamatoztathatóak.

Az adatbázisok összekapcsolásával létrejövő struktúra (ezt a csillagászatban Virtuális Observatóriumnak nevezik) egy „virtuális szervezetet” definiál a hálózaton, melynek tagjai távolsági korlát nélkül képesek kihasználni annak erőforrásait. Az ilyen jellegű szervezetségre a közeljövőben a tudomány mellett elengedhetetlen szüksége lesz a nagyvállalati munkának, az államigazgatásnak, környezetgazdálkodásnak, a nemzetbiztonságnak, de hamarosan még a kulturális szférának is. Egy ilyen szervezet ugyanakkor jelentősen megkönnyítheti a kis- és közép vállalkozások számára is az adatokhoz, valamint az olcsó feldolgozó-kapacitáshoz való hozzájutást, ezáltal javítva esélyeiket a versenyben.

Hol tart a technológia?

A fenti technológiák intenzív fejlődési szakaszban vannak. Léteznek nagyon jól használható alapmodulok, melyekre már lehet építeni, de még nem következett be az a komoly áttörés, amely a fenti problémák hatékony megoldását nyújtaná. Ennek oka egyrészt az, hogy az ipar még nem állt rá egy kiforrott technológia „gyártására”, így a kutatás-fejlesztésnek elsődleges szerepe van ezen a téren. Másrészt, ugyanezen okból, nagyon nagy lehetőségek rejlenek a témában: az, aki az első használható rendszer létrehozásában vezető szerepet játszik majd, könnyen válhat akár kis „garázscégből” is piacvezetővé, mint ahogy annak idején ezt a folyamatot a mobil- vagy internetes technológiáknál, illetve a személyi számítógépek piacán láhattuk.

Egy példa: asztrofizikai virtuális obszervatórium

A csillagászok által létrehozott Virtuális Obszervatórium kitűnő példája az adatbázisok összekapcsolásával létrejövő struktúráknak, a megoldandó problémáknak. Nézzük meg miért, rajta keresztül megérthetjük a lényegét. A csillagászat, hasonlóan más tudományokhoz kihasználta a Moore-törvényt. Az Intel egyik alapítója által még a 60-as években felállított tapasztalati törvény (http://en.wikipedia.org/wiki/Moore's_law) azt mondja ki, hogy a mikroelektronikai elemek sűrűsége nagyjából kétévenként duplázódik. Az exponenciális növekedést kifejező szabály valóban működik immár 4 évtizede, az olvasó könnyen kiszámíthatja, hogy ez hányszorosára növelte az egy chipen lévő elemek számát. A csillagászatot a számítógépek mellett leginkább az egyre olcsóbb és egyre nagyobb felbontású CCD-érzékelők megjelenése forradalmasította. Az előző évszázadok csillagászati Moore-törvénye, amely a távcsövek méretének növekedésében nyilvánult meg, a 90-es évek óta az érzékelők rohamos javulásában folytatódik. A Sloan Digital Sky Survey (SDSS) távcsöve például egy 120 megapixeles kamerát tartalmaz, amely 5 optikai sávban készít felvételeket. Ezek a felvételek azután hatalmas információ mennyiséget eredményeznek. A rengeteg adat kezeléséből adódó nehézségekről mindenkinek lehet sejtése, aki csupán egy pár megapixeles fényképezőgépet birtokol. Az SDSS végül 5 év alatt az égboltról egy 2 millió × 2 millió pixeles képet készít, és ez még csak a 90-es évek közepének technikája. A jelenleg fejlesztés alatt álló távcsövek akár egyetlen éjszaka alatt képesek lesznek erre.

Az adatokat számítógépeken tárolják, és dolgozzák fel. A hagyományos módszerek (fájlok, scriptek használata) mellett azonban az egyre gyorsabb számítógépek is képtelenek megbirkózni a hatalmas adatlavinával. A kutatók kénytelenek igénybe venni a legmodernebb információs technológiákat, így például az adatbázisokat, amelyeket eleve nagy adatmennyiségek kezelésére terveztek. Ezeket viszont inkább az üzleti élet igényeinek megfelelően alkották meg, ezért csak viszonylag egyszerűbb adatstruktúrákat képesek kezelni.

A tudományos adatok viszont szinte mindig komplexek. A galaxisokról például felvételeket készítenek, spektrumokat veszünk fel, kiszámoljuk fényességüket, morfológiai paramétereit, a szupernóvákról időben változó fényességgörbéket veszünk fel. De nemcsak az adatok bonyolultak, hanem a feldolgozás is. Míg egy üzleti példánál leggyakrabban egyetlen azonosító, például bankkártyaszám alapján kell az ügyfél vagy termék adatait előhozni, egy-egy tudományos adatfeldolgozási feladat jóval összetettebb eljárás lehet.

Tovább bonyolítja a helyzetet, hogy ugyanazon objektumhoz tartozó információk általában különböző kutatócsoportok világszerte elszórt adatbázisaiban vannak eltárolva. Van, aki rádió-, mikrohullámú, infravörös-, optikai, ultraibolya-, röntgen- vagy gamma-tartományban vizsgálja az eget. Az űsrobbanás megértéséhez, vagy egy távoli kvazárban, illetve gamma-kitörésben lezajló fizikai folyamatok felderítéséhez az adatokat összegezni kell. Mivel mindenütt egyre gyűlnek az új adatok, azokat gyakran újrakalibrálják. Nem lehet megoldani, hogy minden adat egy helyen legyen, ugyanakkor általában a legfrissebb adatokkal szeretnénk dolgozni. Össze kell tehát kötni az adatbázisokat úgy, hogy közösen lehessen bennük a kívánt adatokat keresni.

A Nemzetközi Virtuális Obszervatórium Szövetség (www.ivoa.net), melynek Magyarország is tagja (hvo.elte.hu), célkitűzéseinek középpontjában elsősorban olyan szabványok kialakítása áll, amelyek az adatbázisoknak ezt a föderációját teszik lehetővé, de a szervezet koordinálja és támogatja más, az eScience témakörébe illeszkedő, széleskörűen használható technológiák és eszközök létrehozását is.

Röviden említünk néhány témakört, amelynek megoldásában magyar szakemberek is részt vesznek. Fontos lenne az adatokat olyan metainformációval kiegészíteni, amely gépek számára is olvasható. Jelenleg minden adatbázishoz létezik emberek számára olvasható dokumentáció. Ideális az lenne, ha a gépek is tudnák, hogy ha az egyik adatbázis egyik oszlopa egy galaxis fényességét fejezi ki, akkor az egy másik adatbázisban levő galaxisfényességgel összevethető, de például égi koordinátával nem. Egy másik gyakori igény a nagy mennyiségű adathalmaz vizualizálása. A jelenlegi technológiák összekapcsolására alapozva ma már lehetséges az SDSS égtérképét is megjeleníteni, a terapixeles kép szabadon mozgatható, zoomolható.

Az új technológia nemcsak a kutatásnak, de az oktatásnak is nagy lehetőségeket nyújt. A <http://skyserver.elte.hu/myskyserver/hu/> „Szakkör” szekciója például olyan kis szakköri projekteket tartalmaz, amelyek akár középiskolában, akár egyetemi bevezető kurzusokban használhatók. A Virtuális Obszervatórium szellemének megfelelően a diákok ugyanazokat az adatokat használhatják, mint a kutatók, és egy ingyenesen elérhető virtuális távcső segítségével maguk észlelhetik az Univerzum tágulását, vagy rajzolhatják fel a csillagok fejlődési diagramját.

Csabai István, Papp Gábor

Eötvös Egyetem

eScience Regionális Egyetemi Tudásközpont