

SHAMAN a hálózati kezelésű, közös, örökségi archívum

BASSA LIA, DR. KISS FERENC

{lia, fek}@itm.bme.hu

Kulcsszavak: keresési módszerek, tárolási rendszerek, archívum-menedzselés

A megőrzés, továbbadás technikai feltételei és szervezeti lehetőségei is sokat változtak az elmúlt idők során. Az eredeti projekt elképzelés egy 2003-ban, Budapesten megszervezett konferencia eredményeiből indult ki, mely szerint szükség volna egy európai virtuális platform kialakítására az örökségek nyilvántartásaihoz. A cél olyan rendszerek és eszközök kifejlesztése, melyek támogatják a tudományos forrásokból származó digitális eredmények felhasználását és hozzáférhetőségét. Elő kell segíteni a digitális kultúra és a tudományos objektumok összességének kialakulását olyan összetett koncepciókkal ellátott bemutatási lehetőségekkel, amelyekhez a hozzáférési rendszert is megfelelően fejlesztik ki.

Az örökség a múltból ránk hagyományozott, vagy általunk az utókor számára megőrzendő ÉRTÉK. Lehet ez természeti vagy ember által készített, épített, létrehozott mű, amelyet valamilyen módon értékelhetővé, élvezhetővé kell tenni a következő generáció számára. Ennek feltétele, hogy megőrizzék. Ez egyszerre jelent megóvást (táj, épület, festmény stb.) illetve rögzítést (zene, tánc, népi hagyományok). A másik feltétel az, hogy az utókor hozzá tudjon férni, a kincs valóban közössé váljék.

Ennek szellemében hozták létre 1972-ben az UNESCO irányításával a Világörökség Egyezményt, melyet most már a Föld csaknem 200 országa aláírt. A 30. évfordulón tartott ünnepségük címe is ezt idézte „Közös örökség – közös felelősség”. A szervezet működési tapasztalatai és a védelmük alá vett helyszínek különbözősége miatt szabályai egyre finomulnak, egyre pontosabban határozzák meg, hogy a tagországok (Részben Államoknak nevezi őket ez a szervezet) milyen kötelezettségekkel rendelkeznek az épített és természeti értékek megőrzését, állagmegóvását, bemutatását illetően. Hamarosan pedig egy új egyezmény fogja meghatározni a nem tárgyi – immateriális – örökségek megőrzésének szabályait. Mindezt annak érdekében, hogy azok az örökségek, melyek egyediek és egyben univerzális értékűek, valóban mindenki számára hozzáférhetőek legyenek és maradjanak.

A különböző oktatási és kutató intézményekben jelenleg még egymástól függetlenül, helyi módszerek szerint archiválják a begyűjtött és regisztrált anyagokat. Egy-egy szakmán belül, de a társtudományok részéről is felmerült az igény egymás kutatási eredményeinek megismerésére. Az előző konferencia óta azonban az új követelmények és a nem szöveg alapú dokumentumok iránti érdeklődés növekedése következtében előtérbe került, hogy a digitális kultúra korában – összhangban az IST munkaprogramjával – a digitalizáláson túlmenően, összetettebb termékeket és szolgáltatásokat lenne kívánatos digitális formában elérhetővé tenni.

Fel kell tárnai, hogyan lehet az új kutatási eredmények, technikák és eszközök segítségével a digitális források hozzáférhetőségét megőrizni.

Több nemzetközi projekt célozta már meg a kulturális értékek digitális feldolgozásának rendszerezését, és mindegyik előrelépett egy-egy területen. Ezek közé tartozik az **ERPANET** (Electronic Resource Preservation Access Network), amely elsősorban az elektronikus adattörzítéssel kíván egy virtuális útirányt és tudásbázist megvalósítani az egyéni és intézményi szakmai információátadás területén a digitális nyilvántartás-fejlesztés eszközeivel.

A másik hivatkozott program, a **BRICKS** három területre összpontosítja munkáját:

- Az elektronikus adatkezelés technológiai fejlesztése a folyamatos fejlődés és növekedés figyelembe vételével.
- A projekt középpontjában az alkalmazások állnak, és az adatkezelés számára hozzájárulnak az alapvető szolgáltatásokat, melyek révén kiderül a potenciális felvevő piac mérete és a létrehozott infrastruktúra használhatósága is.
- Végül, nem utolsó sorban a fenntarthatóságot tűzi ki célul úgy, hogy egy olyan mechanizmust hoz létre, mely egy önmagát eltartó európai eszköztárként funkcionál.

Visszatérve a SHAMAN projektre, amely munkája előkészítéseként először körülnézett Európában, hogy hol merül fel a digitális archívum készítésének és közös-kölcsönös használatának igénye. A válasz már a fentiekből is kiderült. Szem előtt kell tartani azt is, hogy a kulturális örökségek nagy kalapjába tartoznak az épített és műtárgyi örökségeken kívül az immateriális örökségeknek nevezett területek, többek között a nyelvészet, az irodalom vagy a folklór is, melyeknek közös használatára nagy igény van, de nagyon szétszórtnak és változatos technikai szinten archiválják.

Az új projekt iránt Európa szerte nagy volt az érdeklődés: 12 országból 26 intézmény jelentkezett a rész-

vételre. Gondoljunk itt az egyes nyelvek, népszokások, zene, tánc egy nemzeten belüli közös területeire illetve az egyes népek ilyen jellegű tudományainak összefonódásaira. Ezek a kutatások olyan egyetemes értékűek, hogy minden kétséget kizáróan általános igény, hogy nyilvántartási rendszereikbe be lehessen tekinteni, és hozzáférhetőek legyenek az érdekelt szakemberek részére.

Egy ilyen átfogó, kiterjedt, sok szakmát felölelő archívum létrehozása sok technikai problémát vet fel. Természetesen szakmai kutatások eredményeit nem lehet előre meghatározott szabályozási módszer, valamint a jogosultságok megállapítása nélkül, feltétel nélkül közléteni. A közös archívum létrehozása bizonyos közös alaprendszer meglétét is feltételezi, hiszen hasznosítani csak az erre épülő rendszerbe szedett adatokat lehet. Ha sikerül egy olyan alapon megállapodni, amit a résztvevők fel tudnak tölteni a rendelkezésükre álló adatokkal, akkor létrejöhet egy általánosan használható, szakmai, digitálisan hozzáférhető örökségi archívum, a **SHAMAN** (Shared Heritage Archives Management Across Networks). Így válnának különböző típusú kulturális örökségek: épületek, műalkotások, szövegmélek, de még zenei és tánc produkciók is minden résztvevő által használhatóvá, a feldolgozás, az együttműködés, és a további kutatási feladatok megvalósítása céljából.

Egy ilyen jellegű munka megtervezéséhez a műszaki és a szakmai területeken járatos szakemberek sokrétű együttműködésére van szükség. A projekt megtervezése is alapos előkészítést igényel. Az első ilyen jellegű találkozót *Dr. Darányi Sándor a Svéd Könyvtári Informatikai Intézet munkatársa* szervezte meg Borasban 2004. december 6-7. között svéd, magyar, angol, amerikai, portugál, belga, finn, norvég, lett, litván és olasz előadók részvételével. Először minden résztvevő bemutatta saját fejlesztéseit, oktatási és archiválási munkáját, többek közt egy belga zenefelismerő rendszert, észt folklór archívumot, finn nyelvészeti leírást, olasz mozgásfelismerő rendszert, és nem utolsósorban a magyar UNESCO világörökségi adattárát, melyekről alább olvashatunk.

Következő lépésként projektmodulokat vázoltak fel. A különböző egységek tartalmazhatnának egyes műalkotásokat (műveket), meglévő leírásaikat, kapcsolódó szabványokat, beviteli technikákat, jellegzetességi leírásokat, adatbázisokat, hozzáférési feladatokat, tudásfejlesztést, digitális jogkezelést, felhasználói felületeket, rendszerfelépítést és felhasználási forgatókönyveket. A fentiekre láttunk több nemzetközi példát.

Figyelembe véve azt a tényt, hogy a rendszer felépítésével és a felhasználási célkitűzésekkel kapcsolatos döntés alapvető az egész rendszer szempontjából, ezeket pedig az adatgyűjtés módszerei határozzák meg. A felsorolt három szempont indíthatja el a tényleges munkát.

Ezután a decemberi szeminárium két munkacsoportban dolgozott tovább körülírva a rendszerszerkezet felépítésének és felhasználási fejlesztésének leg-

fontosabb szempontjait. E munka eredményeként három próba-kérdőív megtervezésére került sor: egy az anyaggyűjtésre, egy a projekt számára hozzáférhető K+F források eléréséhez és egy az értékelési módszerek kialakításához. Az elsőt Helsinkiben és Tartuban próbálják ki, valamint az érdekeltek segítségével a meglévő anyagok alapján töltik ki. Az alapelv szerint ezeket a kérdőíveket azok fogják elemezni, illetve a projektben való részvételt azoknak fogják felajánlani, akiknek a céljai megegyeznek a projekt alapvető célkitűzéseivel.

Czeslaw Jan Grycz, egy amerikai cég, az Octavo Ltd. vezetője bemutatta, hogyan rögzítettek digitálisan régi kódexeket, értékes, egyedi, könyvtárak által nem kölcsönzött, tudományos műveket, eredeti formájuknak megfelelően. Az olvasó úgy lapozhat és olvashatja a könyvet, mintha a kezében tartaná.

Finnországban elkészült egy olyan nyelvészeti adatbázis, amely nyelvcsaládokról ad meg minden információt földrajzi, nyelvészeti, népességi, történeti szempontok szerint.

Lettországon néprajzi kutatások számítógépes feldolgozása ad lehetőséget a szakma művelői, tanulói és rokonszakmák kutatói számára a kapcsolatok megismerésére. Mindezek mutatják, hogy egymástól függetlenül, de sok területen végeznek ilyen jellegű feldolgozásokat, melyeknek rendszerbe foglalása egyértelműen a következő feladat.

Európa, sőt a világ legtöbb országában foglalkoznak már örökségvédelemmel, ami szükségessé teszi az örökségek minél részletesebb megismerését, vizsgálatát. Hogy mindenki számára érthetővé váljon, miről van szó, az egyik legkülönlegesebb példával kezdünk, amely a mozgás számítógépes rögzítéséről és feldolgozásáról számolt be. E kutatás, amellyel, hogy részese tud lenni a fenti archívumnak, kulturális felhasználásán túl az egészségügyben is hasznosítható: mozgászervi betegségek leírására és a rehabilitációs program kidolgozásához.

A Genovai Egyetemen Antonio Camurri professzor laboratóriumában végzett kísérletek egy MIEE (Multi-sensory Integrated Expressive Environments) alkalmazást használtak az előadóművészeti produkciók leírására, melynek révén multimédiás koncerteket, interaktív múzeumi kiállításokat, tánc produkciókat, videó installációkat, színházi előadásokat dolgoztak fel. A felhasználó-központú rendszer igen magas szintű információt tud befogadni elsősorban a mozgást kifejező mozdulatok természetéről. Így például érzelmi, energia, sebességi és intenzitási tulajdonságairól.

Már számos interaktív rendszer létezik és használatos, melyek képesek audió- és videó anyagok rögzítésére: PureData, Max/MSP (www.cycling74.com), Isadora (www.troikatronix.com). Ezek a rendszerek elsősorban egyfajta interakcióval foglalkoznak, tehát kizárólag audió vagy videó adatokkal, viszont érzékelői és multimodális funkciójuk alig van, vagy egyáltalán nincs. Ennek az igénye az elmúlt évek során merült csak fel. A multimodális kapcsolatok nemcsak azért kerültek elő-

térbe, hogy egyszerre több területen tudjanak dolgozni a kutatók, hanem ezáltal megnőtt a különböző absztrakciós szinteken a kapacitás is, valamint támogatni tudnak különböző csatornákon érkező integrált eljárásokat is. Az EU-IST MEGA (Multisensory Expressive Gesture Applications, www.megaproject.org) projekt határozta meg azt a koncepcionális keretet, amely a multimodális, kifejező mozgásleírásokra vonatkozik, és négy szintje van.

- Az első szint a fizikai jeleké, algoritmusokat tartalmaz, amelyek alapján az érzékelők (videokamera, mikrofon, környezeti, robot és testen lévő érzékelők, például gyorsulásmérő) rögzítik a begyűjtendő adatokat.
- A második szint az alsóbb kategóriájú jellegzetességeké, kiválasztja az érzékelők adatgyűjteményéből azokat a nem elsődleges osztályozás szerinti szempontokat, amelyek leírják a mozdulat véghezvitelét. A tánc esetében ezek a szempontok lehetnek például kinematikai mérések eredményei, vagy akár a testrészek sebessége, gyorsulása, a vizsgált mozgás mennyisége, a test összehúzódása, kinyúlása.
- A harmadik szinten a közép kategóriájú jellegzetességek és térképek két fő témával foglalkoznak: az input anyag – mozgás és zene – szegmentálásával a felépítő mozdulatok szerint valamint ezeknek a mozdulatoknak a megjelenésével az adott térben. Így az első feladat itt az input megfelelő méretű egységekre bontása, és ezek összekapcsolása azokkal a kategóriákkal, melyeket a kommunikáció szempontjából fontosnak tartanak. A tánc elemzésében az előadás egy töredéke mozdulatok sorozatára bontható, ahol a mozdulatok határait a sebesség és irány variációiként lehet lejegyezni. Az egy mozdulaton végzett méréseket vektorokká alakítják, amely egy területen meghatározza azokat a szemantikai jeleket, amelyeket az érzelmekhez és a kifejezésekhez lehet kötni. Az időben és térben lezajló mozdulatsorokat így átalakítják a szemantikai térben ábrázolható adatokká, melyeket aztán megfelelően csoportosítva elemezni lehet.
- A negyedik szint a koncepcióké és szerkezeteké, már közvetlenül az adatelemzéssel foglalkozik. Itt egy olyan szemléletű hálózatról van szó, amely feltérképezi a leírt tulajdonságokat és mozdulatokat, majd ezeket verbális szerkezetekbe alakítja át. Például egy tánc előadását lehet az előadó által kifejezett érzelmek szándékai szerint elemezni. Az alapérzelmek – a düh, a félelem, a bánat és az öröm – szándékai szerint elemezni. Más feldolgozási lehetőségek is vannak, például a mozdulatokra fordított erő szempontjából, mely lehet sikló, sima vagy szaggatott. Végeztek kísérleteket a nézők reakciónak modellezésére is.

A gépi tanulási technikák lehetnek lineárisak vagy akár logikai, esetleg valószínűségi rendszereken át egész különböző ideghálózat típusokig, melyeket vektorokkal ugyanúgy le lehet írni, mint döntési fákkal.

Az elvégzett kísérletek azt mutatták, hogy az érzelmek alapján létrehozott osztályozó rendszer a döntéseknek megfelelően helyes osztályozást nyújtott a nézők besorolásának szempontjából is. Egy hasonló módon végzett, zenei kísérletben azt vizsgáltuk, hogy milyen a hallgató elkötelezettsége, reakciója a zene iránt adott pillanatban, és az eredmények azt igazolták, hogy a vizsgált zenei mű (Scriabin: Etűd) kiemelkedő, mondjuk úgy, kulcspontjai egybeestek az előadó mozgásában tapasztalható sarokpontokkal, és a nézőkben felfedezhető reakciókkal.

Legalább ennyire érdekes a zene digitális archiválásával foglalkozó program, melyet a belgiumi *Gentből érkezett résztvevő, Marc Leman* mutatott be. A genti egyetemmel együttműködik az **IPEM** (Institute of Psychoacoustics and Electronic Music), amely mára kiváló kutatási infrastruktúrára tett szert teljesen digitalizált környezetben a zene- és multimédiakutatás céljainak megfelelően.

Nemzetközi szinten az IPEM a zenetudomány új kutatóközpontjaként jelenik meg, ahol tudományos eredményeik a pszichoakusztikai- és mozgáselméletekre alapulnak az ismeretelmélet és a módszertan eszközeinek felhasználásával. Részt vettek számos nemzetközi szimpóziumon és eszmecsserén a mesterséges intelligencia és a zene, a zene és az informatika, illetve a kognitív zenetudomány tárgykörében, ugyanakkor jelenleg az intézet az egyik központja az ilyen jellegű tudományos munkának. A jövőben folytatni kívánják interdiszciplináris kutatásaikat a zenetudomány területén.

Kutatásaik között szerepel jelenleg a Boras-ban is bemutatott MAMI (Musical Audio Mining) projekt – amely akár dúdolás alapján képes dallamokat felismerni, és rendszerébe besorolni megadva a dallam szerzőjét és egyéb adatait. Egy másik projektjük a DEKKMMA, egy közép-afrikai, népzenei hangarchívum digitalizálásával foglalkozik, mely 2521 óra zenei anyagot tartalmaz, amelyek közül a legrégebbi 1910-ből származik. A belga gyarmatokról származó anyag valószínűleg mind méretét, mind dokumentációs anyagát és minőségét tekintve egyedülálló, és itt is a legnagyobb gondot az jelenti – mint minden hasonló esetben –, hogy a nagyon érzékeny, mulandó, könnyen elenyésző anyagokon rögzített zenét és a hozzá kapcsolódó információt időtálló, digitalizált formában meg tudják őrizni az utókor számára.

A társadalom és művészetek területén végzett kutatásaikhoz, azonban a hatóságok fokozottabb közreműködésére van szükségük. Az általuk elért eredmények azonban nemcsak a múlt megőrzését, az örökségek védelmét szolgálják, hanem oktatási célokra is felhasználják azokat. Foglalkoznak továbbá a számítógép által generált zene kialakulásának ismertetésével a fiatal zeneszerzők részére. Mivel ez egy új területe a zenepedagógiának, új oktatási programot kell kidolgozni hozzá. Mindez együtt feltétlenül hatással van a zenetudomány további fejlődésére, hiszen létezik már olyan interaktív számítógépes zene is, amely a hallga-

tói elemzések alapján intelligens interakciót hoz létre a kialakuló zenei környezetben.

A *Sheffieldi Egyetemen* is foglalkoznak zenekutatással, de ott az adatbázis felépítése, az információkutatás, a kutatási módszertan, a modellezés és az adatintegráció kidolgozása az elsődleges.

A begyűjtött anyag digitalizálásának méreteinek és a terv nagyságának érzékeltetésére bemutatnánk a **Litván Folklor Archívum** eddig már elvégzett munkáját, amelyet egy tudományos társaság kezdett meg 1907-ben, és az Archívum hivatalos, 1935. évi megalapítása óta folyamatosan töltenek fel a változó adathordozók segítségével. Mindez a lehetőségek csak egy töredékét tartalmazza eddig. Kézirattáruk 8400 kézirat több mint 1,5 millió címszavát tartalmazza, többek között 570.000 népdalt, körülbelül 160.000 népi elbeszélést, 350.000 kisebb népi alkotást (közmondást, szólamot, versikét) és még 420.000 be nem sorolt anyagot. Hangtárunkban 350 felvétel van az 1908-1949-es évek közötti időszakból, 7000 hangrögzítés készült 1935 és 1949 között, csaknem 260.000 hangszalagos rögzítésük van az 1952 és 1994-es periódusból, több mint 1000 kazettán rögzített 42 000 anyagot 1971-2004-ig, a 100-at meghaladó DATA minidisc-ek 123 órányi anyagot tárolnak, valamint a vizuális anyagok 8500 fényképet és negatívjaikat tartalmazzák. Megközelítőleg 10-15 ezer cikkel bővül a tár évente.

Magyarország a szervezésen kívül vállalja egy speciális terület digitális archívumának bemutatását is. Az **MTA Néprajzi Kutatóintézetében Hoppál Mihály igazgató és kutatócsoportja** gondozza a sámánizmus archívumot. Számos együttműködő partnerszervezettel együtt (International Society for Shamanistic Research, Magyar Vallástudományi Társaság stb.) több kötetet adtak már ki: a *Bibliotheca Shamanistica* című könyvsorozatot és a *Sámánok és Kultúrák* című könyvet. Emellett hatalmas, mintegy 2000 oldalas, rendszerezett, de eddig elsősorban papíralapú gyűjteményük van, készítették kb. 200 óra filmet, 5000 fényképet és van még 10 000 oldal kéziratuk és szakmai előadásuk.

A digitális feldolgozást a Budapesti Műszaki Egyetem Információ- és Tudásmenedzsment Tanszéke fogja elvégezni, és hozzáférhetővé tenni a projektpartnerek részére.

A SHAMAN projekt alapkövetelménye, hogy bővítésre alkalmas rendszerben kell felépíteni. Tartalmának egyes részei az átlagos nagyközönség részére is elérhetőek lehetnek a jövőben. Fő feladata azonban a szakmai felhasználók kereséseinek kielégítése, amely szintén sokrétű, hiszen az egyetemi hallgatóktól a doktoranduszoktól át a hatóságokig és a szakemberekig mindenki részére gondoskodni kell megfelelő tudományos szintű és jogilag tisztázott hozzáférésről. A metaadat kezeléséhez tématerképeket és thesaurusokat szükséges készíteni, továbbá a koncepcionális szerkezet kialakításához a környezet, a tartalom és az osztályozás rendszerét kell kidolgozni.

A méretek, mennyiségek miatt a kiépítés jelenlegi szakaszában a projekt csak a folklór tanulmányozására terjedne ki, és csak már meglévő, digitalizált anyagokkal dolgozna, tehát újabb digitalizálási művelet nem lenne a jelenlegi projekt része.

A gyűjteményekkel kapcsolatos kérdőíveket január közepéig kellett visszajuttatni Boras-ba véglegesítés, és az érintett partnerek tájékoztatása céljából. A K+F és értékelési kérdőívekkel kapcsolatban a résztvevő svédországi egyetemek a felelősek. A rendszer függőségének csökkentése érdekében a tartalmi és metaadat leírásokat úgy kell kialakítani, hogy függetlenek legyenek eredeti dokumentumaiktól.

A felhasználandó szoftvereket a konzorcium tagjai számára ingyenesen teszik hozzáférhetővé, és így lehetőség nyílik a tesztelés elvégzésére Budapesten, Portóban, Genovában és Gentben. Döntés az eredményről a következő projekttalálkozón várható, melyet májusban Budapesten, a *Budapesti Műszaki és Gazdaságtudományi Egyetem Információ- és Tudásmenedzsment Tanszéke* szervezésében rendeznek meg a hálózati kezelésszerű, közös, örökségi archívum (SHAMAN) résztvevői számára.

Hírek

Az **Interware és a Sun Microsystems** hosszú távú együttműködés keretében kibővítette bérszerver kínálatát azoknak, akiknek nagy teljesítményű, üzleti alkalmazások használatából adódó terhelés mellett is biztonságos és gyors háttérre van szükségük. A szerverhotelben eddig túlnyomórészt webszervereket helyeztek el az ügyfelek, de egyre többen élnek a szerverbérlet lehetőségével. A Sun Microsystems kínálatával az alkalmazásszervereket működtetni kívánó, főként vállalati ügyfélkör igényeit fogja kiszolgálni a szerverhotel.

A magyarországi szerverhoszting piac – elsősorban a szélessávú internetelérések rohamos terjedésének, valamint az egyre szélesebb körben elérhető szélessávú tartalmaknak köszönhetően – az elmúlt időszakban dinamikusabban növekedett. Az egyre több szervert üzemeltető vállalkozások felismerték, hogy a telephelyükön történő szerverüzemeltetés erőforrást von el a munkatársaktól, és költséghatékonyabb módszer, ha kiszolgálójukat egy jól felszerelt adatközpontban helyezik el. Ráadásul az ügyfél telephelyén a szerverek eléréséhez szükséges sáv szélesség biztosítása jóval nagyobb költséget jelent, mint a szerverhotelben való elhelyezés.