

Egy országos méretű orvosi adatbázissal kapcsolatos informatikai kihívások

Pataki Máté, Kovács László, Pataki Balázs

Kivonat

Az Országos Rákregiszter Grid GVOP Projekt (ORG) célja a Nemzeti Rákregiszter számára egységes adatkezelést lehetővé tévő, elosztott rendszerarchitektúrájú, szabványalapú informatikai megoldás elkészítése volt. Különböző tudományterületeken, de az élet szinte minden területén gondot okoz a nagy tömegű, bonyolult szerkezetű adatok összegyűjtése, feldolgozása, biztonságos tárolása, és azok oly módon történő hálózati publikálása, hogy az a kutatás, a politikai, gazdasági vagy más szakmai döntéshozatal és egyéb társadalmi felhasználás számára (adatbányászati, statisztikai módszerekkel) értékelhető legyen. Előadásunkban igyekszünk összegyűjteni és átadni azokat a tapasztalatokat, amelyek nem csak erre a feladatra specifikusak, hanem általánosságban is hasznosak lehetnek mások számára is.

A projektben egy konkrét elosztott adattárházat hoztunk létre, de az informatikai megoldást olyan általánosan alakítottuk ki, hogy az felhasználható legyen egyéb orvosi, illetve más területeken is (pl.: közvélemény-kutatás, társadalom és egyéb tudományos kutatások), adatok gyűjtésére, és azok egymáshoz való összevetésére. A kutatók munkáját egy külön kutatói portál segíti, mely lehetőséget ad az anonimizált adatok statisztikai lekérdezésére, és tetszőleges módon történő kombinálására.

Egy ilyen jelentőségű adatbázis esetén az adatminőség kérdése kulcskérdés volt. Az adatminőség javítására projektünk a totális adatminőség kontroll módszerét dolgozta ki, mely az adatok elsődleges előállításától, azok végső felhasználásáig teljes életciklusban végzi az adatok konzisztenciájának biztosítását, ellenőrzését, hibajelzését, automatikus javítását. Mindezt könnyen bővíthető módon, akár intézetenként testre szabhatóan, úgynevezett konzisztencia-kritériumokkal, melyek olyan egyszerű ellenőrzéstől, hogy a beteg neve megfelel-e a nevének, olyan összetetteket is tartalmazhatnak, hogy az adott beteg korábbi életútja alapján a jelenlegi betegsége mennyire valószínű. Amennyiben kérdés merül fel, azt visszajuttatja a rendszer a kórházba, és az orvos dönthet az adat elfogadásáról, vagy javításáról.

Az adatok biztonsága érdekében minden kommunikáció a központ és a felhasználók, a központ és a kórházak, valamint a kórházi rendszer és az orvosok között is, titkosított csatornákon történik. A hosszú távú tárolást és a későbbi esetleges egyéb rendszerekkel való kompatibilitást, az XML alapú, szabványos formátumok biztosítják.

Linkek

ORG Projekt Honlap: <http://dsd.sztaki.hu/projects/org/en/>

MTA SZTAKI Elosztott Rendszerek Osztály: <http://dsd.sztaki.hu>

Arvato Systems Magyarország: <http://www.arvato-systems.hu/>

Országos Onkológiai Intézet: <http://www.oncol.hu/>