

Intézményi interfész leírás

A felsőoktatási intézmények által az AVIR központi adattárba küldendő adatok interfészének bemutatása

Tartalomjegyzék

| | |
|---|----------|
| TARTALOMJEGYZÉK..... | 1 |
| A FELSŐOKTATÁS ÁGAZATI ADATTÁR (AVIR) ÉS A FELSŐOKTATÁSI INTÉZMÉNYEK KAPCSOLATA..... | 2 |
| AZ INTERFÉSZ TARTALMI LEÍRÁSA, ÉRTELMEZÉSE..... | 2 |
| „INTÉZMÉNYI INTERFÉSZ TÁBLÁK” MUNKALAP (SHEET) | 3 |
| „INTERFÉSZ MEZŐK” MUNKALAP (SHEET)..... | 4 |
| TECHNIKAI HÁTTÉR | 7 |
| ADATOK BETÖLTÉSE..... | 7 |
| ADATOK PUBLIKÁLÁSA | 8 |
| TÁMOGATÁSI TEVÉKENYSÉG | 8 |
| MELLÉKLETEK | 9 |

A felsőoktatás ágazati adattár (AVIR) és a felsőoktatási intézmények kapcsolata

A felsőoktatási ágazati adattár célja az ágazat és a felsőoktatásban érintett intézmények számára szükséges információk biztosítása egységes, integrált módon. Figyelembe véve az adatszolgáltatási kötelezettségeket, arra törekedve, hogy középtávon elérje a rendszeres és az ad hoc adatszolgáltatások kiváltását. Az adattárba az adatok döntően a már jelenleg is működő informatikai rendszerek adatbázisaiból, és a TÁMOP 4.1.1 keretében ezekre épülő intézményi adattárból kerülnek be. Ennek két fő forrása van: különböző központi adatbázisok és adatgyűjtések, valamint az intézményi rendszerek adatbázisai. A TÁMOP 4.1.1 pályázatban nyertes intézmények saját intézményi adattárat fejlesztenek, a különböző intézményi rendszerek adatai ebben az adattárban kerülnek gyűjtésre, integrálásra.

Az intézményi interfész a központi adattár és az intézmények közötti adatkommunikációs felület. Az interfész olyan adattáblákat tartalmaz, amelyeket az intézmények töltenek, és amelyet a központi adattár használ fel, az interfészen keresztül küldött adatok összevezetése és kezelése a központi adattár feladata.

Az adatszolgáltatási rend könnyebbé tétele mellett a központi adattár további célja és feladata, hogy elősegítse a tervezést az ágazat számára, a visszaszolgáltatók által (amit az intézmények kapnak majd a központi adattárból) pedig az intézmények számára segítse ezt a folyamatot.

Mindezek jelenleg csak a TÁMOP 4.1.1-ben nyertes felsőoktatási intézmények számára kötelező feladat és kihívás, de jelenleg zajlik a terület jogszabályi rendezése is. Ennek alapján majd a most még nem fejlesztő intézmények is részt vesznek majd ebben a folyamatban, oly módon, hogy számukra az Educatio Nonprofit Kft. biztosítja majd a megfelelő adatszolgáltatási felületet.

Az interfész adattartalmának kialakításakor a következő irányelvek, szempontok vezérelték a munkát.

- a FIR és felvételi rendszerek (pl. Gólya) által jelenleg is kezelt adatkörök ezekből a rendszerekből kerülnek a központi adattárba, ezeket az intézményeknek nem kell párhuzamosan az AVIR-ba is küldenie.
- az oktatásért felelős minisztérium és háttérintézményei részére történő adatszolgáltatások lehetőség szerint az AVIR-ból történhessenek. Tehát ezek adattartalmát a központi adattárnak juttassa el az intézmény. (A működés kezdetén ez párhuzamos adatszolgáltatást eredményez, egészen addig, amíg a központi adattár feltöltöttsége és adatminősége el nem éri a használhatósághoz szükséges és elégséges szintet.)
- az oktatási ágazaton kívüli adatszolgáltatások tartalmát a központi adattár az adatkérőtől gyűjtse be (pl.: a pénzügyi adatok többségét a Magyar Államkincstártól, a kutatási adatok egy részét a KSH-tól).

Az interfész tartalmi leírása, értelmezése

Az interfész adattartalmát az interfesz.xls fájl tartalmazza. Az excel fájl részletes leírása:

„Intézményi interfész táblák” munkalap (sheet)

Az interfész táblánkénti összefoglaló információkat tartalmazza.

Témakör

Az AVIR projektben használt témaköri besorolás 2. szintje. Az interfész a következő témaköröket érinti:

- oktatás
- kutatás
- pénzügyi és vagyongazdálkodás
- HR gazdálkodás
- létesítmény gazdálkodás
- hallgatói szolgáltatások
- intézményi szolgáltatások

Ezen kívül önálló témakörként megjelenik:

- hároméves megállapodások

Téma

Az adott témakörön belüli téma megjelölése. Pl az oktatás témakörön belül két témát érint az interfész: hallgatók képzései; tantárgyak, kurzusok.

Intézményi interfész tábla

A tényleges interfész tábla logikai neve. Az adott témán belül lehetnek olyan adatok, amelyeket nem lehet 1 táblában megadni, mert azok tartalma vagy részletezettségi szintje nagyon különböző. Az interfész tábla neve jelzi az adott tábla tartalmát. Egy interfésztábla fizikailag is 1 adattáblát is jelent.

Kötelezőség

Különböző típusú intézményektől (állami fenntartású és nem állami fenntartású (pl.: egyházi, magán, alapítványi stb.) nem minden adatot kér be. Minden interfész táblára külön megadásra került, hogy mely intézményeknek kell a táblát elkészíteni és küldeni. Az „OKM fenntartású” táblákat értelemszerűen csak azoknak az intézményeknek kötelező, amelyek állami finanszírozásúak a többi intézmény opcionálisan megadhatja ezeket az adatokat.

Verzió

A teljes interfész fogadó oldali programjai két ütemben készülnek el. Az 1-es verzió jelzi azokat a táblákat, melyek az első fejlesztési körben, azaz **2011 április 30-tól küldendők**. Ezek egyértelmű és előnyben részesített ágazati igényekre épülnek, és megfelelő forrásrendszerek állnak rendelkezésre az intézmények nagy részébenben.

A 2-es verzió tábláinak fogadása csak később készül el, ezeket az első adatszolgáltatások alkalmával nem kell küldeni, de a jelenlegi intézményi projektek keretében meg kell kezdeni a szolgáltatásukat, **legkésőbb 2011. november 30-tól**.

Adatküldés gyakorisága

Az adott táblák elkészítésének és küldésének gyakorisága. A rendszeres adatszolgáltatás a tipikus, azonban van olyan terület, ahol csak meghatározott alkalmakkor kell az táblákat elküldeni, például a Diplomás Pályakövetési Rendszer felmérési adatait a felmérés lezárta után, a kapacitás akkreditációs adatokat pedig csak az akkreditációkor.

Intézményi interfész tábla granularitása

A granularitás az adott tábla részletezettségi szintjét jelöli. Azt adja meg, hogy a tábla 1-1 rekordja milyen mélységben kell, hogy összesítse az adatokat, illetve, hogy elemi adatok esetén mi az elemi adat. Példák:

- Hallgatói térítések tábla; granularitás: tranzakciók; értelmezés: tábla minden rekordja egy konkrét hallgatói befizetést tartalmaz.
- Ingatlan költség tábla; granularitás: ingatlan x közmű; értelmezés: éves összesítésben ingatlanonként annyi sora kell, hogy legyen a táblának, ahány különböző közmű szolgáltat az ingatlan felé.
- Hallgatói kreditek tábla; granularitás: hallgató x tantárgy; értelmezés: egy hallgató annyiszor szerepel a táblában, ahány tantárgyat felvett az adott tanulmányi félévben.

„Interfész mezők” munkalap (sheet)

Az egyes interfész táblák részletes mező leírása

Interfész tábla

Melyik interfész táblánál kell megadni az adatot

Típus

A lehetséges típusok:

- id – táblán belüli azonosító, amely lehet az intézmény által adott azonosító, vagy már korábban ágazati szinten egységesen használt azonosító (pl. FIR azonosító). Az egyedi azonosító használatánál a táblák közötti konzisztenciát biztosítani kell, azaz egységesen ugyanazt a kulcsot kell használni a különböző táblákba az egyazon elemekre.
- mérőszám – numerikus értékek
- dimenzió – karakteres értékek, amelyek a mérőszámokat csoportosítják, osztályozzák, besorolását lehetővé teszik.

Mező

A mező logikai neve, rövid elnevezése

Leírás

Rövid magyarázat a mező tartalmáról

Kapcsolódó fogalomtári definíciók

Mely mezőkhöz tartozik fogalomtárban külön definíció

Mező típusa

Az adott mező milyen típusú: karakteres, dátum, numerikus.

A numerikus típusok jelentése:

| Típus | Értékhatar | Méret |
|----------|--------------------------|--------|
| smallint | $(-2^{15}) - (2^{15}-1)$ | 2 byte |
| int | $(-2^{31}) - (2^{31}-1)$ | 4 byte |
| bigint | $(-2^{63}) - (2^{63}-1)$ | 8 byte |
| float | $(-3.4E+38) - (3.4E+38)$ | 4 byte |

Szótár típus / idegen kulcs

Akkor kitöltött, ha az adott mező tartalma kötött, azaz csak előre meghatározott elemeket tartalmazhat a mező. Ennek 2 fő típusa van:

- szótár: csak azokat az értékeket veheti fel a mező, amelyek a „szótár” munkalapon az itt megadott névhez kapcsolódóan szerepelnek
- idegen kulcs: csak olyan értékeket vehet fel, amelyeket a másik rendszerben használnak. Ez tipikusan a FIR-re vonatkozik. Ebben az oszlopban az kerül definiálásra, hogy a FIR melyik táblájának mely mezője adja meg, hogy milyen értékeket tartalmazhat az interfész mező.

Lehet NULL

A beérkező adatok feldolgozásakor a programok megkövetelik bizonyos adatmezők kitöltését. Ezek üresen hagyása esetén az adatcsomag nem dolgozható fel. A „*lehet NULL*” oszlop adja meg, hogy az adott mező üresen hagyása meggátolja-e a feldolgozást. Lehetséges értékei:

- **nem**: amennyiben a mező üres az intézményi adatcsomagban, akkor a feldolgozó program a teljes adattábla feldolgozását megszakítja. Nem lehetséges tehát olyan feldolgozás, ami a beküldött csomag egyes sorait beépítik az adattárba, másokat pedig nem. Egyes szótár elemek már most is rendelkeznek olyan értékkel, ami a „nem ismert” adatot jelentik. Ezek alkalmazhatók olyan esetben, ahol az intézmény nem tudja előállítani az adott mező értékét.
- **igen**: a mező lehet üres az adatcsomagban. Az üres mező nem akadályozza a feldolgozást.

A fenti információ technikai jelentéssel bír. Egy olyan mező, ami „lehet NULL”, tartalmilag lehet kötelező. A hiányzó érték beengedésének oka az, hogy az információ hiányában is tudjon az intézmény adatot szolgáltatni. Így az országos adattárban az információ egy része rendelkezésre áll majd, és az is tudható, hogy egy adott intézmény mekkora hányadát tudja teljes kitöltöttséggel szolgáltatni az adatoknak.

Verzió

Mező szinten megadja, hogy az adat feldolgozása a fejlesztés melyik ütemétől indul

Mező sorrend

Azt adja meg, hogy az adott interfésztáblán belül az egyes mezőknek milyen sorrendben kell következniük.

„Szótár” munkalap (sheet)

Az egyes mezők engedélyezett értékkészletét tartalmazza.

Tipus

Az „Interfész mezők” munkalap „Szótár típus / idegen kulcs” oszlopával megegyező szótári készletet azonosító kód.

Tipus leírása

A szótári típus leírása.

Rövid név

Az értékkészlet adott elemének rövidítése, melyet az adatküldésnél használni kell.

Jelentés

Az „Rövid név” oszlopban megjelöl kód kifejtése, jelentése

Sorrend

Az AVIR-ban használt rendezési szempont. A központi vezetői rendszer kimutatásain ilyen sorrendben jelennek meg az elemek

„Témakörök adatforrása” munkalap (sheet)

A központi adattár témaköreit sorolja fel, és megadja, hogy melyik témakörnek mi az adatforrása az országos adattárban. Az interfész szempontjából nincs jelentősége, a dokumentumba azért került bele, hogy az adatküldőnek rálátása legyen arra, hogy az egyes adatköröket honnan gyűjti be a központi rendszer.

Témakör

Az „Intézményi interfész táblák” munkalapról már említett témakör szerinti besorolás.

Téma

A központi adattár témái.

Javasolt Interfész rsz

A fejlesztés során mely rendszerből fog tölteni az adott témakör tartalma. Lehetséges értékei:

- APEH: APEH adatkapcsolat
- excel: excel táblák
- FELVI: felvételi adatbázis
- FIR: FIR adatbázis
- FIR, INT: FIR adatbázis és intézményi adatszolgáltatás
- FSZH: FSZH adatkapcsolat
- INT: intézményi adatszolgáltatás
- KSH: KSH adatkapcsolat
- MÁK: MÁK adatkapcsolat

Az üres cella azt jelenti, hogy az adott témakör igényként jelentkezett a felméréseken, de megfelelő minőségű, elérhető adatforrást a tervezés során nem sikerült találni. A fenti adatforrások mindegyikének integrációja nem történik meg a fejlesztés jelen fázisában.

Technikai háttér

Adatok betöltése

Az adatokat a Központi Adattár 3 lehetséges úton képes fogadni, s azokat az intézmények felé eljuttatni:

- Webszolgáltatás (WS) alapú felületen (XML),
- IBM WebSphere Message Queue-n (MQ) keresztül,
- Webes felületen feltöltött CSV fájlok segítségével.

Az első két esetben XML alapú üzenetek közlekednek, amelyek pontos formátumai XSD-kben vannak rögzítve. A CSV esetén is kötött a beküldendő adat formátuma. Az MQ és a WS alapú adatkapcsolat esetén a beküldendő adatokból egy előre meghatározott algoritmus alapján ellenőrző, ún. HASH kódot kell generálni, melynek beküldésével az adattartalom igazolható.

Az intézmény forrásonként egyszerre csak egy időszak adatait töltheti és azt teljes egészében kell töltenie. Hogy melyik forrásnál melyik időszakot tölti, a Vezérlőtábla megfelelő mezőjében adja meg. Ez a tábla tartalmazza azt is, hogy betöltési-, illetve feldolgozási hiba vagy sikeres adatbetöltés esetén milyen e-mailen kér az adatküldő visszaigazolást. Egy adott adatcsomag akkor kerül feldolgozásra, amikor az intézmény megadta az adott adatcsomag ellenőrző kódját, valamint lezárta az adatfeltöltést.

Adatfeltöltés lezárását követően a Központi Rendszer is elvégzi a beküldött adatokon ugyanazon algoritmus alapján az ellenőrző összeg számítását, majd összeveti azt a beküldöttel. Amennyiben eltérés van a két ellenőrző összeg között, hibajelentést küld a megadott e-mail címre, majd a beküldött adatokból generált CSV-t az adott intézmény felé publikálja.

Adatfeldolgozás alatt nem tölthető újabb adatcsomag ugyanazon intézményi forrásból.

Adatok publikálása

Az intézmény a Központi Adattárból kétféle típusú adatot kaphat. Egyrészt a rá vonatkozó intézményi adatokat, másrészt általános, összehasonlító adatokat. Az adatokat a központi rendszer minden esetben automatikusan küldi, ennek kérésére nincs szükség. A kommunikáció ilyen szempontból egyirányú (PUSH). Az adatokat egy, az intézmények által olvasható MQ Q-ba teszi a rendszer, ahonnan az intézmény bármikor elviheti.

Támogatási tevékenység

Támogató szervezet

Az interfész használatát az Educatio Nonprofit Kft. informatikai igazgatósága támogatja. Az interfésszel kapcsolatos bármilyen problémákkal az alábbi elérhetőségen kereshetik a kollégákat:

email cím: avir-techsupport@educatio.hu

Teszt csatornák

Az éles adatküldés előtt az intézménynek tesztelnie kell a kommunikációt. Erre a célra az Educatio biztosít teszt rendszert, külön MQ csatornákkal. Az itt elküldött csomagok feldolgozása ugyanúgy megy végbe, mint az éles rendszerben, tehát az adatküldő pontos visszajelzéseket kap arról, hogy a beküldött adatok feldolgozása sikeres volt-e. Az éles és a teszt rendszerek közti különbségek:

- a teszt rendszerbe bedolgozott adatok nem kérdezhetők le publikus felületen
- a teszt adatküldésre nem vonatkoznak az adatküldési határidők (bármikor, bármilyen csomag küldhető)
- a teszt adatbázist az Educatio szükség esetén kiüríti, így a korábbi csomagokkal történő hivatkozások nem biztosítottak (minden küldésnél szükséges elküldeni a hivatkozott adatokat is)

Utóbbi kitétel a gyakorlatban azt jelenti, hogy pl. a helyiség tábla küldésénél nem számíthat arra a tesztelő, hogy a korábban elküldött ingatlan adatok még léteznek az adattárban. A konzisztencia vizsgálaton egy ilyen esetben hiba keletkezhet. Az Educatio azt biztosítja, hogy adatbázis törlést nem hajt végre nap közben, tehát az egy munkanapon belüli adatküldéseknél a korábbi adatok garantáltan rendelkezésre fognak állni. Ennek megfelelően a teszt csomagokat úgy kell felépíteni, hogy az adott csomag tartalmazzon minden hivatkozott rekordot is.

Az adatbázis törlésekről az Educatio legalább 24 órával a törlés előtt, és a törlés után is értesítést küld az avir-techinfo@educatio.hu levelező listára. Kérjük, hogy a tesztelést végző kollégák erre iratkozzanak fel.

Monitoring felület

Az adatcsere figyelését üzemeltetői riportok segítik, melyet az intézményi felhasználók is megtekinthetnek a központi VIR portálon. Ezek tartalma:

- adatcsomagok küldésének és feldolgozásának tervezett és tényleges időpontja
- feldolgozások naplója, hibák leírása
- a központi adattár által publikált információk leírása

- aktuális interfész specifikáció

Mellékletek

| Fájl neve | Tartalom |
|-------------------|-----------------------------------|
| INT interfesz.xls | Interfész táblák és mezők leírása |