

Kutatási terv

1 A kutatás nemzetközi és hazai előzményei

A kitűzött kutatási célok egy **20 éves** hazai és nemzetközi kooperáció-sorozat **szervesen következő lépését** jelentik. A várható eredmények (elemző robotok) példaképei, s egyben a **kivitelezhetőség referenciái/garanciái: a sakk-automaták és a robot-pilóták/szimulátorok**. A kutatás **nem csak** ezen lehetőségek gazdasági-társadalmi **kiterjesztését/generalizálását** (vö. GPS = general problem solving), hanem **módszertani, technológiai meghaladását** is célozza (vö. konzisztencia-, interpretáció-generálás, osztott rendszerek). Mindezt úgy, hogy az **OTKA** (mint egyik pályázati kiírásért felelős szervezet) által eddig támogatott kutatásokra támaszkodva (**F030664: <http://miau.gau.hu/miau/55/otkaz.doc>, ill. T049013: <http://miau.gau.hu/myx-free/index.php3?x=otka>**), valamint azt **NKTH** (mint másik pályázati kiírásért felelős szervezet) belső (vö. **RIIR/RIIRCORE**) és publikus elvárásait igyekszik maximálisan lefedni a tényalapú szakpolitizálás módszertani és információ-technológiai megalapozása révén (vö.):

- http://itbusiness.hu/felso_menu/magazin/hirhatter/Innovacios_tudasbazis.html, ill.
- <http://www.nkth.gov.hu/portalforum/innovacio-naprakesz/kutatas-fejlesztes> .

1.1 Nemzetközi együttműködésben végzett kutatás

A sok, egymásra épülő modulból álló, stratégia-alapozó modellezés egyik legjelentősebb nemzetközi képviselői az **agrár-szektormodellek**. 1995 és 2007 között számos nemzetközi projektben való részvétel (EU5, EU6-keretprogram: SPEL(GR): <http://miau.gau.hu/miau2009/index.php3?x=e0&string=MSZR>, PIT, IDARA: <http://miau.gau.hu/miau2009/index.php3?x=e0&string=IDARA>, CAPRI), ill. egy tananyagfejlesztést célzó 2 hónapos DAAD ösztöndíj a know how forrásoként funkcionáló, 1980 óta EU-projektet vezető bonni Agrárpolitikai Intézetben (2000, <http://miau.gau.hu/miau/34/aszm3.doc>), valamint egy kínai vendégkutatóval való kooperáció ezen területen (2000/2001) a lehetőségek tárházát jelentette a kulisszák mögé való bepillantásra, a gyenge pontok feltárására, korrekációs javaslatok megfogalmazására (pl. <http://miau.gau.hu/miau/53/autotrend.doc>, http://miau.gau.hu/miau/60/idara_draft.doc).

Bár az EU-csatlakozást követő időszakban (a kihívások csökkenésével) a Bonn-Gödöllő kapcsolatok intenzitása is csökkent, a **Regionet Europe EEIG/EWIV** (német, osztrák, észt, magyar közreműködéssel) 2008-ban alakított szervezet (vö. támogató nyilatkozat) új nemzetközi keretet kínál a stratégiai tervezés vidékfejlesztési célú fejlesztésére: pl. <http://miau.gau.hu/miau2009/index.php3?x=e0&string=Seddin>. Ezen folyamat része a Duna-Ipoly (**DIPO**) vidékfejlesztési stratégia tervezés-módszertani monitoringjának (pl. <http://miau.gau.hu/miau2009/index.php3?x=e0&string=124>) ad hoc felülvizsgálata, mely a pályázat pilot-nézeteként is felfogható (vö. **GIK Kft.**, támogató nyilatkozat).

Emellett a giesseni egyetemmel (**JLU**, Németország) fennálló tradicionális (25 évre visszanyúló) kutatási kapcsolat folyamatos alapot adott egy 1989-es diplomatervtől, az 1990-1993-as DAAD doktori ösztöndíjon, ill. az 1995-os TEMPUS és OTKA támogatáson keresztül arra, hogy ma az agrár-szaktanácsadás megújításának bilaterális előkészítési folyamatában (vö. http://miau.gau.hu/miau/117/aik2008_full.rtf) a 10 legjobb agrárinformatikus (ISZAM BSC) hallgatót (akik a projekt demonstrátor-szintű HR-erőforrását adják, s jelenleg a Regionet Europe EEIG/EWIV gyakornokai) rövid tanulmányút keretében fogadták a potenciális best practice-t képviselő német tanácsadó szervezetek (<http://miau.gau.hu/myx-free/index.php3?x=robots>).

1.2 Hazai kooperációban végzett kutatás

Az EU agrár-szektormodellezés tapasztalatai a hazai, elemzési célú agrárinformatika megreformálására tett kísérletek sorozatát hozták magukkal: **MIMIR-MITIR-MIVIR** (1998-), vagyis a Magyarországi Integrált Mezőgazdasági Információs Rendszer (<http://miau.gau.hu/miau2009/index.php3?x=e0&string=MIMIR>) koncepciójából kiindulva 2002-ben a jelenleg is csatlakozó kutatók egy részével közösen készült az **IIR** (Integrált Igazgatási és Ellenőrzési Rendszer) megvalósíthatósági tanulmánya, ill. ennek részleges megvalósítása a Hewlett Packard konzorcialis vezetésével. Ennek kapcsán a Magyar Turizmus Rt-vel együttműködésben született meg (2002) az **ONTIR** (online turisztikai információs rendszer) terve, ill. már 2000-ben a TeleDataCast Kft.-vel közösen az iKTAbu (online elemzési szolgáltatások) pilot-változata (<http://miau.gau.hu/miau2009/index.php3?x=e0&string=iktabu>). A folyamatot az NKTH számára a GIK. Kft. vezetésével készített **RIIR** (Regionális Innovációs Információs Rendszer) terve vitte tovább, s az ennek modulszerű megvalósítását jelentette az **INNOCSEKK** 156/2006-os **MY-X** projekt (online egyedi hasonlóságelemzési elemzések támogatása: <http://miau.gau.hu/miau2009/index.php3?x=e0&string=myx>), ill. a Baross Gábor Programban konzorcium keretében kialakított **RIIRCORE** szolgáltatás (<http://www.innoreg.hu/>). Ezen, s a publikációs listákban feltüntetett további projektek **kellő tapasztalatot**

adnak a jelen projektben is releváns nagy adatmennyiségek manipulálására, ill. a komplex rendszerek tervezésére fejlesztésére, valamint az online elemző szolgáltatások kialakítására.

2 Az alap kutatás tulajdonságai

Bár az alap kutatás alapvetően interdiszciplináris jellegű, mégis egyszerű (egymásra épülő és egymást feltételező) súlypontokkal jellemezhető:

- a közhasznú adatvagyon-gazdálkodás megreformálása (vö. OLAP = on-line analytical processing – vö. <http://miau.gau.hu/myx-free/index.php3?x=i06>)
- az eddigi (= best practice) ad hoc **elemzési folyamatok** strukturálása (<http://miau.gau.hu/myx-free/index.php3?x=i0100>), majd
- erre alapozó **automatizálása** (pl. http://miau.gau.hu/myx-free/files/myx_fch1hu.png),
- részfolyamatok eredményeinek **automatizált ellentmondás-mentesítése** (<http://miau.gau.hu/miau/111/chf30.doc>), ill.
- a számítások **szöveges értelmezésének automatizálása** (http://miau.gau.hu/myx-free/index_e9.php3?x=e09).

2.1 Jellege

Az alap kutatás kulcsa a jelenleg ismert ad hoc szakértői véleményképzés (vö. sakk-nagymester) problémamegoldó folyamatainak automatizálása. Az így létrejövő **virtuális robotok**, vagyis a mindenkori döntéshozók **műszerfalát jelentő egyensúly-detektorok, fenntarthatóság-műszerek**, azaz egy fajta optimalizált **Balanced Score Card** logika, a szükséges és folyamatosan bővülő (pl. KSH gyorsjelentések saját katalógizálásban:

http://miau.gau.hu/osiris/content/elib/subgroups.php3?focsopid=stat&csopid=ksh_stat&order=i)

adatvagyonon az **egyszer már sikeres, iteratív módon finomított, kombinatorikailag teljes körűvé tett elemzési lépések automatikus futtatását** végzik lehetőség szerint **online**.

A statikusság veszélyének feloldását a saját fejlesztésű **hasonlóságelemzési** modulok biztosítják, melyek a jól ismert, de eddig **egységes szakértői rendszerbe** nem szervezett **matematikai-statisztikai apparátus** számára folyamatosan **új bizonyítási kihívásokat generálnak** az egyre bővülő adatvagyon tanulási mintaként való értelmezésének tetszőleges logikai rétegekre való bontása révén (pl. <http://miau.gau.hu/miau/111/chf30.doc>). A MY-X és a RIIRCORE projekt keretében az **egyedi, kézi vezérlésre tervezett online elemző alkalmazások forráskód szinten a projekt rendelkezésére állnak**.

2.2 Hasznosíthatósága

Az automatizálás ma az ipari hatékonyság alapja. Az elemzés „iparszerűvé” tétele a problémamegoldás/döntéstámogatás (stratégiai tervezés) hatékonyságát kívánja fokozni. Az automatizálás keretében nem csak a már ismert matematikai-statisztikai elemzési lépések kombinatorikailag zárt (azaz minden létező helyzetre felkészített, de mégis statikus) megközelítései termelnek hasznot, hanem magának a tény-alapú mintázat (=tudás) generálásnak **több-rétegű és ellentmondás-mentességre törekvő** automatizálása is, melyet az online/optimalizáló hasonlóságelemzés tesz lehetővé. A hasonlóságelemzés szoros rokonságban áll a már ismert módszerekkel (pl. neurális hálózatok, döntési fák, cluster-elemzések, több-változós modellek), de egyszerűsége és rugalmassága révén quasi korlátlan alkalmazhatóság jellemzi (<http://miau.gau.hu/myx-free/>).

2.3 Eredetisége

Az automatizálás maga látszólag triviális kihívás, csak következetes tudomásul vétele jelent elvileg eredetiséget. A valódi, holisztikus szemléletű know-how egyrészt az **n-rétegű elemzési lépések ellentmondás-mentesítő logikájában** rejlik. Vagyis abban, hogy **nem kell** (s lényegében nem is lehet) **tesztelésre pazarolni tényadatokat**. A klasszikus beválási gyakoriság alapon modell-értékelést végző lépések helyett minden tényadat sok-rétegű (sok részkérdést megválaszoló) megtanulása alapján mindenkor azonnal az éles probléma legvalószínűbb (leginkább ellentmondásmentes) megoldásának megadása az azonnali elemzési cél. Ez a megközelítés biztosítja szinte tetszőleges, eddig kezelhetetlennek tűnő problémák (pl. hosszabb távú meteorológiai, ill. árelőrejelzések) mintázatszerű megértését, **s egyben a 'nem tudom' rendszerválasz automatikus keletkeztetését**. Másrészt a létrejövő hibrid szakértői rendszerek **az attribútumonkénti optimális felbontás mellett saját határaik automatikus felismerését is megtanulják** a rendelkezésre bocsátott tényadatok alapján. A numerikus jellegű modulokat pedig **szövegsablon-**

hozzárendelést végző, vagyis az értelmezést részlegesen automatizálni engedő szakértői rendszerréteg formálja a döntéshozók számára könnyen fogyasztható jelentésekké.

2.4 *Várható elméleti és gyakorlati jelentősége*

Elméleti szinten a **tudásgenerálás, a bizonyítás és értelmezés automatizálásának tervezése** jelent előrelépést. Emellett az alkalmazók számára a **tény-alapú stratégia-alkotás**, vagyis az adott helyzetre ható szívó- és nyomóerőterek **automatikus detektálása** bír jelentőséggel. A feldolgozandó adatvagyon (mely korlátozott mértékben **akár hiányos/hibás is lehet**) kijelölése után a projekt eredményeként előálló elemző és önvezérlő algoritmus-tömeg szöveges jelentés formájában lesz képes felvázolni a vizsgált rendszerben fellelhető erőtereket **indirekt előrejelzéseként**. **Direkt előrejelzések** esetében a rendelkezésre álló adatvagyon alapján leginkább védhető változási irányok és mértékek levezetése történik meg. Amennyiben a feldolgozott jelek ellentmondásossága jelentős, abban az esetben a rendszer nem vállalja fel semmilyen következtetés levonását. **A stratégia-alkotás támogatása abban is megnyilvánul, hogy az így generált tudás képes kilépni az ismert rendszer-határokon (szélső értékeken), vagyis képes a genetikai potenciál fogalmát és tetszőleges inputkombinációk hatását levezetni.**

Gyakorlati szinten maga a megvalósulás, az **egy-gépes és osztott (párhuzamosított) kivitelezés** tekinthető kézzel fogható eredménynek. A megvalósulás rétegei: adatvagyon-kezelésben az **OLAP**-jellegű adatvagyontámogatás; tervezési szinten a **kombinatorikailag zárt bizonyítási módszertan**, vagyis bármilyen adatvagyon esetén előírt próbák és előfeltételek rendszerének egységes kezelése; programozási szinten a jelenleg jórészt manuálisan egymás után illesztett részelemzések egységes, önvezérlő folyamatként való leképezése; vezérlési szinten az adatbázis-műveletek és az **algoritmusok időtakarékoságának** (vö. real time jelleg) biztosítása.

2.5 *Az elvégzendő vizsgálatok újszerűsége*

Egyrészt az adatvagyonok megfelelő hatékonyságú manipulálása ismert technológiák (OLAP) adaptálást, **továbbfejlesztését** tételezi fel. Azonban a jelenlegi közhasznú adatkezelés zavarait látva ezen adaptációk már önmagukban is **jelentős közösségi hasznot** engednek feltételezni (mely hasznosság mértékét egy, a **projekt indulása előtt záródó PhD-kutatás** tárja fel: vö. Pető, KIR = Külső Információs Rendszerek elemzése, mely téma a már sikeresen **zárt PhD-folyamatra** támaszkodik, vö. Szalay, 2009 - ERP-elemzések).

Másrészt a kombinatorikai **teljeskörűség szavatolása** jelenleg nem ismert gyakorlat az elemzési lépések és előfeltételei kusza szövedékében (vö. SPSS tools). Vagyis a jelenlegi ad hoc elemzések alapvető oka az, hogy egyelőre nincs meg az egyes matematikai-statisztikai eljárások (pl. próbák) alkalmazási előfeltételeinek egységes ellenőrző mechanizmusa.

A konzisztens rendszerértelmezést, egyensúlyvesztést, előrejelzést önmagában is újszerű hasonlóságelemzési lépés-láncolatok biztosítják, melyek tetszőleges pontokon ötvözhetők (hibridizálhatók) közismert mesterséges intelligencia modulokkal (pl. döntési fákkal).

Ismét csak kombinatorikailag teljeskörű megoldást kell, hogy kínáljon az ellentmondás-mentesség modellezése és a számszerű eredmények szövegpaneles értelmezése.

3 A pályázatban résztvevő kutatók és segédszemélyzet

3.1 *Létszám*

A projekt a SZIE Gazdaság és Társadalomtudományi Karának **több intézetéből** (pl. Gazdaságelemzési és Módszertani, Informatikai, Regionális gazdaságtani, Vállalat-gazdaságtani) érkező módszertani és elemzési szakértőket (inkl. zárás előtt álló és folyamatos kutatást végző PhD-hallgatók) szervez egységes rendszerbe. Emellett a SZIE **teljes információtechnológiai háttere** (hálózata és üzemeltetési szakértőgárdája) adott a projekthez. A piramis fontos rétege a széles, korábbi TDK és szakdolgozatok kapcsán már bizonyított, leginkább kreatív és **innovatív (levelezős, nappalos, egyetem- és karközi kooperációban dolgozó) hallgatói csapat**, melyen belül a speciális a projekt elvárásai szerint képzett agrárinformatikus csoport a hallgatók közötti belső hierarchia biztosítását is jelenti. **Külső támogatóként** egy személyben az APEH SZTADI, a Controllingportál és a Raabe Kiadó szakértője csatlakozott a folyamathoz. Emellett a korábbi projektpartnerek közül a RIIRCORE konzorcium vezetője (GIK Kft.) biztosít szakmai és infrastrukturális támogatást. Nemzetközi szinten a JLU Giessen párhuzamos kutatása (mely célja a tanácsadási folyamatok automatizálásának megalapozása a vállalkozás szintjén) és az önmagában is több EU-tagországok tömörítő Regionet Europe EEIG/EWIV ad keretet a fejlesztések irányának pontosításához és az eredmények

teszteléséhez. **A speciális adatvagyon-kezelési és programozási feladatok kapcsán ad hoc jelleggel külső (hazai) szakértők bevonására mód lesz.**

3.2 *A tervezett munkaidő ráfordítás arányossága, időmérleg*

Az informális csatornákon beszerzett sarokszámok alapján az egy FTE-re jutó kutatási összeg relatíve előnyös jelen projekt esetében. Az adatvagyonok megfelelő szintű minőségének biztosítása, ill. az állandó frissítés szinte korlátlan emberi erőforrás lekötését jelenti, melyből természetesen jelen projekt keretében csak a marginális elveknek megfelelő szintű minőségbiztosítás szavatolása a cél, vagyis a mindenkori utolsó egységet képviselő **adatmanipulációs feladatok költsége nem haladhatja meg az ettől elvárt pontosság/hitelesség-növekedés piaci értékét.**

A projekt munkaszervezete rel. **lapos**. A projektvezetés közvetlenül támaszkodik a két új kutatóra (tudásmérnökre), akik egyike az adatvagyon, míg a másik az automatizációt felügyeli. A projekt terhére kifizetéseket nem kapó SZIE kutatók saját kutatási feladataik átvilágítását biztosítják, mely egyszerre iránymutató és ellenőrző hatású. A jelentős hallgatói kapacitás elsődlegesen az adatvagyon-kezelést támogatja, míg a külső szakértők a programozás/adatbázis-kezelés speciális problémáinak megoldását szállítják. A külső szakmai partnerek az eredmények értelmezésén keresztül ismét csak minőségbiztosítási hatást fejtenek ki. A partnerek által rendelkezésre bocsátott korábbi projekteredmények (forráskód könyvtárak) alapján a tervezett továbbfejlesztés sikere reális.

3.3 *Szükséges infrastruktúra*

A hálózat, számítókapacitás, programozási környezet (Linux, PHP, PostgreSQL, Javascript, HTML, XML, CSS/HTML-Validators, ill. Windows, VisualBasic) elsődlegesen adott intézményi (SZIE) és partneri (GIK) biztosítékok alapján. A projekt ideje alatt erkölcsi avulás, ill. a mobil-elemző csoport kialakítása a betervezett notebook-állományt szükségszerűvé teszi. **A pályázó a SZIE GTK TATA Kiválósági Központ Informatika Intézet (Információtechnológia Tanszék vezetője), mely intézet jelenleg költözött új telephelyre, így a teljes infrastruktúra új, korszerű.** Az adatvagyonok tekintetében elsődlegesen a **publikus adatvagyonok feldolgozása a cél, vagyis az adatvédelmi problémák eleve elkerülhetők. Az adatvásárlás alatt tehát nem primer adat-felvételezések ellenértékét, hanem a létező adatvagyonok racionális tovább-feldolgozásának (migrációjának, leválogatásának) megalapozása érdekében jogos adatgazdák által megállapított díjakat kell érteni.** Az adatok ellenértéke nem szabályozott, a legutolsó meteorológiai adatsomagban egy rekord mintegy 100Ft/egység értékkel bírt. Jelen projektben ennél **nagyságrendileg olcsóbb** konstrukciók elérése a cél. Az eddigi tapasztalatok alapján 65.000 rekordnyi (kb. 1 Excel-táblázatnyi) adat szakszerű kezelése átlagosan 1 emberhónap hallgatói munkaigényű.

4 **Kockázatkezelés**

A közhasznú adatvagyonok-fragmentumok **viszonylagos ellenőrizetlensége, belső ellentmondásai** jelentős kockázatot jelentenének akkor, ha a hallgatói és kutatói létszám, ill. a projekt belső **minőségbiztosítása** (mely önmagában is konzisztenciára törekvő) nem állna szemben a tény-alapúságot zavaró hatásokkal. A sok résztvevőt mozgató projektek kockázata ezen HR-kapacitások esetleges **migrációja**, melyet azonban a 2 új kutató betervezése és a 24 hónap alatti új hallgatói kapacitások **potenciálja** elviselhető szintre szorít vissza. A projektnek **infrastrukturális és jogi kockázata nincs**. A siker alapja az előzményekből közvetlenül levezethető célokban és a **racionális munkaszervezésben** rejlik. A kutatás kockázatait az adatvagyonok minőségbiztosítása, az elemzési kombinatorikai terek méretének nagysága, ill. a számítások valós idejűségének biztosítása mellett (vö. további innovációs kihívások) magának a helyes elemzési irányoknak, vagyis az elemzések **értelmezhetőségének** kérdése jelenti: adatoldalról mindenkor abból **kell és lehet** kiindulni, **ami adott**, hiszen abban a pillanatban, ha a projekteredmények hasznosításra kerülnek, az adatvagyon minősége semmikor nem lesz jobb. Az **elemzések kombinatorikai tere szakértői szinten tetszőlegesen korlátozható**, hiszen a **valós alkalmazások esetén a valósidejűség mindenkor fontosabb lesz, mint az elemzési minőség**. Tehát a számításmenetek (vagyis az ellentmondások minimalizálást lehetővé tevő) n-rétegűség során az 'n' értéke szintén emberi döntés kérdése. Végül a helyes elemzési célok, s így a potenciális szövegpanelek megtalálását az **interdiszciplináris** kutatói/partneri kör szavatolja. Az együttműködő kutatók eddigi **publikációs igazolják a felkészültséget**. A citációk nagyságrendje és ezek impakt faktora az interdiszciplináris témákat kedvelő szaklapok nagyon alacsony száma és a szakterület gyors fejlődése miatt **semmiképpen nem** értelmezhető projektkockázatként. **A projektnek kommunikációs kockázata nincs**: a partneri kör több induló, színvonalas online szaklapon keresztül (pl. KOINE, Vidék

Hangja, MIAU) minden erre érdemes eredményt gyorsan el tud juttatni a célközönséghez a konferenciák és egyéb szerkesztőségek rel. lassú publikációs stratégiáival szemben.

5 Alkalmazók, alkalmazások, jövőkép

A magyar **EU-elnökség**, ill. az ezzel szoros kapcsolatban lévő kormányprogramok módszertani támogatását központi célként kitűzve a projekt a rendelkezésre álló (tovább-hasznosításra átstrukturált) adatvagyonokból már a tervezési (tesztadat-gyártási) fázisban is értékelhető eredményeket kell, hogy produkáljon a potenciális egyensúlyvesztésekből következő stratégiai irányváltásokra vonatkozóan. A pályázat beadásának pillanatában ilyen országos kérdések pl. valóban az adórendszer, a szociális támogatások, stb. jelenleg kiemelt pontjai esetében vezethető-e le a legnagyobb külső (objektív) nyomás ezek változására? EU-szinten a válságkezelés stratégiai irányainak verifikálása hasonló elemzési kihívásként jelentkezik...

A projekt elsődleges alkalmazása kapcsán a politikai érintettség nehezen tagadható, lévén társadalmi innovációról van benne szó. Így ideális esetben az **OSAP** (Országos Statisztikai Adatgyűjtési Program) keretében (törvényi szinten) szavatolható a statisztikai adatvagyonok elő-elemzése. Amennyiben a mindenkori legfelső döntéshozói szint nem kíván élni a rendszerszemléletű vezérlés-támogatással, úgy az alkalmazás delegálódik regionális, vállalati, személyi (civil és érdekvédelmi) szintre, mely ismét csak nem közvetlenül a klasszikus üzleti modelleket jelent. A stratégiai tervezés támogatása egyedül a (KKV) vállalkozások szintjén tekinthető klasszikus üzletviteli tanácsadásnak.

A jelenleg futó, automatizációs célokat nélkülöző, de ezt megalapozó DIPO-elemzések a projekt támogatása kapcsán remélt pozitív döntésig stabilan megalapozzák a nemzetközi szintű kiterjesztést...

Research Plan

1 The national and international previous history of the research

The declared research goals are the next **essential step** of a **20** year old national and international cooperation-sequence. The role models of the expected results (analytical robots), and the **references/guarantees** that they can be carried out: the **chess automatons** and **robot pilots/simulators**. The aim of the research is **not only** the socio-economic **extension/generalization** (cf. GPS = general problem solving) of these possibilities, but to **exceed the methodology and technology** of them (cf. consistency-, interpretation-generation, shared systems). Doing it so, that the researches supported by the **OTKA** (as one of the organizations responsible for the tender calling) before (**F030664**: <http://miau.gau.hu/miau/55/otkaz.doc> and **T049013**: <http://miau.gau.hu/myx-free/index.php3?x=otka>) and the NKTH's (as the other organization responsible for tender calling) internal (cf. **RIIR/RIIRCORE**) and public expectations are striven to be maximally covered by the methodological and IT foundation of fact-based politics (cf.):

- http://itbusiness.hu/felso_menu/magazin/hirhatter/Innovacios_tudasbazis.html
- <http://www.nkth.gov.hu/portalforum/innovacio-naprakesz/kutatas-fejlesztes>

1.1 Research done in international cooperation

The most significant international representatives of strategy-planning models are the **agricultural-sector models** that consist of many modules that are built on each other. Between 1995 and 2007 a huge volume of workload in international projects (EU5/EU6-framework program: SPEL(GR): <http://miau.gau.hu/miau2009/index.php3?x=e0&string=MSZR>, PIT, IDARA: <http://miau.gau.hu/miau2009/index.php3?x=e0&string=IDARA>, CAPRI), and a DAAD-fellowship (2000, <http://miau.gau.hu/miau/34/aszm3.doc>) with a duration of 2 months (aiming interpretation of EU-literature for the Hungarian education) in the Institute of Agricultural Policy Bonn being responsible for know-how of sector modeling, and the co-operation on the same field with a guest researcher from China (2000/2001) made possible to detect, what kind of logic do work behind the common surface of science, and which points can be identified as weakness, and how can they be corrected (cf. <http://miau.gau.hu/miau/53/autotrend.doc>, http://miau.gau.hu/miau/60/idara_draft.doc).

Although the intensity of the Bonn-Gödöllő relationship diminished after the EU-accession due to the reducing topical challenges, the organization **Regionet Europe EEIG/EWIV** (cf. Letter of intent) that was founded (with German, Austrian, Estonian and Hungarian co-operation) in 2008 offers a new international framework for the development of strategic planning in rural development (e.g. <http://miau.gau.hu/miau2009/index.php3?x=e0&string=Seddin>). The ad hoc controlling of the monitoring on the field of strategic planning methodology for rural development (cf. <http://miau.gau.hu/miau2009/index.php3?x=e0&string=124>) in the Duna-Ipoly (**DIPO**) region can be seen as a part of this framework and also as a pilot-view of this NKTH-OTKA-project plan (cf. Letter of intent from **GIK** Ltd.). The traditional research co-operation (25 years) between the University of Giessen (**JLU**) in Germany and the University of Gödöllő (**SZIE**) made possible (step by step: 1989: thesis study, 1990-1993: PhD-research with DAAD-fellowship, co-operation in frame of TEMPUS/OTKA fellowship) to realize a study trip for the 10 VIP-students (providing HR-capacities for this proposal on the leading level among students, and are the trainees of the Regionet Europe EEIG/EWIV) from Gödöllő, aiming a bilateral preparation of reform in the agricultural extension system (http://miau.gau.hu/miau/117/aik2008_full.rtf). The short study trip enabled consultations with representatives of German institutions providing best practices on the field of agricultural consulting processes (<http://miau.gau.hu/myx-free/index.php3?x=robots>).

1.2 Research done in national co-operation

The agricultural-sector modeling experiences of the EU led to a series of experiments aiming to reform the domestic, analytical agricultural informatics: starting from the concept of **MIMIR-MITIR-MIVIR** (1998-), viz. the Hungarian Integrated Agricultural Information Systems (<http://miau.gau.hu/miau2009/index.php3?x=e0&string=MIMIR>). In 2002, the feasibility study of the **IACS** (Integrated Administration and Controlling System) was made jointly with part of the even now applying researchers, and it was partially worked out with the consortial leadership of Hewlett Packard. In addition to that, the **ONTIR** (online tourism information system) was made (2002) in cooperation with the Magyar Turizmus Zrt.I In 2000, the pilot version of **ikTAbu** (online analytical services) was made in co-operation

with TeleDataCast Ltd (<http://miau.gau.hu/miau2009/index.php3?x=e0&string=iktabu>). The process was furthered by the plan of **RIIR** (Regional Innovation Information System) that was made by GIK Ltd. for the NKTH. The **MY-X** (support of online individual similarity analyses: <http://miau.gau.hu/miau2009/index.php3?x=e0&string=myx>) project (cf. **INNOCSEKK 156/2006**) and the **RIIRCORE** service (<http://www.innoreg.hu/>) created within the frame of the Baross Gábor Program were its modular realization. These, and other projects mentioned in the publication lists **mean enough experience for the manipulation of large databases that are relevant in this project too, and for the planning and development of complex (large scale) systems, and for the creating online analytical services.**

2 The characteristics of the research

Although the research is fundamentally interdisciplinary-natured, it can still be described with simple weights (that are based on each other, and assume the existence of each other):

- the reformation of data assets management for public utility (cf. OLAP = on-line analytical processing – cf. <http://miau.gau.hu/myx-free/index.php3?x=i06>)
- the structuring of the until now (best practice) ad-hoc **analytical processes** (<http://miau.gau.hu/myx-free/index.php3?x=i0100>), then
- **automation** based on it (http://miau.gau.hu/myx-free/files/myx_fchl1en.png),
- making the results of partial processes **consistent automatically** (<http://miau.gau.hu/miau/111/chf30.doc>), and
- **Automation of the textual interpretations** of the calculations (http://miau.gau.hu/myx-free/index_e9.php3?x=e09).

2.1 Features

The key of the research is the automation of the nowadays ad hoc problem-solving processes by human experts (cf. chess grandmaster). The **virtual robots** are going to be (viz. the **control panel** that indicate the **balance-detectors, sustainability instruments** of respective decision-makers, that means some sort of an optimized Balanced Score Card logic) made possible based on the necessary and continuously expanded data assets (e.g. draft reports from CSO offered through their own catalogues: http://miau.gau.hu/osiris/content/elib/subgroups.php3?focsopid=stat&csopid=ksh_stat&order=i), the **automated running of analytical steps** (as soon as realistic: **online**) having characteristics like **proven successful, iteratively fine tuned, combinatorially closed.**

Modules developed by ourselves (based on similarity analysis) ensure the elimination of the danger being static. These modules always generate new challenges of verification for the nowadays' well-known mathematical-statistical methods without a consistent frame from expert system logic. This kind of the multi-layered verification can be provided if the data assets (growing continuously) are interpreted as a series of partial learning patterns (e.g. <http://miau.gau.hu/miau/111/chf30.doc>). The source codes of the MY-X and the RIIRCORE services (designed for individual and manually operated support processes) are available for implementation in the project.

2.2 Usefulness

Automation means the basis of industrial efficiency nowadays. The “industrialization” of analysis is intended to enhance the efficiency of problem solving/decision making (strategic planning). In the frame of automation, not only the combinatorially closed approaches (that are prepared to be able to deal with each possible input situations, but remains still static) of the well-known mathematical-statistical analytical steps produce profit, but the automation of **multi-layered and consistency-aspiring** generation of fact-based patterns (=knowledge), that is enabled by the online/optimizing similarity analysis. The similarity analysis is in close relation with the already known methods (e.g. neural networks, decision trees, cluster analyses, multi-variable models), but it is characterized by quasi indefinite applicability, because of its simplicity and flexibility (<http://miau.gau.hu/myx-free/>).

2.3 Originality

Automation itself may seem to be a trivial challenge, only its coherent approval means originality.

The real, holistic approach of know-how lies in the **inconsistency-aspiring logic of the n-layered analytical steps.** That means that it is **unnecessary** (and it is essentially impossible) to **waste facts for test**

calculations. Instead of the classic model-evaluation logic where the hit rating of models are measured, the goal of the multi-layered analysis is to give the most probable solution for the problem at hand, upon the parallel learning of each fact. This approach ensures the pattern-like comprehension of quasi arbitrary problems that seemed to be impossible to solve until now (e.g. long-term weather and price forecasts), **and the automatic creation of the ‘I don’t know’ system answer too.** On the other hand, these hybrid expert systems that come into existence **learn the automatic recognition of their own limits besides the optimal resolution per attribute,** based on the facts submitted to them. Modules of numeric nature can be formed into reports, which are comprehensible by decision-makers with ease, by a **template-assigning expert system layer that partially allows automation of interpretation.**

2.4 Expected theoretical and practical significance

On theoretical level, **knowledge-generation, the planning of automation of verification and interpretation** means advancement. Besides that, for the appliers, the fact-based strategy-making, viz. the automated **detection** of pushing and pulling **force fields** that apply to certain situations may be significant. After assigning the data assets (that can even be either **incomplete or erroneous** to some extent) that are going to be processed, the analytic and self-monitoring mass of algorithms will be able sketch in the force fields found in the examined system as an **indirect forecast,** in the form textual report. In case of **direct forecasts,** the most defensible directions and proportions of changes will be derived from the available data assets. If the processed data is highly inconsistent, then the system won’t draw any conclusion from them. **The support of strategy-making manifest itself in the ability that the knowledge generated this way can exceed the known system-limits (extreme values), so it is able to derive the definition of genetic potential and the effect of optional input combinations.** On practical level the materialization itself, the **single or shared (paralleled) computing** may be regarded as a tangible outcome. The layers of materialization: in data assets management, the **OLAP-like** data assets support, on design level the **combinatorial closed verification methodology,** viz. the unified handling of a predetermined system for proving of statistical hypotheses and their preconditions in case of any kind of data assets, on programming level, the derivation of the now mostly manually superimposed partial analyses as a unified, self-monitoring process, on control level, the provision of the **time-saving of the algorithms and database-operations.**

2.5 The innovation in the examinations

On the one hand the efficient manipulation of data assets assumes the adaptation, **improvement** of already known technologies (OLAP). However, upon seeing the disorder in current public data management, these adaptations assume **significant public utilities** themselves (and the degree of this utility is revealed by a **PhD-research that might become finalized before the start of the project:** cf. Pető, KIR = analysis of External Information Systems, this topic relies on a **previous successful PhD-thesis,** cf. Szalay, 2009 – ERP-analyses). On the other hand, the **assurance of combinatorial holistics** is currently an unknown practice in the confused web of the preconditions and steps of analysis (cf. SPSS tools). So, the reason for the currently ad-hoc nature of analyses is that, for the time being, there is no unified pre-control mechanism for the application preconditions of the different mathematical-statistical procedures (e.g. tests). **The consistent system interpretation, balance loss, forecast are ensured by chaining similarity analysis steps that are innovative themselves,** and they can be merged (hybridized) with common intelligence modules (e.g. decision trees). **Again, the modeling of consistency and the textual interpretation of numeric results shall offer a combinatorially holistic solution.**

3 The researchers and assistants involved in the proposal

3.1 Staff

The project is geared forward by methodological and analytical experts (including PhD-students who are either doing research persistently, or are near graduation) coming from **diverse institutes** (e.g. Institute of Business Analysis and Methodology, Institute of Regional Economics, Institute of Farm Management, Institute of Computer Sciences) of the SZIU’s Faculty of Economics and Social Sciences. Besides that, **the whole IT-background** (incl. network and maintenance experts) of SZIU is given to the project. An important layer of the pyramid is the most creative and **innovative student’s team (who either attends to distance or regular courses, who work in co-operations between universities and faculties),** within the special agricultural informatics group, that is trained to meet the expectations of the project, ensures the internal hierarchy among students.

As external supporter, an expert of the APEH SZTADI, the Controllingportál, and the Raabe publisher joined the research process. Besides that, from the previous project partners, the leader of the RIIRCORE consortium (GIK Ltd.) provides professional and infrastructural support. On international level, the parallel research work of JLU Giessen (that declared the automation of consultancy processes on enterprise level as a goal), and the Regionet Europe EEIG/EWIV, in which many EU member states are involved, gives the frame to specification of the directions of the research and the testing of results. **It will be possible to hire external (domestic) experts for the special data assets management and programming tasks.**

3.2 Proportionality of planned time allocation, time balance

Based on the data obtained from informal channels, the research cost per FTE is relatively advantageous in case of this project. The assurance of the appropriate quality of the data assets, and the constant updates require employing nearly unlimited human resources. The goal is to ensure a quality assurance suitable for the marginal principles, **so the cost of data manipulation tasks mustn't exceed the market value of precision/validity increase.**

Within the staff of the project, there is a rel. **flat** hierarchy. The project management directly relies on the two new researchers (knowledge engineers). One of them coordinates the data assets, while the other coordinates the automation. The SZIE researchers that are not on the project's budget provides the re-examination of their own research tasks, which has directive and controlling effect at the same time. The significant student capacity primarily supports the data assets management, while the external experts are responsible for programming and database management. The external professional partners are responsible for quality assurance through the interpretation of the results. Based on the results of previous project (source code folders) the success of the planned improvement is real.

3.3 Necessary infrastructure

The network, and calculation capacity and the programming environment (Linux, PHP, PostgreSQL, Javascript, HTML, XML, CSS/HTML-Validators and Windows, VisualBasic) is primarily provided by the institution (SZIE) itself and also by the external (GIK) partners. During the course of the project, obsolescence and the formation of the mobile analytical group makes the procurement of the planned notebooks a necessity. **The applicant is the SZIE-GTK TATA Institute (leader of IT department) that recently moved to a new worksite, so the whole infrastructure is new and modern.** In regard of the data assets, the process of **public** data is the primary goal, **so the data security issues can be avoided. So, the data-purchase doesn't mean primary data recording, but it means the price of rational further processing (migration, selection) of already existing data assets.** The offset of data is unregulated. In the latest meteorological data package a record had a cost of 100 HUF/unit. In this project, the goal is the attainment of **much** (in order of magnitude) **cheaper** constructions. Based on previous experience, the professional handling of 65000 data records (approximately, 1 Excel table ver. MS Office 2003) takes 1 month of work of a student.

4 Risk assessment

The **relatively uncontrolled and inconsistent** fragments of public data assets would mean significant risk, if the student and researcher staff, and the internal **quality assurance** (that aspires to consistency itself) of the project wouldn't contrast the effects that hamper the fact-based nature of the project. The risk of projects that mobilize volume large number of participants is that the HR-capacities may **migrate**, but it is reduced to an endurable level by the planned employment of 2 new researchers and the potential of new student capacities during the course of the 24 months. The project has no infrastructural and legal risks. The key of success lies in the goals that can be directly derived from the premises, and the **rational management of work**. The risks of the research are the following: quality assurance of the data assets, the size of analytical combinatorial spaces, and besides the assurance of real-time calculations (cf. further innovational challenges) the correct analytical directions themselves viz. the **interpretability** of the analyses: in regard of data, **at any time each analysis can be and must be** started from what is **already given**, because at the moment when the results of the project become applied, the quality of data assets won't change significantly. **The combinatorial space of the analyses can be optionally limited on the level of experts, because in case of the real applications, the real-timed nature of analyses will always be more important than the analytical quality.** So, when doing n-layered calculations, the value of 'n' is a question of human decision. The correct analytical goals, and thus the finding of potential templates are assured by the **interdisciplinary** researcher/partner team. The

previous publications of the co-operating researchers are the **guarantees of preparedness**. The volume of citations and their impact factors **can not be considered as a risk** of the project, because of the low number of papers of interdisciplinary topics, and the rapid development on the field. **The project bears no communicational risks**: the partners can rapidly share all news with the target audiences within the upcoming online journals (e.g. KOINE, Voice of the Countryside, MIAU), in contrast with the relatively slow publishing strategies of conferences and editorial boards.

5 Appliers, applications, image

The goal of the project is the methodological support of the Hungarian **EU-presidency**, and the governmental programs that are in close relation with it. Based on the available data assets (restructured for further use), the project shall produce appraisable results even in the planning (test data production) phase for the strategic direction shifts that may be the consequences of potential balance losses. Is it true that in case of the currently emphasized points of national questions like the tax system or social aids, the greatest external (objective) pressure on the change of them can be observed at the moment of the submission of the proposal? On EU-level, the verification of strategic directions of crisis management happens to be a similar analytical challenge...

In concern of the political affection of the primary application of the project can hardly be denied, because there is social innovation in it. So, in ideal case, in frame (in legitimate level) of the **OSAP** (National Statistical Data-Collection Program), pre-analysis of statistical data can be assured.

If the supreme decision-maker level doesn't want to use the system-approach leadership/support, then the application will be delegated to regional, business, and personal (civilians and non-profit/non-governmental organizations) level, that not only mean classis business models. The support of strategic planning can be viewed as classis business consultancy only on the level of (small and medium) enterprises.

The currently running DIPO-analyses (that have no automation goals themselves, but they base the foundations of them) will give a stable basis for international level extension until the (hopefully positive) decision about the project is made.