

# Időszerű és szükségszerű: HEnDA

(Vezetői összefoglaló)

## *Meditor Általános Fejlesztő Iroda*

Alapítva: 1989-ben  
Adószám: 24447775-2-13  
Honlap: [www.meditor.hu](http://www.meditor.hu)  
Email: [meditor@meditor.hu](mailto:meditor@meditor.hu)  
Telefon: 06-30-515-6362  
Kontakt: Végvári Lajos

### A szövegben alkalmazott rövidítések:

Holo-entrópiás adat-analítika:	HEnDA	<i>(Holo-Entropic Data Analytics)</i>
Mérés-adatgyűjtés:	DAQ	<i>(Data Acquisition)</i>
Holografikus Kutatási Stratégia:	HRS	<i>(Holographic Research Strategy)</i>
Mesterséges neuronháló:	ANN	<i>(Artificial Neural Networks)</i>
SmartData konverter:	SDC	<i>(SmartData Converter)</i>
Entrópiás adatsűrítés	EDC	<i>(Entropic Data Compression)</i>

Korunk legfontosabb **nyersanyaga** az adat, pontosabban az a benne rejlő információ, amely lehetővé teszi, hogy döntéseink megalapozottak legyenek. Az adatok logisztikájának és tárolásának módja töretlenül fejlődik, ennek következtében hatalmas adatmennyiségek halmozódnak föl egyéni, vállalati és kormányzati szinten is.

Az adat forrása bármi lehet, viszont az ehhez kapcsolódó **mintavételezésnek** tervezettnek és célirányosnak kell lennie. Bármennyire is igyekszünk azonban kiszűrni a felesleges adatokat, mindig kerülnek a rendszerbe olyan elemek, amelyek a **modellalkotást** megnehezítik, feleslegesen kötnek le számítási kapacitásokat és humán erőforrásokat. Az eredmény többnyire az, hogy adatbázisunk kezelhetetlen méretűvé növekszik.

A HEnDA rendszer olyan **adattányász** eszköz, amely rendkívül hatékonyan támogatja a célirányos döntések előkészítését:

- optimalizálja az adammintavételezés folyamatát
- felveszi/megméri majd tárolja a szükséges adatokat
- az adammintavételezést vizuálisan is nyomonkövethetővé teszi
- kiszűri a felesleges adatforrásokat
- mesterséges intelligencia alapú modellt készít
- iteratív módon alkalmazva az egyes részfolyamatokat optimumközeli állapotban tartja

*A HEnDA rendszer alkalmazásának legnagyobb előnye a döntési folyamatok egyszerűsítésében, automatizálhatóságában és a számítási kapacitásokkal való rendkívül hatékony gazdálkodásában van. A rendszer maga nem érzékeny az alkalmazási területre, bárhol használható, ahol adat keletkezik, azokat rögzíteni és osztályozni kell, majd azokból modellt kell alkotni.*

#### **Esettanulmány:**

Közismert, hogy a nemesfémek igen hatékonyan katalizálnak kulcsfontosságú kémiai folyamatokat. Rendkívül nagy hátrány, hogy a nemesfém bázisú katalizátorok rendkívül drágák. Német partnerünk azzal a feladattal bízott meg, hogy a HEnDA rendszerrel támogassunk meg egy olyan katalizátorfejlesztési folyamatot, amelyben a katalizátor sok komponensből áll és nem tartalmaz nemesfémeket.

Iteratív módon a HRS segítségével sikerült megtalálni az egyes katalizátor-komponensek megfelelő arányát úgy, hogy a katalitikus aktivitás az elméleti határ közelében került. A kisszámú (néhány száz) minta alapján mesterséges neuronhálózattal modellt illesztettünk, amely modell a teljes kombinatorikus térben (több százezer kombináció) vizuálisan elemezhetővé tette a különféle katalizátor-komponensek hatását.

**Összességében** tehát az történt, hogy a szakemberek kiinduló ötleteit a HEnDA rendszerezte, a felesleges, zaj szinten jelenlévő elemeket kiszűrte és a folyamat végén egy **szabadalomképes termék állt elő**. Melléktermékként számos információ keletkezett, amely a további kutatás-fejlesztés során felhasználható volt. A folyamatot teljes egészében egy asztali számítógéppel követtük és kontrolláltuk.

*Köcsk, 2023. március 1.*

Végvári Lajos

További anyagok: *Holo-entrópikus adat-analitika I.*  
*Holo-entrópikus adat-analitika II.*



# Holo-entrópikus adat-analitika I.

## (HEnDA)

### *Meditor Általános Fejlesztő Iroda*

Alapítva: 1989-ben  
Adószám: 24447775-2-13  
Honlap: [www.meditor.hu](http://www.meditor.hu)  
Email: [meditor@meditor.hu](mailto:meditor@meditor.hu)  
Telefon: 06-30-515-6362  
Kontakt: Végvári Lajos

#### A szövegben alkalmazott rövidítések:

Holo-entrópikus adat-analitika:	HEnDA	<i>(Holo-Entropic Data Analytics)</i>
Mérés-adatgyűjtés:	DAQ	<i>(Data Acquisition)</i>
Holografikus Kutatási Stratégia:	HRS	<i>(Holographic Research Strategy)</i>
Mesterséges neuronhálók:	ANN	<i>(Artificial Neural Networks)</i>
SmartData konverter:	SDC	<i>(SmartData Converter)</i>

## 1. Előzmények

A Meditor alapításától kezdve foglalkozik mérés-adatgyűjtő (DAQ) rendszerek fejlesztésével. A 2000-es évek elejétől veszünk részt az akadémiai hálózat kutatás-fejlesztési projektjeiben. Az együttműködés keretében eleinte heterogén hardverkörnyezetben, opensource operációs rendszerek alatt futó mérésadatgyűjtő-vezérlő szoftvereket, később a mérésadatgyűjtés célzottságát és hatékonyságát növelő optimumkereső szoftvereket fejlesztettünk.

Mind az együttműködés, mind a fejlesztés a mai napig folyik. Alapvető cél egy kutatási és/vagy rendszer-üzemeltető automata megvalósítása, amely a konkrét fizikai környezethez adaptívan alkalmazkodik, azt mintegy megtanulja, így képessé válik annak irányítására, az optimumkritériumoknak megfelelő állapotban tartására.

A HenDA 3 fő komponensből áll:

- |               |  |
|---------------|--|
| [1] DAQ modul | (zeneMEA - mérésadatgyűjtő szoftver)   |
| [2] HRS modul | (Holografikus Kutatási Stratégia<br>- optimumkereső és adatmintavételezés irányító szoftver) |
| [3] SDC modul | (henI23 - adatszűrő és mesterséges intelligencia szoftver)                                   |

A heterogén hardverkörnyezetben is működő DAQ modul összegyűjti a szükséges adatokat. Az adat-mintavételezést (milyen fizikai jeleket milyen felbontással mérjünk) a HRS határozza meg. A SDC kiszűri az irreleváns inputokat és egyrészt módosítja a HRS beállításait, másrészt a szűrt adatokból olyan modellt alkot, amelyek a lehető legkisebb zajjal terheltek. A modulok kapcsolatrendszerét az 1. mellékletben tanulmányozhatjuk.

A programok 'C' programnyelven íródtak, Linux környezetben futnak.

## 2. Az egyes modulok működésének rövid ismertetése

### 2.1. A DAQ modul

A DAQ modul maga is modul rendszerű, 4 fő komponensből áll.

- *console*: főmodul, a megosztott memóriát és a rendszerüzeneteket kezeli;
- *driver*: hardverspecifikus adatmintavételező és hardverkontrol modul;
- *scope*: online adatkijelző modul;
- *predict*: mesterséges neuronhálón alapuló előrejelző (adatszűrő) modul.



```
* Config: htc_long * LastCommand: * 2021-ju1-13 06:43:44 *  
  
-----  
Welcome in EasyEnough-world!  
Meditor Digital Strategies since 1989.  
www.meditor.hu * meditor@meditor.hu  
Sol omnibus lucet.  
asy | nough  
-----  
> Type 'exit' or '^C' to leave ZaiMea-system.  
* ZInterface memory O.K.  
* Config file: /etc/htc_long loaded.  
* /meditor/zeneMEr/bin/zm_driver_start  
* DRIVER: shared memory O.K.  
* DAQ hardware opened.  
* DRIVER: New config.  
* DRIVER: Initiated.  
* /meditor/zeneMEr/bin/zm_scope_start  
* Scope: shared memory O.K.  
* Scope: initiated.  
  
[ | config htc_long | killall | exit ]
```

« A *console* főmodul felhasználói felülete

A *driver* modul feladata a különböző hardverelemekből egyetlen esemény-keretbe összegyűjteni az adatokat és azt a többi modul számára a megosztott memóriában hozzáférhetővé tenni. A driver modul felel továbbá a háttértáron való adattárolásért.

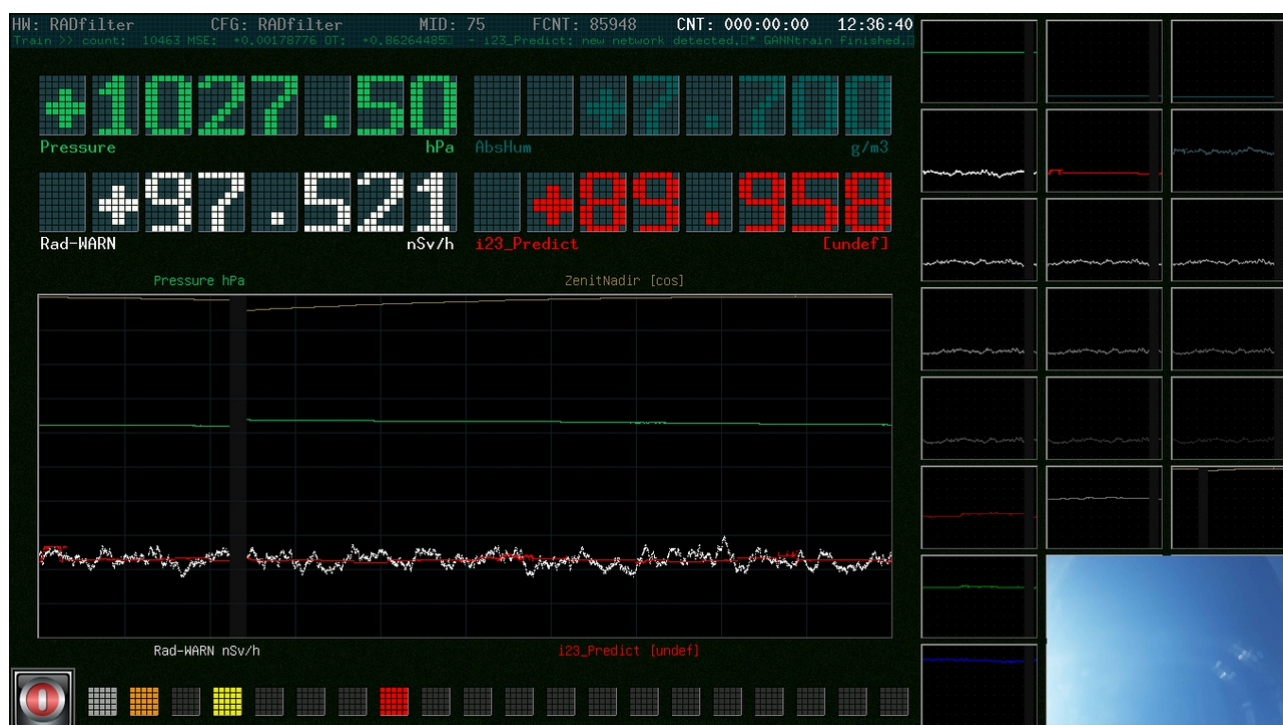
A *driver* modul beállításait felhasználói oldalról egy konfigurációs fájlon keresztül adhatjuk meg. Egy konkrét mérési összeállítást mutatunk meg a [2. mellékletben](#).

A *scope* modul felhasználói felületein követhetjük nyomon a mérési folyamatot. Példákat találunk erre a

[https://www.meditor.hu/meditor\\_ene.html](https://www.meditor.hu/meditor_ene.html)

elérési útvonalon (20 percenként frissül).

A *predict* modul alkalmazástól függően többféle funkciót is elláthat. A mérési adatok alapján felállított modell nemcsak előrejelzésre, hanem szabályozó jel előállítására és online adatszűrésre is alkalmas.



Adatszűrés mesterséges neuronhálóval (fehér: háttérsugárzás, piros: a háttérsugárzás prediktált adata).

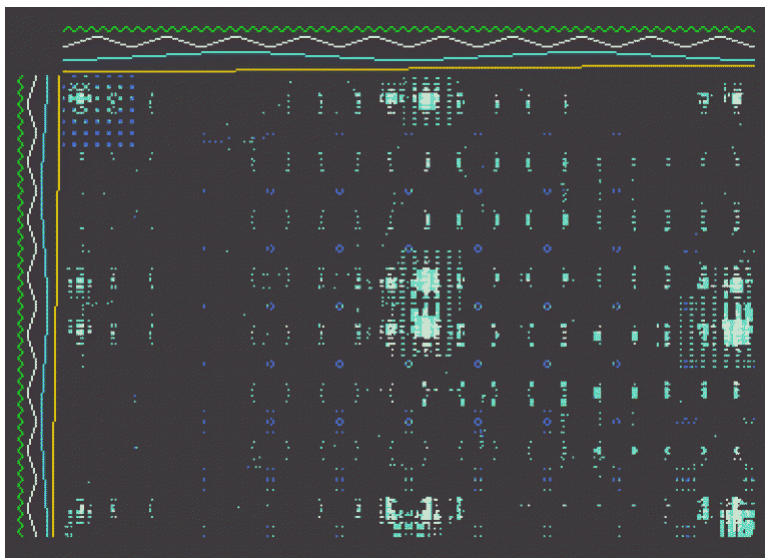
## 2.2. A HRS modul

A modul iteratív módon keresi meg az egyes inputoknak azon értékeit, amelyek az output optimális értékéhez tartoznak. A HRS alkalmas szabadalomképes termékek célzott kifejlesztésére is. Az eljárás publikált, a legfontosabb cikkek és egy szabadalmi leírás itt érhetők el:

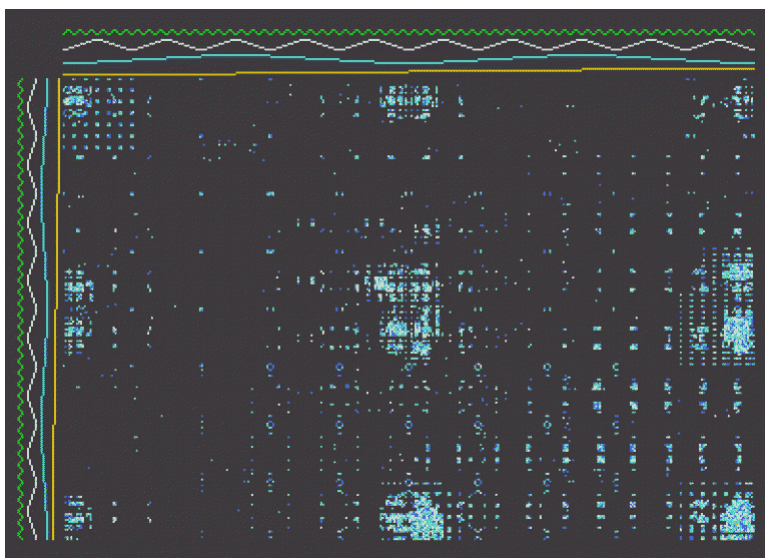
Alapcikk: <https://www.meditor.hu/public/cattoday.pdf>  
HRS összevetése a GA-sal: <https://www.meditor.hu/public/applcat3.pdf>  
HRS és a neurális hálók: [https://www.meditor.hu/public/HM\\_ANNassessment.pdf](https://www.meditor.hu/public/HM_ANNassessment.pdf)  
Legutóbbi: [https://www.meditor.hu/public/TTK-CTNT\\_Claripaper2019.pdf](https://www.meditor.hu/public/TTK-CTNT_Claripaper2019.pdf)  
Szabadalmi leírás: [https://www.meditor.hu/public/patent\\_01.pdf](https://www.meditor.hu/public/patent_01.pdf)

A *HRS* egyik legnagyobb előnye, hogy rendkívül toleráns a zajra. Ugyanannak a térnek egy kisebb és egy nagyobb zajjal terhelt változatában keressük az optimumot:

A világosabb foltok jelzik a 8 dimenziós térben az optimumot. A pesszimum helyeket az algoritmus elkerüli, tehát a mintavételezések helye önmagában informatív!



A fenti tér 250 %-os (!) zajjal terhelt változata. Jól látszik a *HRS* robusztus zajtűrése.



### 2.3. A *SDC* modul

Jelenleg ez a legintenzívebb fejlesztési terület. Elvi háttere még nem publikált, azonban a gyakorlati tesztek és az alkalmazhatóság vizsgálata már 2 éve folyik, a *DAQ* modulhoz való illesztése megtörtént.

Az alapmegközelítés az, hogy egy input-output rendszert mindenféle oksági előfeltételezés nélkül vizsgálunk és keressük a legnagyobb rendezettségű állapotot (entrópia elemzés). Az irreleváns adatkapcsolatokat kizárjuk a modellalkotásból, így egy BigData adathalmaz helyett egy **SmartData** adathalmazt kapunk.

A rendszer matematikája a *Shannon* képlet módosításán alapul, lényegében az entrópia és az információ fogalmát kapcsolja össze újszerűen úgy, hogy az adatbázisból kinyerhető információmennyiség a lehető legnagyobb legyen.

Az adatelemzés során először előállítjuk a **SmartData** adatbázist. A kapott elemzési adatokat ezután többféleképpen használhatjuk fel.

A/ A szűkített (SmartData) adatbázist átadjuk egy tanulórendszernek (ANN). Vegyük észre, hogy egy jól paraméterezett eljárás jelentősen csökkenti a számítási igényt! Ennek különösen a beágyazott rendszerekben vagy más, kis fogyasztású hardverkörnyezetben futó alkalmazások esetében van jelentősége.

B/ Rendkívüli jelentősége van az entrópiikus megközelítésnek a *HRS* iteratív lépéseinek irányítása során: segítségével bármilyen optimumkeresési folyamat sikeres **mohó hegymászó**vá válik!

A *SDC* modul a háttérben dolgozik, interaktív felülete nincs. Látja a rendszer megosztott memóriáját, ezen keresztül kommunikál a többi modullal. Működése konfigurációs fájlon keresztül szabályozható.

Fontos megjegyezni, hogy az egyes modulok hardverszinten szeparálhatók: például a nagyobb számítási igényű adatelemző szoftvereket delegálhatjuk bármilyen más, hálózaton keresztül elérhető, nagyobb teljesítményű gépre, miközben a mérés, a szabályozás és a predikció más gépen vagy gépeken fut.

*Köcsk, 2023. január 4.*

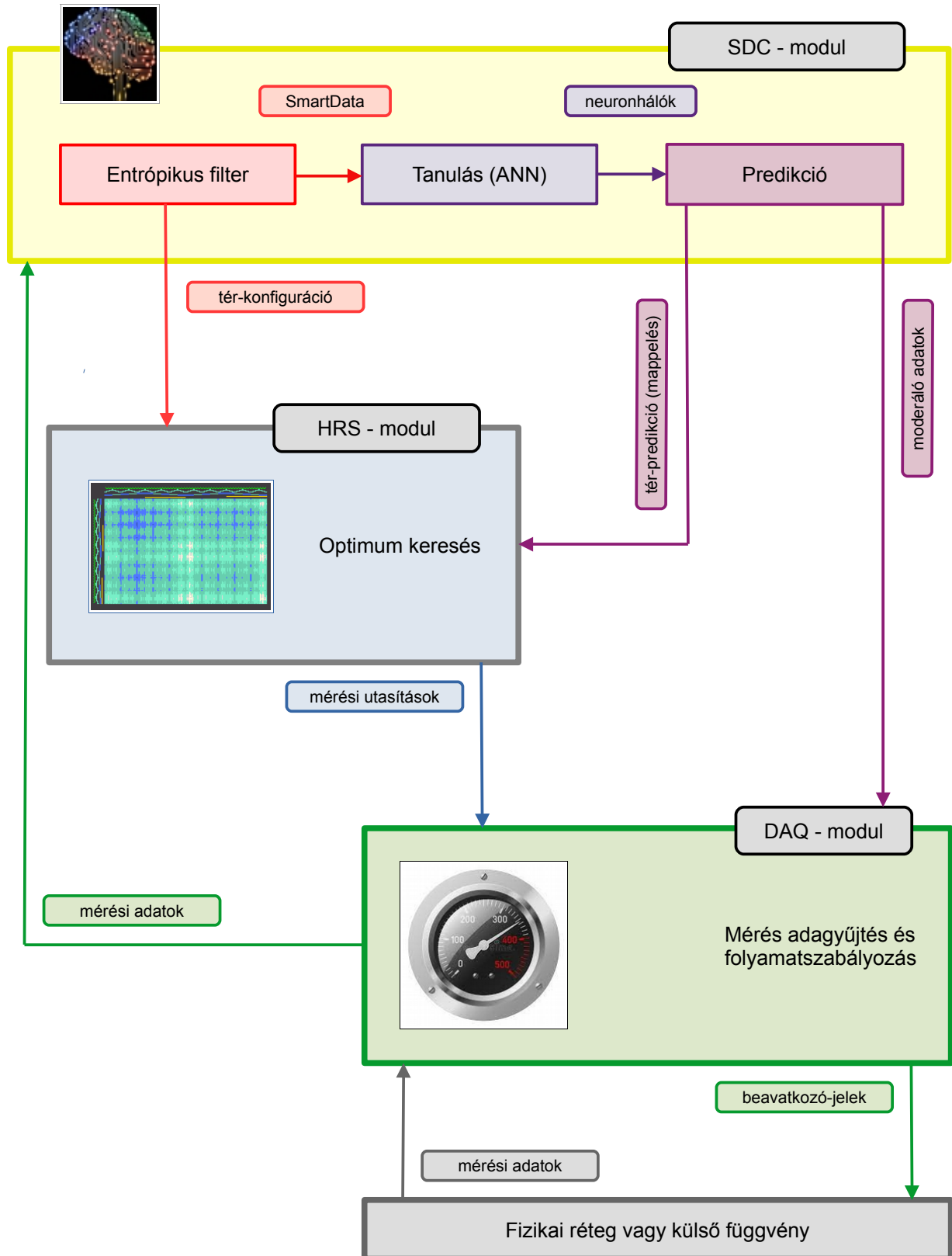


---

**Megjegyzés:** A *predict* modult a kapcsolati hálóban a *SDC* részeként tüntettük fel, mert rendszerszervezési szempontból oda tartozik. Funkcionálisan azonban jelen kiépítettségben a *DAQ* csomag modulja, ezért a szöveges részben ott tárgyaltuk.

---

A HEnDA szuperciklus kapcsolatrendszere



Integrált mérésadatgyűjtő rendszer BNS 98S sugárzás dózis távadóval



A zeneMEA mérőrendszer modbus/TCP protokollon keresztül fogadja az adatokat az Mduino21+ PLC felől.  
ethernet/UTP összeköttetés



Az Mduino21+ PLC kliensként viselkedik a mérőeszközök felé és kiszolgálóként a zeneMEA mérőrendszer felé.



A BNS 98S eszköz RS485 soros vonalon csatlakozik az Mduino21+ PLC-hez.



Az Mduino21+ PLC-hez számos egyéb mérőeszköz csatlakoztatható.

# Holo-entrópikus adat-analitika II.

## (HEnDA)

### *Meditor Általános Fejlesztő Iroda*

Alapítva: 1989-ben  
Adószám: 24447775-2-13  
Honlap: [www.meditor.hu](http://www.meditor.hu)  
Email: [meditor@meditor.hu](mailto:meditor@meditor.hu)  
Telefon: 06-30-515-6362  
Kontakt: Végvári Lajos

#### A szövegben alkalmazott rövidítések:

Holo-entrópikus adat-analitika:	HEnDA	<i>(Holo-Entropic Data Analytics)</i>
Mérés-adatgyűjtés:	DAQ	<i>(Data AcQuisition)</i>
Holografikus Kutatási Stratégia:	HRS	<i>(Holographic Research Strategy)</i>
Mesterséges neuronháló:	ANN	<i>(Artificial Neural Networks)</i>
SmartData konverter:	SDC	<i>(SmartData Converter)</i>
Entrópikus adatsűrítés	EDC	<i>(Entropic Data Compression)</i>

## 1. Bevezetés

A korábbi anyagban (*Holo-entrópikus adat-analitika I., 1. melléklet*) láttuk, hogy az adatok értelmezésének entrópikus megközelítése milyen központi szerepet tölt be a HEnDA rendszerben. A folyamatot, amely a nyers adatok entrópiaviszonyainak feltárásával indul és amelynek során a BigData adatbázisból **SmartData** adatbázis jön létre, *Entrópikus Adatsűrítésnek (EDC)* nevezzük. Minden további adatelemző, modellalkotó és vezérlő tevékenység ezen a SmartData adatbázison alapul.

*Az EDC során lényegében arról van szó, hogy azokat az okokat (inputokat), amelyek nagyobb valószínűséggel hatnak az okozatra (outputra), nagyobb súllyal vesszük figyelembe.*

## 2. Az entrópia fogalma a HEnDA rendszerben

Az entrópia fogalma szakterületeként más és más definícióval rendelkezik. Ezeknek a meghatározásoknak közös vonása, hogy az entrópiát valamilyen fizikailag vagy logikailag létező dolog rendezetlen eloszlásaként rögzítik. Kisebb entrópia = nagyobb rendezettség és fordítva. Egy adatbázis entrópiája esetén ez úgy értelmezhető, hogy kisebb entrópia esetén a vonatkozási rendszer nagyobb információ tartalommal bír.

Az entrópia egyik definíciója az informatikában Shannontól származik. Ez kiválóan rögzíti azokat a szabályokat, amelyek például egy adatátviteli csatorna méretezéséhez szükségesek. Van azonban 2 olyan dolog az alkalmazhatóságával kapcsolatban, ami a mi esetünkben akadályt jelent. Az egyik (1), hogy milyen kvantálási szabályokat alkalmazzunk, a másik (2) több input outputra vonatkozó entrópiájának összehasonlítása. A két eset szorosan összefügg.

### 2.1. A Maximálisan Kinyerhető Információ (MKI)

Az MKI megmutatja, hogy egy adatsort adott számú adatosztályra osztva mennyi információ (hány bit) nyerhető ki, ennek az értéknek a maximumát keressük. Sajnos próbálkozni kell, többféle adatosztályra kell ezt kiszámolni, többféle adathattárral még olyan triviális esetekben, ahol az adatsort *minőségi* változók (pl.: színek) írják le, mert előfordulhat, hogy - a példánál maradva - két vagy több szint célszerű egyetlen adatosztályba sorolni.

Folytonos adatsorokra vonatkozó MKI értékekre mutatunk példát különféle számú adatosztályokra és adatosztály határookra az 1. mellékletben.

Ahhoz, hogy két vagy több különböző adator (outputra vonatkozó) rendezettsége **összehasonlítható** legyen az egyes adatsorok MKI értékeit normálni kell, ugyanis magasabb számú kvantálási osztályokhoz magasabb információ tartalom (= több bit) tartozik. A normálás alapja az adatosztályok számának logaritmus. Ezzel bármely adatsorhoz tartozó MKI érték 0.00 és 1.00 közé kerül. Ez az érték valójában nem más, mint az adott adatsorhoz tartozó adat-**diverzitás** numerikus reprezentációja. Nagyobb érték magasabb fokú diverzitást jelent.

## 2.2 Az entrópia (E), az információnyereség (Y) és a helyettesíthetőség (R)

Egy adatsor-pár egymásra vonatkoztatott entrópiájának kiszámolása valójában egy eloszlásvizsgálat. Az adatsor egyik tagjának adatosztályaiban vizsgáljuk a másik tag adatosztályainak előfordulási valószínűségét és ezeket a valószínűségeket összegezzük. Minél erősebb a két adatsor közötti kapcsolat, annál nagyobb rendezettséget, tehát alacsonyabb entrópiaértéket fogunk kapni. Az 1. mellékelt 2. táblázatának *Entrópia\** oszlopában az egyes inputok outputra vonatkoztatott entrópiaértékeit tüntettük fel.

A kapott entrópiaértékre (E) a következő megállapítások érvényesek:

$$\begin{aligned} (1) \quad 0.00 &\leq E \leq \text{MKI\_output} \leq 1.00 \\ (2) \quad E_{\text{output}} &== 0.00 \end{aligned}$$

A kinyerhető információmennyiség (*yield*) az output paraméter MKI értékének és az output adott inputra vonatkozó entrópiájának a különbsége:  $Y = \text{MKI\_output} - E$ .

$$(3) \quad Y_{\text{output}} == \text{MKI}$$

A számítás C nyelvű kódját a 2. mellékletben találjuk. A Shannon-képlettől való eltérést (normálás az adatosztályok számának logaritmusával) **pirossal** kiemeltük.

A HEnDA rendszer számol még egy hasznos paraméterrel, ez az úgynevezett *helyettesíthetőség (R)*. Ez egy olyan korrelációs mutató, amely megmutatja két paraméter között az összefüggés erősségét, de nem mutatja meg annak előjelét:

$$(4) \quad 0.00 \leq R \leq 1.00$$

Minden paraméter saját magára vonatkoztatott R értéke = 1.00. Értelmszerűen, minél magasabb ez az érték, annál kisebb információvesztéssel hagyható el a két paraméter közül az egyik az adatelemzésből (SmartData!).

A 3. mellékletben megadjuk a vizsgált adatbázis néhány paraméterpárjának az egymásra vonatkoztatott *E*, *Y* és *R* értékét (kereszt-entrópia, kereszt-információhozam, helyettesíthetőség).

## 3. Az EDC eredményeinek felhasználása

Az *EDC* egyik nyilvánvaló előnye, hogy az adatbázisunk redukálódik, ezáltal a modellalkotás **számításigénye** csökken, akár radikálisan is. Általános tapasztalat, hogy egy input-output rendszerben az outputra legfeljebb néhány input van meghatározó hatással, az inputok egyrésze - különösen nagy adatbázisok esetén - csak zajszinten van jelen. Ezek kiszűrésére az *EDC* kimeneti adatai (E, Y, R) hatékonyan használhatók fel. A szűrési feltételek határértékei az *EDC* konfigurációs fájljában állíthatók be. Tekintettel arra, hogy az *EDC* által számított paraméterek standard értékek, a különböző adatbázisok elemzése során szerzett tapasztalatok **hordozhatók**, nem adatbázis-specifikusak.

Külön foglalkozunk az *EDC* felhasználására a *HRS* iterációk során (*Holo-entrópikus adat-analitika I., 2.2 pont*). A hatás kettős: egyrészt az adatbázis méretének csökkentése növeli az áttekinthetőséget, másrészt az iteráció során a változó tér-konfigurációk hatékonyabbá válnak.

Az idézett irodalomban is leírt, illetve a gyakorlatban használt iteratív HRS térkonfigurációk algoritmusát **intuitív** algoritmusnak tekinthetjük. Az algoritmus biztosította, hogy az inputok helye és szomszédviszonyai a téri ábrázolás, illetve az új minták meghatározása során folyamatosan változzanak. Ez módszer minden szimulációban globális optimumhoz vezetett és a gyakorlatban is biztosítható volt az elméleti határhoz való szoros konvergálás.

Az EDC alapú megközelítés legfőbb előnye, hogy egy-egy iteratív lépésben a **térkonfigurációt a minták mérési adatai határozzák meg**. Ez elvi különbség a korábbi eljárással szemben.

*A kiinduló állapot ugyanolyan, mint az intuitív algoritmus esetében. A második lépésben az inputokat sorba rendezzük úgy, hogy a legmagasabb frekvenciával valamelyik EDC kimenet szerinti legnagyobb hatású inputot, legalacsonyabb frekvenciával a legkisebb hatású inputot ábrázoljuk. A további lépésekben, újból és újból elvégezve az EDC vizsgálatot, egy tetszőleges sorba rendezési módszerrel a legnagyobb hatású inputot (ez iterációnként változhat a mérési adatok függvényében!) a legalacsonyabb frekvencia felé toljuk.*

Ez a téri ábrázolás azt fogja eredményezni, hogy a kezdetben sok foltra széteső optimum egyre nagyobb foltokba, végül egyetlen nagy foltba tömörül, amelyben bármilyen, gradiens mentén haladó **mohó hegyászó** algoritmussal meghatározhatjuk az optimumot.

Köcsk, 2023. február 3.



---

**Megjegyzés:** Ez a leírás beleértve a mellékleteket is nem publikus műszaki információkat tartalmaz.

---

## 1. táblázat

A Sound/Noise nevű input MKI (diverzitás) értékei különböző kvantálási szabályoknál					
Adatosztályok száma	LIN	LOG	EXP	SIN	COS
2	0.009905	0.009905	0.025918	0.004547	0.009905
3	0.032289	0.009602	0.094825	0.004325	0.017288
4	0.071483	0.012774	0.229964	0.004270	0.027105
5	0.120031	0.016811	0.471267	0.004130	0.036158
6	0.171934	0.022020	0.514879	0.003866	0.047344
7	0.225262	0.026930	0.575091	0.003830	0.058139
8	0.292471	0.031970	0.566372	0.003584	0.067958
9	0.345755	0.037979	0.560092	0.003392	0.079101
10	0.379198	0.043468	0.549271	0.003236	0.089895
11	0.370803	0.049546	0.534865	0.003108	0.099013
12	0.334127	0.055032	0.521260	0.002999	0.108870
13	0.300046	0.061031	0.506981	0.002905	0.118803
14	0.287465	0.066804	0.493858	0.002824	0.127531
15	0.293038	0.073045	0.481898	0.002752	0.134859
16	0.306726	0.078444	0.471007	0.002688	0.142490
17	0.323415	0.083809	0.461074	0.002630	0.150055
18	0.342177	0.089505	0.452003	0.002578	0.159054
19	0.361667	0.094620	0.443744	0.002531	0.168133
20	0.377910	0.100405	0.436173	0.002488	0.177617
21	0.389172	0.105842	0.429210	0.002448	0.186596
22	0.395152	0.110582	0.422750	0.002411	0.196785
23	0.396518	0.115099	0.416757	0.002377	0.207146
24	0.394142	0.119075	0.411168	0.002345	0.216796
25	0.391237	0.123704	0.405954	0.002315	0.225935
26	0.392607	0.128283	0.401067	0.002287	0.233731
27	0.395997	0.132757	0.396474	0.002261	0.241405
28	0.403502	0.137881	0.392147	0.002236	0.249097
29	0.412184	0.142819	0.388060	0.002213	0.256641
30	0.420870	0.148310	0.384192	0.002191	0.264088

Az adatosztályok határát 5 különböző függvény mentén határoztuk meg, az adatintervallumot 2-30 adatosztályra osztottuk. Minden értéket kiszámoltunk és ebből állapítottuk meg az optimális paramétereket. Jól látszik, hogy nagyon nem mindegy milyen kvantálási szabályokat alkalmazunk az adatelemzés során.

## 2. táblázat

A táblázatból az olvasható ki, hogy az outputra (*Rad-WARN* = háttérsugárzás) legnagyobb hatással (*Információ hozam oszlop*) az *Enthalpy* nevű, környezeti paraméterek alapján számolt input van. Létező összefüggés!

Kvantálási szabályok néhány adatsornál					
Paraméter	Függvény karakterisztika	Adatosztályok száma	MKI (diverzitás)	Entrópia*	Információ hozam*
Thermo	LIN	2	0.99441525	0.99336969	0.00112736
RelHum	EXP	2	0.96532071	0.99347452	0.00102252
Pressure	LOG	2	0.95045309	0.99053129	0.00396576
DewPoint	SIN	3	0.96883063	0.98951188	0.00498517
AbsHum	LIN	2	0.99996453	0.99269184	0.00180521
SpecHum	LIN	2	0.99998226	0.99210597	0.00239108
MixRatio	SIN	3	0.99960500	0.98944964	0.00504741
Enthalpy	SIN	4	0.99178336	0.98920258	<b>0.00529447</b>
RedColor	EXP	2	0.97382292	0.99444759	0.00004946
GreenColor	EXP	2	0.97521941	0.99445364	0.00004341
BlueColor	EXP	2	0.97615982	0.99444763	0.00004942
BlackWhite	EXP	2	0.97575458	0.99445564	0.00004141
Sound/Noise	EXP	7	0.57509120	0.99272081	0.00177624
ZenitNadir	COS	11	0.99997222	0.99402132	0.00047573
Rad-WARN**	LIN	2	0.99449705	0.00000000	0.99449705

*A kvantálási szabályokat minden adatsorra külön meg kell határozni!*

*(\* = magyarázat a szövegben)*

*(\*\* = output)*

```
double MediEntropy(struct ent *op, struct ent *part)
```

```
{unsigned int      count;
 unsigned int      i,j;
 unsigned int      ndx_op,ndx_part;

 double            part_info[MEDI_ENTRO_MAXQUANT];
 unsigned int      part_cnum[MEDI_ENTRO_MAXQUANT];
 double            maxi,tmpd;

 i=0;while(i<MEDI_ENTRO_MAXQUANT)
 {j=0;while(j<MEDI_ENTRO_MAXQUANT){part->cross_table[i][j]=0;j++;}
  i++;
 }

 count=0;while(count<op->data_len)
 {
  ndx_op   =   op->qua[count];
  ndx_part =   part->qua[count];
  part->cross_table[ndx_part][ndx_op]++;
  count++;
 }

 // count part_case_num
 i=0;while(i<part->quant)
 {part_cnum[i]=0;
  j=0;while(j<op->quant){part_cnum[i]+=part->cross_table[i][j];j++;}
  i++;
 }

 // compute part_info
 i=0;while(i<part->quant)
 {part_info[i]=0.00;
  j=0;while(j<op->quant)
  {if(part->cross_table[i][j])
   {
    tmpd   =   ((double)part->cross_table[i][j]/(double)part_cnum[i]);
    maxi   =   tmpd *   (log2(tmpd)/log2(op->quant));
    //maxi  =   tmpd*log2(tmpd);
    part_info[i]-=maxi;
   }
   j++;
  }
  i++;
 }

 // compute entropy
 part->entropy=0.00;
 i=0;while(i<part->quant)
 {
  tmpd=((double)part_cnum[i]/(double)op->data_len)*part_info[i];
  part->entropy+=tmpd;
  i++;
 }

 // compute yield
 part->yield=op->maxinfo-part->entropy;

 return(part->entropy);
 }
```

### 3. melléklet

A táblázatokban meteorológiai adatok szerepelnek. Jól látható az egyes paraméterek közötti összefüggés. Kimutatható a légköri nedvesség és a háttérsugárzás közötti gyenge összefüggés.

Kereszt-entrópiák (E)										
Paraméter	Thermo	RelHum	Pressure	DewPoint	AbsHum	SpecHum	MixRatio	Enthalpy	Sound/Noise	Rad-WARN
Thermo	0.00000000	0.96419977	0.94750067	0.60598531	0.48169323	0.44699787	0.64232070	0.64591446	0.55018660	0.99336969
RelHum	0.99329432	0.00000000	0.93841075	0.91914952	0.94482595	0.96817527	0.95247461	0.96801452	0.56751256	0.99347452
Pressure	0.99146283	0.95327836	0.00000000	0.95407134	0.99801387	0.99905268	0.98828951	0.97571700	0.56773562	0.99053129
DewPoint	0.41931903	0.88657802	0.92706018	0.00000000	0.17277033	0.24200128	0.38482037	0.46881755	0.52443373	0.98951188
AbsHum	0.47614396	0.91018213	0.94850243	0.44692920	0.00000000	0.29540175	0.58771995	0.65899125	0.53929111	0.99269184
SpecHum	0.44143086	0.93351372	0.94952351	0.49059788	0.29538402	0.00000000	0.58772677	0.62575286	0.54285201	0.99210597
MixRatio	0.42813303	0.89062081	0.93251846	0.35404600	0.34714217	0.34717072	0.00000000	0.45728442	0.52841733	0.98944964
Enthalpy	0.30267744	0.91778302	0.91832037	0.30892125	0.33438031	0.26792125	0.32514243	0.00000000	0.52604227	0.98920258
BlackWhite	0.98612121	0.96312255	0.94872540	0.96066284	0.98481210	0.98486171	0.99469125	0.98379035	0.55027276	0.99445564
Sound/Noise	0.92449922	0.94404478	0.92980337	0.87910391	0.89946097	0.90947543	0.91693421	0.92293449	0.00000000	0.99272081
ZenitNadir	0.98139989	0.96068261	0.94843413	0.95831487	0.98084176	0.98111934	0.98950257	0.97213427	0.53936390	0.99402132
Rad-WARN	0.99328790	0.96429818	0.94648734	0.96568534	0.99815932	0.99759119	0.99642044	0.98913613	0.57445849	0.00000000

Kereszt-hozamok (Y)										
Paraméterek	Thermo	RelHum	Pressure	DewPoint	AbsHum	SpecHum	MixRatio	Enthalpy	Sound/Noise	Rad-WARN
Thermo	0.99441525	0.00112093	0.00295242	0.36284532	0.51827130	0.55298439	0.35728430	0.34586891	0.02490459	0.00112736
RelHum	0.00112093	0.96532071	0.01204235	0.04968111	0.05513858	0.03180699	0.04713039	0.02376885	0.00757864	0.00102252
Pressure	0.00295242	0.01204235	0.95045309	0.01475928	0.00195066	0.00092959	0.01131549	0.01606636	0.00735558	0.00396576
DewPoint	0.57509623	0.07874269	0.02339291	0.96883063	0.82719420	0.75798098	0.61478463	0.52296581	0.05065746	0.00498517
AbsHum	0.51827130	0.05513858	0.00195066	0.52190143	0.99996453	0.70458051	0.41188505	0.33279211	0.03580009	0.00180521
SpecHum	0.55298439	0.03180699	0.00092959	0.47823275	0.70458051	0.99998226	0.41187823	0.36603051	0.03223919	0.00239108
MixRatio	0.56628222	0.07469990	0.01793463	0.61478463	0.65282236	0.65281155	0.99960500	0.53449894	0.04667387	0.00504741
Enthalpy	0.69173781	0.04753769	0.03213272	0.65990938	0.66558422	0.73206101	0.67446257	0.99178336	0.04904892	0.00529447
BlackWhite	0.00829404	0.00219815	0.00172769	0.00816779	0.01515243	0.01512056	0.00491375	0.00799302	0.02481844	0.00004141
Sound/Noise	0.06991603	0.02127593	0.02064972	0.08972672	0.10050356	0.09050684	0.08267079	0.06884887	0.57509120	0.00177624
ZenitNadir	0.01301537	0.00463810	0.00201896	0.01051576	0.01912277	0.01886292	0.01010243	0.01964909	0.03572729	0.00047573
Rad-WARN	0.00112736	0.00102252	0.00396576	0.00314529	0.00180521	0.00239108	0.00318456	0.00264724	0.00063271	0.99449705

Helyettesítési értékek (R)										
Paraméterek	Thermo	RelHum	Pressure	DewPoint	AbsHum	SpecHum	MixRatio	Enthalpy	Sound/Noise	Rad-WARN
Thermo	1.00000000	0.00112711	0.00296057	0.45844367	0.52082868	0.55538727	0.45139973	0.49113192	0.04194164	0.00000752
RelHum	0.00112711	1.00000000	0.01243825	0.06465561	0.05707473	0.03290457	0.06133367	0.03476524	0.01314919	0.00000691
Pressure	0.00296057	0.01243825	1.00000000	0.01953384	0.00204791	0.00097651	0.01497659	0.02387420	0.01292364	0.00003951
DewPoint	0.47055013	0.06442138	0.01916331	1.00000000	0.67727921	0.62097380	0.63453866	0.60605855	0.06941322	0.00003535
AbsHum	0.51793836	0.05509738	0.00194652	0.65619212	1.00000000	0.70419270	0.51784701	0.47025951	0.05991543	0.00001319
SpecHum	0.55229537	0.03176403	0.00092814	0.60162911	0.70418021	1.00000000	0.51813305	0.51752188	0.05392326	0.00001888
MixRatio	0.44905616	0.05923005	0.01424017	0.61500341	0.51803326	0.51832860	1.00000000	0.60033064	0.06198332	0.00003484
Enthalpy	0.49243524	0.03383764	0.02287930	0.59203260	0.47413865	0.52180015	0.60506511	1.00000000	0.05847504	0.00003077
BlackWhite	0.00849090	0.00225056	0.00176367	0.01050566	0.01550155	0.01545986	0.00631997	0.01155452	0.04259177	0.00000024
Sound/Noise	0.07252312	0.02207160	0.02135890	0.11693737	0.10418053	0.09376304	0.10773741	0.10084414	1.00000000	0.00001131
ZenitNadir	0.00699095	0.00249152	0.00108136	0.00709662	0.01026445	0.01011903	0.00681741	0.01490307	0.03216943	0.00000144
Rad-WARN	0.00000752	0.00000691	0.00003951	0.00003535	0.00001327	0.00001898	0.00003502	0.00003068	0.00000654	1.00000000

(A Holo-entrópikus adat-analítika I-II. Az ismertetőanyag vége)

# A HEnDA: alkalmazási lehetőség (függelék)

## *Meditor Általános Fejlesztő Iroda*

Alapítva: 1989-ben  
Adószám: 24447775-2-13  
Honlap: [www.meditor.hu](http://www.meditor.hu)  
Email: [meditor@meditor.hu](mailto:meditor@meditor.hu)  
Telefon: 06-30-515-6362  
Kontakt: Végvári Lajos

### A szövegben alkalmazott rövidítések:

Holo-entrópiikus adat-analitika:	<b>HEnDA</b>	<i>(Holo-Entropic Data Analytics)</i>
Mérés-adatgyűjtés:	<b>DAQ</b>	<i>(Data Acquisition)</i>
Holografikus Kutatási Stratégia:	<b>HRS</b>	<i>(Holographic Research Strategy)</i>
Mesterséges neuronháló:	<b>ANN</b>	<i>(Artificial Neural Networks)</i>
SmartData konverter:	<b>SDC</b>	<i>(SmartData Converter)</i>
Entrópiikus adatsűrítés	<b>EDC</b>	<i>(Entropic Data Compression)</i>

A HEnDA rendszer egy olyan informatikai platform, amely moduljainak összehangolásával képes komplex problémák kezelésére, ugyanakkor az egyes modulok önállóan is alkalmazhatók egy-egy részterület hatékony támogatására.

Legegyszerűbben ezt a DAQ modul szerepével kapcsolatban tudjuk bemutatni. Egy gyártási vagy egy kutatási folyamatban, ahol fizikai jelek jelentik az adatforrást, előnyösen használhatjuk a mérés-adatgyűjtő rendszert akár önállóan is, de egy olyan adatelemzési feladatnál, amelyben az adatok más jellegűek, vagy már rendelkezésre állnak egy korábbi mérésből, a DAQ modulra nincs szükség.

Más szavakkal: a HEnDA rendszert nem kell konkrétan fizikailag hozzákapcsolni egy adott berendezéshez, mert adatforrásként bármilyen adatsort képes felhasználni. Az egyetlen megkötés csak annyi, hogy a vizsgált dolognak tulajdonságai (attributumai) legyenek. Nézzünk erre néhány példát!

#### (1) Korábbi kutatási anyagok elemzése

Irodalmi feldolgozással adatbázist állítunk fel, akár úgy is, hogy egymástól távol eső kutatási területeket kapcsolunk össze (interdiszciplináris adatelemzés).

#### (2) Nyilvános adatbázisok elemzése

Ebben az esetben széles körben hozzáférhető adatokra támaszkodunk. Rendszerint akkor merül fel ilyen igény, amikor egy régebbi adatbázist új szempontok szerint vizsgálunk át. Tipikus alkalmazási terület például a klimatikus adatok újraértelmezése, vagy más adatokkal való kölcsönhatásának vizsgálata.

#### (3) Egészségügyi adatbázisok

Probléma orientált adatfelvételezéssel vagy adatrendezéssel adatbázist állítunk fel, majd a HEnDA segítségével feltárjuk azokat az összefüggéseket, amelyek a közösségi kockázatvállalás csökkentésében vagy az egyénre szabott terápiában felhasználhatók.

#### (4) Közvélemény kutatás

Akár marketing, akár politikai céllal gyűjtöttük az adatokat, a HEnDA nagyon pontosan ki fogja jelölni a célcsoport preferenciáit.

#### (5) Hatáselemzés

A HEnDA politikai döntések vagy jogszabályváltozások hatásának követésére is alkalmas. Ebben az esetben idősoros adatokra van szükség a változtatás előttről és az azt követő időszakról. A HRS modul vizualizációs képessége rendkívül könnyen értelmezhetővé teszi a döntési helyzetben lévők számára döntéseik valódi hatását.

#### (6) Stratégiák összevetése

Ebben az esetben a HEnDA adatforrása egy szimulátor lehet, esetleg irodalmi adatokkal kiegészítve. Különleges lehetőség, hogy szimulátorként irodalmi forrásokon és/vagy online adatokon betanult mesterséges intelligenciát használunk. Ez egyben előrejelzési lehetőséget is jelent olyan területeken, mint például a gazdasági (tőzsdei) folyamatok, a politikai vagy katonai stratégiák összevetése és optimalizálása.

*Köcsk, 2023. június 22.*

Végvári Lajos

*További anyagok: Holo-entrópikus adat-analítika I.  
Holo-entrópikus adat-analítika II.*

