

# Machine-readable $\neq$ reusable

FAIR értékelés, metaadat-minőség és mesterséges  
intelligencia

# Mottó

„Ha az adatok a tudás kincsei, akkor a metaadat a térkép, amely segít eligazodni bennük. A rendezett és jól strukturált metaadatok nemcsak a kutatás megbízhatóságát növelik, hanem biztosítják, hogy az adatok a FAIR elvek mentén valóban értéket teremtsenek.”

(Virág Gabriella, OpenAI 2025)




Mit jelent valójában  
a jó metaadat?

**A jó metaadat:**

- értelmezhető
- kontextust ad
- összekapcsolható
- ember és gép számára is használható

A jó metaadat nemcsak megtalálhatóvá,  
hanem értelmezhetővé és újrahasznosíthatóvá  
teszi az adatot.



Technikailag FAIR,  
de valóban  
újrahasznosítható?

Ha egy adat gépileg olvasható, attól valóban  
újrahasznosítható is?

Technikai megfelelés  $\neq$  jó metaadat-minőség

# Mit mér valójában egy FAIR ellenőrző?

## Amit automatizáltan tud mérni

- PID-ek, DOI-k jelenléte
- API / OAI elérés
- szabványos metaadatséma
- gépi olvashatóság
- technikai interoperabilitás

## Amit nem tud mérni

- metaadat jelentése
- kutatási kontextus
- kulcsszavak minősége
- emberi értelmezhetőség
- valódi újrahasznosíthatóság

# Mit mutat a FAIR score és mit nem?

## **Magas FAIR érték**

minimális leírás  
kevés kulcsszó  
nincs DOI, ORCID  
kevés kutatási kontextus



## **Alacsonyabb FAIR érték**

részletes leírás  
gazdag metaadat  
több kapcsolódó elem  
informatívabb rekord

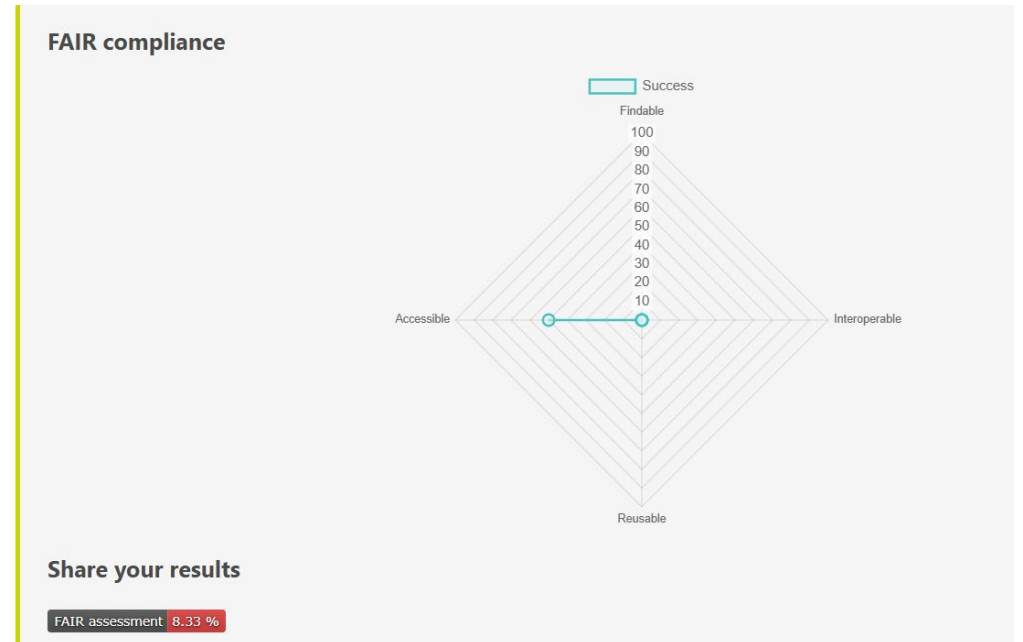
# Saját tapasztalatok a FAIR értékeléssel - 1

Adatcsomag: minimális metaadat, szerző, cím és publikálás dátuma, tárgyszó, DOI nincs, ORCID nincs megadva, az absztrakt két szóból áll

Summary:



Summary:



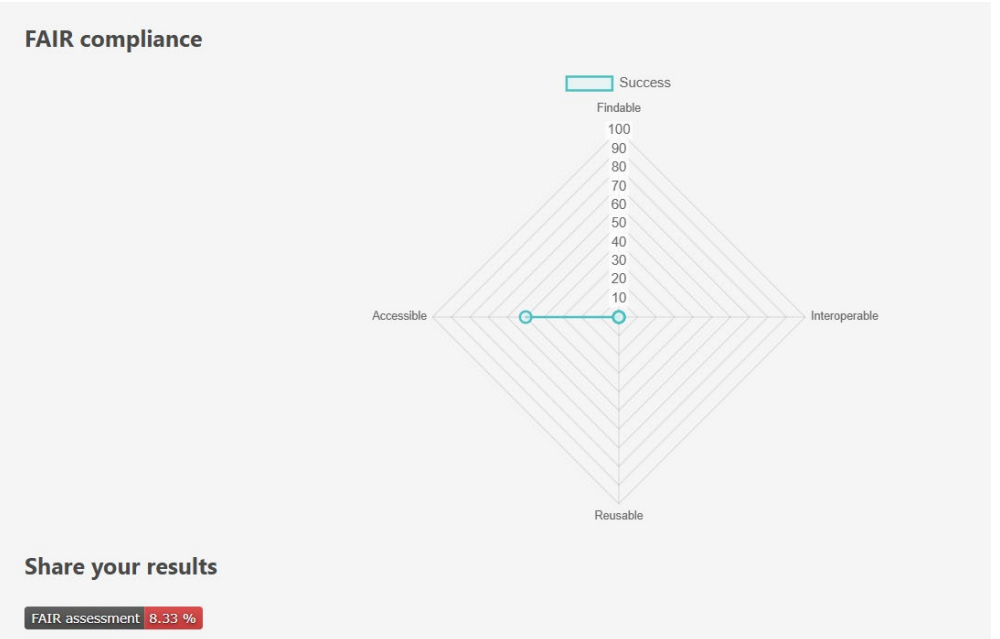
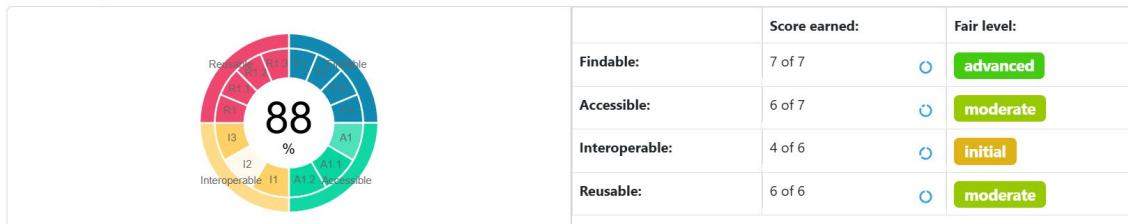
# Saját tapasztalatok a FAIR értékeléssel - 2

Adatcsomag: szerző, cím és publikálás dátuma, tárgyszó, kulcsszavak, bőséges absztrakt, kapcsolódó közlemények, DOI van, de nem linkként szerepel, ORCID van

Summary:



Summary:



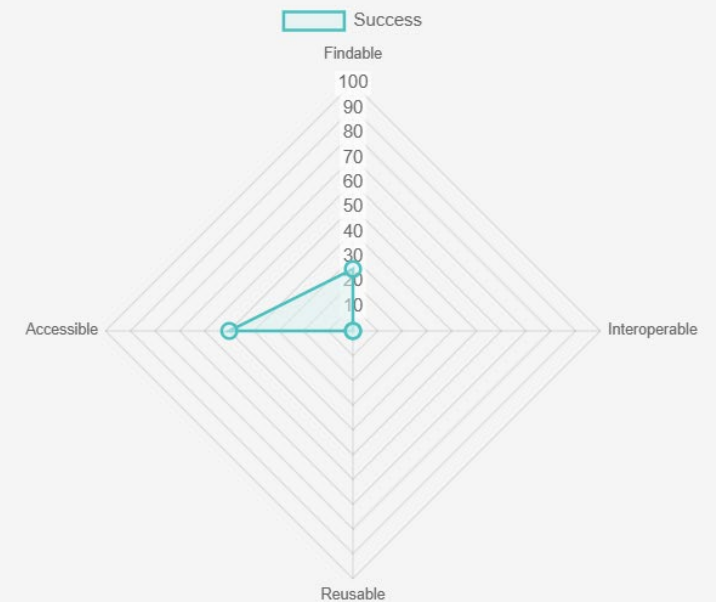
# Saját tapasztalatok a FAIR értékeléssel - 3

Adatcsomag bőséges metaadattal:  
szerző, cím, publikálás dátuma,  
tárgyszó, kulcsszavak, bőséges  
absztrakt, kapcsolódó közlemények,  
DOI, ORCID...

Summary:



## FAIR compliance



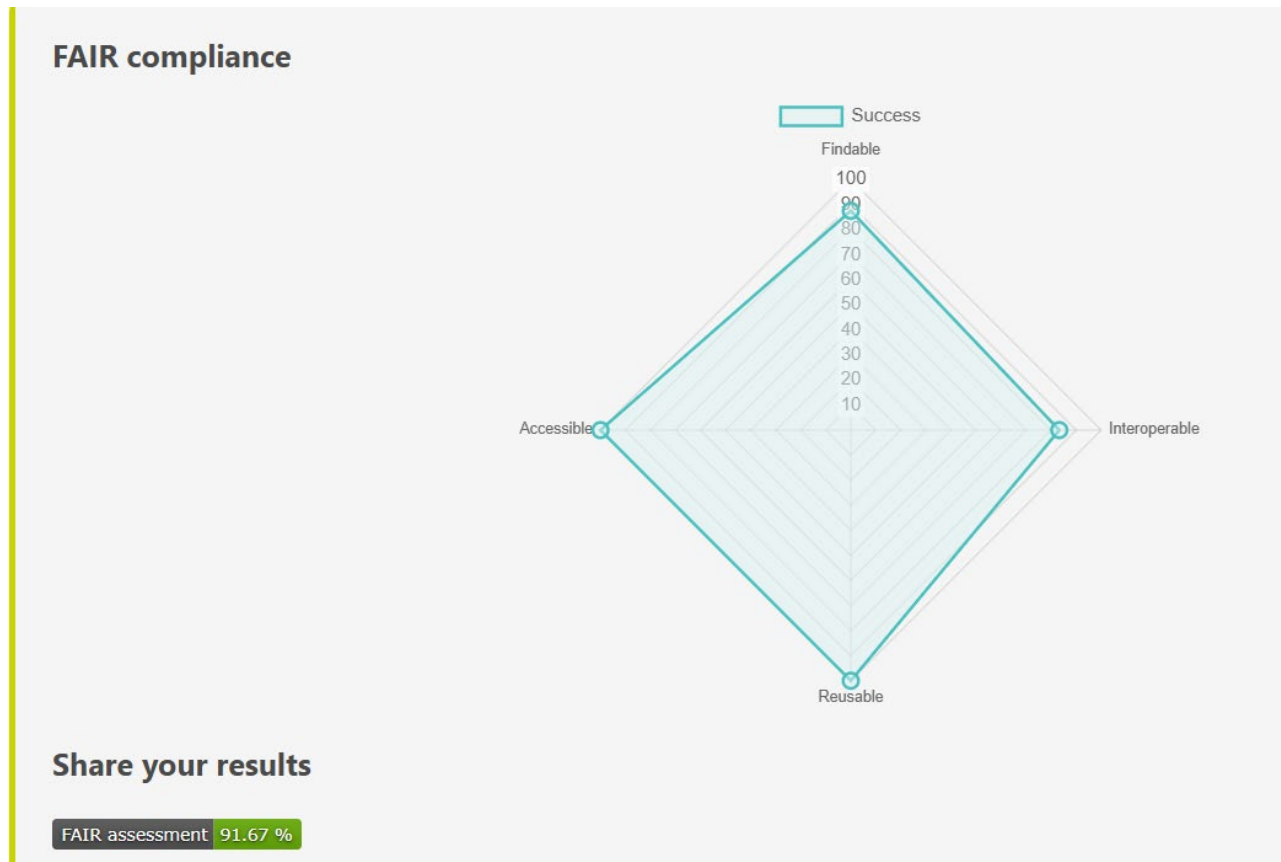
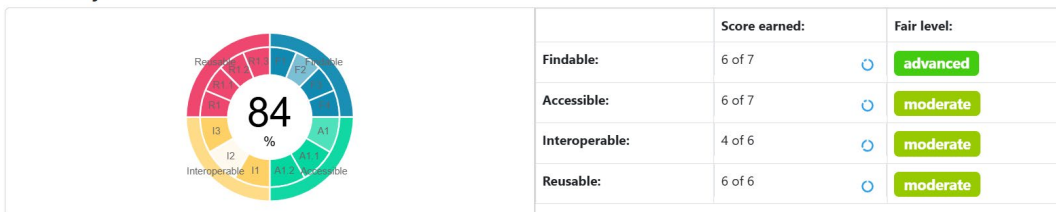
Share your results

FAIR assessment **16.67 %**

# Saját tapasztalatok a FAIR értékeléssel - 4

Zenodo adatrepozitórium  
Adatcsomag minimális metaadattal

Summary:



```

<dc:creator>jaszianna</dc:creator>
<dc:date>2024-11-26</dc:date>
<dc:description>&lt;p&gt;Persistently increased post-stress activity of paraventricular thalamic neurons is
essential for the emergence of stress-induced alterations in behavior.&lt;/p&gt;</dc:description>
<dc:identifier>https://doi.org/10.5281/zenodo.14222815</dc:identifier>
<dc:identifier>oai:zenodo.org:14222815</dc:identifier>
<dc:publisher>Zenodo</dc:publisher>
<dc:relation>https://github.com/jaszianna/PVT_ASD/tree/v1.0.0</dc:relation>
<dc:relation>https://doi.org/10.5281/zenodo.14222814</dc:relation>
<dc:rights>info:eu-repo/semantics/openAccess</dc:rights>
<dc:rights>Creative Commons Attribution 4.0 International</dc:rights>
<dc:rights>https://creativecommons.org/licenses/by/4.0/legalcode</dc:rights>
<dc:title>jaszianna/PVT_ASD: PVT ASD</dc:title>
<dc:type>info:eu-repo/semantics/other</dc:type>
</oai_dc:dc>

```

```

<creators>
<creator>
<creatorName nameType="Personal">jaszianna</creatorName>
<familyName>jaszianna</familyName>
</creator>
</creators>
<titles>
<title>jaszianna/PVT_ASD: PVT ASD</title>
</titles>
<publisher name="Zenodo"/>
<publicationYear>2024</publicationYear>
<dates>
<date dateType="Issued">2024-11-26</date>
<date dateType="Updated">2024-11-26</date>
</dates>
<resourceType resourceTypeGeneral="Software"></resourceType>
<relatedIdentifiers>
<relatedIdentifier relatedIdentifierType="URL" relationType="IsSupplementTo" resourceTypeGeneral="Software">
https://github.com/jaszianna/PVT_ASD/tree/v1.0.0</relatedIdentifier>
<relatedIdentifier relatedIdentifierType="DOI" relationType="IsVersionOf">10.5281/zenodo.14222814</relatedIdentifier>
<relatedIdentifier relatedIdentifierType="URL" relationType="IsIdenticalTo">
https://archive.softwareheritage.org/swh:1:dir:1fb1b08a90babfb7a8f0814c3b7a3448b12b377a;origin=https://doi.org/10.5281/zenodo.14222814;visit=swh:
1:snp:69100ea56bd9e8c42454c0e8aad4c0f4f4337540;anchor=swh:1:rel:726911b726c0327e6efd4da067b7d9b6cc9f8b5a;path=/</relatedIdentifier>
</relatedIdentifiers>
<version>v1.0.0</version>
<rightsList>
<rights rightsURI="https://creativecommons.org/licenses/by/4.0/legalcode" rightsIdentifierScheme="spdx" rightsIdentifier="cc-by-4.0">Creative
Commons Attribution 4.0 International</rights>
<rights rightsURI="info:eu-repo/semantics/openAccess"/>
</rightsList>
<descriptions>
<description descriptionType="Abstract">Persistently increased post-stress activity of paraventricular thalamic neurons is
essential for the emergence of stress-induced alterations in behavior.</description>
</descriptions>
</resource>

```

## Dublin-Core xml DataCite xml

# A FAIR score mögött nem mindig áll újrahasznosítható metaadat

## **Technikai alapok**

Dublin Core, DataCite, schema.org , OAI-PMH , API , DOI, ORCID, ROR

### **Amik támogatják:**

- a megtalálhatóságot
- a hozzáférhetőséget
- az interoperabilitást
- a gépi feldolgozhatóságot és újrahasznosíthatóságot

### **De nem garantálják:**

- a részletes leírást
- a jó kulcsszavakat
- a kutatási kontextust
- a valódi újrahasznosíthatóságot



# Segít az AI

## **Mire használható az AI?**

- automatikus kulcsszó-generálás
- kivonatolás
- többnyelvű metaadat
- entitásfelismerés
- szemantikus gazdagítás
- kapcsolódó rekordok ajánlása
- ...

# VibeARP: AI-alapú adatgazdász asszisztens

Kísérleti AI-ágens a HUN-REN ARP rendszerben

Fejlesztő: HUN-REN SZTAKI

<https://researchdata.hu/sikertortenetek/arp-ai-vibearp-adatgazdasz-asszisztens-első-lepesei>

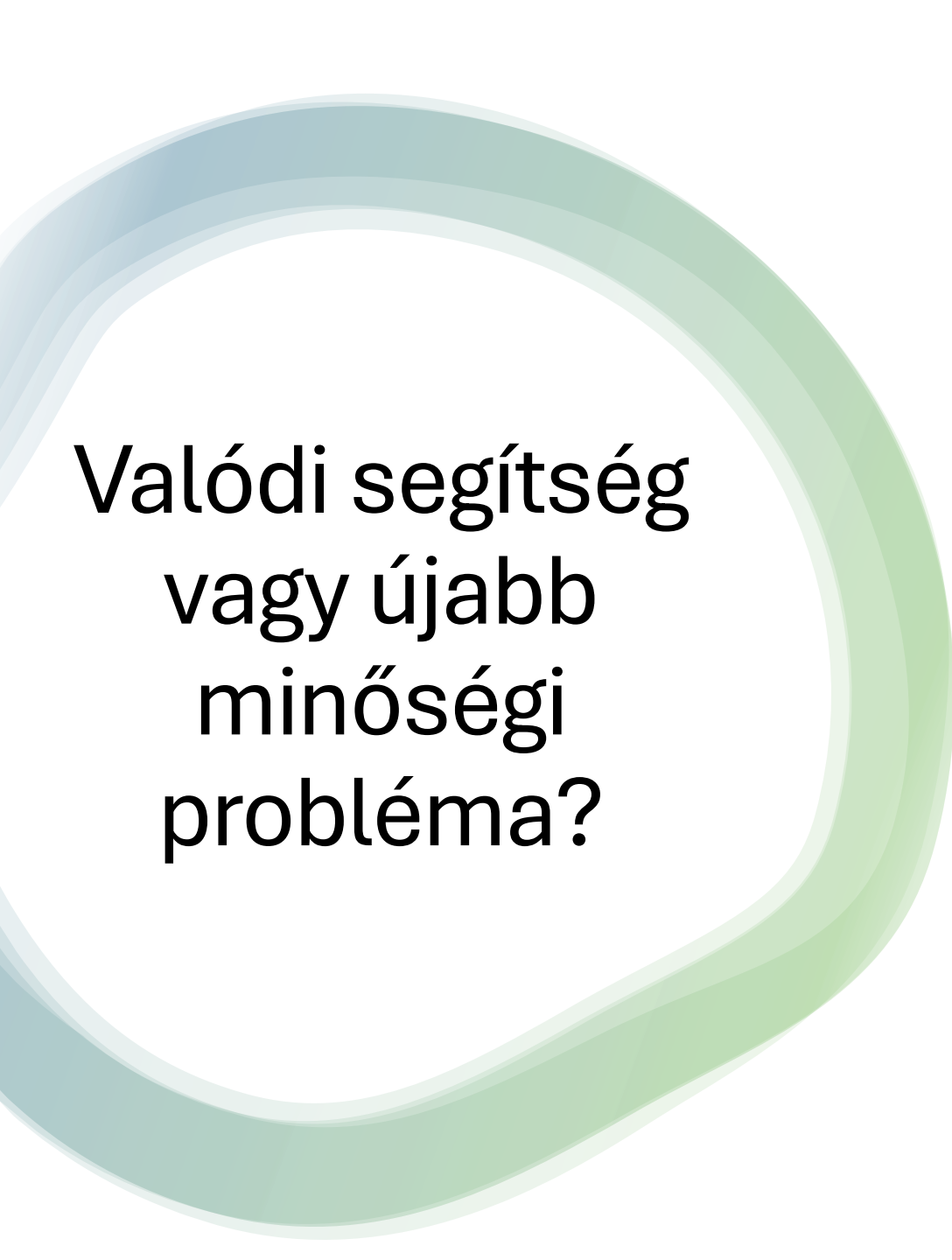
## Mire épít?

- metaadatsémákra
- strukturált metaadatokra
- szemantikus kapcsolatokra
- webes információkinyerésre



## A metaadat minősége közvetlenül befolyásolja:

- a keresés pontosságát
- az AI-válaszok minőségét
- a kapcsolódó rekordok felismerését
- az újrahasznosíthatóságot



Valódi segítség  
vagy újabb  
minőségi  
probléma?

### **AI által okozott vagy felerősített problémák**

- hallucináció
- pontatlan kulcsszavak
- túl általános metaadat
- hibás kapcsolatok
- hamis magabiztosság
- forrás nélküli információk

### **Következtetések:**

A FAIR score nem feltétlenül tükrözi a valódi újrahaznosíthatóságot.

A jó metaadat értelmezhető és kontextust ad.

Az AI segíthet, de nem helyettesíti a minőségi metaadatot.

**Machine-readable  $\neq$  reusable**

**Köszönöm a figyelmet!**