



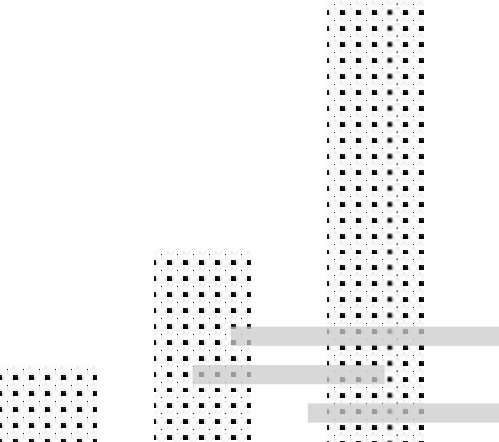


**Ahol a bölcsészettudományi kutatások és a
közgyűjtemények találkoznak: a mesterséges intelligencia**
Maróthy Szilvia

Szakkönyvtári Seregszemle
2022. március 29.

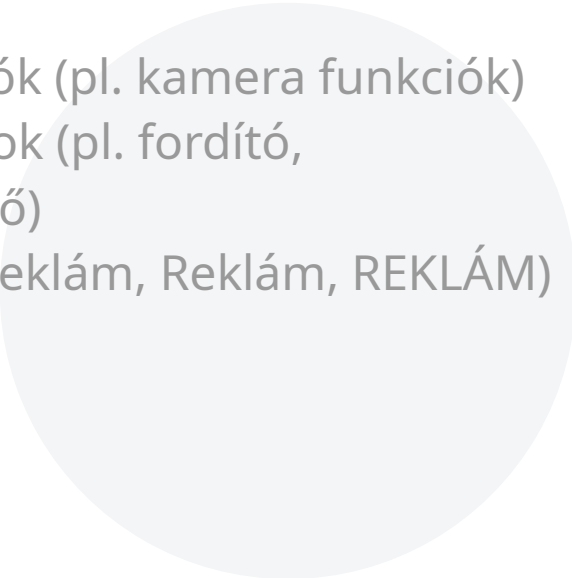
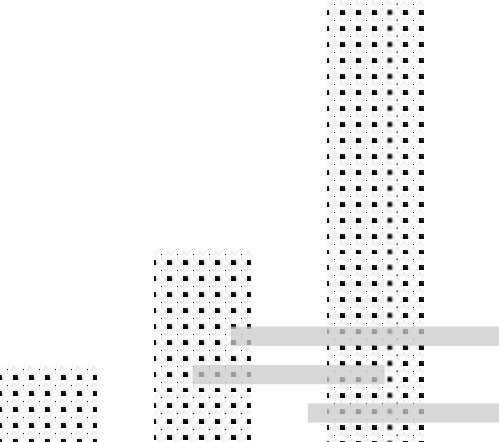


MI kezdetek dióhéjban

- 
- 
- Alan Turing „Computing Machinery and Intelligence” (*Mind*, 1950)
 - 1955 és 1956 J. McCarty, M. Minsky, N. Rochester és C. Shannon: Dartmouth mesterséges intelligenciával foglalkozó kutatói projekt/workshop
- 

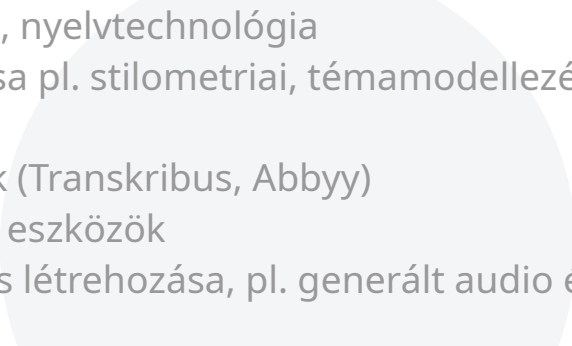

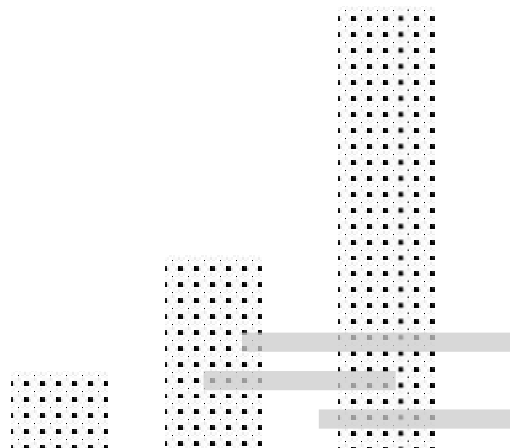


MI ma

- 
- telefonos applikációk (pl. kamera funkciók)
 - webes szolgáltatások (pl. fordító, képfelismerő/kereső)
 - ajánlórendszerek (reklám, Reklám, REKLÁM)
- 

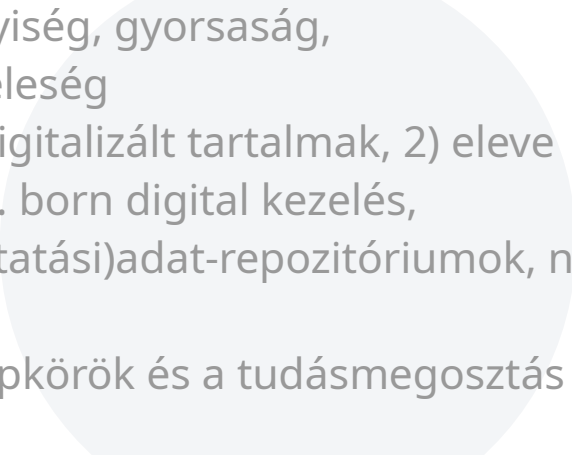
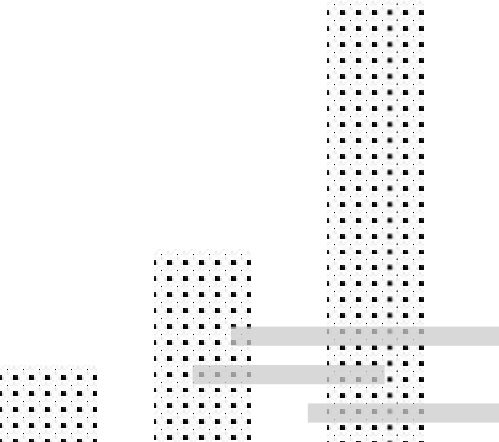


MI és bölcsészettudomány

- 
- 
- számítógépes nyelvészet, nyelvtechnológia
 - NLP eszközök alkalmazása pl. stilometriai, témamodellezéses vizsgálatok
 - OCR és HTR technológiák (Transkribus, Abbyy)
 - képmanipulációs/javítási eszközök
 - műalkotások elemzése és létrehozása, pl. generált audio és/vagy vizuális alkotások
- 


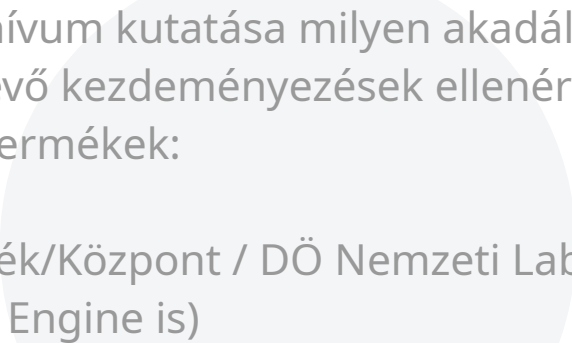
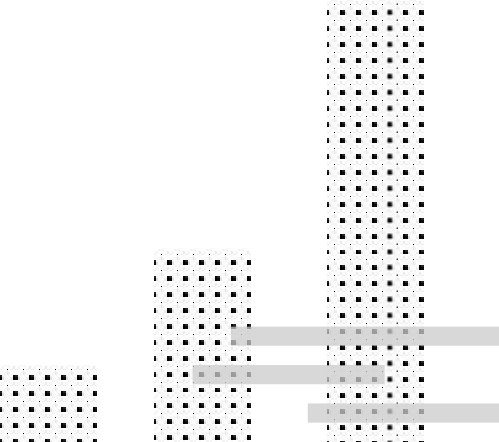


Big data (és makroanalízis)

- 
- *mi a nagy?* adatmennyiség, gyorsaság, változatosság/különféleség
 - *hogyan lesz nagy?* 1) digitalizált tartalmak, 2) eleve digitális tartalmak (vö. born digital kezelés, webarchiválás), 3) (kutatási)adat-repozitóriumok, nyílt (kapcsolt) adat
 - *ki csinálja kinek?* szerepkörök és a tudásmegosztás felelőssége
- 

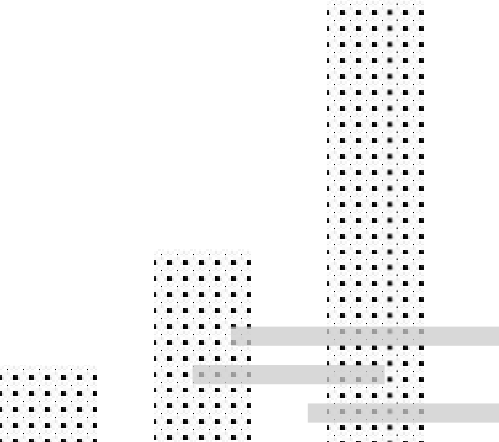



Kitérő: helyzetjelentés

- 
- a web vagy a webarchívum kutatása milyen akadályokba ütközhet a már meglévő kezdeményezések ellenére?
 - webről mentett sajtótermékek:
 - OSZK Webarchívum
 - ELTE BTK DB Tanszék/Központ / DÖ Nemzeti Labor Cikk-kereső (Sketch Engine is)
- 
- 

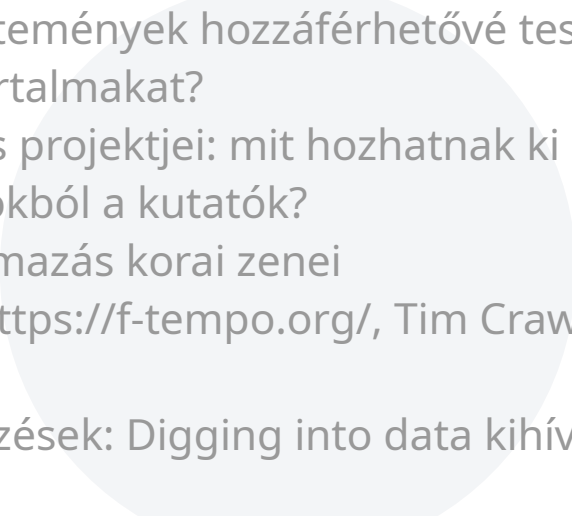
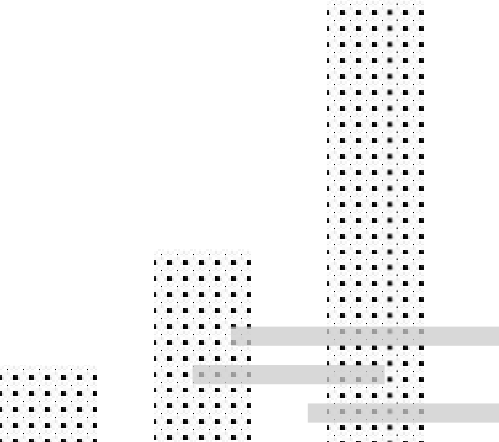


Szerepkörök és határátlépések

- 
- közgyűjtemény: feldolgoz, megőriz, szolgáltató
 - kutatási intézmények: a történeti dokumentumok feltárása, elemzése, értelmezése, narratív bemutatása
 - profitorientált szereplők: kulturális örökség szolgáltatása pénzért
 - léptékváltás a határok átlépésében/elmosódásában: szöveggondozás közgyűjteményben? dokumentum digitalizálás, bibliográfia készítés kutatóintézetben?

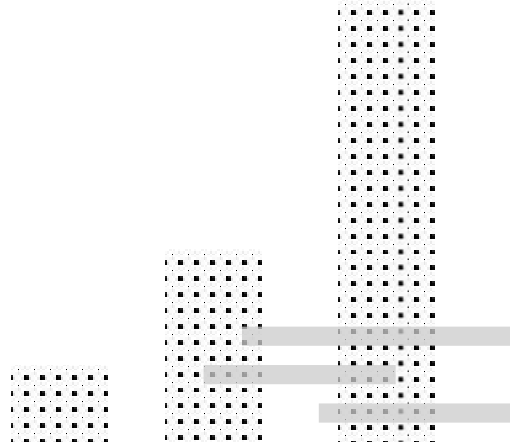


Kitérő: példa

- 
- Miért jó, ha a közgyűjtemények hozzáférhetővé teszik a digitalizált/digitális tartalmakat?
 - pl. British Library Labs projektjei: mit hozhatnak ki a közgyűjteményi adatokból a kutatók?
 - F-TEMPO webes alkalmazás korai zenei nyomtatványokhoz (<https://f-tempo.org/>, Tim Crawford, 2019)
 - Hasonló kezdeményezések: Digging into data kihívás, Europeana Grants
- 



MI a következő?

- Legyenek az intézmények, tartalmaik nyitottak, rendezzenek versenyeket, workshopokat a tartalmak megismerésére/megismertetésére!
 - Fontos a szerepkörök átgondolása, közös (közgyűjteményi és kutatási) ökoszisztémák megalkotása!
 - Nem az a fő kérdés, mit hogyan digitalizáljunk. Az a kérdés, mit kezdünk a digitális tartalmakkal? mit hogyan digitalizáltunk már? (a teljesség, a hitelesség, a pontosság...)
- 



MI a következő?

- Digitális és digitalizált dokumentumok tartalmi feltárása (OCR, HTR, képfelismerés, képi elemzés, audiovizuális tartalmak elemzése)
 - Keresési lehetőségek finomítása
 - Kutatási ökoszisztémák létrehozása
 - A kutatásban pl. szövegek kritikai apparátusa, szövegemendáció, archaikus nyelvű szövegek modernizálása, szerzői attribúció, tartalomelemzés, osztályozás, tulajdonnevek azonosítása.
- 