

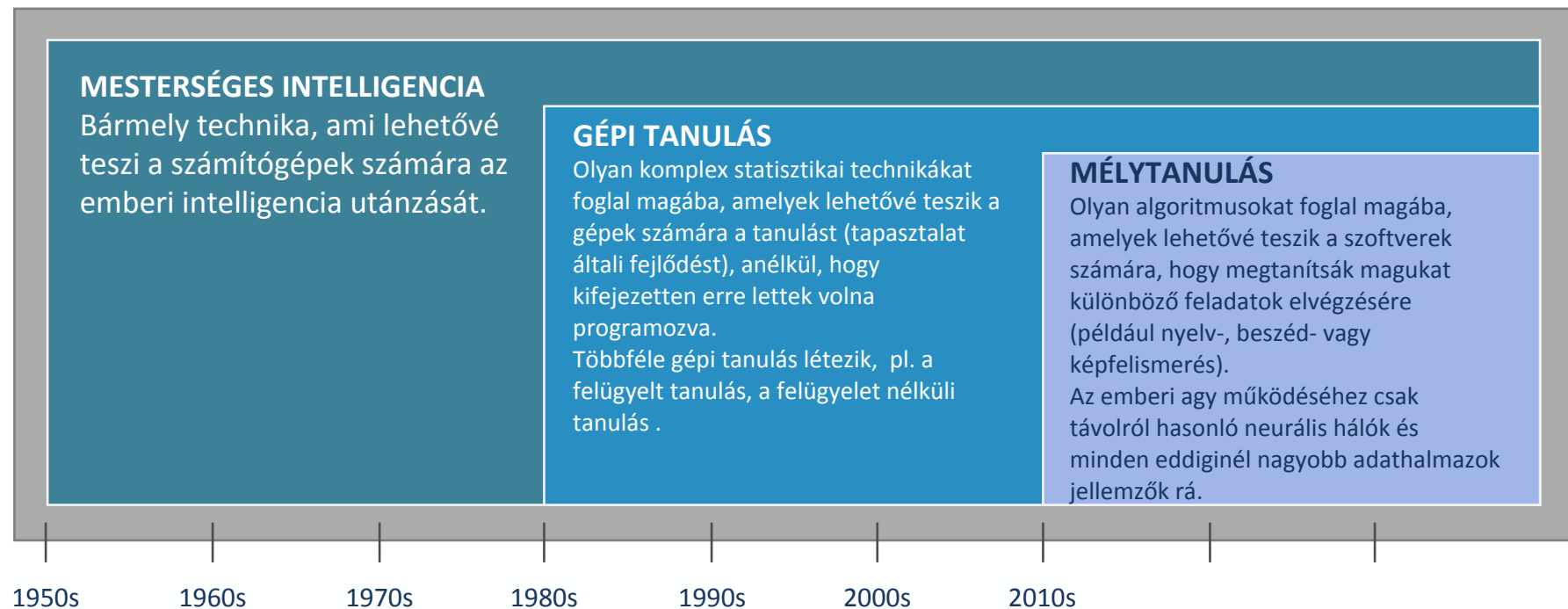
A NYELVTECHNOLÓGIA ÉS A MESTERSÉGES INTELLIGENCIA LEHETŐSÉGEI – HOL TARTUNK MOST, ÉS MIK A CÉLOK?

PRÓSZÉKY GÁBOR

NYELVTUDOMÁNYI KUTATÓKÖZPONT



A MESTERSÉGES INTELLIGENCIA, A GÉPI TANULÁS ÉS A MÉLYTANULÁS VISZONYA^[3]



A MESTERSÉGES INTELLIGENCIA (MI) MA

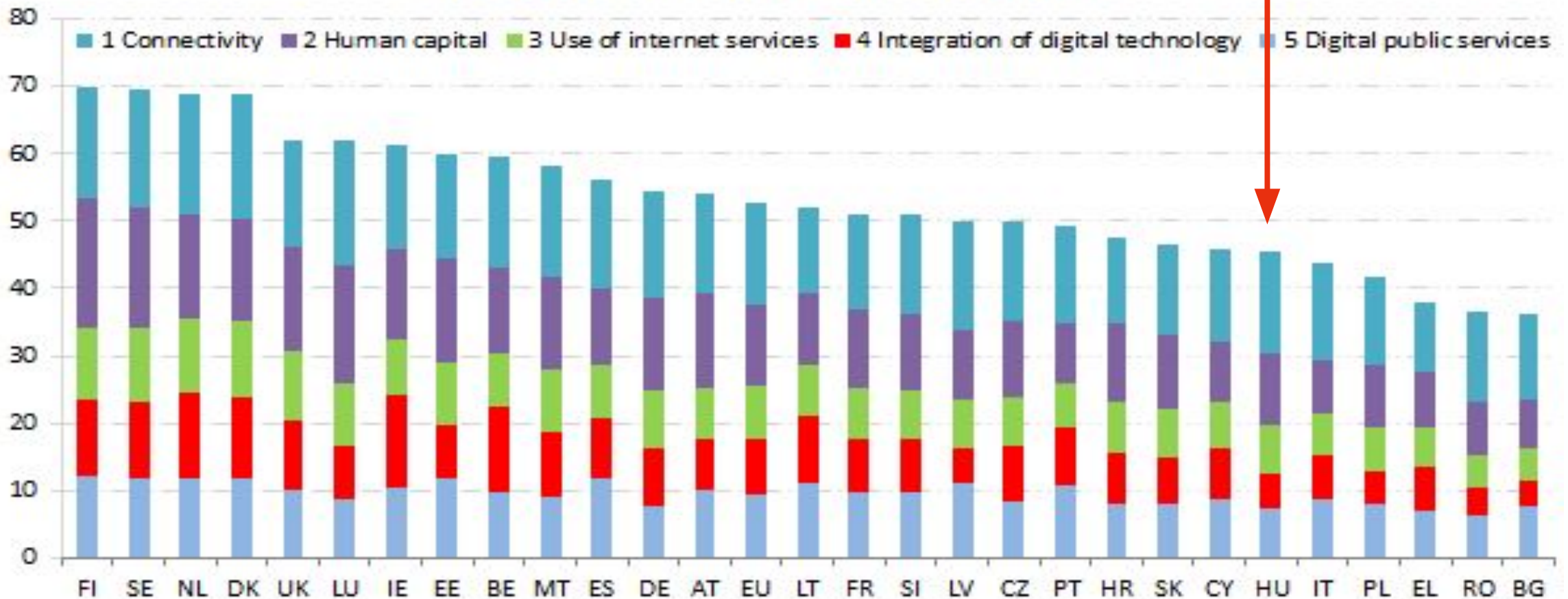
- A mesterséges intelligencia nagy lehetőség^[1]
- Európában egyre több új projektben a mesterséges intelligencia áll a középpontban
- Már ma is számos példa vesz körül minket a mindennapi életben...
- Itthon is megindultak a mesterséges intelligenciával (és azon belül a nyelvtechnológiával is) kapcsolatos akciók és tevékenységek: MI Koalíció, MILab, ...
- ... és a NYTK nyelvtechnológiai kutatásai

A MI A MINDENNAPOKBAN

- Marketing (pl. Netflix, Amazon, ...)
- Chatbotok (pl. banki, ügyfélszolgálati és menedzsment-területen)
- Pénzügyek (pl. tőzsdei kereskedés)
- Mezőgazdaság (pl. az erőforrások ideális felhasználása)
- Egészségügy
- Űrkutatás
- Intelligens autók
- Digitális személyi asszisztensek (Siri, Alexa, ...) – bár magyarra nincsenek
- ...
- Nyelvészet

MENNYIRE VAGYUNK DIGITÁLISAK?

Digital Economy and Society Index (DESI) 2019 ranking



NYELV ÉS INTELLIGENCIA

- *„Ha az eszközzel folytatott beszélgetést nem lehet megkülönböztetni egy emberrel folytatott hasonló beszélgetéstől, akkor azt az eszközt intelligensnek lehet nevezni.”*
(Alan Turing)
- **Probléma:** az emberi nyelvek hatékonyak, rugalmasak és összetettek
 - A szavaknak több jelentésük is lehet
 - Sokféle módon is el lehet mondani ugyanazt
 - Minden jelentés a (nyelvi és nem nyelvi) kontextustól függ
 - Sok a metafora, ui. van szó szerinti és átvitt értelem

MÉLYTANULÁS A NYELVTECHNOLÓGIÁBAN

A mélytanulás a nyelvtechnológia számos területén jelen van:

- Automatikus beszéd felismerés (ASR)
- Beszéd generálás (TTS)
- Dialógus rendszerek / chatbotok
- Szentimentelemzés
- Névelem-felismerés (NER)
- Relációkinyerés (RE)
- Összefoglaló-generálás
- Kérdés-válasz rendszerek (QA)
- Gépi fordítás (MT)
- ...

MÉLYTANULÁS A NYELVTECHNOLÓGIÁBAN 2022

Működő megoldások

Levélszemétszűrés

Let's go to Agra! ✓

Buy VIAGRA ... ✗

Szófaji egyértelműsítés

ADJ ADJ NOUN VERB ADV

Colorless green ideas sleep furiously.

Névkifejezés-felismerés

PERSON ORG LOC

Einstein met with UN officials in Princeton

Működő megoldások

Véleménybányászat

Best roast chicken in San Francisco! 👍

The waiter ignored us for 20 minutes. 👎

Koreferencia-feloldás

Carter told Mubarak he shouldn't run again.

Jelentésegértelműsítés

I need new batteries for my *mouse*.

Szintaktikai elemzés

I can see Alcatraz from the window!

Gépi fordítás

第13届上海国际电影节开幕...

The 13th Shanghai International Film Festival...

Információkivonatolás

You're invited to our dinner party, Friday May 27 at 8:30

Party
May 27
add

Elfogadható megoldások

Kérdés-megválaszolás

Q. How effective is ibuprofen in reducing fever in patients with acute febrile illness?

Parafrázis-készítés

XYZ acquired ABC yesterday

ABC has been taken over by XYZ

Szövegösszegzés

The Dow Jones is up

The S&P500 jumped

Housing prices rose

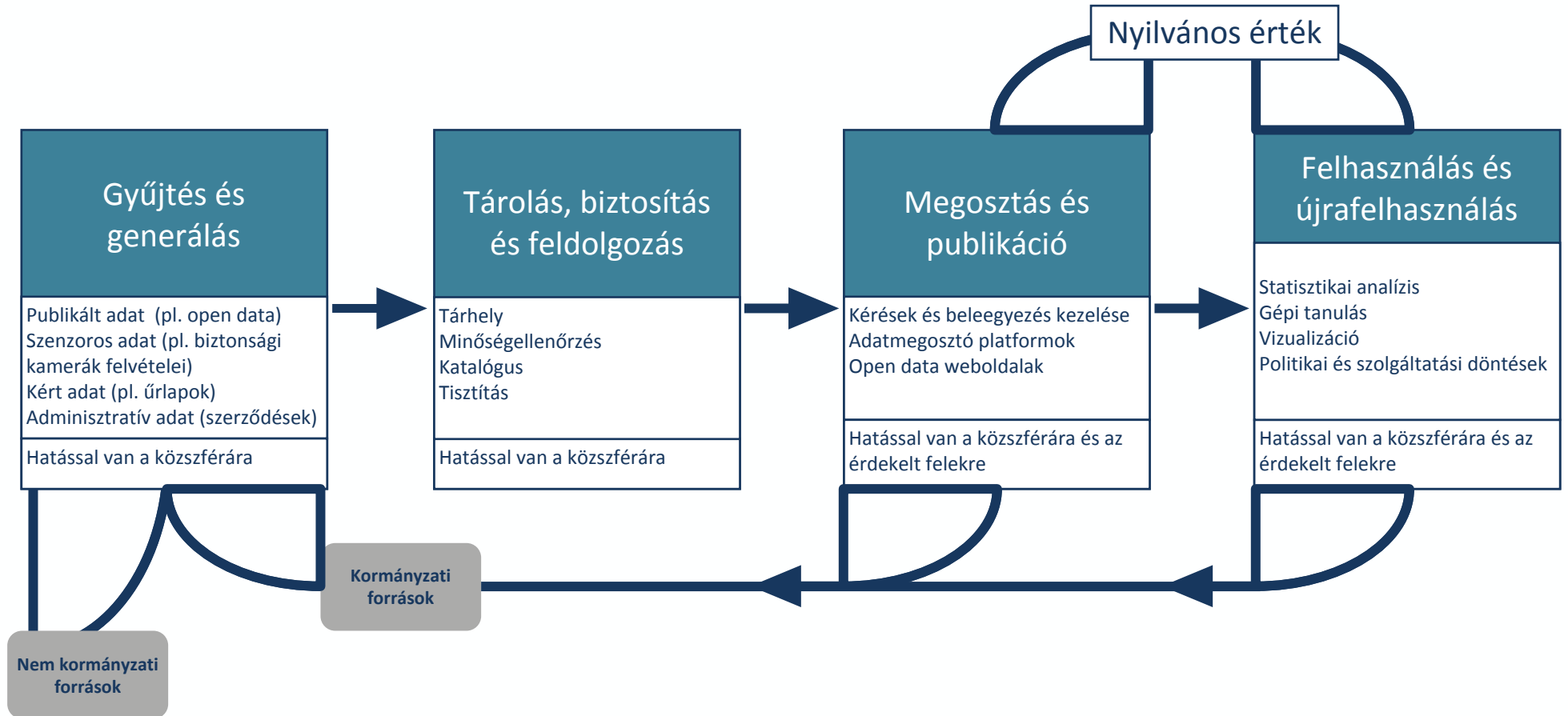
Economy is good

Dialogus

Where is Citizen Kane playing in SF?

Castro Theatre at 7:30. Do you want a ticket?

AZ ADAT A MESTERSÉGES INTELLIGENCIA „ÜZEMANYAGA” [4]



A DIGITÁLIS KORMÁNYZATI TECHNOLÓGIA 2019-ES GARTNER HYPE-CIKLUSA^[5]



A NYTK HOZZÁJÁRULÁSA AZ MI NYELVTECHNOLÓGIAI FELHASZNÁLÁSHOZ

- Elkészült a HiBERT (NYTK / PTE / Microsoft Hungary): a magyar nyelv olyan digitális nyelvmodellje (BERT-Large), amelyet eddig csak tíz nyelvre fejlesztettek ki a világon
- További elkészült nyelvmodellek: HIL-ELECTRA, HIL-ALBERT, HIL-RoBERTa, HIL-SBERT, HiBART
- A fentiek néhány finomhangolással készült demóverziója már elérhető:
 - ékezetesítő: <https://juniper.nytud.hu/demo/diacritic>
 - névelem-felismerő és anonimizáló: <https://juniper.nytud.hu/demo/ner>
 - szentimentelemző: <https://juniper.nytud.hu/demo/sentana>
 - szövegosztályozó: <https://juniper.nytud.hu/demo/tagger>
 - kivonatoló (extraktív és absztraktív): <https://juniper.nytud.hu/demo/summarization>
 - gépi fordító: <https://juniper.nytud.hu/demo/translation>
- A java pedig csak ezután jön...

FORRÁSOK

[1]

https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[2] <https://ec.europa.eu/digital-single-market/en/desi>

[3] Jamie Berryhil, Kévin Kok Heang, Rob Clogher, Keegan McBride (2019): „Hello, World: Artificial intelligence and its use in the public sector“, OECD Observatory of Public Sector Innovation, p.23. Available online at <https://oecd-opsi.org/projects/ai/>

[4]

<https://www.oecd-ilibrary.org/docserver/09ab162c-en.pdf?expires=1583412591&id=id&accname=guest&checksum=DE3AC27D9E81DA0C3A35448B23FCC54E>

[5]

<https://www.gartner.com/en/newsroom/press-releases/2019-08-28-gartner-2019-hype-cycle-shows-cloud-office-has-hit-ma>

TOVÁBBI IRODALOM

<https://trends.oecd-opsi.org/embracing-innovation-in-government-global-trends-2019.pdf>

https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

KÖSZÖNÖM A FIGYELMET!

WWW.NYTUD.HU

Website: www.lr-coordination.eu

Twitter: [@LR_Coordination](https://twitter.com/LR_Coordination)

Facebook: www.facebook.com/EuropeanLanguageResourceCoordination

LinkedIn: www.linkedin.com/in/lrcoordination

Email: info@lr-coordination.eu

