Eszter Baranyai
Senior Education and Research Expert, MNB

THE IMPACT OF LARGE LANGUAGE MODELS ON THE LABOUR MARKET: SPATIAL EVIDENCE FROM JOB ADS IN HUNGARY

The Impact of AI on the Macroeconomy and Monetary Policy: Joint conference of ESCB ChaMP Research Network and BdE

24 October 2024

Co-authors: Marcell Granát (MNB) and Mór Szepesi (MNB, Yale University)*

*With thanks to Malatinszky Gábor (MNB)

Disclaimer: The views expressed are those of the authors and do not necessarily reflect the official view of The Central Bank of Hungary.

OVERVIEW OF RESULTS

- 10% of workload could be substituted by LLMs that are at least twice as fast as humans and without a negative impact on quality. In the US this is 15%.
- LLMs are complementary for all job ads. Rarely does exposure exceed 30 per cent.
- Spatial differences in exposure: Of the factors investigated it is industry mix that matters the most. Positive correlation between LLM exposure on the one hand and proportion of young adults or share of population living in cities, on the other.

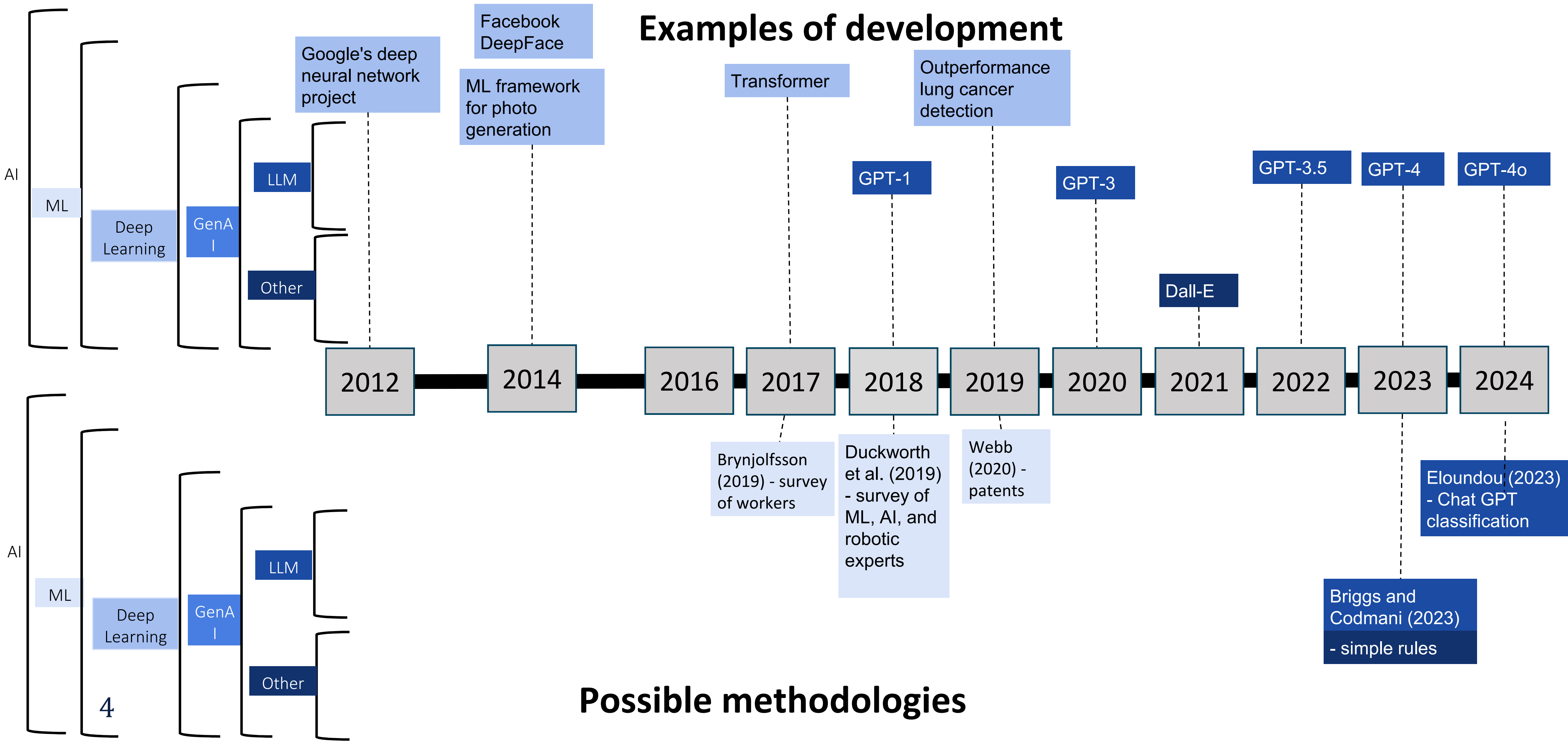
RESEARCH QUESTIONS

- What portion of current job vacancy work could be substituted by AI?
- What are the spatial patterns? What factors are the spatial patterns associated with?

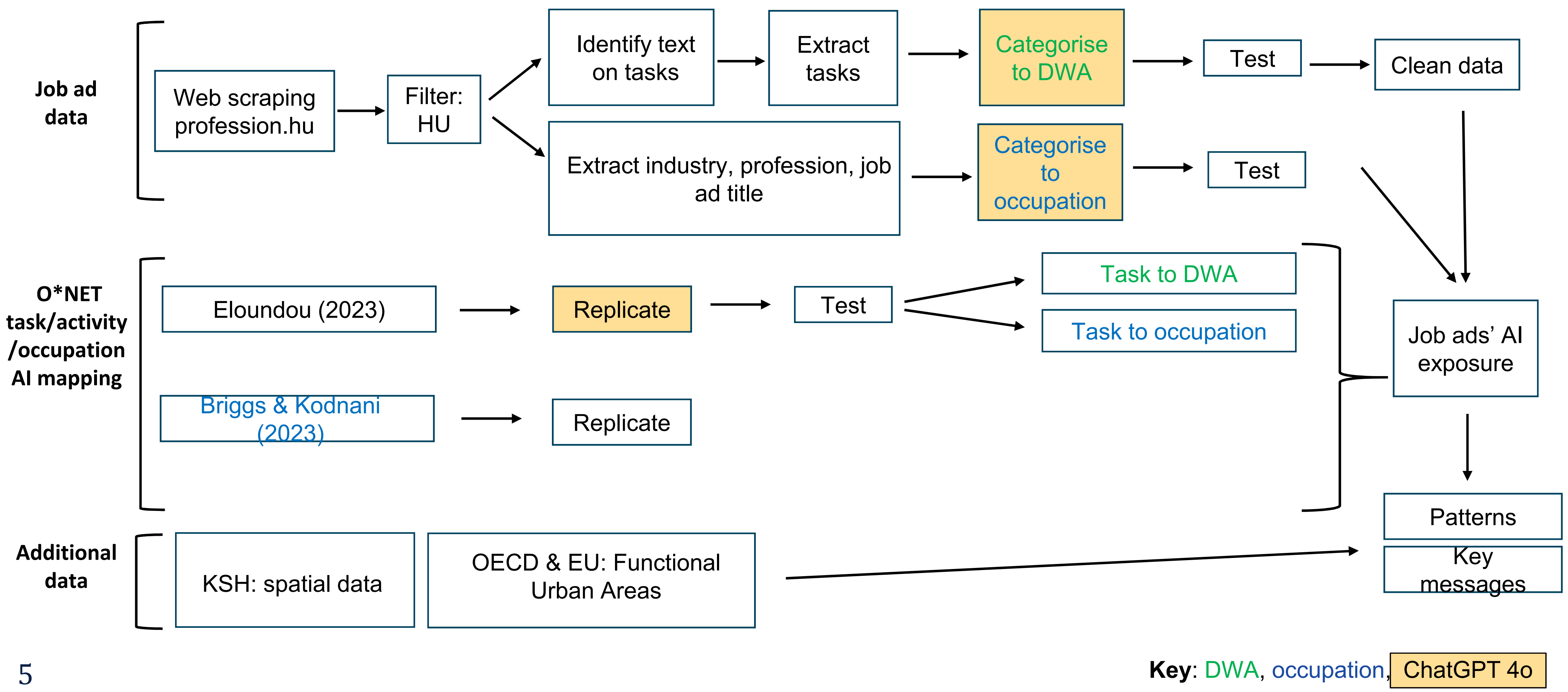
Novelty and use of our research:

- Focus on LLMs – much of the literature predates their rapid take-off
- Spatial focus
- Uses detailed job postings data from largest job portal
- Emerging Markets have been less studied

AI HISTORY OVERVIEW



RESEARCH PROJECT FLOW CHART



DATA – JOB PORTALS

Job portal	No of ads (16 Jan 2024)
Profession.hu	**13,850**
Linkedin (ads in HU)	11,045
EURES (ads in HU)	5,774
CVOnline.hu	5,703
Jófogás	2,633
Jobline	1,561
Jooble (ads in HU)	12,240 (uses other portals)

Job vacancies (2023)	78,975

Job portal websites, KSH

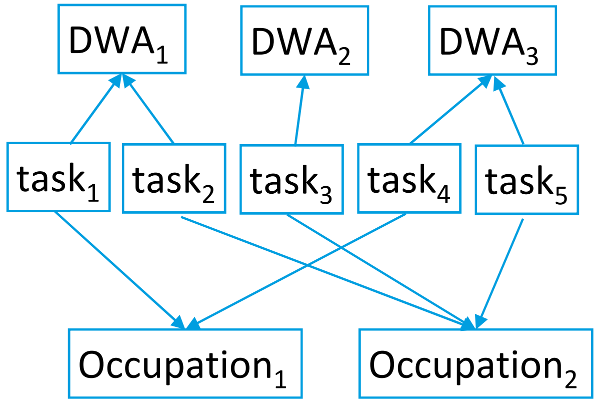
O*NET DATA

Detailed Work Activity example: Classify organisms based on their characteristics or behavior.

c. 2,000

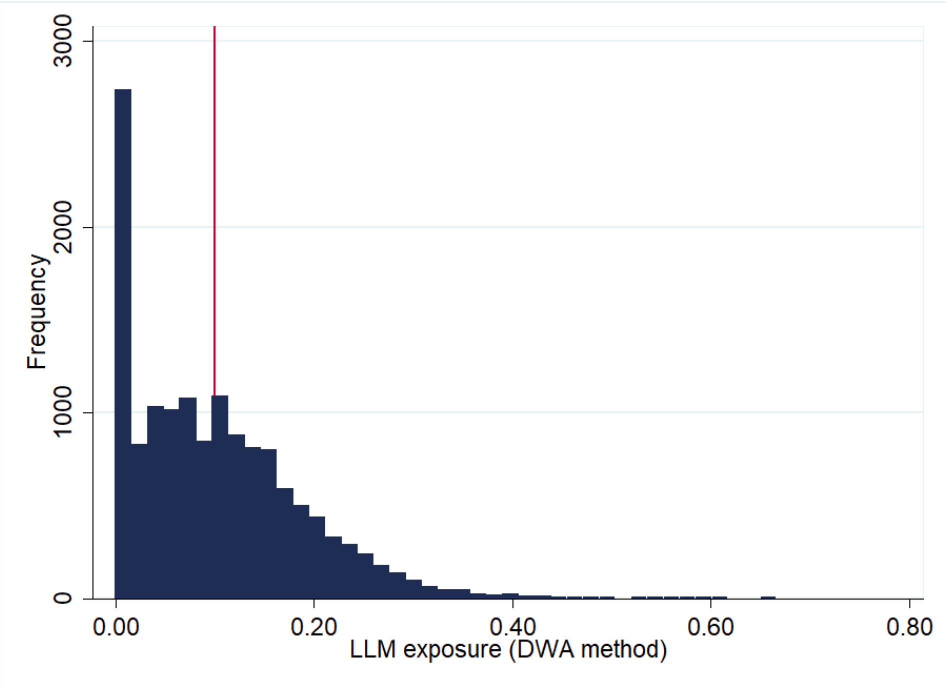
Task example: Review, classify, and record survey data in preparation for computer analysis. > 20,000

Occupation example: Survey researcher c. 1,000

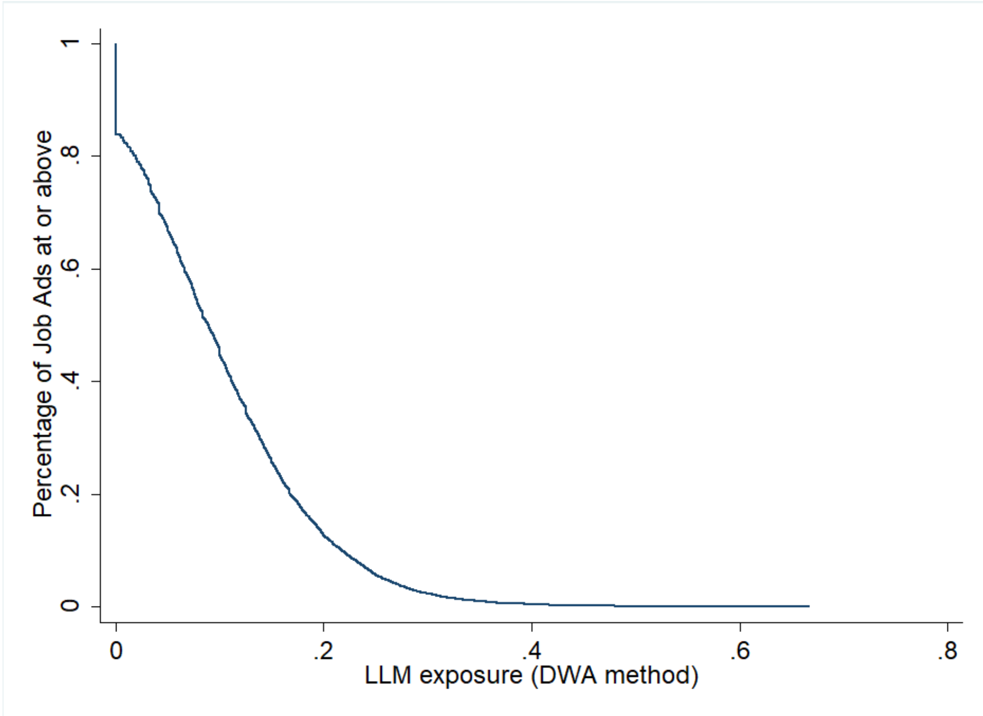


RESULTS: LLM COMPLEMENTS

Distribution of LLM exposure

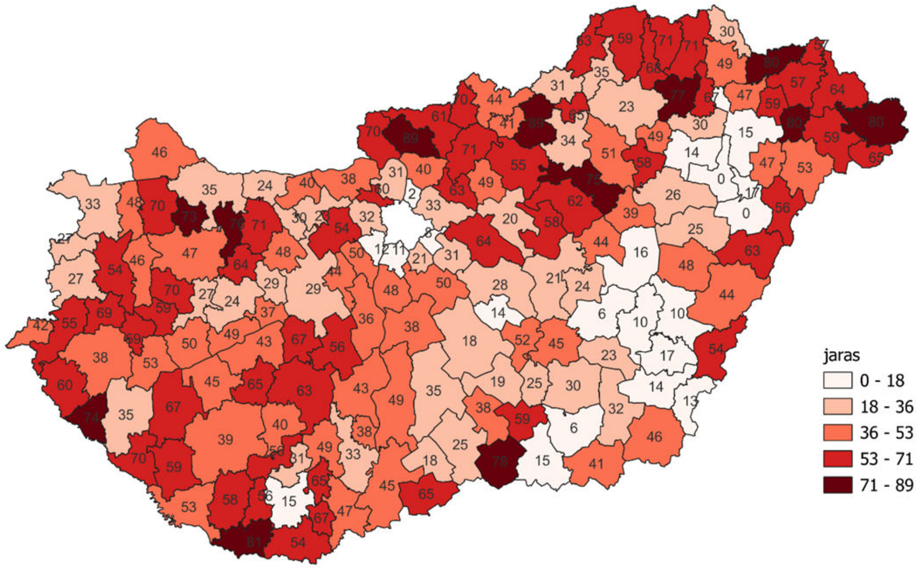


Share of Job ads with LLM exposure> x



LLM EXPOSURE IS CORRELATED WITH THE TYPE OF SETTLEMENT

Share of villages by district



Source: KSH

LLM exposure by type of settlement*

*statistically significant differences at 0.01 level between:
i) villages and cities and ii) capital and other cities.

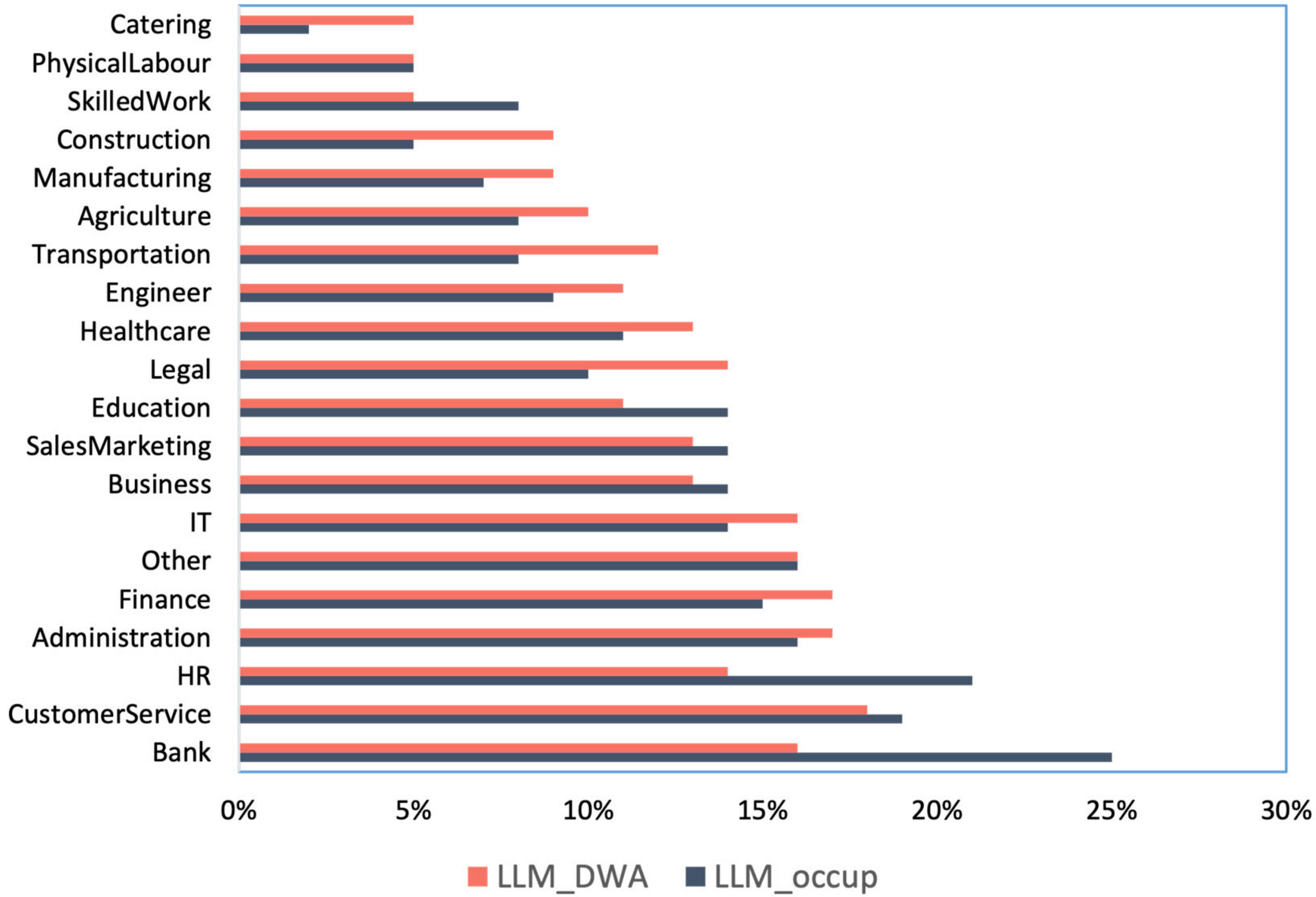
RESULTS: INDUSTRY MATTERS

Linear regression with OLS estimation: LMM exposure (DWA method)

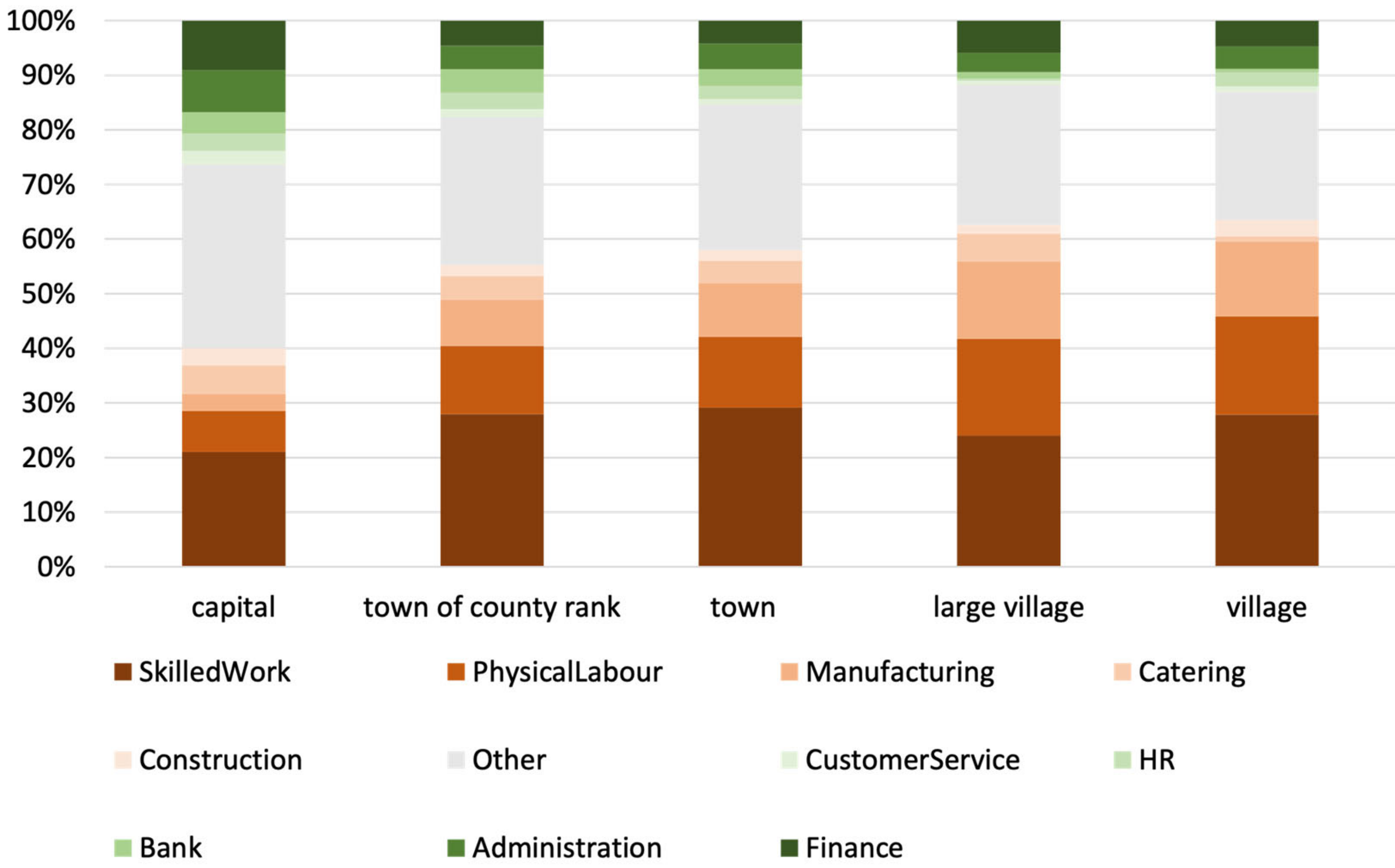
	(1)	(2)	(3)	(4)
Village dummy	**-0.69**** (0.30)	**-0.73*** (0.38)	-0.33 (0.25)	
Young% (20-30 over 20-60)	**0.19***** (0.04)	**0.17***** (0.05)	0.00 (0.03)	
LogEarnings	**3.94***** (0.47)	-1.18 (1.08)	0.12 (0.42)	
Constant	-46.19*** (6.2)	21.06 (14.11)	7.56 (5.54)	9.10*** (1.01)
Industry dummies	N	N	Y	Y
Budapest FUA excluded?	N	Y	N	N
Observations	13 200	4 821	13 200	13 200
Prob>F	0.00	0.00	0.00	0.00
R2	0.01	0.00	0.29	0.29

Robust standard errors in brackets

DIFFERENCES IN EXPOSURE TO LLM ACROSS INDUSTRIES

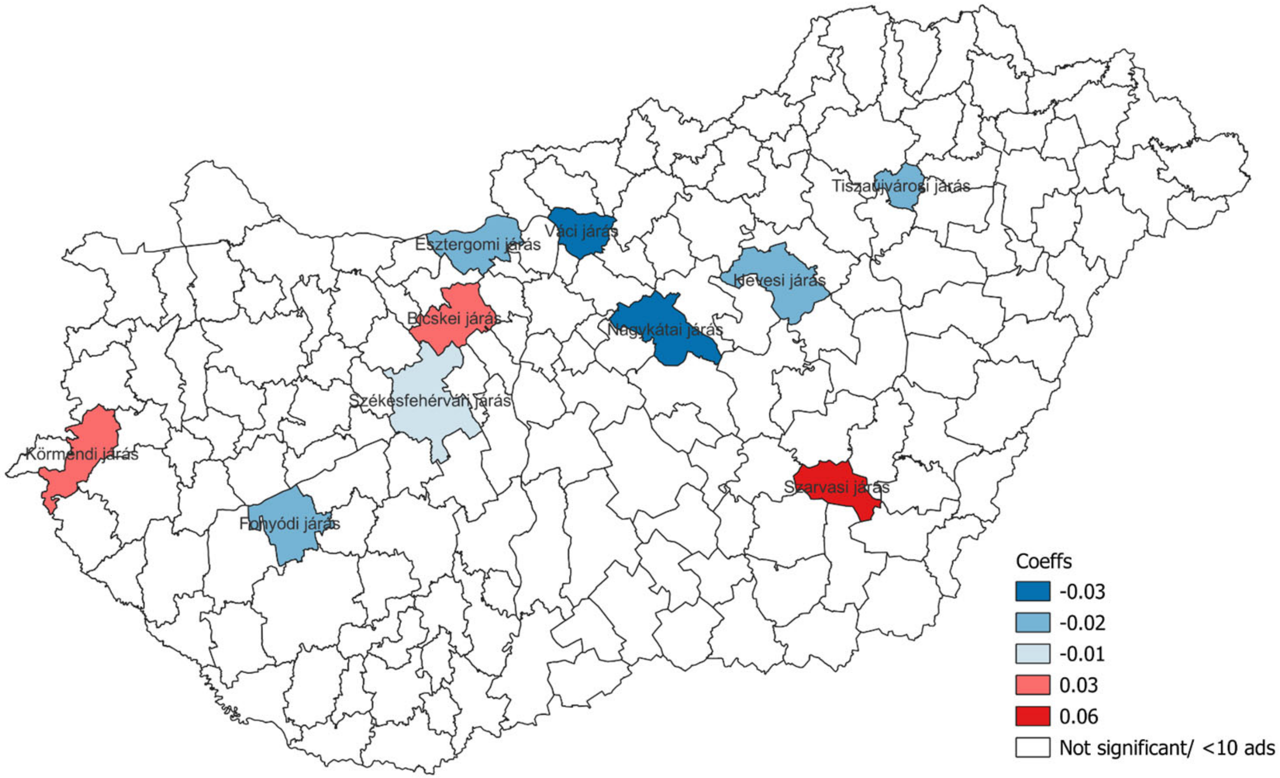


SETTLEMENT TYPE & INDUSTRY MIX



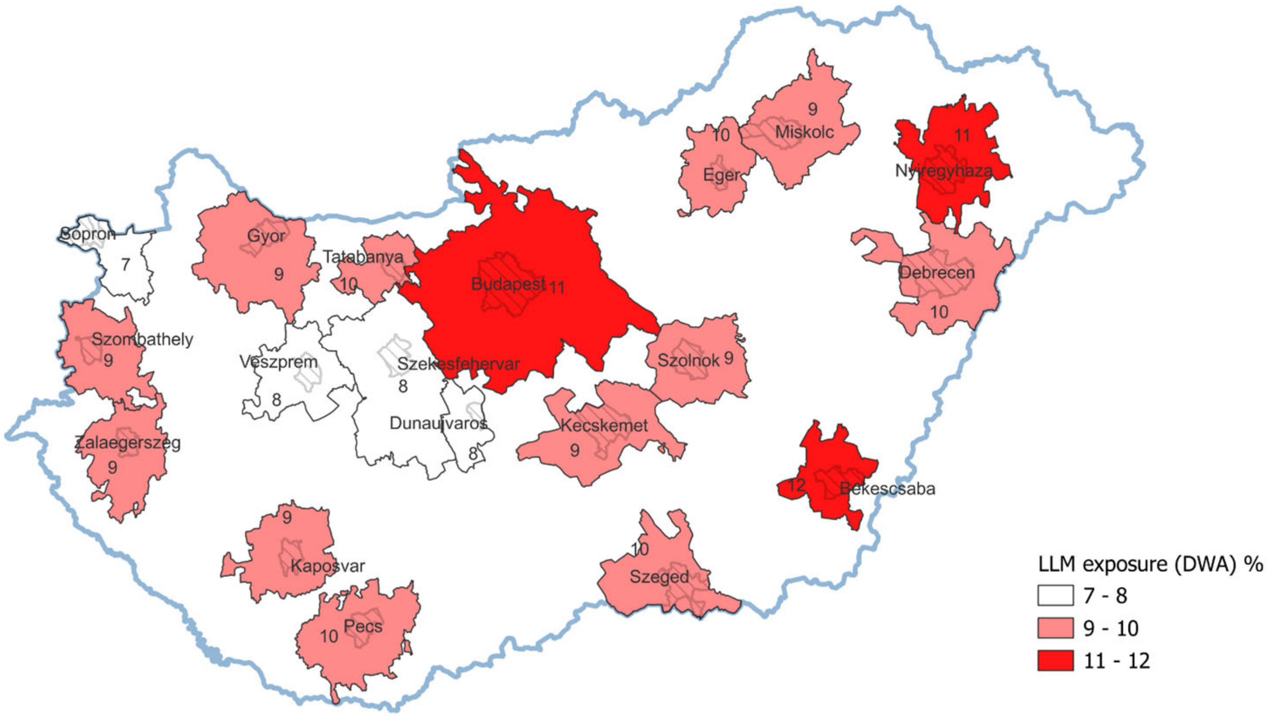
DIFFERENCES IN EXPOSURE VS INDUSTRY MIX

Where is LLM exposure statistically different from what the district's industry mix would imply?



Residuals from a regression where y= LLM exposure and x= industries, are then regressed on districts. This map depicts coefficients from the second regression.

EXPOSURE BY FUNCTIONAL URBAN AREA



- Some cities with a complex industrial profile in the Western part of Hungary close to highways leading to Austria have significant production operations – and higher share of sectors with comparatively low LLM exposure.
- Some cities in the East have comparatively more job offers in more LLM-exposed sectors such as Administration and Banking.

Functional Urban Areas use population density and travel-to-work flows to demarcate areas where labour market is highly integrated. Source: OECD & European Commission and own calculations

COMPARISON WITH THE LITERATURE

Topic	Study	Country	Statement	Our findings
Complement vs substitute	Eloundou et al. (2023) Briggs & Kodnani (2023)	US	Rare to find any occupation for which LLMs could do nearly all the work.	Agree.
% Exposure	Eloundou et al. (2023)	US	At least 10% (50%) of work tasks affected by LLMs for 80% (20%) of US workforce.	This is true for 45% (0%) of the Hungarian workforce.
	Briggs & Kodnani (2023)	World	18% of work globally could be automated by ***genAI.***	We also find 18% for ***genAI.***
	Briggs & Kodnani (2023)	US	2/3 of US occupations are exposed to ***genAI***, most have a 25-50% exposure.	83% are exposed to ***genAI*** >0.05, 69% exposed>0.1. Almost all between 0-0.4.
What/who exposed	Eloundou et al. (2023)	US	Information processing industries exhibit high exposure, while manufacturing, agriculture, and mining demonstrate lower exposure.	Similar. Catering, physical labour and construction also low.
Geographical patterns	Hamaguchi (2018)	Japan	Women especially in larger cities more exposed to ***computerization*** (receptionist, clerical work, sales).	Larges cities higher exposure (**LLM**)
	Frank (2018)	US	Lower potential for **automation** in big cities rather than small (due to managerial, technical professions)	Larger settlement types more exposed to **LLM**s
	Hat (2020)	Austria	urban areas and small towns are relatively less exposed than rural areas to **digitalisation**	Larger settlement types more exposed to **LLM**s

LIMITATIONS

- Task/ DWA aggregation to job – mostly simple add-up of tasks, or core/supplementary (no sophisticated weighing)
- Based on current technology (may change soon given rate of development)
- Largely one technology (LLM)
- Looks at technological feasibility, not whether it is economically feasible, doesn’t consider security concerns, etc
- Job portal data not representative of jobs available especially rural blue-collar jobs
- Current job ad task descriptions may reflect intention to hire humans. This may change.

RESULTS

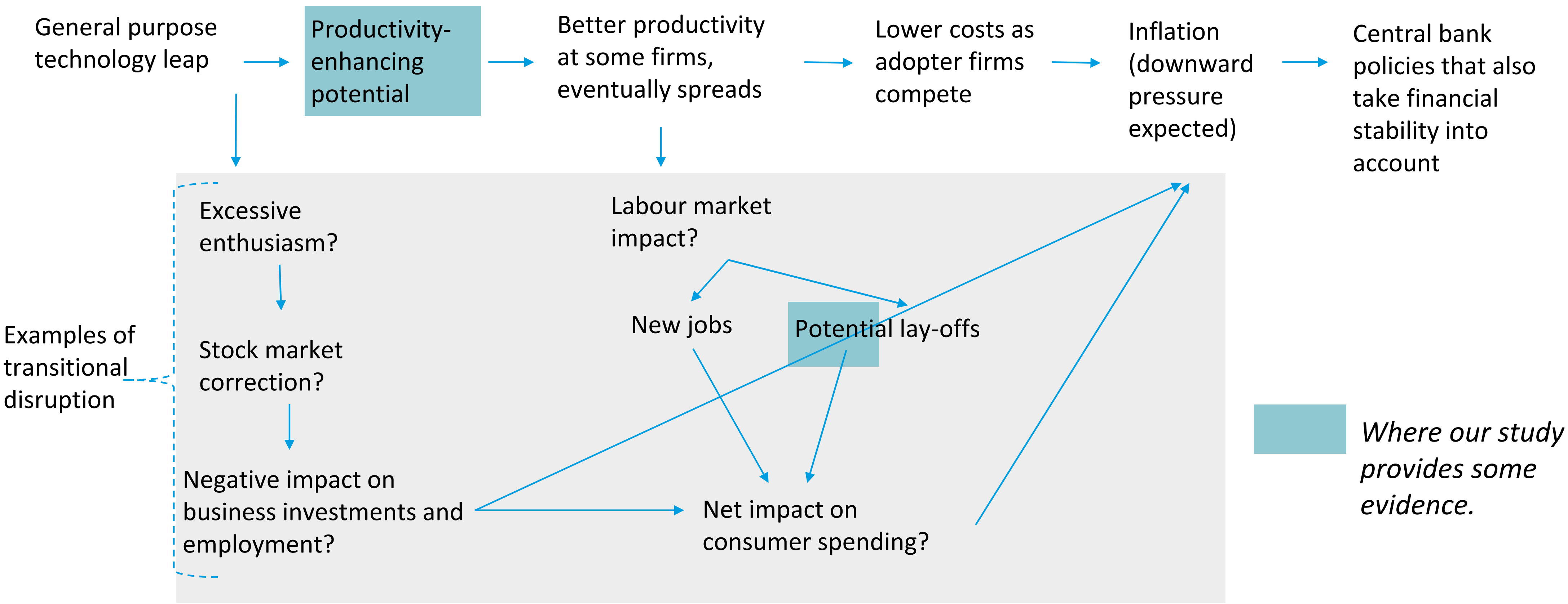
- 10% of workload could be substituted by LLMs that are at least twice as fast as humans and without a negative impact on quality. In the US this is 15%.
- LLMs are complementary for all job ads. Rarely does exposure exceed 30 per cent.
- Spatial differences in exposure: Of the factors investigated it is industry mix that matters the most. Positive correlation between LLM exposure on the one hand and proportion of young adults or share of population living in cities, on the other.

POSSIBLE IMPLICATIONS

- LLM could be used to improve the productivity of workers.
- Labour market: net effect unclear but LLM complementary to all jobs. Labour market and education policies to ensure benefits are reaped and impact on employment managed.
- Follow-on project: calculate estimate for possible productivity effect -> monetary policy implications.
- Development across many technologies, not just LLM. Nonetheless, do spatial differences in LLM exposure translate to an impact on regional productivity trends (within and across nations)?
- Labour market policies: importance of industries as industry mix is what LLM exposure appears most closely related to.

TECHNOLOGICAL LEAP AND MONETARY POLICY

Stylised chart



Based on Poloz (2021) and authors

100 éve Magyarország gyarapodásáért

THANK YOU FOR YOUR ATTENTION!

ADDITIONAL SLIDES

100 éve Magyarország gyarapodásáért

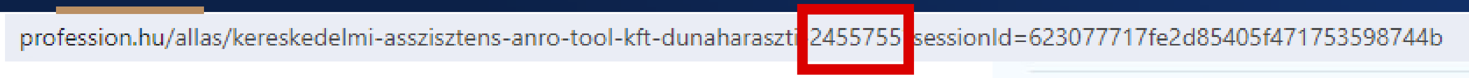
DATA - SAMPLE

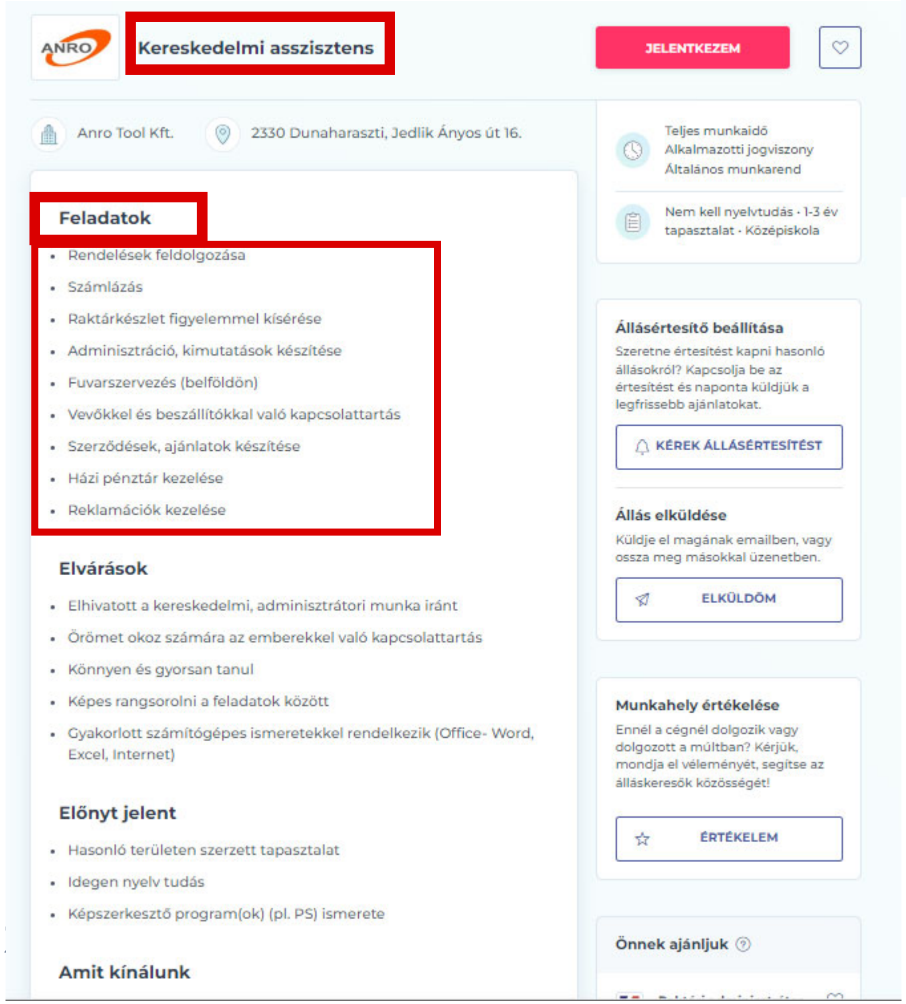
area name	% sample	% national	delta
Central Hungary	**60**	**55**	**5**
Budapest	46	45	1
Pest	14	10	4
Southern Transdubania	**4**	**4**	**0**
Baranya	1	1	0
Somogy	1	2	-1
Tolna	1	1	0
Northern Hungary	**5**	**5**	**0**
Borsod-Abaúj-Zemplén	2	3	-1
Heves	2	1	1
Nógrád	0	1	-1
Northern Great Plain	**7**	**8**	**-1**
Hajdú-Bihar	3	3	0
Jász-Nagykun-Szolnok	2	3	-1
Szabolcs-Szatmár-Bereg	2	2	0

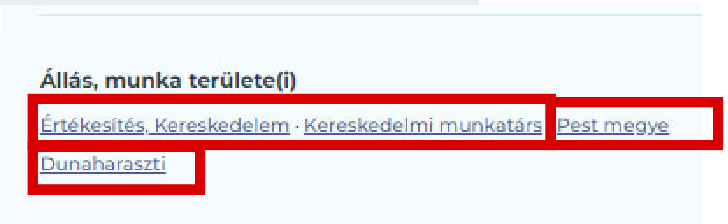
area name	% sample	% national	delta
Western Transdubania	**7**	**9**	**-2**
Győr-Moson-Sopron	4	5	-1
Vas	2	2	0
Zala	1	2	-1
Southern Great Plain	**5**	**7**	**-2**
Bács-Kiskun	2	3	-1
Békés	1	2	-1
Csongrád-Csanád	2	2	0
Central Transdubania	**9**	**12**	**-3**
Fejér	4	6	-2
Komárom-Esztergom	3	4	-1
Veszprém	2	2	0
Missing county data	**4**	**-**	**-**
Overall	**100**	**100**	**-**
n	**15,124**	**78,975**	**-**

Profession.hu, KSH

DATA – WEB SCRAPING AND TEXT EXTRACTION



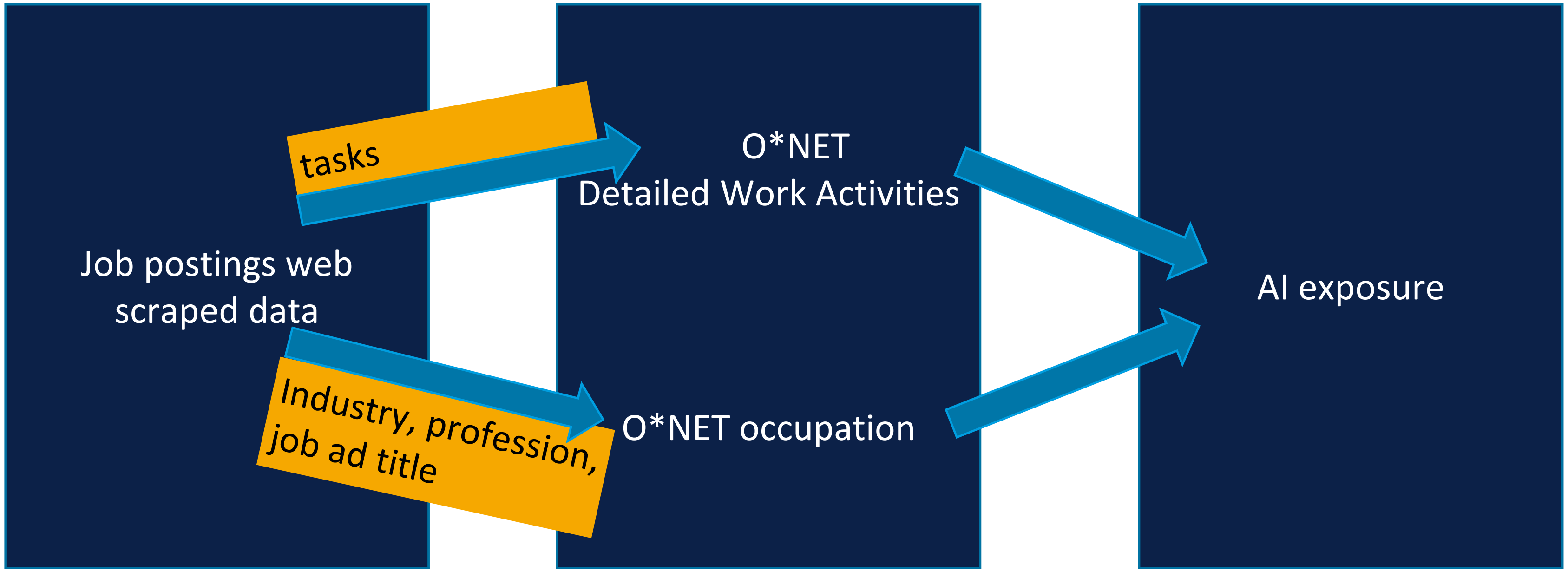




- Extract: county, settlement, job ID, industry, profession, job ad title
- Identify tasks headings: manual identification of 467 titles from c. 2000
- Extract text under titles

Chart 3: overview of mapping

MAPPING: VIA TWO ROUTES



JOB AD TASKS TO DWA-S: EXAMPLE

Job ad description of tasks

['Fermentációval előállított fehérjeoldat kicsapással, ultraszűréssel végzett tisztítási lépéseinek végrehajtása'
'A tisztított oldat steril szűrése és letöltése zárt letöltő rendszerrel'
'Az automata szekvenciákkal működő gyártóberendezések kezelése'
'Mobil tartályok mozgatása, csatlakoztatások és automata tartályműveletek irányítása elektronikus kezelőpanelen'
'Gyártásközi minőség-ellenőrzési minták levétele és az előírt tesztek elvégzése és dokumentálása (számítógépes rendszerekben is)'
'Az elektronikus illetve papír alapú gyártásdokumentáció egyidejű és pontos vezetése'
'A gyártóterület tisztán tartása, a rendezett és tiszta állapot ellenőrzése'
'Munkarend:'
'12 órás nappalos műszak, reggel 6 órától este 6 óráig tartó munkavégzéssel (3 nap munkanap, 3 nap szabadnap)']

ChatGPT →

O*NET DWA

- Test materials, solutions, or samples.
- Clean equipment or facilities.
- Collect samples of materials or products for testing.
- Document operational procedures.
- Operate industrial equipment.
- Perform manual service or maintenance tasks.

Manual testing and correction on random sample (600 job ads, 4% of population)

1. Corrected sample – override (c. 450)
2. Correct all data based on patterns identified *(erroneous "nones", hallgatói jogviszony, grant, conduct market research, manage professional relationships, perform clerical work in medical settings, direct operations of correctional facilities)*

JOB AD TASKS TO DWA-S: RESULTS OF TESTING

Tester	How correct (%)?*	N
Overall	88	600
Tester 1	94	300
Tester 2	82	300

*Includes the removal of nones but no other adjustments

Typical errors include:

- when work activity is right but context wrong (e.g. perform clerical activities in a medical setting): correct for most common mistakes
- Sometimes erroneously categorises as “None” – review all Nones manually and erroneous ones recategorised partly manually (post), partly by ChatGPT
- “Develop professional relationships or networks” often missing – corrected based on words

JOB AD TO OCCUPATION

Job ad title, industry, profession

Telefonos közvéleménykutató – Diákmunka ; Marketing; Marketing

Mérlegképes könyvelő; Pénzügy; Könyvelés

ChatGPT →

O*NET occupation

Market Research Analysts and Marketing Specialists

Accountants

↓

Manual testing on random sample (10% of population)

Confidence Level	Tester 1			Tester 2			Average	
	Disagree	Sample	%	Disagree	Sample	%	Disagree	%
High	11	1513	1	9	1513	1	10	1
Med	33	1513	2	17	1513	1	25	2
Low	35	1513	2	93	1513	6	64	4
Total	79	1513	5	119	1513	8	99	7

Some examples of disagreement (M and H Confidence):

- Robothegesztő : Welders, Cutters and Welder Fitters (does not include “robot");
- Uszodagépészeti-, vízgépészeti szerelo: Plumbers; (deemed to broad)

LLM EXPOSURE

Definition: can decrease the time required to complete the DWA or task by at least half (50%) with no deterioration in quality

Occupation, task

Education Teachers, Postsecondary: Write grant proposals to procure external research funding.

ChatGPT →

Exposed

Testing results of 921 tasks 5% of O*NET task population

Disagree with Chat GPT with at least M confidence			
	Not exposed to exposed	Exposed to not exposed	Total
Tester 1	2.5%	0.3%	2.8%
Tester 2	3.5%	1.4%	4.9%
Tester 3	0.1%	0.7%	0.8%
Average	2%	0.8%	2.8%

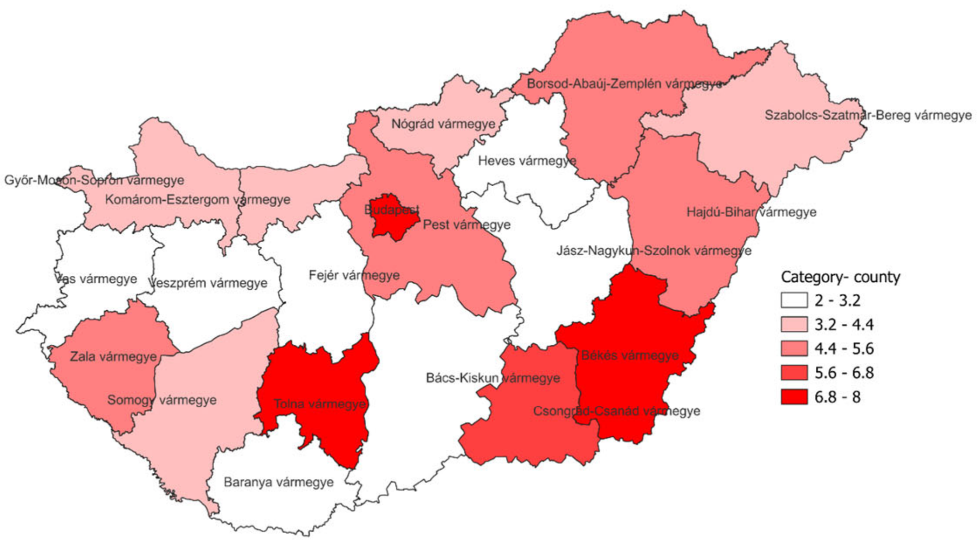
Testers not LLM experts.

Other examples of exposed:

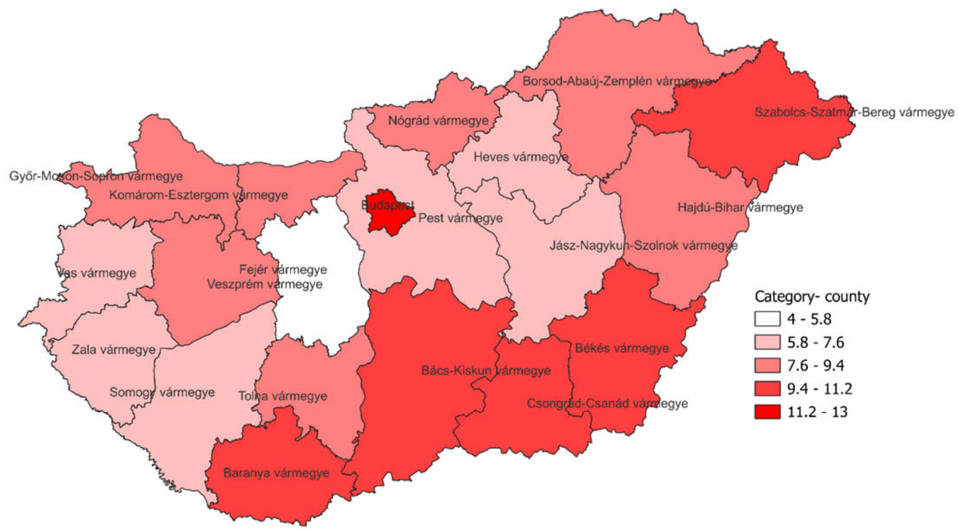
- Communicate with customers, employees, and other individuals to answer questions, disseminate or explain information, take orders, and address complaints.
- Collect business intelligence data from available industry reports, public information, field reports, or purchased sources.

SECTORS WITH HIGHER LEVELS OF LLM EXPOSURE

Share of Administration in sample

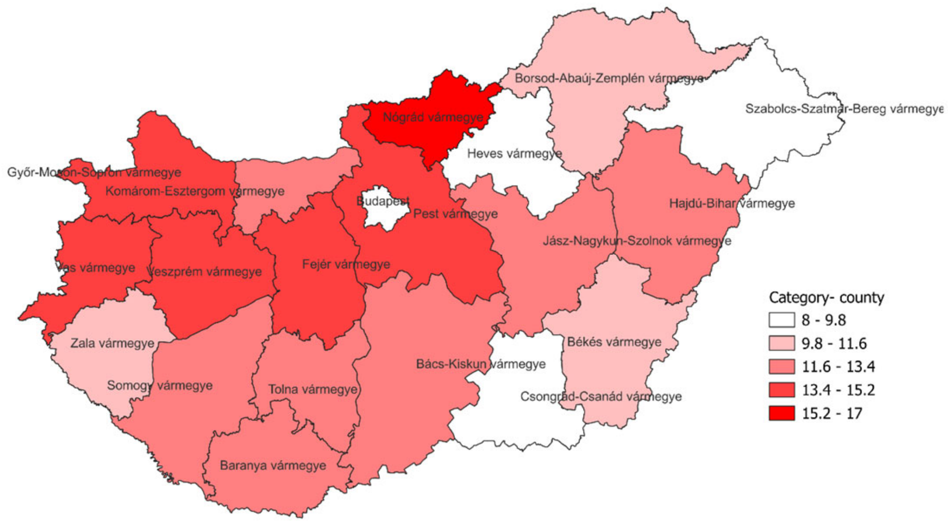


Share of Bank and Finance in sample

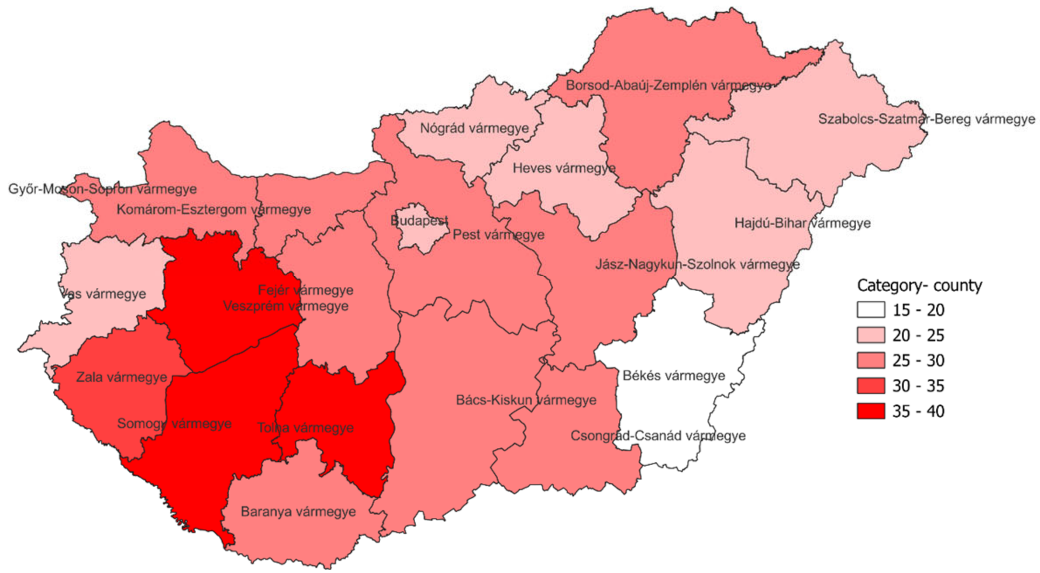


SECTORS WITH LOWER LEVELS OF LLM EXPOSURE

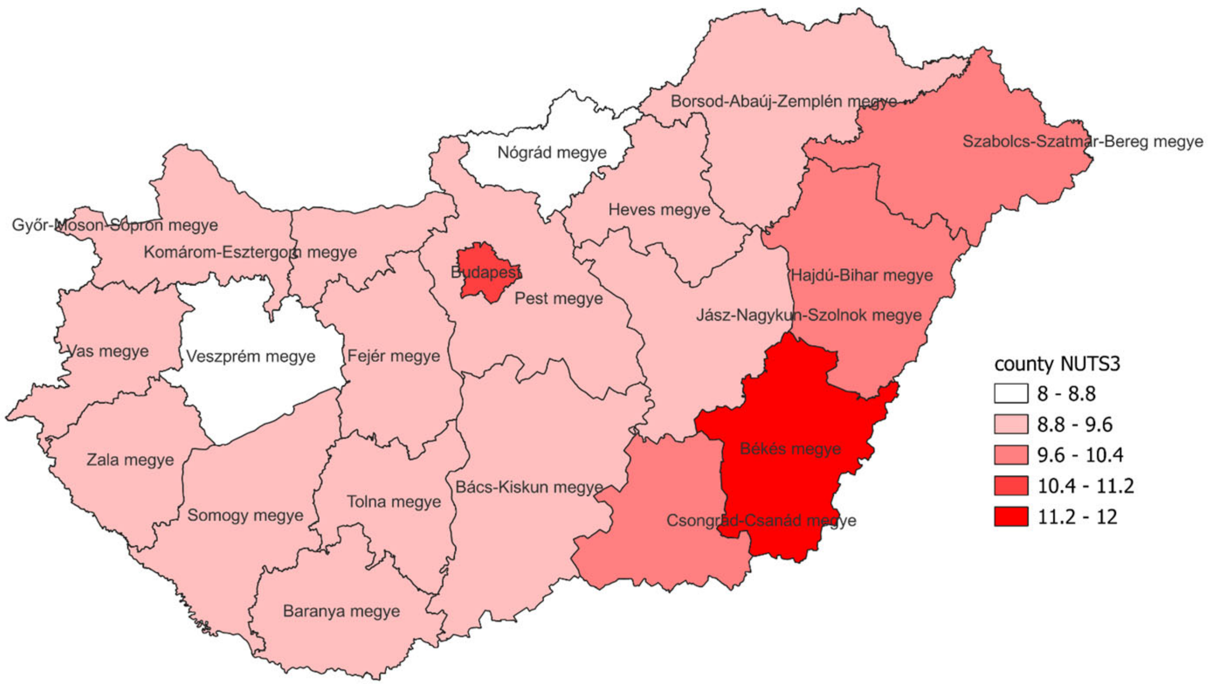
Share of Physical Labour in sample



Share of Skilled work in sample

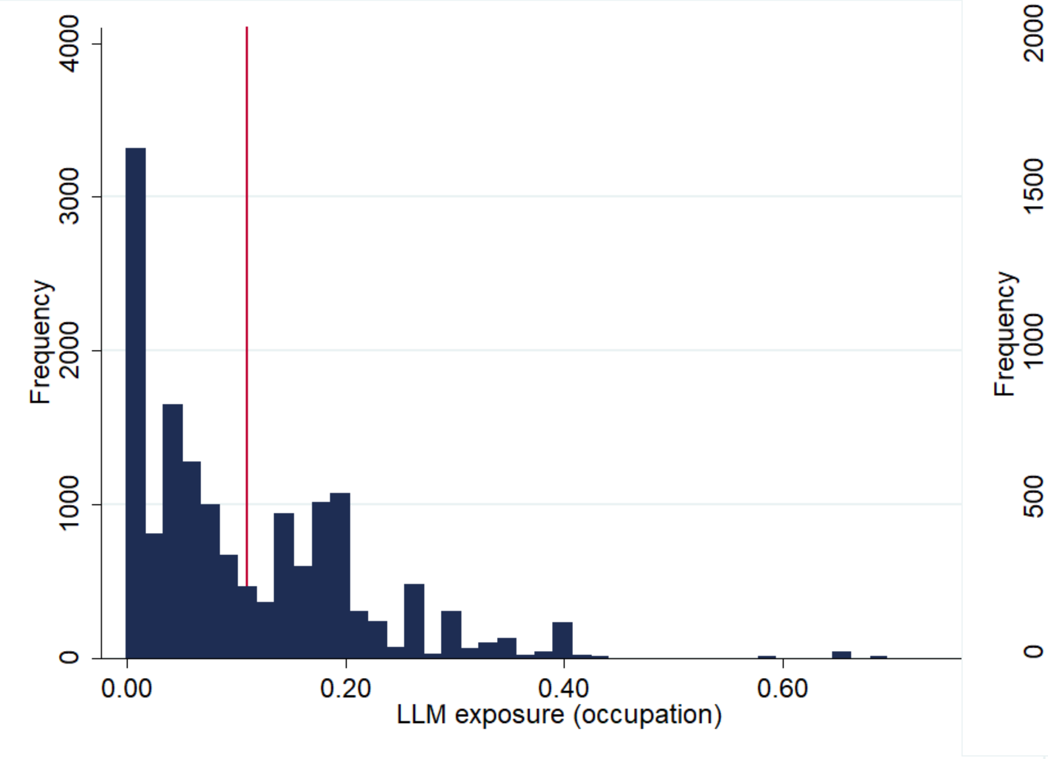


LLM EXPOSURE (DWA METHOD)

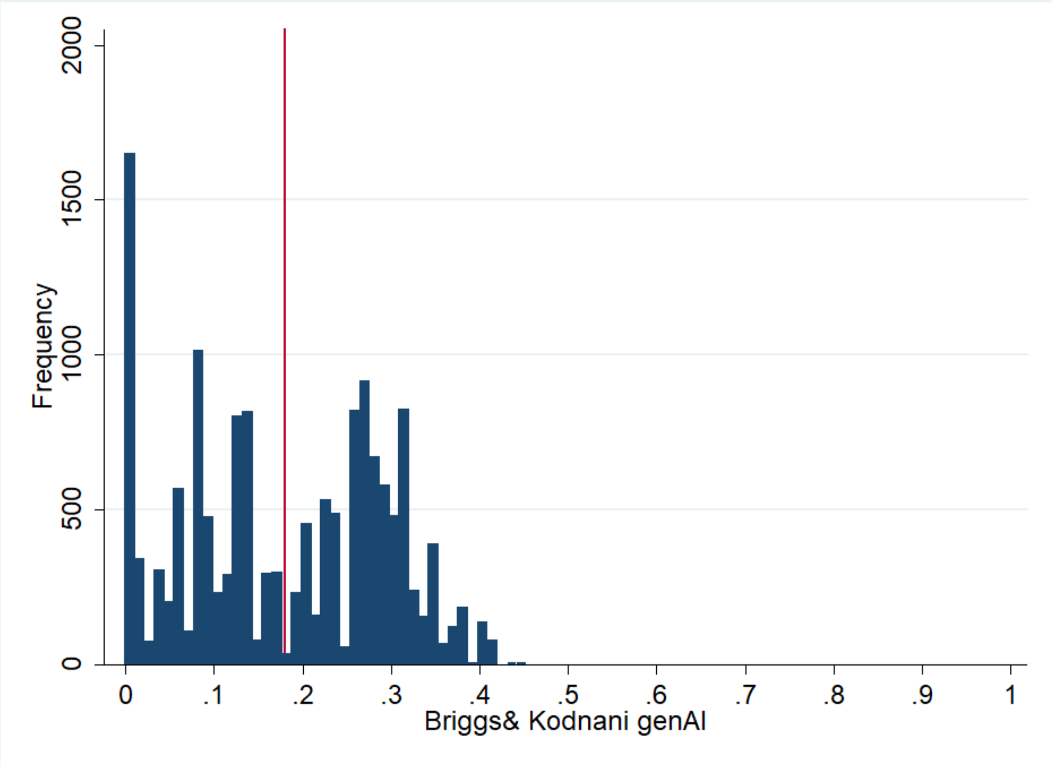


ROBUSTNESS CHECKS

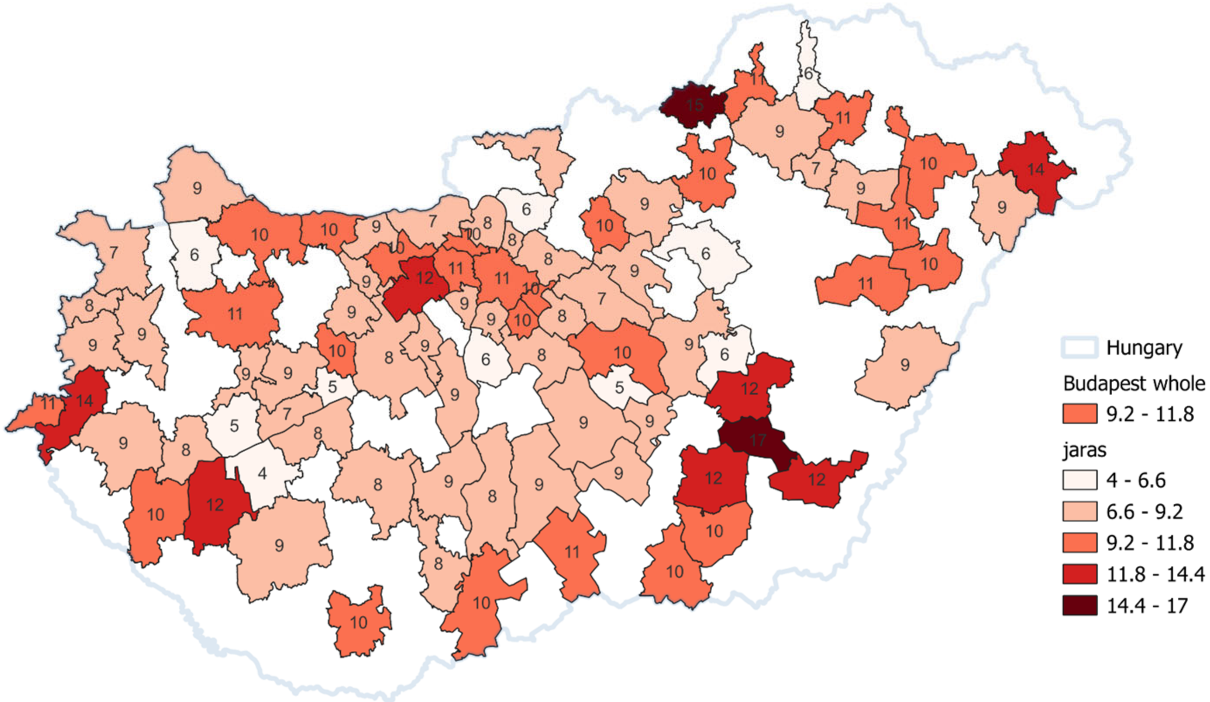
LLM using occupation method



Broader genAI using simple method



LLM EXPOSURE (DWA METHOD)



SOURCES

Briggs and Kodnani (2023). The Potentially Large Effects of Artificial Intelligence on Economic Growth.

Brynjolfsson, E., Mitchell, T., & Rock, D. (2018, May). What can machines learn and what does it mean for occupations and the economy?. In AEA papers and proceedings (Vol. 108, pp. 43-47). 2014 Broadway, Suite 305, Nashville, TN 37203: American Economic Association.

Brynjolfsson, E., Rock, D., & Syverson, C. (2019). Artificial intelligence and the modern productivity paradox. The economics of artificial intelligence: An agenda, 23, 23-57. https://www.nber.org/system/files/working_papers/w24001/w24001.pdf

Csomós, Gy (2015). The ranking of cities as centres of the Hungarian economy, 1992-2012. Regional Statistics 5 (1): 66–85. DOI: 10.15196/RS05104

Duckworth, P., Graham, L., & Osborne, M. (2019, January). Inferring work task Automatability from AI expert evidence. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 485-491). https://www.robots.ox.ac.uk/~mosb/public/pdf/4904/Duckworth%20et%20al.%20-%202019%20-%20Inferring%20Work%20Task%20Automatability%20from%20AI%20Expert%20.pdf

Eisfeldt, A. L., Schubert, G., & Zhang, M. B. (2023). Generative AI and firm values (No. w31222). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w31222/w31222.pdf

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, 2023, Gpts are gpts: An early look at the labor market impact potential of large language models, arXiv preprintarXiv:2303.10130 . https://arxiv.org/abs/2303.10130

Frank, M. R., Sun, L., Cebrian, M., Youn, H., & Rahwan, I. (2018). Small cities face greater impact from automation. Journal of the Royal Society Interface, 15(139), 20170946.

Hamaguchi, N., & Kondo, K. (2018). Regional employment and artificial intelligence in Japan. RIETI.

Hat, K., & Stoeglehner, G. (2020). Spatial dimension of the employment market exposition to digitalisation—The case of Austria. Sustainability, 12(5), 1852.

Poloz, Stephen S. (2021) Technological progress and monetary policy: managing the fourth industrial revolution. Journal of International Money and Finance 114, 102373.

Webb, M. (2019). The impact of artificial intelligence on the labor market. Available at SSRN 3482150. https://www.michaelwebb.co/webb_ai.pdf

METHODS WE USE FOR O*NET TO AI EXPOSURE

Study	Source	What it measures	Basis
Eloundou et al. (2023); Eisfeldt et al. (2023)	Chat GPT	Gen AI. Exposure to **large language models** (LLMs) or future applications based on LLMs	Task (to be mapped to DWA and Occupation)
Briggs and Kodnani (2023)	Author categorisation, simple (13 broad categories where difficulty <= 4)	**Gen AI**	Occupation

HOW TO INTERPRET SCALES?

Study	Scale	Interpretation
Eloundou et al. (2023); Eisfeldt et al. (2023)	Task: 1 or 0. DWA: % of tasks “1” Occupation*: weighted % of tasks “1”	% of workload where LLM reduces time by at least 50pc with no deterioration in quality
Briggs and Kodnani (2023)	(0 to 100) %	% of each occupation’s workload that genAI has the potential to replace

*Core tasks carry double weight compared to supplementary tasks

WHAT WE LOOK AT (STYLISED EXAMPLE)