

AI-FOR-HEALTH MATURITY IN EU MEMBER STATES A WHO-INFORMED EUROSTAT ANALYSIS

Piotr JANKE^{1*}, Mirosław MATUSEK²

¹ Silesian University of Technology, Faculty of Organization and Management; piotr.janke@polsl.pl,
ORCID: 0000-0001-8065-9013

² Silesian University of Technology, Faculty of Organization and Management; miroslaw.matusek@polsl.pl,
ORCID: 0000-0002-6681-8265

* Correspondence author

Purpose: This paper operationalizes AI-for-health maturity in EU Member States by translating WHO-informed evidence into a reproducible analytical outcome linked to harmonized Eurostat indicators.

Design/methodology/approach: A country-year panel for the EU27 covering 2021-2024 was constructed by combining a WHO-informed maturity outcome with Eurostat-based predictors capturing innovation and R&D, data capability, health-system resources and spending, and digital access and readiness. Missing values were handled using a deterministic three-stage procedure based on within-country interpolation, forward/backward fill, and year-specific EU median imputation. The analytical workflow included descriptive screening, correlation analysis, Kruskal-Wallis tests, and a cautious predictive extension using grouped cross-validation by country.

Findings: The final panel comprised 108 country-year observations, 15 numeric predictors, and two WHO-derived outcomes: a continuous maturity score and an ordinal maturity class. The strongest descriptive associations were found for innovation- and R&D-related indicators, data-system capability, digital access, and selected health-system financing variables. Several predictors were highly collinear, indicating a tightly related innovation-capacity block. Elastic Net performed best in prediction, but explained variance remained moderate.

Research limitations/implications: Limitations include the short time horizon, small sample, missingness, deterministic imputation, and an author-defined scoring procedure.

Practical implications: The workflow provides a transparent basis for benchmarking AI-for-health maturity across European health systems.

Social implications: The study may inform policy discussion on equitable digital transformation and trustworthy AI adoption in health.

Originality/value: The paper offers a replicable EU27 framework for comparative research on AI readiness in European health systems.

Keywords: AI for health; digital maturity; Eurostat; EU27; health system readiness.

Category of the paper: Research paper.

1. Introduction

Artificial intelligence (AI) is increasingly shaping how health systems plan services, allocate resources, and deliver care, while simultaneously raising challenges related to governance, ethics, responsibility, and safeguards for trustworthy adoption. Across the WHO European Region, countries differ substantially in strategic and operational context, stakeholder engagement and workforce capacity, regulatory and legal frameworks, health data governance, AI applications, and perceived barriers to uptake. While such heterogeneity is well documented in narrative reporting and country profiles, descriptive evidence is not always immediately usable for quantitative comparative analysis or for predictive modelling workflows that require harmonized, machine-readable inputs (*Artificial Intelligence for Health in the WHO European Region*, 2025).

To bridge this gap, researchers often rely on “maturity” constructs that translate complex, multi-domain system readiness into structured indicators that can be compared across units and over time (Liaw et al., 2023). Digital health capability maturity models emphasize staged development and socio-technical foundations, and they also highlight that maturity can evolve dynamically, requiring clearly defined and reproducible measurement approaches (Liaw et al., 2022). At the same time, the practical analytic value of a maturity outcome depends on linking it to harmonized predictors that capture national differences in human capital, innovation capacity, health system resources, and digital infrastructure (*Artificial Intelligence for Health in the WHO European Region*, 2025; *Artificial Intelligence Is Reshaping Health Systems: State of Readiness across the WHO European Region*, 2025).

In parallel, the regulatory environment for AI in health is rapidly evolving and increasingly salient for cross-country comparisons, particularly in Europe where legal and policy instruments shape the conditions for deployment, accountability, and risk management (Frontiers in Public Health, 2024). This makes transparent operationalization especially important: even when an outcome such as AI-for-health “maturity” is defined, results can be sensitive to indicator definitions, dimension filters (e.g., units, sectors, population groups), and data availability constraints (*Artificial Intelligence Is Reshaping Health Systems: State of Readiness across the WHO European Region*, 2025; Eurostat, 2017).

For European comparative work, Eurostat provides a particularly suitable source of candidate predictors because it disseminates standardized indicators with metadata and programmatic access through official SDMX APIs, enabling transparent extraction and reproducible data pipelines (Eurostat, 2022). Building a country–year dataset from multiple Eurostat domains nevertheless involves practical challenges, including incomplete coverage for some indicators and inconsistent availability across countries and years; documenting missingness and applying explicit inclusion rules are therefore necessary steps prior to exploratory analysis and model development (Eurostat, 2017, 2022).

This study develops a reproducible EU27 country–year dataset for 2021–2024 that combines a WHO-informed AI-for-health maturity outcome (continuous score and ordinal maturity class) with candidate predictors retrieved from Eurostat using pre-specified dimension filters to preserve cross-country comparability (*Artificial Intelligence for Health in the WHO European Region*, 2025; Eurostat, 2022). We then characterize missingness and perform exploratory data analysis (EDA) to describe distributions, outliers, correlation structure, and potential separability of maturity classes, providing a transparent methodological foundation for subsequent work on country-generalizable classification and/or regression modelling of AI-for-health maturity (Eurostat, 2017; Liaw et al., 2023).

2. Methods

Recent peer-reviewed studies that aim to explain or predict country-level outcomes increasingly rely on multi-source data integration rather than single repositories. A common methodological pattern is to construct a “hybrid” dataset by merging heterogeneous national indicators (e.g., demographics, health status, policy measures, mobility, digitalization proxies) from multiple public sources into a single analytical table, followed by feature preprocessing, missing-data handling, and predictive modelling. For example, cross-European machine-learning work on COVID-19 outcomes explicitly builds a hybrid dataset from multiple public sources and then applies a modelling pipeline (feature preprocessing, dimensionality reduction/selection, predictive modelling, and explainability) to identify the country-level drivers of observed variation across waves (Moustakidis et al., 2022).

In organizational settings, analogous “two-dataset” designs are common: authors merge an external maturity measure with independent quality or performance outcomes and then quantify associations using regression or predictive models. A recent study of hospital digital maturity illustrates this approach by linking digital maturity metrics to quality and safety outcomes, using a multi-source design where the maturity construct and the outcomes originate from different data systems, thereby increasing interpretability while also introducing integration challenges (e.g., alignment of units, time windows, and missingness across sources) (Snowdon et al., 2024).

A third methodological stream focuses on data linkage as a way to enrich analysis beyond what a single dataset can provide. European evidence on linkage between survey data and administrative registers shows that linkage can materially increase analytic value, but it comes with governance constraints, cross-country heterogeneity in identifier availability, and systematic differences in coverage—issues that mirror the practical challenges of cross-country comparability and missingness encountered in macro indicator panels (Meltzer et al., 2022). In practice, this literature motivates explicit documentation of linkage keys (or, in macro

panels, merge keys such as country and year), as well as transparent reporting of completeness and potential bias introduced by non-uniform data availability.

Finally, some quantitative policy and readiness studies explicitly evaluate and optimize AI-related policy choices using formal quantitative frameworks and multi-source inputs, reinforcing the idea that methodological credibility in this area depends on (i) transparent integration of heterogeneous indicators, (ii) explicit missing-data strategies, and (iii) reproducible preprocessing decisions that can be audited and re-run when data are updated (Zhou et al., 2025).

Taken together, these peer-reviewed methodological examples support the core design choices in this study: integrating multiple sources into a country-year panel, documenting and controlling missingness, and using reproducible preprocessing to create a modelling-ready dataset that is suitable for comparative analysis and subsequent predictive workflows.

2.1. Datasets and scope.

We constructed a country-year panel for 27 EU Member States covering 2021–2024, combining an outcome variable derived from WHO-related digital health maturity evidence with candidate predictors retrieved from Eurostat. The architecture is shown in Figure 1.

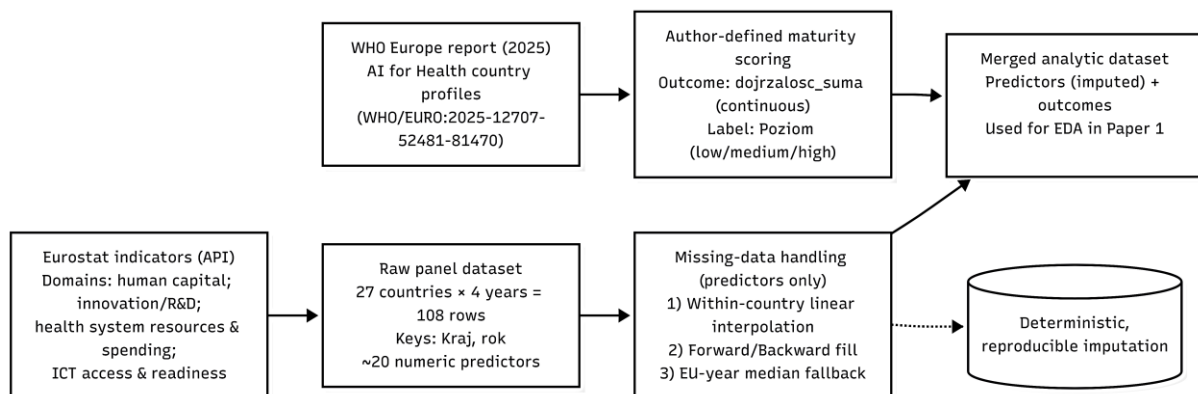


Figure 1. Data construction workflow for the analytic dataset, from WHO-based AI maturity scoring and Eurostat predictors to imputation and final dataset assembly.

Source: Own work.

Outcome definition (WHO-based maturity). Digital health maturity was operationalized as (i) a continuous score (*dojrzalosc_suma*) computed using an author-defined scoring procedure applied to WHO-sourced maturity evidence, and (ii) an ordinal class label (*Poziom*: low/medium/high) obtained by thresholding the score according to pre-specified cut-offs.

The WHO classification, which includes three stages of artificial intelligence institutionalization—informal adoption, pilot implementation, and established use—was applied. For each country, the indicator was calculated as the weighted sum of the number of AI applications assigned to these categories, with weights of 1 for informal adoption, 2 for pilot implementation, and 3 for established use. This operationalization made it possible to capture

both the scale of AI use and the degree of its institutional embedding within the healthcare system.

Subsequently, based on the obtained index values, countries were classified according to three levels of AI implementation maturity: low, medium, and high. Compared to a binary classification, this three-category approach allows for a more nuanced presentation of differences between countries by distinguishing not only between systems with low and high levels of maturity but also an intermediate category of countries occupying a middle position in terms of the advancement of AI implementation.

Predictor selection (Eurostat). Predictor variables were selected a priori to cover human capital, innovation/R&D, health system resources and spending, and ICT access and readiness, and were extracted from Eurostat using fixed dimension filters (e.g., unit, sector, age/sex) to ensure cross-country comparability.

Data access and reproducibility. Eurostat data and metadata can be accessed programmatically via the Eurostat dissemination API; dataset codes and filters are reported to ensure reproducibility.

Dataset overview (EU27, 2021-2024)

We constructed a country-year panel for the 27 EU Member States covering 2021-2024, combining a WHO-informed AI-for-health maturity outcome with candidate predictors retrieved from Eurostat. The analytical objective was to create a reproducible dataset suitable for descriptive screening, between-class comparison, and cautious predictive modelling of AI-for-health maturity across countries.

Digital health maturity was operationalized as (i) a continuous score (*dojrzalosc_suma*) derived from WHO-informed evidence and (ii) an ordinal class label (*Poziom*: low, medium, high) obtained using pre-specified thresholds. The WHO-based operationalization applied a weighted scoring logic reflecting three stages of AI institutionalization in health systems: informal adoption, pilot implementation, and established use.

Predictor variables were selected a priori to represent innovation and R&D, data capability, health-system resources and spending, and digital access. Eurostat indicators were extracted using fixed dimension filters to preserve cross-country comparability and reproducibility.

The final analytical panel retained 15 numeric predictors:

- *Otoczenie_innowacyjne_i_BR, Presja_popytowa.*
- *Zdolnosc_systemowa_dane.*
- *System_ochrony_zdrowia.*
- *Wydatki_firm_na_BR_BERD_wedlug_dzialalnosci.*
- *GERD_sektor_i_dziedziny_BR.*
- *GERD_sektor_i_rodzaj_instytucji.*
- *Lata_zycia_w_zdrowiu_HLY.*
- *Lozka_szpitalne_wedlug_regionow.*

- *Urządzenia_do_obrazowania_medycznego.*
- *Wydatki_na_opieke_zdrowotna_wedlug_dostawcow.*
- *Wydatki_przedsiębiorstw_BR_zaawansowane_tehnologie.*
- *Wydatki_przedsiębiorstw_BR_sektor_ICT.*
- *Dostep_do_internetu_gospodarstwa_domowe,*
- *Urządzenia_do_dostepu_do_internetu_osoby.*

Four candidate variables were excluded before the final analytical panel was assembled: *Poziom_wdrozenia_AI_w_firmach*, *Analityka_danych_wedlug_dzialalnosci*, *Zatrudnienie_w_sektorach_tech_i_wiedzy*, and *Bezpieczenstwo_ICT_polityka_srodki_ryzyko_swiadomosc*. In the final imputed dataset, no missing values remained across the retained analytical variables.

Before imputation, however, missingness varied substantially across retained predictors. The highest pre-imputation missingness among retained variables was observed for *Urządzenia_do_dostepu_do_internetu_osoby* (55.6%), *Wydatki_przedsiębiorstw_BR_zaawansowane_tehnologie* (54.6%), *Wydatki_przedsiębiorstw_BR_sektor_ICT* (42.6%), and *Lozka_szpitalne_wedlug_regionow* (38.0%), while several other retained variables showed lower but non-negligible incompleteness.

The retained variables and their pre-imputation missingness are summarized in Table 1. This distinction between the pre-imputation candidate panel and the final imputed analytical panel is important for interpreting both the completeness diagnostics and the subsequent modelling results.

Table 1.

Missingness table -Missing % = $NA_in_108 / 108 \times 100$.

variable	n	%
<i>Urządzenia do dostępu do internetu osoby</i>	60	55.6
<i>Wydatki przedsiębiorstw BR zaawansowane technologie</i>	59	54.6
<i>Bezpieczenstwo ICT polityka srodki ryzyko swiadomosc</i>	54	50.0
<i>Wydatki przedsiębiorstw BR sektor ICT</i>	46	42.6
<i>Lozka szpitalne według regionow</i>	41	38.0
<i>Lata zycia w zdrowiu HLY</i>	28	25.9
<i>Urządzenia do obrazowania medycznego</i>	27	25.0
<i>Poziom wdrozenia AI w firmach</i>	27	25.0
<i>System ochrony zdrowia</i>	27	25.0
<i>Wydatki na opieke zdrowotna według dostawcow</i>	27	25.0
<i>Dostep do internetu gospodarstwa domowe</i>	5	4.6
<i>Zdolnosc systemowa dane</i>	2	1.9
<i>Zatrudnienie w sektorach tech i wiedzy</i>	2	1.9

Source: Own elaboration.

Data gaps are mainly due to incomplete reporting for certain years/countries and the need to apply harmonized filter settings (e.g., entity, sector, indicator) to ensure comparability between Member States. For some indicators on digitization and ICT security, data availability is limited to subsets of countries/years, leading to a higher number of gaps.

2.2. Missing data & imputation

Missingness handling. Missing values in predictors were handled using a deterministic, three-stage procedure applied within each country over time: (1) linear interpolation, (2) forward/backward fill, and (3) year-specific EU median imputation as a fallback.

$$\hat{x}_{c,t} = \begin{cases} x_{c,t_0} + \frac{x_{c,t_1} - x_{c,t_0}}{t_1 - t_0} (t - t_0), & \text{if there exist } t_0 < t < t_1 \text{ and } x_{c,t_0}, x_{c,t_1} \text{ are observed,} \\ x_{c,t^-}, & \text{if interpolation is not possible and there exists } t^- < t \text{ with observed } x_{c,t^-} \text{ (forward fill),} \\ x_{c,t^+}, & \text{if interpolation and forward fill are not possible and there exists } t^+ > t \text{ with observed } x_{c,t^+} \text{ (backward fill),} \\ \text{median}_{c' \in EU} (x_{c',t}), & \text{otherwise (fallback: EU median in year } t). \end{cases} \quad (1)$$

Where c denotes the country and t the year; t_0 and t_1 are the nearest observed years bracketing the missing value within country c , t^- is the closest earlier observed year, and t^+ is the closest later observed year.

3. Results

The analytical dataset comprised 108 country-year observations for the EU27 in 2021–2024, including 15 numeric Eurostat predictors and two WHO-derived outcomes: the continuous maturity score “*dojrzalosc_suma*” and the ordinal maturity class “*Poziom*”. Integrity checks identified no duplicate country-year rows and no missing cells after preprocessing, which indicates that the final analytical panel was internally consistent and complete for the reported analyses. The distribution and balance of classes are shown in Figure 2.

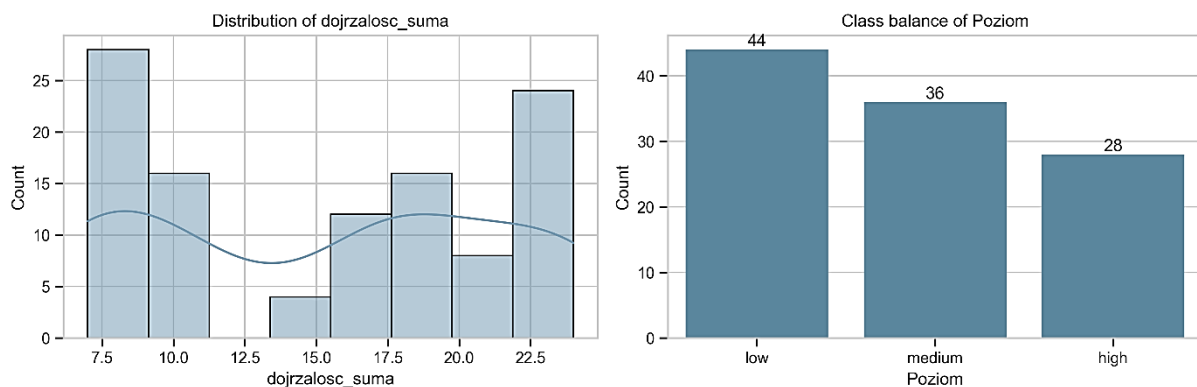


Figure 2. Targets overview.

Source: Own work.

The ordinal outcome was moderately imbalanced, with 44 observations classified as low maturity (40.7%), 36 as medium maturity (33.3%), and 28 as high maturity (25.9%). This distribution shows that the panel covered the full maturity spectrum, while low- and medium-maturity observations remained more common than high-maturity cases. Table 2 shows the best predictors according to Spearman's and Pearson's correlations.

Table 2.
Top predictors

Variable	spearman r	pearson r
<i>Otoczenie innowacyjne i BR</i>	0,51	0,47
<i>Presja popytowa</i>	0,51	0,47
<i>Wydatki firm na BR BERD według działalności</i>	0,51	0,47
<i>GERD sektor i dziedziny BR</i>	0,51	0,47
<i>Zdolność systemowa dane</i>	0,47	0,26
<i>Dostęp do internetu gospodarstwa domowe</i>	0,45	0,37
<i>Wydatki na opiekę zdrowotną według dostawców</i>	0,44	0,47
<i>System ochrony zdrowia</i>	0,42	0,40
<i>GERD sektor i rodzaj instytucji</i>	0,41	0,00
<i>Wydatki przedsiębiorstw BR zaawansowane technologie</i>	0,41	0,00
<i>Lata życia w zdrowiu HLY</i>	-0,33	-0,35
<i>Łóżka szpitalne według regionów</i>	0,31	0,33
<i>Urządzenia do dostępu do internetu osoby</i>	0,25	0,25
<i>Wydatki przedsiębiorstw BR sektor ICT</i>	0,10	0,22
<i>Urządzenia do obrazowania medycznego</i>	0,09	0,15

Source: Own work.

Descriptive screening showed that the strongest monotonic associations with the continuous maturity score were observed for a cluster of innovation and R&D-related indicators, including *Otoczenie innowacyjne i BR*, *Presja popytowa*, *Wydatki firm na BR BERD według działalności*, and *GERD sektor i dziedziny BR* (all Spearman $r = 0.510$), followed by *Zdolność systemowa dane* (Spearman $r = 0.466$). Additional positive associations were observed for *Dostęp do internetu gospodarstwa domowe*, *Wydatki na opiekę zdrowotną według dostawców*, and *System ochrony zdrowia*, suggesting that innovation capacity, data-system capability, digital access, and selected health-system financing indicators aligned with higher maturity scores in descriptive terms.

Between-class comparisons further supported this pattern. The largest class differences by Kruskal–Wallis screening were observed for:

- *Wydatki na opiekę zdrowotną według dostawców* ($H = 30.44$).
- *Otoczenie innowacyjne i BR* ($H = 27.99$).
- *Presja popytowa* ($H = 27.99$).
- *Wydatki firm na BR BERD według działalności* ($H = 27.99$).
- *GERD sektor i dziedziny BR* ($H = 27.99$).
- *Zdolność systemowa dane* ($H = 26.32$).

indicating that these variables most clearly differentiated low-, medium-, and high-maturity observations. The differences between the classes are shown in Figure 3.

At the same time, the predictor correlation matrix showed substantial redundancy among several retained variables. In particular, some innovation and R&D indicators were perfectly rank-correlated in the current panel, including *Otoczenie innowacyjne i BR*, *Presja popytowa*, *Wydatki firm na BR BERD według działalności*, and *GERD sektor i dziedziny BR* ($r = 1.000$), which suggests that the descriptive signal was partly concentrated in a tightly collinear predictor block rather than in fully independent dimensions. Table 3 shows the complete results.

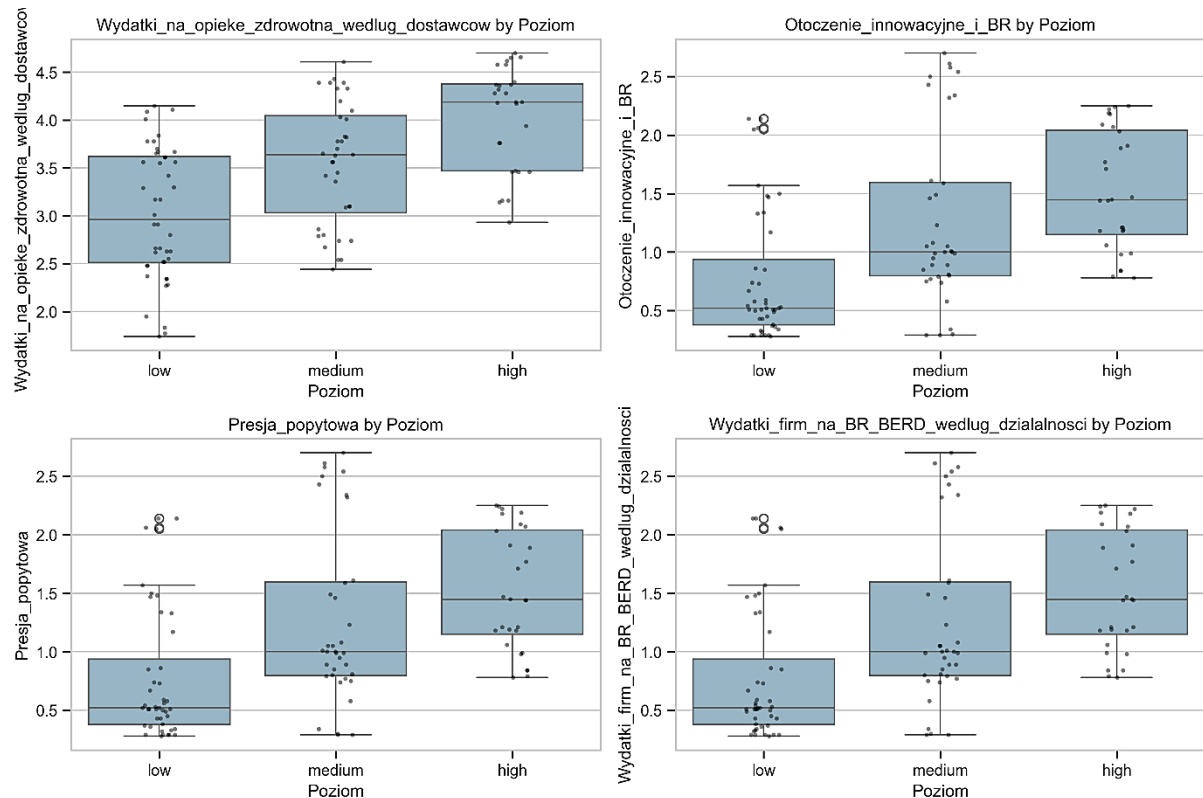


Figure 3. Top boxplots by outcome.

Source: Own work.

Table 3.
Group differences by maturity class

variable	kruskal_H	kruskal_p
Wydatki na opieke zdrowotna wedlug dostawcow	30,44	0,00000
Otoczenie innowacyjne i BR	27,99	0,00000
Presja popytowa	27,99	0,00000
Wydatki firm na BR BERD wedlug dzialalnosci	27,99	0,00000
GERD sektor i dziedziny BR	27,99	0,00000
Zdolnosc systemowa dane	26,32	0,00000
System ochrony zdrowia	23,93	0,00001
GERD sektor i rodzaj instytucji	22,75	0,00001
Wydatki przedsiebiorstw BR zaawansowane technologie	22,75	0,00001
Dostep do internetu gospodarstwa domowe	15,46	0,00044
Lata zycia w zdrowiu HLY	11,63	0,00298
Lozka szpitalne wedlug regionow	7,27	0,02639
Urzadzenia do dostepu do internetu osoby	6,38	0,04121
Urzadzenia do obrazowania medycznego	5,76	0,05610
Wydatki przedsiebiorstw BR sektor ICT	1,94	0,37877

Source: Own work.

As a cautious predictive extension, grouped cross-validation by country was used to test (e.g., Table 4) whether the retained predictors contained out-of-country signal for the continuous maturity score. Among the evaluated regression models, Elastic Net performed best, with MAE = 4.337 and RMSE = 5.230, outperforming both Ridge and the baseline mean model; however, the explained variance remained modest ($R^2 = 0.297$), indicating partial but limited predictive generalisability across countries.

Table 4.
Grouped regression modelling

	model	MAE	RMSE	R2
0	ElasticNet	4.336912	5.230291	0.297355
1	Ridge	4.452667	5.284934	0.282596
2	Baseline mean	5.921998	6.665395	-0.141133

Source: Own work.

The partial predictive fit between the observed and predicted values is shown in Figure 4.

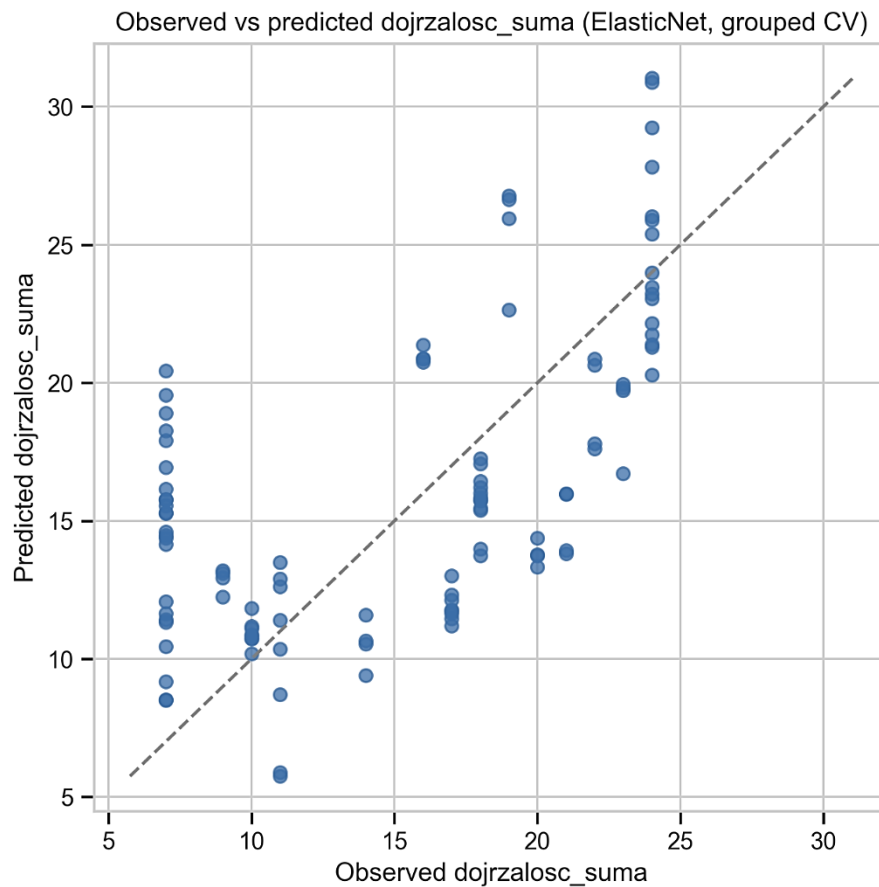


Figure 4. Regression-observed vs predicted.

Source: Own work.

The observed-versus-predicted plot suggested a possible non-linear pattern, particularly at the lower and upper ends of the outcome range; however, polynomial extensions did not improve grouped cross-validated performance over ElasticNet or Ridge. Therefore, the apparent curvature was treated as a useful diagnostic signal rather than sufficient evidence to replace the main linear specification.

4. Discussion

The present findings should be interpreted within a wider body of research showing that digital maturity in health care is a multidimensional construct that extends beyond infrastructure and software availability. Recent work on digital health communication maturity, data and AI governance, and national hospital maturity assessment indicates that meaningful readiness depends on organizational capability, interoperability, communication processes, implementation routines, and governance structures (Cresswell et al., 2025; Kim, Namkoong, 2025; Morley et al., 2022).

From this perspective, the strong descriptive role of innovation- and R&D-related indicators is consistent with comparative research suggesting that AI maturity is embedded in broader ecosystem capacity. Cross-country analysis of OECD health systems argues that countries differ not only in their rate of adoption but also in the maturity of their policy, strategic, and implementation environments, while OECD policy work similarly emphasizes that scaling AI in health requires system-level preparedness rather than isolated pilots or fragmented projects (Castonguay et al., 2024).

The importance of data capability and digital access in the present analysis is also closely aligned with previous maturity and interoperability research. The Health Information Systems Interoperability Maturity Toolkit and its application in Kenya illustrate that digital initiatives may coexist with relatively weak operational interoperability, which helps explain why macro-level readiness varies substantially across settings and why information exchange remains a limiting factor in many systems (Evaluation, 2017; Nyangena et al., 2021). Related international frameworks from the World Bank and the Global Digital Health Partnership likewise position data exchange, standards, and implementation governance as core dimensions of digital maturity assessment (World Bank, n.d.; World Health Organization, n.d.)

The present results are also compatible with hospital-level evidence indicating that digital advancement tends to be uneven across dimensions. Findings from the German DigitalRadar show that maturity may be stronger in structures and systems than in patient participation, telemedicine, or exchange of information, which suggests that digital transformation is layered rather than uniform (Geissler et al., 2024). Similar conclusions emerge from comparative work on national approaches to assessing hospital digital maturity and from systematic reviews of hospital maturity dimensions, both of which describe digital development as internally differentiated and only partly synchronized across organizational domains (Cresswell et al., 2025; Nurhayati et al., 2025).

This interpretation is further reinforced by telemedicine and telehealth maturity research. Existing models of telemedicine service maturity conceptualize maturity as staged organizational development involving workflow integration, leadership support, service design, and long-term scaling capacity rather than as the simple presence of remote technologies

(Development of a Maturity Model for Telemedicine Services, 2024; Ingenium Digital Health, 2021; van Dyk, Schutte, 2013). Evidence from the Brazilian telehealth maturity model reaches a similar conclusion by showing that institutional diagnostics are needed before maturity can be meaningfully assessed, which supports the view that readiness is deeply embedded in organizational context (Silva et al., 2024).

At the same time, the present findings should not be taken to mean that general digital maturity is equivalent to AI-specific readiness. Work on data and AI governance in health care has shown that trustworthy AI deployment requires explicit arrangements for accountability, stewardship, oversight, and institutional responsibility, while broader OECD analyses stress the importance of evaluation, policy alignment, and implementation safeguards for scaling AI safely in practice (Morley et al., 2022; Organisation for Economic Co-operation & Development, 2026). This distinction is reinforced by recent discussions of digital maturity as a possible but incomplete gatekeeping concept for health AI, as well as by work arguing that accreditation-linked requirements may help build the organizational foundations needed for effective AI adoption (As'ad, Faran, 2025; Bellei et al., 2026).

The observed concentration of signal within an innovation-related block is therefore theoretically plausible. Previous maturity research suggests that governance, infrastructure, communication capability, interoperability, and organizational routines frequently co-evolve, meaning that empirical studies often capture a broader latent readiness dimension rather than a set of fully independent mechanisms (Global Digital Health Partnership, 2026; Kim, Namkoong, 2025; Mapping Digital Health Maturity Models and Accreditation-Linked Standards, 2025). From that perspective, multicollinearity may reflect substantive interdependence among readiness domains rather than only a technical modelling issue.

The moderate predictive generalizability observed in this type of work also supports a cautious interpretation of explanatory models. Differences in framework scope, indicator design, implementation context, and data completeness make strong out-of-sample prediction difficult across heterogeneous health systems, even when descriptive structures remain informative (As'ad, Faran, 2025; Cresswell et al., 2025; World Bank, n.d.). For this reason, maturity-oriented models may be especially valuable as strategic and comparative tools, whereas their predictive performance should be interpreted with methodological caution.

Overall, the present study fits a literature showing that AI-for-health readiness is embedded in broader national capabilities related to innovation, interoperability, implementation maturity, and governance. Future research would benefit from extending macro-level models with more explicit indicators of standards, accreditation, patient participation, safety capability, and AI-specific oversight so that digital maturity and AI readiness can be distinguished more clearly in comparative health-system research (Geissler et al., 2024; Kutza et al., 2022).

5. Conclusion

This study developed a reproducible EU27 country-year dataset that combines a WHO-informed AI-for-health maturity measure with harmonized Eurostat predictors for 2021-2024. The results show that the constructed panel is analytically coherent and that AI-for-health maturity is descriptively associated with innovation and R&D capacity, data-system capability, digital access, and selected health-system financing indicators.

At the same time, the modelling results indicate only moderate predictive generalizability part across countries. For that reason, the predictive component should be interpreted as a transparent proof-of-concept rather than a final forecasting benchmark.

The main contribution of the paper is therefore methodological as much as substantive. By operationalizing a WHO-informed maturity construct and embedding it in a reproducible Eurostat-based workflow, the study provides a practical basis for future comparative analyses of AI readiness in European health systems.

References

1. *Artificial intelligence for health in the WHO European Region* (2025). <https://www.who.int/europe/publications/i/item/WHO-EURO-2025-12707-52481-81470>
2. *Artificial intelligence is reshaping health systems: state of readiness across the WHO European Region* (2025). <https://www.who.int/europe/publications/i/item/WHO-EURO-2025-12707-52481-81028>
3. As'ad, M., Faran, N. (2025). Digital maturity scores as gatekeepers for health AI: useful proxy or false comfort? *Proceedings (Baylor University Medical Center)*, 38(5), 779-782. <https://doi.org/10.1080/08998280.2025.2524301>
4. Bellei, E.A., Debon, R., Costa, E., De Marchi, A.C.B. (2026). How hospital accreditation requirements bridge enablers for AI readiness: interpretative analysis of intersections in framework standards. *Frontiers in Digital Health*. <https://doi.org/10.3389/fdgth.2026.1769181>
5. Castonguay, A., Wagner, G., Motulsky, A., Pare, G. (2024). AI maturity in health care: An overview of 10 OECD countries. *Health Policy*, 140, 104938. <https://doi.org/10.1016/j.healthpol.2023.104938>
6. Cresswell, K., Jahn, F., Silsand, L., Woods, L., Postema, T., Logan, M., Malkic, S., Ammenwerth, E. (2025). Assessing Digital Maturity of Hospitals: Viewpoint Comparing National Approaches in Five Countries. *Journal of Medical Internet Research*, 27, e57858. <https://doi.org/10.2196/57858>

7. *Development of a Maturity Model for Telemedicine Services* (2024). <https://re.public.polimi.it/bitstream/11311/1288139/1/Olive%20et%20al%202024.pdf>
8. Eurostat (2017). *How to decode information and data included in the SDMX files*. https://ec.europa.eu/eurostat/documents/728032/8982469/How_to_decode_information_and_data_included_in_the_SDMX_files.pdf
9. Eurostat (2022). *Detailed guidelines — SDMX 2.1 API — data query*. <https://ec.europa.eu/eurostat/web/user-guides/data-browser/api-data-access/api-detailed-guidelines/sdmx2-1/data-query>
10. Geissler, A., Hollenbach, J., Haring, M., Amelung, V.E., Thun, S., Haering, A. (2024). A nationwide digital maturity assessment of hospitals – Results from the German DigitalRadar. *Health Policy and Technology*, 13(4), 100904. <https://doi.org/10.1016/j.hlpt.2024.100904>
11. Global Digital Health Partnership (2026). *Framework for Selecting Digital Health Maturity Assessments*. https://gdhp.health/wp-content/uploads/GDHP-Digital-Health-Maturity-White-Paper_final_v1.1_26032026.pdf
12. Ingenium Digital Health (2021). *A Telehealth Program Maturity Model*. <https://ingeniumdigitalhealth.com/wp-content/uploads/2021/01/936-Ingenium-Whitepaper-012-A-Telehealth-Program-Maturity-Model.pdf>
13. Kim, G.J., Namkoong, K. (2025). Developing the Digital Health Communication Maturity Model: Systematic Review. *Journal of Medical Internet Research*, 27, e68344. <https://doi.org/10.2196/68344>
14. Kutza, J.-O., Hübner, U., Holmgren, A.J. (2022). Patient Safety Informatics: Criteria Development for Assessing the Maturity of Digital Patient Safety in Hospitals. *Studies in Health Technology and Informatics*, 294, 885-889. <https://doi.org/10.3233/SHTI220618>
15. Liaw, S.T. et al. (2022). Digital health and capability maturity models—a critical literature review. *Frontiers in Digital Health*, 4, 1045685. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9846694/>
16. Liaw, S.T. et al. (2023). Digital health and capability maturity models—a critical thematic review and conceptual synthesis of the literature. *Journal of the American Medical Informatics Association*. <https://pubmed.ncbi.nlm.nih.gov/36451257/>
17. *Mapping digital health maturity models and accreditation-linked standards* (2025). <https://www.springermedizin.de/mapping-digital-health-maturity-models-and-accreditation-linked-/52169580>
18. MEASURE Evaluation (2017). *Health Information Systems Interoperability Maturity Toolkit*. <https://www.measureevaluation.org/tools/health-information-systems-interoperability-toolkit.html>
19. Meltzer, H.M., Jensen, T.K., Májek, O., Moshhammer, H., Tolonen, H., Mørck, T.A., Weihe, P., Åkesson, A. (2022). Enhancing Human Biomonitoring Studies through Linkage to

- Administrative Registers—Status in Europe. *International Journal of Environmental Research and Public Health*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9103287/>
20. Morley, J., Murphy, L., Mishra, A., Joshi, I., Karpathakis, K. (2022). Governing Data and Artificial Intelligence for Health Care: Developing an International Understanding. *JMIR Formative Research*, 6(1), e31623. <https://doi.org/10.2196/31623>
 21. Moustakidis, S., Kouloulis, C., Tsironis, D., Sarigiannidis, P., Tsikrika, S., Vrochidis, S., Zacharopoulou, T.E., Papadopoulos, D. (2022). Identifying Country-Level Risk Factors for the Spread of COVID-19 in Europe Using Machine Learning. *Viruses*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8955542/>
 22. Nurhayati, S., Suratman, Ngadiman (2025). Analysis of Digital Maturity Aspects in Hospitals: A Systematic Literature Review. *International Journal of Science and Society*, 7(3), 155-164. <https://doi.org/10.54783/ijssoc.v7i3.1496>
 23. Nyangena, J., Rajgopal, R., Ombech, E.A., Oloo, E., Luchetu, H., Wambugu, S., Kamau, O., Nzioka, C., Gwer, S., Ndiritu Ndirangu, M. (2021). Maturity assessment of Kenya's health information system interoperability readiness. *BMJ Health & Care Informatics*, 28(1), e100241. <https://doi.org/10.1136/bmjhci-2020-100241>
 24. Organisation for Economic Co-operation, & Development (2026). *Scaling Artificial Intelligence in Health*. https://www.oecd.org/en/publications/scaling-artificial-intelligence-in-health_a436e12d-en.html
 25. Silva, A.B., Pisa, I.T., Messina, L.A., Pelogi, A.P.S., Tenório, J.M., Sousa, F.S., Santos, D.L., Maia, J., Andrade Filha, I.G. de, Guedes, A.C.C.M., Lopes, P.R. de L., Maia, P.R. da S. (2024). Diagnostic evaluation of institutions as a basis for designing the Brazilian maturity model of telehealth services. *BMC Health Services Research*, 24(1), 372. <https://doi.org/10.1186/s12913-024-10723-8>
 26. Snowdon, A. et al. (2024). Digital Maturity as a Predictor of Quality and Safety Outcomes in US Hospitals: Cross-Sectional Observational Study. *Journal of Medical Internet Research*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11336495/>
 27. van Dyk, L., Schutte, C. (2013). *The Telemedicine Service Maturity Model: A Framework for the Measurement and Improvement of Telemedicine Services*. <https://www.intechopen.com/chapters/45211>
 28. World Bank (n.d.). *Digital Health Maturity Assessments and Toolkits*. Retrieved from: <https://openknowledge.worldbank.org/server/api/core/bitstreams/95128705-06f6-43b8-9236-9e7f8abf301d/content>
 29. World Health Organization (n.d.). *Global Digital Health Monitor - Countries Dashboard*. Retrieved <https://data.who.int/dashboards/gdhm/countries>
 30. Zhou, Z., Xiang, Y., Liu, C., Huang, X., Fu, Y. (2025). Quantitative evaluation and optimization of AI policy and regulatory texts for smart healthcare. *Frontiers in Public Health*, 13, 1731013. <https://doi.org/10.3389/fpubh.2025.1731013>

Appendix/Supplementary Material

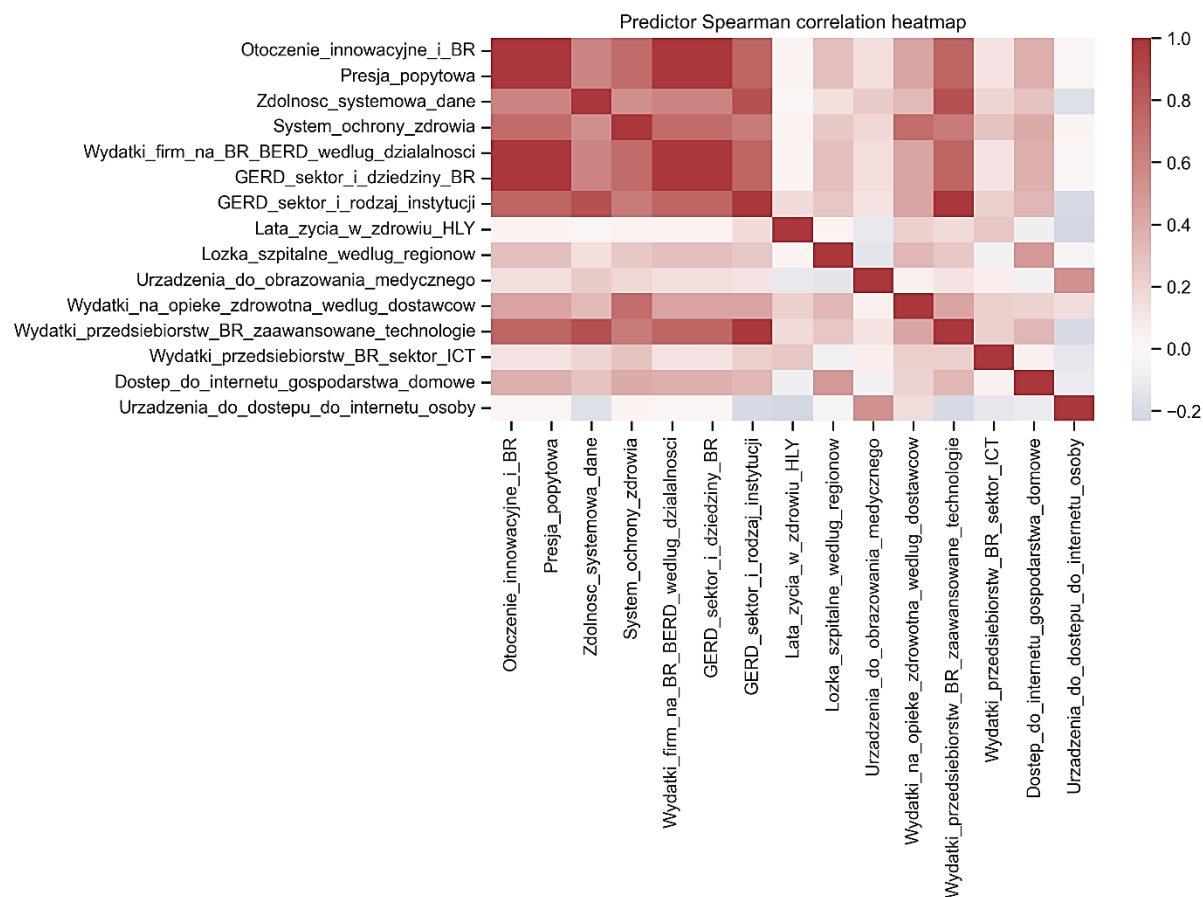


Figure 5. Spearman correlation heatmap.

Source: Own work.