

Appendix

Database	Mental-health terms ¹	ML / AI terms ²	Prediction terms ³	Social-media terms ⁴	Field / Filter
PubMed	PTSD, Post-Traumatic Stress Disorder, Moral Injury, Depression, Depressive Symptoms	Machine Learning, Deep Learning, Transfer Learning, Artificial Intelligence, Natural Language Processing	Prediction, Predictive Model Method, Algorithm*, Forecast*, Machine Prediction	Social Media, Web 2.0, Twitter Messaging, Mobile Social Network, Social Tagging, Folksonomy	Title/Abstract; MeSH; 2022–2025
PsycINFO	PTSD, Post-Traumatic Stress Disorder, Depression, Depressive Symptoms	Artificial Intelligence, Deep Learning, Natural Language Processing, Predictive Modeling	Prediction, Predictive Model	Social Media, Social Networking Site	Empirical / Quantitative filters; 2022–2025
ACM DL	PTSD, Post-Traumatic Stress Disorder, Depression, Depressive Symptoms	Machine Learning, Deep Learning, Artificial Intelligence, Natural Language Processing	Prediction Model, Forecasting Algorithm	Social Media, Twitter, Facebook, Reddit, Weibo, SNS	Boolean “All fields”; 2022–2025
IEEE Xplore	PTSD, Post-Traumatic Stress Disorder, Depression, Depressive Symptoms	Machine Learning, Deep Learning, Artificial Intelligence, Natural Language Processing	Prediction Model, Forecasting Algorithm, Predict*	Social Media, Twitter, Facebook, Reddit, Weibo	Abstract; 2022–2025
Web of Science	PTSD, Post-Traumatic Stress Disorder, Depression, Depressive Symptoms	Machine Learning, Deep Learning, Artificial Intelligence, Natural Language Processing	Prediction Model, Forecasting Algorithm, Predict*	Social Media, Twitter, Facebook, Reddit, Weibo	TS= Topic; 2022–2025
Scopus	PTSD, Post-Traumatic Stress Disorder, Depression, Depressive Symptoms	Machine Learning, Deep Learning, Artificial Intelligence, Natural Language Processing	Prediction Model, Forecasting Algorithm, Predict*	Social Media, Twitter, Facebook, Reddit, Weibo	TITLE-ABS-KEY; PUBYEAR > 2021

Table A1: Literature Search

a. Year-2024 vs. pre-2024 performance jump

Metric	Main OLS β (p)	Overall t (p)	User-level t (p)	Post-level t (p)	W/o baseline model t (p)	Controlled β (p)
Pearson r	0.158, $p < .001$	+6.40, $p < .001$	—	+5.05, $p < .001$	+6.70, $p < .001$	—
Accuracy	0.073, $p < .001$	+6.73, $p < .001$	+3.84, $p < .001$	+7.60, $p < .001$	+7.91, $p < .001$	0.088, $p < .001$
F1-score	0.059, $p < .001$	+3.49, $p < .001$	+2.03, $p = .045$	+3.60, $p < .001$	n.s.	0.080, $p < .001$

b. Post-2024 vs. pre-2024 performance jump

Metric	Main OLS β (p)	Overall t (p)	User-level t (p)	Post-level t (p)	W/o baseline model t (p)	Controlled β (p)
Pearson r	0.158, $p < .001$	+6.40, $p < .001$	—	+5.05, $p < .001$	+6.70, $p < .001$	—
Accuracy	0.028, $p = .031$	+2.11, $p = .036$	+3.84, $p < .001$	n.s.	n.s.	0.031, $p = .017$
F1-score	0.029, $p = .088$	+1.72, $p = .086$	+2.03, $p = .045$	n.s.	n.s.	0.038, $p = .027$

Table A2: Robustness checks for the step-change

Note. OLS = ordinary least squares with a 2024 dummy; “Controlled” OLS adjusts for task, level, and baseline status. n.s. = not significant at $p < .05$. “—” indicates insufficient data to compute the statistic.

Metric	Overall t (p)	User-level t (p)	Post-level t (p)	Controlled β (p)
Accuracy	+4.11, $p < .001$	+3.64, $p < .001$	+2.98, $p = .003$	0.05, $p < .001$
F1-score	+4.64, $p < .001$	+3.77, $p < .001$	+3.37, $p < .001$	0.05, $p < .05$
Pearson r	+3.00, $p = .003$	+3.15, $p = .010$	+3.65, $p < .001$	0.06, $p < .05$

Table A3: Effect of deep / hybrid processing versus traditional ML

Note. Positive t indicates higher performance for deep / hybrid models. “Controlled β ” comes from an OLS model that adjusts for task, level, baseline status, platform, sample size, feature family, modality, and annotation method.

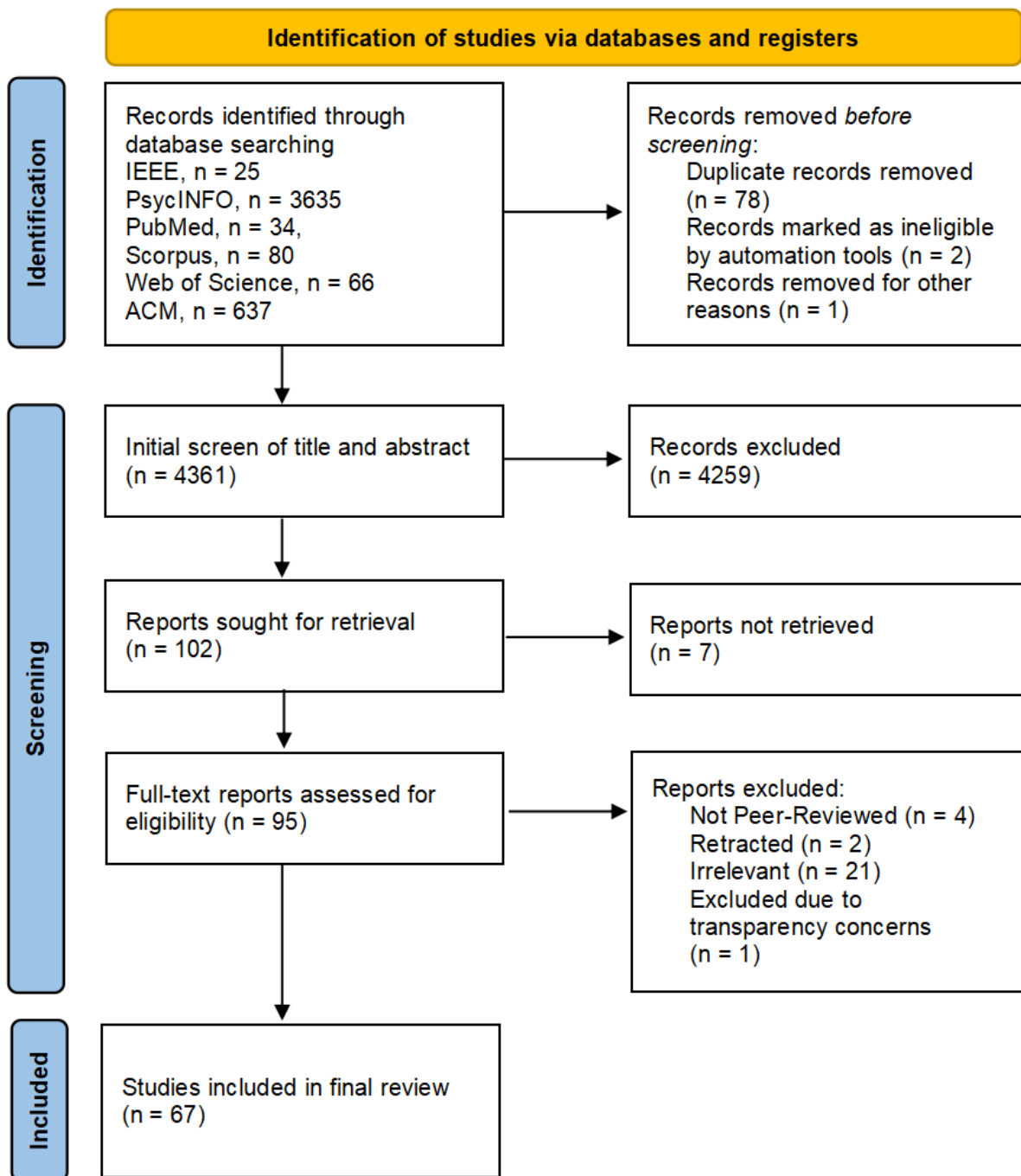


Figure 1: PRISMA Flowchart