Model-Centric Governance for AI Modules in Multimodal Telemonitoring Platforms: A Conceptual Framework for the NeuroPredict Ecosystem

Lidia BĂJENARU[1,2], Cristian PETRACHE[1], Dragoș NICOLAU[1]
[1]National Institute for Research and Development in Informatics - ICI Bucharest, Romania
[2]National University of Science and Technology Politehnica, Bucharest, Romania
lidia.bajenaru@ici.ro, cristian.petrache@ici.ro, dragos.nicolau@ici.ro

Abstract: Digital health platforms that combine Internet of Medical Things (IoMT) devices, cloud–edge infrastructures, and artificial intelligence (AI) analytics are increasingly used to support remote monitoring of patients with chronic and neurodegenerative conditions. Within this context, NeuroPredict integrates heterogeneous physiological, cognitive, behavioural, and environmental data streams with embedded AI modules designed to generate indicators for clinical monitoring and decision support. Although prior work has addressed platform architecture, cybersecurity, role-based access control, and multimodal data integration, the governance of AI models deployed in such ecosystems remains insufficiently explored. This paper proposes a model-centric governance framework for AI-enabled telemonitoring platforms, organized around six complementary dimensions: model integrity and provenance, controlled deployment and update governance, threshold and calibration governance, inference traceability through prediction records, multimodal consistency validation, and subgroup/fairness monitoring under human oversight. The framework consolidates existing model-management and traceability capabilities and extends them within a unified governance layer. By linking infrastructure-level security with model-level governance, this work contributes to trustworthy AI deployment in digital health ecosystems and provides a foundation for reliable, auditable, and transparent AI-driven telemonitoring and future real-world deployment.
Keywords: AI governance, digital health, Internet of Medical Things, cybersecurity, NeuroPredict platform.

INTRODUCTION

Neurodegenerative diseases such as Alzheimer's disease (AD), Parkinson's disease (PD), and multiple sclerosis (MS) are characterized by progressive, heterogeneous, and long-term trajectories that may affect cognition, motor function, behaviour, sleep, autonomy, and activities of daily living. Their management increasingly requires

https://doi.org/10.54851/v8i2y202604

longitudinal and personalized monitoring capable of capturing subtle changes that may occur outside episodic in-clinic evaluations. This need has accelerated the development of digital health ecosystems based on wearable devices, smartphones, home sensors, Internet of Medical Things (IoMT) infrastructures, cloud–edge processing, and artificial intelligence (AI)-based analytical modules (Muurling et al., 2021; Powers et al., 2021; Lipsmeier et al., 2022; Omberg et al., 2022; Butler et al., 2025).

Recent studies indicate a transition from isolated device-based measurements toward multimodal telemonitoring architectures. In Parkinson's disease, smartphone-based assessments, smartwatch inertial sensors, and wearable/mobile applications have demonstrated the feasibility of remote monitoring for motor symptoms and treatment-related fluctuations (Powers et al., 2021; Lipsmeier et al., 2022; Omberg et al., 2022). In Alzheimer's disease, remote monitoring studies have combined smartphones, wearables, home-based sensors, cognitive tasks, and functional assessments to support longitudinal observation of cognitive and daily-living changes (Muurling et al., 2021; Muurling et al., 2024; Butler et al., 2025). Similar trends are emerging in multiple sclerosis, where mobile, wearable, electronic health record, and self-reported data have been used to model clinically relevant disease dimensions in naturalistic settings (Gashi et al., 2024). These developments confirm the increasing importance of multimodal data acquisition and AI-assisted analysis in neurodegenerative disease monitoring.

However, the transition from multimodal data acquisition to AI-enabled telemonitoring raises a second-order governance problem. Once AI modules become embedded components of a digital health platform, their reliability no longer depends only on their initial predictive performance, but also on how model artifacts, preprocessing configurations, feature schemas, thresholds, validation evidence, update procedures, and inference outputs are controlled throughout their lifecycle. Recent healthcare AI literature emphasizes the need for structured quality assessment, data-quality governance, continual performance monitoring, recurring local validation, transparency, and post-deployment oversight for AI-based prediction models (de Hond et al., 2022; Feng et al., 2022; Youssef et al., 2023; Gilbert et al., 2025). Nevertheless, these requirements are rarely translated into a platform-level governance layer for multimodal telemonitoring systems.

This gap is particularly relevant for the NeuroPredict platform, a laboratory-tested proof-of-concept multimodal telemonitoring platform developed for neurodegenerative disease monitoring. The NeuroPredict platform integrates heterogeneous physiological, behavioural, cognitive, emotional, and contextual data streams through a modular edge–cloud IoMT architecture. Prior work has described its secure data acquisition, semantic harmonization, multimodal synchronization, encrypted communication and storage, role-based access control, token-based authentication, and role-specific dashboards for patients, caregivers, physicians, and specialist physicians (Băjenaru et al., 2024; Sandulescu et al., 2024; Băjenaru et al., 2025 Ianculescu et al., 2025; Vevera et al., 2025). In parallel, several AI-oriented components previously reported at different stages of development are now implemented within the current NeuroPredict platform, including fall detection, Parkinson's disease-related screening based on spiral analysis, Alzheimer's disease-oriented explainable machine learning, and voice-based emotion recognition (Ianculescu et al., 2025; Tîrziu et

al., 2024; Tîrziu et al., 2025; Băjenaru et al., 2026 Mihăescu et al., 2026).

Previous NeuroPredict studies addressed infrastructure security, multimodal data processing, user roles, privacy, model version tracking, inference logging, and individual AI components. However, these capabilities were introduced as separate technical mechanisms rather than as part of a unified model-centric governance structure. As the AI modules are now integrated into the platform, a broader governance perspective is required to coordinate model provenance, deployment and update control, thresholds and calibration, structured inference traceability, multimodal consistency, and subgroup-aware monitoring throughout the analytical lifecycle. This distinction is especially important in multimodal telemonitoring, where analytical outputs may depend on changing sensors, preprocessing pipelines, calibration settings, data-quality conditions, and evolving model configurations.

Against this background, the present paper proposes a model-centric governance framework for the integrated AI modules of the NeuroPredict platform. Rather than introducing model versioning or inference logging as new standalone mechanisms, the framework consolidates existing capabilities and extends them with additional lifecycle-oriented governance controls. The framework is structured around six interrelated dimensions: model integrity and provenance, controlled deployment and update governance, threshold and calibration governance, inference traceability through prediction records, multimodal consistency validation, and subgroup/fairness monitoring under human oversight.

Existing frameworks address lifecycle quality assessment (de Hond et al., 2022), continual monitoring and updating (Feng et al., 2022), data quality (Schwabe et al., 2024), transparent model documentation (Gilbert et al., 2025), and responsible-AI governance practices (Papagiannidis et al., 2025). The proposed framework adapts these foundations to NeuroPredict's multimodal edge–cloud architecture by linking model configurations and inference records with modality availability, cross-modal consistency checks, and subgroup-aware human oversight through role-specific dashboards. Its contribution lies in coordinating these controls with existing infrastructure security. The architectural mapping distinguishes existing, partially implemented, and proposed governance mechanisms, consistent with the platform's proof-of-concept status.

NEUROPREDICT PLATFORM ARCHITECTURE AND AI-RELATED SECURITY CHALLENGES

Overview of the NeuroPredict Telemonitoring Architecture

NeuroPredict is a continuous multimodal telemonitoring and health-assessment platform designed to support neurodegenerative disease management. Its architecture integrates heterogeneous sensing devices, centralized processing services, and AI-based analytical modules within a layered cloud ecosystem supporting monitoring and decision-support functions. Data acquired from medical and wearable devices are transmitted through secure communication channels and processed using preprocessing and cross-modal standardization procedures (Hemalatha et al., 2025; Katz et al., 2025). The platform integrates heterogeneous longitudinal patient data into a structured representation of the patient's condition. At the analytical level, ML and XAI modules convert multimodal inputs into structured indicators supporting risk stratification, trend analysis, and longitudinal profiling. These outputs are intended to support clinical decision-making.

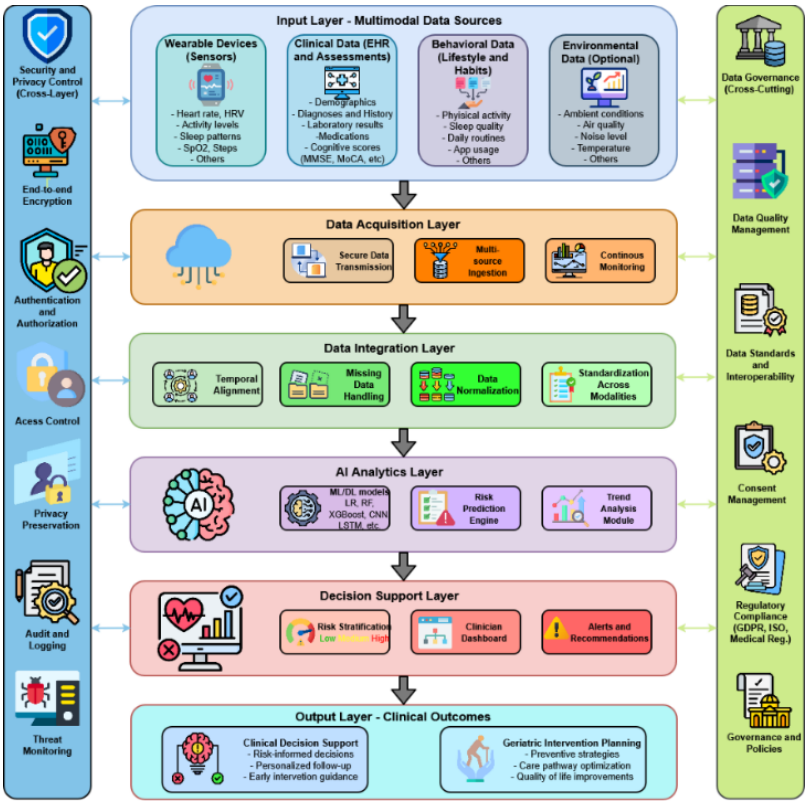


Figure 1. *NeuroPredict System Architecture Overview*

Figure 1 presents the main architectural layers and data flows of the NeuroPredict platform.

Multimodal Data Processing Pipeline

The multimodal data processing pipeline represents how heterogeneous data streams are preprocessed into analysis-ready representations in the NeuroPredict ecosystem. Because data sources are heterogeneous, the pipeline ensures consistency, temporal alignment, and semantic harmonization. The first step of the pipeline is data ingestion and synchronization. Since data are generated at different sampling rates and with different completeness levels, temporal alignment is necessary in order to generate coherent patient-level timelines. This is especially relevant in telemonitoring environments, where continuous physiological measurements need to be integrated with episodic clinical assessments to allow for longitudinal analysis. In more complex scenarios, representation learning techniques may be used to derive latent features from high-dimensional inputs. Fusion can be performed at the feature or decision level depending on the analytical goal, which influences interpretability, robustness and sensitivity to missing data. The main steps of the NeuroPredict multimodal processing pipeline are summarized in Figure 2.

Embedded AI Modules in the Monitoring Ecosystem

AI technologies are integrated within the NeuroPredict monitoring architecture as modular analytical units that operate with harmonized multimodal data produced by the processing pipeline. The modules work together to estimate patient risk, explain prediction results, and monitor patient evolution over time. The predictive layer is designed to support AD and PD related risk estimation and longitudinal analysis by combining conventional ML methods with

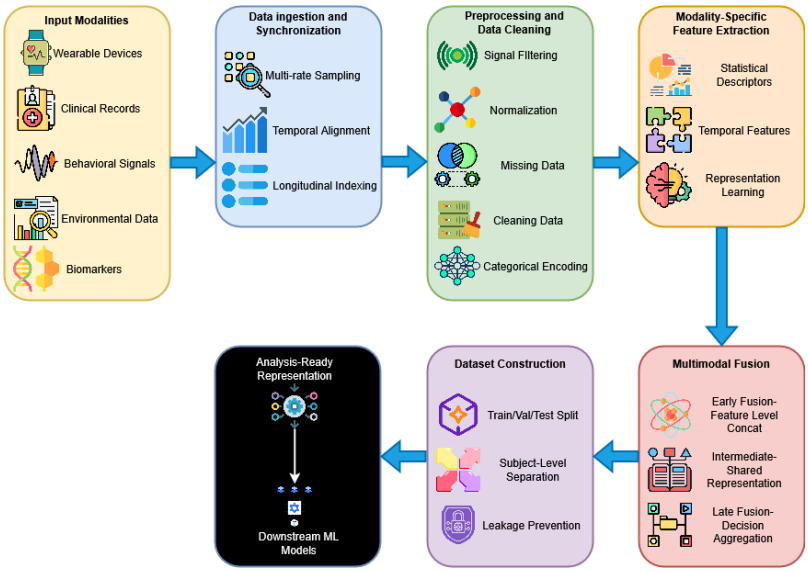


Figure 2. *Multimodal Data Processing Pipeline in the NeuroPredict Ecosystem*

more advanced nonlinear models. Different ML models can be employed, including logistic regression, ensemble methods, and deep neural networks, using physiological, cognitive, behavioural, and demographic data as input. Explainability methods are used to examine model behaviour from a global perspective. Explainability methods (Lundberg & Lee, 2017) help identify the variables influencing individual predictions, and population-level analyses examine the importance of risk factors across the monitored patients. The predicted probabilities are calibrated and classified into predefined risk categories for screening and patient monitoring (Aravindkumar et al., 2026). Figure 3 provides a conceptual overview of the analytical functions

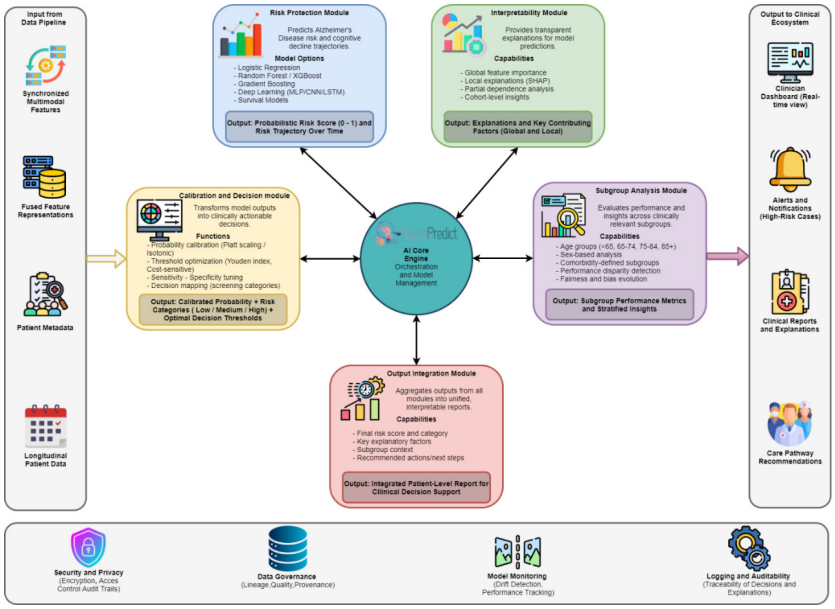


Figure 3. *Conceptual organization of AI-related analytical functions in the NeuroPredict ecosystem*

associated with NeuroPredict. It does not imply that all depicted functions have been fully implemented or clinically validated.

Security Challenges in AI-Enabled Telemonitoring Platforms

In platforms such as NeuroPredict, patient data are collected from distributed devices, continuously monitored, and analyzed using AI methods. Protecting these data requires security mechanisms throughout acquisition, transmission, processing, and storage. Unlike conventional medical information systems centered primarily on data storage and processing, NeuroPredict relies on a distributed ecosystem of wearable devices, mobile communication channels, edge/cloud components, and centralized AI services (Mitan et al., 2025). The platform analyzes longitudinal signals that may reflect daily routines, behavioural patterns, lifestyle characteristics, and aspects of clinical status. Secure data transfer between distributed devices and processing services is therefore essential. ML models are dynamic analytical components whose performance may affect decision-support outputs. The platform may be affected by adversarial manipulation of input data (Brohi & Mastoi, 2025), changes in patient populations over time, and reduced prediction accuracy when new data differ from those used for model training. Figure 4 presents the main security risks associated with each layer of the NeuroPredict platform.

Infrastructure Security vs. AI Model Security

The use of ML models in clinical applications introduces security and assurance concerns that extend beyond the underlying infrastructure. Protecting the platform and protecting the models are closely related, but they address different classes of threats. Infrastructure security focuses on the technological ecosystem used for data collection, transmission, storage, and processing. Typical threats include unauthorized access, credential compromise, network interception, ransomware, and denial-of-service attacks. Because telemonitoring systems rely on distributed data collection from external devices and remote users, failures in connectivity or breaches of data-transmission paths can affect downstream analyses and potentially compromise decision-support functions. Infrastructure security (Narula et al., 2025) and AI model security are therefore

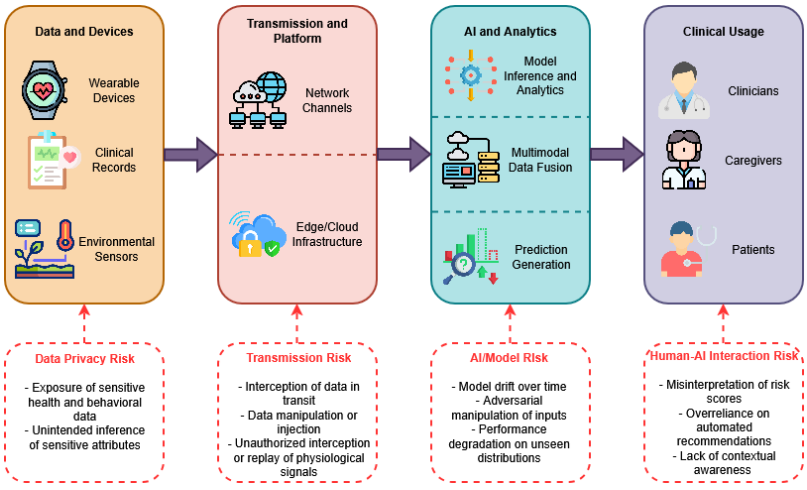


Figure 4. *Security and risk landscape in AI-enabled telemonitoring systems*

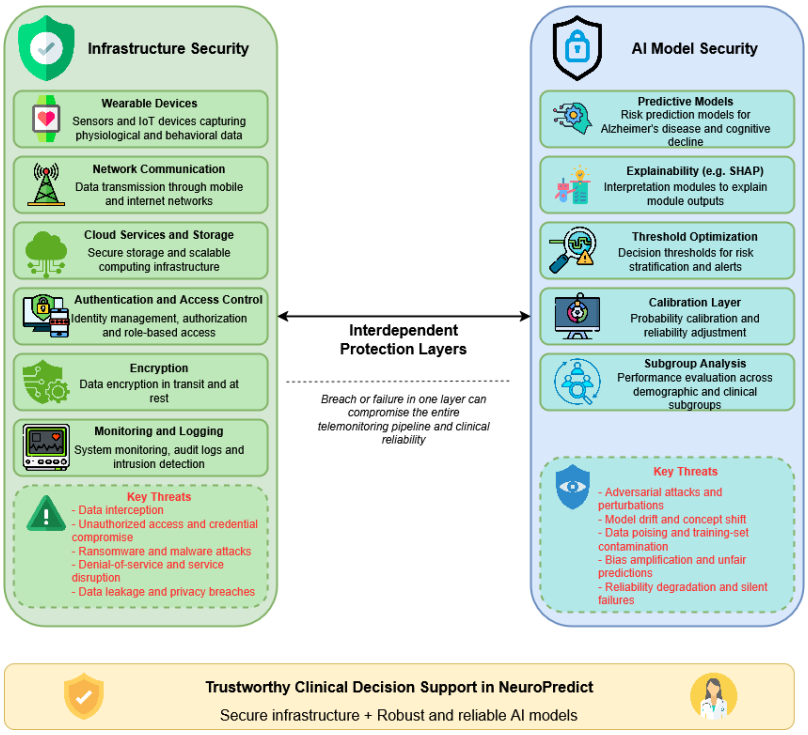


Figure 5. *Conceptual distinction between infrastructure security and AI model security and assurance in the NeuroPredict platform*

interconnected but operate at different layers. A secure infrastructure does not by itself guarantee reliable predictions: model bias, subgroup bias, poor calibration, adversarial manipulation, model drift, or unauthorized model changes may reduce predictive reliability. For AI-based telemonitoring systems, security therefore extends beyond protection from external attacks and includes mechanisms that preserve model trustworthiness (Wirz et al., 2025). Figure 5 summarizes these complementary security and assurance concerns, which motivate the governance measures presented in the *Model-Centric Governance Framework for AI Modules* section.

Risks Associated with AI Model Deployment and Multimodal Analytics

The role of artificial intelligence in telemonitoring ecosystems such as the NeuroPredict platform introduces a set of operational, analytical and decision-support risks that are beyond the typical software-related problems. NeuroPredict continuously collects and analyzes multimodal patient data. As a result, prediction quality depends not only on the technical infrastructure but also on the quality of the input data and the behaviour of the AI models. Poor signal quality, asynchronous measurements, sensor noise, and missing values may affect subsequent processing steps and ultimately reduce prediction accuracy. In addition, errors may be difficult to trace because several analytical components contribute to the final prediction. For this reason, the platform is designed to support model governance through the documentation of model version, lifecycle stage, analytical configuration, inference context, and data quality, as described in the *Model-Centric Governance Framework for AI Modules* section; which of

these elements are currently implemented and which remain proposed is specified in the *Integration and Implementation Status of the Governance Framework within the NeuroPredict Architecture* section.

MODEL-CENTRIC GOVERNANCE FRAMEWORK FOR AI MODULES

Framework Derivation and Conceptual Governance Layer

The *NeuroPredict Platform Architecture and AI-Related Security Challenges* section showed that the NeuroPredict platform combines distributed IoMT-style data acquisition, edge–cloud processing, multimodal harmonization, and embedded AI modules that generate role-specific monitoring outputs for patients, caregivers, physicians, and specialist physicians. In such an architecture, cybersecurity controls alone are not sufficient. Even when communication channels, storage, authentication, and access control remain secure, the reliability of the platform can still be compromised by model drift, model replacement, unstable preprocessing, ungoverned threshold changes, weak traceability of AI outputs, or inconsistent interpretation across modalities. The NeuroPredict platform therefore requires a governance layer that operates specifically at model level and complements, rather than replaces, the infrastructure protections already described in the aforementioned section.

The framework was derived through a conceptual mapping of the AI-related risks identified in the sections spanning *Security Challenges in AI-Enabled Telemonitoring Platforms* and *Risks Associated with AI Model Deployment and Multimodal Analytics* to corresponding model-lifecycle governance requirements. These requirements were then mapped onto the existing NeuroPredict architecture, resulting in the six governance dimensions described below. Specifically, undocumented model replacement and configuration mismatches require artifact identification and configuration linkage through model integrity and provenance; drift and unreviewed updates require validation, approval, monitoring, and rollback through controlled deployment and update governance; and threshold changes or calibration degradation require documented review through threshold and calibration governance. Limited output traceability requires analytical-context preservation through prediction records; asynchronous or incomplete inputs require alignment and completeness checks through multimodal consistency validation; and subgroup performance disparities require disaggregated assessment and accountable review through subgroup/fairness monitoring under human oversight. These risk–requirement relationships underpin the six components in Figure 6. The proposed framework (Figure 6) introduces a model-centric governance layer positioned between the multimodal processing pipeline, embedded AI modules, and role-based outputs of NeuroPredict. This layer consolidates governance-relevant metadata to support documentation, approval, monitoring, review, and suspension when required. Its lifecycle-oriented approach follows health-AI guidance across development, validation, implementation, maintenance, and updating, extending beyond one-time technical evaluation (de Hond et al., 2022; Feng et al., 2022; Papagiannidis et al., 2025).

Figure 6 details the six governance dimensions, their cross-cutting principles, and their links to the model lifecycle and model registry. The layer comprises six governance dimensions: *model integrity and provenance*, *controlled deployment and update governance*, *threshold and calibration governance*, *inference traceability through prediction records*, *multimodal consistency validation*, and *subgroup/fairness monitoring under human oversight*.

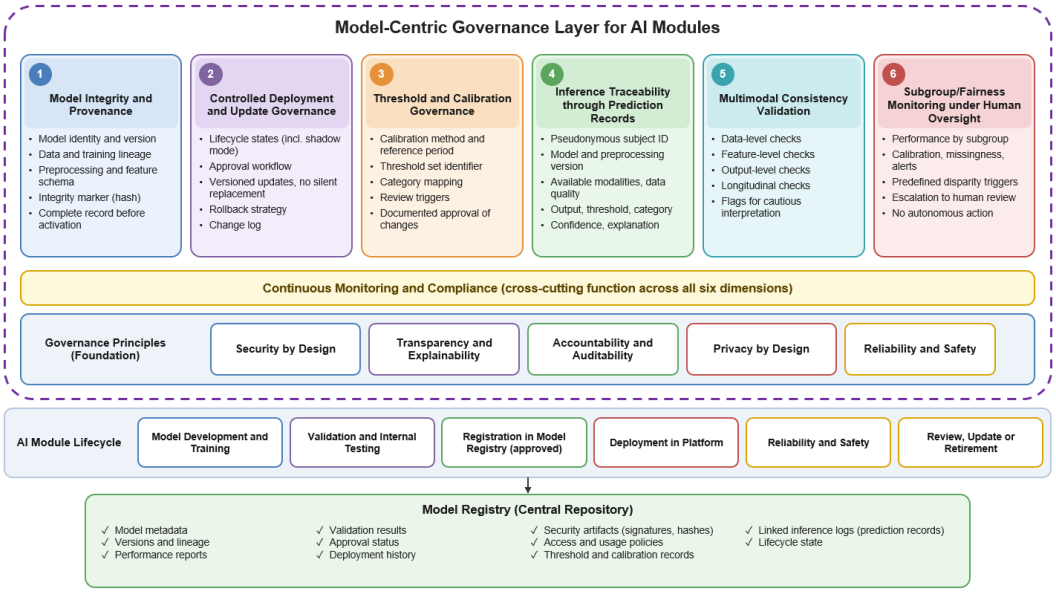


Figure 6. *Model-Centric Governance Layer for AI Modules*

At the current maturity of the NeuroPredict platform, this framework should be understood as a governance scaffold for a laboratory-tested proof-of-concept platform and for future deployment readiness. It does not imply clinical validation, regulatory clearance, or autonomous clinical decision-making. Its function is to ensure that AI modules in the NeuroPredict platform remain identifiable, traceable, reviewable, and accountable as the platform evolves technically and analytically.

Model Integrity, Provenance, and Documentation

Model integrity in the NeuroPredict platform means that an AI module is authentic, approved for a defined analytical purpose, and executed only together with the preprocessing pipeline, feature schema, calibration configuration, and thresholds under which it was evaluated. In a multimodal telemonitoring platform, the model cannot be treated as weights or code alone. It is a compound analytical object that includes modality assumptions, missing-data strategy, feature definitions, output type, explanation method where applicable, and evidence regarding its validated operating context. This is relevant because the *NeuroPredict Platform Architecture and AI-Related Security Challenges* section describes heterogeneous integrated AI modules supporting fall detection, gait analysis, cognitive-risk estimation, and emotion-related analysis, each relying on different data streams and validation assumptions.

For this reason, each AI module should be associated with a provenance record containing, at minimum, the following metadata fields: model identifier; version; developer/owner; intended analytical function; target condition or monitoring task; input modalities; training and validation data sources; temporal scope of the data; preprocessing version; feature schema; missing-data handling strategy; output type; calibration method; performance summary; subgroup performance summary where available; explanation method; current lifecycle state; approval date; review date; known limitations; and integrity marker for

the approved artifact (for example, hash or equivalent checksum). Recent work on health-AI model cards and regulatory transparency supports layered, role-appropriate but verifiable disclosure of intended use, limitations, performance, data characteristics, subgroup evidence, and local maintenance requirements rather than minimal or marketing-oriented summaries (Gilbert et al., 2025; Mehta et al., 2025). Data-quality evidence should be part of the same provenance object because the quality, representativeness, and fitness-for-purpose of the training and validation datasets strongly shape model behaviour and trustworthiness (Schwabe et al., 2024).

The governance requirement derived from the aforementioned section is therefore straightforward: no AI module should be activated in the NeuroPredict platform unless its provenance record is complete, internally consistent, and linked to the approved artifact and preprocessing configuration. Any mismatch between the registered version, deployed artifact, feature schema, or threshold set should trigger review and should be able to prevent activation. This requirement directly answers the risks identified in that section, namely undocumented model change, preprocessing instability, and reduced auditability.

Controlled Deployment, Update Governance, and MLOps-Oriented Lifecycle Control

The deployment issue in the NeuroPredict platform is not only whether a model performs well under laboratory conditions; it is whether the model remains fit when incoming data change because of user heterogeneity, altered sensor behaviour, pipeline modifications, or evolving disease trajectories. The *NeuroPredict Platform Architecture and AI-Related Security Challenges* section already identifies these risks through model drift, distributional change, multimodal incompleteness, and device variability. Accordingly, model activation should be treated as a governed lifecycle transition rather than as a conventional software release. Post-2020 healthcare AI literature consistently argues that model quality degrades in dynamic environments and that monitoring, maintenance, and updating must be built into the operating model rather than added after the fact (Feng et al., 2022; Youssef et al., 2023).

For the current NeuroPredict maturity level, a proposed set of lifecycle states includes: development, internal validation, technical testing, shadow mode, laboratory activation, monitored laboratory use, suspended, and archived. These labels are intentionally aligned with a proof-of-concept platform and avoid implying routine clinical deployment. Transition from one state to the next should depend on predefined acceptance criteria covering predictive performance, calibration, compatibility with the active feature schema, robustness to missing data, stability across relevant subgroups where sample size permits, and consistency with the intended role-specific output presentation. Shadow mode is particularly appropriate for the NeuroPredict platform because it allows a candidate model to process real incoming multimodal data without affecting dashboards or alerts, thereby exposing the model to realistic data variability while limiting user-facing risk.

Updates should be treated as governed versions, never as silent replacement. Each update should carry a documented rationale, such as additional training data, recalibration, detected drift, changed sensor characteristics, altered feature extraction, or evidence of degraded performance. Cross-cutting MLOps-oriented governance requirements should include immutable versioning of models and preprocessing assets, change logs, approval records, rollback capability, separation of development/testing/activation environments, periodic review cadence, and monitored indicators for data

quality, performance, calibration, and alert behaviour. In the *Model-Centric Governance Framework for AI Modules* section these are governance requirements; their concrete platform realization is appropriately deferred to the *Integration and Implementation Status of the Governance Framework within the NeuroPredict Architecture* section.

Threshold and Calibration Governance

Thresholds and calibration require explicit governance in the NeuroPredict platform because some embedded AI modules generate probabilistic outputs that may be converted into monitoring or risk categories, as stated in the *NeuroPredict Platform Architecture and AI-Related Security Challenges* section. In such a setting, a change in threshold or calibration may materially alter platform behaviour even when the underlying predictive model remains unchanged. Governance must therefore treat thresholds, calibration objects, and category mappings as controlled analytical assets rather than as informal parameter settings.

At minimum, the governed threshold/calibration record should include: associated model identifier and version; output definition; calibration method; calibration dataset or reference period; threshold set identifier; category labels; validation rationale; date of approval; review date; and conditions under which recalibration or threshold review is triggered. Review triggers should include materially shifted alert rates, evidence of calibration degradation, changed prevalence in the monitored population, or substantive changes in modality availability relative to the validated setting. This requirement is consistent with broader health-AI guidance that emphasizes careful evaluation, transparent documentation, and continuing maintenance across the full implementation pathway, including local calibration and maintenance information for end users and deployers (de Hond et al., 2022; Feng et al., 2022; Gilbert et al., 2025).

Inference Traceability and Prediction Records

Inference traceability is the governance ability to reconstruct the analytical context in which a specific AI output was generated. In the NeuroPredict platform this is essential because outputs are longitudinal, role-specific, and potentially reinterpreted over time while models, preprocessing pipelines, thresholds, or modality availability evolve. Each AI output should therefore be treated as an auditable inference event rather than as a standalone score.

The minimum prediction record schema should include: pseudonymous subject identifier; timestamp; model identifier and version; preprocessing version; feature schema version; available modalities at inference time; key data-quality indicators; output value or probability vector; applied threshold or threshold-set identifier; resulting category; confidence or uncertainty indicator if available; explanation summary if an explanation method is used; and lifecycle state of the model at the time of inference. For role-based presentation, the same underlying record may support differentiated views for patients, caregivers, physicians, and specialist physicians, but the technical record should remain common and immutable. Recent transparency work in health AI supports the need for accessible yet verifiable model information, while broader governance literature indicates that accountability depends on being able to connect outputs back to the structured practices and controls under which they were produced (Gilbert et al., 2025; Mehta et al., 2025; Papagiannidis et al., 2025).

Prediction records serve three governance purposes in the NeuroPredict platform. First, they support retrospective review when a monitoring output appears implausible or changes abruptly. Second, they preserve comparability over time by distinguishing

genuine subject evolution from changes caused by model version, preprocessing update, recalibration, or threshold change. Third, they support data minimization because traceability does not require indiscriminate retention of all raw signals; it requires preservation of the minimum analytical context necessary for traceability and retrospective review.

Multimodal Consistency Validation and Subgroup/Fairness Monitoring under Human Oversight

The final two governance dimensions respond directly to the multimodal and heterogeneous character of the NeuroPredict platform. The *NeuroPredict Platform Architecture and AI-Related Security Challenges* section emphasizes that the platform combines physiological, cognitive, behavioural, environmental, and potentially affective data, and that inconsistency can emerge from asynchronous acquisition, missing values, noisy sensors, changed device behaviour, and subgroup heterogeneity. Governance must therefore move beyond single-model validation and examine the coherence of the platform's analytical ecosystem as a whole.

The proposed framework defines four levels of multimodal consistency validation. Data-level checks verify modality presence, temporal alignment, basic completeness, and whether current missingness remains within the range tolerated by the validated model. Feature-level checks assess plausibility of derived variables and compatibility with the expected feature schema. Output-level checks do not require agreement across modalities; instead, they identify technically implausible combinations, unexplained discontinuities, or relevant cross-modal discordance that may require further review. Longitudinal checks examine whether observed changes are plausibly attributable to subject evolution or more likely reflect pipeline discontinuity, device replacement, altered adherence, or model instability. This system-level view is consistent with recent work showing that trustworthy medical AI depends on data-quality scrutiny, bias awareness, and local recurring validation under the actual deployment context (Youssef et al., 2023; Schwabe et al., 2024).

Subgroup monitoring should be treated as a formal governance obligation rather than as an optional analytical add-on. The NeuroPredict platform targets heterogeneous users differing in age, sex, comorbidities, functional status, digital literacy, disease stage, and adherence to sensing devices. Recent fairness literature shows that inequities can arise from data acquisition, labelling, population differences, and workflow context, and may remain hidden if performance is only assessed in aggregate (Chen et al., 2023). Accordingly, subgroup review should be triggered whenever sufficient data exist and when predefined disparities emerge in discrimination, calibration, confidence, missingness, or alert frequency. Exact trigger thresholds are implementation-specific and should be defined in the *Integration and Implementation Status of the Governance Framework within the NeuroPredict Architecture* section or in the platform's operating procedures, whereas the governance obligation to monitor them falls within the scope of the *Model-Centric Governance Framework for AI Modules* section.

Finally, multimodal consistency and subgroup monitoring should feed into human oversight rather than into autonomous action. Outputs that are weakly supported, low confidence, or inconsistent across modalities should be flagged for cautious interpretation in role-specific dashboards. Within the current laboratory-tested NeuroPredict platform, this preserves the appropriate balance between AI-supported monitoring and accountable human review without overstating the platform's validation status.

INTEGRATION AND IMPLEMENTATION STATUS OF THE GOVERNANCE FRAMEWORK WITHIN THE NEUROPREDICT ARCHITECTURE

The proposed governance framework incorporates the mechanisms already available in NeuroPredict and complements them with additional governance controls. Model registry and version-management capabilities are currently available, while deployment governance, structured prediction records, and multimodal consistency mechanisms are implemented to different extents. Threshold and calibration governance and formal subgroup/fairness monitoring remain proposed components for subsequent implementation. At the architectural level, the governance layer (Papagiannidis et al., 2025) operates as a transversal control structure positioned between the infrastructure domain and the analytical AI domain (Roberts et al., 2026). The governance layer coordinates the operation of the NeuroPredict platform by working with the existing mechanisms for identity management, access control, secure communication, and data traceability. Beyond infrastructure security, it also supports the management and oversight of AI models integrated into the current laboratory-tested proof-of-concept platform. NeuroPredict maintains a centralized model registry containing information such as the model identifier, version, analytical task, preprocessing pipeline, input modalities, performance metrics, and operational status. The registry also links deployed models to inference records for traceability. The proposed governance framework extends the current model versioning mechanism with lifecycle management functions, including controlled deployment, update documentation, review, suspension, and rollback.

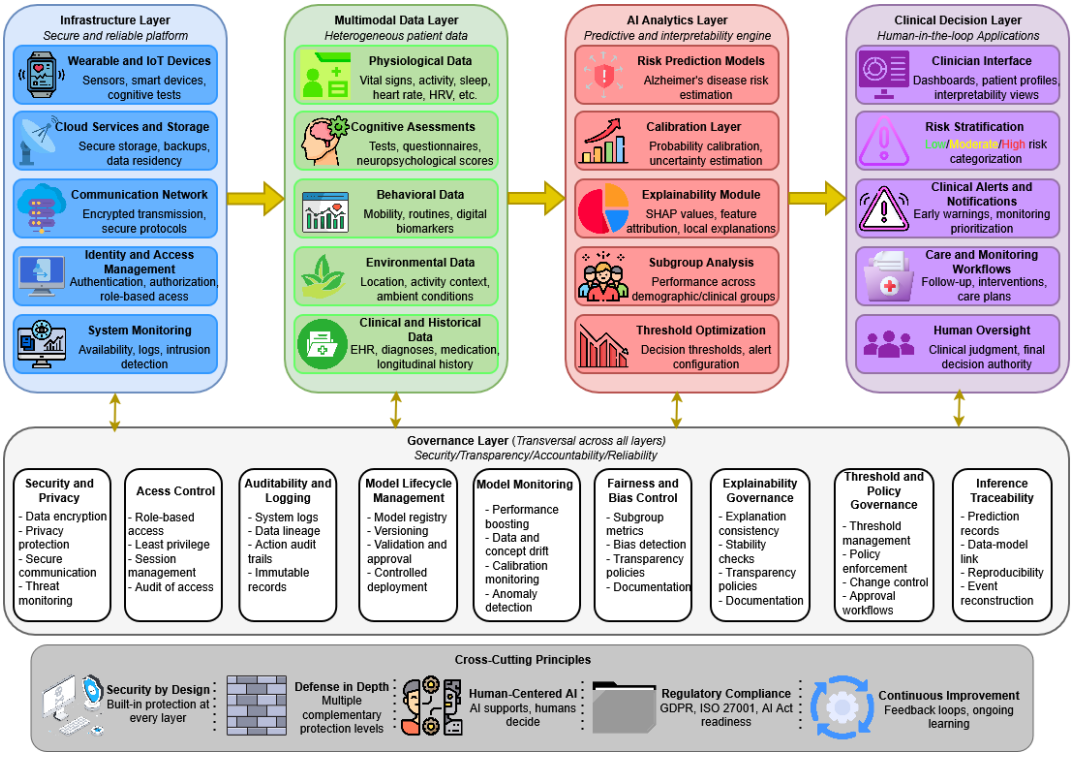


Figure 7. *Proposed integration of the model-centric governance framework within the NeuroPredict architecture*

The proposed governance framework would extend the current NeuroPredict implementation by linking each prediction to its preprocessing settings, calibration parameters, confidence information, and available explainability metadata. At present, only some of these data are recorded, and the complete prediction record is still under development. The framework is designed to be compatible with the platform's existing security mechanisms, namely secure storage, authentication, access control, encrypted communication, and device monitoring. Figure 7 summarizes the proposed governance implementation which aims to improve traceability and reliability while remaining consistent with the proof-of-concept status of the platform and supporting future deployment readiness.

DISCUSSION AND CONCLUSIONS

The proposed framework addresses a governance layer that complements conventional infrastructure cybersecurity in AI-enabled telemonitoring platforms. While secure communication, authentication, access control, and encrypted storage protect the technical infrastructure, they do not ensure that AI outputs are generated by approved model versions, validated preprocessing pipelines, controlled thresholds, and traceable inference configurations. The proposed model-centric governance framework addresses these requirements by treating AI models and their associated configurations as lifecycle-managed analytical assets.

Benefits for AI-Enabled Telemonitoring Platforms

The framework connects infrastructure-level cybersecurity with model-level assurance by consolidating existing model-management and traceability capabilities with additional lifecycle controls within a unified governance structure. Its contribution therefore lies not in introducing model versioning or inference logging as standalone mechanisms, but in coordinating these capabilities with provenance, deployment control, threshold and calibration governance, multimodal validation, and subgroup-aware human oversight.

Limitations and Future Research Directions

The framework has not yet been implemented and evaluated as a complete governance subsystem. It combines mechanisms already available within NeuroPredict with partially implemented and proposed governance controls. Current laboratory evaluations cannot establish the effectiveness of the proposed mechanisms for detecting model drift, calibration degradation, subgroup-specific performance differences, or clinically relevant multimodal inconsistencies under real-world conditions. Moreover, exact inference reproducibility would require additional mechanisms for preserving or securely referencing input data, feature representations, runtime configurations, and model artifacts. Therefore, the framework should be regarded as an architectural foundation for further implementation and validation rather than evidence of clinical deployment readiness.

Future work will focus on extending the existing model registry toward broader governance functions, completing structured prediction records and controlled lifecycle transitions, and implementing explicit threshold/calibration governance and subgroup/fairness monitoring. Further evaluation will address model drift, calibration stability, multimodal discordance, and subgroup-level performance. These mechanisms will subsequently be assessed together with the clinical validation of NeuroPredict to determine their technical feasibility, operational usefulness, and contribution to transparent and auditable AI-supported monitoring.

Conclusions

This study formulated a conceptual model-centric governance framework for AI modules in multimodal telemonitoring platforms and mapped its requirements to the NeuroPredict architecture, distinguishing existing capabilities from partially implemented and proposed controls. The framework comprises six complementary governance dimensions: model integrity and provenance, controlled deployment and update governance, threshold and calibration governance, inference traceability through prediction records, multimodal consistency validation, and subgroup/fairness monitoring under human oversight. Together, these dimensions specify requirements for controlling model identity, analytical configuration, lifecycle transitions, and the preservation of inference context.

The principal contribution is the architectural coordination of existing infrastructure security, model-management, and inference-logging capabilities with additional model-lifecycle governance requirements. The proposed approach links model artifacts and preprocessing configurations to deployment approvals, controlled threshold and calibration settings, structured prediction records, multimodal consistency checks, and subgroup-specific performance review. This coordination extends infrastructure-level protection with model-level assurance and accountable human oversight.

At the current laboratory-tested proof-of-concept stage, the integrated governance subsystem remains to be completed and empirically evaluated. The framework provides a structured basis for staged implementation and systematic technical assessment of configuration integrity, inference traceability, calibration stability, multimodal consistency, and subgroup performance. These evaluations will inform further refinement of the governance mechanisms and support subsequent clinical validation of the NeuroPredict platform.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contribution of the Romanian Ministry of Research, Innovation, and Digitization for funding the development of the NeuroPredict platform inside the project “Advanced Artificial Intelligence Techniques in Science and Applications” (contract no. 13N/2023 (PN 23 38 05 01)) for the period 2023–2026.

REFERENCE LIST

Aravindkumar, A., Ramadoss, M., Fakhruddin Ahmed, S. A., Sampath, V., & Lakshminarayanan, K. (2026). Explainable AI in healthcare: a systematic review of XAI use cases in imaging, diagnostics, and rehabilitation. *Frontiers in Artificial Intelligence*, *9*. https://doi.org/10.3389/frai.2026.1749527

Băjenaru, L., Săndulescu, V., Petrache, C., & Nicolau, D. (2025). Role-Based Secure Telehealth in IoMT Ecosystems: Design Insights from NeuroPredict Platform. *Romanian Cyber Security Journal*, *7*(2), 29–41. https://doi.org/10.54851/v7i2y202503

Băjenaru, O. L., Băjenaru, L., Ianculescu, M., Constantin, V. Ștefan, Gușatu, A. M., & Nuță, C. R. (2024). Geriatric Healthcare Supported by Decision-Making Tools Integrated into Digital Health Solutions. *Electronics 2024, Vol. 13, Page 3440*, *13*(17), 3440. https://doi.org/10.3390/ELECTRONICS13173440

Băjenaru, L., Valeanu, L., Ianculescu, M., Gușatu, AM., Alexandru, A. (2026). Bridging Emotion Recognition and Neurodegenerative Care: A Voice-Based AI Module within the NeuroPredict Platform. In: Costin, HN.,

Magjarevic, R., Petroiu, G.G. (eds) Advances in Digital Health and Medical Bioengineering II. EHB 2025. IFMBE Proceedings, vol 143. Springer, Cham. https://doi.org/10.1007/978-3-032-24045-3_43

Brohi, S., & Mastoi, Q. (2025). From Accuracy to Vulnerability: Quantifying the Impact of Adversarial Perturbations on Healthcare AI Models. *Big Data and Cognitive Computing*, *9*(5), 114. https://doi.org/10.3390/bdcc9050114

Butler, P. M., Yang, J., Brown, R., Hobbs, M., Becker, A., Penalver-Andres, J., Syz, P., Muller, S., Cosne, G., Juraver, A., Song, H. H., Saha-Chaudhuri, P., Roggen, D., Scotland, A., Silveira, N., Demircioglu, G., Gabelle, A., Hughes, R., Erkkinen, M. G., … Belachew, S. (2025). Smartwatch- and smartphone-based remote assessment of brain health and detection of mild cognitive impairment. *Nature Medicine*, *31*(3), 829–839. https://doi.org/10.1038/s41591-024-03475-9

Chen, R. J., Wang, J. J., Williamson, D. F. K., Chen, T. Y., Lipkova, J., Lu, M. Y., Sahai, S., & Mahmood, F. (2023). Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nature Biomedical Engineering*, *7*(6), 719–742. https://doi.org/10.1038/s41551-023-01056-8

de Hond, A. A. H., Leeuwenberg, A. M., Hooft, L., Kant, I. M. J., Nijman, S. W. J., van Os, H. J. A., Aardoom, J. J., Debray, T. P. A., Schuit, E., van Smeden, M., Reitsma, J. B., Steyerberg, E. W., Chavannes, N. H., & Moons, K. G. M. (2022). Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *Npj Digital Medicine*, *5*(1), 2. https://doi.org/10.1038/s41746-021-00549-7

Feng, J., Phillips, R. V., Malenica, I., Bishara, A., Hubbard, A. E., Celi, L. A., & Pirracchio, R. (2022). Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. *Npj Digital Medicine*, *5*(1), 66. https://doi.org/10.1038/s41746-022-00611-y

Gashi, S., Oldrati, P., Moebus, M., Hilty, M., Barrios, L., Ozdemir, F., Kana, V., Lutterotti, A., Rätsch, G., & Holz, C. (2024). Modeling multiple sclerosis using mobile and wearable sensor data. *Npj Digital Medicine*, *7*(1), 64. https://doi.org/10.1038/s41746-024-01025-8

Gilbert, S., Adler, R., Holoyad, T., & Weicken, E. (2025). Could transparent model cards with layered accessible information drive trust and safety in health AI? *Npj Digital Medicine*, *8*(1), 124. https://doi.org/10.1038/s41746-025-01482-9

Hemalatha, S., Adavala, K. M., Kumar, P. S. V. V. S. R., Kumaravel, P., Pillai, N. M., & Mohan, G. K. (2025). An Architectural Framework for Telemedicine Systems: Components, Roles, and Implementation Challenges. Telehealth and Medicine Today, 10(3). https://doi.org/10.30953/thmt.v10.596

Ianculescu, M., Băjenaru, L., Vasilevschi, A.-M., Gheorghe-Moisii, M., & Gheorghe, C.-G. (2025). Advancing Neurodegenerative Disease Management: Technical, Ethical, and Regulatory Insights from the NeuroPredict Platform. *Future Internet*, *17*(7), 320. https://doi.org/10.3390/fi17070320

Ianculescu, M., Petean, C., Petrache, C., & Băjenaru, L. (2025). An Embedded Facial Emotion Recognition Module for Monitoring Affective States in Multiple Sclerosis. *2025 29th International Conference on System Theory, Control and Computing (ICSTCC)*, 607–612. https://doi.org/10.1109/ICSTCC66753.2025.11240267

Katz, C., Ruiz, J. M., Saigí-Rubió, F., & Novillo-Ortiz, D. (2025). The State of the Art of Telemedicine Implementation Architecture: Rapid Umbrella Review of Systematic Reviews. *Journal of Medical Internet Research*, *27*, e70276. https://doi.org/10.2196/70276

Lipsmeier, F., Taylor, K. I., Postuma, R. B., Volkova-Volkmar, E., Kilchenmann, T., Mollenhauer, B., Bamdadian, A., Popp, W. L., Cheng, W.-Y., Zhang, Y.-P., Wolf, D., Schjodt-Eriksen, J., Boulay, A., Svoboda, H., Zago, W., Pagano, G., & Lindemann, M. (2022). Reliability and validity of the Roche PD Mobile Application for remote monitoring of early Parkinson's disease. *Scientific Reports*, *12*(1), 12081. https://doi.org/10.1038/s41598-022-15874-4

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

Mehta, V., Komanduri, A., Bhadouriya, R. S., Mehta, V., Johnson, M. D., Shrestha, P., Nikolov, M., Jain, B., Shah, N., & Schulman, K. (2025). Evaluating transparency in AI/ML model characteristics for FDA-reviewed medical devices. *Npj Digital Medicine*, *8*(1), 673. https://doi.org/10.1038/s41746-025-02052-9

Mihăescu, A.-G., Petrache, M.-C., & Bica, O. (2026). Explainable Machine Learning Approaches for Alzheimer's Disease Detection. In: Costin, HN., Magjarevic, R., Petroiu, G.G. (eds) Advances in Digital Health and Medical Bioengineering II. EHB 2025. IFMBE Proceedings, vol 143. Springer, Cham. https://doi.org/10.1007/978-3-032-24045-3_44

Mitan, E., Neacșu, D., Ovreiu, S., & Savu, D. (2025). A Multidisciplinary Approach to Trustworthy AI in the Digital Society. Romanian Cyber Security Journal, 7(2), 43–66. https://doi.org/10.54851/v7i2y202504

Muurling, M., de Boer, C., Hinds, C., Atreya, A., Doherty, A., Alepopoulos, V., Curcic, J., Brem, A.-K., Conde, P., Kuruppu, S., Morató, X., Saletti, V., Galluzzi, S., Vilarino Luis, E., Cardoso, S., Stukelj, T., Kramberger, M. G., Roik, D., Koychev, I., … Visser, P. J. (2024). Feasibility and usability of remote monitoring in Alzheimer's disease. *DIGITAL HEALTH*, *10*. https://doi.org/10.1177/20552076241238133

Muurling, M., de Boer, C., Kozak, R., Religa, D., Koychev, I., Verheij, H., Nies, V. J. M., Duyndam, A., Sood, M., Fröhlich, H., Hannesdottir, K., Erdemli, G., Lucivero, F., Lancaster, C., Hinds, C., Stravopoulos, T. G., Nikolopoulos, S., Kompatsiaris, I., Manyakov, N. V., … Visser, P. J. (2021). Remote monitoring technologies in Alzheimer's disease: design of the RADAR-AD study. *Alzheimer's Research & Therapy*, *13*(1), 89. https://doi.org/10.1186/s13195-021-00825-4

Narula, S., Ghasemigol, M., Carnerero-Cano, J., Minnich, A., Lupu, E., & Takabi, D. (2025). Exploring Research and Tools in AI Security: A Systematic Mapping Study. *IEEE Access*, *13*, 84057–84080. https://doi.org/10.1109/ACCESS.2025.3567195

Omberg, L., Chaibub Neto, E., Perumal, T. M., Pratap, A., Tediarjo, A., Adams, J., Bloem, B. R., Bot, B. M., Elson, M., Goldman, S. M., Kellen, M. R., Kieburtz, K., Klein, A., Little, M. A., Schneider, R., Suver, C., Tarolli, C., Tanner, C. M., Trister, A. D., … Mangravite, L. M. (2022). Remote smartphone monitoring of Parkinson's disease and individual response to therapy. *Nature Biotechnology*, *40*(4), 480–487. https://doi.org/10.1038/s41587-021-00974-9

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, *34*(2), 101885. https://doi.org/10.1016/j.jsis.2024.101885

Powers, R., Etezadi-Amoli, M., Arnold, E. M., Kianian, S., Mance, I., Gibiansky, M., Trietsch, D., Alvarado, A. S., Kretlow, J. D., Herrington, T. M., Brillman, S., Huang, N., Lin, P. T., Pham, H. A., & Ullal, A. V. (2021). Smartwatch inertial sensors continuously monitor real-world motor fluctuations in Parkinson's disease. *Science Translational Medicine*, *13*(579). https://doi.org/10.1126/scitranslmed.abd7865

Roberts, H., Taddeo, M., & Floridi, L. (2026). A Framework for Evaluating Global Governance Initiatives. *Global Policy*, *17*(2), 189–199. https://doi.org/10.1111/1758-5899.70164

Sandulescu, V., Ianculescu, M., Valeanu, L., & Alexandru, A. (2024). Integrating IoMT and AI for Proactive Healthcare: Predictive Models and Emotion Detection in Neurodegenerative Diseases. *Algorithms*, *17*(9), 376. https://doi.org/10.3390/a17090376

Schwabe, D., Becker, K., Seyferth, M., Klaß, A., & Schaeffter, T. (2024). The METRIC-framework for assessing data quality for trustworthy AI in medicine: a systematic review. *Npj Digital Medicine*, *7*(1), 203. https://doi.org/10.1038/s41746-024-01196-4

Tîrziu, E., Vasilevschi, A.-M., Alexandru, A., & Tudora, E. (2024). Enhanced Fall Detection Using YOLOv7-W6-Pose for Real-Time Elderly Monitoring. *Future Internet*, *16*(12), 472. https://doi.org/10.3390/fi16120472

Tîrziu, E., Vasilevschi, A.-M., Alexandru, A., & Tudora, E. (2025). Real-Time Fall Monitoring for Seniors via YOLO and Voice Interaction. *Future Internet*, *17*(8), 324. https://doi.org/10.3390/fi17080324

Vevera, A.-V., Ianculescu, M., & Alexandru, A. (2025). A Digital Health Platform for Remote and Multimodal Monitoring in Neurodegenerative Diseases. *Future Internet*, *17*(12), 571. https://doi.org/10.3390/fi17120571

Wirz, C. D., Demuth, J. L., Bostrom, A., Cains, M. G., Ebert-Uphoff, I., Gagne, D. J., Schumacher, A., McGovern, A., & Madlambayan, D. (2025). (Re)Conceptualizing trustworthy AI: A foundation for change. *Artificial Intelligence*, *342*, 104309. https://doi.org/10.1016/j.artint.2025.104309

Youssef, A., Pencina, M., Thakur, A., Zhu, T., Clifton, D., & Shah, N. H. (2023). External validation of AI models in health should be replaced with recurring local validation. *Nature Medicine*, *29*(11), 2686–2687. https://doi.org/10.1038/s41591-023-02540-z



This is an open access article distributed under the terms and conditions of the
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.