Smart CHANGE
Empowering Youth with AI for healthier lives

ai·prognosis

trustroke

Umbrella

Trustworthy AI for Transformative Healthcare:

Experience from Four European Innovation Projects

WHITE PAPER - SEPTEMBER 2026

Funded by the European Union

The four projects are co-funded by the European Union under the Horizon Europe research and innovation programm.

Executive Summary

Artificial intelligence (AI) is poised to revolutionise healthcare by enabling earlier diagnosis, more personalised treatment, and efficient prevention strategies. Across Europe, a cluster of EU-funded research projects is advancing this transformation through applications in cardiology, neurology, metabolic health, and digital platforms. This white paper synthesises insights from these projects, drawing attention to shared challenges, solutions, and pathways for scaling. Key cross-cutting themes include the trustworthiness and explainability of AI, the need for federated and privacy-preserving data use, and the critical importance of user-centred design. A market study and reflections on strategic positioning conclude the white paper.

Glossary of Terms

In this section, key technical terms and concepts utilised by the projects are defined to ensure reader clarity, alignment and consistency with the Trustworthy AI theme.

Artificial Intelligence (AI). A field of computer science dedicated to creating systems capable of performing tasks that typically require human intelligence, such as recognising patterns, making decisions, and solving complex problems. In healthcare, AI utilises advanced algorithms and machine learning to process vast amounts of medical data—ranging from genetic sequences to lifestyle habits—to identify health trends, automate diagnostics, and forecast individual patient risks with greater speed and precision than traditional clinical methods.

Machine learning (ML). In medical applications, ML is a subset of AI that uses statistical and mathematical modelling techniques to enable computer systems to automatically discover patterns and learn from historical medical data without being explicitly programmed for every specific task.

Federated Learning (FL). A decentralised machine learning approach that enables multiple healthcare institutions to collaboratively train a shared artificial intelligence model without exchanging or centralising raw patient data.

Clinical Decision Support System (CDSS). An interactive software application designed to assist healthcare professionals in making clinical decisions.

Non-communicable Disease (NCD). Also known as chronic diseases, NCDs are medical conditions that are not transmissible directly from person to person. They typically progress slowly and last for a long duration, arising from a complex combination of genetic, physiological, environmental, and behavioural factors. Most NCD-related deaths are attributed to cardiovascular diseases (e.g., heart attacks and stroke), cancers, chronic respiratory diseases (e.g., asthma), and diabetes.

Risk prediction. In the context of healthcare and NCDs, risk prediction is defined as the process of estimating the probability that an individual will experience a specific health outcome—such as developing a chronic condition—based on their unique combination of biological, environmental, and behavioural factors. AI-driven risk prediction excels at identifying complex relationships between modifiable risk factors—such as tobacco use and physical inactivity—and non-modifiable factors like genetics to forecast disease trajectories more accurately than traditional statistical methods.

Explainable AI (XAI). A set of methods and techniques that make the outputs and internal decision-making processes of AI systems transparent and understandable to humans. In healthcare, XAI is critical because it allows clinicians to verify the rationale behind a diagnostic or treatment recommendation rather than relying on an opaque "black box" model.

Counterfactual. A counterfactual is a "what-if" scenario that describes the smallest change in input features required to flip a model's prediction to a different, typically more desirable, outcome. In health-related applications, counterfactuals act as personalised clinical prescriptions, showing what specific changes in a person's behaviour would flip a high-risk prediction to a low-risk one.

FAIR Principles. A set of guiding standards—Findable, Accessible, Interoperable, and Reusable—designed to improve data management and stewardship. In a medical context, these principles ensure that clinical and research data are structured for "machine-actionability," allowing different systems to find and use the data efficiently to accelerate innovation and collaboration.

Synthetic Data: Artificially generated information that mimics the statistical properties and patterns of real-world health data without containing any information from actual individuals. It is primarily used to protect patient privacy while providing high-quality datasets for training machine learning models or conducting clinical research when real data is scarce or sensitive.

Trustworthy AI. A framework for developing and deploying AI systems that are reliable, ethical, and aligned with human values.

1. Introduction: Trustworthy Artificial Intelligence in Healthcare

The spectacular advances in artificial intelligence technologies in recent years have naturally led to an interest in its application in a variety of sectors. This is a white paper written by members of the **StayHealthy Cluster** of projects that are investigating the use of artificial intelligence (AI) in healthcare.

Healthcare is an example of a *mission-critical* sector (along with others such as transportation), whereby the **trustworthiness** of the technologies is paramount. In line with established EU frameworks, trustworthiness can be understood as resting on lawfulness, ethical alignment, and technical robustness, with **explainability** acting as one (important) means to support trust rather than as an end in itself.

In this context, these dimensions can be further specified within the EU regulatory and ethical framework. Lawfulness refers to compliance with applicable legal regimes, including the General Data Protection Regulation (notably requirements on lawful basis, the processing of health data under Article 9, data minimisation and data protection impact assessments), the AI Act (in particular obligations applicable to high-risk AI systems, such as risk management, data governance and human oversight) and, where relevant, the Medical Device Regulation. Ethical alignment relates to respect for fundamental rights as set out in the Charter of Fundamental Rights of the European Union and further developed in the Ethics Guidelines for Trustworthy AI, including in particular the mitigation of bias, the prevention of discriminatory outcomes and the meaningful involvement of human oversight. Technical robustness concerns the reliability and safety of systems, including their accuracy, resilience and auditability, especially in safety-critical clinical environments. These aspects are also reflected in recurring sources of public mistrust, such as concerns over biased outcomes (e.g. along gender, ethnicity or age), limited transparency and risks linked to the processing and secondary use of health data.

Trustworthiness is at the very heart of healthcare in general, and AI-based healthcare in particular. The projects deal with a variety of key stakeholders, including health care professionals (ranging from doctors to clinical staff) and patients (ranging from infants and children up to mature adults). Some of the projects deal with families. All of these stakeholders need to be able to trust the technologies that are being deployed on their behalf, and to be able to do so they need to trust the institutions and people that have developed such technologies. Doctors need to be able to trust the diagnoses of an AI-based clinical decision support system. Parents need to be able to trust the AI-based recommendations of an application for their child's healthy lifestyle, perhaps by receiving a thorough explanation of the reasoning that led to those recommendations. Patients in critical care need to be able to trust the AI-based technology supporting their care not only to be technically sound, but to guarantee respect for their basic human rights to being treated ethically and within the bounds of law.

The primary objective of this white paper is to elicit and discuss common challenges, obstacles, solutions, and insights among the projects around the overarching theme of **Trustworthy AI in healthcare**. The projects are following the legal and ethical frameworks put in place by the European Union both for the management of artificial intelligence and for the use of digital technologies in medical devices where applicable.

This is the first of two white papers authored by the cluster: in this first white paper, the cluster presents its constituent projects (their objectives, uses, roadmaps, and challenges); whereas the second Whitepaper will focus on the results obtained, experiences, recommendations, and impact.

2. Methodology

In an initial round, the projects provided detailed descriptions of their objectives and experiences to date, familiarising each other with the overall thematic content present within the cluster. Then, in a second round of analysis, emerging common themes were identified for further description by the involved partners. Based upon this identified thematic landscape, more specific contributions from the involved partners took place, particularly in critical areas such as ethical and clinical considerations, where the challenges and potential approaches were identified and analysed. An overview of the exploitation landscape – both in market-oriented and social / public service contexts – was undertaken to arrive at a common understanding of needs for future activities beyond the scope of the individual projects in the cluster.

3. Thematic Landscape of the Cluster Projects

AI is reshaping the future of healthcare, with the potential to enhance clinical decision-making, improve health outcomes, and alleviate system-wide burdens. Yet, this promise is tempered by complex challenges—fragmented data systems, lack of trust in AI predictions, privacy concerns, and the slow pace of regulatory adoption to emerging technologies. These projects respond with co-creation methodologies, FAIR data principles, and rigorous certification processes. Stakeholder engagement and transparent validation are key strategies for turning prototypes into trusted tools.

The following sections are at the heart of this white paper, reporting on project experiences in the main thematic areas addressed by the cluster.

3.1 Personalised Risk Modelling Across the Lifespan

Although digitisation (e.g., Personal Electronic Health Records) had already enabled the promise of personalised healthcare, the advent of powerful AI technologies has amplified that promise many times over. Projects such as **SmartCHANGE**, **AI-PROGNOSIS**, **UMBRELLA**, and **TRUSTroke** are leveraging diverse datasets—genetic, clinical, behavioural, imaging—to create personalised risk profiles. These efforts support a vision of proactive healthcare, where risks related to an individual's personal characteristics are identified and addressed long before critical illness manifests itself.

One of the major challenges of working with personalised risk profiles is the datasets themselves: how to obtain the personalised data? The **SmartCHANGE** project is working with a combination of mobile devices and wearables in this regard, an approach which is gaining considerable traction in recent years, due also to the popularity of mobile devices among youth. On the mobile device, an app called "HappyPlant" presents the user with a traditional questionnaire that collects general personal data such as height, weight, and relevant habits such as estimated nightly hours of sleep. The app connects with a wearable device (wristwatch) that monitors physical parameters over the course of the day. This data is utilised to drive a gamified set of health challenges for the user – growing a "happy and healthy plant", with visually pleasing graphics. The app communicates in turn with the risk prediction module to generate personalised health risk predictions and proposed mitigations over the lifetime of the user.

SmartCHANGE has encountered numerous technical challenges in working with this approach:

- The wearables are commercial devices that come in many different configurations. In the specific case of SmartCHANGE, the wearables are the Garmin brand whose technical characteristics can vary across the brand itself but also across place of manufacture and sale. It has proven challenging to accommodate all of the different technical parameters.
- Producing a compelling gamification experience is challenging with respect to both technical and psychological factors. Much care must be invested to ensure that the experience is motivating rather than discouraging for the user.
- Even when the local personal data is successfully obtained from the mobile devices, it remains challenging to combine the data with suitable historical data to generate viable risk predictions through AI. The historical datasets can have gaps which need to be filled (perhaps through synthetic data) and may be deficient in some longitudinal areas, necessitating critical decisions (such as which age ranges over the lifetime to take into consideration in the risk predictions).

AI-PROGNOSIS shows that personalised AI in healthcare is not limited to estimating an individual's risk of future disease but can also extend to modelling how a condition may progress over time and how different individuals may respond to treatment. Working across these different prediction tasks required the project to integrate multimodal longitudinal data, including clinical, genetic, behavioural and wearable-derived information, in order to move from static population-level estimates toward

more dynamic and personalised profiles. This, however, brought a number of practical challenges: heterogeneous datasets had to be harmonised despite differences in variables, follow-up structures and data quality; real-world sensor data had to be transformed into stable and clinically meaningful features; and model performance had to be assessed not only in terms of accuracy, but also calibration, fairness and generalisability across independent datasets. The project's experience suggests that the real challenge in personalised modelling is not only to build predictive models, but to ensure that these models remain reliable, interpretable and clinically meaningful when deployed in realistic settings. In that sense, personalised modelling ultimately converges with the overarching theme of Trustworthy AI in healthcare: useful models must be built on robust data pipelines, evaluated for bias and uncertainty, aligned with user needs, and communicated in ways that support confident human decision-making rather than opaque automation.

UMBRELLA addresses personalised risk modelling within the context of stroke care, where timely and accurate predictions can have a direct impact on treatment decisions and patient outcomes. The project is integrating retrospective and prospective data from multiple European centres through a federated infrastructure, enabling the development and validation of trustworthy AI models without centralising sensitive health data. A particular challenge lies in harmonising heterogeneous clinical protocols, imaging data and care pathways across healthcare systems while ensuring that predictive models remain clinically relevant, transparent and transferable across different settings. These efforts highlight the close relationship between personalised medicine, interoperability and trustworthy AI.

3.2 Explainable AI and Clinical Decision Support

Clinicians' reluctance to adopt AI often stems from concerns over black-box systems and lack of integration into workflows. **All projects** directly address these concerns through explainable AI (XAI), transparent design, and user co-creation. These projects are developing clinical decision support systems (CDSS) that assist rather than replace clinical judgment.

SmartCHANGE approaches the challenge of Explainable AI through the technique of *counterfactuals*, whereby the desirable effects of changes to behaviour (such as "more exercise") are estimated. This technique is considered a promising way to interact with health professionals to explain the results produced by the AI models. Through highly visual interaction, a doctor can investigate the subject's counterfactual cohort to better understand their situation and the various potential directions they could take. Nonetheless, the visual explanation of AI model behaviour to health practitioners remains mostly uncharted territory. **SmartCHANGE** is carefully monitoring the results of its feasibility studies across multiple countries to gather feedback directly from the health professionals, which will be used to modify and improve the user interface and the underlying principles of operation of the Visual Explainer.

An important part of the user onboarding experience designed to create trust in the AI support mechanisms is the **SmartCHANGE** *participatory design protocol*, whereby users directly participate in the process of entering the system while interacting with our personnel to ensure that the trust of the participant in the overall experience is maintained. A challenge in this regard has been to create a participatory design protocol that can be deployed in multiple countries and multiple cultural contexts.

AI-PROGNOSIS addresses explainability in the context of clinical decision support, where AI outputs are intended to inform rather than replace medical judgment. Explainability is therefore treated less as a purely technical feature and more as a way to support transparency and meaningful human oversight, in line with the requirements of the General Data Protection Regulation and the broader framework of Trustworthy AI. In practice, this means ensuring that model outputs can be interpreted by clinicians in a way that fits their professional responsibilities, while avoiding over-reliance on automated results. This is particularly important in preserving the clinician's role as the final decision-maker and in ensuring that AI systems remain supportive tools within existing clinical workflows.

UMBRELLA similarly places clinical decision support at the centre of its AI strategy. In stroke care, clinical decisions often need to be made rapidly and on the basis of complex and heterogeneous information. The project is therefore exploring how trustworthy AI models can support clinicians in diagnosis, risk assessment and treatment planning while remaining aligned with existing clinical workflows. Particular

emphasis is placed on ensuring that AI-generated outputs can be validated locally within participating healthcare institutions and deployed in a manner that supports, rather than replaces, professional judgement. This approach reflects the broader objective of creating scalable and trustworthy decision-support tools that can be adopted across diverse European healthcare environments.

3.3 Data Privacy, Ethics, and Regulation

Privacy-preserving and ethical data use is a cornerstone across the cluster. **SmartCHANGE** and **TRUSTroke** employ FL and anonymisation to ensure security and compliance with European standards. The emphasis on FAIR data, GDPR alignment, and explainability ensures that innovation aligns with public trust and legal frameworks.

The regulatory environment is increasingly dense, with the convergence of the GDPR, Medical Device Regulation, and the AI Act. **SmartCHANGE** is addressing issues of privacy with FL and techniques of anonymisation. **SmartCHANGE** is also examining the emerging regulatory environment concerning different kinds of applications, such as "wellness applications" and "medical devices", which have different levels of regulatory implications, to optimise the allocation of resources dedicated to compliance. A major challenge in this respect is certainly the rapidly evolving character of the regulatory context.

AI-PROGNOSIS employs a procedural ethics approach. Part of this means dynamically aligning with developments of frameworks and regulations, such as the EU AI Act, the best practices outlined in FDA guidance on AI-enabled medical devices, the FUTURE-AI international consensus guidelines, and ISO/IEC standards relevant to Trustworthy AI development, and promoting usage of the TRIPOD-AI checklist as a primary standard for transparent and comprehensive reporting. As another part of this procedural approach, AI-PROGNOSIS adopts systematic engagement with key stakeholders, and the identification, assessment, addressing in system design, and updating of ethical concerns not necessarily covered by current legal and policy frameworks, as an iterative and circular process.

Building on this approach, **AI-PROGNOSIS** adopts a structured framework for data protection and ethical governance, reflecting the challenges associated with processing health data in a clinical research setting. This includes alignment with the General Data Protection Regulation through clearly defined legal bases, primarily explicit consent under Article 9, together with purpose limitation and data minimisation, supported by a joint controllership framework under Article 26 GDPR. The project also recognises that legal compliance does not in itself ensure ethical adequacy and therefore integrates ongoing ethical assessment throughout the research lifecycle. This approach is further reflected in the implementation of appropriate technical and organisational measures, including pseudonymisation, controlled access and secure data processing environments. It also supports a risk-based approach to data governance, considering the sensitivity of health data and the potential impact on data subjects' rights and freedoms.

In **UMBRELLA**, privacy-preserving and ethical data use are central to project, which is developing both a federated real-world data platform and a federated learning infrastructure to support cross-border collaboration while preserving the privacy of stroke patients and survivors and keeping sensitive data under local control. UMBRELLA is also addressing the practical challenges associated with regulatory readiness through the development of roadmaps towards future compliance and certification. By combining privacy-preserving technologies with early consideration of governance and regulatory requirements, the project demonstrates how trustworthy AI development must integrate technical design, interoperability and compliance from the outset.

3.4 Market Overview

The global AI healthcare market was valued at over $36 billion in 2025 and is projected to grow to over $500 billion by 2033. Europe's strong public health infrastructure and regulatory leadership make it a prime region for AI-health deployment. However, regulatory bottlenecks (including complex certification processes) and an often research-driven rather than entrepreneurial culture require strategic alignment.

3.5 Segments of Opportunity

Segment	Market CAGR	Relevant Projects
AI-based risk prediction in health	35–40%	SmartCHANGE, AI-PROGNOSIS
Clinical decision support systems (CDSS)	10-13%	TRUSTroke, UMBRELLA
Mobile digital health apps	25-30%	AI-PROGNOSIS, TRUSTroke
Synthetic data generation	Circa 40%	SmartCHANGE, TRUSTroke, UMBRELLA

3.6 Barriers to Entry

Common barriers include low clinician trust, lack of interoperability, data silos, and regulatory burden. These can be addressed through explainable AI, FL, and CE-oriented development, though hospitals across Europe present heterogeneous technical and regulatory contexts.

3.7 Strategic Positioning

Successful exploitation in both market and public service contexts depends on:

- Trustworthiness as a differentiator
- Scalability through interoperability
- Targeting both clinical and citizen users
- Validation via real-world evidence

4. Conclusion

Together, these projects demonstrate how responsible, explainable, and privacy-preserving AI can transform European healthcare. Through technical innovation and human-centred design, they pave the way for tools that are not only powerful but also trustworthy. With continued collaboration and support, these initiatives can shape a future where AI is a cornerstone of ethical, effective healthcare.

Appendix: Participating projects

SmartCHANGE: Early Risk Prediction for Lifelong Health

Programme: Horizon Europe

Duration: 2023–2027

Focus: AI-driven early risk prediction and preventive health tools for children and adolescents

Contribution to Trustworthy AI in Healthcare

SmartCHANGE advances trustworthy artificial intelligence in preventive healthcare by developing explainable, privacy-preserving risk prediction models targeting cardiovascular and metabolic diseases in children and youth. The project contributes directly to the cluster's shared priorities: FL, counterfactual explainability, participatory design, and regulatory-aware development.

Core Expertise

- Longitudinal AI-based risk modelling across the lifespan
- FL infrastructure for cross-border health data
- Counterfactual-based explainable AI for clinical interaction
- Participatory co-design with families and healthcare professionals
- Deployment validation across multiple European healthcare systems

Key Assets and Outputs

- AI-based lifetime risk prediction engine trained on 15 heterogeneous datasets
- Clinical Decision Support System (CDSS) for healthcare professionals
- Mobile health application ("HappyPlant") integrating wearables and gamified lifestyle feedback
- Visual explainability interface for clinicians
- Multi-country feasibility studies generating real-world evidence

AI-PROGNOSIS: Personalised AI for Parkinson's Disease Detection and Management

Programme: Horizon Europe

Duration: 2023–2027

Focus: AI-based Parkinson's disease risk assessment and prognosis

Following a trustworthy and inclusive approach to AI development and based on multidisciplinary expertise and broad stakeholder engagement, AI-PROGNOSIS aims to advance PD diagnosis and care by:

- Developing novel, predictive AI models for personalised PD risk assessment and prognosis (in terms of time to higher disability transition and response to medication) based on multi-source patient records and databases, including in-depth health, phenotypic and genetic data,
- Implementing a system of biomarkers informing the AI models by tracking key risk and progression markers such as REM behaviour disorder and key motor symptoms, in daily living based on smartphone and wearable data, and
- Translating the models and digital biomarkers into a validated, privacy-aware AI-driven toolkit, supporting healthcare professionals (HCPs) in disease screening, monitoring and treatment optimisation via quantitative, explainable evidence, and empowering individuals with/without PD with tailored insights for informed health management.

TRUSTroke: Enhancing Stroke Management with Ethical AI Platforms

Programme: Horizon Europe

Duration: 2023–2027

Focus: A platform for improved ischaemic stroke patient management

EU contribution: € 6 076 437,50

Project Website: https://trustroke.eu/

Cordis: https://cordis.europa.eu/project/id/101080564

TRUSTroke is creating a trustworthy, federated AI platform for stroke prediction and management across acute and chronic phases. By combining clinical and patient-reported data and applying FAIR principles, TRUSTroke supports recurrence prediction, care personalisation, and cross-hospital learning. Its user-centred design ensures accessibility and clinical alignment.

Project objectives (OB):

- OB1: Develop a FL network and enable robust and secure distributed AI training.
- OB2: To create integrated Trustworthy and validate AI solutions for stroke risk assessment.
- OB3: To create a framework for patient empowerment and communication with integrated prognostic models and stroke outcomes.
- OB4: To develop the TRUSTroke platform, a flexible and scalable platform integrating O1, O2 and O3 and clinically validated in relevant application environments.
- OB5: Improve stroke patients' pathway.

UMBRELLA: Revolutionising stroke care in Europe

Programme: Innovative Health Initiative Joint Undertaking (IHI JU)

Duration: 2024–2029

Focus: Comprehensive, holistic and patient-centred stroke management, enabling faster, more accurate and personalised diagnosis, treatment and outcome prediction.

EU contribution: € 14 791 678,93

Project Website: www.umbrella-ihi.eu

Cordis: https://cordis.europa.eu/project/id/101172825

UMBRELLA aims to revolutionise stroke care by improving and standardising the entire stroke care pathway, with a focus on rapid treatment access, accurate diagnosis, personalised management, and rehabilitation.

The project objectives are:

- OB1: To integrate, harmonise and standardise in a federated manner local retrospective and prospective data collection, as well as stroke protocols and procedures across the whole stroke care pathway.
- OB2: To build and enhance a multi-country federated Real World Data platform (U-platform) that allows the local creation and validation of trustworthy AI models to advance personalised stroke diagnosis, risk prediction, and decision-making.
- OB3: To implement and exploit a Federated Learning framework (FL-platform) that allows the training of the U-platform-AI models across various centres while preserving data security and privacy.
- OB4: To improve stroke patients' outcomes and interoperable silos' communication with novel digital technologies for patients' engagement, real-time monitoring, data visualisation, and decision-making.
- OB5: To develop a roadmap to regulatory compliance and future certification for the selected improved stroke cases' management and the UMBRELLA FL-platform.
- OB6: To develop a roadmap for the long-term sustainability and exploitation of UMBRELLA advancements in stroke care and a plan for dissemination.