Explanation-Driven Visual Interventions in Genomics

Andrea Esposito[1,*], Paolo Buono[1] and Rosa Lanzilotti[1]

[1]Department of Computer Science, University of Bari Aldo Moro, Via E. Orabona 4, 70125 Bari, Italy

Abstract

Artificial Intelligence (AI) is reshaping medical diagnostics, yet its adoption in high-stakes genomic applications remains constrained: the opacity of AI models prevents clinicians from meaningfully verifying, contesting, or learning from automated decisions. This paper presents the GenXAI vision, a system architecture that combines Genomic Language Models (GLMs) with Explainable AI (XAI) and physician-facing visual interfaces to support trustworthy, reproducible microbial species identification from metagenomic data. Applying the paradigm of explanation-driven interventions and Symbiotic AI (SAI), GenXAI goes beyond passive explainability: clinicians can inspect, annotate, and correct XAI attributions. Each correction feeds back into model retraining, creating a continuous mutual improvement loop in which human expertise and AI capability co-evolve. The system is designed from the outset to comply with the EU AI Act and to ensure long-term auditability. We argue that GenXAI instantiates the principles of trustworthiness, reproducibility, and intelligent visual analytics called for by TRI-IVIS in the genomic diagnostics domain.

Keywords

Genomic Language Models, Explainable AI, Symbiotic AI, Explanation-Driven Interventions, Metagenomics, Visual Analytics, Human-Centered AI

1. Introduction

Microbial identification from metagenomic samples is a cornerstone of modern clinical diagnostics, with particular relevance in conditions such as inflammatory bowel disease (IBD), where the composition of the gut microbiome is both a diagnostic marker and a target for personalized therapy [1]. Traditional culture-based methods are slow, incomplete, and unable to capture the full diversity of microbial communities [2, 3]. High-throughput sequencing now makes it possible to interrogate entire microbial ecosystems from patient samples, but the resulting data (vast, high-dimensional, and complex) requires computational interpretation that lies far beyond clinical routine [4]. AI, and specifically Genomic Language Models (GLMs), has emerged as a powerful approach to automated microbial identification. By treating nucleotide sequences as a language and learning contextual embeddings from large genomic corpora, GLMs can encode a genome into a compact representation whose similarity to a reference library can be efficiently computed via cosine distance [5].

Yet the clinical adoption of GLM-based diagnostics faces challenges, including *opacity*. When a model identifies a bacterial species from a patient's metagenomic read, a physician receives a label and possibly a confidence score, but has no account of which genomic regions drove the classification, nor any means of interrogating whether the decision reflects genuine sequence similarity or an artifact of the training distribution. In high-stakes diagnostic contexts, this opacity is a barrier to accountability, to regulatory compliance, and to the calibrated trust that safe clinical AI requires [6]. Explainable AI (XAI) techniques such as SHAP [7] and LIME [8] can expose the feature contributions underlying a model's predictions. However, the prevailing use of XAI in medical AI is passive: explanations are generated and displayed, but clinicians cannot act on them. As Esposito et al. [9] discusses in the context of rhinocytology, this positions physicians as passive reviewers rather than active collaborators: a human-on-the-loop paradigm that falls short of the *symbiotic* relationship needed for safe, evolving AI systems.

Proceedings of TRI-IVIS 2026, the First International Workshop on Trustable, Reproducible, and Intelligent Information Visualization Systems, June 2026, Venice, Italy

*Corresponding author.

andrea.esposito@uniba.it (A. Esposito); paolo.buono@uniba.it (P. Buono); rosa.lanzilotti@uniba.it (R. Lanzilotti)

0000-0002-9536-3087 (A. Esposito); 0000-0002-1421-3686 (P. Buono); 0000-0002-2039-8162 (R. Lanzilotti)

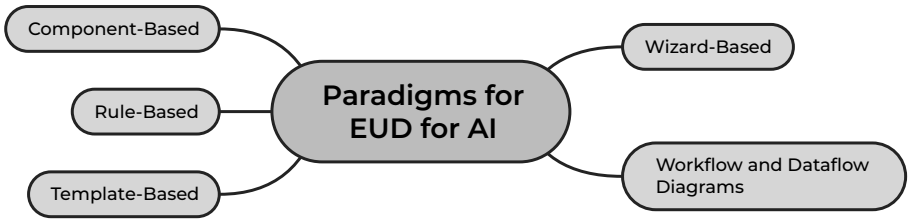

Figure 1: A classification framework for the existing EUD paradigms for AI. Adapted from Esposito et al. [15]

This paper presents *GenXAI* (Genomics eXplainable Artificial Intelligence), a vision for an explainability-driven visual system within a genomic diagnostic platform. Its contribution is to extend explanation-driven interventions [9] to the genomic domain: clinicians can not only inspect XAI attributions rendered as genome-specific visualizations, but can also annotate, override, and correct them. Each intervention is logged and fed back into model retraining, instantiating the definition of Symbiotic AI (SAI) proposed by Desolda et al. [10]: a system in which humans and AI mutually augment their capabilities, with AI benefiting from a continuous stream of user-provided data and humans benefiting from AI's continuously improving performance and knowledge.

The idea described in this paper is hembrional, but consist in the proposal of a system architecture for GenXAI integrating GLM inference, XAI attribution, and an interactive visual interface, favoring a clinical workflow integration model which fosters physician autonomy and accountability. In particular, we position GenXAI as an instantiation of the Symbiotic AI paradigm [11] in genomic diagnostics.

2. Background and Related Work

The analogy between natural language and genomic sequences has motivated a productive line of research in applying transformer-based language models to DNA and RNA. Models such as DNABERT [12] and the Nucleotide Transformer [13] are pre-trained on large genomic corpora using masked token objectives, learning rich contextual representations of nucleotide sequences. A genome (or a metagenomic read) can be tokenized into overlapping k-mers and encoded by the model into a dense embedding vector that captures not only local sequence composition but also long-range functional relationships [12, 13, 2]. Species identification then reduces to a nearest-neighbor search in embedding space: the cosine similarity between the query embedding and a library of reference genome embeddings yields a ranked list of candidate species. This approach has several advantages over classical alignment-based methods: it is computationally efficient at inference time, robust to sequencing noise, and capable of handling partial or fragmented reads common in metagenomic contexts.

In this context, and in general in the medical domain, XAI has been applied with growing frequency, showing to support clinician trust calibration and error detection [6]. Two main approaches may be adopted to provide explanations in genomics: SHAP and LIME. *SHAP (SHapley Additive exPlanations)* [7] provides a theoretically grounded framework for attributing a model's output to its input features by computing Shapley values from cooperative game theory. In the context of sequence models, SHAP can assign an importance score to each token or genomic region, yielding a global picture of which parts of the genome are most discriminative for a given species classification. *LIME (Local Interpretable Model-agnostic Explanations)* [8] takes a complementary, local approach: it perturbs the input sequence and fits a locally linear surrogate model to the resulting predictions, identifying which subsequences most influence a specific decision. Together, SHAP and LIME provide both population-level and instance-level accounts of model reasoning.

However, the dominant paradigm remains one of static, read-only explanations [14]—a limitation that constrains both usability and the potential for human-AI collaboration.

2.1. End-User Development for AI and Explanation-Driven Interventions

End-User Development (EUD) encompasses methods that allow non-technical domain experts to customize, adapt, and refine software artifacts without programming expertise [16], including AI systems [15]. A systematic literature review by Esposito et al. [15] identifies five paradigms for EUD for AI: component-based, rule-based, wizard-based, template-based, and workflow-based customization (Figure 1). A common limitation across these paradigms is that they focus on configuring AI systems before deployment or overriding their outputs post hoc, without enabling experts to reshape the model's internal reasoning.

Explanation-driven interventions [9] address this gap by using XAI outputs as the interface through which users refine model behavior. Through this paradigm, users can inspect the AI's outcomes, edit the explanations to align them with expert knowledge, and mark classifications as correct or incorrect. These *interventions* are logged and used to refine the model, creating a progressive adaptation loop. The result is a human-in-the-loop refinement paradigm that requires no technical expertise—domain knowledge alone is sufficient to steer AI behavior [9]. This paradigm represents a significant advance over human-on-the-loop validation-based collaboration, where experts review AI predictions but have no mechanism to modify the AI's reasoning process [17, 9]. GenXAI aims to extend explanation-driven interventions from the cytological image domain to genomic sequence data, where the intervention interface must be reimagined around sequence-native visualizations rather than image annotations.

2.2. Symbiotic AI

Symbiotic AI [11] is a specialization of Human-Centered AI (HCAI) [18] that moves beyond augmentation (the one-directional enhancement of human capabilities by AI) towards a bidirectional co-evolution [10, 19]. In particular, Desolda et al. [10] define Symbiotic AI (SAI) systems as *"HCAI systems powered by a continuing and deeper collaboration between humans and AI, i.e., a symbiosis of human intelligence and artificial intelligence. Humans and AI mutually augment their capabilities, balancing each other's strengths and weaknesses without hampering either the autonomy of humans or the performance of AI, leaving humans in control of the system's decisions. In an SAI system, AI benefits from a continuous stream of new user-provided data to refine itself, while humans benefit from AI's improved performance and knowledge."* Therefore, in an SAI system, humans and AI mutually augment their capabilities, balancing each other's strengths and weaknesses without hampering either the autonomy of humans or the performance of AI, leaving humans in control of the system's decisions. AI benefits from a continuous stream of new user-provided data to refine itself, while humans benefit from AI's performance and knowledge.

GenXAI is proposed as an SAI system: users' interventions on XAI visualizations are the mechanism through which the AI receives new training signals, and the resulting model improvements are the mechanism through which physicians gain a more reliable and trustworthy diagnostic partner.

2.3. Visualization in Bioinformatics and Clinical AI

Existing bioinformatics visualization tools such as the Integrative Genomics Viewer (IGV) [20], the UCSC Genome Browser [21], and Bandage [22] provide powerful interfaces for exploring genomic data. However, they are designed for genomic research rather than clinical diagnostic workflows, and they have no integration with XAI attribution systems. The resulting visualizations expose raw sequence and alignment data but do not communicate model reasoning or support physician intervention. GenXAI aims to fill this gap by designing visualizations that are explicitly shaped by XAI outputs and that support active physician engagement. Others focus on report interpretation by the end-user [23], but do not consider aspects such as trust and reproducibility.

Trust and reproducibility in clinical AI are increasingly subject to regulatory scrutiny. The EU AI Act classifies medical diagnostic AI as a high-risk application, requiring conformity assessments, transparency mechanisms, human oversight provisions, and technical documentation [24]. Long-term archivability of decisions and their supporting evidence is an additional requirement for post-market surveillance and regulatory audit.

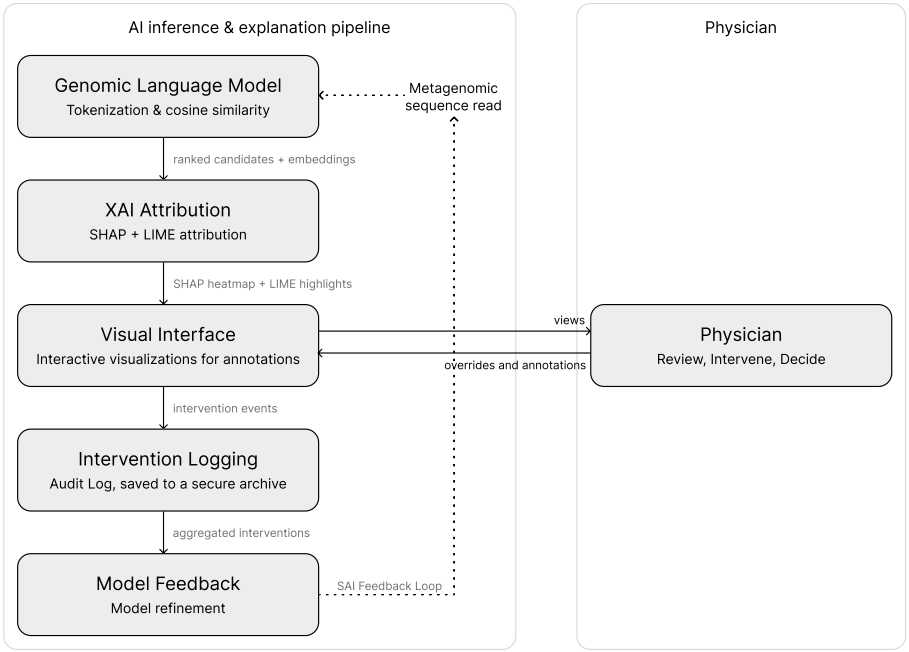

Figure 2: The 4-tier architecture of GenXAI

3. The GenXAI Vision

GenXAI is conceived as an explainability and intervention layer within a broader genomic diagnostic platform, which is developed to provide AI-supported microbiome profiling for patients with inflammatory bowel disease. The system architecture consists of five interconnected components, described in the subsequent paragraphs and depicted in Figure 2.

The *Genomic Language Model* component accepts a metagenomic sequence read as input and tokenizes it into overlapping k-mers, which are then encoded into a high-dimensional embedding vector. The embedding is compared against a versioned reference genome library (e.g., using cosine similarity), producing a ranked list of candidate bacterial species with associated similarity scores. This information is passed on to a *XAI Attribution* component, which applies both SHAP and LIME to the similarity computation: this allows to identify which genomic regions and token positions are most systematically discriminative for the top-ranked species while highlighting which regions of the genome most strongly support or undermine the proposed species identification. A *Visual Interface* renders and presents these attributions. The interface is not read-only: each visualization is interactive, allowing the physician to flag, annotate, correct, or override AI outputs. An *Intervention Logging* component records every physician action—along with the session context, the XAI attribution state at the time of intervention, the patient query sequence, and the final physician decision—in a structured, timestamped audit log persisted to a secure archive. The set of collected interventions is finally provided to a *Model Feedback* component, which uses them as a training signal for iterative model refinement. Override decisions provide implicit re-labeling of query-candidate pairs; annotation of incorrectly weighted genomic regions provides token-level supervision; and pattern analysis of systematic disagreements can identify distributional gaps in the reference library or the embedding model. Techniques from online learning [25] and reinforcement learning from human feedback (RLHF) [26] are candidates for the feedback mechanism. The system can be developed as a cloud-native architecture for clinical metagenomic diagnostics [27] or as a stand-alone application. This paper presents the logical components, regardless of where the system runs.

This architecture instantiates the SAI definition of Desolda et al. [10] in a concrete diagnostic workflow through explanation-driven interventions [9]: the physician can evaluate AI outcomes using clinical expertise that the model cannot encode through interventions (corrections, annotations, and overrides),

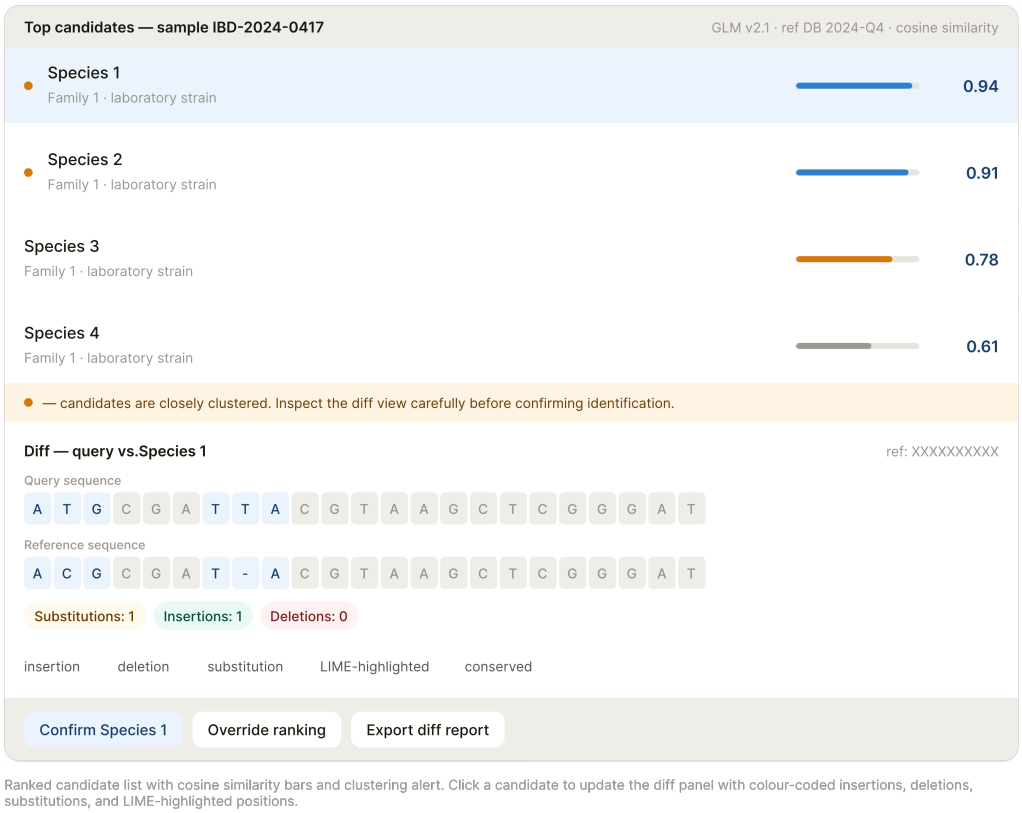

Figure 3: A prototype of the GenXAI interface: ranked similarity list with interactive diff view.

which constitute the continuous stream of user-provided data that allows the model to refine itself for more reliable, better-calibrated identifications subsequent outputs. Crucially, this loop preserves physician autonomy: the system never enforces its suggestions, and every decision is explicitly made by the clinician and progressively adapting the AI to the specific user, patient population, and diagnostic context of the clinical environment in which it is deployed [10].

4. Visualization Design for Genome Comparison

We propose that the GenXAI visual interface presents a set of coordinated views, each designed to expose a different facet of the AI's reasoning and to afford a different mode of physician intervention. This separation allows users to only focus on what matters most to them at any specific moment, thus following an *overview-first, detail-on-demand* principle [28] to minimize cognitive load. The views are linked: selecting a genomic region in any one view highlights the corresponding region in the others.

The primary panel of the interface presents the top-N candidate species ordered by cosine similarity score to the query sequence, displayed as a ranked list with color-coded confidence tiers (Figure 3). Each candidate entry shows the species name, the cosine similarity score, and a provenance indicator showing which reference genome contributed the match. Selecting a candidate opens a comparison view in which the query sequence and the reference genome are rendered as parallel horizontal strips, with insertions, deletions, and substitutions highlighted using a colour scheme borrowed from classical sequence comparisons conventions—additions in green, deletions in red, substitutions in amber. The view is is filtered and annotated by LIME attribution scores, so that the most explanation-relevant divergence regions are visually emphasised.

The intervention affordance in this view is a ranking override: the physician can drag a candidate to

a different position in the ranked list if clinical knowledge contradicts the model's confidence ordering. The re-ranking is logged as an intervention event, with the original similarity scores and the physician's revised ordering preserved in the audit log.

4.1. Interaction Design Principles

All intervention affordances must be available as first-class interface actions (rather than being buried in settings menus) to reduce the friction of user engagement [17]. The interface provides immediate visual feedback for every intervention, making the connection between physician action and AI reasoning transparent. No technical expertise is required: intervention vocabulary is expressed entirely in genomic and clinical terms (regions, species, virulence factors, sequencing artefacts) rather than in machine learning terminology.

The proposed designs aim to avoid the two failure modes identified in the XAI literature: *automation bias*, where clinicians accept AI recommendations without critical evaluation, and *algorithm aversion*, where clinicians dismiss AI recommendations regardless of their quality [29]. By making disagreement easy and rewarded the interface cultivates a posture of calibrated engagement rather than uncritical acceptance or blanket rejection.

5. Clinical Workflow Integration

While in systems like GenDAI [30], the primary target users also involve the patients, the primary target user of GenXAI is the clinical microbiologist or laboratory physician responsible for interpreting metagenomic diagnostic reports. This user has deep expertise in microbial taxonomy, pathogen biology, and the clinical significance of species identifications in the context of specific patient presentations, but should not be assumed to have training in machine learning, embedding geometry, or XAI methodology. The interface is designed accordingly: XAI outputs are rendered as genomic visualisations whose interpretation leverages existing clinical knowledge rather than requiring new technical competencies.

Within the diagnostic workflow, GenXAI occupies the role of a decision support system rather than an autonomous decision-maker. The system proposes a ranked species identification supported by visual evidence; the physician evaluates the proposal against their clinical knowledge of the patient, the sample characteristics, and the epidemiological context; and the physician makes the final diagnostic decision, which may confirm, modify, or reject the AI's recommendation. This positioning preserves full physician accountability for diagnostic outcomes while providing AI-derived evidence that can substantially accelerate and augment the evaluation process.

A particular challenge in embedding-based species identification is cosine similarity clustering: when the top-N candidates have similar similarity scores, the model cannot confidently distinguish among them, and a single-point identification would be misleading. GenXAI handles this case by visually flagging candidate clusters (e.g., color-coding the ranked list entries with a proximity indicator when similarity scores are within a configurable threshold) and by increasing the visual salience of divergent regions in the alignment view for the top two or three candidates. This communicates uncertainty to the physician without suppressing the AI's output and invites more careful inspection precisely in the cases where it is most needed.

5.1. Three-Step Intervention Workflow

The clinical interaction with GenXAI follows a three-step workflow inspired by the intervention paradigm of Esposito et al. [9] and adapted to the genomic modality.

1. Review The physician is presented with the ranked similarity list and the three coordinated visualisations for the patient's metagenomic sample. The interface loads with the top-ranked candidate pre-selected and its diff view and SHAP heatmap pre-rendered. The physician reviews the visual evidence at whatever level of detail their clinical confidence requires: a high-confidence, clearly

motivated top identification may warrant only a brief scan of the heatmap, while an ambiguous or unexpected identification may prompt detailed inspection of the alignment highlights and the diff view.

2. **Intervention** If the physician's clinical assessment diverges from the AI's recommendation, they intervene using one or more of the affordances described in Section 4: reordering the ranked list, annotating tokens in the alignment view, or flagging incorrectly weighted regions in the SHAP heatmap. Interventions do not require the physician to formulate an alternative hypothesis in computational terms: the interface accepts clinical annotations and translates them into model-compatible feedback internally.

3. **Decision and Logging** The physician confirms or overrides the diagnostic conclusion. The confirmed species identification, along with the complete intervention record, the XAI attribution state, the query sequence, and the reference match, is written to the structured audit log. The physician's decision is the authoritative diagnostic output; the AI's original ranking is preserved alongside it to support post-hoc review.

6. Regulatory and Reproducibility Considerations

The EU AI Act classifies medical diagnostic AI as a high-risk application, subject to a conformity assessment process that includes requirements for technical documentation, transparency, human oversight, accuracy, robustness, and cybersecurity [24]. GenXAI is designed to meet these requirements by construction rather than as a post-deployment compliance exercise.

Transparency is addressed through the XAI visualizations described in Section 4, which provide clinicians with a human-interpretable account of every classification decision. Human oversight is addressed through the intervention workflow described in Section 5, which ensures that every diagnostic conclusion is explicitly confirmed by a qualified physician rather than automatically issued by the model. The intervention logging architecture ensures that the system's decision-making process (including both the AI's reasoning and the physician's evaluation) is fully documented and auditable.

Regarding long-term archivability, every diagnostic session generates a structured record comprising: the raw query sequence, the GLM model version and embedding, the reference genome library version, the cosine similarity scores for all candidates, the SHAP and LIME attribution outputs, the complete physician intervention log, and the final diagnostic decision with timestamp and physician identifier. This record is stored in immutable form and is retrievable for regulatory audit years after the original diagnostic event.

Reproducibility of AI decisions over time poses a specific challenge for explanation-based interventions: LIME's perturbation-based approach involves stochastic sampling, which means that two runs of LIME on the same input may yield slightly different attribution scores. GenXAI addresses this by storing the random seed used in each LIME computation alongside its output, enabling exact reproduction of the attribution used in any historical diagnostic session.

Model versioning is handled at both the GLM and the reference genome database levels. Each versioned combination of model and database is assigned a unique identifier that is recorded in every diagnostic session log. When the model is updated through the SAI feedback loop, the new version does not overwrite the previous one: historical records remain linked to the model version that produced them, ensuring that past decisions can be simulated and evaluated against the model's state at the time.

These design choices collectively respond to the characteristics that put emphasis on trustable, reproducible, and intelligent information systems: trustability through XAI visualisations and intervention affordances; reproducibility through versioned models, fixed seeds, and immutable audit logs; and intelligence through a GLM-based inference engine that learns continuously from physician expertise.

7. Discussion and Future Work

GenXAI represents, to our knowledge, the first application of the explanation-driven intervention paradigm to genomic sequence data. By transposing a framework developed for cytological image classification [9] to the one-dimensional, token-structured domain of nucleotide sequences, it raises a set of research questions that are both technically novel and practically significant.

The most immediate open challenge concerns the *computational scalability* of LIME to long genomic sequences. LIME's perturbation-based explanation requires running the model on many perturbed versions of the input; for genomic reads of hundreds or thousands of tokens, this creates a computational bottleneck that may be prohibitive at clinical throughput. Mitigations include hierarchical LIME approaches that first identify high-level genomic regions for perturbation and then refine within them, or surrogate gradient-based local explanation methods that do not require repeated forward passes.

A second challenge concerns the *quality of physician interventions* as training signal. As Esposito et al. [9] discusses, erroneous interventions (arising from fatigue, misremembering, or genuine clinical uncertainty) can degrade model performance if they propagate uncritically into retraining. GenXAI will need mechanisms for detecting and down-weighting inconsistent interventions, for example, by flagging cases where a physician's annotation on one session contradicts the aggregate pattern of annotations from multiple sessions, or by requiring a threshold of agreement across multiple physician reviews before an intervention is incorporated into model fine-tuning.

A third future direction concerns the *extension* of GenXAI from single-species identification to polymicrobial community profiling. IBD and other microbiome-associated conditions involve complex communities of dozens or hundreds of co-occurring species, and a complete diagnostic picture requires not only species identification but also relative abundance estimation and characterization of inter-species functional relationships. Visualizing XAI attributions at the community level—for example, as a two-dimensional embedding space plot in which species cluster by functional role, with physician-selectable similarity overlays—represents a substantial visualization design challenge that remains open.

Finally, especially considering the prototypical nature of this proposal, the system requires prospective evaluation through user studies with clinical microbiologists. Key evaluation dimensions include comprehension of the visualisations (do physicians correctly interpret what the SHAP heatmap and alignment highlights communicate?), trust calibration (does the intervention interface lead to better-calibrated reliance on AI recommendations compared to a static display?), diagnostic accuracy (does physician-AI collaboration via GenXAI improve species identification accuracy compared to AI alone or physician alone?), and cognitive load (does the three-view interface add acceptable overhead to the clinical workflow?). These studies are planned as a next phase of the project.

8. Conclusion

This paper has presented GenXAI, a vision for an explainability-driven visual system for genomic diagnostics that embodies the principles of Symbiotic AI, explanation-driven interventions, and trustworthy, reproducible, intelligent information systems. By combining Genomic Language Models with SHAP and LIME attribution, rendering XAI outputs through three coordinated genome-native visualizations, and enabling physicians to intervene on those outputs through clinically grounded interaction affordances, GenXAI creates a continuous mutual improvement loop in which human expertise and AI capability co-evolve.

The system extends the explanation-driven intervention paradigm [9] to a new modality and grounds it in the formal definition of Symbiotic AI provided by Desolda et al. [10]: a system in which AI benefits from a continuous stream of new user-provided data to refine itself, while humans benefit from AI's improved performance and knowledge, with humans remaining in control at all times. In this project's context, this symbiosis connects the clinical expertise of laboratory physicians with the representational power of genomic language models in a feedback loop designed to simultaneously improve diagnostic

accuracy, clinical trust, and regulatory compliance.

We hope that GenXAI demonstrates the fertility of the genomics domain as a new frontier for the intelligent visualisation systems research community, and invite collaboration on the open challenges of scalable genomic XAI, intervention quality management, and community-level microbiome visualization.

Acknowledgments

This project is funded by the European Union, MSCA SE, Grant Agreement No 101182801.

Funded by the European Union

Declaration on Generative AI

During the preparation of this work, the author(s) used Grammarly in order to: Grammar and spelling check. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] J. Lloyd-Price, A. Mahurkar, G. Rahnavard, J. Crabtree, J. Orvis, A. B. Hall, A. Brady, H. H. Creasy, C. McCracken, M. G. Giglio, D. McDonald, E. A. Franzosa, R. Knight, O. White, C. Huttenhower, Strains, functions and dynamics in the expanded Human Microbiome Project, Nature 550 (2017) 61–66. URL: https://www.nature.com/articles/nature23889. doi:10.1038/nature23889.

[2] T. Krause, J. T. Wassan, P. Mc Kevitt, H. Wang, H. Zheng, M. Hemmje, Analyzing Large Microbiome Datasets Using Machine Learning and Big Data, BioMedInformatics 1 (2021) 138–165. URL: https://www.mdpi.com/2673-7426/1/3/10. doi:10.3390/biomedinformatics1030010.

[3] O. Prakash, M. Parmar, M. Vaijanapurkar, V. Rale, Y. S. Shouche, Recent trend, biases and limitations of cultivation-based diversity studies of microbes, FEMS Microbiology Letters 368 (2021) fnab118. URL: https://academic.oup.com/femsle/article/doi/10.1093/femsle/fnab118/6359716. doi:10.1093/femsle/fnab118.

[4] C. Y. Chiu, S. A. Miller, Clinical metagenomics, Nature Reviews Genetics 20 (2019) 341–355. URL: https://www.nature.com/articles/s41576-019-0113-7. doi:10.1038/s41576-019-0113-7.

[5] M. Leske, J. A. FitzGerald, K. Coughlan, F. Bottacini, H. Afli, B. G. N. Andrade, Alignment-free Bacterial Taxonomy Classification with Genomic Language Models, 2025. URL: http://biorxiv.org/lookup/doi/10.1101/2025.06.27.662019. doi:10.1101/2025.06.27.662019.

[6] C. Combi, B. Amico, R. Bellazzi, A. Holzinger, J. H. Moore, M. Zitnik, J. H. Holmes, A Manifesto on Explainability for Artificial Intelligence in Medicine, Artificial Intelligence in Medicine (2022) 102423. URL: https://www.sciencedirect.com/science/article/pii/S0933365722001750. doi:10.1016/j.artmed.2022.102423.

[7] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17, Curran Associates Inc., Red Hook, NY, USA, 2017, pp. 4768–4777.

[8] M. T. Ribeiro, S. Singh, C. Guestrin, "Why Should I Trust You?": Explaining the Predictions of Any Classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco California USA, 2016, pp. 1135–1144. URL: https://dl.acm.org/doi/10.1145/2939672.2939778. doi:10.1145/2939672.2939778.

[9] A. Esposito, M. Calvano, A. Curci, F. Greco, R. Lanzilotti, A. Piccinno, Explanation-Driven Interventions for Artificial Intelligence Model Customization: Empowering End-Users to Tailor Black-Box AI in Rhinocytology, in: C. Santoro, A. Schmidt, M. Matera, A. Bellucci (Eds.), End-User Development, volume 15713 of *Lecture Notes in Computer Science*, Springer Nature Switzerland, Cham, 2025, pp. 161–170. URL: https://rdcu.be/er7Di. doi:10.1007/978-3-031-95452-8_10.

[10] G. Desolda, A. Esposito, R. Lanzilotti, A. Piccinno, M. F. Costabile, From human-centered to symbiotic artificial intelligence: A focus on medical applications, Multimedia Tools and Applications 84 (2024) 32109–32150. URL: https://rdcu.be/d1RF4. doi:10.1007/s11042-024-20414-5, [Q1 Media Technology].

[11] S. S. Grigsby, Artificial Intelligence for Advanced Human-Machine Symbiosis, in: D. D. Schmorrow, C. M. Fidopiastis (Eds.), Augmented Cognition: Intelligent Technologies, volume 10915 of *Lecture Notes in Computer Science*, Springer International Publishing, Cham, 2018, pp. 255–266. URL: https://link.springer.com/10.1007/978-3-319-91470-1_22. doi:10.1007/978-3-319-91470-1_22.

[12] Y. Ji, Z. Zhou, H. Liu, R. V. Davuluri, DNABERT: Pre-trained Bidirectional Encoder Representations from Transformers model for DNA-language in genome, Bioinformatics 37 (2021) 2112–2120. URL: https://academic.oup.com/bioinformatics/article/37/15/2112/6128680. doi:10.1093/bioinformatics/btab083.

[13] H. Dalla-Torre, L. Gonzalez, J. Mendoza-Revilla, N. Lopez Carranza, A. H. Grzywaczewski, F. Oteri, C. Dallago, E. Trop, B. P. De Almeida, H. Sirelkhatim, G. Richard, M. Skwark, K. Beguir, M. Lopez, T. Pierrot, Nucleotide Transformer: Building and evaluating robust foundation models for human genomics, Nature Methods 22 (2025) 287–297. URL: https://www.nature.com/articles/s41592-024-02523-z. doi:10.1038/s41592-024-02523-z.

[14] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A Survey of Methods for Explaining Black Box Models, ACM Computing Surveys 51 (2019) 1–42. URL: https://dl.acm.org/doi/10.1145/3236009. doi:10.1145/3236009.

[15] A. Esposito, M. Calvano, A. Curci, G. Desolda, R. Lanzilotti, C. Lorusso, A. Piccinno, End-User Development for Artificial Intelligence: A Systematic Literature Review, in: L. D. Spano, A. Schmidt, C. Santoro, S. Stumpf (Eds.), End-User Development, volume 13917 of *Lecture Notes in Computer Science*, Springer Nature Switzerland, Cham, 2023, pp. 19–34. URL: https://rdcu.be/dd32o. doi:10.1007/978-3-031-34433-6_2, presented by Andrea Esposito.

[16] B. R. Barricelli, F. Cassano, D. Fogli, A. Piccinno, End-user development, end-user programming and end-user software engineering: A systematic mapping study, Journal of Systems and Software 149 (2019) 101–137. URL: https://doi.org/10.1016/j.jss.2018.11.041. doi:10.1016/j.jss.2018.11.041.

[17] G. Desolda, G. Dimauro, A. Esposito, R. Lanzilotti, M. Matera, M. Zancanaro, A Human–AI interaction paradigm and its application to rhinocytology, Artificial Intelligence in Medicine 155 (2024) 102933. URL: https://linkinghub.elsevier.com/retrieve/pii/S0933365724001751. doi:10.1016/j.artmed.2024.102933, [Q1 Artificial Intelligence; Q1 Medicine].

[18] B. Shneiderman, Human-Centered AI, 1 ed., Oxford University Press, Oxford, 2022. URL: https://academic.oup.com/book/41126. doi:10.1093/oso/9780192845290.001.0001.

[19] P. Buono, N. Berthouze, M. F. Costabile, A. Grando, A. Holzinger, Special issue on Human-Centered Artificial Intelligence for One Health, Artificial Intelligence in Medicine 156 (2024) 102946. URL: https://linkinghub.elsevier.com/retrieve/pii/S093336572400188X. doi:10.1016/j.artmed.2024.102946.

[20] J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer, Nature Biotechnology 29 (2011) 24–26. URL: https://www.nature.com/articles/nbt.1754. doi:10.1038/nbt.1754.

[21] W. J. Kent, C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, D. Haussler, The Human Genome Browser at UCSC, Genome Research 12 (2002) 996–1006. URL: http://genome.cshlp.org/lookup/doi/10.1101/gr.229102. doi:10.1101/gr.229102.

[22] R. R. Wick, M. B. Schultz, J. Zobel, K. E. Holt, Bandage: Interactive visualization of *de novo* genome assemblies, Bioinformatics 31 (2015) 3350–3352. URL: https://academic.oup.com/bioinformatics/article/31/20/3350/196114. doi:10.1093/bioinformatics/btv383.

[23] P. Buono, P. Tamla, T. Krause, F. De Luzi, F. Monti, M. Mecella, Revisiting data visualizations in diagnostic reports, in: 2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2025, pp. 7200–7206. doi:10.1109/BIBM66473.2025.11356146.

[24] European Parliament, Council of the European Union, Regulation of the European Parliament and

of the Council Laying down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 2024.

[25] S. Amari, A Theory of Adaptive Pattern Classifiers, IEEE Transactions on Electronic Computers EC-16 (1967) 299–307. URL: http://ieeexplore.ieee.org/document/4039068/. doi:10.1109/PGEC.1967.264666.

[26] T. Kaufmann, P. Weng, V. Bengs, E. Hüllermeier, A survey of reinforcement learning from human feedback, Transactions on Machine Learning Research (2025). URL: https://openreview.net/forum?id=f7OkIurx4b.

[27] P. Tamla, T. Krause, M. Hemmje, F. Monti, F. De Luzi, M. Mecella, B. Andrade, P. Buono, A. Molinari, The gendai cloud-native infrastructure and data stewardship for clinical metagenomic diagnostics, in: 2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2025, pp. 7229–7236. doi:10.1109/BIBM66473.2025.11356811.

[28] B. Shneiderman, The eyes have it: A task by data type taxonomy for information visualizations, in: Proceedings 1996 IEEE Symposium on Visual Languages, IEEE Comput. Soc. Press, Boulder, CO, USA, 1996, pp. 336–343. URL: http://ieeexplore.ieee.org/document/545307/. doi:10.1109/VL.1996.545307, shneiderman's Visualization Mantra!

[29] T. A. Bach, A. Khan, H. Hallock, G. Beltrão, S. Sousa, A Systematic Literature Review of User Trust in AI-Enabled Systems: An HCI Perspective, International Journal of Human–Computer Interaction 40 (2024) 1251–1266. URL: https://www.tandfonline.com/doi/full/10.1080/10447318.2022.2138826. doi:10.1080/10447318.2022.2138826.

[30] T. Krause, P. Tamla, A. Leoni, F. Monti, F. De Luzi, J. F. Gerald, B. Andrade, H. Afli, M. Mecella, P. Buono, A. Molinari, M. Hemmje, From reads to reports: A vision for a gfm-powered genomic diagnostic platform, in: 2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2025, pp. 7207–7213. doi:10.1109/BIBM66473.2025.11356528.