AI Health Agents: Longevity, Pathway2vec, ReflectE, and Category Theory

Melanie Swan[1, 2], **Takashi Kido**[3], **Eric Roland**[4], **Renato P. dos Santos**[5]
[1]DIYgenomics
[2]University College London
[3]Teikyo University
[4]RedBud AI, LLC
[5]Lutheran University of Brazil
melanie@DIYgenomics.org, kido.takashi@gmail.com, eric.roland@gmail.com, renatopsantos@ulbra.edu.br

Abstract

Health Agents are introduced as personalized AI health advisors for "healthcare by app" instead of "sickcare by appointment," especially to target Healthy Longevity as a global wellness priority with two billion people estimated to be over 65 in 2050. Health Agents could allow physicians to oversee thousands of patients simultaneously, addressing the 50% of the world's population still not covered by essential health services. As AI Health Interfaces shift to continuous health monitoring (1000x/minute) with medical-grade smartwatches, pins, and wearables, individuals can customize the level of information viewed. As any genAI agent system, Health Agents "speak" natural language to humans and formal language (as Math Agents) to the computational infrastructure, possibly outputting the mathematics of personalized homeostatic health as part of their reinforcement learning agent behavior. Health Agents could deliver precision medicine as a service. Longevity may be achieved 80% with sleep, diet, exercise, and stress reduction, and 20% by medical intervention (metformin, rapamycin, NAD+/sirtuins, alpha-ketoglutarate, taurine), measured quantitatively with aging clocks, biomarkers, and hallmarks. Health Agents are a web3 genAI tool for automated health management, via Personalized Aging Clocks, digital-biological twins, and pathway2vec approaches, for human-AI intelligence amplification towards healthy longevity for global well-being.

The AI Longevity Mindset

The AI Mindset

The AI Stack. The AI infrastructure is evolving rapidly, particularly with genAI (generative AI which creates new data based on what it has learned from a training dataset). Activity can be ordered in the four tiers of human-interface AI assistants, reinforcement learning (RL) agents (self-driving, robotics), knowledge graphs, and artificial neural network architectures (ANNs). AI assistants and RL agents (embodied through prompting) are an intelligence amplification tool for human-AI collaborative access to the vast range of knowledge and computational resources now available.

ANNs. The first neural network architecture to deliver genAI at scale is transformers (GPTs, generative pretrained transformer neural networks), Large Language Models (LLMs) which use attention as the mechanism to process all connections in a dataset simultaneously to perform next word (any token) prediction (OpenAI 2023). LLMs treat a data corpus as a language, with syntax, semantics, and grammar, whether natural language, mathematics, computer code, or proteins. These kinds of Foundation Models are trained on broad internet-scale data for application to a wide range of use cases. Transformers are so-called because they "transform" vector-based data representations during the learning phase (using linear algebra methods).

Transformer architectures are being extended with state-of-the-art LLMs released for multimodal VLMs (vision-language models) (Gemini 2023), larger context windows (e.g. genome-scale training, 1 million base pair size context window (HyenaDNA, Nguyen et al. 2023)), and longer sequential data processing with various convolutional and other methods such as SSMs (structured state space models (Mamba, Gu and Dao, 2023)) and model grafting (hybrid network architectures evolving during training, StripedHyena-7b (7 billion parameters (learned weights between data elements), Poli et al. 2023).

GPTs to GNNs: 2D to 3D+. An advance in digital biology is GNNs (graph neural networks, technically a form of transformer) to process 3D data such as molecules (Bronstein et al. 2021) with attention or message-passing. The early success of GPTs is credited to the "traditional" machine learning recipe (Halevy et al. 2009) of a small set of algorithms operating on a very large dataset, with substantial computational power. GNNs require a more extensive implementation of physics to treat 3D environments. The transformations of data representations in GNNs are more closely tied to the three main symmetry transformations in physics: translation (displacement), rotation, and reflection, and the notions of invariance (output unchanged per transformation) and equivariance (output changes consistently with transformation). For example, AlphaFold2's Invariant Point Attention models the displacement and rotation of amino acids as triangles in space to identify pairwise combinations based on angle and torsional force (Jumper et al. 2021). Also used in GNNs is beyond-Euclidean hyperbolic space to efficiently represent large datasets, for example hierarchical tree-structured data (Zhou et al. 2022).

Knowledge Graph Embedding. Knowledge Graph vector Embedding (KGE) methods also employ a full range of hyperbolic space and symmetry transformations, with founding algorithms TransE (translation embedding) (Bordes et al. 2013), RotatE (rotation embedding) (Sun et al. 2019), and ReflectE (reflection embedding) (Zhang et al. 2022). More capacious number systems expand from the everyday real numbers (1D numbers) to 2D complex numbers with ComplEx (Trouillon et al. 2016) and 4D quaternion numbers with QuatE (Zhang et al. 2019). Quantum formulations are in development, e.g. quantum embedding (Li et al. 2023a) and baqprop (quantum backpropagation of errors) (Verdon et al. 2019). Temporal KGEs are a discovery domain with temporal symmetry, antisymmetry, and inversion deployed via Lorentz transformation (LorenTzE, Li et al. 2023b), tensor factorization (TSimplE, He et al. 2023), and eGNN neural operator temporal dynamics (Xu et al. 2024). Finally, KGE efforts are abstracted to mathematical formalism with geometric algebras (GeomE, Xu et al. 2020), group-theoretic semigroups (SemE, Yang et al. 2022), and Riemannian optimization (OrthogonalE, Zhu and Shimodaira, 2024).

The Formalization Turn. ANNs and KGE methods highlight the implementation of mathematical physics in the AI infrastructure, notably quantum-classical-relativistic models, real-complex-quaternionic (1D-2D-4D) numbers, and beyond-Euclidean space (spherical, hyperbolic) and time (Lorentz invariance, imaginary (complex-valued) time, and time reversal symmetry). A second aspect of the "formalization turn" continues the project of integrating disciplinary fields by finding mathematical structure underlying them (Wigner 1960). For example, combinatoric and geometric structure in particle scattering amplitudes (Arkani-Hamed et al. 2024), a category-theoretic account of double-entry bookkeeping (Katis et al. 2008), and the formal axioms of blockchains (Goncharov and Nechesov 2023).

Category theory is a meta-mathematics for investigating the relationships between different types of mathematical objects (e.g. sets, groups, vector spaces) using categories and functions between them, composing their relations. Category-theoretic methods are seen in deep learning (Gavranovic et al. 2024), information theory (Katsumata et al. 2023), and genomics (Wu 2023), and generally proposed for treating the complexity of biosystems (Rosen 1991).

Math Agents. Math Agents are an "AI math layer" for various mathematical and computational tasks (Swan et al. 2023). Math is the data corpus processed with vector embedding and visualized in equation clusters to view the mathscape (set of equations) of a paper or sector at once. The implication of embedding as a standard genAI method is treating "big data" (entire data corpora) at the level of embedding (a mathematical formulation) to deliver a clean abstract view of very-large datasets. Embedding spaces allow not only data viewing but novel discovery. For example, a Universal Cell Embedding foundation model (representing every cell state and type) was used to identify new developmental lineages (Rosen et al. 2023), and a UMAP (compression) visualization of zero-shot embedding was used to find novel mouse kidney cell types (Kragsteen et al. 2023). GenAI means content generation, and math agent systems may be prompted to write the mathematics of the underlying knowledge graphs "for free" as part of their output. Not only is the content-level prediction obtained (e.g. folded protein structure), but also its mathematical description. AI is a method for interacting with reality at the level of math (a composite view of "big data" and "big math").

Health Agents. Health Agents are envisioned as personalized AI health advisors for "healthcare by app" instead of "sickcare by appointment." Health Agents have two audiences: human and AI in generating the content-level predictions of personalized health and longevity interventions, together with the formal-level of mathematics describing homeostatic health. Health agent systems could operate by wearable apps from underlying blockchain-based healthcare digital-biological twin platforms allowing physicians to oversee thousands of patients simultaneously.

Digital Biology. Digital Biology is the extension of computational biology informatics with genAI methods. Given the complexity of biology, mathematics as a discovery tool has been infeasible. It has not yet been possible to write the robust mathematics of biology as formalizations explaining pathology and homeostasis. However, Health Agents could bring "mathematics as a method" to biological theorizing.

Agile Mindset. The AI mindset suggests a constant first-principles stance in the era of Digital Biology. The molecular biology dogma of the DNA-RNA-protein synthesis chain is being reversed from top-down to also include bottom-up protein structure to DNA (e.g. in AlphaMissense (Cheng et al. 2023)). Drug design is replacing drug discovery in the idea of simply designing molecules with needed properties instead of performing trial-and-error drug searches (Stokes et al. 2020). Treating the pathway not the condition is a new ethos in systems biology (Gschwind et al. 2023).

The Longevity Mindset

Another mindset shift is considering aging as a treatable disease instead of as a natural and inevitable condition of life.

The World Health Organization updated its International Classification of Diseases (ICD-11) in 2022 for a diagnostic category of "ageing associated decline in intrinsic capacity" (Rabheru et al. 2022). Reducing suffering through longevity therapies could be a worldwide priority given the "senior tsunami" of 20% of the world's population estimated to be 60 or older in 2050 (United Nations 2017). Longevity is defined as a healthy lifespan of vitality, energy, and wellness, contra aging as an exponential decline in capabilities leading to age-related diseases and death. The aim is population-scale interventions to slow, reverse, and prevent aging.

Longevity Clocks. One tool for measuring age and interventional impact is aging clocks which compare biological age to chronological age. Such Personalized Aging Clocks include a variety of epigenetic, transcriptomic, glycan, metabolomic, and telomere length clocks at the organism and organ level (Polidori 2024). Blood tests are used measure these factors, for example, the plasma protein signature for eleven organ-specific aging clocks (brain, muscle, artery, heart, lung, immune, liver, kidney, pancreas, adipose, intestine) to find 20% of 5,676 healthy adults already having accelerated aging in at least one organ (Oh et al. 2023). Epigenetic clocks confirmed the rejuvenation of six tissues getting younger as measured by DNA methylation values (in an animal parabiosis model) (Horvath et al. 2024).

Aging clocks may be used in concert with biomarkers of aging (Moqri et al. 2023) and ageotypes (phenotypic age-typing by metabolic, immune, liver, and kidney health) (Ahadi et al. 2020) in targeted longevity interventions. The aim is turning the biological clock back in 10-year periods (e.g. a 70–80-year-old having the muscle health of a 60–70-year-old), and then possibly maintaining people at a desired biological age which may be 20-40 (Bischof et al. 2023).

Longevity Interventions. The experimental evidence for longevity interventions continues to grow (Orr et al. 2024, Blagosklonny 2023, Barzilai et al. 2016, Matysek et al. 2023, Soh et al. 2023, Fahy et al. 2019). Suggested geroprotective medications and supplements include rapamycin, metformin, senolytics, acarbose, spermidine, NR/NAD+ enhancers, NSAIDs, lithium, glucosamine, glycine, and alpha-ketoglutarate (Guarente et al. 2024, Gyanwali et al. 2022, Partridge et al. 2020). The two leading interventions with demonstrable results are rapamycin and metformin, in combination activating AMPK and decreasing mTORC1 signaling which may optimize the allocation of energy resources towards the maintenance of proteostasis (protein homeostasis) (Wolff et al. 2020).

Semaglutide Boom. Adding to aging clocks and aging biomarkers as actionable approaches to longevity is the surprise that 3% of Americans may already be taking an anti-aging drug without knowing it. Named *Science*'s 2023 breakthrough of the year, semaglutide weight loss drugs (GLP-1 agonists such as Wegovy, Ozempic) may also have cardiac benefits and an anti-inflammatory role in the brain-gut axis (Wong et al. 2023). Semaglutide is a medication which mimics the GLP-1 (glucagon-like peptide-1) hormone released in the gut to help the body feel full, producing insulin and reducing blood sugar (glucose). The digital health divide is a pressing concern as on the one hand, Wall Street analysts estimate that worldwide spending on semaglutide, mostly not covered by insurance, could reach $100 billion by 2035 (Adegbesan 2023). On the other hand, the World Health Organization notes that more than half of the global population is still not covered by essential health services (Taylor 2023). Implementing longevity protocols to extend healthy lifespan could become an ethical imperative and a matter of equity, access, and business model.

Longevity – *There's an app for that~!* The longevity revolution could be by app – implemented with Health Agent wearables, sensors, patches, apps, and 3D printers, monitored by longevity physicians, with digital twin partners (virtual patient simulations). Wearables capturing temperature, sleep quality, and heart rate variability provide addressable early-warning signs for various pathologies (Alavi et al. 2022), for example sleep quality predicting type 2 diabetes onset by ten years (Komine et al. 2016). Technology-savvy populations suggest an uberized (widespread accessible via mobile technology) approach to healthcare and longevity therapy delivery, as Deloitte confirms 90% worldwide mobile phone penetration in 2017 (smartphones 80%; 81% in emerging markets) (Wigginton et al. 2017).

Forward-looking countries are targeting longevity as a government policy initiative with goals for healthy citizenry with +5-year healthspans in Singapore (the sixth "Blue Zone" country) and Arab states (Kalin 2023). Health is emerging as a competitive currency and basic human right for human potentiality (Nussbaum 2003). The XPRIZE Longevity Prize was announced in November 2023 for a therapeutic intervention to restore muscle, cognitive, and immune function 10-20 years in 65–80-year-old populations within one year. Longevity venture capital, although down from 2021 peaks, reported 101 deals and USD $1.1 billion for the first three quarters of 2023, sector tracker *Longevity.Technology* reported (Newman and Belleza 2023) and eleven dedicated Longevity venture funds were profiled by *Forbes* (Predin 2023). "AI for social good" has important uses in improving health outcomes (Tomasav et al. 2020).

Generative AI and Biology

Natural language is the first area of demonstrable genAI progress, however, biology may be orders of magnitude more complex including because the "ruleset" is unknown. Whereas one of the largest open-source foundation models, LLaMA, has 65 billion parameters (learnable weights between entities) (MetaAI 2023), state-of-the-art protein models have 100 billion parameters (Chen et al. 2023), and genome language models may require even more. Biological computational complexity classes (protein, genome, pathway) could be formalized from earlier graph visualization starting points (Cirillo et al. 2018, Kugler et al. 2010).

Protein Language Models. Life science AI foundation models include BioMap's xTrimoPGLM (cross-modal interactome and multiomics transformer) with 100 billion parameters as mentioned (Chen et al. 2023) and MetaAI's ESM-2(15B) (evolutionary stochastic model) with 15 billion parameters (Lin et al. 2022). They represent the protein language internally as opposed to beginning with evolutionary MSA (multiple sequence alignment) reads as in AlphaFold2 and RoseTTAFold. Digital biology platforms such as NVDIA's BioNeMo (biological neural modeling) offer pretrained drug discovery models as a cloud service.

Genome Language Models. AI foundation models are being developed in genomics for sequencing and analysis. In sequencing, DeepVariant is a Stanford-led transformer project that holds the Guiness Book of World Records for the fastest human genome sequenced (5 hours 2 min, on 16 Mar 2021, still unbeaten as of February 2024). The DeepConsensus project uses a gap-aware sequence transformer to reduce read errors by 42% as compared with hidden Markov models as the traditional sequence-reading method (Baid et al. 2023). DNAGPT is a transformer performing sequence classification through numerical regression and a comprehensive token language (Zhang et al. 2023).

In genomic analysis, there are various projects focused on building genome-scale language models such as HyenaDNA, training on whole-genome datasets to model individual mutations (Nguyen et al. 2023). The idea is to be able to prompt ChatGPT with an entire human genome to identify mutational profile risks and interventions (e.g. which aging clocks to start with in precision health programs). Other projects include DNABERT for making predictions about transcription factor binding sites, scBERT trained on scRNA-Seq (single cell RNA sequencing) data to predict gene-gene interactions, and Enformer for making predictions about long-range interactions in the genome. AlphaMissense uses protein structures to predict pathogenic missense mutations (of 71 million human mutations, 32% are pathogenic) (Cheng et al. 2023). Missense mutations are ~58% of mutations; then nonsense (10%), frameshift (8%), splice (6%), insertion-deletion (5%), and other (13%).

Digital Twins and Biology

A digital twin is a virtual representation of a physical object, person, or process, estimated by McKinsey to be a $48.2 billion industry in 2026 (Borden 2023). Digital twins are used to model manufacturing operations and infrastructure. Singapore completed the first digital twin of an entire country (Virtual Singapore) in 2022. In healthcare, virtual patient models are used for procedure simulation, medical education, clinical research, and drug development. Clinical trial simulations of millions of genAI-created virtual patients might be routine in the future.

On the one hand, a long-term vision supports the idea of there being biomedical digital twins for the world's 8.1 billion humans. Each person could have an ID number, a MAC address (phone), and a secure web3 digital twin (Akash and Ferdous 2022). On the other hand, healthcare digital twins are not an immediate possibility given the complexity of biological systems, the need for large-scale, high-quality data, and the potential for model inaccuracies.

Longevity Twins. Homeostatic health could be Turing-complete (a format running on any platform, biological or machine; digital twin or biological counterpart). Digital twins started for longevity medicine could extend to future BCI (brain-computer interface) and connectome projects.

Web3

Web3 refers to the current third phase of the internet's development in expanding the interaction mode from the passive *read-only web* (1990s) to the interactive *read-write web* (2000s) to the secure and remunerative *read-write-own web* (2020s enabled by web3 blockchain technologies). Digital transformation continues as many industries become increasingly digitally instantiated. First was the "ready" conversion of content (1990s dot-com news, media, and entertainment), then followed by the more complicated implementation of money and economics, digital art and intellectual property (IP), supply chain, manufacturing, transportation, and science. These latter require more complex features such as non-fungibility and contracts, capabilities provided by blockchains. Blockchains are secure distributed ledger systems providing a database for resource allocation and an immutable record of event histories. Blockchain ecosystems are a foundational information technology using secure properties for digital exchange and economics more broadly as a design principle to produce non-economic outcomes with a broader society benefit (Swan 2015).

Blockchains in Biology

Blockchains are used in health and biology for secure data transfer, supply chain logistics, chain-of-custody tracing, and clinical trials. MediLedger is a global pharmaceutical supply chain blockchain (led by Pfizer, Amgen, and Gilead) completing a pilot program with the U.S. FDA in 2023 towards the 2023 Drug Supply Chain Security Act. Triall is a clinical trials blockchain platform conducting a two-year multi-center pulmonary arterial hypertension clinical trial with the Mayo Clinic. BloodChain is a blood donation network managed with blockchains.

Genome Blockchains. The scale of contemporary science (million-patient studies involving 30 GB whole-human genome files from sixteen sites (Bellenguez et al. 2022)) implies new models for its conduct. For example, the direct-to-consumer whole-human genome sequencing company Nebula Genomics mints a new user-controlled NFT for each sequenced genome (with genomes.io).

DeSci and Longevity DAOs. Web3 methods facilitate DeSci (decentralized science) for carrying out internet-scale bioinformatics. Example projects include LabDAO (an open science collaboration community) and VitaDAO (specific to longevity research). (A DAO is a distributed autonomous organization, an entity formed with blockchain-based smart contracts, with some level of automated administration.)

Blockchain Healthcare Digital Twins

The healthcare delivery system of the future could be one orchestrated by health agents and healthcare digital twins, as physicians oversee the smart health ecosystem by app. Blockchains are suggested as a platform for healthcare digital twins for several reasons. First is the usual notion of blockchains for secure multi-party access to a single unchangeable event history. Second is proof attestation to track the efficacy of interventions (insurance companies are already considering using aging clocks in actuarial tables). Third is an interoperable overlay for integrating multiple omics data streams and EHRs. Fourth is blockchain design principles for modeling non-economic aspects such as homeostasis as its own "bioeconomy." Fifth is the ability to add genAI technologies to the secure health stack with Health Agents as the "user" of the blockchain healthcare digital twin; blockchains both track and facilitate the deployment of AI agents. Health Agents could help the digital twin learn its own longevity protocol.

Blockchain Healthcare Digital Twins for Longevity. The complexity of pathways and processes in the human body can be modeled in the blockchain healthcare digital twin as a multiscalar homeostatic economy of wellness and disease. A blockchain system can instantiate the body, labeling entities as wallet addresses, giving them relevant biocurrency balances, and modeling their activity with smart contract transactions. The schema allows any level of drill-down and roll-up for views of the system per the hashing structure (e.g. organ, tissue, cellular level). One top-level Merkle root can call the entirety of the body. One lowest-level transaction could record the amount of insulin-facilitated glucose release into a cell. The complex pathways of the systems biology of aging (Furber 2019) and their related interventions could be modeled with a blockchain smart contract system. As Virtual Singapore's first digital twin of a country, the first full digital twin of the human can be imagined. The eleven-organ aging clocks could have avatorial representation sitting around "the conference table of the body" as the user interface, genAI stating their agenda based on real-life biomarker levels (analogous to other non-human entities AI-voiced through sensor output).

Longevity State Machine. A smart contract system is implicated to automate blockchain healthcare digital twins. This could be via the Polkadot blockchain ecosystem. Various chains federate for interoperability in the overall structure of "block space as a service" (secure computation and event-recording). Proof-of-stake (randomized participant-based) mining systems provide a greener alternative to costly Bitcoin proof-of-work mining.

Smart contract state machines update wallet balances per transactions. Wallets may contain any cargo such as cryptocurrencies, NFTs (digital assets), computer memory, identity documents (digital passport), or biocurrencies used by Health Agents to manage longevity programs in healthcare digital twins. Independent observers (oracles) take readings (e.g. from wearables, blood tests, patches, toilets, and apps), sending attested measurements to the smart contract system.

Biocurrency Transactions. In a biosystem, there are many different "biocurrencies" circulating to conduct homeostatic activities. The longevity twin has wallet addresses for hundreds of blood biomarkers (e.g. glucose, insulin, homocysteine, HS-CRP, Hb1ac, lymphocytes); everything seen on a blood test. For example, a homocysteine wallet might have an initial balance of 9.0, which is too high, as the desired level of the biomarker may be 5.0. The target protocol is written as a smart contract with logic about the course of action, for example, reducing homocysteine with folate. Each intervention, supplement, or prescription drug could have a wallet address and balance that is managed with smart contracts. Given the homocysteine wallet balance of 9.0, the instruction is for a daily regimen of 500 mg of folate. The folate wallet is activated with a transaction to dispense 500 mg into the smart smoothie, decrementing the overall monthly balance of 50,000 to 49,500. Wallet entities may have their own biocurrencies (e.g. homocysteine, folate), all convertible to BioCoin or some other universal currency.

Biowallets and Longevity Payment Channels. Health Agent smart contracts could manage longevity interventions in off-chain payment channels (contractual interaction sequences) for daily interactions settled to the main health chain on a weekly or monthly basis. Smart contracts would obtain measures of pathway biocurrencies from connected sensors (oracles) and distribute intervention-currencies as a result. The payment channel orchestrates the folate balance acting on the homocysteine pathway to reduce the homocysteine balance. HomocysteineCoin and FolateCoin are readily convertible to meta-tokens LongevityCoin and BioCoin.

Bioconsensus. Health Agents can finetune the different biocurrencies with payment channel smart contracts to learn the optimal personalized longevity protocol. Ideally the system can self-learn its homeostasis. Smart contract coordinated biowallets could be rewarded by Health Agent reinforcement learning mechanisms to reach their own bio-consensus as to the truth state constituting optimal longevity. The biowallet is a useful binary mechanism to target positive and negative behavior, for example reversing disease as "bad player behavior" by assigning negative wallet balances to missense mutations and transposon activations that are targeted by the "good actor behavior" of epigenetic methylation to earn token and pay down negative wallet balances. Health Agent blockchain longevity twins could be tested with Lightning Network payment channels.

Aging Clocks Registered at Birth. As biological processes peak and decline at various stages, some even before birth, a full life-cycle longevity program could be used to register aging clocks as wallets at birth. As life proceeds, balances can be updated to compare biological and chronological age and apply interventions. Blockchain healthcare digital twins are further implicated to manage the potential future suite of wearables, patches, and on-board electronics of brain-computer interfaces, connectome maps, and medical nanorobots. Security is crucial as 5 million people have implanted pacemakers, some internet-connected for real-time remote monitoring, automated pacing therapies, and software updates.

Health Agents

Health Agents are the concept of an AI system tasked with precision human health and longevity. Such a personalized health system app-tool entails dialogue and data capture at the human-AI interface level and agent-based activity at the AI infrastructure level to access information, model health, and orchestrate intervention. Health Agents are qualitative and quantitative, interacting at both the human-consumable content-level of personalized health and longevity intervention, and the AI-usable formal-level of mathematics describing homeostatic health (possibly involving category theory and other formalisms). A digital twin system is implicated as a virtual model of individual health, secured by blockchain smart contracts, user-permissioned for agents to learn by sharing federated data (as in vehicular blockchains of self-driving networks) and for health care plan or societal level aggregation. For precision longevity, the aim is to have a formal statement of health (homeostasis) which can be learned by AI Health Agent systems. The results could be connected to aging clock, biometrics, and tracking data for longevity physicians to review in the Longevity App.

Pathway2vec

Aging clocks research has identified pathways not traditionally targeted by longevity interventions (Oh et al. 2023), hence the first project for Health Agents is implementing Personalized Aging Clocks in a pathway2vec program. Pathway2vec is representing biological pathways as vectors for input to ANNs, in a precedence of approaches beginning with word2vec (Mikolov et al. 2013), and including math2vec (vector representations of equations), gene2vec, SNP2vec, mut2vec, and disease2vec. Related "cancer2vec" approaches use cluster embedding to identify individual cancer risk (Choy et al. 2019) and patient-specific molecular patterns of cancer (Pfeifer et al. 2022), which could likewise model personalized aging-clock targets in longevity. A pathway2vec project defines a vector representation of drug pairs and pathway genes (Yamagiwa et al. 2023).

Longevity Protocol Testing

The next steps in the realization of precision longevity medicine could be testing and tailoring proposed protocols for personalized intervention (e.g. Partridge et al. 2020 p. 516, Houston 2018 pp. 86-87, and Sinclair 2019 p. 304). This entails blood biomarker tests to assess an individual's ageotype profile of accelerated aging using eleven-organ system aging clocks (Oh et al. 2023) and other aging clocks (e.g. epigenetic) (Horvath et al. 2024).

Risks and Limitations

The biggest risk in the AI Health Agents proposal is that although technology development is proceeding quickly, the implementation of healthy longevity programs may be too late to adequately respond to aging demographics. Longevity scholars therefore promote the "escape velocity" idea of deploying immediate interventions to buy enough time until more complete solutions are available (de Grey 2004). Although a contemporary lens sees longevity as a technology, it may not be easy to reprogram biology. There are social challenges to traditional ideas of health, life, and care, along with regulatory hurdles and healthcare system adoption constraints. Further, the use of genAI in biology and medicine has a new set of risks related to data quality, interpretability, and ethics (hallucinations and bias), which requires attention before Health Agents are deployed.

Conclusion

The build-out of the AI infrastructure is proceeding quickly, particularly to facilitate the study of biology. Biocomplexity poses a formidable challenge, but AI methods have encouraging results in the third wave of digitization. The goal is to harness the new tools of digital biology to access a larger scope of investigation with synthesis between areas and actionable steps. This work introduces the concept of Health Agents (personal AI health stewards realized with blockchain digital health twins) as a new idea in human-AI collaboration and intelligence amplification. Health Agents provide an interface for not only qualitative knowledge access, but more importantly "non math speaker" access to the growing intensity of formal methods in the computational infrastructure. The immediate application of Health Agents is healthy longevity, starting to be conceived as a technology, readily implementable and optimizable like any other.

Healthy Longevity and Society. Longevity medicine aims to extend the healthy lifespan of humans by preventing and treating age-related diseases. There are many potential benefits for society. First is improving the quality of life and well-being of older adults and families by reducing the burden of chronic diseases, disability, and dependency (Fried et al. 2022). Second is enhancing the productivity and contribution of older adults to the economy and society, by enabling people to work longer, volunteer, mentor, and participate in civic activities (Accius et al. 2022). Third is reducing the healthcare costs associated with aging, by shifting the focus from treating diseases to promoting health and preventing disease (OECD 2019). Fourth is creating new opportunities for various industries, such as biotechnology, pharmaceuticals, education, and entertainment that cater to the needs of older adults (Ng and Indran 2024). Longevity medicine realized with AI tools such as Health Agents is a mindset and health program that could help individuals and society to realize a longer, healthier, and happier existence.

References

Accius, J.; Ladner, J.; and Alexander, S. 2022. The Global Longevity Economy. AARP.

Adegbesan, A. 2023. Obesity Drug Market Set to Hit $100 Billion by 2035, BMO Says. *Bloomberg News*.

Ahadi, S.; Zhou, W.; Rose, S. M. S.; et al. 2020. Personal aging markers and ageotypes. *Nat Med* 26(1): 83–90.

Akash, S. S.; and Ferdous, M. S. 2022. A Blockchain Based System for Healthcare Digital Twin. *IEEE Access* 10:50523–50547.

Alavi, A.; Bogu, G. K.; Wang, M.; et al. 2022. Real-time alerting. *Nat Med* 28(1): 175–184.

Arkani-Hamed, N.; Cao, Q.; Dong, J.; et al. 2024. NLSM ⊂ Tr(φ 3). *arXiv preprint arXiv:2401.05483*.

Baid, G.; Cook, D. E.; Shafin, K.; et al. 2023. DeepConsensus. *Nat Biotechn* 41(2): 232–238.

Basher, A.; and Hallam, S. J. 2021. Leveraging heterogeneous network embedding for metabolic pathway prediction. *Bioinform* 37(6): 822–829.

Bellenguez, C.; Kucukali, F.; Jansen, I. E.; et al. 2022. New insights into the genetic etiology of Alzheimer's disease. *Nat Genet* 54(4): 412–436.

Blagosklonny, M. V. 2023. Towards disease-oriented dosing of rapamycin for longevity *Aging* 15(14): 6632–6640.

Barzilai, N.; Crandall, J. P.; Kritchevsky, S. B.; et al. 2016. Metformin as a Tool to Target Aging. *Cell Metab* 23(6): 1060–1065.

Bischof, E.; Zhavoronkov, A.; Moskalev, A.; et al. 2023. Longevity Medicine 101. Lecture 9.

Borden, K. 2023. What is digital-twin technology? *McKinsey*.

Bordes, A.; Usunier, N.; Garcia-Duran, A.; et al. 2013. Translating embeddings for modeling MM data. *NeurIPS* 2787–2795.

Bronstein, M. M.; Bruna, J.; Cohen, T.; and Velickovic, P. 2021. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*.

Chen, B.; Cheng, X.; Geng, Y.; et al. 2023. xTrimoPGLM: Unified 100B-Scale Pre-trained Transformer for Deciphering the Language of Protein. *bioRxiv doi: 10.1101/2023.07.05.547496v1*

Cheng, J.; Novati, G.; Pan, J.; et al. 2023. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* 381(6664): eadg7492.

Choy, C. T.; Wong, C. H.; and Chan, S. L. 2019. Infer related genes from large scale gene expression. *Front Genet Sec CG* 9:682.

Cirillo, D.; Ponce-de-Leon, M.; and Valencia, A. 2018. Algorithmic complexity in computational biology: basics, challenges and limitations. *arXiv preprint arXiv:1811.07312*.

Couzin-Frankel, J. 2023. Obesity meets its match. *Science*.

de Grey, A. D. N. J. 2004. Escape Velocity: Why the Prospect of Extreme Human Life Extension Matters. *PLoS Biol* 2(6): e187.

Fahy, G. M.; Brooke, R. T.; Watson, J. P.; et al. 2019. Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell* 18(6): e13028.

Fried, L. P.; Wong, J. E. L.; and Dzau, V. 2022. A global roadmap for healthy longevity. *Nat Aging* 2: 1080–1083.

Furber, J. D. 2019. Systems Biology of Human Aging – Network Model. *Innov Aging* 3(Suppl 1): S973.

Gavranovic, B.; Lessard, P.; Dudzik, A.; et al. 2024. Categorical Deep Learning. *arXiv preprint arXiv:2402.15332*.

Gemini Team, Google. 2023. Gemini: A Family of Highly Capable Multimodal Models. *arXiv preprint arXiv:2312.11805v1.*

Goncharov, S.; and Nechesov, A. 2023. Axiomatization of BC Theory. *Math* 11: 2966.

Gschwind, A. R.; Mualim, K. S.; Karbalayghareh, A.; et al. 2023. An encyclopedia of enhancer-gene regulatory interactions in the human genome. *bioRxiv doi: 10.1101/2023.11.09.563812.*

Gu, A.; and Dao, T. 2023. Mamba: Linear-Time Sequence Modeling with Selective State Spaces. *arXiv preprint arXiv:2312.00752.*

Guarente, L.; Sinclair, D. A.; and Kroemer, G. 2024. Human trials exploring anti-aging medicines. *Cell Metab* 36(2): 354-376.

Gyanwali, B.; Lim, Z. X.; Soh, J.; et al. 2022. Alpha-Ketoglutarate dietary supplementation to improve health in humans. *Trends Endocrinol Metab* 33(2): 136–146.

Halevy, A.; Norvig, P.; and Pereira, F. 2009. The unreasonable effectiveness of data. *IEEE intelligent systems* 24 (2): 8–12.

He, P.; Zhou, G.; Zhang, M.; et al. 2023. Improving temporal knowledge graph embedding using tensor factorization. *Appl Intell* 53:8746–8760.

Houston, M. 2018. The role of noninvasive cardiovascular testing. *Ther Adv Cardiovasc Dis* 12(3): 85–108.

Horvath, S.; Singh, K.; Raj, K.; et al. 2024. Reversal of biological age in multiple rat organs by young porcine plasma fraction. *Geroscience* 46(1): 367–394.

Ioannidis, V. N.; Song, X.; Manchanda, S.; et al. 2020. DRKG: Drug Repurposing Knowledge Graph for Covid-19.

Jumper, J.; Evans, R.; Pritzel, A.; et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–89.

Kalin, S. 2023. Saudi Arabia Research on Aging. *WSJ.*

Katis, P.; Sabadini, N.; and Walters, R. F. C. 2008. On partita doppia. *arXiv preprint arXiv:0803.2429v1.*

Katsumata, S.; Rival, X. and Dubut, J. 2023. A Categorical Framework for Program Semantics. *arXiv preprint arXiv:2309.08822.*

Komine, T.; Takadama, K.; and Nishino, S. 2016. Toward Next-Generation Sleep Monitoring. *AAAI Spring Symposia 2016*: Well-Being Computing: AI Meets Health and Happiness Science.

Kragesteen, B. K.; Giladi, A.; David, E.; et al. 2023. The transcriptional identity of erythropoietin cells. *Nat Med* 29: 1191–1200.

Kugler, H.; Larjo, A.; and Harel, D. 2010 Biocharts: a visual formalism for complex biological systems. *JRSI* 7(48):1015–1024.

Li, L.; Zhang, X.; Jin, Z.; et al. 2023a. KG completion: quantum embedding and quaternion interaction. *Info Sci* 648(119548).

Li, M. M.; Huang, K.; and Zitnik, M. 2022. Graph representation learning in biomedicine and healthcare. *Nat Bio Eng* 6: 1353–69.

Li, N.; Haihong, E.; Shi, L.; et al. 2023b. LorenTzE: Temporal KGE. *ICANN* pp. 472–484.

Matysek, A.; Sun, L.; Kimmantudawage, S. P.; et al. 2023. Targeting impaired nutrient sensing via the sirtuin pathway. *Ageing Res Rev* 90:102029.

MetaAI. 2023. LLaMA: A foundational, 65-bn-parameter LLM. https://ai.meta.com/blog/large-language-model-llama-meta-ai/.

Mikolov, T.; Sutskever, I.; Chen, K.; et al. 2013. Distributed Representations: Words and Phrases. *arXiv preprint arXiv:1310.4546.*

Moqri, M.; Herzog, C.; Poganik, J. R.; et at. 2023. Biomarkers of aging for the identification and evaluation of longevity interventions. *Cell* 186(18): 3758–3775.

Newman, P.; and Belleza, C. 2023. Q3 2023: Longevity Investment Report. *Longevity.Technology*.

Ng, R.; and Indran, N. 2024. Innovations for an Aging Society. *Gerontologist* 64(2):gnad015.

Nguyen, E., Poli, M., Faizi, F.; et al. 2023. HyenaDNA: Long-range genomic sequence. *arXiv preprint arXiv:2306.15794v1.*

Nussbaum, M. C. 2003. Capabilities as Fundamental Entitlements: Sen and Social Justice. *Feminist Economics* 9(2-3): 33–59.

OECD. 2019. The Heavy Burden of Obesity. oe.cd/obesity2019.

Oh, H. S. H., Rutledge, J., Nachun, D.; et al. 2023. Organ aging signatures in the plasma proteome track health and disease. *Nature* 624, 164–172.

OpenAI. 2023. GPT-4 Technical Report. *arXiv preprint arXiv:2303.08774v4.*

Orr, M. E.; Kotkowski, E.; Ramirez, P.; et al. 2024. A randomized placebo-controlled trial of nicotinamide riboside. *Geroscience* 46:665–682.

Partridge, L.; Fuentealba, M.; and Kennedy, B. K. 2020. The quest to slow ageing through drug discovery. *Nat Rev Drug Discov* 19(8): 513–532.

Pfeifer, B.; Saranti, A.; and Holzinger, A. 2022. GNN-SubNet: disease subnetwork explainable GNNs. *Bioinform* 38: ii120-ii126.

Poli, M.; Massaroli, S.; Nguyen, E.; et al. 2023. Hyena Hierarchy: Larger Convolutional LMs. *arXiv preprint arXiv:2302.10866.*

Polidori, M. C. 2024. Aging hallmarks, biomarkers, and clocks for personalized medicine. *Free Radical Biol and Med* 215: 48-55.

Predin, J. M. 2023. Venture Capital Giants in The Billion-Dollar Quest for Longevity Breakthroughs. *Forbes*.

Rabheru, K.; Byles, J. E.; and Kalache, A. 2022. How "old age" was withdrawn from ICD-11. *The Lancet* 3(7): e457–e459.

Rosen, R. 1991. *Life itself: a comprehensive inquiry into the nature, origin, and fabrication of life*. New York: Columbia UP.

Rosen, Y.; Roohani, Y.; Agarwal, A.; et al. 2023. Universal Cell Embeddings: A Foundation Model for Cell Biology.

Sinclair, D. A. 2019. *Why We Age*. New York: Atria.

Soh, J.; Raventhiran, S.; Lee, J. H.; et al. 2023. The effect of glycine administration on the characteristics of physiological systems. *Geroscience* 46, 219–239.

Stokes, J. M.; Yang, K.; Swanson, K. 2020. Deep Learning Approach to Antibiotic Discovery. *Cell* 180(4): 688-702.e13.

Sun, Z.; Deng, Z.; Nie, J.; and Tang, J. 2019. RotatE: KGE Rotation in Complex Space. *arXiv preprint arXiv:1902.10197.*

Swan, M. 2015. *Blockchain: Blueprint for a New Economy*. Sebastopol CA: O'Reilly Media.

Swan, M.; Kido, T.; Roland, E.; and dos Santos, R. P. 2023. Math Agents. *arXiv preprint arXiv:2307.02502*.

Taylor, L. 2023. More than half the world's population lack basic health services as progress stalls. BMJ 382: 2160.

Tomasev, N.; Cornebise, J.; Hutter, F.; et al. 2020. AI for social good. *Nat Commun* 11: 2468.

Trouillon, T.; Welbl, J.; Riedel, S.; et al. 2016. Complex embeddings for simple link prediction. *arXiv preprint arXiv:1606.06357.*

United Nations. 2017. World population projected to reach 9.8 billion in 2050, and 11.2 billion in 2100. The 2017 Revision.

Verdon, G.; McCourt, T.; and X. 2019. Quantum Graph Neural Networks. *arXiv preprint arXiv:1909.12264*.

Wigginton, C.; Curran, M.; and Brodeur, C. 2017. Global mobile consumer trends, 2nd edition. *Deloitte*.

Wigner, E. P. 1960. The unreasonable effectiveness of mathematics in the natural sciences. *Commun Pure Appl Math* 13(1): 1–14.

Wolff, C. A.; Reid, J. J.; Musci, R. V.; et al. 2020. Differential Effects of Rapamycin and Metformin. *J Gerontol* 75(1): 32–39.

Wong, C. K.; McLean, B. A.; Baggio, L. L.; et al. 2023. Central glucagon-like peptide 1. *Cell Metab* S1550–4131(23): 00420–5.

Wu, Y. 2023. A Category of Genes. *arXiv preprint arXiv:2311.08546v1.*

Xu, C.; Nayyeri, M.; Chen, Y.-Y.; and Lehmann, J. 2020. KGE in Geometric Algebras. *arXiv preprint arXiv:2010.00989.*

Xu, M.; Han, J.; Lou, A.; et al. 2024. Equivariant Graph Neural Operator for 3D Dynamics. *arXiv preprint arXiv*:2401.11037v1.

Yamagiwa, H.; Hashimoto, R.; Arakane, K.; et al. 2023. Analogy Tasks in BioConceptVec Biological Pathways. *IEICE* 123(91).

Yang, T.; Wang, Y.; Sha, L.; et al. 2022. Knowledgebra. *ML and Knowl Extr* 4(2): 432–445.

Zhang, D.; Zhang, W.; and Zhao, U. 2023. DNAGPT: A Generalized Pre-trained Tool for DNA Sequence Analysis. *arXiv preprint arXiv:2307.05628.*

Zhang, Q.; Wang, R.; Yang, J.; and Xue, L. 2022. KGE Reflection transformation. *Knowledge-Based Systems* 238 (107861).

Zhang, S.; Tay, Y.; Yao, L.; and Liu, Q. 2019. Quaternion Knowledge Graph Embeddings. *arXiv preprint arXiv:1904.10281.*

Zhou, H.; Wang, W.; Jin, J; et al. 2022. GNN for PPIs. *Molecules* 27(18): 6135.

Zhu, Y.; and Shimodaira, H. (2024). Block-Diagonal Orthogonal Relation for KGE. *arXiv preprint arXiv:2401.05967.*