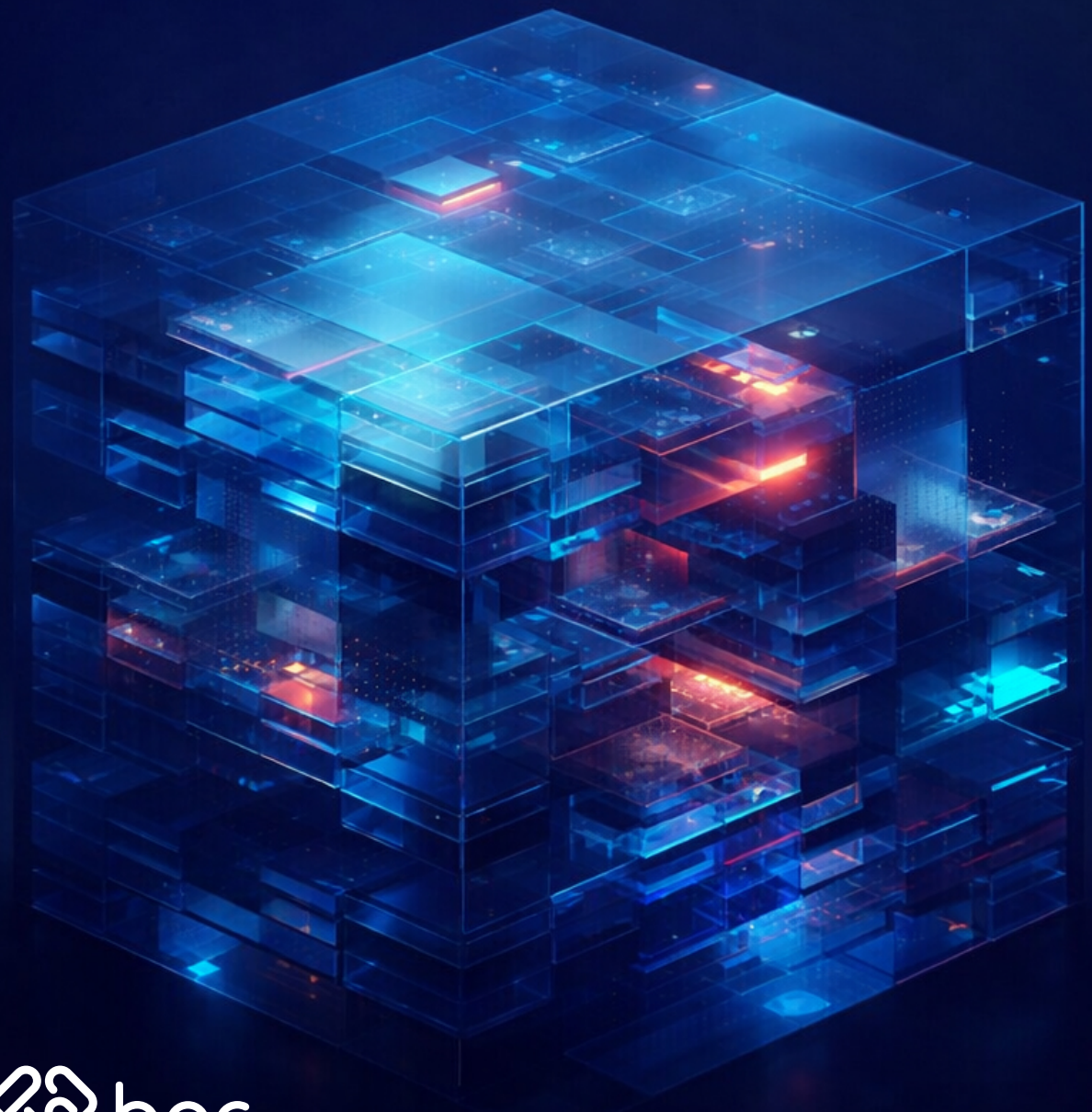
HPC ECHO

From Science to Innovation



hpc

2026

Content

ABOUT KOMONDOR	3
An HPC and What Lies Behind It...	4
Magnetic chains on superconductors, the quest for a fault-tolerant quantum bit	5
Ultrafast Hungarian Large Language Model Trained on Komondor HPC	8
The Secret Life of Small Molecules in Chemical Industry Processes	10
In search of the inhibition of HIV and SARS-CoV-2 virus proteins	14
Peering into the microbial universe within: Large-scale numerical simulations of the gut ecosystem	17
Komondor in the service of dog genetics	20
How supercomputers teach us to "think like systems engineers"?	23
Computer simulation and virtual reality	27
When Artificial Intelligence Restores the Joy of Movement: How a Hungarian Startup is Revolutionizing Knee Surgery	30
TENGR.AI: a secure AI alternative built on Hungarian infrastructure	32

HPC Services at the Peak

Welcome to the pages of our annual publication, in which—now for the fifth time—we summarize the latest achievements of Hungary's HPC (High Performance Computing) infrastructure. The past period has once again demonstrated that supercomputing is no longer merely a tool of a narrow scientific elite, but one of the key driving forces of economic competitiveness and innovation. Whether it is domain-specific research or the optimization of industrial processes, the strategic importance of HPC infrastructure has become undeniable.

It is also evident that artificial intelligence (AI) and HPC operate in symbiosis. HPC, designed from the ground up for peak performance, provides an excellent foundation for AI computations. This new technology enables unprecedented discoveries, while also introducing new requirements. Energy efficiency of hardware and sustainable operation are now just as critical as increasing computational capacity. I am proud to say that we are at the forefront in this regard, and the long-awaited capacity expansion is also just around the corner.

It is important to emphasize, however, that behind the performance of machines lies human expertise and the dedication of the research community. With the support of the National HPC Competence Center—now entering its fifth year of operation—a community has been built that, in record time, has fully utilized Komondor with meaningful research tasks. Komondor, currently Hungary's most powerful supercomputer, now operates continuously at peak capacity. Although the institution operating HPC changed in 2025, professionalism and objectives remain the same. This publication is a tribute to the work of experts and researchers who push the limits of both themselves and the infrastructure every day, producing internationally noteworthy results. I trust that the articles collected here will not only inform but also inspire you in addressing the professional challenges of the coming year.

Zoltán Kiss
Head of HPC Technology and Services Office

ABOUT KOMONDOR

Supercomputing, or High Performance Computing (HPC), is one of the cornerstones of modern research and innovation. HPC systems provide computational capacity that exceeds that of traditional computers by orders of magnitude. This makes it possible to perform complex simulations, high-resolution modeling, analysis of massive datasets, or training of artificial intelligence models in significantly shorter timeframes—hours or even minutes instead of months. Today, HPC is not only a tool for scientific research, but also a defining technological backbone for industrial innovation, drug development, energy, financial modeling, and governmental data processing.

Hungary's largest-capacity and highest-performance-density supercomputer, the Komondor supercomputer, is the flagship of the national HPC infrastructure. The system is capable of more than 6 petaflops of computational performance, features an internal network speed of 200 Gbps, and includes 400 TB of ultra-fast storage, while achieving an outstanding energy efficiency of 28.166 GFlops/watt. The purpose of Komondor is to ensure the fast and reliable execution of large-scale computational tasks across various scientific fields.

Komondor also competes on the international stage: in the June 2025 edition of the TOP500 list—which ranks the world's most powerful general-purpose supercomputers based on the LINPACK (HPL) benchmark—it was ranked 381st. It also entered the Green500 ranking, achieving 76th place, proving that it is competitive not only in computational performance but also in energy efficiency. This is particularly important for sustainable, modern HPC systems, where the balance between performance and energy consumption is a key issue.

Komondor is operated and made accessible by Digital Government Development and Project Management Ltd. (DKF), Hungary's sole HPC service provider and development authority. The system's capacities are available not only to higher education, public education, research centers, and government institutions, but also to market players.

In the academic sector, the system is currently used in several hundred ongoing projects, primarily in natural sciences such as molecular biology, chemistry, pharmaceutical research, and quantum physics. Among corporate users, artificial intelligence model training is becoming increasingly dominant, while governmental projects mainly focus on processing large volumes of data. Thanks to the professional support of the HPC Competence Center, Komondor's utilization reached 100% in the second half of 2025, with a continuous queue of computational tasks.

Komondor is therefore not just a high-performance computer, but a strategic infrastructure for Hungary's research, innovation, and digital development, representing both international competitiveness and a sustainable technological future.

An HPC and What Lies Behind It...

HPC services and the available infrastructure in Hungary celebrate their 25th anniversary this year. The past five years of this long journey have been shaped by the ecosystem-building activities of the National HPC Competence Center (NCC Hungary).

From the very beginning, a key objective has been to provide the most advanced technologies and systems. At the same time, developments and expanding opportunities that are equally important for users have accelerated significantly with the establishment of the National HPC Competence Center.

The range of services available with the support of the National HPC Competence Center is truly extensive, from providing simple HPC access and authentication to developing customized solutions and performing data analysis tailored to specific clients. The service matrix offered by the Center covers the entire spectrum of potential Hungarian clients: from startup enterprises to academic users with centuries-long research traditions. Its network of connections is equally broad, ranging from domestic individual entrepreneurs and international research institutes with hundreds of staff to the highest-level European HPC communities.

With the help of the National HPC Competence Center, anyone can gain access to simpler, faster, and more flexible HPC usage, acquire the necessary theoretical knowledge and practical skills, establish international connections, or even enter the global market.

While using a supercomputer alone can provide a significant competitive advantage to almost anyone, those who encounter this opportunity through the Competence Center gain far more.

In this publication, readers will discover successful research projects and effective solutions that have become truly outstanding and worthy of presentation through the use of services provided by the Competence Center.

Mihály Terjék
Head of the Hungarian HPC Competence Center



Magnetic chains on superconductors, the quest for a fault-tolerant quantum bit

Author: *Bendegúz Nyári*

As a member of the **Computational Magnetism and Superconductivity** research group, created through the collaboration of researchers from BME and the Wigner RCP, I have been studying the electronic structure of magnetic materials, and later of hybrid systems that combine magnets and superconductors, for almost ten years. Currently, our work focuses on topological superconductor systems that could revolutionize information technology. We are living at the dawn of the quantum computing era, and together with many others, we are working to identify physical systems that could function as quantum bits, the fundamental building blocks of future computers.

The results presented here would not have been possible without the work of the other members of our research group. I would like to thank András Lászlóffy, Levente Rózsa, and Balázs Újfalussy from the Wigner RCP, as well as László Szunyogh from BME, for their valuable contributions.

Magnetic atomic chains on superconductors represent one of the most promising model systems for exploring new quantum states. These hybrid nanostructures – built atom by atom on superconducting surfaces – can host exotic quantum states known as Majorana zero modes, which are expected to be key building blocks for future topological quantum computers. The ability to design, model, and understand such systems brings us closer to creating quantum bits that are inherently protected from environmental noise.

When a single magnetic atom is placed on a superconductor, it breaks the surrounding Cooper pairs and creates localized quantum states inside the superconducting gap, known as Yu–Shiba– Rusinov (YSR) states. If several atoms are assembled into a chain, these localized states hybridize into one-dimensional YSR bands. Under suitable conditions of magnetism, symmetry, and spin–orbit coupling, these YSR bands can acquire a topological character and host zero-energy states localized at the chain ends, the Majorana zero modes, as illustrated in Fig. 1.

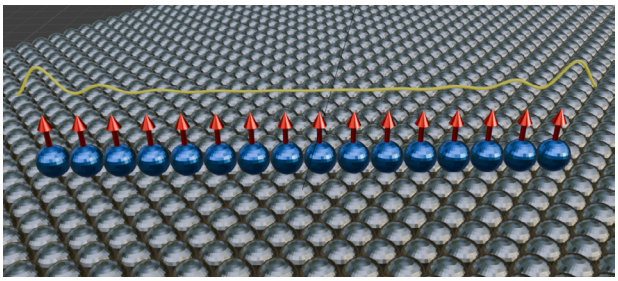


Figure 1: *The illustration of a magnetic chain on a superconductor showing the Majorana zero modes localized to the ends of the chain.*

To explore these effects in detail, we combined atomic-scale experiments and supercomputer-based first-principles calculations. The theoretical framework was developed by researchers at the Budapest University of Technology and Economics (BME) and at the HUN-REN Wigner Research Centre for Physics using the screened Korringa–Kohn–Rostoker (SKKR) method [1] and employing the Dirac–Bogoliubov–de Gennes equation [2], which allows for the unified treatment of superconductivity, magnetism, and relativistic effects like spin–orbit coupling in a large variety of systems, like surfaces, atomic clusters [3] and one-dimensional wires [4, 5]. Carrying out these calculations required the computational power of the Komondor supercomputer, enabling us to model the elec- tronic structure of superconducting nanostructures in a material-specific way.

In our recent work [6], we studied manganese (Mn) atomic chains fabricated on two superconducting materials: niobium [Nb(110)] and tantalum [Ta(110)]. These elements are closely related chemically but differ in their spin–orbit coupling strength, tantalum being heavier and thus exhibiting stronger relativistic effects. Our results reveal that multiple bands with different topological proper- ties can coexist in these systems. The theoretical analysis and experiments both show that these are even and odd with respect to the mirror plane containing the chain. On the niobium substrate, the even-symmetry states display a pronounced minigap opened by spin–orbit coupling but remain topo- logically trivial, meaning that no localized end states appear within this gap. The comparison of the experimental and the theoretical bandstructure is shown in Fig. 2. In contrast, the odd-symmetry states show low-energy features, which with enhanced spin-orbit coupling approach zero energy and be- come partially localized with increasing chain length, predicted to be precursors of Majorana modes.

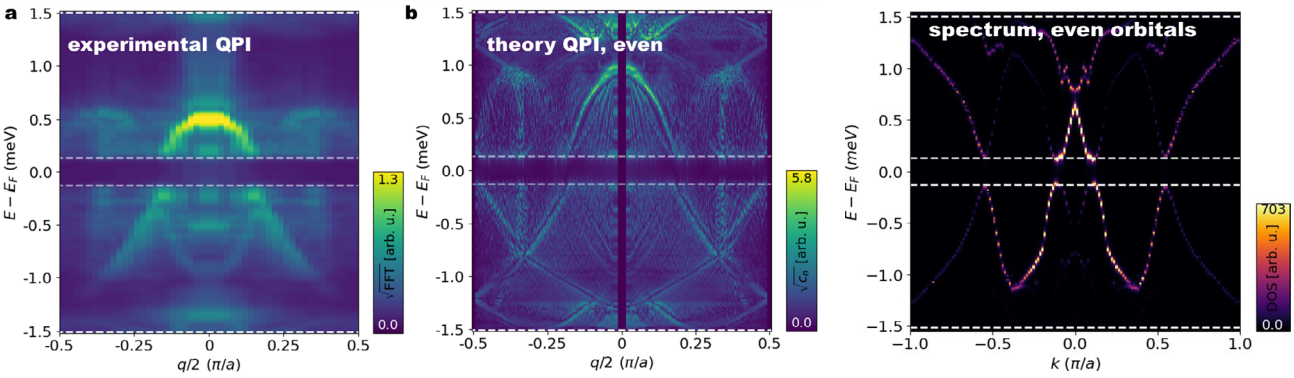


Figure 2: The comparision of experimental and theoretical data for Mn chains on Nb(110), a The quasiparticle interference (QPI) spectrum measured in STM experiments, b the QPI spectrum ob- tained from theory, c the spectral function.

Replacing niobium with tantalum strengthens the spin–orbit coupling and modifies the topological character of the system. While the enhanced coupling increases the hybridization between the even and odd states, it also makes the appearance of Majorana-like end states highly sensitive to the direction of magnetization. Our calculations show that rotating the magnetic moments from the chain direction, the ground state, to an out-of-plane orientation significantly enlarges the minigap and re- sults in clearly localized end states, providing theoretical conditions under which robust Majorana modes could emerge.

These findings demonstrate how crystal symmetry and relativistic effects jointly determine the emergence of topological superconducting phases in magnetic nanostructures. The work also high- lights the importance of high-performance computing in modeling complex quantum materials, where electronic and magnetic properties must be treated self-consistently.

By combining state-of-the-art first-principles theory with atomic manipulation experiments, we have identified the fundamental symmetry mechanisms that govern Yu–Shiba–Rusinov bands in magnetic chains on superconductors. Understanding these mechanisms is essential for guiding the design of materials that could one day host controllable Majorana quasiparticles, bringing us closer to realizing scalable topological quantum devices.

[1] L. Szunyogh, B. Újfalussy, P. Weinberger, and J. Kollár, "Self-consistent localized KKR scheme for surfaces and interfaces," Physical Review B, vol. 49, no. 4, p. 2721, 1994.

[2] G. Csire, A. Deák, B. Nyári, H. Ebert, J. F. Annett, and B. Újfalussy, "Relativistic spin-polarized KKR theory for superconducting heterostructures: Oscillating order parameter in the Au layer of Nb/Au/Fe trilayers," Physical Review B, vol. 97, no. 2, p. 024514, 2018.

[3] B. Nyári, A. Lászlóffy, L. Szunyogh, G. Csire, K. Park, and B. Ujfalussy, "Relativistic first- principles theory of Yu-Shiba-Rusinov states applied to Mn adatoms and Mn dimers on Nb (110)," Physical Review B, vol. 104, no. 23, p. 235426, 2021.

[4] B. Nyári, A. Lászlóffy, G. Csire, L. Szunyogh, and B. Újfalussy, "Topological superconductivity from first principles. I. Shiba band structure and topological edge states of artificial spin chains," Physical Review B, vol. 108, no. 13, p. 134512, 2023.

[5] A. Lászlóffy, B. Nyári, G. Csire, L. Szunyogh, and B. Újfalussy, "Topological superconductivity from first principles. II. Effects from manipulation of spin spirals: Topological fragmentation, braiding, and quasi-Majorana bound states," Physical Review B, vol. 108, no. 13, p. 134513, 2023.

[6] B. Nyári, P. Beck, A. Lászlóffy, L. Schneider, K. Palotás, L. Szunyogh, J. Wiebe, B. Újfalussy, and L. Rózsa, "Symmetry aspects of Yu-Shiba-Rusinov bands in magnetic atomic chains on a superconductor," Physical Review B, vol. 112, no. 11, p. 115414, 2025.

Ultrafast Hungarian Large Language Model Trained on Komondor HPC

The members of our research team (in alphabetic order):
Zsolt Csibi (PhD Student – AI Dept.)
Natabara Gyöngyössy (PhD Student – AI Dept.)
Kornél Nagy (Research Manager – DH-Lab)
Dávid Nemeskey (Research Fellow – DH-Lab)
Gábor Palkó (Head of DH-Lab)
Martin Sallai (AI Engineer – DH Lab)
András Simonyi (NLP Expert – AI Dept.)
András Szekeres (Research Fellow – DH-Lab)

The Racka research team is a collaboration at ELTE Eötvös Loránd University between the Department of Digital Humanities (Faculty of Humanities) and the Department of Artificial Intelligence (Faculty of Informatics). Together, we explore how AI and digital heritage can enrich each other through shared research and teaching.

> **The idea was born at an evening NLP reading club discussion, while we compared LLM scaling laws and the limited size of Hungarian datasets, sparking the thought that Hungary needed its own small, optimally scaled language model to help preserve and strengthen the digital life of our language.**

Our goal is to democratize access to AI and enhance local language capabilities, supporting the growth of Hungary's digital independence. Training such a model brought its own hurdles: adapting to the constraints of HPC infrastructure and fine-tuning the tokenizer to improve efficiency. Still, these challenges became milestones, motivating us as we learned to navigate new territory.

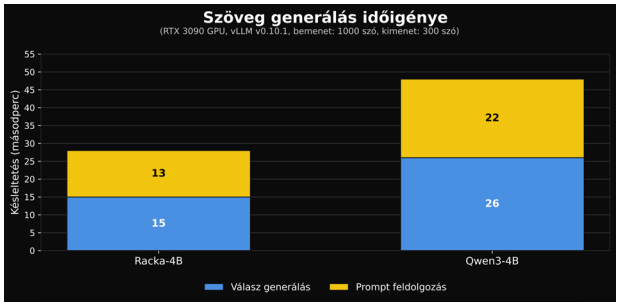
Looking ahead, we hope our work will add to the groundwork for future university-led LLM training, inspire similar initiatives, and encourage more Hungarian language models to join the Racka "flock."

Racka-4B is a lightweight, continually pretrained large language model (LLM) that we created at Eötvös Loránd University (ELTE) to strengthen Hungarian language representation while preserving competitiveness in high-resource languages such as English and German. Based on the Qwen-3 4B backbone, we adapted the model using parameter-efficient continual pretraining with Low-Rank Adaptation (LoRA). Racka-4B can generate text, reason, and operate in chat mode, retaining these abilities from its base model. Following international research practice, we named it after the Racka-4B, a native Hungarian long-haired sheep breed. Despite its modest size, it is one of the smallest Hungarian chat-capable models and already demonstrates considerable knowledge. With further optimization and quantization, it could even run offline on mobile devices, enabling secure and accessible Hungarian AI without expensive hardware.

The foundation of Racka-4B lies in our 160 billion-token multilingual dataset, assembled from both internet sources and curated materials. The dataset consists of 50% Hungarian, 20% English, 20% German, and 10% code. This composition ensures that while Hungarian is prioritized, catastrophic forgetting of high-resource languages is minimized, and reasoning abilities are preserved through exposure to code. In total, we used 130 million Hungarian, English, and German documents, along with code files for adaptation. This large-scale corpus represents a state-of-the-art resource for Hungarian-centric language modeling.

A key practical contribution in our work is the adaptation of the tokenizer. Most multilingual tokenizers are optimized for English thus they split Hungarian words into too many subword units, which inflates sequence lengths and reduces efficiency. To address this, we trained a new tokenizer specifically on Hungarian data, then merged it with the Qwen-3 tokenizer. This hybrid vocabulary introduced 32,000 new Hungarian-specific tokens while preserving the Latin alphabet and existing English/German tokens. Using a vocabulary-transfer method, we initialized embeddings for the new tokens semantically, ensuring better performance from the start. As a result, our model can process longer texts more efficiently and makes fewer grammatical errors in Hungarian.

We trained the Racka-4B on Komondor HPC, Hungary's largest supercomputer, using 64 NVIDIA A100 (40GB) GPUs across 16 nodes. We chose Distributed Data Parallel (DDP) over Fully Sharded Data Parallel (FSDP), as DDP provided superior scaling on Komondor's interconnect. Training lasted 12 days, corresponding to about 2 years and 1 month of cumulative GPU time, during which the model successfully processed the entire dataset once.



Our results highlight the effectiveness of both tokenizer adaptation and resource-aware training. Subword fertility was substantially reduced, enabling faster processing, lower computational costs, and improved representational capacity. Hungarian performance improved compared to the base model, while English and German abilities were preserved. The model also retained reasoning and chat capabilities, showing promise as a versatile Hungarian-centric LLM. However, improvements in Hungarian downstream tasks remain moderate at this stage, underscoring the difficulty of adapting large language models for medium-resource languages.

In future work, we will focus on systematic hyperparameter optimization, comprehensive evaluations across a broader range of Hungarian NLP benchmarks and supervised or preference fine-tuning for downstream applications. Given the upcoming Hungarian AI legislation allows, we also plan open sourcing our model and tokenizer to strengthen linguistic digital sovereignty for Hungarian and provide a replicable methodology for other underrepresented languages.

Acknowledgements

The authors acknowledge the support of the National Laboratory for Digital Heritage. Project No. 2022-2.1.1-NL-2022-00009 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2022-2.1.1-NL funding scheme. This research was supported by the EKOP-24 University Excellence Scholarship Program of the Ministry for Culture and Innovation, funded by the National Research, Development and Innovation Fund.

We acknowledge the Digital Government Development and Project Management Ltd. for awarding us access to the Komondor HPC facility based in Hungary.

The Secret Life of Small Molecules in Chemical Industry Processes

Author: *Milán Szőri*

Overview of Research Results of the Supercomputing Modeling Research Group at the Institute of Chemistry, University of Miskolc

The Northern Hungary region is one of the key centers of the Hungarian chemical industry: it accounts for roughly one-fifth of domestic sales, and it is home to large-scale manufacturers of essential raw materials for polyurethane production. The Supercomputing Modelling Research Group of the Institute of Chemistry at the University of Miskolc, in cooperation with industrial partners in the region, investigates the chemical processes involved in polyurethane production using high-precision quantum chemical calculations.

Polyurethanes are versatile plastics found in many everyday products, such as furniture foams, thermal insulation materials, and synthetic leather coatings. Virtual experiments conducted on the Komondor supercomputer reveal the mechanisms and energetics of reactions at the molecular level, supporting the development of more efficient and sustainable technologies.

Investigation of the reaction mechanism of polyurethane production

Among Hungarian chemical industry clusters, the North Hungarian region accounts for approximately 20% of national chemical industry sales. Two large-scale chemical producers located in the region: BorsodChem-Wanhua's Isocyanate Plant in Kazincbarcika produces 350,000 tons of methylene diphenyl 4,4'-diisocyanate (MDI, Figure1) annually, while MOL's Polyol Plant in Tiszaújváros has an annual capacity of 200,000 tons of polyether polyol. These are the two key components of polyurethane production.

At the HPC Modeling Research Group of the Institute of Chemistry, University of Miskolc, research addressing the needs of the regional chemical industry is a central priority. A particular emphasis is placed on elucidating the reaction mechanisms of industrial urethane production, an effort that has fostered close and ongoing collaboration between the university and industrial partners.

The development of chemical technologies depends on understanding the chemical reactions and energy changes that occur during production, which are often hard to study using experiments alone. Calculations performed on the Komondor supercomputer provide virtual measurements that explain macroscopic behavior at the molecular level, helping to understand processes, identify root causes, and design more efficient and sustainable technologies that support the circular economy.

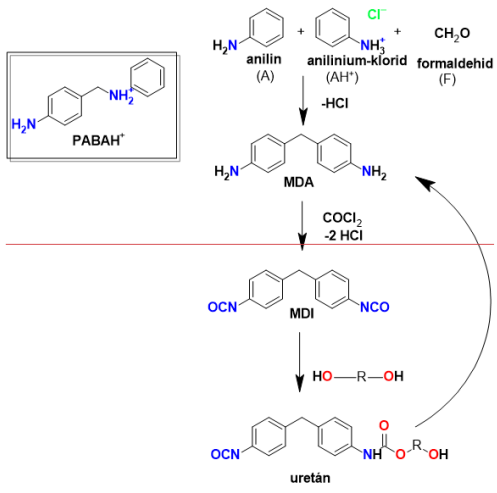


Figure 1. *From Linear to Circular Economy: Key Reaction Steps in Polyurethane Production and Urethane Degradation.*

Our research group used the G3MP2B3 quantum chemistry method to map the eight-step, proton-catalyzed diamine formation reaction mechanism of the aniline-formaldehyde system, which leads to the formation of MDA (Figure 1). The most probable side reactions were identified, and the acid dissociation constants of key intermediates were determined. We demonstrated that the protonated form of p-aminobenzylamine (PABAH+, Figure 1) represents the lowest-energy intermediate; its accumulation reduces product yield, while its structural instability promotes by-product formation. The reaction mechanism was examined in both aqueous and aniline-rich media, and analysis of the corresponding energy profiles revealed that the aqueous pathway is energetically more favorable.

The conversion of MDA to methylenediphenyl diisocyanate (MDI) can proceed (Figure 1) via two alternative reaction pathways : either through phosgeneation of both amine groups via a dicarbamic acid dichloride intermediate ("phosgeneation first"), or through a sequential isocyanate formation mechanism. Our calculations indicate that the former pathway is energetically preferred, with the first phosgene addition exhibiting pseudo-negative activation energy. However, the HCl generated may inhibit subsequent phosgeneation steps through the formation of amine hydrochloride intermediates. We determined the thermochemical properties of all relevant reactants, intermediates, and products in the gas phase, and proposed enthalpy contributions for NCO and NHCOCl substituents on the phenyl ring. These parameters enable reliable estimation of the overall enthalpy balance of industrial MDI production.

To model urethane formation (Figure 1), we investigated the reaction between phenyl isocyanate and 1-propanol. Using G4MP2/SMD quantum chemical calculations, the reaction mechanism was explored under both alcohol- and isocyanate-rich conditions. In the case of isocyanate excess, we proposed a novel two-step mechanism involving an allophanate intermediate. Activation energies were also determined experimentally using HPLC, and showed excellent agreement with the computational results.

The chemical stability of urethane bonds is a major obstacle to the recycling of polyurethanes, so we examined the decomposition pathways of a model compound, methyl N-phenylcarbamate (MPCate), using the G3MP2B3 protocol. Energy profiles were determined for decomposition via hydrogenation, hydrolysis, methanolysis, peroxidation, glycolysis, ammonolysis, methylamine reduction, and reaction with dimethyl phosphite. Among these, decomposition with H_2, H_2O_2, and CH_3NH_2 proved particularly promising—especially in aqueous media—due to their relatively low activation barriers and exothermic character. The resulting products (e.g., aniline

and dimethyl carbonate) are industrially valuable, aligning well with circular economy principles. In addition, ionic decomposition pathways were investigated, revealing alkaline hydrolysis as the most favorable route.

Permeability of Biological Membranes to Small Chemicals

Our research group investigates not only the chemical reactions underlying chemical technologies, but also the interactions between industrially produced molecules and biological systems, allowing potential environmental and health risks to be assessed. The entry of many industrial chemicals into the human body remains largely unexplored; however, some alarming trends are illustrated in Figure 2, particularly given that accumulation of such compounds may lead to serious health effects. This lack of knowledge is often attributed to the difficulty of developing compound-specific analytical methods.

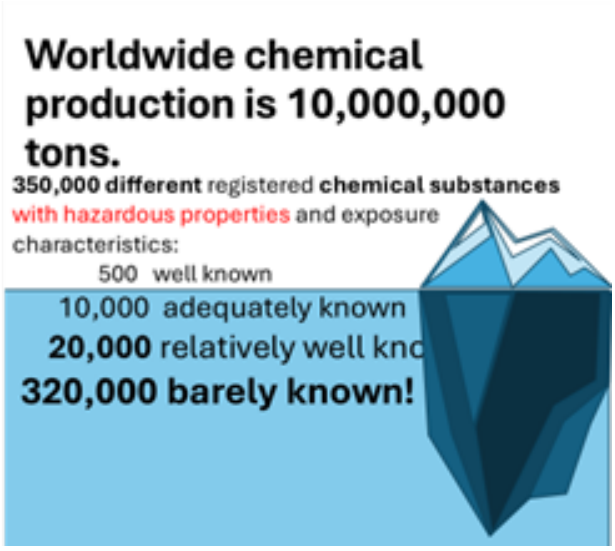


Figure 2. *Global Chemical Production and Awareness of Chemical Hazards.*

Molecular dynamics (MD) simulations provide an effective approach for studying the transport of environmentally released pollutants across biological membranes. We studied four typical small molecules used in the chemical industry: 1,4-dioxane (a widely used solvent), phenol (which occurs in large quantities in the wastewater from isocyanate production), oxane (a promising next-generation fuel component), and N-methyl-2-pyrrolidone (NMP, a solvent used in lithium battery production). As shown in Figure 3 for the case of NMP, passive membrane transport was analyzed in detail, with particular emphasis on membrane structural characteristics and transport energetics.

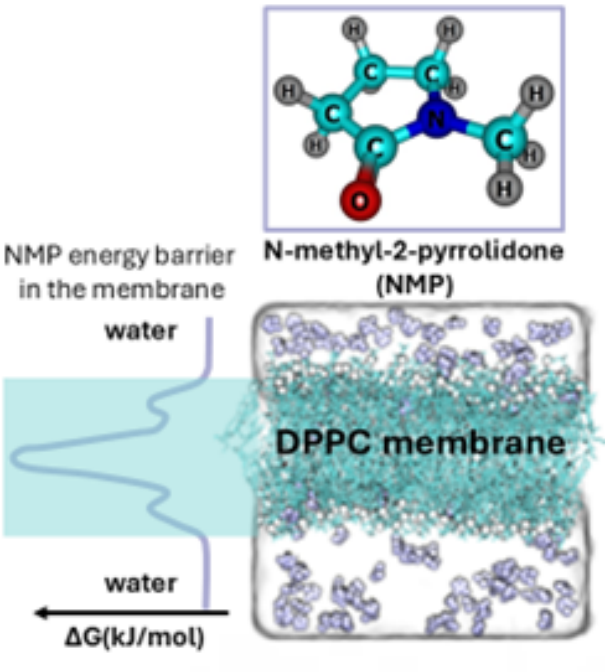


Figure 3. *Structure of N-Methyl-2-pyrrolidone (NMP), Free Energy Profile of Its Dissolution in a DPPC Membrane, and the Corresponding Characteristic Configuration.*

In contrast to NMP, oxane preferentially accumulates in the center of the membrane bilayer from a thermodynamic standpoint, whereas dioxane actively facilitates the ingress of water molecules into the membrane interior. These findings clearly demonstrate that the chemical nature of the lipid headgroups—namely their polarity, size, and charge—in the DPPC, DPPE, DPPA, and DPPG membranes decisively governs both the extent of molecular incorporation and the resulting membrane structural response (Figure 3) fundamentally determines how deeply the compounds studied are incorporated and what structural changes they induce. Although none of the studied compounds encounter a substantial free energy barrier during membrane translocation, phenol and oxane stand out as particularly harmful, inducing pronounced membrane disruption. Such effects strongly suggest a potential for compromised membrane integrity in human cells.

https://cefic.org/landscape-of-the-industry/hungary/?utm (Accessed: September 26, 2025)

https://borsodchem.com/uploads/documentitem/30/bor-sustainability-report-2021-2022-eng-1707925684.pdf (Accessed: September 26, 2025)

https://molgroupchemicals.com/en/news/press-releases/mols-1-3-billion-polyol-complex-in-tiszaujvaros-inaugurated (Accessed: September 26, 2025)

https://fiek.uni-miskolc.hu/ (Accessed: September 26, 2025)

Boros, R. Z. Z.; Farkas, L.; Nehéz, K.; Viskolcz, B.; Szori, M. An Ab Initio Investigation of the 4,4'-Methlylene Diphenyl Diamine (4,4'-MDA) Formation from the Reaction of Aniline with Formaldehyde. Polymers (Basel). **2019**, 11, 398. https://doi.org/10.3390/polym11030398.

Boros, R. Z.; Koós, T.; Wafaa, C.; Nehéz, K.; Farkas, L.; Viskolcz, B.; Szőri, M. A Theoretical Study on the Phosgenation of Methylene Diphenyl Diamine (MDA). Chem. Phys. Lett. **2018**, 706, 568–576. https://doi.org/10.1016/j.cplett.2018.06.024.

Cheikh, W.; Rózsa, Z. B.; López, C. O. C.; Mizsey, P.; Viskolcz, B.; Szori, M.; Fejes, Z. Urethane Formation with an Excess of Isocyanate or Alcohol: Experimental and Ab Initio Study. Polymers (Basel). **2019**, 11, 1543. https://doi.org/10.3390/polym11101543.

Horváth, T.; Kecskés, K.; Jordán Csábrádiné, A.; Szőri-Dorogházi, E.; Viskolcz, B.; Szőri, M. Searching for the Achilles' Heel of Urethane Linkage—An Energetic Perspective. Polymers (Basel). **2024**, 16, 1126. https://doi.org/10.3390/polym16081126.

Rózsa, Z. B.; Horváth, T.; Viskolcz, B.; Szőri, M.; Influence of Membrane Composition on the Passive Membrane Penetration of Industrially Relevant NSO-Heterocycles Int. J. Mol. Sci. **2025**, 26, 7427; https://doi.org/10.3390/ijms26157427.

Rózsa, Z. B.; Horváth, Thangaraj, R.; Viskolcz, B.; Szőri, M. Foreseeing the future of green Technology. Molecular dynamic investigation on passive membrane penetration by the products of the CO2 and 1,3-butadiene reaction J. Mol. Liquids, **2022**, 361, 119581 https://doi.org/10.1016/j.molliq.2022.119581.

Rózsa, Z. B.; Szőri-Dorogházi, E.; Viskolcz, V.; Szőri, M. Transmembrane penetration mechanism of cyclic pollutants inspected by molecular dynamics and metadynamics: the case of morpholine, phenol, 1,4-dioxane and oxane. Phys. Chem. Chem. Phys., **2021**, 23, 15338-15351.

In search of the inhibition of HIV and SARS-CoV-2 virus proteins

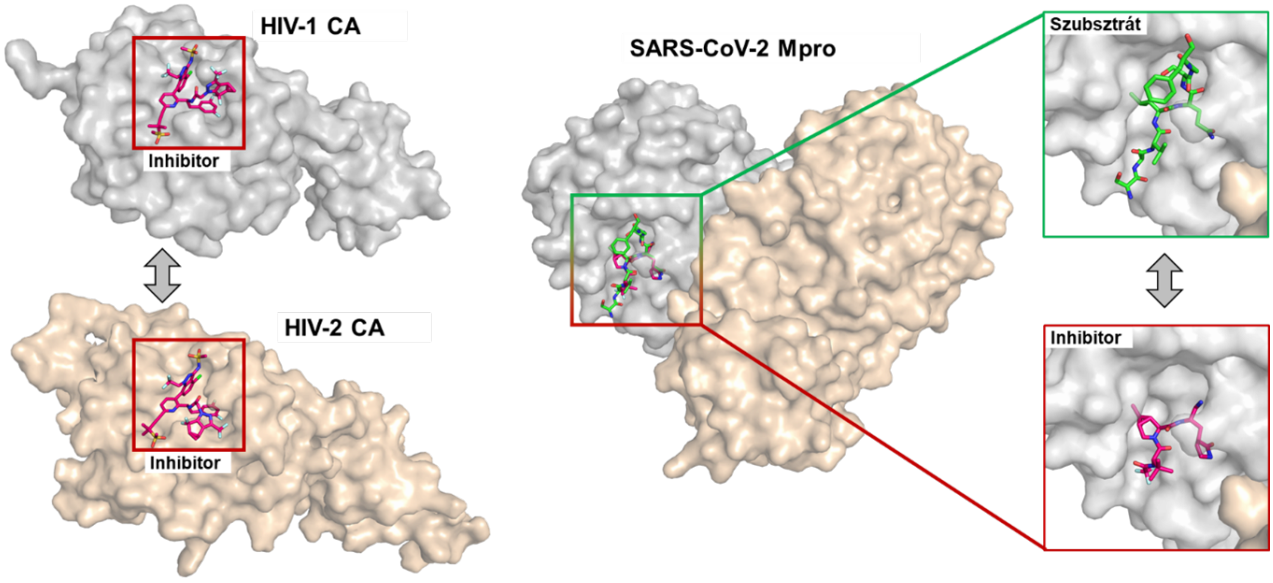
Author: *Dr. János András Mótyán*

The goal of the Laboratory of Retroviral Biochemistry research group is, among others, the biochemical and enzymological investigation of viral proteins by using both experimental and computational methods. We used the resources of Komondor to study several different proteins using computational chemistry approaches. For example, we have studied two viruses of significant health importance, the human immunodeficiency virus (HIV) and the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The proteins of both viruses are considered to be crucial drug targets, therefore the investigation of inhibitor molecules is important to understand their mechanism of action, and is expected to provide valuable information for the development of (more) effective inhibitors, and understand molecular background of drug resistance.



A HIV a legismertebb vírusok egyike, mely a szerzeHIV is one of most well-known viruses. It is the causative agent of acquired immunodeficiency syndrome (AIDS). There are two main types, HIV-1 and HIV-2. The HIV-1 is responsible for the vast majority of infections, there are approximately 40 million people living with HIV worldwide. The number of infections caused by HIV-2 is significantly smaller; estimates from more than a decade ago put the number of people infected with HIV-2 at around 2-3 million. This is caused by the differences of HIV-1 and HIV-2; a lower viral load, a longer latency period and a slower rate of development of AIDS is characteristic for HIV-2. Since HIV-1 is primarily responsible for the vast majority of infections, and thus the global epidemic, this type of the virus has been studied in much more detail, while we have significantly less information about HIV-2. Viral proteins are the main targets of drugs used in the treatment of infected people, and HIV virus proteins are also among the most thoroughly studied molecules. Although many drugs have been developed to treat HIV infection, there is a continuous need to study and develop drug molecules, as viruses can easily become resistant to these drugs. The therapeutic drugs that are applied to treat HIV infections were typically developed against HIV-1. As compared to HIV-1, HIV-2 is understudied, so in many cases the effectiveness of these drugs in the case of HIV-2 is not known or our knowledge is not sufficiently detailed.

In the frame of our project we aimed to determine the effectiveness of drugs - developed to inhibit HIV-1 - against HIV-1 and HIV-2 viral proteins. We used both experimental and computational methods, and we utilized the resources of Komondor to study the interactions between proteins and drugs at the atomic level. The large capacity provided by the supercomputer allowed us to obtain a more detailed picture of the behaviour of the proteins and inhibitors, by using long timescale simulations and softwares that require high computational capacity. Inhibitors targeting two different proteins of HIV viruses, the so-called integrase and capsid were also studied. Our experimental studies revealed that the tested, commercially available inhibitors are able to inhibit both viral proteins, even with comparable efficiency. The findings were also supported by structural studies, according to which the tested drug molecules - developed against HIV-1 - are able to bind to the integrase and capsid proteins of both HIV-1 and HIV-2 in a similar way, which may explain their effectiveness in inhibiting HIV-2 proteins [**Kiarie et al., 2024**].



Investigation of enzyme-inhibitor complexes. *Left panel: interaction of lenacapavir inhibitor with HIV-1 and HIV-2 capsid (CA) proteins. Right panel: structure of SARS-CoV-2 main protease (Mpro) and the binding of a substrate and an inhibitor molecule to the active site of the enzyme.*

The first cases of SARS-CoV-2 infections were identified in 2019, and since then, nearly 800 million cases of infection have been registered worldwide. After the outbreak of the global pandemic in 2020, our research group decided to use its experimental and computational infrastructure to study SARS-CoV-2 proteins. Since then, in 2025, we still study coronavirus proteins. The most important targets of our studies are such proteins of SARS-CoV-2 virus that serve as drug targets, i.e. the spike protein and the main protease, with a special emphasis on the development of resistance against drug molecules [**Mahdi et al., 2025**]. For this purpose, we utilized the experiences of the research laboratory obtained from decades of studying viral proteins, including those of HIV-1. To compare HIV-1 and SARS-CoV-2 viruses, we used computational methods to map the interactions between proteins and drug molecules, which revealed fundamental similarities in the mechanisms behind the development of resistance. Interestingly, an active ingredient of the therapeutic drug that targets the SARS-CoV-2 main protease was developed previously to inhibit HIV-1 protease. As a result of the studies, we identified those parts of the main protease that are responsible for the binding the drug molecule (nirmatrelvir) and whose changes during the evolution of the virus, i.e. in newly emerging virus variants, could potentially contribute to the development of resistance [**Mótyán et al., 2022**]. The currently ongoing studies are based on these observations and aim to study the molecular mechanism of resistance development in a more detailed manner. We intend to compare the accuracy of the data obtained from supercomputer calculations with the results of experimental studies for each variant. The goal is to develop a method that can be used reliably to identify and characterize new, potentially resistant variants and to aid in the design of effective inhibitors.

In addition to studying HIV and SARS-CoV-2 proteins, we have successfully applied the computational capabilities of Komondor to investigate other proteins, as well. For example, we used computational chemical and structural bioinformatics methods to examine such proteins which are of special importance of health and biotechnology perspective [**Hoffka et al., 2023**] [**Golda et al., 2024**], and programs run on the supercomputer were applied to study of a human digestive enzyme [**Hoffka et al., 2024**] and to design an artificially produced enzyme [**Listov et al., 2025**].

Golda M, Hoffka G, Cherry S, Tropea JE, Lountos GT, Waugh DS, Wlodawer A, Tőzsér J, Mótyán JA. P1′ specificity of the S219V/R203G mutant tobacco etch virus protease. Proteins. 2024; 92(9): 1085-1096.

Hoffka G, Lountos GT, Needle D, Wlodawer A, Waugh DS, Tőzsér J, Mótyán JA. Self-inhibited State of Venezuelan Equine Encephalitis Virus (VEEV) nsP2 Cysteine Protease: A Crystallographic and Molecular Dynamics Analysis. J Mol Biol. 2023; 435(6):168012.

Hoffka G, Mhana S, Vas M, Toldi V, Mótyán JA, Tőzsér J, Szabó A. Modeling protease-sensitive human pancreatic lipase mutations in the mouse ortholog. J Biol Chem. 2024; 300(10): 107763.

Kiarie IW, Hoffka G, Laporte M, Leyssen P, Neyts J, Tőzsér J, Mahdi M. Efficacy of Integrase Strand Transfer Inhibitors and the Capsid Inhibitor Lenacapavir against HIV-2, and Exploring the Effect of Raltegravir on the Activity of SARS-CoV-2. Viruses. 2024; 16(10): 1607.

Listov D, Vos E, Hoffka G, Hoch SY, Berg A, Hamer-Rogotner S, Dym O, Kamerlin SCL, Fleishman SJ. Complete computational design of high-efficiency Kemp elimination enzymes. Nature. 2025; 643(8074): 1421-1427.

Mahdi M, Kiarie IW, Mótyán JA, Hoffka G, Al-Muffti AS, Tóth A, Tőzsér J. Receptor Binding for the Entry Mechanisms of SARS-CoV-2: Insights from the Original Strain and Emerging Variants. Viruses. 2025; 17(5):691.

Mótyán JA, Mahdi M, Hoffka G, Tőzsér J. Potential Resistance of SARS-CoV-2 Main Protease (Mpro) against Protease Inhibitors: Lessons Learned from HIV-1 Protease. Int J Mol Sci. 2022; 23(7): 3507.

Peering into the microbial universe within: Large-scale numerical simulations of the gut ecosystem

Authors:
Máté Jakab, Gergely Boza, István Scheuring

The community of microorganisms living in the human intestinal tract, known as the gut microbiome, is an extemely complex ecosystem that fundamentally influences the host's health and metabolism. Microbes play a key role in breaking down dietary components that the human body cannot process on its own, while also producing vitamins and other essential molecules. Our research group has developed a new computational model to study the functioning and development of microbial communities, taking into account metabolic interactions between species and their different degradation strategies. Simulations show that a more diverse nutrient composition promotes the formation of richer and more stable microbial communities. The model may contribute to a better understanding of microbiome function and, in the future, support the development of personalized medicine and microbiome-based therapies.

The community of microbes living in the intestinal tract of humans—and vertebrates in general—the gut microbiome, is an extremely complex ecosystem comprising billions of microorganisms and hundreds of species. These tiny living organisms fundamentally influence the host's health, behavior, and overall condition. Within the spatially structured and highly diverse environment of the gut, countless microbes interact with one another—often including species about which we still know very little.

During digestion, complex food components and polymers that the host organism cannot break down are processed by the microbial community—especially in the large intestine—through multi-step degradation pathways. The smaller molecules produced in this process (such as oligo-, di-, and monomers) are partly utilized by the microbes themselves, but the host organism also benefits from these compounds. In addition, many microbial species produce vitamins and other compounds essential to the host's health.

Understanding the structure, stability, and operating mechanisms of this hidden ecosystem represents a major scientific challenge. Although an increasing number of studies are uncovering the relationship between the microbiome and health, a deeper understanding of the system's dynamics is still in its early stages. Our research group is therefore developing new modeling approaches to capture and describe the complex dynamics of these processes.

Birth of a new model

After conducting an extensive review of the scientific literature and identifying key gaps in existing models, we began designing a new model that captures important but often overlooked aspects of biological reality. Our goal was to represent the multi-directional complexity of the gut microbiome while keeping the model simple enough to analyze and draw meaningful insights from.

With this kind of models, we can better understand how dietary habits influence the composition of the microbiome and explore how it recovers after antibiotic treatment through computer simulations. One of our biggest challenges was figuring out how to represent metabolic degradation cascades or networks in a way that reflects real biological processes as accurately and generally as possible. After many discussions—both within our team and with international experts in microbiome research—we arrived at a solution that we found both practical and scientifically satisfying. The development of this new model system itself was primarily inspired by insights that emerged from years of interdisciplinary discussions within the H2020 FindingPheno Project, involving theoretical and experimental biologists, data scientists, and physicists.

Model description

Our model simulates how the gut microbiome develops over time by introducing microbial species one after another, mimicking a process of sequential biological invasion. Each species interacts with others by consuming shared resources within the gut and releasing metabolic byproducts into the environment. These byproducts can, in turn, be used by other microbes, forming intricate networks of interactions mediated by metabolites.

A key innovation of our approach is the introduction of a hierarchical system of metabolites, where each resource is assigned an energy value that microbes can extract. Another new element of the model is the distinctions between specialists and generalist, the latter having to pay higher maintenance cost of the bigger set of genes. In our simulations, species perform only catabolic reactions—they break down complex resources into simpler ones, releasing energy for their own metabolism. Whenever a resource is consumed, one or more new compounds are produced, each containing less energy. Species that can harvest enough energy to meet their metabolic needs survive and grow, while those that cannot eventually disappear.

Another distinctive feature of our model is its inclusion of multiple, partly separated local sub-communities. These represent food particles within the gut that microbes can colonize and digest. Each sub-community begins with a mix of primary resources that are continuously available, while all other compounds appear only as byproducts of microbial metabolism. Microbes can migrate between these local communities, but resources themselves remain confined to each specific microenvironment.

Using this framework, we explore how differences in the composition of various food particles—essentially, the diversity of available nutrients—affect the overall diversity of microbial species. To quantify this, we calculate the compositional diversity of the resource profiles of local communities. The same metric, along with the traditional Shannon diversity index, allows us to assess species diversity on multiple levels. The compositional diversity is sensitive for the differences between the local communities, while Shannon diversity measures the diversity of the whole system.

Our numerical results show a clear trend: as food diversity increases, so does microbial diversity, while both specialist and generalist species coexist stably within the ecosystem. This simple yet fundamental outcome supports the validity of our modeling framework and helps guide future refinements that better reflect real-world microbial behavior.

The next step in this work is to test our model's predictions experimentally, in collaboration with colleagues who study gut microbiomes in laboratory systems. These collaborations will allow us to compare simulation outcomes with actual biological data and improve the model's accuracy.

Ultimately, modeling efforts like this help transform the immense complexity of microbial ecosystems into quantitative understanding. They reveal how microbial interactions shape the stability of communities, influence metabolism, and determine how the microbiome responds to changes in diet, medication, or disease. Our current model, which uses computer-generated virtual species, lays the groundwork for exploring these fundamental principles—and as we integrate real-world data, it will become an even more powerful tool for studying the dynamics of the gut microbiome.

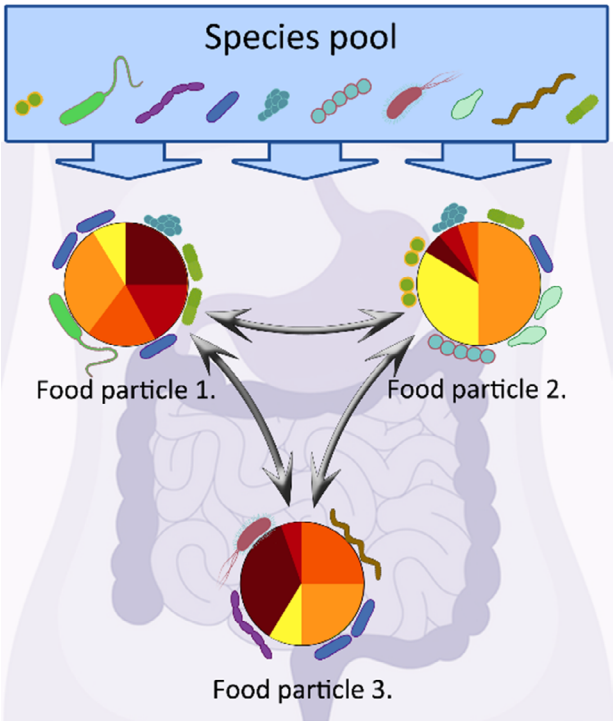


Figure 1. *An overview of the model system. Local communities, i.e. food particles, are connected by the migration of species (grey arrows). Each local community has a different resource composition indicated by the different pie charts. The invading species are picked randomly from a common pool of species.*

The future of medicine and the microbiome

The future of medicine is increasingly turning toward the microbiome as a key to personalized health and precision medicine. By understanding and harnessing these microbial ecosystems, doctors may soon be able to prevent or treat diseases by restoring microbial balance rather than simply targeting symptoms. Emerging therapies could include tailored probiotics, microbiome transplants, and dietary interventions designed to reshape gut communities for optimal health.

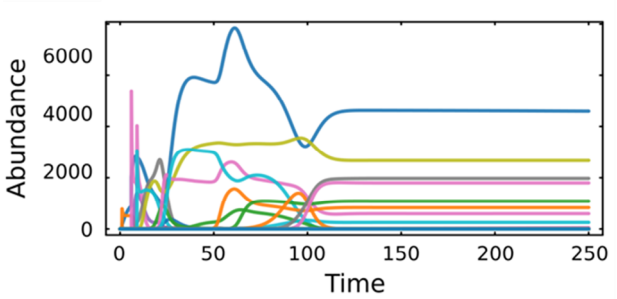


Figure 2: *Species abundances over time in a local community. Invasion events stop at Time=100, followed by a relaxation period. Each line represents a different species.*

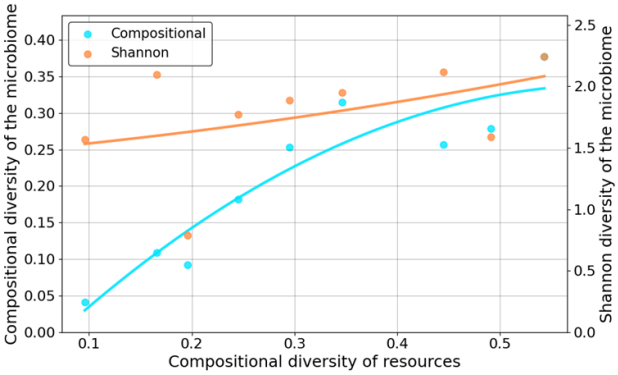


Figure 3. *Different species diversity measures as influenced by the compositional diversity of resource supply. Simulations were run until relaxation after adding 100 invaders. Compositional diversity and Shannon diversity measures were calculated from the resulting equilibrium abundances. Points shown are averages of ten different replicate simulations with the same parameters.*

Komondor in the service of dog genetics

Author:
Dávid Jónás

Members:
Dávid Jónás; Zsófia Rékasi; Hanga Szobonya; Dóra Tfirst; Tamara Bucska; Bence Morvay; Virág Tarcsa

Dávid Jónás is a bioinformatician, who is working at the Department of Ethology of ELTE and has more than 15 years of experience in genetic data analysis. He previously worked on artificial selection projects, pedigree and whole-genome sequence data analysis, on RNA- and ChIP sequencing experiments on genome-wide association studies and on copy number variation detection as well.

Currently he is working as assistant professor at ELTE and he investigates the genetic background of aging with his research team, including thesis-, TDK- and PhD students. Every student's topic involves a research as well, all of which is implemented using the Komondor high-performance supercomputer.

Our research team focuses on the investigation of dog (Canis familiaris) genetics from the widest perspective: from dog domestication through the genetic background of several traits and disorders to the genetics of cognitive aging. Within these topics, our primary aim is to better understand biological aging in dogs as well as to put dogs into the wider perspective of aging research.

The companion dog as a model

Many aspects of aging can be studied on classical laboratory model animals, but the complex phenomenon of aging as it occurrs in humans cannot. Aging is simultaneously affected by genetic constituents – genes, regulatory elements, etc. –, many environmental factors and their multi-level interactions, which cannot be simultaneously modelled in a laboratory, where individual elements of aging are investigated by themselves.

An appropriate translational animal model is therefore required before we can apply results and knowledge from the classical experimental laboratory animal models to humans (e.g. in biomedical research), and the companion dog is an ideal choice for many reasons (Kaeberlein et al., 2016).

Komondors required

In total, we have generated 1,5 Tb of compressed, raw genetic data from 745 samples and downloaded an additional 4 Tb of data from thousands of samples from public sequence repositories. The storage and analysis of this dataset would not be possible without supercomputers and therefore we used Komondor, the Hungarian high-performance computer for our research.

Numerous experiments – one aim

In the past years we have implemented a large number of genetic/bioinformatic studies on Komondor. A wide range of methods were applied, including genome-wide association studies, comparative genetic studies based on both whole-genome and RNA sequence data, genetic imputation from low/medium density SNP arrays to whole-genome data, to mention a few.

Using these tools and resources, we studied the DNA of extremely old dogs (defined as dogs who lived at least 1,5 times their respective breed's average age), and identified tens of thousands of novel single nucleotide polymorphysms (SNP) and short indel mutations as well as copy number variations, all of which are potentially related to the donors' long and healthy lifespan. In a preliminary study we published ~80,000 novel SNP mutations in dogs of extreme age, 670 of which were missense variants at 472 genes (Jónás et al., 2020); a manuscript with a larger dataset is in preparation.

At RNA level we searched for genetic markers of healthy aging from both brain and blood samples: although we identified 3,436 genes to be differentially expressed in the prefrontal cortical region of the brain (Sándor et al., 2024), we could not identify such biomarkers of aging in a more accessible tissue, the blood (Jónás et al., 2025). We also described a shift in alternative splicing events as a function of chronological age in 204 genes from the cortex samples (Jónás et al., In prep.).

Our results from both DNA and RNA experiments indicate that the regulation of gene expression play an important role in healthy aging in dogs. Multiple genes with missense mutations were related to gene expression regulation, while at the RNA level we observed that the overwhelming majority of genes with significantly different gene expression levels between young adult and old adult dogs changed in expression magnitude and were not turned on/off. Several gene ontology terms of the 3,436 significant genes were also related to genetic regulation mechanisms. Currently we are comparing RNA sequence data from the prefrontal cortical region of the brain of classical laboratory rodents, dogs, foxes and humans with the hypothesis that the general gene expression profile observed in dogs will be more similar to those of humans at least partly due to the similar environmental effects. We are also researching the long-standing hypothesis that retrotransposon activity increases with chronological age, with the help of high-throughput DNA sequence data.

By comparing a genome-wide gene expression profiles of wolves and dogs, we investigated the evolution of the dog species from its most recent common ancestor with wolves and identified 1,567 genes with statistically significantly different gene expression levels in blood between the studied species (Figure 1). We also explored the genetic background of several traits: the rare word-learning[1] ability of dogs (Figure 2), the freezing behaviour of laboratory animals – a known defensive reaction of animals in face of danger –, and olfaction performance of dogs as well, which is

[1] *Words in this context refer to object names and not to action-related command words.*

one of the most important perception ability of dogs, used in many situations (e.g. in hunting or communication). Using whole-genome sequence data, we recently finished a project plan to identify the causative mutation underlying the freezing behaviour and whole-genome sequencing of 50 dogs is currently ongoing for this project.

Finally, the Komondor HPC is used in higher education by our team at Eötvös Loránd University and it was also used for 10 BSc and MSc theses (5 finished, 5 ongoing) and for 3 TDK projects (1 finished: I. place at ELTE's Faculty of Sciences TDK competition's Bioinformatics section; 2 ongoing).

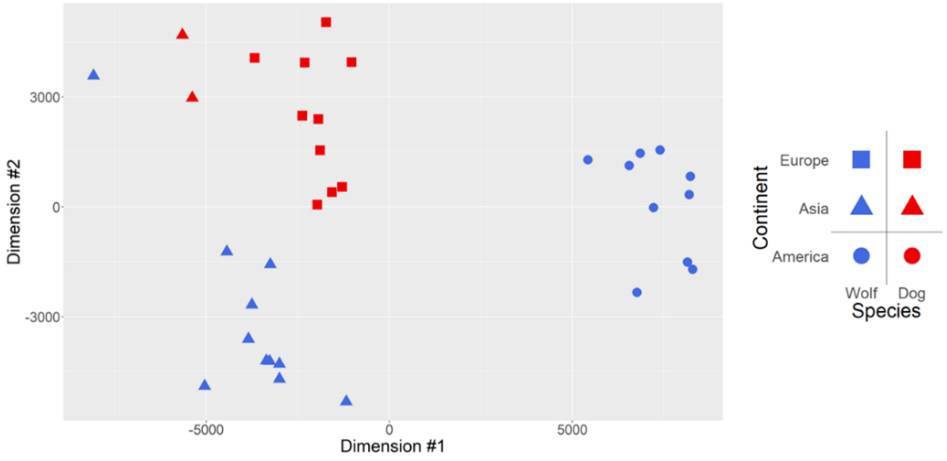


Figure 1
Multidimensional scaling of 33 canids from three continents: Europe (n=10), Asia (n=13) and the Americas (n=10). Figure title and figure from Jónás et al. (2024).

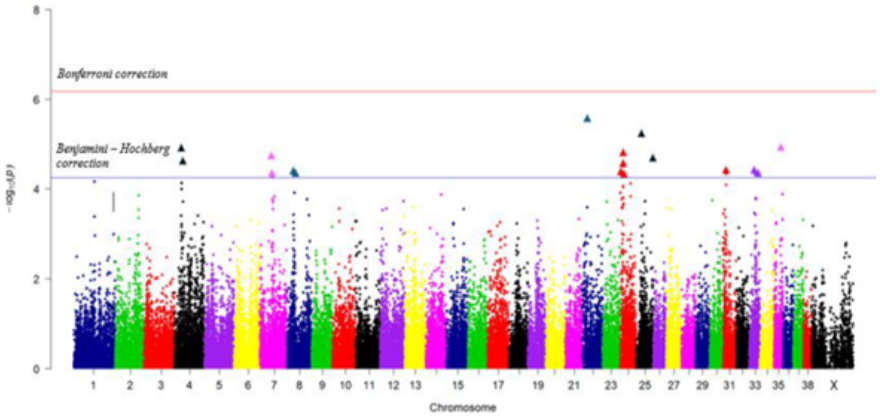


Figure 2
Manhattan Plot of the implemented GWAS analysis. The chromosomes are ordered from chromosome 1-38 (not all axis labels are visible). Chromosomes next to each other appear in different colours for better visual. Triangles mark the significant SNPs in the same colour as the corresponding chromosome. Figure 6 from Bence Morvay's M.Sc. thesis (2025).

Kaeberlein et al. (2016). The Dog Aging Project: Translational Geroscience in Companion Animals. doi: 10.1007/s00335-016-9638-7.

Jónás et al. (2020). A Preliminary Study to Investigate the Genetic Background of Longevity Based on Whole-Genome Sequence Data of Two Methuselah Dogs. doi: 10.3389/fgene.2020.00315.

Jónás et al. (2024). Dog Domestication Strongly Relied on Translation Regulation According to Differential Gene Expression Analysis. doi: 10.3390/ani14182655.

Jónás et al. (2025). Differential Gene Expression Analysis of Whole Blood Transcriptome Between Young and Old Border Collie Dogs. doi: 10.3390/vetsci12020086.

Jónás et al. (In prep.). Differential transcript usage of canine RNA-seq samples from brain cortical region.

Sándor et al. (2022). Poly(A) RNA sequencing reveals age-related differences in the prefrontal cortex of dogs. doi: 10.1007/s11357-022-00533-3.

Morvay B. 2025. Investigation of the word learning ability of dogs using GWAS methods. M.Sc. thesis.

How supercomputers teach us to “think like systems engineers”?

Komondor, language models, and industry convergence for a new era of troubleshooting

Author:
Márk Dénes

In modern IT operations, a large number of incident tickets must be handled daily under tight deadlines, while documentation often contains only the final solution and not the underlying engineering reasoning. The aim of our research was to reconstruct solution strategies from the text of incident tickets using artificial intelligence.

Using a training dataset generated in a supercomputing environment, we fine-tuned several large language models. Among these, the larger models were already able to propose solution steps with accuracy approaching human performance.

In modern IT operations, fast and reliable troubleshooting is one of the main challenges. An ITSM provider receives dozens of error tickets every day – typically 150-300 at our company – which must be handled within a short SLA timeframe. Experienced engineers know the „how," but their worksheets usually only contain the „what"– the conclusion, or solution. I am an engineer myself, and when I realized how much value lies behind the lines, I decided to try to capture the engineering mindset itself. As an industry player, I had the rare opportunity to talk about this last November in Budapest and this year at the ASHPC25 conference in Slovenia – I would like to briefly present where we have gotten since then.

DAM IT has a huge number of previously solved bug tickets available, but the decisions, dead ends, and risk analyses behind them are often only verbally communicated information and part of engineering knowledge. The question is quite simple, yet bold: **Can we train an engineer's solution strategy based solely on the text of the ticket and worksheet?** First, we tried it out with colleagues to get some perspective. We asked three Windows system engineers to write solution strategies for ten tickets each that they had not solved themselves and where the worksheet only contained the final result. We then rated them on a scale of 0 to 3 (0: dead end, 1: wrong direction, but the conclusion is logical, 2: right direction, but does not contain the real solution, 3: correct strategy, contains the real solution). The average score for the small survey was 2.46, meaning that the strategy can indeed be „extracted" from the text. In most cases, the right direction was found, and often the correct solution as well.

The next question was: **if an engineer can do this, can a language model do it too?** However, fine-tuning the model requires not only a large amount of data, but also high-quality data, which was not available in our case. We tried a different approach: using a multi-step system prompt, we instructed a large, reasoning-capable model (DeepSeek R1) to create a solution strategy based on the report and worksheet. We optimized the prompt in several rounds—with engineering feedback, this time involving not only Windows but also network, Linux, Service Desk, and system administrator participants, as well as examples related to these topics. The final 5-shot system prompt covered a sufficiently wide range of areas and report types, ultimately generating 2057 examples. DeepSeek R1 scored an average of 2.4 on the previous 0-3 scale test with this prompt, surprisingly close to human performance. An important part of this result was that we first had the model prepare a short summary and problem interpretation, as well as a risk analysis, and only then could it focus on writing the strategy.

The short answer to why all this had to be run on HPC is that thinking models write long texts and have enormous memory requirements. DeepSeek R1 is a giant Mixture of Experts (MoE) with 671B parameters; a single example generated 8,000-12,000 tokens (mostly due to the internal Chain-of-Thought monologue), of which the summaries and strategy accounted for approximately 3,000 tokens. The Komondor supercomputer AI node had 8 NVIDIA A100-40GB cards, totaling 320 GB, which allowed us to load the Q2_K_XL quantized variant with the llama.cpp backend. The generation speed was 10-15 tokens/s, depending on the context size, and it took 8-10 minutes to generate one example. This means that generating 1000 examples would take about a full week and require roughly 1300 GPU hours. We experimented with other backends (vLLM, exl2/exl3), but a quantized version of the right size was not really available at the time, and llama.cpp does not scale well for batch inference greater than 1, so Komondor became our choice.

The training data opened up a new possibility in that we had a relevant solution strategy for each ticket: **is it possible to train the model to generate a relevant solution strategy purely from tickets, without using the conclusion (worksheet)?** We performed LoRA SFT on multiple models in our Etalon training system, which we developed and optimized for multiple GPUs and HPC. Training took between 8 and 24 GPU hours, depending on the model size; out of 6,000 runs, checkpoints 4,000-5,000 typically showed the best performance. The small models (Gemma 3 1B/4B and DeepSeek R1-based 8B) did not result in a usable strategy; Gemma 3 12B already started to „think" as expected, but the content did not perform as expected. In the end, the larger models proved successful: Gemma 3 27B and Qwen 3 30B A3B both achieved an average of 2.2 based solely on the ticket text, often with a strategy that included the correct solution. An important difference from an operational point of view is that the Qwen model generated the strategy 5-6 times faster than Gemma thanks to its MoE architecture, which is a key advantage in ticket sales.

Results (scale of 0–3, based solely on the ticket text, without knowledge of the worksheet):

Model	Parameters	Result
Gemma 3	1B	0,0
Gemma 3v	4B	0,0
Gemma 3	12B	1,4
Gemma 3	**27B**	**2,2**
Qwen 3	**30B A3B**	**2,2**
Qwen 3	32B	2,1
DeepSeek R1 0528 Qwen 3	8B	0,0

Example of a strategy generated for an imaginary partner's ticket: (Translated from Hungarian)

Ticket Text:

Partner: Boltzmann Kft. | Name: Mark Denes | Operation Service Level: Server operation | Type: Incident | Priority: 3 (normal) | Support Grade: System administrator | Subject: DC (192.168.1.30) WARNING SERVICE: Volume C: | Message: DC (192.168.1.30) WARNING SERVICE: Volume C: WARNING: Disk space usage below 10%

Summary:

Normal priority monitoring incident to be handled by a system administrator-level specialist at the server operation service level. The reporter, Mark Denes, is an employee of Boltzmann Kft. The problem concerns disk space usage on the C: drive of the DC (Domain Controller) server, where the alert indicates a value below 10%.

Problem interpretation:

On the Domain Controller, the C: drive is typically associated with the Windows operating system, applications, and temporary files. The low free space of 10% can pose a serious risk to server operation:

- Installation or running of operating system updates may be hindered, which can leave security vulnerabilities open.
- The Windows fast startup (Hybrid Boot) feature may be interrupted, which can slow down system startup.
- The operation of Exchange or other critical services may be temporarily affected if there is insufficient space to store log files or perform temporary data management.

The reasons for disk space reduction can include rapidly growing logs, accumulation of temporary files, undeleted update packages, or simply exhaustion of physical disk capacity. It is important to note that in the case of DCs, the system discovery (SYSVOL) directory and GPO subdirectories can also take up significant space.

Solution Strategy:

The primary goal is to free up disk space immediately to avoid falling below the 10% threshold, while investigating the underlying cause is secondary. The solution process must ensure stable server operation and uninterrupted critical services.

1. Detailed analysis of disk space usage with tools (e.g., WinDirStat, TreeSize) to identify the largest file types and directories. This can help filter out unnecessary or obsolete data.

2. Delete temporary Internet files (Temp Internet Files) and remove items in the „Recycle Bin" folder, which often accumulate.

3. Manually delete Windows Update updates (using Disk Cleanup or by forcing a cleanup of the „Component Store"), as these can take up significant space.

4. For Exchange services, review message queues and the latest logs, paying particular attention to „ActiveSync" or „Mailbox Transport" services, which can sometimes accumulate large amounts of data.

5. If the above steps are not sufficient, it is recommended to consult with the user about the existence of large files (e.g., videos, large backup data) that can be moved to other storage devices.

6. Finally, monitor server performance and temperature during operational status assessment, as lack of disk space may indicate other problems.

Built into the corporate system, the model automatically provides a three-part output for newly arrived tickets in a matter of seconds: a brief summary, problem interpretation (risk/impact), and solution strategy. This greatly assists Service Desk colleagues when they need to delegate tickets, and also saves engineers significant time when determining the scope of the problem.

As for the continuation, in our project running in 2025, we are teaching our own developed and managed agents (LLM Agent) to solve problems automatically, where strategy generation was the very first step and opened the doors. The curriculum development method presented above has proven itself in many other domains, so it is not necessarily only useful in an industrial or business context.

Whether it is research into elegant engineering solutions to operational problems or other everyday phenomena, the Komondor supercomputer provides a reliable and ideal environment for these, giving space for breakthroughs to happen.

Thanks *to the engineering team for their evaluations, the domestic HPC community, and the Komondor operators for their reliable capacity. I feel honored to be able to participate in this as a „guest researcher" – it is a great experience to see these two worlds come together, and I am excited to see what the future holds.*

Márk László Dénes, MS-HW Systems Engineer, AIOPS Team Leader • DAM Invisible Technology Ltd. • denes.mark@damit.hu

Computer simulation and virtual reality

Author: *Dezső Boda*

One of the fundamental questions of science is how ordered behavior emerges from the vast number of interactions between atoms and molecules. In our research, we studied nanoscale channels, so-called nanopores, through which ions migrate under the influence of an electric field. Experiments show that in certain cases the pore does not behave as expected: for example, in a $CaCl_2$ solution it does not exhibit ion selectivity. The cause of this phenomenon was revealed using large-scale molecular simulations that model the system atom by atom on a supercomputer.



Humans are curious by nature, and scientists are especially so. We want to understand why the Universe works the way we observe it works. Our curiosity is not limited to things we can see with our eyes or a telescope; we are just as interested in the microscopic world of atoms and molecules that make up everything around us, including our own bodies. The fundamental question is this: how does the spontaneous behavior of an unimaginably large number of atoms lead to the kind of order that we can observe with our eyes or instruments?

What do we mean by "order" here? Essentially, anything in which we can detect patterns can be considered an ordered system. A liquid is more ordered than a gas. A crystal is more ordered than a liquid. Galaxies arrange themselves into structures that are visibly different from randomness. Interestingly, these structures (networks) look strikingly similar to the networks formed by brain cells. From a technology perspective, imagine we have a system and we hit it with some signal (an electrical pulse, light, pressure—it could be anything). How does the system respond—what is the output signal? If we discover some regularity between the input and the output, that indicates a kind of order in the system. Our goal is to understand why the system responds to a stimulus the way it does.

However, the connection between input and output may be determined down at the level of atoms. How can we study that level? By turning to the systems of interest: nanopores are channels whose diameters lie in the nanometer range. This is the size scale of atoms. A diameter of a water molecule is about 0.3 nm (0.3 × 10-9 m). What happens inside such a narrow pore? What methods can we use to study such a nanopore?

First, what is a nanopore? Imagine a thin membrane separating two electrolyte solutions, with a tiny hole in the membrane—the nanopore. In an electrolyte solution, the solvent is typically water, and ions "float" in it. The best-known electrolyte solution is table salt (NaCl) dissolved in water—seawater. If an external electric field is applied to this system, ions will migrate through the nanopore from one side to the other. A key question is which ions pass in greater numbers—that is, what is the nanopore's selectivity. Many factors influence this, but for instance, if the pore wall is negatively charged, we would expect the pore to be selective for positively charged ions (cations, such as Na+). That is indeed what happens for NaCl, but not for $CaCl_2$. Why not? That is the question we asked.

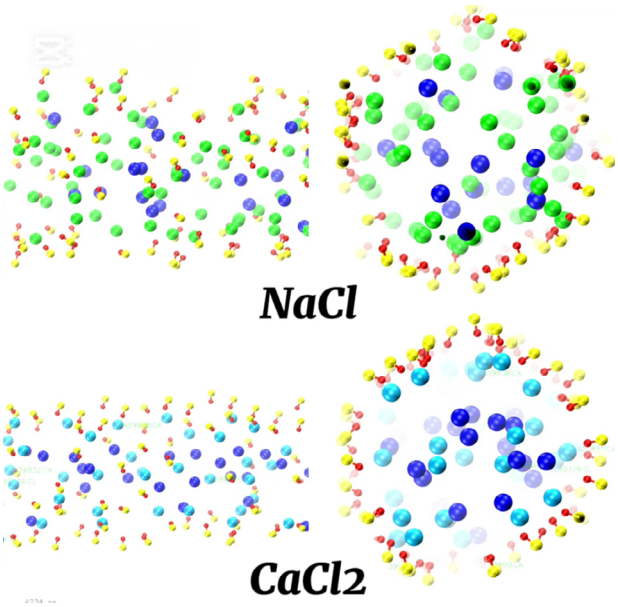
Experiments can demonstrate that in the case of $CaCl_2$ the pore is not selective; it allows Ca^{2+} and Cl^- to pass in roughly similar proportions [1]. However, experiments do not answer why this is so—what the molecular mechanism behind this behavior is. To answer that, we would need a microscope that not only shows us atoms, but also accounts for their motion. No such microscope exists. Or does it? In a sense, it does—computer simulation. The idea is to model the particles that make up the system (the ions, the water molecules, the atoms of the membrane, and everything else), along with the interactions between them, and then simulate their behavior on a computer [2,3].

This is pretty similar to how the epic battle scenes in The Lord of the Rings were created: the behaviors of orcs and elves were modeled, and then the computer generated the battle. The computer "knew" what would happen when an orc met an elf. Molecular simulation is the same idea, except the orcs and elves are atoms and molecules. In this way, one can make impressive movies (https://youtu.be/SBHBw7req98?si=36FjRDBWG510xQEB), but we must remember at every moment that we are not watching reality but a virtual reality. We do have one powerful tool to distinguish between them: if the simulation produces the same result as the experiment, we accept the model. There is nothing mysterious about this—that's how physics works. Physics can only study a model, not reality itself. As Professor Sándor Papp liked to say: "a model is an amputated reality."

These simulations, however, are extremely computationally intensive. We need to model tens of thousands of particles, millions sometimes. That requires powerful computers. With the Komondor supercomputer, we were able to carry out these simulations, analyze the results, and draw conclusions from the analysis.

Let's conceptually split the nanopore into two regions: a surface layer next to the wall and a central region. The figure and the video show that cations accumulate near the wall (Na+ as green spheres, Ca^{2+} as light blue spheres) to neutralize the negative charges on the wall (small red spheres). Anions (Cl^-, dark blue spheres) do not really enter this surface region. In this zone, NaCl and $CaCl_2$ behave very differently. Na+ ions can move within the surface region, and because there are more of them than Cl^- ions, the pore becomes Na+-selective. In contrast, Ca^{2+} ions (because of their double charge) stick strongly to the negative charges on the wall and thus become

paralized—their mobility becomes very small. As a result, Ca^{2+} ions in the surface region do not contribute to the current. But Cl^- ions in that region do not contribute either because they are not present there. Therefore, in $CaCl_2$ the current is carried only in the central, non-selective region, and so the pore as a whole is non-selective.



The snapshots on the top are from a simulation of a nanopore containing NaCl, shown from the side and along the longitudinal axis; the bottom is the same for $CaCl_2$.
Video link: https://youtube.com/shorts/sNgINE4XPgM?si=opkDPiZovIkL8YZw

Of course, we can support these statements with various quantitative data and functions extracted from the simulations, but the simple reasoning above offers an intuitive description of what we observed. Today, computer simulations are used for many purposes and have matured into easy-to-use software packages that can be adopted just like a measuring instrument. Simulation techniques have outgrown their infancy and now stand alongside experimental techniques as powerful, complementary methods of investigation.

[1] Y. He, D. Gillespie, D. Boda, I. Vlassiouk, R. S. Eisenberg, and Z. S. Siwy. Tuning transport properties of nanofluidic devices with local charge inversion. JACS, 131(14):5194-5202, 2009.

[2] M. Valiskó, B. Matejczyk, Z. Ható, T. Kristóf, E. Mádai, D. Fertig, D. Gillespie, D. Boda. Multiscale analysis of the effect of surface charge pattern on a nanopore's rectification and selectivity properties: from all-atom model to Poisson-Nernst-Planck. J. Chem. Phys. 150(14):144703, 2019.

[3] S. Shabbir, D. Boda, Z. Ható: Transport properties of CaCl2 are determined by the balance of ion-ion and ion-water interactions as revealed by molecular dynamics simulations. J. Mol Liq. 426:127308, 2025

When Artificial Intelligence Restores the Joy of Movement: How a Hungarian Startup is Revolutionizing Knee Surgery

Authors:
Peter Szoldan, Prof. Dr. Laszlo Hangody, Endre Szabo, Dr. Zsofia Egyed

The Team: Where Medicine Meets Artificial Intelligence

GenuForm AI is composed of professionals who uniquely combine medical and technological expertise:

Peter Szoldan (CEO) - AI engineer, earned his degree at Harvard University, previously worked as a consultant at McKinsey & Company

Prof. Dr. Laszlo Hangody (CMO) - Orthopedic surgeon with 30+ years of clinical practice, world-renowned inventor of mosaicplasty

Endre Szabo (CTO) - Mathematician and AI expert, researcher in combinatorial geometry

Dr. Zsofia Egyed - Radiologist with 30+ years of clinical practice



A Problem Affecting Millions

Imagine being able to squat, run, or simply walk without pain, something many of us take for granted. Yet approximately 30% of the population over 55 suffers from some degree of knee cartilage degradation, and half of them experience painful symptoms. The knee joint is one of the most stressed points in the human body: it not only carries our body weight but is subjected to tremendous forces with every step, jump, and squat.

The current solution, total knee replacement, while widely adopted, has significant limitations. After surgery, patients often regain only limited range of motion, and after 15-20 years, the implant loosens, requiring another operation. But what if there was a better solution?

Donor Cartilage

The answer has actually been known for a long time: transplanting cartilage tissue from donors. This method, unlike artificial prosthetics, can restore full range of motion and serve the patient for life. Moreover, it's immunologically ideal: there is no organ rejection when transplanting cartilage tissue.

However, there is a critical obstacle: geometric matching. Cartilage is a rigid tissue; if the shape of the donor's and recipient's knees doesn't match perfectly, the success of the surgery becomes questionable. Until now, this matching was done manually, which was extremely time-consuming and imprecise. That's why only about 10,000 such surgeries are performed worldwide annually, while the potential demand is twenty-five times greater.

Artificial Intelligence

GenuForm AI solves precisely this bottleneck. The Hungarian-developed system matches donors and recipient patients in four steps using artificial intelligence:

1. **Patient Mapping:** Based on MRI scans, the artificial intelligence determines the exact geometry of the patient's existing cartilage.

2. **Digital Reconstruction:** The system „restores" the healthy state: it determines what the damaged area would look like when intact.

3. **Donor Mapping:** When donor tissue becomes available, the artificial intelligence immediately analyzes its geometry.
4. **The Match:** A combinatorial geometry algorithm finds the perfect match between donor and recipient patient. Often, multiple patients can be served from a single donor cartilage, and the algorithm optimally divides the available tissue**.**

Training our artificial intelligence models and performing the complex geometric calculations requires enormous computational capacity. During development, we relied on the resources provided by the Komondor supercomputer, without which the research could not have progressed at this pace. High-performance computing (HPC) enabled us to train our 3D image processing algorithms on large volumes of medical imaging data and solve combinatorial matching tasks in real time

The Success of the First Surgery

In 2024, GenuForm AI technology was applied to a patient for the first time, under a research license. The results were excellent. The rapid surgical timeline of 24-48 hours ensures the tissue remains viable, dramatically increasing the surgery's success rate.

Why We Do This

The GenuForm AI team is driven by the conviction that the intersection of artificial intelligence and medicine holds enormous, yet untapped potential. 3D computer vision technology has reached the level of maturity in recent years that enables solving such complex geometric matching tasks, and we want to put this opportunity in service of people.

Our goal is to apply our decades of medical and technological experience for the benefit of healthcare. We believe that cartilage transplantation can become as routine a procedure as prosthetic implantation within 5-10 years, just with better outcomes, full range of motion, and lifelong solutions. We want to ensure that people don't have to give up the joy of movement: gardening, walks, sports, everything that defines quality of life.

TENGR.AI: a secure AI alternative built on Hungarian infrastructure

Author: *Tibor Horváth, CEO of Tengr.ai*

In generative AI, the last few years have been shaped as much by access to computing power as by better models. The biggest players usually rely on large-scale cloud GPU infrastructure, and that creates a serious barrier to entry, both technologically and financially. The pressure is especially visible in Europe, where development teams often must work with tighter resources while still competing in a global market.

In that context, the more interesting approaches are the ones that rethink how generative systems are run and how resources are allocated. The aim is straightforward: get a better cost-performance balance without giving up control over the system. TENGR.AI fits that pattern. It combines self-managed infrastructure with Hungarian high-performance computing resources, creating a hybrid setup that works differently from the usual cloud-first model.



Introducing TENGR.AI

TENGR.AI is an uncensored platform focused on media generation and editing. It produces content from both text and visual inputs and does not employ hard bans or blanket refusals. Instead, it turns user requests into forms the system can process and answer within safe boundaries.

The platform's evolution is governed by the idea that the real value of generative models often lies less in one-off image creation and more in workflows that are repeatable, controlled, and able to run at scale. In that sense, TENGR.AI is not just a creative tool. It can also be seen as a visual processing pipeline, with the stages of generation laid out in a structured way.

One of the key elements of that approach is the use of so-called "recipes", which capture multi-step generation workflows. These recipes combine different models, parameters, and processing steps, making it possible to reproduce a given visual result consistently. That matters in use cases where consistency and iteration are essential, such as marketing materials, creative campaigns, or product visualization. The logic is close to a classic software pipeline, where reproducibility and reuse matter as much as the final output.

Data security and European sovereignty

Data handling and user control are central to how the platform positions itself. The architecture aims to minimize the amount of user data it has to handle while keeping control over generated content primarily in the user's hands. That fits broader European goals around digital sovereignty and privacy by design. At the same time, the real strength of those benefits always depends on implementation details, so they should not be treated as universal guarantees.

That philosophy is reinforced by the Vault feature, one of TENGR.AI's new defining components. Vault is based on client-side encryption: content is encrypted and decrypted on the user's device, so anything stored on the servers remains encrypted. In practice, that means the provider still stores the data, but cannot access it at the content level because it does not hold the keys needed to decrypt it.

This gives users more control and lowers the risk that sensitive or commercially valuable material becomes accessible to unauthorized parties. Still, the security of a system like this depends heavily on key management and on the quality of the client-side implementation. In that sense, Vault should be seen as a strong architectural privacy model rather than an absolute guarantee. It can provide a high level of isolation in practice, but not unlimited protection.

Comparable solutions do exist elsewhere in the market, but client-side encrypted storage is still relatively uncommon as a central feature of generative AI platforms. That makes Vault less of a novelty claim and more of a deliberate product choice, one that puts user data handling alongside functionality as a core design concern.

Taken together, this makes TENGR.AI more than just another generative platform. It positions itself as a system designed to fit European expectations around privacy and digital self-determination.

The role of Hungary's supercomputer

One of the defining elements of TENGR.AI's operation is its hybrid infrastructure model, which combines self-managed systems with the resources of the Komondor supercomputer. This setup makes it possible to separate different kinds of workloads and handle them more efficiently, which is particularly important in generative AI systems. Day-to-day development and inference tasks typically run on the company's own infrastructure, where lower latency and more direct operational control are easier to maintain. More compute-intensive processes, such as model training or large-scale experimentation, are moved into the HPC environment, where large-scale parallelism can be used more effectively.

This kind of workload separation can also help fine-tune the balance between cost and performance. HPC resources make it possible to handle larger models and more complex training processes, while self-managed infrastructure preserves operational flexibility and supports rapid iteration. The result is a form of resource orchestration, with each component running where it fits best.

Domain-specific fine-tuning with E.ON

In generative AI systems, there is a fundamental difference between the training phase and the serving phase. Training usually requires large, highly parallelizable workloads, while inference depends more on low latency and stable operation. TENGR.AI's architecture reflects that distinction directly: the HPC environment is the space for development and experimentation, while the production environment is tuned to serve user needs reliably. Separating the two can also reduce interference between development and live systems, leading to more predictable operation.

A practical example of this architecture can be seen in a joint project with E.ON focused on water polo. The goal of the initiative was to make generative models represent the visual characteristics of a specific sport more accurately, since general-purpose models often struggle in these kinds of contexts. The first step was to assemble a domain-specific dataset that included both image and text sources. The models were then fine-tuned to better reproduce the sport's characteristic movement patterns and visual structures.

That work was supported by an iterative validation process in which the results were continuously compared against real visual material from sports events. Projects like this show that practical generative AI often goes beyond using general-purpose models off the shelf. In many cases, it requires domain-specific fine-tuning, and that in turn demands substantial computing resources. In this context, HPC is not just a supporting layer. It is a key part of the development process itself.

The hybrid model also matters from a collaboration standpoint. HPC environments are typically shared infrastructures that can be accessed by multiple actors, creating space for research and industry partnerships. In that setting, computing capacity, development experience, and domain knowledge can accumulate in one place, which can speed up progress over time. This is especially valuable in projects where both computing power and expert knowledge are critical.

Outlook

One likely direction for future development is reducing the gap between traditional HPC environments and modern container-based software architectures. Conventional supercomputing systems are often built around batch-oriented operation, which does not always fit the iterative and interactive nature of AI development. Approaches such as extended Slurm-based scheduling or the integration of containerization layers can make these systems more flexible for AI workloads. That matters because it can narrow the gap between development and production and shorten time to market.

At the same time, it is important not to overstate the case. HPC-based architectures are not a universal solution for every generative AI application. Integrating and operating these systems brings its own complexity, especially when heterogeneous environments and different ways of operating must work together. Effective use often requires specialized expertise, and in some cases cloud-based solutions may still be simpler or faster to scale. The right choice depends heavily on context and on the requirements of the specific application.

What TENGR.AI shows is that generative AI development does not have to rest entirely on cloud infrastructure. Well-integrated national or regional HPC resources can offer a viable alternative, especially where cost efficiency, control, and domain-specific development all matter at once. This is likely to attract growing attention in Europe, where technological sovereignty has strategic importance. The generative AI ecosystem will probably not organize itself around a single dominant infrastructure model. It is more likely to evolve through the coexistence of different architectures and resource strategies, with HPC-based approaches emerging as a credible, if context-dependent, option.

DKF | hpc



EURO²

Funded by
the European Union

The EuroCC2 project is supported by the European Union's Horizon 2020 research and innovation framework program in accordance with Grant Agreement No. 101101903

hpc.dkf.hu

@HPC CC Hungary

@HPC_hu