Complex systems

https://doi.org/10.1038/s43588-025-00929-z

Discovering the laws behind complex networked systems

Iacopo Iacopini & Eugenio Valdano

A recent study shows that neural symbolic regression offers a route to automated discovery of governing equations for network dynamics across high-dimensional complex systems.

Complex systems are all around us, from gene regulatory networks controlling cellular function, to interacting ecological communities and the spread of diseases in human and animal populations. Describing these systems and understanding how they evolve requires mathematical laws that capture the interactions between the system's units and the dynamical processes unfolding on them. The increased abundance of observational data combined with unprecedented computational power has given us access to the heterogeneous and high-dimensional nature of complex networked systems. Yet identifying general and interpretable mathematical laws governing them remains a fundamental challenge. Writing in this issue of *Nature Computational Science*, Zihan Yu and colleagues present ND^2 (Neural Discovery of Network Dynamics), a method that can automatically discover mathematical formulas governing how complex systems evolve over networks[1].

The history of science is punctuated by moments when the right mathematical law unlocked profound understanding, leading to widespread revolutions that cross disciplines (Fig. 1). While Kepler's laws could predict planetary motion, we only started to explain it with Newton's law of gravitation. Such mathematical relationships do more than predict. Each formula becomes a lens through which scientists interpret observations, design experiments, and build new theories. This is why the predictive power of machine learning and the generalizability — and interpretability — of mathematical laws are giving rise to a new artificial intelligence (AI) field: automated scientific discovery[2].

However, the high dimensionality — and apparent lack of symmetries — of complex networked systems pose additional challenges to discovering simple governing laws: the search space for symbolic formulas grows super-exponentially with network size. ND^2 solves this by introducing a new encoding of the system that makes this search independent of the number of nodes in the underlying network. A transformer then guides the search to identify the most suitable mathematical formula describing the encoded dynamics, having been pre-trained on randomly generated dynamical systems and networks.

Notably, the success of ND^2 supports the recently proposed low-rank hypothesis of complex systems, which posits that low-dimensional mathematical representations can indeed capture high-dimensional network dynamics[3]. It also illustrates that transformers trained on simple synthetic data can help identify such representations in real and noisy systems.

ND^2 recovers established dynamics across several synthetic benchmarks and, more importantly, it is applicable to contexts where

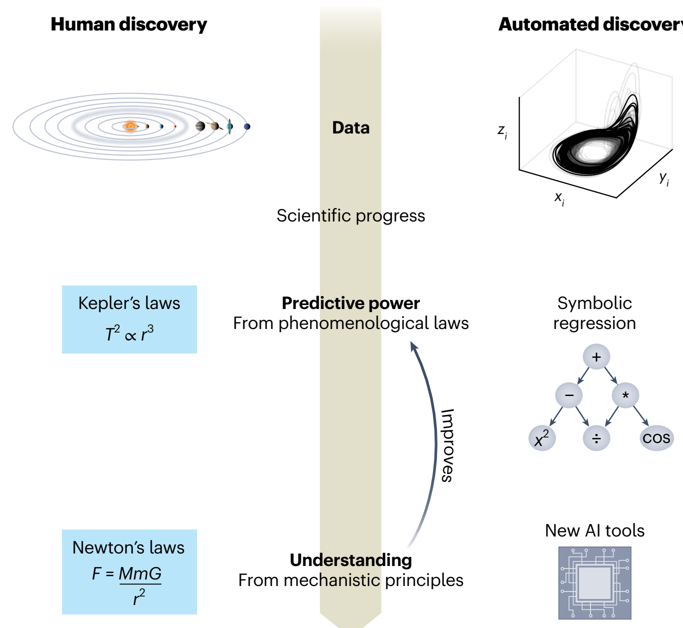

Fig. 1 | Scientific discovery, from prediction to understanding. Historical example: astronomical data led to Kepler's phenomenological laws of planetary motion, later explained by Newton's law of gravitation. Modern analogue: symbolic regression yields predictive formulas. Future progress in AI-driven automated scientific discovery will need to tackle Newton-like mechanistic understanding.

governing laws and network structures are unknown or poorly known, as in microbial communities and gene regulatory networks. In the latter, ND^2 found that a gene's effect on another is modulated by additional genes, in line with recent evidence of higher-order interactions[4].

The authors emphasize that their method does not rely on prior external knowledge about what the mathematical formulas should look like, unlike other recent methods that rely on predefined libraries of candidate functions[5,6]. This highlights a broader discussion in automated discovery: what role should prior knowledge play? While Yu and colleagues demonstrate the power of avoiding strong priors, their approach represents one epistemological extreme. Alternative stances, like Guimerà's Bayesian machine scientist, intentionally incorporate prior, human-generated, scientific knowledge[7]. So how far should automation go? On the one hand, only agnostic models can discover laws that do not depend on the contingent path of scientific progress. On the other hand, embedding domain-specific knowledge via physics-informed pre-training or inductive bias[2] may be necessary to guarantee that discovered formulas do represent new scientific understanding – instead of being a sophisticated form of regression that maximizes predictive power.

This tension between constraint-free and informed search will shape automated scientific discovery. Do methods such as ND[2] reveal new frameworks or merely fit phenomenological patterns? Two challenges lie ahead. First, predictive power can trap you in a 'local optimum'. In 1928, the British physicist P. A. M. Dirac derived the equation that now bears his name to describe spin-1/2 particles such as electrons. The equation implied the existence of negative energy states that made no sense at the time. Dirac, however, believed that[8]: "It is more important to have beauty in one's equations than to have them fit experiment." He was right, and those seemingly unphysical solutions led to the later discovery of antimatter. Now, what is the role of simplicity and beauty in automated discovery and how can models encode it? This question may have been futile in the past, but the idea of automating aesthetic criteria may now be within the reach of generative AI. Second, groundbreaking scientific discoveries often require using the right mathematical framework and identifying the symmetries of the system. Take the example of special relativity. Albert Einstein opened his famous 1905 paper with the words[9] "It is known that Maxwell's electrodynamics leads to [...]", referring to the modern description of electromagnetism by J. K. Maxwell in 1861–1862. Einstein did not derive relativity by optimizing formulas on a large body of data on high-speed motion, which did not exist at the time. He built it from the only possible formulas for the change of coordinates (Lorentz transformations) that were compatible with Maxwell's theory. Will future models for scientific discovery be able to discover the hidden symmetries and constraints in high-dimensional, heterogeneous systems, leading to mathematical formulas that are beyond phenomenological tools? In this sense, the low-dimensional representation underpinning ND[2] is promising, as is the growing body of work on automating the discovery of symmetries and conservation laws from data[10,11].

As automated discovery matures facing these questions, it may complement rather than replace human intuition in revealing mathematical structures underlying complex systems. Traditional approaches often refine established models incrementally, adding parameters based on new observations: a path dependency in scientific discovery that may constrain theoretical exploration. Automated methods could help break free from these established routes, potentially revealing unexpected mathematical relationships. Whether such discoveries represent genuine scientific breakthroughs or sophisticated pattern recognition is, however, is yet to be shown.

Iacopo Iacopini[1,2] ✉ **& Eugenio Valdano**[3] ✉

[1]Network Science Institute, Northeastern University London, London, UK. [2]Department of Physics, Northeastern University, Boston, MA, USA. [3]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Paris, France.

✉e-mail: i.iacopini@northeastern.edu; eugenio.valdano@inserm.fr

Published online: 05 January 2026

References

1. Yu, Z., Ding, J. & Li, Y. *Nat. Comput. Sci.* https://doi.org/10.1038/s43588-025-00893-8 (2025).
2. Wang, H. et al. *Nature* **620**, 47–60 (2023).
3. Thibeault, V., Allard, A. & Desrosiers, P. *Nat. Phys.* **20**, 294–302 (2024).
4. Battiston, F. et al. *Phys. Rep.* **874**, 1–92 (2020).
5. Gao, T. T. & Yan, G. *Nat. Comput. Sci.* **2**, 160–168 (2022).
6. Hu, J., Cui, J. & Yang, B. *Nat. Commun.* **16**, 6226 (2025).
7. Guimerà, R. et al. *Sci. Adv.* **6**, eaav6971 (2020).
8. Dirac, P. A. M. *Sci. Am.* **208**, 45–53 (1963).
9. Einstein, A. *Ann. Phys.* **322**, 891–921 (1905).
10. Liu, Z. & Tegmark, M. *Phys. Rev. Lett.* **126**, 180604 (2021).
11. Lu, P. Y. et al. *Nat. Commun.* **14**, 4744 (2023).

Competing interests

The authors declare no competing interests.