THE AUTHOR FILE

Olga Troyanskaya

YETI emerges from a dance between genomics and computation.

"My goal is to solve biological problems," says Olga Troyanskaya, who is on the Princeton University faculty in computer science and at the Lewis-Sigler Institute for Integrative Genomics. She is also deputy director for genomics at the Flatiron Institute, which focuses on data-driven analysis and methods for the sciences and is funded by the Simons Foundation.

Troyanskaya's latest platform is Your Evidence Tailored Integration (YETI), the idea for which germinated early in her career. "It used to be that all the biological knowledge was either in the literature or, in some ways, in the head of the smartest biologist you know," she says. Such "gurus" still exist, but the available data amounts are now so vast that biologists usually cannot readily place their data in the context of existing data mountains to find answers to their question of interest. This is true even for biologists with advanced programming skills. "It's incredibly non-trivial beyond programming, storage and all the technical issues," she says.

YETI has a full computational belly: 237 weighted data networks that are functional maps built from all existing 'omics data on biological pathways. Troyanskaya and her group generated these networks using a machine-learning approach called context-sensitive regularized Bayesian integration.

Also in YETI's gut is an algorithm called Lasso, which is the user's advocate in that it computationally detects and selects the networks most relevant to the user's datasets. Lasso takes signals in the user's dataset to identify pertinent functional maps. Next is a computational hand-off in which a network is built tailored for the user's further analysis. Lasso has tasks: 100 steps are built into this algorithm.

For now, YETI lives on Princeton servers and is maintained by Troyanskaya's group. Eventually, it will become part of HumanBase, a larger, free-to-use Flatiron Institute resource that Troyanskaya directs. HumanBase is set up for scientists to do data-driven predictions of gene expression, function and regulation.

Although Troyanskaya's work is mainly computational, she adores biology, which

Olga Troyanskaya. Credit: R. Dannenfelser

drives her to build ways to make it easier for biologists to more easily ask their questions of data. "I've just always been fascinated by biological questions but I'm honestly, at the end of the day, just much better at the computational aspect of things," she says.

As a teenager, she discovered books about genetics on her family bookshelf. "It was the most interesting thing I'd read, I very clearly remember that," she says. "I just fell in love with it." She gave a lecture on Down syndrome in her class, which was not exactly typical in the Soviet Union of her youth.

As an undergraduate in the US, Troyanskaya double-majored in computer science and biology and minored in math, and she completed her PhD at Stanford University. "I feel so lucky," she says. Early in her career, she worked on yeast and *Caenorhabditis elegans*. Since then, she has been working on mice and projects about people: neurodevelopment, neurodegenerative disease, kidney disease and breast cancer, to name a few.

"Because I'm a methods person so I actually can do something useful in each of those areas," she says.

All of Troyanskaya's team members have long-term collaborations with external scientists. "They love it because really that's how you make an impact," she says. Together, the researchers consider biology and computation, but it's not easy. "A lot of it is teaching each other," she says. "They have to be patient with us, we have to be

patient with them and, in the end, you can do something that you could never have possibly done separately," she says.

Beyond the lab, Troyanskaya spends time with family. She has two children, one of whom just took part in a national swim meet. He is 7. He is committed to swimming and she supports him, "as long as he has fun." Troyanskaya used to swim competitively and she danced.

"I used to dance pretty seriously," she says—mainly ballroom and vintage, which is the re-creation of historic dance styles.

Her lab works hard, but she does not expect her trainees to only be in the lab and "never ever think about anything else," she says. "I feel like being a well-rounded person helps you think more creatively and be more excited about your work."

"Olga, she is my better self," says Vessela Kristensen, a cancer researcher at the University of Norway, a longtime collaborator who also spent a sabbatical year in Troyanskaya's lab. Troyanskaya is, for example, smarter, faster, more intense, more energetic. She is someone who delves into tough subjects and is "an awesome member" of the "ever so male-dominated field of computer science, tough and analytical, but also soft and emotional and caring."

"My goal is to solve biological problems."

Troyanskaya's office is filled with breast milk collection bottles, exotic teas, chili-flavored chocolates and the latest computer gadget, says Kristensen. She cares for her family as much as her current and former postdocs, now faculty members on both sides of the Atlantic. "Olga is ahead of us in many ways, and I am so happy to be on the ride from time to time," she says. □

Vivien Marx

Published online: 30 November 2018
<https://doi.org/10.1038/s41592-018-0226-5>

Reference

Lee, Y.-s. et al. Interpretation of an individual functional genomics experiment guided by massive public data. *Nat. Methods* <https://doi.org/10.1038/s41592-018-0218-5> (2018).