

# Allen Goodman

Senior Principal Machine Learning Engineer · Genentech (Frontier Research)

New York, NY · [goodmaa3@gene.com](mailto:goodmaa3@gene.com) · [0x00b1.github.io](https://0x00b1.github.io)

[github.com/0x00b1](https://github.com/0x00b1) · Google Scholar · ORCID: 0000-0002-6434-2320 · @0x00B1

---

## Research Interests

My research centers on differentiable programming and its application to the life sciences. I am interested in how end-to-end differentiable systems—spanning representation learning, generative modeling, and uncertainty-aware optimization—can accelerate the design of biological molecules and the interpretation of high-dimensional experimental measurements. My current work focuses on machine learning for therapeutic antibody and protein design, including protein language models, concept-based interpretability, and lab-in-the-loop optimization that tightly couples predictive and generative models with wet-lab experimentation. This builds on a longer line of work in computational microscopy and image-based profiling, where I helped develop widely used open-source tools for quantifying cellular phenotypes at scale. Across these areas, I am motivated by building performant, reliable, and reusable scientific software, and by methods that remain faithful to the structure of the underlying measurement and biology.

## Experience

### Genentech · Frontier Research (gRED)

New York, NY & San Francisco, CA

*Senior Principal Machine Learning Engineer*

May 2025 – Present

*Principal Machine Learning Engineer*

Feb 2024 – May 2025

*Senior Machine Learning Engineer*

Nov 2022 – Feb 2024

- Machine learning for therapeutic antibody and protein design within Genentech Research and Early Development (gRED).
- Contributions span protein language modeling and efficient pretraining, property-conditioned and generative design, model interpretability (concept-bottleneck language models), and uncertainty quantification.
- Lab-in-the-loop optimization coupling predictive and generative models with experimental feedback for functional antibody engineering.

### Meta · Fundamental AI Research (FAIR)

Cambridge, MA

*Research Engineer (E6)*

Sep 2021 – Nov 2022

- Research on data compression, optical neural networks, and first-class differentiability.
- Core contributor to *NeuralCompression*, *PyTorch*, and *Segment Anything*.

### Broad Institute of MIT and Harvard · Imaging Platform

Cambridge, MA

*Principal Software Engineer, Anne Carpenter Laboratory*

Sep 2015 – Jul 2021

- Lead engineer and co-author of *CellProfiler*, a widely used open-source platform for high-throughput biological image analysis (releases 3.0 and 4).
- Developed methods and tooling for image-based profiling and deep-learning cell and nucleus segmentation; contributed data and analysis to the 2018 Data Science Bowl.
- Chan Zuckerberg Initiative Engineering Fellow.

### Progress Chef

Seattle, WA

*Senior Software Engineer*

Sep 2013 – May 2015

- Infrastructure-automation software and tooling.

## Simple Finance

Portland, OR

Director of Products

Jan 2010 – Aug 2013

- Early team member and product lead at Simple, a consumer banking startup.

## Education

---

Amherst College

Amherst, MA

B.A. in Mathematics

2012

## Technical Skills

---

**Research areas** Differentiable programming · Machine learning · Computer vision · Numerical analysis · Signal processing

**Applications** Therapeutic protein & antibody design · Protein language models · Image-based profiling · Biological image analysis

**Engineering** Python · PyTorch · C/C++ · CUDA · Open-source scientific software

## Honors & Awards

---

- **Chan Zuckerberg Initiative Engineering Fellow** — Broad Institute of MIT and Harvard.
- **Multiple Scene.org Awards & Meteoriks Awards** — international demoscene competitions; the only American programmer to have won multiple awards in each.

## Selected Software & Open-Source Contributions

---

- **CellProfiler** — high-throughput biological image analysis; core developer & co-author (v3, v4).
- **Pycytominer** — reproducible image-based profiling; co-author.
- **PyTorch** — deep learning framework; contributor.
- **Segment Anything** — promptable image segmentation; contributor.
- **NeuralCompression** — neural data-compression research library; contributor.
- **Keras R-CNN** — cell detection in biological images; lead author.
- **Deepometry** — deep learning for imaging cytometry; co-author.

## Publications

---

Author name in **bold**. Full list: [Google Scholar](#) · [ORCID 0000-0002-6434-2320](#).

Serrano, E., Chandrasekaran, S. N., Bunten, D., Brewer, K. I., Tomkinson, J., Kern, R., . . . , **Goodman, A.**, et al. (2025). Reproducible image-based profiling with Pycytominer. *Nature Methods*, 22(4), 677–680.

Ismail, A. A., Oikarinen, T., Wang, A., Adebayo, J., Stanton, S., Joren, T., Kleinhenz, J., **Goodman, A.**, Corrada Bravo, H., Cho, K., et al. (2025). Concept bottleneck language models for protein design. *International Conference on Learning Representations (ICLR)*.

Frey, N. C., Hötzel, I., Stanton, S. D., Kelly, R., Alberstein, R. G., Makowski, E. K., . . . , **Goodman, A.**, et al. (2025). Lab-in-the-loop therapeutic antibody design with deep learning. *bioRxiv* (preprint).

Tagasovska, N., Ludwiczak, J., Loukas, A., Makowski, E. K., Mohammadi-Peyhani, H., Mahajan, S. P., . . . , **Goodman, A.**, et al. (2025). Property Enhancer: a data-efficient multi-objective approach for functional antibody optimization. *bioRxiv* (preprint).

- Moshkov, N., Bornholdt, M., Benoit, S., Smith, M., McQuin, C., **Goodman, A.**, Senft, R. A., Han, Y., Babadi, M., Horvath, P., et al. (2024). Learning representations for image-based profiling of perturbations. *Nature Communications*, 15(1), 1594.
- Frey, N. C., Joren, T., Ismail, A. A., **Goodman, A.**, Bonneau, R., Cho, K., & Gligorijević, V. (2024). Cramming protein language model training in 24 GPU hours. *bioRxiv* (preprint).
- Lee, S., Abini-Agbomson, S., Perry, D. S., **Goodman, A.**, Rao, B., Huang, M. Y., Diedrich, J. K., Moresco, J. J., Yates III, J. R., Armache, K.-J., et al. (2023). Intrinsic mesoscale properties of a Polycomb protein underpin heterochromatin fidelity. *Nature Structural & Molecular Biology*, 30(7), 891–901.
- Stirling, D. R., Swain-Bowden, M. J., Lucas, A. M., Carpenter, A. E., Cimini, B. A., & **Goodman, A.** (2021). CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics*, 22(1), 433.
- Doan, M., Barnes, C., McQuin, C., Caicedo, J. C., **Goodman, A.**, Carpenter, A. E., & Rees, P. (2021). Deepometry, a framework for applying supervised and weakly supervised deep learning to imaging cytometry. *Nature Protocols*, 16(7), 3572–3595.
- Doan, M., Sebastian, J. A., Caicedo, J. C., Siegert, S., Roch, A., Turner, T. R., Mykhailova, O., Pinto, R. N., McQuin, C., **Goodman, A.**, et al. (2020). Objective assessment of stored blood quality by deep learning. *Proceedings of the National Academy of Sciences*, 117(35), 21381–21390.
- Hung, J., **Goodman, A.**, Ravel, D., Lopes, S. C. P., Rangel, G. W., Nery, O. A., Malleret, B., Nosten, F., Lacerda, M. V. G., Ferreira, M. U., et al. (2020). Keras R-CNN: library for cell detection in biological images using deep neural networks. *BMC Bioinformatics*, 21(1), 300.
- Hollandi, R., Szkalitsky, A., Toth, T., Tasnadi, E., Molnar, C., Mathe, B., . . . , **Goodman, A.**, et al. (2020). nucleAIzer: a parameter-free deep learning framework for nucleus segmentation using image style transfer. *Cell Systems*, 10(5), 453–458.
- Doan, M., Case, M., Mašić, D., Hennig, H., McQuin, C., Caicedo, J., Singh, S., **Goodman, A.**, Wolkenhauer, O., Summers, H. D., et al. (2020). Label-free leukemia monitoring by computer vision. *Cytometry Part A*, 97(4), 407–414.
- Caicedo, J. C., **Goodman, A.**, Karhohs, K. W., Cimini, B. A., Ackerman, J., Haghighi, M., Heng, C., Becker, T., Doan, M., McQuin, C., et al. (2019). Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nature Methods*, 16(12), 1247–1253.
- Caicedo, J. C., Roth, J., **Goodman, A.**, Becker, T., Karhohs, K. W., Broisin, M., Molnar, C., McQuin, C., Singh, S., Theis, F. J., et al. (2019). Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. *Cytometry Part A*, 95(9), 952–965.
- Kulikov, V., Guo, S.-M., Stone, M., **Goodman, A.**, Carpenter, A., Bathe, M., & Lempitsky, V. (2019). DoGNet: a deep architecture for synapse detection in multiplexed fluorescence images. *PLoS Computational Biology*, 15(5), e1007012.
- Rueden, C. T., Ackerman, J., Arena, E. T., Eglinger, J., Cimini, B. A., **Goodman, A.**, Carpenter, A. E., & Eliceiri, K. W. (2019). Scientific Community Image Forum: a discussion forum for scientific image software. *PLoS Biology*, 17(6), e3000340.
- McQuin, C., **Goodman, A.**, Chernyshev, V., Kamentsky, L., Cimini, B. A., Karhohs, K. W., Doan, M., Ding, L., Rafelski, S. M., Thirstrup, D., et al. (2018). CellProfiler 3.0: next-generation image processing for biology. *PLoS Biology*, 16(7), e2005970.
- Caicedo, J. C., McQuin, C., **Goodman, A.**, Singh, S., & Carpenter, A. E. (2018). Weakly supervised learning of single-cell feature embeddings. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 9309–9318.