

AI Compiler Engineer

at Modular

United States / Canada

About Modular

We believe that AI is a net positive force in the world. Our vision and mission are to help rebuild AI infrastructure to advance humanity and our environment. We will do whatever it takes to empower our customers, team, and company to benefit from that pursuit. You can read about [our culture](#) and [careers here](#) to understand how we work and what we value.

We are owners and advocates for the underlying technologies, developer platforms, product components, and infrastructure. These essential building blocks form the high-quality and coherent experiences our users expect. We aim to drive the pace of innovation for every AI/ML developer.

About the role:

Modular is looking for AI Compiler engineers to work on multi-framework, multi-hardware and user-extensible AI compiler infrastructure. We are looking for candidates with strong communication and teamwork skills who excel in collaborating across team boundaries. In this role you will own the design and implementation of compiler types, operations and passes related to machine learning models, and work across teams to integrate them with the rest of Modular's product.

LOCATION: Candidates based in the US or Canada are welcome to apply. You can work out of one of our offices in Seattle, WA, or remotely from home.

What you will do:

- Build an MLIR-based machine learning compiler with scalable and high quality infrastructure
- Implement generic and extensible optimizations for machine learning models from multiple frameworks and on multiple devices
- Work with other teams to provide a balanced stack that fully utilizes today's complex server and mobile systems.
- Collaborate with machine learning researchers and engineers to guide compiler development for future ML trends.

What you bring to the table:

- 2+ years of compiler engineering experience.
- Experience working with compilers for machine learning, such as XLA, Glow, TVM, nGraph, TensorRT, IREE, etc.
- Experience with machine learning graph optimizations
- Creativity and curiosity for solving complex problems, a team-oriented attitude that enables you to work well with others, and alignment with our culture.
- Strong knowledge of core compiler algorithms and data structures.
- In-depth knowledge of C++, as well as, knowledge of basic GitHub workflows like pull requests.

Helpful but not required:

- Strong knowledge of and experience working with MLIR and LLVM.
- Experience with 8 bit and lower model quantization
- Advanced degree in Computer Science or a related area

The estimated base salary range for this role to be performed in the US, regardless of the state, is \$157,500.00 - \$220,000.00 USD. The salary for the successful applicant will depend on a variety of permissible, non-discriminatory job-related factors, which include but are not limited to education, training, work experience, business needs, or market demands. This range may be modified in the future. The total compensation for a candidate will also include annual target bonus, equity, and benefits, with equity making up a significant portion of your total compensation.

For candidates who fall outside of the listed requirements, we nevertheless encourage you to apply as we may have openings that are lower/higher level than the ones advertised.

What Modular brings to the table:

- **Amazing Team.** We are a progressive and agile team with some of the industry's best

engineering and product leaders.

- **Flexible Location.** We want you to work where you are most productive and we are a remote-first company. We have offices in Palo Alto, CA and Seattle, WA for collaboration, and will open more next year.
- **World-class Benefits.** In order to attract the best, we need to offer the best. Premier insurance plans, up to 5% 401k matching, very generous WFH setup stipends, flexible paid time off, and more are available to you! *Please note that specific benefit packages may vary based on your location.*
- **Competitive Compensation.** We offer very strong compensation packages, including stock options. We want people to be focused on their best work and believe in tailoring compensation plans to meet the needs of our workforce.
- **Team Building Events.** We organize regular team onsite and local meetups in different cities.

Working at Modular will enable you to grow quickly as you work alongside incredibly motivated and talented people who have high standards, possess a growth mindset, and a purpose to truly change the world.

Modular is proud to emphasize an equal opportunity, safe environment for people to do their best work. Modular is an affirmative action employer. We are committed to providing equal employment opportunities regardless of race, color, ancestry, religion, sex, national origin, sexual orientation, age, citizenship, marital status, disability, gender identity, or Veteran status. We also consider qualified applicants regardless of criminal histories, consistent with legal requirements.

This employer participates in E-Verify and will provide the federal government with your Form I-9 information to confirm that you are authorized to work in the U.S. If E-Verify cannot confirm that you are authorized to work, this employer is required to give you written instructions and an opportunity to contact Department of Homeland Security (DHS) or Social Security Administration (SSA) so you can begin to resolve the issue before the employer can take any action against you, including terminating your employment. Employers can only use E-Verify once you have accepted a job offer and completed the Form I-9.

Apply for this Job

* Required

Apply With LinkedIn

First Name *

Last Name *

Email *

Phone

Resume/CV * [Attach, Dropbox, or enter manually](#)
(File types: pdf, doc, docx, txt, rtf)

Cover Letter [Attach, Dropbox, or enter manually](#)
(File types: pdf, doc, docx, txt, rtf)

LinkedIn Profile

Website

Where are you currently residing? (City and Country) *

Which countries are you currently legally authorized to work in? *