

Scientific Data Engineer

Job ID
REQ-10081250
июл 06, 2026
США
Available in: English

Сводка

#LI-Hybrid
Location: Cambridge, MA

Job Description Summary

The Oncology Data Science (OncDS) team in Biomedical Research provides computational biology, AI, and data expertise and brings together projects spanning the entire pre-clinical to clinical development pipeline across a wide and unique range of novel therapeutics. OncDS specializes in using high-throughput genomic and biomarker data for target identification, drug discovery, and clinical development.

We are seeking a motivated and curious Scientific Data Engineer at the beginning of their professional data journey to support the development, maintenance, and modernization of FAIR oncology reference datasets. This role combines hands-on data curation with workflow automation, AI-enabled process improvement, dataset refresh and quality control, and fit-for-purpose data engineering to help ensure that key oncology data assets are reliable, reproducible, and ready for downstream scientific use.

This position will work closely with senior team members and cross-functional collaborators to modernize data curation processes, identify opportunities for AI-enabled automation, and contribute to high-value reference datasets that support Oncology-wide data strategy and FAIR data goals. The role is designed to build practical experience in scalable data practices, operational excellence, and automation-driven continuous improvement.

Why join us: This is an opportunity to work at the intersection of data engineering, computational biology, and oncology research, contributing to data assets that enable important scientific decisions across the drug discovery pipeline. You will join a collaborative team that values curiosity, learning, and continuous improvement, and you will have the chance to grow your technical and scientific skills while helping build reliable, scalable data foundations for cutting-edge oncology research.

About the Role

Key responsibilities:

- Work with senior team members to identify opportunities for simplification, automation, and continuous improvement in operational data processes, including modernization of data curation workflows
- Support the development and improvement of automated workflows for data processing, quality control, recurring dataset updates, and AI-enabled process improvement
- Collaborate with cross-functional teams to curate, maintain, and improve key oncology reference datasets, with an emphasis on quality, correctness, reproducibility, and operational reliability
- Contribute to reference dataset lifecycle management, including documentation, versioning, and traceable update processes
- Assess dataset structure, integrity, and fitness for downstream scientific use, including attention to identifiers, primary keys, missing values, consistency, and related data quality issues, and help resolve data issues in collaboration with relevant partners
- Contribute to best practices for reproducible workflows, data handling, and scalable dataset operations
- Help keep oncology reference datasets current, usable, and responsive to evolving scientific and organizational priorities

Essential Requirements:

- Degree in bioinformatics, computational biology, data science, computer science, or a related field, or equivalent relevant experience
- Experience working with biological, experimental, or omics datasets in a research environment
- Proficiency in one or more programming languages used in data science or bioinformatics, such as Python or R
- Experience with data wrangling, data quality control, and reproducible analysis workflows
- Familiarity with workflow automation and scripting for recurring data processing tasks
- Familiarity with database principles and sound data modeling practices, including normalization, primary keys, joins, and thoughtful handling of missing values and data integrity issues
- Knowledge of version control and Unix / Linux-based working environments
- Strong organizational skills and attention to detail, with a demonstrated interest in operational excellence
- Ability to learn quickly, manage multiple tasks, and work effectively in a collaborative environment
- Excellent written and verbal communication skills, and the ability to work effectively across technical and scientific teams

Desirable Requirements:

- Experience with oncology datasets, translational research data, or biomedical reference data resources
- Familiarity with FAIR data principles, metadata curation, or scientific data stewardship
- Experience maintaining reference datasets or supporting data products used by multiple stakeholders
- Experience collaborating with experimental scientists or lab-based partners

Novartis Compensation and Benefit Summary:

The salary for this position is expected to range between \$103,600 and \$192,400 per year.

The final salary offered is determined based on factors like, but not limited to, relevant skills and experience, and upon joining Novartis will be reviewed periodically. Novartis may change the published salary range based on company and market factors.

Your compensation will include a performance-based cash incentive and, depending on the level of the role, eligibility to be considered for annual equity awards.

US-based eligible employees will receive a comprehensive benefits package that includes health, life and disability benefits, a 401(k) with company contribution and match, and a variety of other benefits. In addition, employees are eligible for a generous time off package including vacation, personal days, holidays and other leaves.

Why Novartis: Helping people with disease and their families takes more than innovative science. It takes a community of smart, passionate people like you. Collaborating, supporting and inspiring each other. Combining to achieve breakthroughs that change patients' lives. Ready to create a brighter future together?
<https://www.novartis.com/about/strategy/people-and-culture>

Benefits and Rewards: Learn about all the ways we'll help you thrive personally and professionally.
[Read our handbook \(PDF 30 MB\)](#)

EEO Statement:

The Novartis Group of Companies are Equal Opportunity Employers. We do not discriminate in recruitment, hiring, training, promotion or other employment practices for reasons of race, color, religion, sex, national origin, age, sexual orientation, gender identity or expression, marital or veteran status, disability, or any other legally protected status.

Accessibility & Reasonable Accommodations

The Novartis Group of Companies are committed to working with and providing reasonable accommodation to individuals with disabilities. If, because of a medical condition or disability, you need a reasonable accommodation for any part of the application process, or to perform the essential functions of a position, please send an e-mail to us_reasonableaccommodations@novartis.com or call +1(877)395-2339 and let us know the nature of your request and your contact information. Please include the job requisition number in your message.

Дивизион
Biomedical Research
Business Unit
Research
Место
США
Состояние
Massachusetts
Сайт
Cambridge (USA)
Company / Legal Entity
U175 (FCRS = US175) Novartis Institutes for BioMedical Research, Inc.
Functional Area
Data and Digital
Job Type
Full time
Employment Type
Regular
Shift Work
No

Job ID
REQ-10081250

Scientific Data Engineer

[Apply to Job](#)

Job ID
REQ-10081250

Scientific Data Engineer

[Apply to Job](#)

Source URL: <https://novartis.ru/careers/career-search/job/details/req-10081250-scientific-data-engineer>

List of links present in page

1. <https://www.novartis.com/about/strategy/people-and-culture>
2. https://www.novartis.com/sites/novartis_com/files/novartis-life-handbook.pdf
3. <mailto:us.reasonableaccommodations@novartis.com>
4. https://novartis.wd3.myworkdayjobs.com/en-US/Novartis_Careers/job/Cambridge-USA/Scientific-Data-Engineer_REQ-10081250-1
5. https://novartis.wd3.myworkdayjobs.com/en-US/Novartis_Careers/job/Cambridge-USA/Scientific-Data-Engineer_REQ-10081250-1