**[Data Science Initiative](https://datascience.uoregon.edu/)**

The Data Science Initiative is a hub of collaboration and resources for researchers in every discipline who work with large datasets. The initiative includes topic areas such as environmental big data, business analytics, biomedical data science, and data science of social interactions and social impacts. The core of the program includes data collection, database creation and curation, and statistical and computational analysis.

The UO now has an undergraduate degree in Data Science, and nearly 100 faculty who are doing work in data science, from departments of psychology, music, physics, biology, school of business, college of education, digital arts, English, linguistics, and many more. Examples include:

* In *biology*, one faculty uses computational and mathematical approaches to leverage genome sequence data for evolutionary inference.
* In *computer and information sciences*, one researcher draws on the subfields of AI such as multi-agent systems, game theory, machine learning, and optimization, as well as fields outside of AI such as cognitive modeling and conservation biology to create solutions to issues of public safety, cyber security, sustainability, and public health.
* In *economics*, one researcher is focused on using remote sensing and other big data products to address environmental policy questions.
* In *earth sciences*, one faculty uses real-world seismic data to his research to reveal the physics of earthquakes and other seismic shifts. Analyzing data helps the research team understand the impacts of future earthquakes, and helps them identify potential ruptures in specific areas around the globe.
* In the *education*, computational approaches to large-scale educational research, including national and geographic trends and variations in achievement gaps, as well as community-level features that relate to this variation, is an area of focus.
* In *physics*, one focus of research is on applying Bayesian forward-modeling techniques to astronomical data sets.
* In *philosophy*, a focus of research mobilizes concepts from political philosophy and historiography to explore the political dimensions of ways in which data is increasingly defining our selfhood—such as how the emergence of mass-scale data systems that store social security numbers and birth certificates, data techniques for categorizing personality traits and measuring intelligence, and the role of data in the shadow of racial inequality.
* UO and [Oregon Health & Science University](https://news.ohsu.edu/2019/06/13/ohsu-uo-join-forces-to-combat-cancer-with-data-science) also are working together to create a joint center in biomedical data science that would empower researchers at both institutions to attack cancer with big data. The center initially involves as many as 20 researchers and their teams.

**Resources and Tools:**

UO has powerful resources for researchers in data science, including one of the fastest academic supercomputers in the Northwest, known as Talapas. The high-performance computing center is able to perform more than 250 trillion calculations per second. Researchers at the University of Oregon can access Talapas, and other resources available through Research Advanced Computing Services, for everything from statistical studies to genomic assemblies to quantum chemistry.

Data Services, a program of UO Libraries, can help researchers organize, manage, and curate their data, including creating a data management plan, finding the right kinds of data for their research, and software training and assistance.

The Oregon Data Science Collaborative is an NSF-funded collaborative institute between Oregon State University, the University of Oregon, and Portland State University. The Collaborative’s goal is to advance research across disciplines by facilitating collaboration between researchers and data scientists and by building a community of practice to disseminate familiarity, expertise, and training in data science. They offer data science workshops and research consulting for researchers across Oregon and the wider scientific community.

UO/OHSU Joint Center for Biomedical Data Science (CBDS). Funded through two $10M lead gifts ($20M total), the University of Oregon and Oregon Health & Science University have joined forces to create a new joint center in biomedical data science, empowering researchers at both institutions to attack cancer with big data. The partnership combines efforts at the UO’s Phil and Penny Knight Campus for Accelerating Scientific Impact with those at the OHSU Knight Cancer Institute to detect and fight deadly forms of cancer and other common diseases. The center is facilitating applied, translation, and clinical research, developing innovative new approaches to quickly analyze large sets of data, allowing researchers to “listen in” on cell development for early detection of lethal diseases. The CBDS leverages UO’s strengths in biological sciences with OHSU’s efforts in precision medicine. Jointly, the center will initially involve as many as 20 researchers and their teams and bolster opportunities for students and researchers by increasing clinical research opportunities. UO is in the processing of hiring several additional faculty and is now recruiting top researchers. Additionally, the Presidential Initiative in Data Science collaborated on tenure-track faculty searches in Math and CIS for faculty to teach in the new data science undergraduate degree and to be research members of CBDS. CAS is currently in final negotiations with a world expert in the analysis of genomic data. The CBDS is poised to rapidly expand research and education activities in the areas of computational genomics and data science. A 3,000 sq ft cutting edge collaborative space is being designed as the home for the CBDS in the second building of the Knight Campus, bringing in close proximity computational and experimental scientists who will work together to tackle cancer genomic challenges.

Cloud Computing. The next phase of research computing at UO aims to provide a user experience where students and researchers can seamlessly utilize infrastructure hosted locally, in the cloud, or some combination of both. This includes the ability to ‘burst’ to the public cloud to support workloads with varying demand patterns or spikes in demand.

The UO sees the future of academic computing as increasingly leveraging the public cloud (e.g., Azure, AWS, Google Cloud). Over this past year, UO partnered with Data Science Faculty to develop scalable, cloud-native, curated computational environments for data science courses; these easily deployable, containerized environments make use of modern, open-source tools ubiquitous in the professional world to prepare students for their future in the academe or industry. The UO continues to invest in cloud computing including new cloud positions, enhancing cloud skillsets, providing training workshops and educational resources for the campus community, and maximizing vendor partnerships.

Consolidated paragraph:

The Data Science Initiative was launched in 2018 with the goal of developing significant new research and education in data science across the University through a methodological core for the implementation of data science approaches such as data collection, database creation and curation, as well as statistical and computational analysis. In addition, several aspects of design, communication, ethics and law are components of the core. Domain spokes are areas where the research and educational activities focused primarily on a particular field or problem. To date, the UO has hired over 20 faculty, with computational genomics being one of the primary areas. A particularly important development is of a data science undergraduate program that has a ‘data science + domain’ structure. Students study core quantitative methods and apply those methods to a chosen area of domain emphasis. There are presently a dozen domain areas that cover most schools and colleges on campus, with biology being one of the largest. In addition to engaging new students in data science training, DSI supports existing and novel research at UO, acting as a university-wide research institute to help develop new research connectivity and impact. Several Research Professors in Data Science act as catalysts to help accelerate research by bringing data science tools and approaches to bear on the problems being tackled. By building connections with sister institutions across Oregon, Pacific Rim research universities, and government and industry, DSI helps to garner more resources and secure infrastructure to advance data science research and practice across several domain areas.