Health Data Warehouse (HDW)

Contact
Director: Prof. Bilal Khan, PhD
Email: inhdw@lehigh.edu
Address: Health Science and Technology Building, Room 160

Core Faculty and Research Staff

Bilal Khan, Director (Biostatistics and Health Data Science, Computer Science)

Chad Meyerhoefer. Associate Director (Economics)

Mohamed Saad (Senior Research Scientist)

Affiliated Faculty

- John Hughes (Biostatistics and Health Data Science, Mathematics)
- Bridget Dever (Educational Psychology)
- Tom McAndrew (Biostatistics and Health Data Science)
- Lucy Napper (Psychology)
- Michael Rivera (Business Information Systems)
- Jong Shin (Population Health)
- Gabrielle String (Population Health, Civil and Environmental Engineering)

Affiliated Graduate Research Assistants

- Gavin Fox (Ph.D. program in Psychology)
- Ojaswi Joshi (M.Sc. program in Data Science)
- Jess Lathrop (Ph.D. program in Educational Psychology)
- Nathaniel Rhodes (Ph.D. program in Economics)

Mission

The Health Data Warehouse (HDW) is a jointly funded service center of the College of Health and the College of Business at Lehigh University. It serves as a secure, high-performance computing environment that supports health data research, AI-driven analytics, and software innovation.

Beyond providing a highly secure, air-gapped data storage environment, HDW enables faculty, researchers, and collaborators to:

- Store and compute data securely in an ultra-secure, HIPAA-compliant facility designed for sensitive research.
- Utilize SALUS, a curated archive of public health-related datasets for education and research.
- Leverage DANTZIG, an AI-powered “live” data warehousing system that continuously uncovers new hypotheses and patterns.
- Deploy ODIN, a real-time behavioral research platform supporting ecological momentary assessment (EMA) and just-in-time interventions.
- Access machine learning and AI consulting services, assisting research teams in applying advanced data-driven analytics.
- Develop custom software applications, including forecasting tools, interactive data dashboards, and AI-powered research platforms.

By integrating secure health data storage, cutting-edge AI analytics, and full-stack software solutions, HDW accelerates high-impact research, enhances funding competitiveness, and fosters industry-academic partnerships.

As a strategic investment by Lehigh University, HDW generates return on investment (ROI) by:

- Attracting external funding and research grants.
- Facilitating partnerships with industry, government, and healthcare organizations.
- Enhancing education and workforce development in health data science and AI.
- Translating research into real-world healthcare and business applications.

With data security as a foundation and innovation as a mission, HDW enables faculty and research partners to push the boundaries of health, business, and technology.

RESEARCH ACTIVITIES

The Health Data Warehouse (HDW) supports a wide range of interdisciplinary research projects that utilize secure data storage, AI-driven analytics, and custom software development. Faculty and researchers across Lehigh University use HDW’s infrastructure to analyze restricted datasets, develop predictive models, and create AI-driven tools for health and business applications.

Secure Data Storage and Analysis

HDW provides high-security computing environments for faculty handling restricted or sensitive datasets. HDW also offers in-house expertise in machine learning modeling and biostatistics which can be leveraged by HDW affiliates.

- Chad Meyerhoefer (Economics) and his student Nate Rhodes use HDW to house and analyze data from the National Center for Education Statistics (NCES).
- Bridget Dever (Educational Psychology) and her student Jess Lathrop use HDW to manage and analyze datasets from The Inter-university Consortium for Political and Social Research.

AI-Powered Behavioral and Public Health Research

ODIN, HDW’s real-time behavioral research platform, is supporting faculty conducting ecological momentary assessment (EMA) and just-in-time interventions:

- Jong Shin (Population Health) is preparing to use ODIN for a longitudinal study on sleep quality.
- Lucy Napper (Psychology) and her student Gavin Fox are using ODIN to study alcohol consumption among college students, linking behavioral patterns with environmental and physiological markers.

Data-Driven Business and Health Research

HDW’s DANTZIG platform provides live AI-powered data mining for faculty conducting research in public health and business analytics:

- Gabrielle String is using DANTZIG to analyze WASH (Water, Sanitation, and Hygiene) data from the Ivory Coast, uncovering patterns in sanitation access and public health outcomes.
- Michael J. Rivera (Business Information Systems) is using DANTZIG for data mining and hypothesis generation in business and economic forecasting.

Custom Software Development for Research

HDW provides full-stack software development services for interactive data platforms and predictive modeling tools:

- Tom McAndrew (College of Health) is collaborating with HDW to build a crowdsourced forecasting system that aggregates expert predictions for influenza epidemic trajectory modeling.

By supporting a diverse portfolio of research activities, HDW enhances Lehigh’s capacity for data-driven discovery, providing faculty with infrastructure, expertise, and computing power to conduct cutting-edge research.

FUNDING

The Health Data Warehouse (HDW) strengthens Lehigh University’s ability to secure external research funding by providing a scalable, secure, and compliant computing environment that enhances faculty grant applications.

HDW’s infrastructure has supported research funded by federal, state, and private agencies, including projects using restricted datasets from the National Center for Education Statistics (NCES) and The Inter-university Consortium for Political and Social Research.

HDW fosters partnerships with industry, healthcare organizations, and government agencies, offering trusted infrastructure for collaborative research, sponsored projects, and interdisciplinary initiatives. To sustain growth, HDW is exploring grant opportunities, institutional

partnerships, and cost-recovery models to expand its capabilities while maximizing return on investment.

EDUCATIONAL OPPORTUNITIES

HDW provides transformative educational experiences for students and faculty engaged in health data science, artificial intelligence, and computational research.

Through SALUS, HDW maintains a curated archive of over 3,000 public datasets, supporting:

- Coursework integration, allowing faculty to use real-world data in big data analytics, predictive modeling, and machine learning instruction.
- Independent study projects, giving students hands-on experience with data science, AI, and computational modeling.
- Research assistantships, where students work directly with faculty using secure computing environments, advanced analytics, and research software development.

The Health Data Warehouse (HDW) serves as a hub for interdisciplinary training and workforce development, equipping students with in-demand technical skills needed for careers in health informatics, biomedical data science, business analytics, and AI-driven research.