

Ben Bubnick

ADVISORY DATA SCIENTIST · DATA INTEGRATION SPECIALIST

3419 Superior Park Drive, Cleveland Heights, Ohio, 44118, U.S.

☎ (+1) 216-556-1612 | ✉ ben.bubnick@gmail.com | 🌐 benbubnick

"I want to use my 15+ years experience to reimagine healthcare by creating, assembling, and improving our business relationships through data science. I believe that when I inspire people to do the things that inspire them, we can improve the relationship between caregiver and patient."

Summary

Professional Summary	AI & Machine Learning, Generalist/Utility Roles, Data Curation/Wrangling, Data Analytics Training, Scientific Analysis
Big Data Platform	Cloud infrastructure using Hadoop & Azure, SSIS/SQL/T-SQL/PostgreSQL, NoSQL, DB2, IMS, Hive/Impala/Pig/kafka
Health Care Analytics	Payer/Provider data structures: EMR, HL7(C-CDA, CCD, FHIR), 837, CMS & Flat File Architecture: XML, JSON, CSV, etc
Demonstrated Past Success	IBM Band Promotion, Subject Matter Expert, Lab Services Award Winner, Eminence and Excellence Award Winner

- Add some sort of skill list
- Another skill
- Third skill
- A fourth skill
- And a final skill

Experience

Data Scientist

Merative

PROVIDER ANALYTICS

Jul. 2022 - Present

Leads end to end delivery of, large & complex AI and Cognitive led transformation programs and works with Merative and Partners to bring the best tools and delivery methods. Identify, analyze, and extract key data from heterogeneous systems; and lead technical requirements gathering, data discovery, documentation, and implementation of assigned client integrations. Manages client resources and drive technical solutions to timely completion. Works closely with the delivery teams to develop Cognitive, Automation & AI led programs. Lead the technical delivery of assigned client integrations which has included the ordering, installation, and configuration of required hardware, including VPN devices to support client integrations. Configuration and development of software necessary to support client integrations, including ETL and Data Validation.

Advisory Data Scientist

IBM

WATSON HEALTH

Nov. 2020 - Jul. 2022

Lead member of Provider Analytics for data integration of disparate and centralized data into data marts, data mining, statistical model development, advanced analytic methods. Worked closely with senior leaders and colleagues to identify opportunities, set objectives, and formulate analytic strategies that result in measurable impact. Technical lead on diverse data sources, such as electronic health record data, insurance claims data, scheduling and financial data, social determinants of health data, and any other relevant sources.

- HDP Data Migration
- Explorys Decomm
- GP Support for CCF
- Jira Migration Validation
- GP Migration
- DS Certification REACH
- Surgical Optimization development
- IHM Validation

Senior Curation Data Architect

IBM

WATSON HEALTH

May. 2020 - Nov. 2020

Senior member of Data Onboarding & Governance for integration of healthcare data and enrichment of metadata. Configuration of data models, analytics, reports, and visualizations. SME to project data architects, delivery, analytics, integration, etc. Directed the configuration of the data onboarding platform and associated infrastructure, and developed improvements to advance stakeholder's business priorities using tailored analytics.

- Led cross-function technical team as clinical data analyst for 2020/2021 measure updates project
- Created machine learning Covid-19 tool that reduced need for hospitalization identification time by up to two weeks
- Created AI Surgical Scheduling Optimization tool currently evolving into new offering
- Developed large clinical data integration in excess of several 100 billion structured and semi-structured records for life sciences analytics projects
- Led HIPAA patient deidentification project from more than 100 data sources for content analytics and measures
- Served in an advisory capacity for new data platform development project
- Closed documentation gaps, organizing more than 100 past projects

AI Data Curation Architect

IBM

WATSON HEALTH

Jan. 2018 - May. 2020

Lead member of Data Onboarding & Governance, member of the Watson Foundation for Health Solutions Architecture team. Remediation of data issues during UAT. Built the data onboarding platform and associated infrastructure, and developed improvements to reducing customer implementation cycle time and data triage activities. Reported on the “health of the data” during data onboarding. Tuned machine learning classification algorithms for patient matching, and led the internal review of similar work by my peers.

- Reduced mapping rework by 45% with through new knowledge-driven initiatives
- Developed new data validation and review process that reduced total rework by 25%
- Managed and developed ambitious multi-org enterprise with over 70 interconnected, project-sized components
- Earned AI certification to peer review tuning algorithm and reviewed over 50% of projects
- Curated 25 pages of AI workflow documentation and mentored new hires in AI tuning process
- Reduced time spent on knowledge transfer by ≈80% by reorganizing project documentation
- Spearheaded new-hire mentoring program that led to ≈4,500 increase in total billable hours
- Awarded Lab Services Award for coordinating internal project teams to deliver client needs on time

Data Scientist

Explorys, an IBM Company

EPM APPLICATION SUITE

Jun. 2015 - May. 2018

Data Ingestions, pre-analytic enrichments, analytic augmentation, and data governance. Identified and extracted key data from multiple heterogeneous systems; and led technical requirements gathering, data discovery, documentation, and implementation of assigned client integrations. Ordered and oversaw installation of & configuration of required hardware, including VPN devices to support client integrations. Configuration and development of software necessary to support client integrations.

- Principal integration specialist on over 60 projects for 20 different clients
- Eliminated need for a \$30,000 annual contract by developing data curation tool
- Chosen as team lead for ≈12 remote data science contractors for 36 integration projects throughout 2015 - 2017
- Created over 125 new knowledge base articles, making up 30% of the team's online content
- Eliminated average of 3 weeks from employee onboarding process
- Reduced cost-per-hire by \$5,210 by re-developing education program
- Awarded Eminence and Excellence Award in 2017 for analytics project

Software Engineer

AmTrust Financial Services

COMMERCIAL PACKAGE POLICIES

Mar. 2014 - Jun. 2015

Developed web forms for insurance industry using ISO standards and created multi-tier MVC components, Web APIs, and UIs using VB, ASP, Javascript, & HTML5

- Development on two software architecture projects in 2015 and principal developer on offering product in 2014
- Developed over 65 ISO rating based web forms for insurance industry
- Steered development group into a unit testing project on multi-tier MVC components in 2014

Color Scientist

PPG

AUTOMOTIVE REFINISH

Apr. 2012 - Jun. 2013

Developed software and radiative transfer models on complex paint mixtures to determine formulation methods and for spectroscopic profile analysis on color match samples.

- Improved color matching classification algorithm performance by 40% through new numerical methods
- Developed cross-instrument correlation software for risk analysis project
- Awarded Color Lab Automation Award in 2012 for expanding data collection, analyzing automotive pigments and applicable color formulation

Adjunct Professor

various colleges

LORAIN COUNTY COMMUNITY COLLEGE, MIAMI UNIVERSITY, CINCINNATI STATE TECHNICAL AND COMMUNITY COLLEGE

Aug. 2010 - May. 2019

Developed curricula in introductory physics & astronomy courses, and technical presentations of complex physical concepts for major, non-major, and remedial students. Wrote concise laboratory instruction manuals, trained in precise measurement techniques & data analysis using DataStudio, and coordinated analysis documentation.

- Exceeded feedback evaluation goals by an average of 16% in 2017 and by 17% in 2018
- Tracked a 28% reduction in student confusion by reorganizing online education materials
- Uploaded over 700 hours of lecture material for convenient consumption

Research Assistant

University of Cincinnati

DEPARTMENT OF PHYSICS

Sep. 2006 - Sep. 2010

Gathered data using SpeX, a 0.8–5.4 μm spectrograph, at the NASA Infrared Telescope Facility. Observations in the short-wavelength, crossdispersed mode SXD. Multiple analysis projects comparing 2MASS with measured data, statistical classification without independent photometry for cluster properties, and photometric analysis bootstrapped to independent measurements.

- Volunteered 120 hours of Quality Control for the Cambridge Astronomical Survey Unit Variable Stars database
- Published 3 scholarly articles on near-infrared astrophysics over a 24-month period
- Awarded Mary J. Hanna Research Fellowship in 2010

Writing

Surgical scheduling optimization and procedure duration prediction

IBM

TECHNICAL DEVELOPMENT TEAM - TOTAL PROCEDURE TIME ESTIMATION

Dec. 2020

Improvement on traditional fixed ratio methods for total procedure time estimation, using linear regression models based on relevant variables.

Process and Requirements for CCD Data Integration

IBM

LEAD INTEGRATION DEVELOPER

Oct. 2017

Technical white paper on processes and requirements, overcoming the many difficulties with flat file data curation of CCDs.

AI Paint Formula Prediction from Spectroscopic Measurements

Johns Hopkins Bloomberg School of

Public Health

STUDENT DEVELOPER

May. 2015

Naive bayesian color and effects matching algorithm for the formula of complex automotive paint formula.

Characterization of Surface Coatings using Radiative Transfer Models for Color Matching

PPG

RADIATIVE TRANSFER PHYSICS

Dec. 2010

Approximation of Chandrasekhar radiative transfer model for atmospheric scattering is applied to color matching. (U.S Patent Application)

Massive Stellar Clusters in the Disk of the Milky Way Galaxy

OhioLINK ETD

GRADUATE RESEARCH ASSISTANT

Dec. 2010

Using spectroscopic and photometric analysis of two open clusters to determine the structure of the plane of the Milky Way Galaxy.

Near-Infrared Spectroscopy of Candidates Members of the Galactic Cluster [BDS2003] 107

PASP

UNDERGRADUATE RESEARCH ASSISTANT

Feb. 2008

New near-infrared classification spectra for nine candidate members and comparative 2MASS photometry for several cluster members.

Talks

IBM Watson Care Insights Reach Initiative

Watson Health

TRAINER

Apr. - Sep. 2019

Watson Care Insights aims to analyze longitudinal patient records and real-time data points through HL7 feeds to provide practitioners insights into the patient's data, and address these challenges by integrating within the physician's workflow in the electronic health record system.

Content & Data Analytics Learning Exchange

Watson Health

PEER PRESENTER

Apr. & May 2018

Introduction of implementation of machine learning algorithms and their applications using real-world health data analytics examples.

2018 IRI Fall Networks Conference

Innovation Research Interchange

PRESENTER FOR THE INFORMATION SERVICES/INFORMATION TECHNOLOGY NETWORK TALKS

Sep. 2018

Cognitive computing can use NLP to extract insights from the unstructured data that holds most of the relevant information, but also it can also perform the necessary transformation accurately at the earliest stages of the data integration pipeline.

Cross Team Training Series

Explorys

LEAD PRESENTER FOR THE METL GROUP

Sep. 2016

Seven part series in data ingestion approach using the Hadoop software framework. Data can be aggregated much more quickly and cost-effectively using a data lake framework than in traditional data warehouses, because of the speed and low-cost of massively parallel computing.

Awards

2022 **Gold Champion**, Awarded four consecutive years for 1,477 classroom hours and 43 digital credentials

IBM

2020 **GDP Bonus**, Improving revenue per SaaS client by 3.77%

IBM

2019 **GDP Bonus**, Improving revenue per SaaS client by 1.34%

IBM

2018 **Lab Services Award**, Coordinating Internal Teams to deliver client needs on time

IBM

2018 **GDP Bonus**, Improving revenue per SaaS client by 1.17%

IBM

2017 **Eminence and Excellence Award**, Analytics project

Explorys

2012 **Color Lab Automation Award**, Something about it

PPG

Volunteer

After School Computer Science Supplemental Education

CSforCLE/CMSD

INSTRUCTOR

Oct. 2018 - May. 2019

With CMSD, Cuyahoga County, Cleveland State University, and Envision Excellence in STEM Education, the Cleveland Foundation established the CSforAll initiative. The goal is to provide computer science training in all of the District's high schools, with a particular focus on kids who do not have access to a computer or a dependable Internet connection at home.

Data Science High School Competition

HIT in the CLE

COACH

Mar. 2017 - Mar. 2019

The High School Data Science Competition assigns an interesting problem to five student teams to tackle using a large data set. The teams must analyze the data in depth with the help of industry mentors and then make the best recommendation to solve the problem. The HIT in the CLE talent program was created by BioEnterprise, CoverMyMeds, and the Cleveland Foundation to help students realize the possibilities of a career in health IT and to help IT firms in Northeast Ohio construct a talent pipeline.

RTAHeatMap Civic Hacking Project

Code for America

ETL GUILD

Jan. 2015 - Jun. 2015

Historically, transportation and planning organizations have had to rely significantly on costly and unreliable manual data collection systems to manage and evaluate their services. The volume, kind, and quality of data available to planners and managers is changing as Information and Telecommunication Technology advances. The goal of this project is to use such technologies to enable better and more full access to the public record, and to depict this data as an isochrone map.

After School Supplemental Education Program

T.C.P. World Academy

TUTOR

Sep. 2008 - May. 2009

Through involvement in society simulation activities utilizing technology for higher learning, students at T.C.P. World Academy become academically involved and independent learners. T.C.P. World Academy educates students how to grasp the purpose of their education as well as the value of this experience and their ability to participate in a global society.

Certifications

2022 **Data Science**, Capstone - Surgical scheduling optimization and procedure duration prediction

IBM

2017 **Agile Advocate**, Achieved 1,000 engagement hours in an Agile Team

IBM

2015 **Data Science**, Capstone - AI Paint Formula Prediction from Spectroscopic Measurements

Johns Hopkins University

Education

Master of Science in Physics

University of Cincinnati

McMICKEN COLLEGE OF ARTS AND SCIENCES

2008 - 2010

Massive Stellar Clusters in the Disk of the Milky Way Galaxy

Astronomers have a lot of information on globular clusters. They're easy to spot because they're mostly outside the galaxy's plane in the halo. Extinction is rare, the star population is dense, and they are rather numerous. However, nothing is known about the open cluster population in the galaxy's plane. Open clusters in the galactic disk have been mapped in this thesis using near-infrared broad-band photometry and spectroscopy.

Bachelor of Fine Arts in Printmaking

Ohio University

COLLEGE OF FINE ARTS

2000 - 2004

A Novel Method for Low-Toxic Photo-Etch Printing

This thesis demonstrates how aquatint etching techniques, which have long been employed in intaglio printmaking, may be applied to give artists with a unique method for photo-etch procedures. The photo etching process was enhanced using scientific approaches, and then compared to copper etched in ferric chloride.