

# Ben Bubnick

ADVISORY DATA SCIENTIST · TRAINING & EDUCATION 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

<b>Professional Summary</b>	Generalist/Utility roles, Data Curation, Data Science, Analytics Training, Machine Learning, Scientific Analysis
<b>Big Data Platform</b>	Cloud infrastructure using Hadoop, Hive, Impala, Spark, kafka, Oracle, SQL Server, Flink, NoSQL, DB2, IMS, PL/SQL
<b>Health Care Analytics</b>	HEDIS measures, Tailored analytics & custom tooling, Payer/Provider data structures: EMR, CCD, HL7, 837, CMS
<b>Demonstrated Past Success</b>	IBM Band Promotion, Subject Matter Expert, Lab Services Award Winner, Eminence and Excellence Award Winner

I manage client resources and drive technical solutions to timely completion. I identify, analyze, and extract key data from heterogeneous systems; and lead technical requirements gathering, data discovery, documentation, and implementation of assigned client integrations. I also 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. I create and lead teams that create tailored analytics and AI solutions to our customers and internal clients.

## Experience

### Advisory Data Scientist

IBM

WATSON HEALTH

Nov. 2020 - Present

Senior member of Data Onboarding & Governance 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. Directs the configuration of the data onboarding platform and associated infrastructure, and develop improvements to advance stakeholder's business priorities using tailored analytics.

### Senior Curation Data Architect

IBM

WATSON HEALTH

May. 2020 - Nov. 2020

Senior member of Data Onboarding & Governance 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. Directs the configuration of the data onboarding platform and associated infrastructure, and develop improvements to advance stakeholder's business priorities using tailored analytics.

### 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, and report on the "health of the data" during data onboarding. Tune machine learning classification algorithms for patient matching, and led the internal review of similar work by my peers.

### Data Scientist

IBM, Explorys

WATSON HEALTH, EXPLORYS AN IBM COMPANY

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.

### 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

### Color Scientist

PPG Industries

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.

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

### 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 RESEARCHER

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.

## Volunteer

---

2018/19 **Instructor**, Computer Science supplemental high school course hosted by BioEnterprise

CSforCLE/CMSD

2018/19 **Volunteer Coach**, Data Science High School Competition hosted by Case Western Reserve University

HIT in the CLE

2015 **ETL Volunteer**, RTAHeatMap project for Open Cleveland, a Code for America Brigade

Open Cleveland

2008/09 **Tutor**, On-site after school tutoring for T.C.P. World Academy

Community School

## Education

---

### Master of Science in Physics

University of Cincinnati

McMICKEN COLLEGE OF ARTS AND SCIENCES

2008 - 2010

- Thesis - Massive Stellar Clusters in the Disk of the Milky Way Galaxy

### Bachelor of Fine Arts in Printmaking

Ohio University

COLLEGE OF FINE ARTS

2000 - 2004

- Thesis - A Novel Method for Low-Toxic Photo-Etch Printing