Ben **Bubnick**

AI DATA CURATION ARCHITECT · ANALYTICS TRAINING SPECIALIST

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

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

"With 12+ years experience in data analysis and curation, I have leadership aspirations because I enjoy helping others grow and succeed. I find problem solving and working with other business units engaging. I have several years of experience doing this, and if the company sees value in me, I would be able to work in a leadership role and hit the ground running."

Summary

Professional Summary	Data Curation, Data Science, Analytics Training, Machine Learning, Scientific Analysis
Core Competencies	Cognitive Practitioner, Metrics Improvement, Agile Development, Training Development
Demonstrated Past Success	Subject Matter Expert, Lab Services Award Winner, Eminence and Excellence Award Winner

Experience

AI Data Curation Architect
IBM

Cleveland, Ohio
Jan. 2018 - PRESENT

- Reduced mapping rework by 45% with through new knowledge-driven initiatives
- Developed new data validation and review process that reduced total rework by 25%
- Organized 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 in the last 2 quarters
- Earned 7 badges during IBM Re-skilling Initiative and logged over 200 hours of training in 2018
- Curated 25 pages of AI workflow documentation and mentored new hires in AI tuning process
- Awarded Lab Services Award for coordinating internal project teams to deliver client needs on time
- Awarded GDP bonus for improving revenue per SaaS client by at least 1.34%

Data Scientist
IBM, EXPLORYS

Cleveland, Ohio
Jun. 2015 - May. 2018

- Eliminated need for a $30,000 annual contract by developing data curation tool
- Principal integration specialist on over 60 projects for 20 different clients
- Chosen as team lead for ≈12 remote data science contractors for 36 integration projects throughout 2015 - 2017
- Went above my duties to close project documentation gaps, organizing more than 100 past projects
- Created over 125 new knowledge base articles, making up 30% of the team's online content
- Reduced time spent on knowledge transfer by ≈80% by reorganizing project documentation
- Awarded Eminence and Excellence Award in 2017 for analytics project
- Awarded GDP bonus for improving revenue per SaaS client by at least 1.17%

Data Integration Training Specialist
IBM

Cleveland, Ohio
Sep. 2016 - PRESENT

- Eliminated average of 3 weeks from employee onboarding process
- Reduced cost-per-hire by $5,210 by re-developing education program
- Spearheaded new-hire mentoring program that led to ≈4,500 increase in total billable hours
- Trained 60% of current data curation teams, including 2 managers

Software Engineer
AMTRUST FINANCIAL SERVICES

Cleveland, Ohio
Mar. 2014 - Jun. 2015

- 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
- Collaborated with developers and business analysts to prioritize needs
- Steered development group into a unit testing project on multi-tier MVC components in 2014

Color Scientist
PPG INDUSTRIES

Cleveland, Ohio
Apr. 2012 - Jun. 2013

- 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 color matching laboratory opening

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

Cleveland, Ohio
Aug. 2010 - PRESENT

- 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 pre-recorded lecture material for convenient consumption

Scientific Research

UNIVERSITY OF CINCINATI

Cincinnati, Ohio

Sep. 2006 - Sep. 2010

- 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 Hanna Research Fellowship in 2010

Talks

IBM Content & Data Analytics Learning Exchange

PRESENTER FOR THE INFORMATION SERVICES/INFORMATION TECHNOLOGY NETWORK TALKS

Cleveland, Ohio

Apr. & May 2018

- Bayesian Statistics using the Monty Hall problem example
- Naive Bayesian classification for patient matching
- Splines regression for population vitals ditributions

2018 IRI Fall Networks Conference

PRESENTER FOR THE INFORMATION SERVICES/INFORMATION TECHNOLOGY NETWORK TALKS

Cleveland, Ohio

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. Combined with the scalability and speed of the data lake architecture, this and other cognitive tools can effectively close the gaps on some of the most formidable flat file curation challenges.

IBM Cross Team Training Series

LEAD PRESENTER FOR THE METL GROUP

Cleveland, Ohio

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. All data is stored in the data lake in its native format until it is needed, and each data element has a metadata tag for easy retrieval. As a result, responses are ad-hoc rather than predetrmined, and reports based on new requirements can be delivered in days or weeks rather than months or years.

Writing

Process and Requirements for CCD Data Integration

LEAD INTEGRATION DEVELOPER

IBM

Oct. 2017

- White 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

RADIATIVE TRANSFER PHYSICS

PPG

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

GRADUATE RESEARCHER

OhioLINK ETD

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

UNDERGRADUATE RESEARCH ASSISTANT

PASP

Feb. 2008

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

Volunteer

2018/19	**Instructor,** Computer Science supplemental high school course hosted by CSforCLE/CMSD	*Cleveland, Ohio*
2018	**Volunteer Coach,** Data Science High School Competition hosted by HIT in the CLE	*Cleveland, Ohio*
2015	**ETL Volunteer,** RTAHeatMap project for Open Cleveland, a Code for America Brigade	*Cleveland, Ohio*
2008/09	**Tutor,** On-site after school tutoring for T.C.P. World Academy	*Cincinnati, Ohio*

Education

University of Cincinnati

M.S. PHYSICS

Cincinnati, Ohio

2008 - 2010

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

Ohio University

B.F.A. PRINTMAKING

Athens, Ohio

2000 - 2004

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

Johns Hopkins University   *Baltimore, MD*

Data Science Certificate (MOOC)   *2015 - 2016*

- Capstone - A Machine Learning Algorithm for Paint Formula Prediction from Spectroscopic Measurements

Johns Hopkins University   *Baltimore, MD*

Data Science Certificate (MOOC)   *2015 - 2016*

- Capstone - A Machine Learning Algorithm for Paint Formula Prediction from Spectroscopic Measurements