

Ben Bubnick

DATA SCIENCE · DATA CURATION · DATA INTEGRATION

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

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

Experience

Merative

DATA SCIENTIST

Tools: Python, Azure, Postman

Clinical Operations

Jul. 2022 - Present

Technical lead on the delivery of large & complex AI and Cognitive led transformation programs and works with Merative & Partners to optimize tools and delivery methods.

- Led HIPAA patient deidentification project from 200+ million individual patients for content analytics
- Created testing framework for ETL & implementation processes on new data platform

IBM

ADVISORY DATA SCIENTIST

Healthcare Data Analytics

Jun. 2015 - Jul. 2022

Technical lead on delivery of Cognitive & AI led programs using 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.

Advisory Data Scientist

Nov. 2021 - Jul. 2022

Tools: Python, GoLang, Azure, Docker, K8s, C#, SQL/PostgreSQL, Tableau

- Created machine learning Covid-19 tool that reduced need for hospitalization identification
- Led delivery on tailored analytics and reporting project for client administrative data
- Developed data migration validation analysis on 150+ billion patient records
- Developed surgical smart case card using combined CNN/RNN model

Promotion Senior Data Curation Architect

Dec. 2020 - Nov. 2021

Tools: Gradle, Python/R, Hadoop/HDFS, Hive/Impala, Spark, kafka, Docker, K8s, Elasticsearch, Kibana, SQL

Senior member of Data Onboarding & Governance for integration of healthcare data and enrichment of metadata. Configuration of data models, analytics, reports, and visualizations.

- Developed clinical data integration of 100+ billion structured & semi-structured records
- Developed topic modelling project on nearly 6 million semi-structured provider notes data
- Curated 25 pages of AI workflow documentation and mentored new hires in AI tuning process
- Created AI Surgical Scheduling Optimization tool currently evolving into new offering

Data Curation Architect

Nov. 2017 - Nov. 2020

Tools: Ruby, Pig, Java, Gradle, Perl, Hadoop/HDFS, SQL/SSIS/Oracle, NoSQL, Hive/Impala, kafka, Elasticsearch, Kibana

Lead member of Data Onboarding & Governance, member of the Watson Foundation for Health Solutions Architecture team.

- Managed multi-org enterprise integration with over 70 interconnected data sources
- Developed new data validation and review process that reduced total rework by 25%
- Reduced mapping rework by 45% with through new knowledge-driven initiatives

Software Developer

Sep. 2016 - Sep. 2017

Tools: Ruby, Pig, Java, Perl, Jenkins, Hadoop/HDFS/CDH, SQL/mysql/Oracle, DB2, IMS, PL/SQL

Technical delivery lead for client integrations which has included the ordering, installation, and configuration of required hardware, including VPN devices to support client integrations.

- Reduced time spent on knowledge transfer by ~80% by reorganizing project documentation
- Eliminated need for a \$30,000 annual contract by developing data curation tool
- Reduced cost-per-hire by \$5,210 by re-developing education program

Data Scientist

Nov. 2015 - Sep. 2016

Tools: Ruby, Pig, Hadoop/HDFS/CDH, Shell, SQL/mysql/Oracle, Jenkins, Java

Data Ingestions, pre-analytic enrichments, analytic augmentation, data governance, and implementation of assigned client integrations.

- Managed team of 12 remote data science contractors for 36 integration projects
- Created over 125 new knowledge base articles, making up 30% of the team's online content
- Developed automated data integration tool to cut out 28% unnecessary development time

Data Scientist Consultant

Jun. 2015 - Nov. 2015

Tools: Ruby, Pig, Hadoop/HDFS, React, NodeJS, JavaScript, Shell, SQL, Oracle, Jenkins

Identified clinical, financial, and operational data elements within different health provider systems to develop data transformations and providing meaning to the data.

- Developed cron-based widget to define blackout periods for client data pulls
- Performed troubleshooting on production issues across multiple environments
- Principal integration specialist for 10+ client database warehouses

Software Engineer

AmTrust Financial Services

COMMERCIAL PACKAGE POLICIES

Mar. 2014 - Jun. 2015

Tools: VB.NET, ASP.NET, C#, CSS, Javascript, HTML5, SQL, PL/SQL

Developed web forms for commercial insurance using ISO standards and created multi-tier MVC components, Web APIs, and UIs.

Color Scientist

AUTOMOTIVE REFINISH

Tools: VB.NET, C#, SQL, C++, FORTRAN

PPG

Apr. 2012 - Jun. 2013

Developed software and radiative transfer models, spectroscopic profile analysis on color match samples, and cross-instrument correlation software.

Research Assistant

DEPARTMENT OF PHYSICS

Tools: FORTRAN, Dataplot, IRAF, IDL, JMP, LabVIEW, Maple, Mathematica, MATLAB, PDL, SAS/GRAPH

University of Cincinnati

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.