Ben Bubnick

DATA SCIENCE · DATA CURATION · DATA INTEGRATION

Cleveland Heights, Ohio, 44118 | (216) 556-1612

✉ ben.bubnick@gmail.com | 🔗 benbubnick

Summary

Results-driven Data Scientist with a strong background in software engineering and research, bringing extensive experience in the healthcare industry, AI, and cognitive-led transformation programs. Adept at leading cross-functional teams, optimizing processes, and developing innovative data-driven solutions. Strong communicator and mentor with a proven track record of delivering results on time and within budget.

Technical Skills

Data Processing & Analytics	SQL, PostgreSQL, Hive/Impala, Spark, kafka, Kibana, Gradle, Jenkins
Programming Languages	Python, GoLang, Ruby, Pig, Perl, VB.NET, ASP.NET, C#, C++, FORTRAN
Tools & Frameworks	Azure, Hadoop/HDFS, Postman, Pandas, NumPy, Scikit-learn, XGBoost

Experience

Merative *Jul. 2022 - Present*

DATA SCIENTIST *Python, Azure, Postman*

Technical lead on AI and Cognitive-led transformation programs.

- 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 *Jun. 2015 - Jul. 2022*

Nov. 2021 - Jul. 2022

ADVISORY DATA SCIENTIST *Python, GoLang, Azure, Docker, SQL/PostgreSQL, Tableau*

Technical lead on Cognitive & AI-led programs using diverse data sources.

- 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

Dec. 2020 - Nov. 2021

Promotion SENIOR DATA CURATION ARCHITECT *Gradle, Python, Hive/Impala, Spark, kafka, Kibana, SQL*

Senior member of Data Onboarding & Governance.

- 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

Nov. 2017 - Dec. 2020

DATA CURATION ARCHITECT *Ruby, Pig, SQL/SSIS/Oracle, Hive/Impala, kafka, Kibana*

Lead member of Data Onboarding & Governance.

- 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 - Nov. 2017*
                                    Ruby, Pig, Perl, Jenkins, SQL/mySQL/Oracle, DB2, PL/SQL
Technical delivery lead for 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*
                                    Ruby, Pig, Jenkins, Hadoop/HDFS/CDH, SQL/mySQL/Oracle
Technical delivery lead for 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*
                                    Ruby, Pig, React, NodeJS, JavaScript, SQL, Oracle, Jenkins
Developed data transformations and provided meaning to the data.

AmTrust Financial Services   *Mar. 2014 - Jun. 2015*
SOFTWARE ENGINEER   *VB.NET, ASP.NET, C#, CSS, Javascript, HTML5, SQL, PL/SQL*
Created multi-tier MVC components and web forms for commercial insurance policy generating software.

PPG Industries   *Apr. 2012 - Jun. 2013*
COLOR SCIENTIST   *VB.NET, C#, SQL, C++, FORTRAN*
Developed radiative transfer models, spectroscopic profile analysis, and cross-instrument correlation software.

University of Cincinnati   *Sep. 2006 - Sep. 2010*
RESEARCH ASSISTANT   *FORTRAN, IRAF, IDL, JMP, LabVIEW, Maple, Mathematica, PDL*
Gathered data using a 0.8-5.4 μm spectrograph at NASA IRTF.

Awards

2018 - 2022	**Gold Champion**, For 1,477 classroom hours and earning 43 digital credentials	*IBM*
2018	**Lab Services Award**, For coordinating internal teams to deliver on time	*IBM*
2017	**Eminence and Excellence Award**, For developing a new data curation tool	*IBM*
2012	**Color Lab Automation Award**, For expanding data collection and analysis reports	*PPG Industries*
2010	**Mary J. Hanna Fellowship**, For STEM students pursuing research-based Masters	*UC*

Education

2022	**Certification**, Data Science	*IBM*
2015	**Certification**, Data Science	*Johns Hopkins University*
2010	**Master of Science**, Astrophysics	*University of Cincinnati*
2004	**Bachelor of Fine Arts**, Printmaking	*Ohio University*