**Shawn Becker**

**Lehi, UT 84043**

**sbecker@linkedin/in/shawn.becker**

**Machine Learning / Data Engineering / Data Architecture**

**Driven ML Developer and Data Engineer seeking challenges to excel in ML Ops and advanced data modeling.**

**Experience**

**Senior Data Engineer**

**Fannie Mae / Risk & Corporate Systems · Contract**

**Feb 2024 - Jul 2024 · 6 mos**

**Remote**

**• Documented processes to build, test, and deploy data pipeline**

**components for in-house ETL framework.**

**• Extensive work with SQL, AWS Redshift, Glue, S3, IAM, Lambda, REST,**

**API Gateway, Postman, and SNS.**

**• Utilized Agile practices with Jira, including backlog refinements, sprint**

**planning, daily scrums, bi-weekly sprint reviews, and end-of-sprint**

**retrospectives. Enabled the product owner to review each shipped**

**product increment, allowing for potential revision or re-prioritization of**

**backlog items.**

**Skills: Daily Scrum · REST APIs · Continuous Integration and Continuous**

**Delivery (CI/CD) · Apache Airflow · Docker · Sprint Planning · Data Build**

**Tool (DBT) · SQL · Git · GitHub · JavaScript · User Acceptance Testing ·**

**Agile Methodologies · Extract, Transform, Load (ETL) · Amazon**

**CloudWatch · Data Architecture · Sprints · Software Development Life**

**Cycle (SDLC) · Amazon S3 · Sprint Retrospective · Kubernetes · Data**

**Modeling · Microsoft Excel · Node.js · AWS Glue · Key Performance**

**Indicators · Amazon EC2 · Linux · Data Engineering · Databases · Scrum ·**

**AWS Lambda · Unit Testing · Jira · Jenkins · Docker Compose · Sprint**

**Review · Amazon Redshift · OpenAI Products · Statistical Data Analysis ·**

**Python (Programming Language)**

**Senior Data Engineer**

**The Cigna Group - Data CyberSecurity Division · Contract**

**May 2023 - Dec 2023 · 8 mos**

**Remote**

**• Modernized applications via Jenkins CI/CD pipeline upgrade,**

**integrating SetupTools, Artifactory/PyPI, SonarQube, and Xray.**

**• Investigated and implemented preparation of legacy ETL data pipeline**

**components, migrating from on-prem Unity IoC apps to the AWS cloud**

**using change data capture.**

**• Engineered a Python REST API enabling credential retrieval from**

**CyberArk’s identity management platform using mutual TLS/SSL**

**authentication via AWS API Gateway**

**• Initiated CyberArk service updates to extract credentials at runtime,**

**avoiding the need to access locally encrypted files and eliminating**

**engineering effort to satisfy cybersecurity password rollover**

**requirements, reducing costs by 95% for each password rollover event.**

**Skills: Daily Scrum · REST APIs · Continuous Integration and Continuous**

**Delivery (CI/CD) · Application Migrations · Apache Airflow · Azure**

**Databricks · Sprint Planning · SQL · Git · GitHub · User Acceptance**

**Testing · Agile Methodologies · Extract, Transform, Load (ETL) · Amazon**

**CloudWatch · Data Architecture · Sprints · Software Development Life**

**Cycle (SDLC) · Amazon S3 · Sprint Retrospective · Data Modeling ·**

**Node.js · AWS Glue · Key Performance Indicators · Amazon EC2 · Linux ·**

**Data Engineering · Databases · Scrum · AWS Lambda · Unit Testing · Jira**

**· Jenkins · Cyber Security · Sprint Review · Amazon Redshift · Integration**

**Testing · OpenAI Products · Python (Programming Language)**

**Senior Data Engineer**

**Warner Bros. Interactive Entertainment · Contract**

**Sep 2022 - Apr 2023 · 8 mos**

**Remote**

**HomeMy NetworkJobsMessaging NotificationsMe For BusinessLear**

**24• Utilized PySpark for ETL/ELT processes and Python for third-party**

**integrations and dev-ops collaborations with Jenkins, DataDog, and**

**ZenDesk. Leveraged Google BigQuery and AWS services, including**

**PostgreSQL, Glue, Lambda, and SalesForce.**

**• Developed high-volume pipeline ingress and integrations, enabling**

**efficient and secure game telemetry and user PII data transfer between**

**WB-distributed games and leading marketing platforms using Segment**

**CDP, Kafka, Redshift, Glue, and Airflow.**

**• Conducted exploration and statistical analysis of game marketing data**

**using Jupyter, Python, Pandas, NumPy, Sci-kit Learn, and Keras for**

**analytical processing and Seaborn, Plotly, and Matplotlib for data**

**visualization.**

**• Practiced Agile SDLC in Jira with spring planning, daily scrums, and bi-**

**weekly sprint reviews.**

**Skills: Google BigQuery · Daily Scrum · REST APIs · Continuous**

**Integration and Continuous Delivery (CI/CD) · Apache Airflow · Sprint**

**Planning · SQL · Amazon QuickSight · Martech · Git · GitHub · Ansible ·**

**User Acceptance Testing · Extract, Transform, Load (ETL) · Tableau ·**

**Amazon CloudWatch · Sprints · Software Development Life Cycle (SDLC)**

**· Amazon S3 · Sprint Retrospective · Data Modeling · Snow**

**flake · AWS Glue · Key Performance Indicators · Amazon EC2 · Linux · Data**

**Engineering · Databases · Scrum · AWS Lambda · Unit Testing · Jira ·**

**Jenkins · Sprint Review · Amazon Redshift · Integration Testing · OpenAI**

**Products · Statistical Data Analysis · Pandas (Software) · Python**

**(Programming Language) · Amazon Aurora · Apache Kafka**

**Senior Data Engineer**

**Angel Studios · Full-time**

**Dec 2021 - Aug 2022 · 9 mos**

**Provo, Utah, United States**

**• Developed a production-ready CNN using AWS SageMaker, Python,**

**PyTorch, and Keras for classifying movie frames, enabling a new revenue**

**opportunity through automated content tagging.**

**• Earned certifications in Advanced Learning Algorithms, Advanced SQL**

**for Data Scientists, and Supervised Machine Learning: Regression &**

**Classification, enhancing professional skills.**

**• Conducted data exploration and machine learning algorithms using**

**Jupyter, Python, Pandas, NumPy, Sci-kit Learn, Pytorch, and Keras for**

**processing and Seaborn, Plotly, and Matplotlib**

**for data visualization,**

**significantly enhancing analytical capabilities.**

**• Developed effective business intelligence strategies using Snowflake,**

**Looker, and Tablea for comprehensive sales and finance reporting.**

**Skills: Daily Scrum · Continuous Integration and Continuous Delivery**

**(CI/CD) · Docker · Sprint Planning · Convolutional Neural Networks (CNN)**

**· SQL · Amazon QuickSight · Martech · Git · GitHub · User Acceptance**

**Testing · Extract, Transform, Load (ETL) · Tableau · Amazon CloudWatch ·**

**Algorithms · Sprints · Software Development Life Cycle (SDLC) · Amazon**

**S3 · Sprint Retrospective · Data Modeling · Snowflake · Microsoft Excel · Node.js · Key Performance Indicators · Linux · Data Engineering · Databases · AWS SageMaker · Scrum · Unit Testing · Jira · Jenkins · Sprint Review · Statistical Data Analysis · Pandas (Software) · Python (Programming Language) · Machine Learning: Classification**

**Senior Data Engineer**

**Greenseed · Full-time**

**Nov**

**2020 - Nov**

**2021 · 1 yr 1 mo**

**Orem, Utah, United States**

**• Conducted advanced statistical exploration of real-estate sales data**

**using Python, Pandas, NumPy, SciPy, and Scikit-learn.**

**• Created insightful data visualizations with Seaborn, Plotly, and**

**Matplotlib.**

**• Implemented a CI/CD pipeline using GitHub Actions, Coverage,**

**SonarQube, and Xray.**

**• Migrated applications from on-prem PostgreSQL to AWS Redshift**

**warehouse using Docker, DockerHub, ECR, CloudWatch, and ECS with**

**Fargate.**

**• Enhanced machine learning skills with tutorials for TensorFlow,**

**PyTorch, and Keras using Kaggle datasets.**

**Skills: GitHub Actions · Daily Scrum · Continuous Integration and Continuous Delivery (CI/CD) · Application Migrations · Docker · Sprint**

**Planning · SQL · Amazon QuickSight · DevOps · Git · GitHub · Ansible ·**

**Extract, Transform, Load (ETL) · Tableau · Amazon CloudWatch ·**

**Algorithms · Sprints · Software Development Life Cycle (SDLC) · Amazon**

**S3 · Terraform · Kubernetes · Dimensional Modeling · Data Modeling ·**

**AWS CloudFormation · Microsoft Excel · Node.js · Amazon EC2 · Linux ·**

**Data Engineering · Leadership · Databases · Scrum · Software as a**

**Service (SaaS) · Database Design · Unit Testing · Jira · Jenkins · Docker**

**Compose · PostgreSQL · Sprint Review · Amazon Redshift · Statistical**

**Data Analysis · Pandas (Software) · Python (Programming Language) ·**

**Amazon Aurora**

**Senior Full Stack Engineer**

**Nu Skin · Contract**

**Nov 2019 – Nov 2020 · 1 yr 1 mo**

**Provo, Utah, United States · On-site**

**• Formulated storyboards, use cases, and unit tests to streamline project**

**development.**

**• Enhanced site registration and login pages by designing workflow and**

**wireframes.**

**• Innovated Vue Vuetify components with NodeJS SCSS for improved**

**f**

**unctionality and reuse.**

**• Documented and launched new packages for company-wide use,**

**enhancing efficiency.**

**• Internationalized content using Adobe Experience Cloud.**

**Skills: Daily Scrum · Continuous Integration and Continuous Delivery**

**(CI/CD) · Full-Stack Development · Sprint Planning · HTM**

**L · Git · GitHub ·**

**JavaScript · User Acceptance Testing · Agile Methodologies · Sprints ·**

**Software Development Life Cycle (SDLC) · Sprint Retrospective · Node.js**

**· Linux · Scrum · Unit Testing · Jira · Jenkins · Vue.js · Front-End**

**Development · Sprint Review · Visio**

**Senior Data Engineer / Data Architect**

**Seniorlink / Vela · Full-time**

**Mar 2017 - Nov**

**2019 · 2 yrs 9 mos**

**Boston, MA · On-site**

**• Designed and deployed an AWS data pipeline for the Vela platform,**

**which provided messaging, communication, and collaboration tools for**

**the healthcare industry.**

**• Managed data ingress by queueing API Gateway-delivered message**

**payloads into Amazon Kinesis Data Stream shards. Utilized SNS-**

**triggered Python jobs running on serverless Lambdas to aggregate shard**

**data into date-partitioned Parquet files in an S3 data lake.**

**• Performed ETL processes, extracting, transforming, and loading**

**Parquet data from the data lake to Redshift via scheduled PySpark batch**

**jobs on an EM**

**R cluster orchestrated by Data Pipeline.**

**• Generated daily business intelligence reports using Tableau**

**with**

**Redshift OLAP.**

**• Designed and built an in-house star-schema data warehouse using**

**dimensional modeling, featuring slowly changing type-2 dimension**

**tables sharing a common streaming facts table in PostgreSQL.**

**• Coordinated with domestic and offshore development and quality**

**assurance teams.**

**Skills: Daily Scrum · REST APIs · Continuous Integration and Continuous Delivery (CI/CD) · Sprint Planning · AWS DynamoDB · SQL · DevOps · Git · GitHub · User Acceptance Testing · Extract, Transform, Load (ETL) · Tableau · Amazon CloudWatch · Data Architecture · Sprints · Software Development Life Cycle (SDLC) · Amazon S3 · Terraform · Dimensional Modeling · Data Modeling · Microsoft Excel · Amazon EC2 · Linux · Data Engineering · Databases · Scrum · Spring · Software as a Service (SaaS) · Database Design · PySpark EMR · AWS Lambda · Unit Testing · Jira · Jenkins · PostgreSQL · Data Governance · Amazon Redshift · Integration Testing · OpenAI Products · Statistical Data Analysis · Pandas (Software) · Python (Programming Language) · Amazon Kinesis · Visio**

**Technical Project Manager / Solutions Architect / Co-Founder**

**ClipFile Corporation - Full-time**

**Feb 2011 - Mar 2017 · 6 yrs 2 mos**

**Newton Center, MA – on-site**

**• Managed a team to create and launch a pioneering SaaS on Amazon**

**Web Services, empowering readers and content creators to search and**

**share mindsets seamlessly.**

**• Designed and implemented patented technology, creating a**

**comprehensive consumer-facing CMS that facilitated precise matching**

**among selected items.**

**• Applied advanced techniques such as principal component analysis,**

**eigenvector decomposition, dimensionality reduction, vectorization,**

**cosine similarity, K-means, Bayesian clustering, and collaborative filtering**

**to implement fuzzy word matching and word clustering. This enhanced**

**end-user search results using NoSQL store SimpleDB and Hibernate.**

**• Utilized technologies including Java, J2ee, JSP, Servlets, Spring,**

**Maven, JavaScript, JQuery, HTML5, CSS, JUnit, JMeter, and JProfile.**

**• Developed a RESTful API layer using Swagger Shiro, Eclipse, Maven,**

**Ant, and Git.**

**Skills: REST APIs · Mentoring · Full-Stack Development · Sentiment**

**Analysis · AWS DynamoDB · HTML · SQL · DevOps · Git · GitHub · JavaScript · Spring MVC · Amazon CloudWatch · Project Risk · Algorithms· Data Architecture · Java · Software Development Life Cycle (SDLC) · Spring Boot · Microsoft Excel · Amazon EC2 · Linux · Leadership · Databases · Spring · Software as a Service (SaaS) · Database Design · AWS Lambda · Microsoft Project · PostgreSQL · Front-End Development ·**

**Pattern Recognition · Visio**

**Technical Program Manager / Solutions Architect / Co-Founder**

**Sierra Vista Group, LLC · Full-time**

**Nov 2002 - Feb 2011 · 8 yrs 4 mos**

**Boston, MA**

**• Identified profitable opportunities in product development, software**

**engineering, and data modeling and successf**

**ully negotiated budgets**

**and project milestones with C-level management.**

**• Recruited high-value independent consultants specializing in graphic**

**design, user experience, f**

**ull-stack development, quality assurance,**

**DevOps, and database administration.**

**• Developed and aligned comprehensive project schedules and detailed**

**technical specifications with specific business requirements within strict**

**budgets.**

**• Conducted implementation evaluations to ensure flawless delivery of**

**technical solutions.**

**• Managed IT strategies customized for clients in the entertainment,**

**medical services, manu**

**facturing, insurance, and cyber security**

**industries, including AMI, Rowe Jukeboxes, Eleven Systems, Coca-Cola**

**Corp. Europe, Medical Services Corp. (now**

**One Call), and Intrusic CyberSecurity.**

**• See SIERRA VISTA GROUP PROJECTS below**

**Skills: Application Migrations · Mentoring · Full-Stack Development ·**

**HTML · SQL · DevOps · Git · Waterfall Project Management · GitHub ·**

**JavaScript · Spring MVC · Project Risk · Algorithms · Data Architecture ·**

**Java · Software Development Life Cycle (SDLC) · Microsoft Excel ·**

**Amazon EC2 · Linux · Leadership · Databases · Spring · Software as a**

**Service (SaaS) · Database Design · Unit Testing · Microsoft Project ·**

**PostgreSQL · Data Governance · Business Requirements · Presentation**

**Skills · Visio · Stakeholder Management**

**Technical Project Manager / Solutions Architect**

**One Call Care Management (formerly MSC Care Management) ·**

**Contract**

**2007 - 2010 · 3 yrs**

**Jacksonville, FL · Hybrid**

**• Coordinated prioritization of**

**business requirements among**

**stakeholders in 10 different departments to customize an enterprise-**

**scale call center system for workers' compensation claim intake, vendor**

**selection, and patient care (One Call: Workers Compensation and**

**Healthcare Solutions).**

**• Created Functional Requirements and Technical Specifications**

**documents to guide the creation and testing of new**

**features.• Defined user stories and test cases to ensure clear communication and**

**understanding of requirements.**

**• Created project plans and schedules to effectively manage and track**

**progress.**

**• Oversaw quality assurance and User Acceptance Testing with**

**participants from all affected divisions.**

**• Performed code reviews to ensure code quality and adherence to best**

**practices.**

**• Mentored team members to develop their skills and ensure consistent**

**knowledge sharing.**

**• Identified and mitigated issues that could affect successfully hitting**

**milestones with existing resources without exceeding the negotiated**

**budget.**

**• A SIERRA VISTA GROUP PROJECT**

**Skills: Mentoring · Full-Stack Development · HTML · SQL · DevOps · Git ·**

**Waterfall Project Management · GitHub · JavaScript · Project Risk ·**

**Algorithms · Data Architecture · Java · Strategic Planning · Software**

**Development Life Cycle (SDLC) · Microsoft Excel · Amazon EC2 · Linux ·**

**Leadership · Databases · technical project manager · Software as a**

**Service (SaaS) · Database Design · Microsoft Project · Data Governance ·**

**Business Requirements · Presentation Skills · Visio · Stakeholder Management**

**Senior Software Architect**

**Trimble Inc. · Contract**

**2005 - 2007 · 2 yrs**

**Cambridge, MA – on-site**

**• Spearheaded the development and implementation of a custom**

**message-oriented middleware and data architecture to integrate the**

**mobile devices of a fleet of drivers to the legacy ERP systems of multi-**

**national clients, primarily Coca Cola Corp. Europe**

**• A SIERRA VISTA GROUP PROJECT**

**Skills: Application Migrations · Mentoring · Full-Stack Development · SQL**

**· DevOps · GitHub · Data Architecture · Java · Strategic Planning ·**

**Software Development Life Cycle (SDLC) · Data Modeling · Linux ·**

**Leadership · Databases · Database Design · Business Requirements ·**

**Presentation Skills · Stakeholder Management**

**Technical Project Manager / Solutions Architect**

**AMI Entertainment Network, LLC · Contract**

**2004 - 2005 · 1 yr**

**Grand Rapids, MI – hybrid**

**• Managed the design and development of system for encrypted music**

**and video content distribution, digital rights management, performance**

**and mechanical usage event tracking, and royalty payments.**

**• A SIERRA VISTA GROUP PROJECT**

**Skills: Mentoring · Full-Stack Development · HTML · SQL · DevOps · Waterfall Project Management · JavaScript · Project Risk · Data Architecture · Java · Strategic Planning · Software Development Life Cycle (SDLC) · Data Modeling · Microsoft Excel · Key Performance Indicators · Amazon EC2 · Linux · Leadership · Databases · technical project manager · Software as a Service (SaaS) · Database Design · Microsoft Project · Business Requirements · Presentation Skills · Visio · Stakeholder Management**

**Senior Software Architect**

**Intrusic Cybersecurity LLC**

**2003 - 2004 · 1 yr**

**Waltham, MA – on site**

**• Helped design an enterprise-level network compromise detection and**

**reporting system with real-time UDP/TCP packet tracking, storage,**

**analysis, and reporting using C++ on BSD Unix.**

**• A SIERRA VISTA GROUP PROJECT**

**Skills: REST APIs · SQL · DevOps · Algorithms · Data Architecture · Java · Strategic Planning · Software Development Life Cycle (SDLC) · Data Modeling · Linux**

**Technical Project Manager / Technical Lead**

**Rowe Entertainment (and Internet Jukeboxes) · Contract**

**2002 - 2003 · 1 yr**

**Grand Rapids, MI – hybrid**

**• Managed design and full-stack implementation of network,**

**database and software architecture to support new enterprise**

**application**

**• A SIERRA VISTA GROUP PROJECT**

**Skills: Application Migrations · Full-Stack Development · HTML · SQL ·**

**DevOps · Waterfall Project Management · JavaScript · Project Risk · Data**

**Architecture · Java · Software Development Life Cycle (SDLC) · Data**

**Modeling · Amazon EC2 · Linux · Leadership · Databases · technical**

**project manager · Software as a Service (SaaS) · Database Design ·**

**Microsoft Project · Front-End Development · Business Requirements ·**

**Presentation Skills · Visio · Stakeholder Management**

**Technical Lead / co-founder**

**HomePortfolio, Inc. · Full-time**

**1997 - 2002 · 5 yrs**

**Newtonville, MA – on-site**

**• Promoted over 10,000 manufacturers and retailers by building a**

**strongly attributed content management system and vendor pricing**

**exchange**

**• Helped secure over $70M in venture financing**

**Skills: Full-Stack Development · HTML · SQL · DevOps · JavaScript ·**

**Data Architecture · Java · Strategic Planning · Software Development Life**

**Cycle (SDLC) · Data Modeling · Microsoft Excel · Amazon EC2 · Linux ·**

**Leadership · Oracle Database · Software as a Service (SaaS) · Database**

**Design · Front-End Development · Pattern Recognition · Presentation**

**Skills**

**Photogrammetric 3D Scene Reconstruction**

**Autodesk · Contract**

**Apr 1996 - Aug 1996 · 5 mos**

**Remote**

**Cambridge, Massachusetts, United States**

**• Used self-authored photogrammetric scene reconstruction software to**

**build textured 3D models of interior architecture from one uncalibrated**

**image per room.**

**• Reconstructed kitchen models are featured in Autodesk's "Picture This**

**Kitchen" product, allowing users to visualize each room with different**

**flooring, cabinetry, appliances, lighting, and walls from a palette of**

**options.**

**Skills: C (Programming Language) · Linux · Machine Learning:**

**Regression · Computer Vision**

**Front-end Software Developer**

**Cimetrix Incorporated · Part-time**

**Jan 1989 - Aug 1990 · 1 yr 8 mos**

**Provo, Utah, United States – on-site**

**• Implemented GUI for multi-axis reverse kinematics robotic path**

**planning and simulation software.**

**Skills: HTML · C (Programming Language) · Linux · Full-stack development**

EDUCATION:

Massachusetts Institute of Technology

Doctorate of Philosophy (Ph.D.)

Media Arts and Sciences

Media Laboratory

1997

Discertation:

"Vision-assisted modeling for model-based video representations"

Publications:

\* https://www.researchgate.net/profile/Shawn-Beckersee

\* https://alumni.media.mit.edu/~sbeck/

Focus:

Reconstructing textured 3-D scenes from one or more 2-D images

· Algorithms · Statistical Data Analysis · Mentoring · C (Programming Language) · Presentation Skills · Computer Vision · Machine Learning: Regression · Linux · Machine Learning: Classification · Pattern Recognition · Applied Machine Learning · DevOps · HTML

Brigham Young University

Master of Science (MS)

Computer Science

College of Math and Sciences

1990

Thesis:

"Interactive Morphometrics with Depth Buffer and Linear Probe"​

Publications:

\* https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1732&context=facpub

Focus:

Algorithms · Mentoring · C (Programming Language) · Presentation Skills · Linux

Brigham Young University

Bachelor of Science (BS)

Design Engineering Technology

1987

Focus:

Machine Design - CAD/CAM/CAE - Industrial Engineering - C (Programming Language) · Linux