

Luke Acosta

Big Data Architect / Senior Developer

☎ (630) 555-1847 ✉ luke.acosta.bi@example.com 💼 linkedin.com/example/lukeacosta 📍 Naperville, IL 60540

SUMMARY

Big Data Architect and **Senior Developer** with 8+ years shaping **Business Intelligence** and **analytical data integration** solutions across structured and unstructured environments. Hands-on delivery covers Hadoop, Hive, Impala, Sqoop, HBase, SQL, HiveQL, Informatica, Python, Unix shell, metadata management, **data modeling**, source profiling, and data quality controls. Translates business information requirements into practical architectures, source-to-target mappings, scalable models, and reliable load processes. Leads discovery sessions with executives, business users, developers, Engineering, and Information Technology partners. Guides architecture reviews, customer meetings, data quality investigations, and competing delivery priorities with clear communication. Ready to help Northstar DataWorks improve trusted analytics, metadata visibility, and data integration outcomes across changing business needs.

EXPERIENCE

Big Data Architect / Senior Developer

Cedar Ridge Analytics 📅 September 2023 - Present 📍 Oak Brook, IL

Directs structured and unstructured integration architecture across Hadoop, relational platforms, metadata repositories, and reporting ecosystems. Leads discovery, design decisions, **data quality** governance, and customer-facing delivery for multiple concurrent initiatives.

- Directed **Hadoop** and relational integration architecture supporting finance, operations, and customer reporting ecosystems.
- Led executive discovery workshops, converting **information requirements** into models, mappings, and prioritized delivery plans.
- Established metadata standards across databases, Hadoop, Informatica, OLAP, and reporting catalogs for stronger lineage visibility.
- Designed Python and **Unix shell** validators for counts, null patterns, relationships, and domain values.
- Guided Informatica and **Sqoop** redesigns after source analysis exposed inconsistent keys and late-arriving records.
- Served as architecture contact during **customer meetings**, quality escalations, and competing delivery priorities.

Senior Business Intelligence Developer

Blue Mesa Systems 📅 June 2021 - August 2023 📍 Downers Grove, IL

Delivered enterprise **Business Intelligence solutions** joining SQL Server warehouse structures with Hadoop data sets. Owned modeling, reconciliation, ingestion controls, metadata organization, reporting collaboration, and customer communication across changing analytical requirements.

- Delivered **SQL** Server and **Hadoop** integration supporting combined transactional and semi-structured reporting analysis.
- Modeled dimensional, normalized, and wide-table structures for changing requirements and documented design tradeoffs.
- Built **HiveQL** and SQL reconciliations across landing, curated, and presentation layers before reporting cycles.
- Maintained **Informatica** workflows and Unix shell orchestration with control totals, restart logic, and exception notifications.
- Organized OLAP dimensions, **metadata** definitions, and catalog descriptions with reporting analysts for trusted measures.
- Presented design options and delivery status, translating technical constraints into aligned stakeholder decisions.

Business Intelligence Developer

Prairie Peak Consulting 📅 July 2019 - May 2021 📍 Chicago, IL

Developed analytical integration components for retail and distribution clients across governed warehouses and Hadoop environments. Combined source analysis, **data modeling**, scripting, metadata documentation, testing, and quality investigation within client delivery teams.

- Developed SQL, **Hive**, **Impala**, Sqoop, and Informatica components for governed analytical integration programs.
- Profiled databases and flat files, documenting semantics, relationships, duplicate patterns, and source limitations.
- Created **Python** and **Unix shell** utilities for validation, schema checks, sequencing, and operational handoffs.
- Converted reporting needs into logical models, transformation specifications, and test cases for application owners.
- Recorded ownership, transformation rules, refresh timing, and field definitions within metadata repositories.
- Traced quality incidents through **ETL** stages with developers and business representatives, documenting corrective actions.

Junior Data Integration Developer

Riverbend Health Services 📅 June 2018 - June 2019 📍 Evanston, IL

Built foundational integration assets for operational dashboards and recurring **Business Intelligence** reporting. Supported Hadoop experimentation, source reconciliation, metadata documentation, scripting, customer meetings, defect triage, and supervised release activities.

- Built SQL and Informatica mappings for structured systems feeding operational dashboards and recurring reports.
- Supported Hadoop proof of concept loads through Sqoop, Hive, and Impala under senior guidance.
- Performed source-to-target comparisons, relationship checks, and record reconciliations for scheduled load testing.
- Documented definitions, semantics, transformation rules, and refresh dependencies in a team metadata catalog.
- Wrote Python and **Unix shell** checks that produced exception summaries for engineering review.

- Participated in **customer meetings**, defect triage, and release walkthroughs with clear technical documentation.

PROJECTS

Enterprise Metadata Lineage Program 📅 2024

Created a cross-platform metadata framework spanning databases, Hadoop repositories, Informatica workflows, OLAP structures, and reporting catalogs. Improved source ownership, transformation visibility, and support handoffs.

Analytical Data Quality Control Framework 📅 2022

Built validation patterns with Python, Unix shell, SQL, and HiveQL for counts, nulls, relationships, domains, and reconciliation across analytical load stages.

LEADERSHIP & AWARDS

- Architecture Leadership Recognition, Cedar Ridge Analytics, 2025
- Client Delivery Excellence Award, Blue Mesa Systems, 2023
- Data Quality Improvement Recognition, Prairie Peak Consulting, 2021

EDUCATION

Bachelor's Degree in Information Systems

University of Illinois Chicago 🎓 GPA: 0 📅 2018 📍 Chicago, IL

***Coursework:** Database systems, data modeling, business intelligence, systems analysis*

CERTIFICATIONS

- Cloudera Certified Associate Data Analyst 📅 2020
- Informatica PowerCenter Developer Professional 📅 2021

TECHNICAL SKILLS

- **Big Data Platforms:** Hadoop, Hive, Impala, HBase
- **Data Transfer Tools:** Sqoop, Informatica PowerCenter, ETL
- **Query Languages:** SQL, HiveQL, relational querying
- **Scripting Languages:** Python, Ruby, PHP
- **Shell Automation:** Unix shell, Bash, file validation
- **Data Modeling:** Dimensional modeling, normalized models, wide tables
- **Source Analysis:** Data profiling, relationship analysis, source semantics
- **Metadata Systems:** Metadata repositories, data catalogs, business glossary
- **Business Intelligence:** OLAP, reporting tools, analytical data stores
- **Data Quality:** Reconciliation, control totals, exception analysis
- **Integration Design:** Source-to-target mapping, load sequencing, transformation rules
- **Documentation Tools:** Technical documentation, lineage records, test cases

SKILLS

- | | | | |
|----------|---------------|-----------------|---------------------------|
| • Hadoop | • Sqoop | • ETL | • Data Profiling |
| • Hive | • HBase | • Python | • Metadata Management |
| • HiveQL | • SQL | • Unix Shell | • Data Quality Management |
| • Impala | • Informatica | • Data Modeling | |

PROFESSIONAL AFFILIATIONS

- Data Management Association, Chicago Chapter
- Association for Computing Machinery, Professional Member

LANGUAGES

- English (Native)
- Spanish (Intermediate)

ADDITIONAL INFORMATION

Work Status : Authorized to work in United States. No sponsorship required.

REFERENCES

AVAILABLE ON REQUEST