

# 2. Data Engineering Role Framework



[2.1 What Data Engineers Actually Do](#)

[2.2 Core Technologies & Skills to Look For](#)

[2.3 Experience Level Indicators](#)

[2.4 Real LinkedIn Search Tips for Data Engineers](#)

[2.5 Beyond Keywords: Evaluating Real Data Engineering Expertise](#)

Overview.

## 2.1 What Data Engineers Actually Do

Data engineers build and maintain the systems that allow organizations to collect, store, access, and analyze data at scale. They're the architects and plumbers of the data world - creating the infrastructure that makes everything else possible.

Their day-to-day work includes:

- Building data pipelines that extract, transform, and load (ETL) data from various sources
- Designing data models and schemas for efficient storage and retrieval
- Implementing systems for data quality and governance
- Optimizing performance of data systems
- Collaborating with data scientists, analysts, and business stakeholders

The basics.

## 2.2 Core Technologies & Skills to Look For

### ✓ Essential Skills (Must-Have)

- **Python:** The lingua franca of data engineering
- **SQL:** Deep knowledge of database querying and optimization
- **Data Pipeline Tools:** Experience with tools like Airflow, Luigi, Prefect
- **Cloud Platforms:** AWS, Azure, or GCP experience
- **Version Control:** Git for code management
- **Data Warehousing:** Understanding dimensional modeling and data warehouse concepts

### + Valuable Additions (Nice-to-Have)

- **Scala/Java:** For Spark and other JVM-based data tools
- **Streaming Technologies:** Kafka, Kinesis, Spark Streaming
- **Container Orchestration:** Docker, Kubernetes
- **Infrastructure as Code:** Terraform, CloudFormation
- **Data Governance:** Experience with data cataloging, lineage tracking, quality monitoring

## 2.3 Experience Level Indicators

# Junior Data Engineers (0-2 years)

### What to look for in LinkedIn profiles

- Recent CS/Engineering graduates with data-focused coursework
- Bootcamp graduates with data engineering focus
- 1-2 professional projects or internships with data components
- Certifications in cloud platforms (AWS/Azure/GCP) or data tools

### Projects should demonstrate

- Basic ETL pipelines for small datasets
- Simple data transformations in Python
- Database creation and basic SQL queries
- Familiarity with at least one cloud service (S3, BigQuery, etc.)

### Red flags

- No hands-on project work
- Only theoretical knowledge without practical implementation
- No experience with Python or SQL

**Why this level fits certain teams:** Junior engineers bring fresh perspectives and are eager to learn. They're ideal for teams with established processes and senior engineers who can mentor them. They can take on well-defined tasks while growing into the role.

## 2.3 Experience Level Indicators

# Mid-Level Data Engineers (2-5 years)

### What to look for in LinkedIn profiles

- Multiple data projects across different domains or companies
- Specific mentions of tech beyond the basics (Airflow, Kafka, etc.)
- Contributions to multiple phases of the data lifecycle
- Evidence of increasing responsibility over time

### Projects should demonstrate

- End-to-end data pipeline implementation
- Multi-source data integration
- Performance optimization work
- Cloud-based data systems
- Data quality implementation
- Automated testing and monitoring

### Red flags

- Still working primarily with basic tools after 3+ years
- No experience with orchestration or scheduling
- Limited understanding of data modeling principles
- Inability to articulate data quality approaches

**Why this level fits certain teams:** Mid-level engineers can work independently on substantial components of your data infrastructure. They should be able to own features end-to-end and make sound architectural decisions within their domain. They're the backbone of productive data teams.

## 2.3 Experience Level Indicators

# Senior Data Engineers (5+ years)

### What to look for in LinkedIn profiles

- Architecture-level responsibilities
- Leadership of significant data initiatives
- Cross-functional collaboration experience
- Mentions of mentoring or team leadership
- Optimization and scaling work

### Projects should demonstrate

- Data architecture design at scale
- Performance tuning for large datasets
- Implementation of data governance frameworks
- Cross-team or cross-company data integration
- Technical leadership and strategic planning
- Business impact of data initiatives

### Red flags

- Still primarily working at implementation level without architectural responsibilities
- Limited experience with large-scale data challenges
- No evidence of cross-functional collaboration
- Unable to connect technical solutions to business outcomes

**Why this level fits certain teams:** Senior engineers should shape the technical direction of your data infrastructure and mentor more junior team members. They understand both the technical and business sides of data engineering and can translate between them. They're essential for teams building complex, scalable systems.

Get searching.

## 2.4 Real LinkedIn Search Tips for Data Engineers

### > Basic Boolean Search Examples

```
("data engineer" OR "ETL developer" OR "data pipeline engineer")  
AND (Python OR SQL)  
AND (AWS OR Azure OR GCP OR "Google Cloud")  
AND (Airflow OR Kafka OR Spark OR dbt)
```

### > Finding Junior Candidates

```
("data engineer" OR "ETL developer" OR "data analyst")  
AND (Python OR SQL)  
AND ("recent graduate" OR "junior" OR "associate" OR "bootcamp")  
AND (project OR internship OR certification)
```

### > Finding Senior Candidates

```
("senior data engineer" OR "lead data engineer" OR "principal data  
engineer" OR "data architect")  
AND (architecture OR design OR "data modeling" OR governance)  
AND (scale OR enterprise OR "data lake" OR "data warehouse")  
AND (team OR mentor OR lead)
```

### Pro Tip

Check which LinkedIn groups they're part of. Active members of data engineering communities often stay current with best practices.

Final tips.

## 2.5 Beyond Keywords: Evaluating Real Data Engineering Expertise

When reviewing profiles, look beyond mere technology mentions. Here's what separates real data engineers from those who just list tools:

- 1. Project scale descriptions:** They should mention data volumes, pipeline complexity, or performance metrics
- 2. Problem-solving narratives:** Look for descriptions of challenges overcome, not just technologies used
- 3. End-to-end ownership:** Evidence they built complete solutions, not just components
- 4. Business impact:** Mentions of how their work affected business outcomes
- 5. Technical depth:** Detailed explanations of architecture decisions or optimization approaches

### ✓ Example of a strong LinkedIn description

"Led redesign of customer analytics pipeline processing 5TB daily, reducing processing time by 70% and enabling real-time fraud detection that saved \$2M annually. Architected cloud migration from on-premise Hadoop to AWS EMR/S3 data lake with Airflow orchestration."

### ✗ Example of a weak LinkedIn description

"Worked with Python, SQL, AWS, Airflow, and Spark to build data pipelines for the marketing team."