

Data Engineering

204426

Challenges in Data Engineering

Challenges in Data Engineering

- Data Engineers Must Learn On Their Feet
- Too Much Data To Handle
- Poor Performance
- Can't Get To Data
- Data Pipeline Maintenance
- Data Governance

Challenges in Data Engineering

Data Engineers Must Learn On Their Feet

- Data engineering is a relatively new and dynamic discipline.
- While engineers of the past focused more on creating data pipelines and collecting data into warehouses.
- Now the work is far more complicated, with added responsibilities in data analytics and building algorithms.
- Advancements are industry-changing, and that change is still ongoing.
- Diversity in Skill Sets
- Data engineering role to keep changing with it, and where it ultimately ends up remains to be seen.

Challenges in Data Engineering

Too Much Data To Handle

- Data engineers today must work with:
 - A huge amount of data
 - Fast grows
 - High variation

} **Big Data**

Poor Performance

- A huge amount of data → Slow to generate report and model
- Move to data warehouses
- Cloud data warehouses have several advantages:
 - More scalable and elastic than the traditional warehouse
 - Save time and resources on database management

Challenges in Data Engineering

Can't Get To Data

- Pull in data sets fast enough
- ETL technology → code-heavy and delay your process further
- ELT technology → working with the data on an as-needed basis
 - It can be useful in developing a bigger picture of the data and guiding you toward better data sets for your core models.

Data Pipeline Maintenance

- Demand for more data pipelines and Big Data
- → Keep existing pipelines in working order
- Low-code/no-code systems → reduces the maintenance burden

Challenges in Data Engineering

Data Governance

- Data governance is a collection of processes, roles, policies, standards, and metrics that ensure the effective and efficient use of information in enabling an organization to achieve its goals.
- Prevent bad data quality and protect the data
- Balance between getting data quickly and “good enough” and keeping the data accurate enough to make sound business decisions.

References

- <https://www.theseattledataguy.com/9-challenges-that-data-engineers-face-data-engineering-consulting>
- <https://towardsdatascience.com/how-to-break-into-data-engineering-and-why-its-so-hard-afe023922ef6>