

# ELT

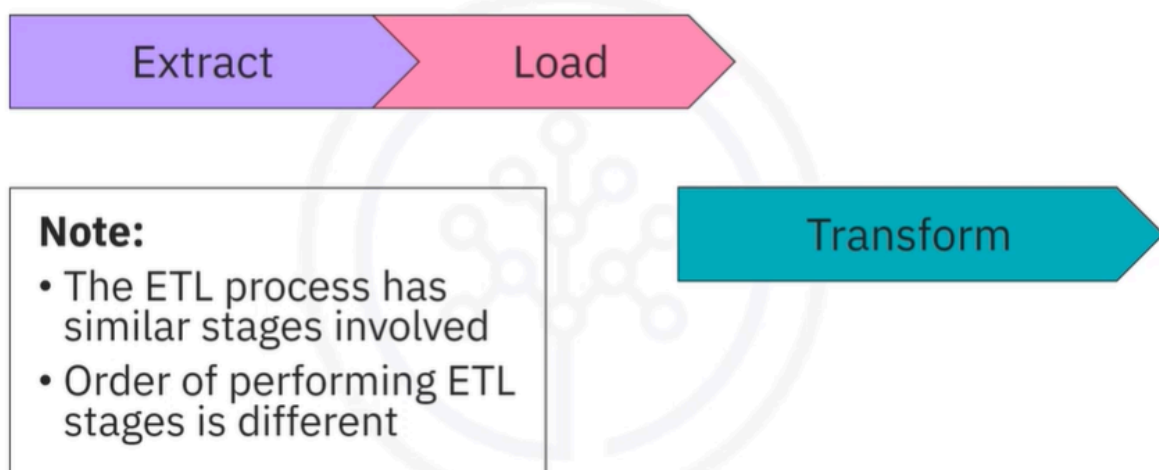
Sure! Let's talk about the ELT process, which stands for Extract, Load, and Transform.

In simple terms, ELT is a method used to handle data. Imagine you have a big box of LEGO bricks (this represents your data) that you want to build something cool with. First, you **extract** the bricks from different places, like your closet or a friend's house. Next, you **load** all those bricks into a big container (this is like loading the data into a storage system). Finally, instead of building your creation right away, you can take your time to **transform** the bricks into whatever you want, whenever you want. This means you can change your mind and build something different without losing any bricks!

This process is especially useful when dealing with large amounts of data because it allows for flexibility and speed. You can explore and analyze your data in various ways without worrying about losing any important information.

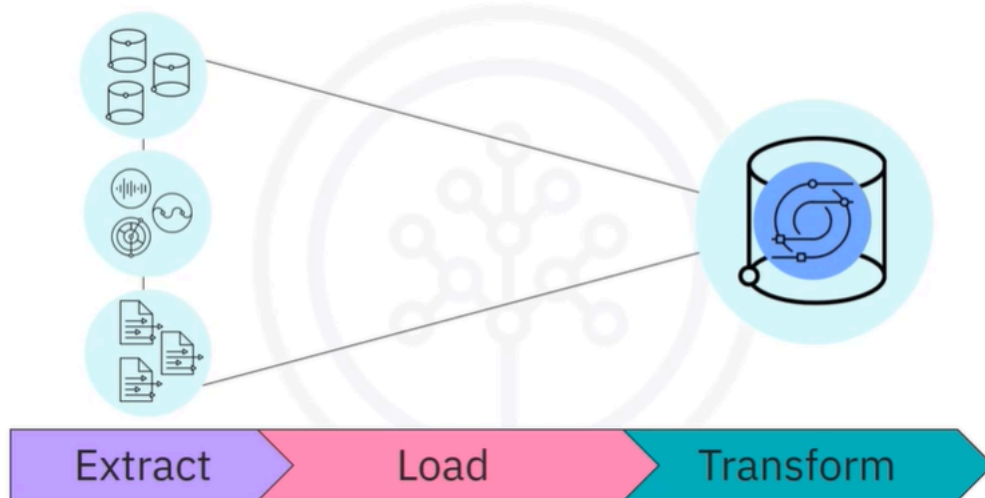
## What is an ELT process?

---



What is an ELT process? ELT stands for extract, load, and transform. ELT is an acronym for a specific automated data pipeline engineering methodology. ELT is similar to ETL in that similar stages are involved, but the order in which they are performed is different.

## What is an ELT process?

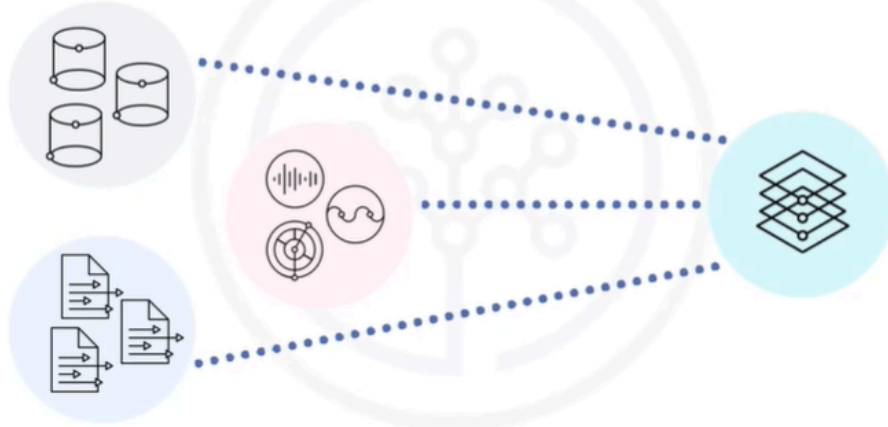


- For ELT processes, data is acquired and directly loaded as is into its destination environment. From its new home, usually a sophisticated analytics platform such as a data lake, it can be transformed on demand and however users wish.

## What is the Extraction process?

---

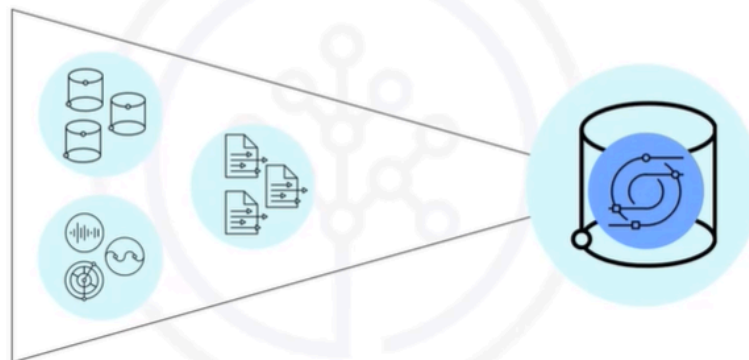
E=Extract: Extracting data from sources



## What is the Loading process?

---

L=Load: Loading data as-is into destination system



# What is the Transform process?

---

T=Transform: Transforming data on demand



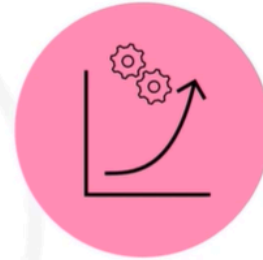
- Like ETL, the first stage in the ELT process is extraction. The extraction process obtains the data from all sources and reads the data often in an asynchronous fashion into an application.
- The loading process takes the raw data as is and loads it into its new environment where modern analytics tools can then be used directly.
- The transformation process for ELT is much more dynamic than it is for conventional ETL. Modern analytics tools in the destination environment enable interactive, on-demand exploration and visualization of your data, including advanced analytics, such as modeling and prediction.

## ELT use cases

---

Cases include:

- Demanding scalability requirements of Big Data
- Streaming analytics
- Integration of highly distributed data sources
- Multiple data products from the same sources



- Use cases for ELT processes typically fall within the high performance computing and big data realms. Cases include dealing with the massive swings and scale that come with implementing big data products, calculating real-time analytics on streaming big data, and bringing together data sources that are highly distributed around the globe. In terms of speed, moving data is usually more of a bottleneck than processing it. The less you move it, the better. Therefore, ELT may be your best bet when you want flexibility in building a suite of data products from the same sources.

# ELT is an emerging trend

---

- Big Data -> Cloud computing



- ELT separates the data pipeline from the processing
  - More flexibility
  - No information loss
- 
- Why is ELT emerging? Firstly, cloud computing solutions are evolving at tremendous rates due to the demands of big data. They can easily handle huge amounts of asynchronous data, which can be highly distributed around the world.
  - Cloud computing resources are practically unlimited and they can scale on demand. Unlike traditional on-premises hardware, you only pay for the computing resources you use. You do not have to worry about under utilizing resources, that is, overspending on equipment. With ELT, you can have a clean separation between moving data and processing data.
  - Of course, cloud computing is equally prepared to handle the most challenging cases for either of these two tasks. There may be many reasons to transform your data and just as many ways to do it. Thus, ELT is a flexible option that enables a variety of applications from the same source of data.
  - Because you are working with a replica of the source data, there is no information loss. Many kinds of transformations can lead to information loss, and if these happen somewhere upstream in the pipeline, it may be a long time before you can have a change request met. Worse yet, the information may be forever lost if the raw data is not stored.