

Enhance Data Pipeline Reliability with MANTA's Time Slicing Feature

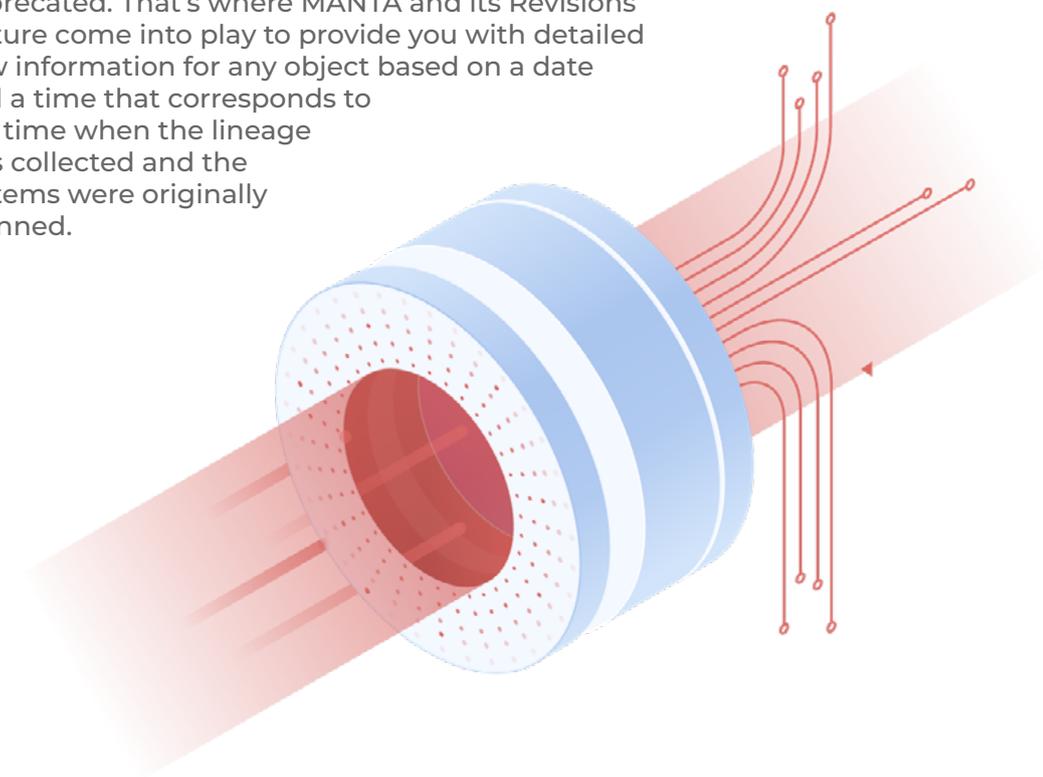


Enhance Data Pipeline Reliability with MANTA's Time Slicing Feature

Keeping complex data pipelines healthy and making the most out of data analytics while reducing cycle time without harming the quality of the analytics solution is not an easy task. Data pipeline observability is nearly impossible to achieve when there are manual tasks involved. Data teams run into numerous challenges caused by a lack of collaboration and erroneous results that harm business analytics and lead to wrong decisions.

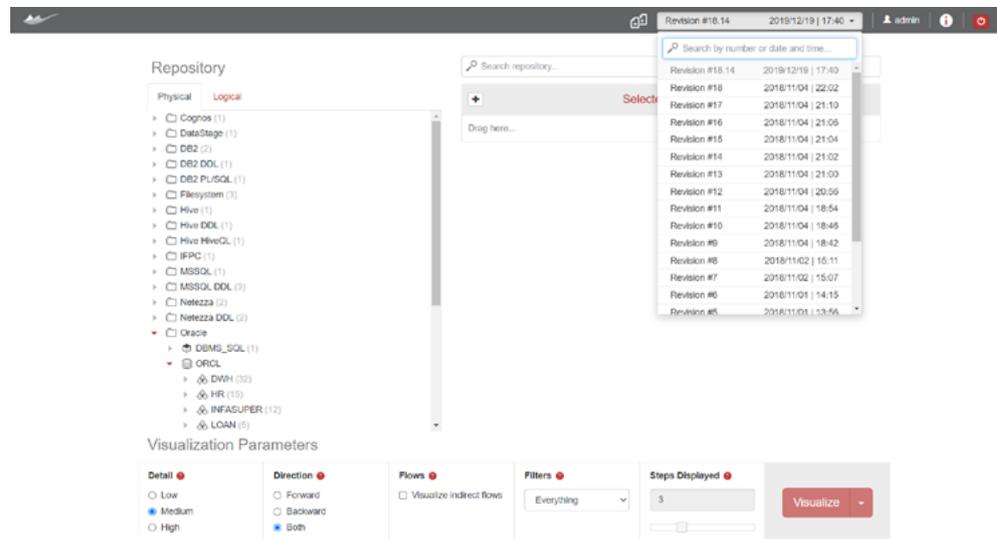
To make the most out of analytics, it's not enough to know the current state of the data. For various purposes, such as reporting as well as impact and root cause analyses, it's crucial to know what the environment looked like in the past and how that differs from its current state.

As time goes by, even if data lineage is efficiently collected and reported, it can start to get stale. Parts and pieces go missing if the solution cannot report on lineage from dates in the past. Manually configured lineage is almost impossible to recreate, as your ETL developer might have moved on, sources might have been archived, or entire applications may have been deprecated. That's where MANTA and its Revisions feature come into play to provide you with detailed flow information for any object based on a date and a time that corresponds to the time when the lineage was collected and the systems were originally scanned.



What Is a Revision?

Every time you run a scan in MANTA, the results are saved in our repository that you can access via the revisions dropdown. A Revision is simply a slice in time showing you how the system looked at the time of the selected scan.

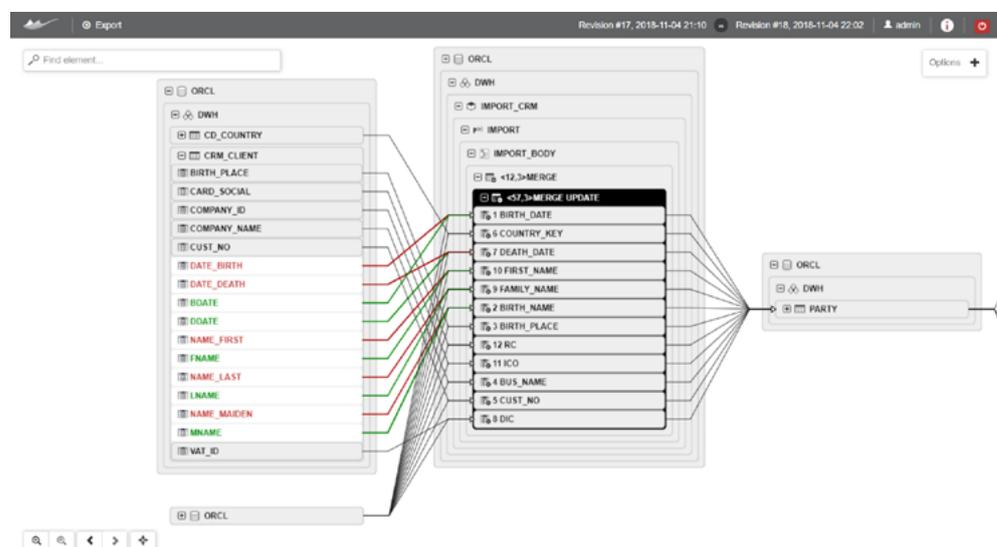


Select an older Revision to visualize the assets within the selected graph and see how they looked at that moment in time. It will help you understand how the data flows looked earlier.

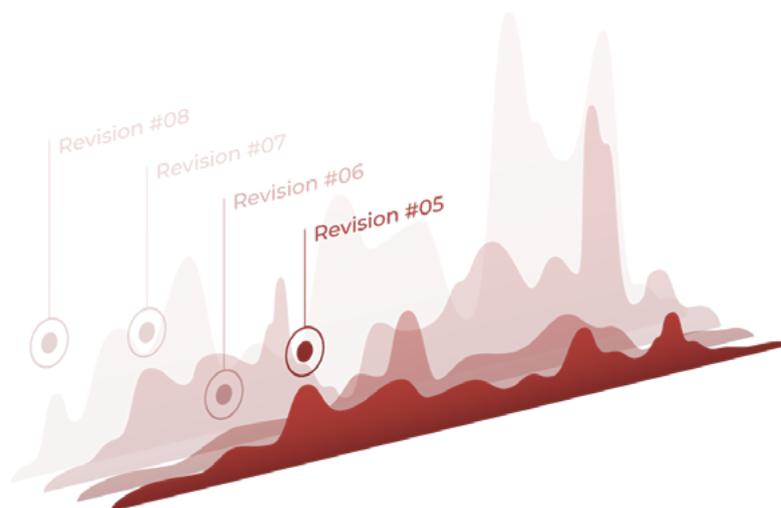
Comparing Revisions to Achieve Data Pipeline Observability

Knowing how lineage looked in the past paves the way for DataOps implementation, and the ability to **compare historical revisions** is key to achieving a holistic view of the data landscape.

You can do it easily in MANTA. Simply select two different Revisions in the repository to see what new elements were added (marked in green) and what elements were deleted (marked in red). You can compare the latest Revision with historical ones or compare historical Revisions to each other.



Comparing two different time slices provides you with powerful insights into how the lineage has developed. You can easily identify what changed in the flow and what could have possibly caused discrepancies or breakage in the flow. It will also help you answer numerous questions such as:



- Are there additional (or fewer) sources and targets?
- Have the SQL statements changed?
- Are the individual column mappings or calculations the same?
- Can we locate the source of a problem that occurred several months ago and trace exactly what happened and in what piece of code?
- How have the characteristics of an asset (column lengths, descriptions, or added-value metadata regarding data quality or sensitivity) changed over time?

Easily Monitor, Review, and Control Changes in Data Flows

What else can you do with lineage and historical comparisons?

You can **monitor the increase in total items** you are governing across the environment or **review trends in code changes or the number of touchpoints that access a particular element or table in the data lake**. You can also **view the distribution of recognized personal data to external systems** and then document the changes in those flows. With this ability, historical lineage will allow you to **illustrate the progress made towards controlling where your data is flowing** (or illustrate a growing lack of such control!).

MANTA Revisions are a powerful feature that helps your organization implement DataOps and achieve data pipeline observability. Lineage is collected and time slices are compared in an automated fashion, which delivers results faster and is more accurate than manual efforts. Revisions also solve collaboration issues. What normally would have involved hours of reconstructing and comparing past data flows can be achieved with a few clicks.

Would you like to know more about our time slicing feature and how it can enhance your organization's DataOps efforts? Get in touch with us at manta@getmanta.com. We will be more than happy to show you how Revisions work in real life.