



## Reading: ETL-specific tools and their applications

In a previous reading, you were given a list of common business intelligence tools and some of their uses. Many of them have built-in pipeline functionality, but there are a few ETL-specific tools you may encounter. Creating pipeline systems—including ETL pipelines that move and transform data between different data sources to the target database—is a large part of a BI professional's job, so having an idea of what tools are out there can be really useful. This reading provides an overview.

Tool	Uses
Apache Nifi	<ul style="list-style-type: none"><li>• Connect a variety of data sources</li><li>• Access a web-based user interface</li><li>• Configure and change pipeline systems as needed</li><li>• Modify data movement through the system at any time</li></ul>
Google DataFlow	<ul style="list-style-type: none"><li>• Synchronize or replicate data across a variety of data sources</li><li>• Identify pipeline issues with smart diagnostic features</li><li>• Use SQL to develop pipelines from the BigQuery UI</li><li>• Schedule resources to reduce batch processing costs</li><li>• Use pipeline templates to kickstart the pipeline creation process and share systems across your organization</li></ul>
IBM InfoSphere Information Server	<ul style="list-style-type: none"><li>• Integrate data across multiple systems</li><li>• Govern and explore available data</li><li>• Improve business alignment and processes</li><li>• Analyze and monitor data from multiple data sources</li></ul>
Microsoft SQL SIS	<ul style="list-style-type: none"><li>• Connect data from a variety of sources integration</li><li>• Use built-in transformation tools</li><li>• Access graphical tools to create solutions</li></ul>

	<p>without coding</p> <ul style="list-style-type: none"> <li>• Generate custom packages to address specific business needs</li> </ul>
Oracle Data Integrator	<ul style="list-style-type: none"> <li>• Connect data from a variety of sources</li> <li>• Track changes and monitor system performance with built-in features</li> <li>• Access system monitoring and drill-down capabilities</li> <li>• Reduce monitoring costs with access to built-in Oracle services</li> </ul>
Pentaho Data Integrator	<ul style="list-style-type: none"> <li>• Connect data from a variety of sources</li> <li>• Create codeless pipelines with drag-and-drop interface</li> <li>• Access dataflow templates for easy use</li> <li>• Analyze data with integrated tools</li> </ul>
Talend	<ul style="list-style-type: none"> <li>• Connect data from a variety of sources</li> <li>• Design, implement, and reuse pipeline from a cloud server</li> <li>• Access and search for data using integrated Talend services</li> <li>• Clean and prepare data with built-in tools</li> </ul>

---