



HITACHI
Inspire the Next

Pentaho Data Integration Review From A Customer



From IT Central Station, the leading review site for enterprise technology solutions.

Review by a Real User

Verified by IT Central Station



Senior Consultant at a financial services firm with 10,001+ employees

it_user402600

VALUABLE FEATURES:

It allows for rapid prototyping of a wide array of ETL workloads.

ROOM FOR IMPROVEMENT:

Support for common Hadoop utilities can be expanded, such as bulk load with composite row keys for HBase, and include drivers for Impala out-of-the-box. A richer interface to Hive could also be beneficial as we currently have to go through a raw connection and execute SQL scripts, for which some syntax is not respected.

As of version 6, there are also some new issues introduced that pose a bit of an annoyance:

1) ON KETTLE'S RAMP UP - LOG4J ERRORS

2) IBM WEBSHERE MQ PRODUCER - VARIABLE SUBSTITUTION FOR THE URL DOES NOT WORK - YOU HAVE TO HARDCODE.

3) SHARED.XML FOR DB CONNECTIONS - VARIABLE SUBSTITUTION FOR CONNECTION PROPERTIES DOES NOT WORK - HAVE TO HARDCODE THINGS LIKE KERBEROS PRINCIPAL FOR A HIVE/IMPALA CONNECTION.

Deployment Issues:

WE HAD NO ISSUES DEPLOYING IT.

Scalability Issues:

THE ROBUSTNESS OF THIS SOLUTION IN A PRODUCTION CLUSTER (>30 NODES) REMAINS TO BE SEEN.

[Read 1 reviews of Pentaho Data Integration](#)