

# DATA-FITNESS

You've heard of it. **Now measure it.**

Real data-fitness is foundational. It runs wide and deep, through your people, processes and technology. It gets you firing on all cylinders and performing at the top of your game. Here's how we define it.

## INDICATORS:

- ✓ Handling massive data volumes and variety in an easily scalable system.
- ✓ New functionality layered on as required with little or no coding.
- ✓ Automatic ingestion and curation of any data format or source.

# STRENGTH

## INHIBITORS:

- ✗ Heavy manual workload with each new data format.
- ✗ High-cost integration work to connect point solutions.
- ✗ Bottlenecks and systems buckling under large data loads.

## INDICATORS:

- ✓ Open architecture platform with custom/community-built plug-ins.
- ✓ A high level of automation for onboarding new data sources and formats.
- ✓ Core-based and pay-as-you-go pricing models for maximum scalability and flexibility.

# AGILITY

## INHIBITORS:

- ✗ Stitching together an ever-increasing number of point solutions.
- ✗ Consumption models which scale poorly and lead to hyperscaler lock-in.
- ✗ Manual coding for ingestion and novel transformations delays data readiness.

## INDICATORS:

- ✓ Quality monitoring with an AI/ML-enabled metadata management tool.
- ✓ Rapid root-cause analysis and with full observability.
- ✓ Clear upstream/downstream lineage of data while preserving IT oversight.

# PRECISION

## INHIBITORS:

- ✗ Multiple manual checks and second guessing of untrusted data.
- ✗ Significant engineer time spent spotting defects and bottlenecks.
- ✗ High compliance risk with no way to identify or trace most PII data.

## INDICATORS:

- ✓ Tiering data between cloud/on-prem and storage/compute for economy.
- ✓ Democratize pipeline management across business units leaving oversight with IT.
- ✓ Policy-based governance for archiving and optimizing infrastructure.

# EFFICIENCY

## INHIBITORS:

- ✗ Adding ad hoc pipelines as an ineffective way to scale up to new volumes.
- ✗ No way to manage duplicate data without manual discernment and parsing.
- ✗ No way to view whole data landscape or collapse data silos for cost-benefit.

## INDICATORS:

- ✓ Express response time to requests with rapid analysis and actionable intelligence.
- ✓ Accelerated innovation cycle and a culture of data-driven experimentation.
- ✓ Expedited time to value with low-latency, high-volume, high-velocity data pipelines.

# SPEED

## INHIBITORS:

- ✗ Multiple manual checks and second guessing of untrusted data.
- ✗ Significant engineer time spent spotting defects and bottlenecks.
- ✗ High compliance risk with no way to identify or trace most PII data.

That defines it. **Now let's deliver it.**

Pentaho+ is a modular data intelligence platform comprising four highly evolved solutions. It grows with you as you develop your data-fitness.

**ORGANIZE** with Pentaho Data Catalog

Understand and organize data across your business.

**SCALE** with Pentaho Data Optimizer

Reduce operating costs and get seamless access to data.

**ORCHESTRATE** with Pentaho Data Integration

Easily manage data volume, variety and velocity.

**EXPLORE** with Pentaho Business Analytics

Monetize and unlock insights with embedded analytics.

## Pentaho+ Platform

Get data-fit.  
Pentaho

### Pentaho

Corporate Headquarters  
2535 Augustine Drive  
Santa Clara, CA 95054 USA  
pentaho.com

Hitachi® and Hitachi Vantara® are registered trademarks of Hitachi, Ltd. in the U.S. and other countries. Pentaho® is a registered trademark of Hitachi Vantara LLC in the U.S. and other countries.