

# Netflix reimagines discovery and governance at scale

A unified global catalog powers self-serve discovery, governance, and soon AI agents.



“Now, our data teams can discover what they need, understand the full impact of changes before they make them, and self-serve governance questions that used to require deep legacy knowledge. It's transformed how we operate at scale, and it's positioning us perfectly for the agentic AI future we're building toward.”

**Nitin Sarma**

Sr. Engineering Manager, Data Discovery, Governance & Experiences at Netflix

SIZE

**10,000+ employees**

DATA STACK

**AWS and homegrown data platform**

SOLUTION

**DataHub OSS**

USE CASE

**Discovery, Governance**

## The topline

### Challenge

Expansion into new verticals created exponential data growth, fragmented discovery, siloed onboarding, and complex impact analysis.

### Solution

Built a unified global catalog on DataHub to connect all data, ML, and software entities.

### Impact

Achieved unified discovery, cross-domain impact analysis, self-serve governance, and a scalable foundation for agentic AI.

## Deep dive

### Challenge

Netflix's expansion into ads, live events, and games created exponential data growth and three critical problems along with it.

- 1. Fragmented discovery** forced new engineers to rely on colleagues, Slack threads, and tribal knowledge just to find reliable data.
- 2. Governance gaps** left ownership, classification, and compliance accountability unclear as teams grew and reorganized.
- 3. Siloed impact analysis** meant engineers had no visibility into downstream effects when making changes, a growing risk as GenAI systems became first-class citizens at Netflix.

# With DataHub, Netflix built a single source of truth.

The company needed a fundamental shift: from siloed, team-specific thinking to a unified, global approach that served all technical practitioners.

"DataHub has become the central nervous system for discovery and governance at Netflix... , And we're really looking forward for more and more to be built on top of it."

**Nitin Sarma**

Sr. Engineering Manager, Data Discovery, Governance & Experiences at Netflix



## Solution

Netflix chose DataHub as the foundation for a global catalog spanning data, ML, and software entities across the company.

Four capabilities made DataHub the clear choice: entities and relationships modeled as first-class citizens; a unified ingestion framework that eliminated custom builds; automatic lineage surfacing through explicit relationship modeling; and coverage insights that translate directly into governance improvements.

To balance centralized governance with team autonomy, Netflix implemented GraphQL federation, preventing the catalog from becoming a bottleneck while maintaining unified discovery across domains.

The result: a single, authoritative view of every technical asset at Netflix.

## DataHub Impact

Key outcomes included:

- Eliminated support bottlenecks with self-serve discovery.
- Accelerated onboarding and time to productivity for new engineers.
- Established continuous governance monitoring with proactive gap identification.
- Positioned Netflix to deploy AI agents on a trusted, unified context layer.
- Reduced storage costs through automated retention policies and dataset cleanup.
- Unified cost ownership visibility across systems, data, and ML.
- Accelerated migrations, system upgrades, and ownership transfers.

## Start your own success story with DataHub

### Join our open source community

Connect with thousands of practitioners at [datahub.com/community](https://datahub.com/community)



### See the magic of DataHub Cloud

Schedule a custom demo at [datahub.com/demo](https://datahub.com/demo)

