

ESCAPE DATA PITFALLS: MAKING AI WORK IN HR TECH

A guide to building the data foundation
your AI needs to succeed.

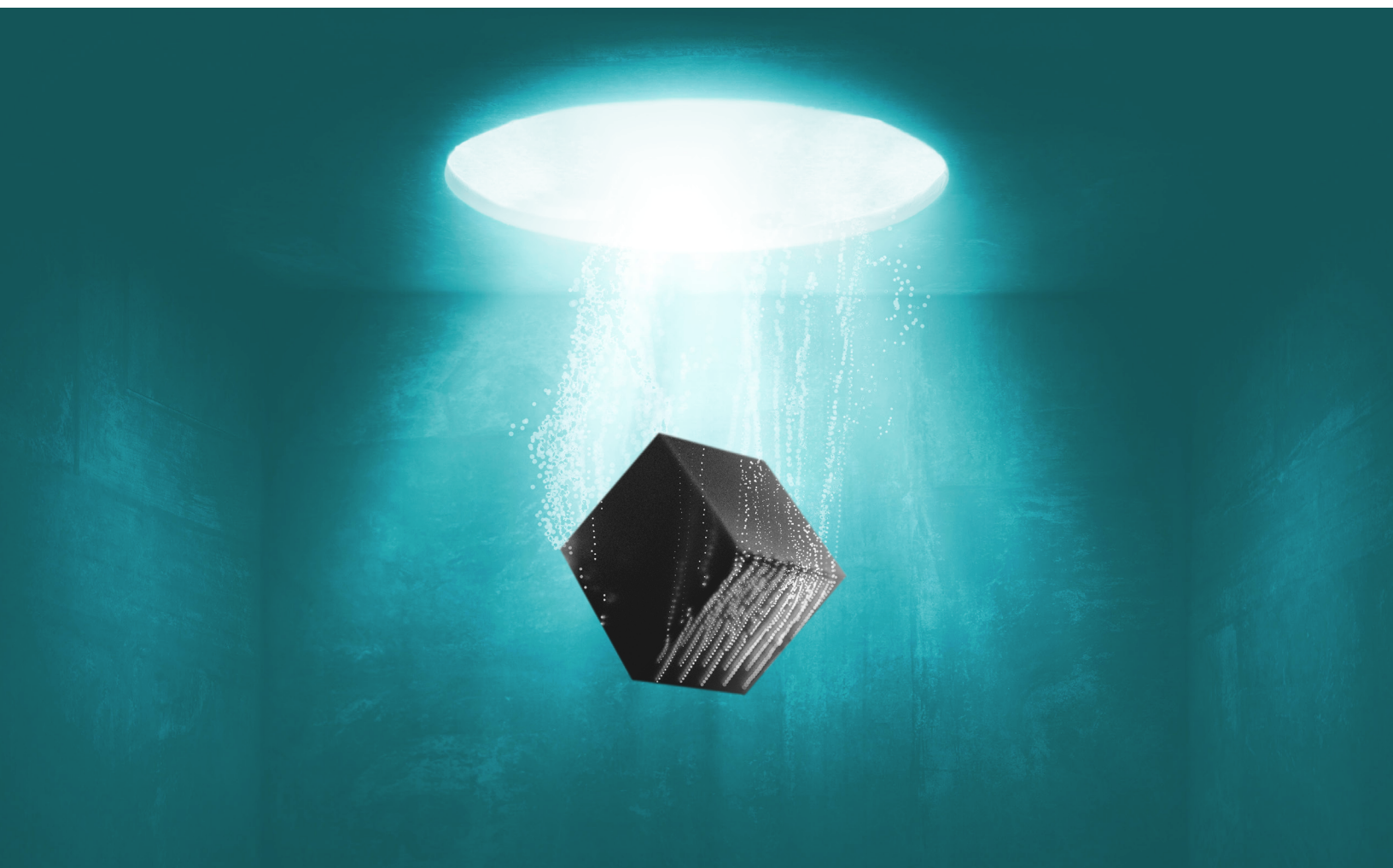




Table of contents

Executive summary	02
AI demands a new approach to data	04
Why AI projects fail in HR tech	05
The 5 data pitfalls holding back AI in HR tech	06
The consequences of an incomplete data foundation	09
Benefits of a strong data foundation	14
Visier's approach to AI	15



Executive summary

HR is moving fast toward conversational analytics, copilots, and AI agents.

Leaders expect immediate answers about people and the work they do, managers want guidance in the flow of work, and HR technical teams want reliability and control.

But AI doesn't start with chatbots or agents. It starts with data.

While general AI data readiness is widely discussed, very few resources focus on the specific implications with people data, which has its own nuances and challenges.

Having worked with over 70,000 customers and supported more than 2 million AI users, Visier brings unmatched expertise in workforce analytics. We understand what works—and what fails—when it comes to making AI successful within HR.

This guide exposes the readiness gap many organizations discover when they try to deploy AI within HR and lays out a pragmatic roadmap to fix it. The future of HR decision-making depends not just on what your AI says, but on what it actually understands about your workforce.

THE WORKFORCE DATA AI-READINESS GAP





Everyone wants AI, but few are ready

Over the past two years, AI has dominated every boardroom. ChatGPT has redefined user expectations. HR teams want faster insights, end users want personalized guidance, and CEOs expect productivity gains.

Naturally, technical teams who support HR are feeling the pressure. The roadmap is always shifting. Conversations that used to be about dashboards and integrations are now about copilots, assistants, and AI agents.

But launching AI capabilities is not as simple as plugging into an LLM. AI isn't magical, it's mechanical. It's only as good as the data it works with.

HRIT, HRIS, and data teams must answer two questions: How do we prepare our enterprise workforce data for AI?, and how do we account for the unique requirements and challenges of people and work data?

What makes workforce data different from regular business data

The data stored by HR tech providers about people and work is:

- ✓ Highly sensitive
- ✓ Often fragmented across several modules
- ✓ Stored in different formats
- ✓ Sometimes contradictory
- ✓ Poorly described (or not described at all)
- ✓ Missing key relationships and metadata
- ✓ Governed inconsistently, if at all

AI demands a new approach to data

For years, workforce data was used primarily for reporting and dashboards—descriptive, backward-looking insights, and in some cases, forward-thinking predictive analytics (think machine learning models).

While it wasn't ideal, the way insights were consumed was always predictable. Dashboards were built with specific use cases in mind, and natural language questions were never even a possibility.

Those use cases tolerated data inconsistencies, duplication, or technical debt, because your team could always define the inputs and outputs and have complete control over how reports and insights are delivered.

Now, AI demands something else entirely. Just in the last couple of years, we've already seen UX shift drastically from rigid interfaces to:

- Natural language chat
- Conversational decision support
- Predictive, contextual insights
- Automated workflows and assistants

These new experiences are live, interactive, and data-hungry. They rely on structured, contextual, trustworthy data.



Why AI projects fail in HR tech

You may not think your data is a problem now, but once you get into your AI project, you may discover deeper issues that stand in the way of your progress.

AI project failures don't happen in a vacuum. They're not the result of one bad sprint or a buggy integration. They're the byproduct of systemic weaknesses, and data is often the weakest link.

While organizational factors like unclear objectives, inadequate infrastructure, or unrealistic timelines all contribute to AI project risk, the data layer is where most HR tech vendors stumble. [Gartner](#) reports that 60% of AI projects lacking AI-ready data will be abandoned, while [The Center for Business Analytics](#) found that 90% of AI projects fail in analytically immature organizations where data quality, availability, and structure are insufficient. Furthermore, Deloitte, McKinsey, Forrester, and KPMG consistently identify data quality and governance as the primary obstacles to successful AI implementation.

These numbers aren't just alarming, they're a flashing red warning for technical teams under pressure to "add AI" to their tech stack. Without an AI-ready foundation, these initiatives are doomed before the first line of code is written.

60%

ABANDONMENT RATE OF
AI PROJECTS THAT LACK
AI-READY DATA

90%

FAILURE RATE OF AI
PROJECTS IN ANALYTICALLY
IMMATURE ORGS





The 5 data pitfalls holding back AI in HR tech

1 Poor data quality

According to [Fivetran](#), underperforming AI models built using low-quality or inaccurate data cost companies up to 6% of annual revenue, on average.

Your model is only as good as the data it learns from. Within HR, data quality problems are magnified:

- Headcount data from manual updates leads to inaccurate turnover predictions
- Incomplete job descriptions or performance histories can result in irrelevant coaching recommendations
- Duplicate employee records provide for distorted organizational views and double counting

This isn't just a data science problem. It's a user trust problem. If a manager asks a genAI assistant for insights on headcount, and the answer is clearly wrong, that feature is dead on arrival.

Headcount

			Headcount	
Q3				
Q4			308	✓
Q1				

What was headcount in Q4 last year?

Last year's headcount at the end of Q4 was 297 ⚠️

Ask me anything...

| AI showing inaccurate data leads to instant distrust.

NO TRUST = NO USAGE = NO SECOND CHANCES



Who are my top performers this quarter?

 **Sorry, I don't have enough data to answer that question**

Ask me anything...

Insufficient data leads to dead ends, or in the worst case, AI hallucinations.

2 Insufficient data volume

Most HR tech stacks have the data volume, but it's typically scattered, siloed, and unstructured. For example:

- One database holds attendance records. Another defines the org chart. Surveys live elsewhere.
- Key context, like reason codes, relationships, or event history, is missing or fragmented.
- Vendors can't access the volume of data they need to train or ground AI responsibly because their datasets are limited to their customer base.

This is why having an integrated, semantic model with access to benchmarks from millions of employee records is critical. It connects the dots across workforce domains, providing clean, connected, queryable data for AI to reason over.

3 Outdated or unavailable data

Even if your HR data is clean and well-modeled, AI will fail if it can't access the right data at the right time.

When data pipelines lag behind business activity, AI outputs quickly become inaccurate or irrelevant. For example, while weekly updates may be sufficient for high-level workforce planning or long-term trend analysis, nightly updates are often needed for headcount or organization structure, and even more frequent updates are essential for time and attendance or workforce scheduling.

If your LLMs and agents can't access fresh, permissioned data streams, they're forced to "guess," leading to hallucinations, stale insights, and broken user trust.

 **Staffing Shortage**
✕

Scheduled staff for **Day Shift** at **Columbus, Indiana Assembly Facility** is below the coverage threshold.

View Dashboard

 **Stale data can trigger misfired alerts and prompt the wrong actions.**

Out-of-date data leads to bad decisions resulting in lost trust, lost productivity, and lost revenue.



⚠️ Red flag: Recommendations reinforce historical inequities

Who should I promote on my team?

Based on performance history and promotion likelihood scores, Kyle, Steve, and John are the best candidates for promotion.

Ask me anything...

Skewed or incomplete data results in inherited AI biases, influencing recommendations for the end user.

4 Inherited data bias

Nearly half of employed U.S. job seekers (49%) believe AI tools used in recruiting are more biased than their human counterparts, according to a [recent poll](#) conducted by the American Staffing Association Workforce Monitor.

AI bias doesn't begin with the model. Its foundations are in skewed or legacy HR data. For example:

- Historical hiring decisions encoded with bias get baked into AI recommendations.
- Systems trained on incomplete demographic or performance data amplify inequality.
- Representation gaps in training data produce biased output, reputational risk, legal exposure.

Addressing bias requires visibility and control over data lineage, definitions, and intent. This is only possible with a structured metadata layer and domain-aware models grounded in a massive dataset of employee and company records—far more than any single HR tech vendor would realistically possess among their customer base.

5 Weak data governance

HR data is among the most sensitive data an organization handles. If AI is going to work, it must respect:

- **Dynamic role-based access:**
Who can see what, as roles shift
- **Security and compliance rules:**
GDPR, HIPAA, internal policies
- **Explainability requirements:**
How a recommendation was made

Too many technical teams skip this. They simply plug in a genAI model and hope for the best. The result? Misfired insights, data breaches, or compliance nightmares.

Visier's platform is built to avoid these pitfalls with explainable metrics, anonymized data from over 35 million employees, permission-aware querying, and AI that never answers outside of what a user is allowed to see.

🚫 Sensitive data exposed — user lacks necessary permissions

How many employees are on medical leave?

There are 3 employees on medical leave: Angela, Priya, and Mark.



Ask me anything...

Improperly securing data and permissions at the LLM level can result in serious data breaches.



The consequences of an incomplete data foundation

When technical teams rush to ship AI features without the right foundation, they're putting themselves at risk of several consequences:

- **Misinformed insights:** Generic LLMs hallucinate when they don't understand org structure, job taxonomy, or HR context.
- **Lost trust:** Users get answers that don't match what they see in their dashboards or workflows.
- **Security risks:** Without role-based access and lineage tracking, users could inadvertently be exposed to sensitive or non-permissible data.
- **Wasted investment:** Time and budget go into building assistants users don't trust and customers actively disable.
- **Slower roadmaps:** Technical teams get buried in rework and data cleanup instead of building new capabilities.

These occur largely because most HR data is not AI-ready.

Differences between traditional and AI-ready workforce data

TRADITIONAL WORKFORCE DATA	AI-READY WORKFORCE DATA
Flat tables from disconnected modules	Unified, cross-domain model with job, org, and event history
Siloed by function (e.g., performance, recruiting, time)	Linked across the employee lifecycle
Poor data labeling (e.g., attrition code)	Metadata-rich, semantically tagged (e.g., voluntary attrition)
No context on reporting lines or org structure	Contextualized with full hierarchy and team dynamics
Stale and outdated snapshots	Up-to-date and current, in some cases, real-time data
No audit trail for data changes	Full lineage and explainability for every metric
Inconsistent access controls	Role-aware governance (e.g., manager vs. admin views)
Generic LLM prompts with no grounding	AI grounded in real employee data from a massive dataset



The unpredictability of AI's UIs is a compounding factor

To add to the data problem, today's implementation of AI is still far from what the world will settle on. We started with text-based chat because LLMs are built around words, but chat-based interaction is not the final destination.

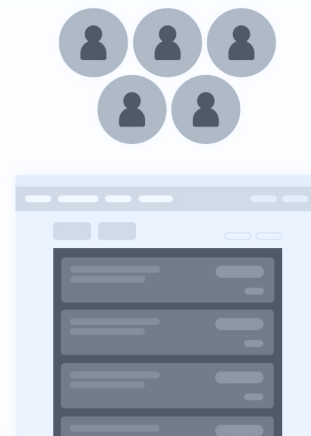
Today, widely accepted user experience (UX) principles are, generally, a collection of static controls on a page serving up data and facilitating interaction, tied to a bolted-on chatbot. Concepts like generative UI are breaking these rules.

The uncertainty around how users will interact with AI and apps in general makes it increasingly difficult —and risky— to make any big bets on UI for AI.

One thing is for certain, though. Future UI principles will rely on a solid data foundation to serve up AI to users. That data foundation is completely within your technical team's control.

TODAY

Same interface for everyone



FUTURE WITH GENUI

Personalized interface for each user



One of many theories of the future of UI is GenUI, a concept coined by [Nielsen Norman Group](#).



The 4 Cs of AI-ready workforce data

To avoid these traps and ensure AI projects succeed, data must meet four conditions:

1. CONSOLIDATED

All relevant people and work data consolidated in one data model, at a level of freshness acceptable to the use case

2. CLASSIFIED

Data is described semantically for AI to understand it

3. CONTROLLED

Secure, permission-aware, and explainable

4. CONTEXTUALIZED

Enriched with relationships, a large sample of anonymized data, time logic, and business rules

VISIER'S 4 C'S OF AI-READY WORKFORCE DATA



visier



4 steps to AI-ready workforce data

1

Unify and normalize your data

Your data probably lives in at least half a dozen places across a number of databases developed at different times, or perhaps acquired as part of your suite strategy. This presents a challenge when delivering both analytics and AI. If the data isn't combined, AI will return erroneous results.

To make data usable for AI:

- Use a prebuilt ETL framework to extract and transform data across silos
- Standardize fields, resolve duplicates, and normalize formats
- Align to a semantic workforce data model (jobs, orgs, locations, events, etc.)
- Employ the appropriate synchronization method (e.g. nightly updates for headcount)

Visier's transformation engine is optimized for workforce data and the temporal nuances that come along with it, giving you full control over how data is ingested into and defined in the model.

Create rule

Display name*

Pay Level Increase

Rule type*

Business rule

Multi-subject rule

Description

Whenever a change to the Pay_Level property happens and the current value of the Pay_Level property is greater than what it was perviously, insert a Promotion event with the Promotion_Reason property set to "Pay Level Increase"

Formula

1

when changeIn(Pay_Level) and (currentValueOf(Pay_Level) > previousValueOf(Pay_Level))

2

add Promotion(Promotion_Reason := "Pay Level Increase")

CANCEL

CREATE

Business rule definitions inside Visier

2

Apply a workforce-specific data model

Generic data models don't work for HR. You need a structure that reflects:

- Roles and reporting lines
- Employee types and relationships
- Unique events like onboarding, promotion, attrition, or leave
- Time-based analysis (point-in-time vs. period-end)

This lets your HR tech stack understand the world of work, not just store it.

Visier's extensible data model has:

- Predefined objects, metrics, and relationships
- Built-in business logic and calculations
- Ability to be customized for your specific domain extensions

Analytic Objects > Employee

Object Type

All

Properties

Dimensions

Concepts

507

483

15

9

DISPLAY NAME	OBJECT TYPE	DATA AVAILABLE
Employee ID	Simple Property	Yes
Age	Calculated Property	Yes
Full Name	Calculated Property	Yes
Sales Quota	Simple Property	Yes
Territory	Leveled Dimension	Yes

Visier's data model covers every aspect of the employee lifecycle and beyond



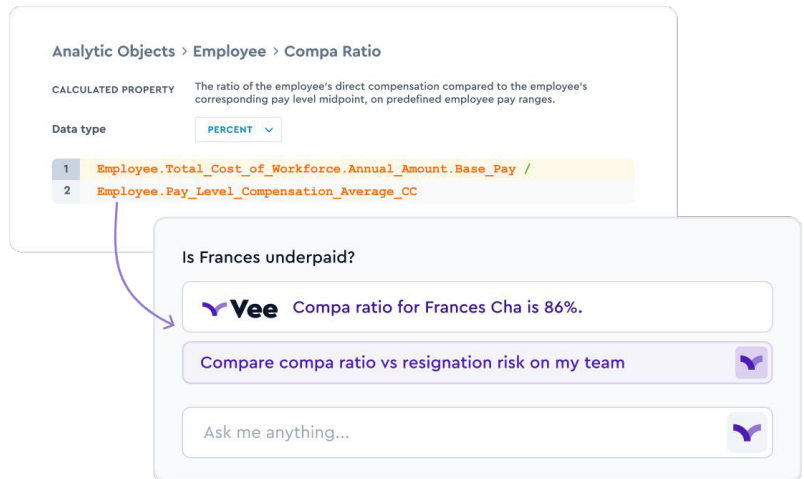
3

Enrich data with context and metadata

AI doesn't just need clean data. It needs described data, and a lot of it. This requires:

- Tagging columns and entities with human-readable labels
- Mapping synonyms and intents for natural language understanding
- Capturing lineage and transformations for traceability
- Using anonymized, benchmarked data to ground more reliable, bias-resistant models

Visier's model includes rich metadata and historical benchmarks drawn from over 35 million anonymized employee records, giving your AI an in-depth, grounded understanding of how people data behaves.



Metadata inside Visier's data model shown through Vee, Visier's genAI Assistant for HR

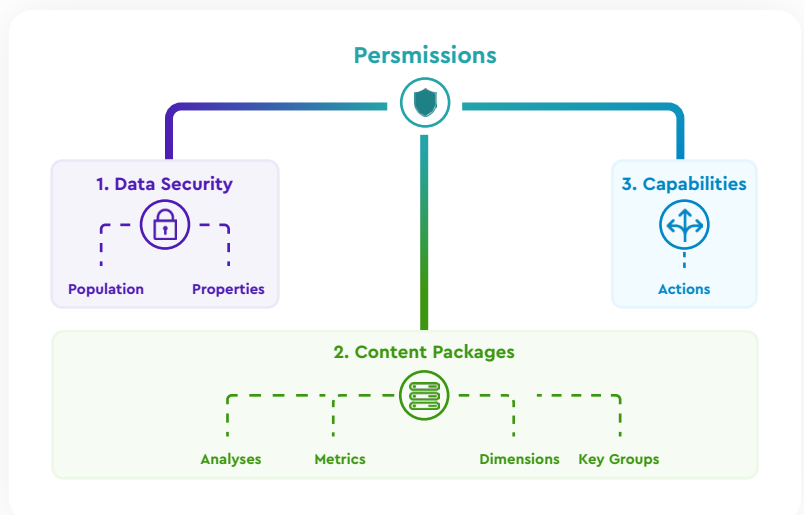
4

Enable AI responsibly through a trusted platform

AI readiness isn't just about structure, it's about control. Especially in HR, where data privacy and governance stakes are high, data must be:

- **Permission-aware:** Answers should reflect what each user is allowed to see
- **Explainable:** Every metric and insight needs a clear, traceable origin
- **Compliant:** Must be compliant with internal policies, GDPR, HIPAA, and other industry standards
- **Monitored and auditable:** Changes in data should be logged and reviewable

Visier's end-to-end security model ensures every query respects role-based access, security controls, and lineage with explainability baked into every AI response.



Visier's dynamic role-based security model



BENEFITS OF A STRONG DATA FOUNDATION



FOR HR TECHNICAL TEAMS

- ✓ Ship AI capabilities faster with fewer false starts
- ✓ Focus engineering on innovation, not data cleanup
- ✓ Confidence that AI is grounded in definitions & permissions
- ✓ Data is ready for AI, whatever the future of UX becomes



FOR END USERS

- ✓ Workforce questions can be asked in plain language
- ✓ Managers and employees get instant answers they trust
- ✓ HR leaders act confidently on predictive insights
- ✓ UX feels intuitive, contextual, and intelligent



FOR THE BUSINESS

- ✓ Faster time to value
- ✓ Proven ROI
- ✓ Higher workforce productivity
- ✓ Make better people decisions



What led us down this path was the realization that data is critical to everything we do. We knew we needed to make significant strides in this area, and given our resource constraints, we believed Visier would be a quick way to accelerate our progress. It offered an effective means to integrate all our data sources and tell a comprehensive story.



Shannon Rutledge
Director, HR Data & Analytics



T.RowePrice



Visier's approach to AI

Implementing AI in HR technology isn't just about building features, it's about making sure those features work consistently, safely, and accurately across thousands of unique organizations, hierarchies, and use cases.

That's exactly what Visier was built to solve.

Over the past 15 years, we've developed a data, analytics, and AI platform purpose-built for people and work data, solving the foundational problems most HR tech vendors now face as they try to add AI to their roadmap.

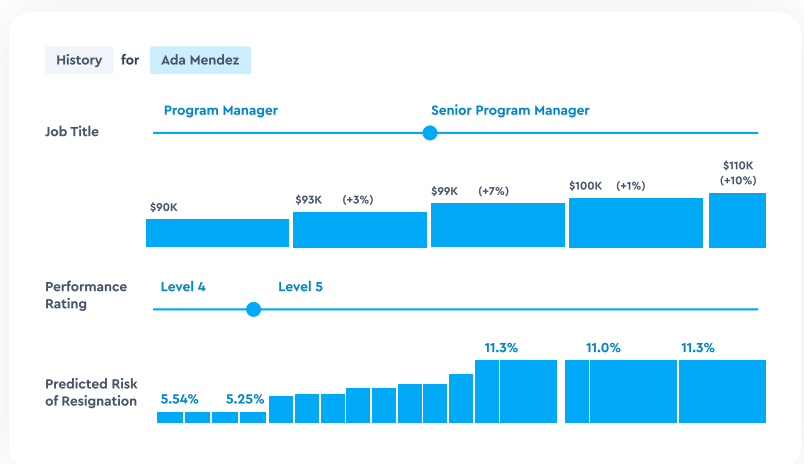
Here's how we help you go from messy source data to trusted, AI-ready insights your stakeholders will actually believe and act on.



Data engineering built for HR complexity

Visier connects to any database or data warehouse and unifies the data through intelligent ETL pipelines using:

- ✓ Prebuilt connectors and data-in APIs
- ✓ Customizable batch data sync schedule
- ✓ Automated transformation, de-duplication, and value inference
- ✓ Time-aware logic built in (slowly changing dimensions, event-based snapshots)



Visier maps complex time-based people journeys out of the box.

WHY DOES IT MATTER?

You can't build reliable AI from raw, siloed data. Visier ensures your source data is clean, normalized, and ready for downstream modeling, removing the first major blocker to AI readiness in HR systems.



Workforce-specific data model

Visier's people data model understands the HR domain out of the box with thousands of metrics, defined objects (like job, location, and tenure), and built-in business logic.

- ✓ 2,000+ metrics across 11 workforce domains
- ✓ Predefined relationships (e.g., manager to direct reports)
- ✓ Event tracking and time-based analysis (e.g., "as of last quarter")



A few elements of Visier's data model.

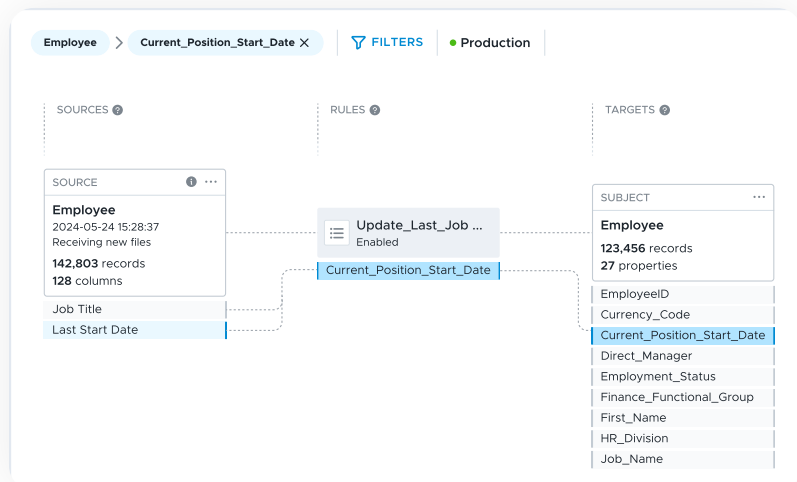
WHY DOES IT MATTER?

AI needs structure. Without a workforce-specific model, your AI doesn't know what "attrition" means or how to calculate "time to fill." This layer gives your AI the vocabulary and logic of HR—no manual mapping needed.

An AI-ready data warehouse built for people and the work they do

Visier comes with a governed, performant warehouse optimized for people data, with embedded logic for metric computation, filtering, and cohort definition.

- ✓ High-speed querying and pre-computation for responsiveness
- ✓ Fully governed tables and API access for embedded use
- ✓ Scalable across millions of employee records and tenants
- ✓ Cost predictability with PEPM-based billing



Visier's data warehouse comes with all of the relationships and mappings needed to deliver workforce AI out of the box.

WHY DOES IT MATTER?

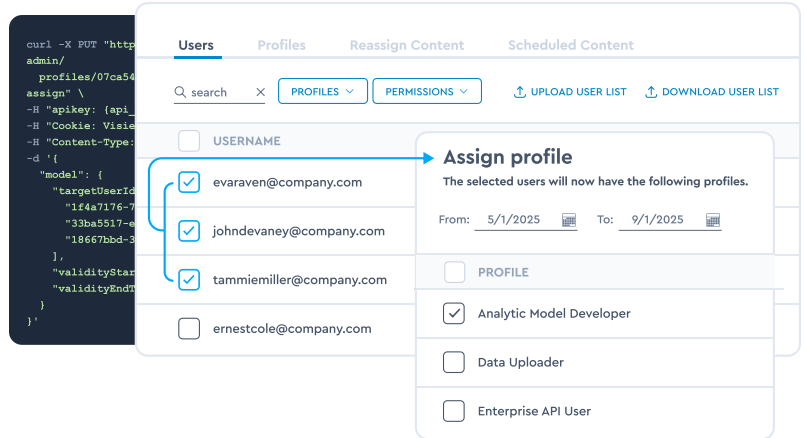
Most AI use cases fail on latency or scale, especially when querying large, complex datasets. Visier ensures your AI can deliver reliable, permission-safe answers without straining your core systems or ballooning costs.



Enterprise-grade security and governance

End-to-end governance is baked into the platform, including access controls, auditability, and regulatory compliance.

- ✓ Dynamic role-based and row-level permissions
- ✓ Separation of data and LLM
- ✓ SOC2, GDPR, HIPAA-aligned controls
- ✓ Explainability and lineage are built in



Users can manage their Visier deployment via Studio or through automated API calls. Shown above: user management.

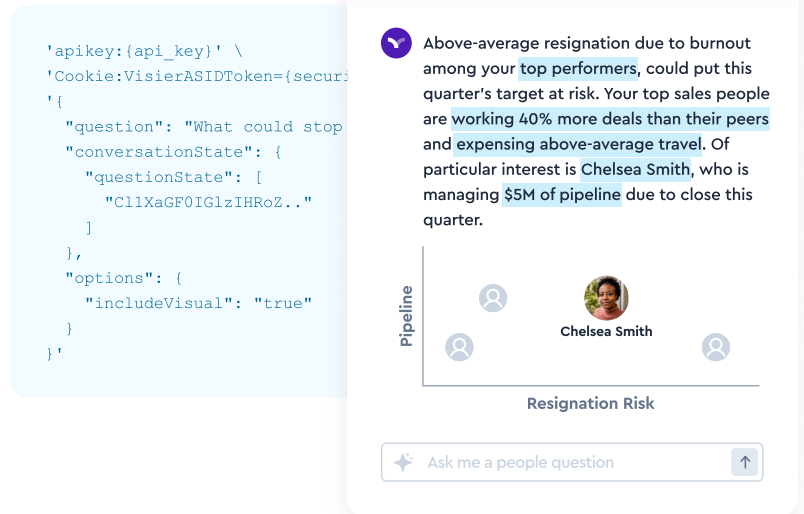
WHY DOES IT MATTER?

If AI answers something it shouldn't—even once—your compliance and security is at risk. Visier ensures your AI only answers what users are authorized to see, with full traceability.

Embedded AI that actually works

Visier powers genAI experiences that are grounded in workforce data, not just chatbots bolted onto generic LLMs. Here's how:

- ✓ **Vee**, an AI assistant that understands HR analytics
- ✓ Answer generation grounded in real metrics, not hallucinated logic
- ✓ AI agents for proactive decision support and automation (coming soon)
- ✓ Directly accessible in the flow of work via Microsoft Copilot or Slack



| Vee Visier's GenAI assistant delivering trusted insights.

WHY DOES IT MATTER?

A chatbot that guesses is worse than no chatbot at all. Visier makes your AI useful from day one because it's grounded in real definitions, real permissions, and real business context.

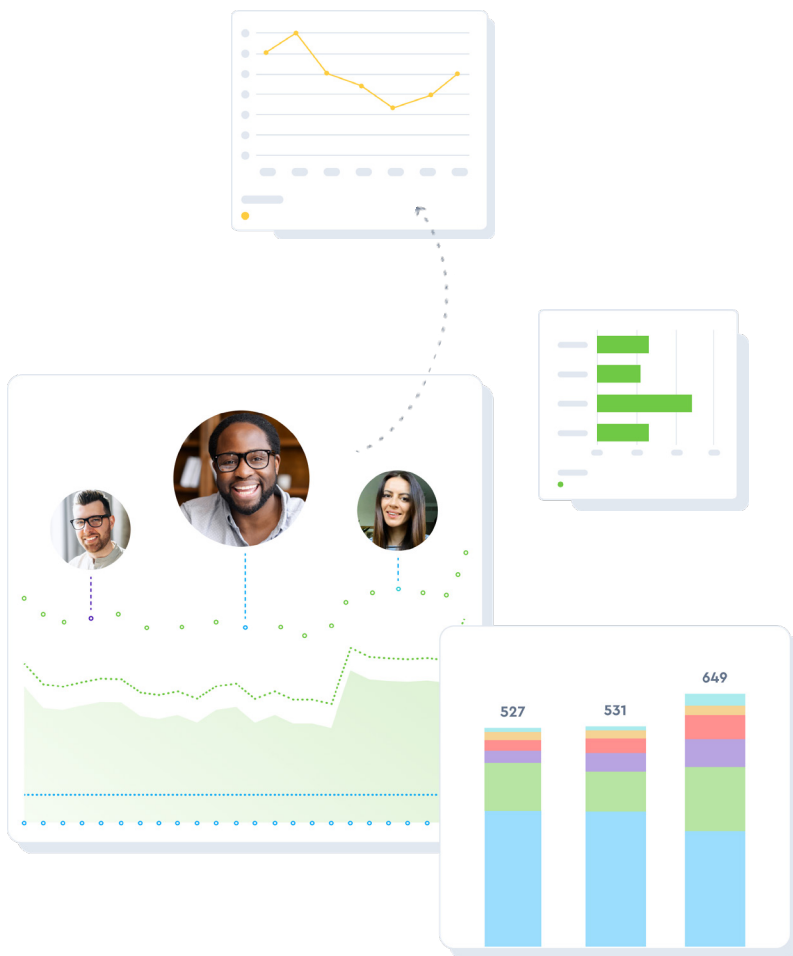


Why HR tech vendors choose Visier

We've spent over a decade solving the data problems most AI initiatives are just now uncovering and we've done it at scale, across thousands of organizations and tens of millions of employee records:

- ✓ 36M+ anonymized employee records for benchmarking, bias mitigation, and training
- ✓ 65,000+ organizations' data flowing through our platform
- ✓ Proven security, reliability, and extensibility at enterprise scale
- ✓ A complete platform: ingestion, modeling, analytics, AI, governance in one stack

You don't need to build your AI foundation from scratch. With Visier, you get to market faster, avoid common pitfalls, and deliver AI features your stakeholders can actually trust and use.



Let's make your data AI-ready

AI is here. Expectations are growing. And you can't afford to get it wrong.

Let's talk about how to:

-  Unify and model your people data
-  Build metadata and governance into your foundation
-  Enable genAI, AI agents, and intelligent UX without the data debt

Contact us to learn how we help HR technical teams accelerate their AI journey: www.visier.com/demo



About Visier

Visier is the global leader in Workforce Intelligence that powers every people decision. Our award-winning, agentic AI technology surfaces the insights leaders need to plan, decide, and act with confidence in the moments that matter most. As the market leader in people analytics, workforce planning, organizational design, and manager effectiveness solutions, we fuel smarter decision-making for organizations across the globe. Our mission is to help businesses lead with insight at scale as they continuously transform.

Founded in 2010 by the pioneers of business intelligence, we have over 70,000 customers in 75 countries—including enterprises like BASF, Panasonic, Domino's Pizza, Experian, Amgen, eBay, and Ford Motor Company.

To learn more about Visier, visit www.visier.com.



visier.com | info@visier.com | 1.888.277.9331

©2025 Visier Solutions Inc. All rights reserved.

Visier, Visier People, and the Visier logo are the trademarks and/or registered trademarks of Visier, Inc. and Visier Solutions Inc. All other brand names and logos are the trademarks of their respective owners.