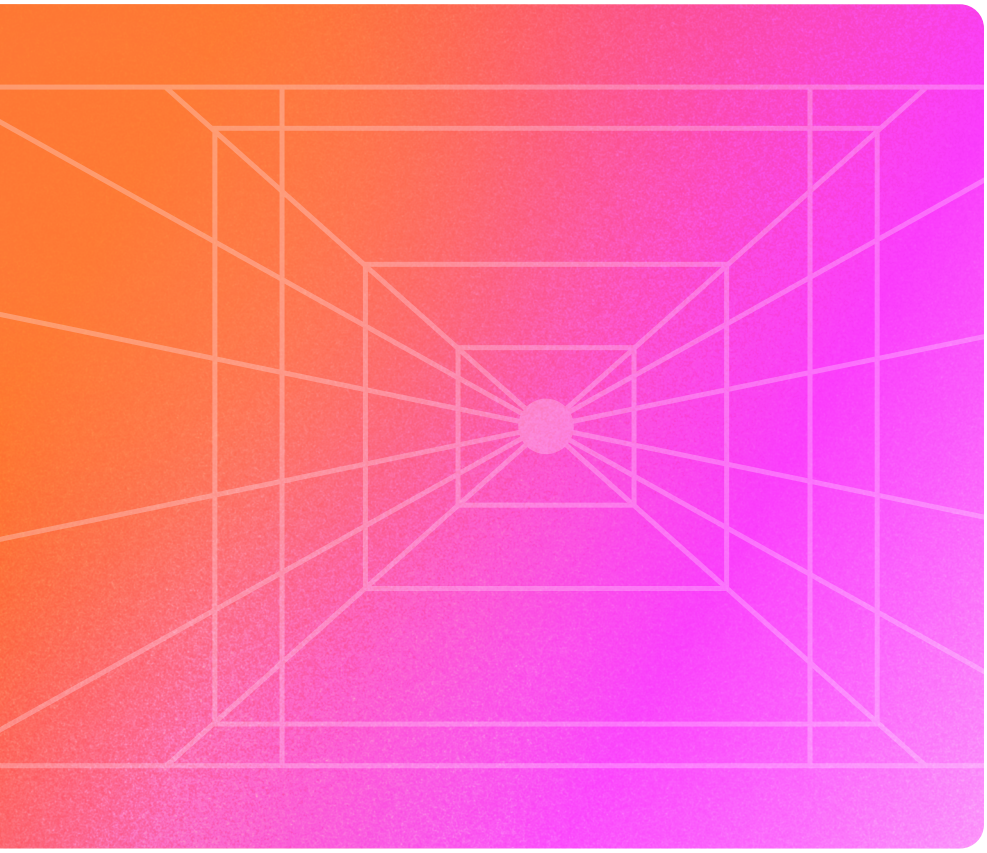
Lumos

AGENTIC ROLE MINING

A Path to Smarter Access Management



Who this guide is for

Security, IT, and IAM teams who want to automate role mining.

What you'll learn

- Why you need role mining
- What makes role mining hard
- Why traditional role mining falls short
- A three-step approach to accelerate role mining
- How Lumos and Albus (our AI identity agent) automate the role mining process

Stop Working in the Dark

Role mining promised clean RBAC, but it has not delivered.

On paper it sounds easy: write rules that say what a person should get based on team, location or level. In real life, it is more difficult than it looks. This is where roles get messy, policies become over-permissioned, audits drag on and security gaps happen.

This short guide does three things: explains the foundations of RBAC and role mining, highlights the flaws of traditional role mining, and shares a process to make role mining continuous and intelligent.

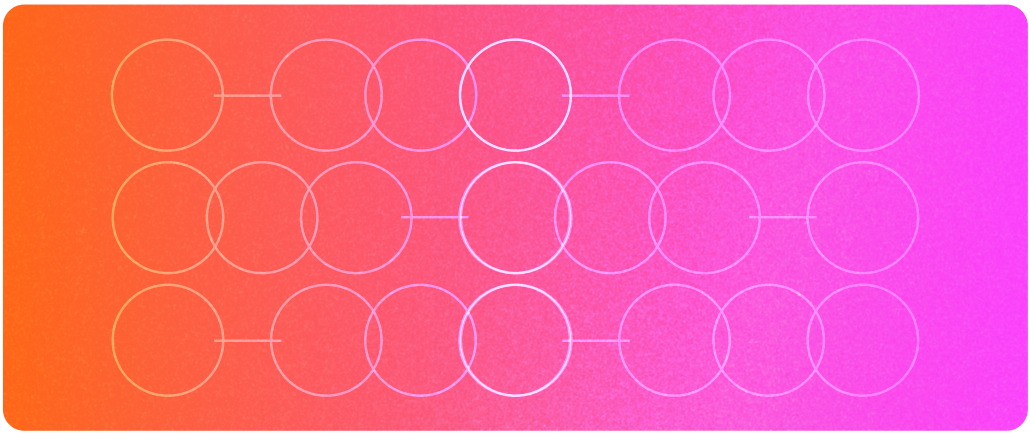
By the end, you will know how to go from brittle rules to living roles that reflect your business today; and prove least privilege, speed up access, and walk into audits with confidence.

Why You Need Role Mining

Most access environments aren't architected with precision. They've evolved over years of quick fixes, urgent requests, and "just give them access so they can get their work done" decisions. This leads to layers of rubber-stamped permissions with no context, role bloat, and access sprawl.

Role mining is how you flip the switch.

It gives you full visibility into your access landscape. It reveals where your access structure is solid, where permissions have sprawl, and where outdated access creates unnecessary risks.



Why Role Mining Is Hard

If you've ever tried to build a clean RBAC model and ended up knee-deep in conflicting titles, inconsistent access data, and overlapping entitlements – you're not alone. On paper, role mining sounds like a data exercise: group similar users and define reusable roles. However, in reality it is a complex, high-effort process.

You interview managers for business context. You export data from your systems. You merge CSVs together. You try to spot patterns through noise. The business shifts, HR data is messy, and the rules you wrote six months ago no longer match the current organization. You pay consultants to help and projects stretch from weeks to quarters.

A large enterprise can spend seven figures a year just to maintain roles, while smaller teams stall with their RBAC initiative because they cannot staff it. Meanwhile, access drifts – people keep what they no longer need, new hires wait, and auditors ask simple questions that take days to answer.

Role mining remains one of the hardest challenges in identity governance: data is fragmented, human context is missing, and manual effort to keep roles current rarely scales.

Messy HR Data

Your HRIS should be the source of truth, but if titles are inconsistent, reporting lines are unclear, or worker types are mislabeled, it becomes a source of confusion. Role suggestions built on messy data aren't usable.

Access Sprawl

Over time, entitlements accumulate and never get turned off. People move and access stays. The fix isn't just listing who has access, it's knowing what's actually used and right-sizing accordingly. Keeping least-privilege intact is critical.

Dynamic RoleChanges

Teams evolve, organizations grow and shrink, responsibilities flex, and hybrid roles appear. Static roles in a dynamic business become instantly obsolete. Policies age out quickly and need continuous assessment and enforcement.

Ownership Gaps

Without clear ownership of roles, everything bottlenecks in IT. This quickly becomes impossible to maintain as roles overlap and you're unable to update your own access patterns. Role mining isn't one-and-done cleaning; it's a continuous and risk-aware governance practice that needs intelligence, oversight and smart automation.

What is RBAC?

Role-Based Access Control (RBAC) groups access based on job functions. When implemented correctly, RBAC ensures that everyone can access what they need, while minimizing potential vulnerabilities through well-designed controls.

What is ABAC?

Attribute-Based Access Control (ABAC) adds flexibility to managing who can access your company's systems, apps, and data. ABAC considers multiple attributes to make access decisions.

Why Are RBAC and ABAC Important?

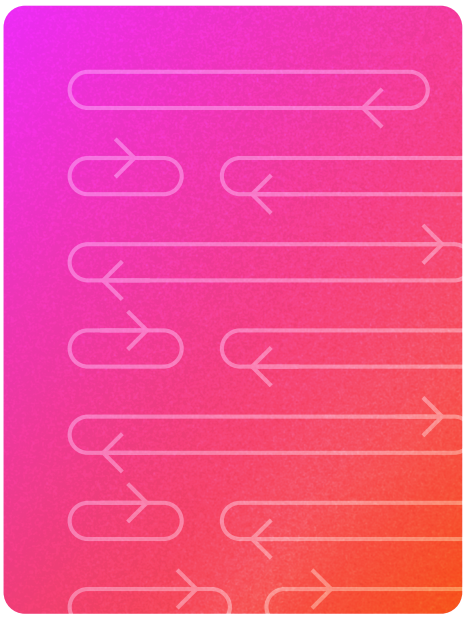
RBAC gives you order and structure. ABAC gives you flexibility and context. Together, they make least privilege possible at enterprise scale.

THESE TWO CONTROL MODELS:

- Reduce Lateral Movement
- Shrink Attack Surface
- Simplify Access Reviews
- Make Users Productive

Legacy Solutions Leave You in the Dark

Legacy role mining tools promised you clean roles, smooth governance, and a faster path to least privilege. But these solutions often fall short of accelerating the role mining process.



Hard to Ingest the Full Picture

Legacy platforms struggle to ingest fragmented data environments. Even when synced, data is stale or partial, such as: group-level without clean entitlements, assignments without usage, and complex and one-off legacy role architecture with missing context.

Incomplete Picture with Assignment-Only Insight

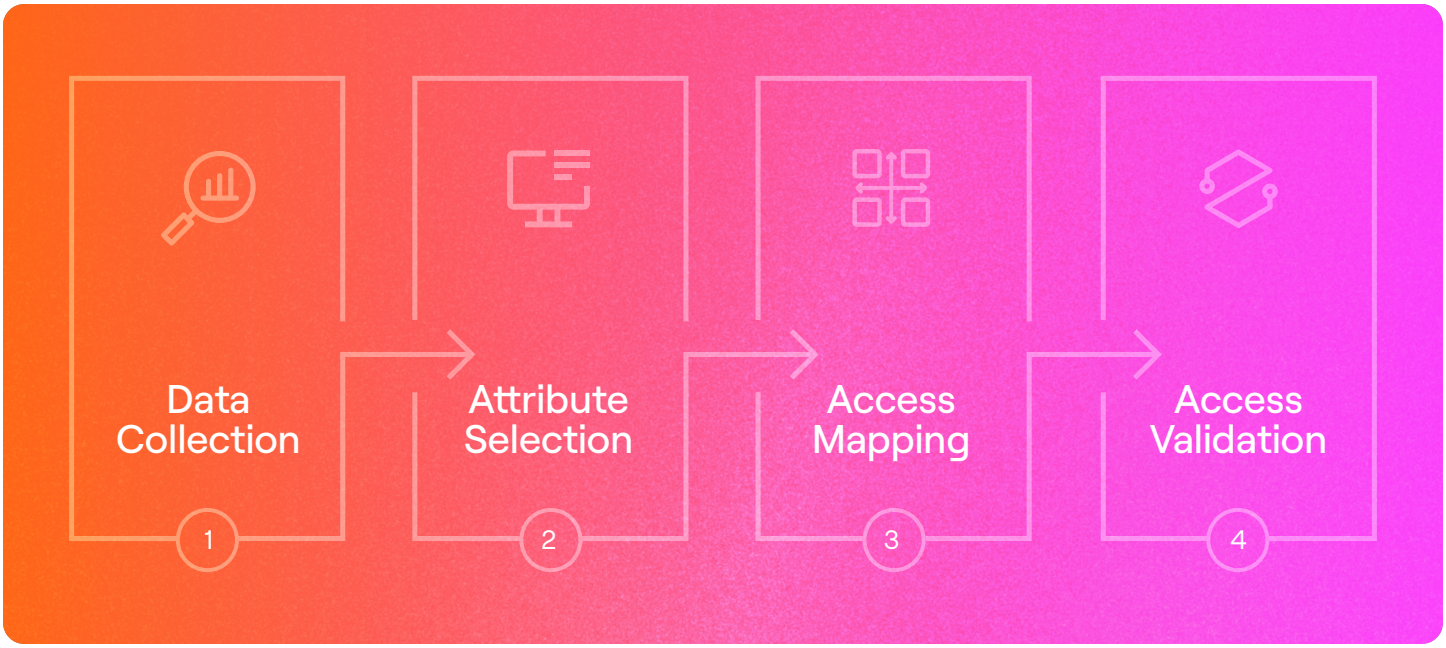
Without usage context and privilege depth, you're stuck guessing at relevance and risk. Incomplete role mining does not lead to strong access controls.

Missing Business Context

Legacy tools completely overlook the business logic and what your critical initiatives are – you get data that shows who, but never why. Tools that ignore business context suggest "cleanups" that break workflows and users end up spending time filing IT tickets.

Stuck in a Manual Loop

Legacy tools force you into endless manual work. No clustering, no recommendations, no intelligent grouping. Just rows of users and mountains of entitlements and permissions.



4-STEP BLUEPRINT TO ROLE MINING

Agentic AI Approach

Traditional role mining requires tedious manual data wrangling and spreadsheet-heavy analysis.

GenAI and agents change this approach. They make sense of messy data, manage complex actions at scale, and stay relevant by learning over time. Go from manual access hand-holding to autonomous governance.

STEP 1

Data Collection: Build a Unified Access Dataset

Before analysis or role mining begins, start by collecting all identity and access data into a single, queryable graph. This step establishes visibility across your organization's systems and entitlements to make informed decisions.

Technical Considerations:

Ensure data inputs with the right sources and at the right granularity level for contextual analysis.

HR Systems

Core attributes such as title, department, worker type, and location.

Identity Providers / Directories

Group memberships and hierarchies from Active Directory, Okta, or EntraID, etc.

Application Entitlements

Fine-grained exports from systems like GitHub (repositories), AWS (IAM roles), Salesforce (permission sets), or Datadog (RBAC roles).

Usage Data (Recommended)

Login frequency, last activity, and access telemetry for contextual insights.

Risk and Metadata (Recommended)

Permission descriptions, business criticality, and risk scores.

Manual

Teams manually export CSVs from each system, normalize fields, and reconcile inconsistencies. Visibility gaps and stale data are common.

vs.

Agentic

Continuously pull normalized data from live connectors, automatically resolve identities, and enrich entitlements with usage and risk context.

STEP 2

Attribute Selection: Identify the Right Access Drivers

Analyze access patterns across the unified dataset spanning HR, IdP, SaaS, cloud, and on-prem systems and determine which attributes best explain access behavior. By evaluating attributes across different factors like coverage, granularity, stability, and manageability, identify the most effective combination of attributes for role definition.

Manual

Teams trial different attribute combinations in spreadsheets, relying on static snapshots and heuristics. Role drift and bloat are common.

vs.

Agentic

Runs multi-dimensional correlation analyses at scale, scoring each attribute set and recommending optimal segmentation automatically with auditable explanations.

Technical Considerations:

Ensure data inputs with the right sources and at the right granularity level for contextual analysis.

Evaluate

Evaluate attribute density and reliability to avoid sparse policy segmentation.

Correlate

Correlate attributes with entitlement clusters to detect strong access drivers.

Prioritize

Prioritize combinations that maintain least-privilege while minimizing administrative overhead.

STEP 3

Access Mapping: Build a Smart Access Matrix

Once the right attributes are selected, generate an ABAC access matrix that visualizes relationships between different cohorts (departments, worker types) and their access across applications. Identify over- and underused entitlements, privilege concentrations, and potential optimizations. This helps with determining birthright vs. self-service access allocation.

Manual

Teams manually pivot large datasets to discover overlaps, often missing nuanced privilege relationships. Excess permission and stale sets creep in.

vs.

Agentic

Continuously regenerate a dynamic access matrix with justifications, business context, and anomaly detection built-in.

Technical Considerations:

Map

Map each entitlement to its attribute-driven access rule.

Integrate

Integrate usage telemetry to distinguish active vs. dormant access.

Flag

Flag high-risk clusters and shadow admin roles for review.

Quantify

Quantify potential impact of least-privilege optimization (e.g., entitlement reduction ratios).

STEP 4

Access Validation: Align with Business Owners

Once draft policies are generated, Lumos facilitates validation with business and application owners before promotion into production. Policies are mapped to lifecycle automation (LCM) workflows, ensuring that provisioning, deprovisioning, and just-in-time access are enforced seamlessly.

Manual

Teams manually pivot large datasets to discover overlaps, often missing nuanced privilege relationships. Excess permission and stale sets creep in.

vs.

Agentic

Continuously regenerate a dynamic access matrix with justifications, business context, and anomaly detection built-in.

Technical Considerations:

Conduct

Conduct validation reviews with app owners to confirm policy intent and compliance.

Capture

Capture exceptions, conditional access, and time-bound entitlements for auditability.

Assign

Assign ownership and review cadences to maintain role hygiene

Auto

Automatically push approved roles into enforcement engines (e.g., Okta, AWS, ServiceNow).

Lumos | Albus

At Lumos, we built an autonomous identity platform that does the heavy lifting with role mining.

Meet Albus, a multi-agent system. It pulls the right signals from your stack, understands how people actually work, drafts clean roles, and learns dynamically to keep them current.

With Albus, you don't just get an assistant. You get an intelligent teammate that brings:

Fine-Grained Data (Beyond AD/IdP Groups)

Entitlement-level detail across SaaS, cloud, on-prem, and custom systems plus usage signals so you can see not just who has access, but how it is used.

Scale at AI Speed

Generate role architecture and policy drafts, validate with human owners, and approve or reject in real-time. Decisions feed future recommendations and fuel better role architecture.

Business Context

Transparent annotations, privilege labels, onboarding manuals, compliance mappings so role suggestions reflect how your business really runs.

Integrated Enforcement (LCM)

Discovery flows directly into provisioning, deprovisioning, JIT access, and updates. What you illuminate, you can immediately control.

Keep the Lights On – Continuously

Role mining shouldn't be a one-time exercise. With Albus, it becomes a continuous practice: discover roles from reality, validate them with context, and enforce them automatically as your organization evolves.

Want to see your own access landscape under the lights? We can turn these steps into a tailored policy matrix and walk through your data with Albus end-to-end.

Ready to flip the switch? Request a free assessment. →