M&Aで拡大し続けるGENDAのデータ活用を促すためのDatabricks権限管理

uma-chan

2025-12-22

1. 自己紹介

Mawatari Daiki / uma-chan

株式会社GENDA IT戦略部 データチーム

データエンジニア / MLOpsエンジニア

1.1. 今日話すこと

M&Aで拡大するGENDAで、多様なユーザーが使いやすいDatabricks権限管理をどう実現しているか

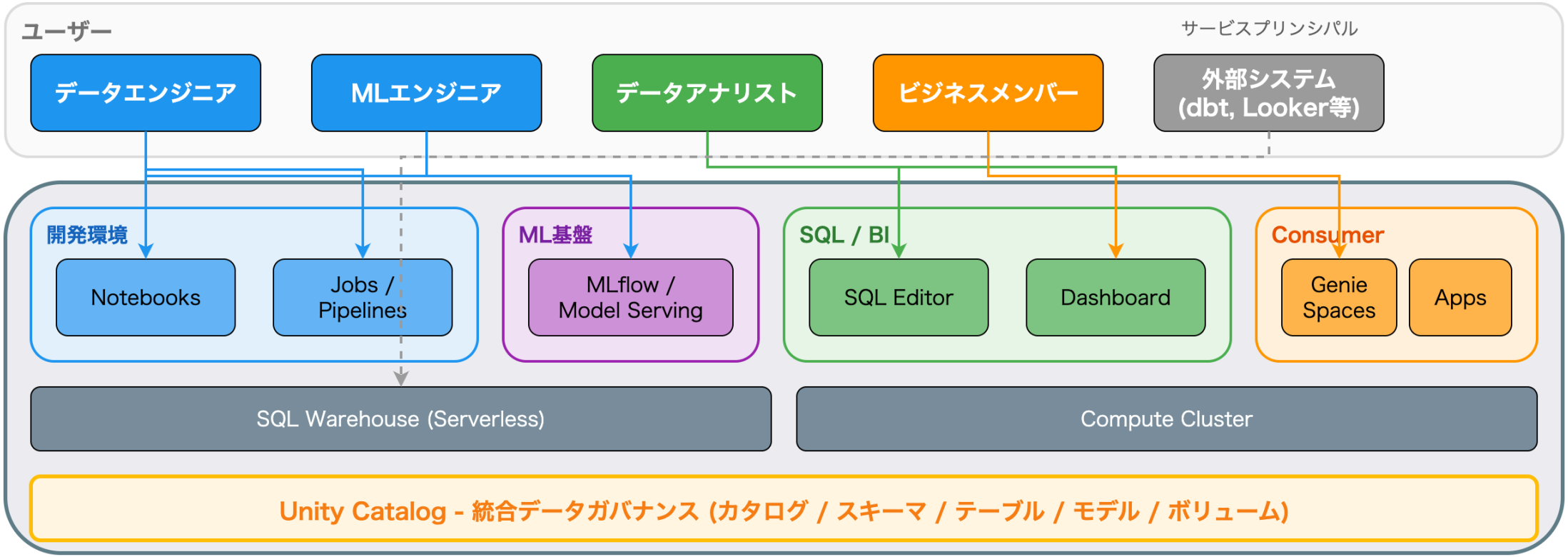
- データエンジニア、MLエンジニア、アナリスト、ビジネスメンバー
- それぞれに適切な権限を、シンプルに管理したい
- Unity Catalog + グループベースの権限管理で解決
- まだWIP (改善を続けている段階)

2. GENDAとデータチーム

2.1. GENDAについて

- 「世界中の人々の日常にもっと『楽しい』を届けたい」
- M&Aによりグループ拡大中
- GiGO、カラオケBanBan、ヒューマックス 等

2.2. Databricks活用の全体像



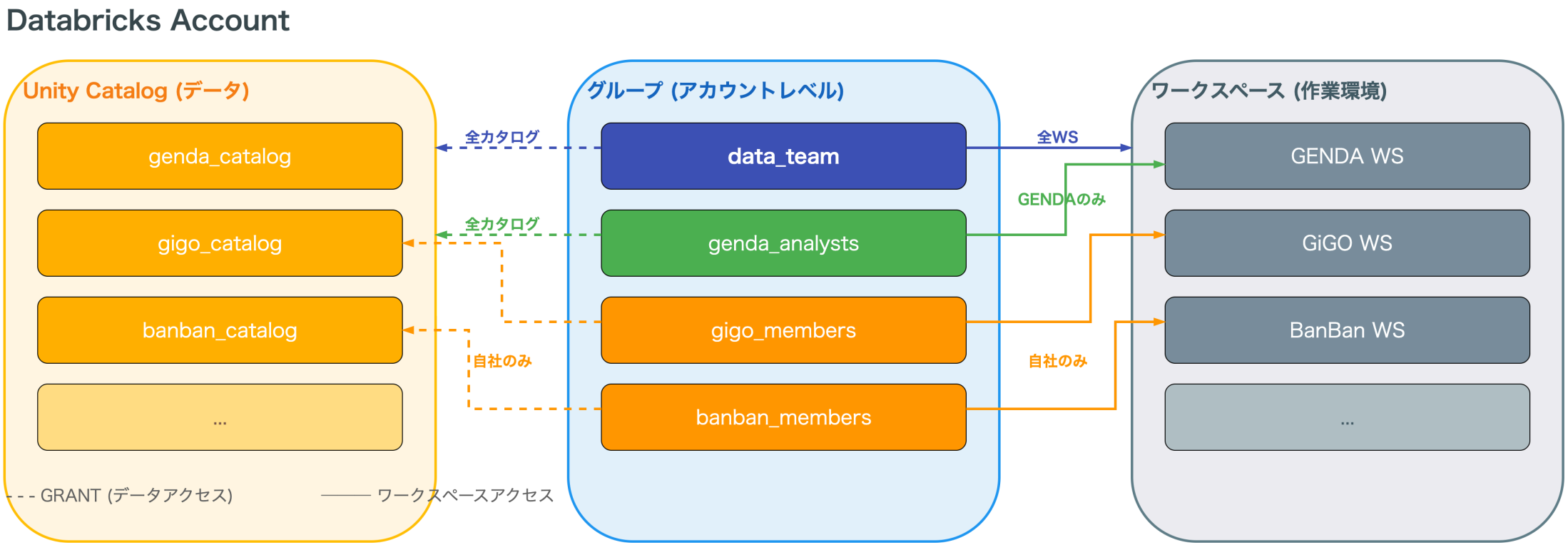
GENDA Databricks活用全体像

2.3. Unity Catalog とは (Snowflake比較)

Snowflake	Unity Catalog
DATABASE.SCHEMA.TABLE	CATALOG.SCHEMA.TABLE
ROLE へ GRANT	GROUP へ GRANT
Stage (ファイル)	Volume (ファイル)

- Databricksのデータガバナンス機能
- テーブル、ビュー、Volume、MLモデルを一元管理

2.4. マルチワークスペース構成



ポイント: genda_analysts は GENDAワークスペースのみ利用だが、全社のデータを分析可能（ワークスペースとデータアクセスが独立）

マルチワークスペース構成

2.5. マルチワークスペースのポイント

- 各グループ企業ごとにワークスペースを分離
  - リソース競合を回避
  - 見えるものを減らして権限管理をシンプルに
- GENDAデータチームは全ワークスペースにアクセス可能 (Superset)
- Unity Catalogはアカウントレベルで統一管理
- カタログ単位でアクセス制御

2.6. 多様なユーザー

- データエンジニア - パイプライン構築
- MLエンジニア - モデル開発・運用
- データアナリスト - SQL分析・可視化
- ビジネスメンバー - レポート閲覧

3. 課題

3.1. 多様なユーザーへの権限管理

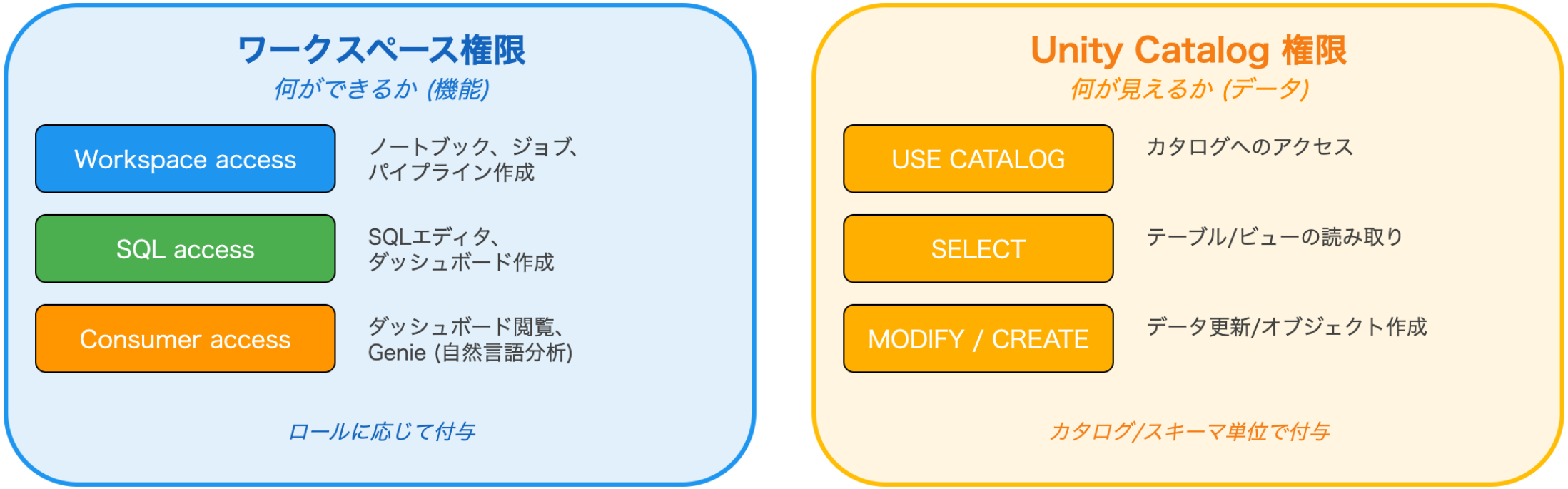
- 各ロールで必要な機能が異なる
- セキュリティと利便性のバランス
- M&Aで新しい会社・チームが増加

3.2. ロールごとのニーズ

ロール	ニーズ
データエンジニア	パイプライン構築、ジョブ管理
MLエンジニア	モデル開発、実験管理
データアナリスト	SQL分析、ダッシュボード作成
ビジネスメンバー	ダッシュボード閲覧、レポート確認

4. 解決策

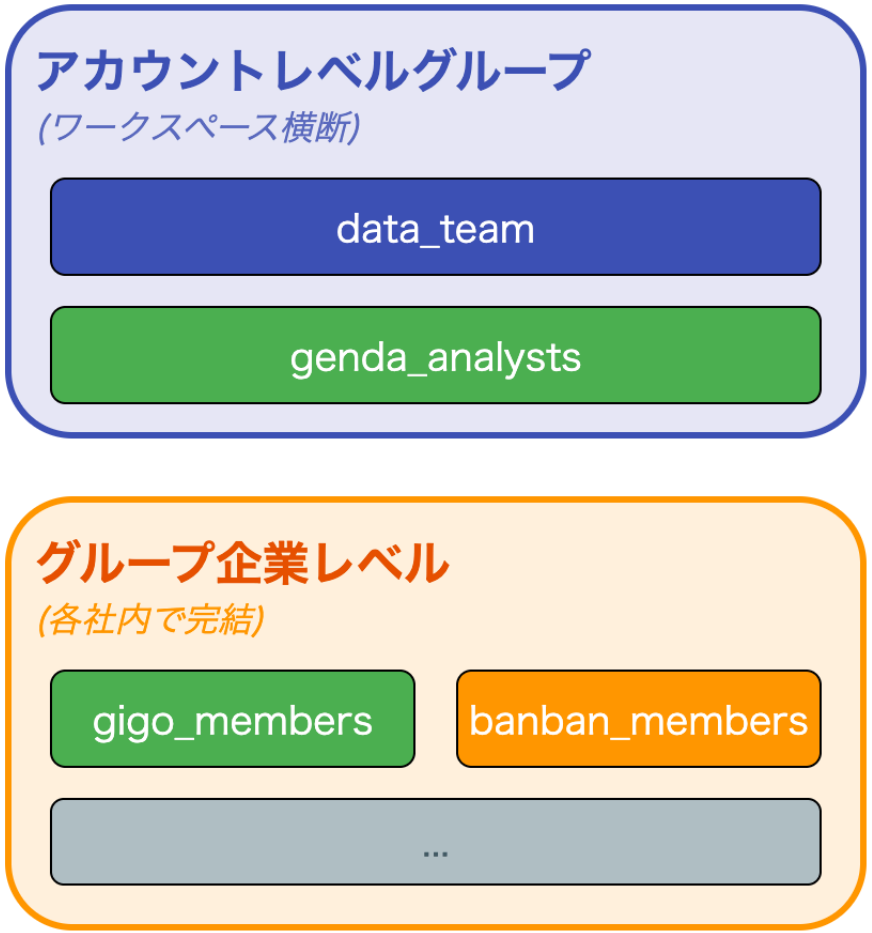
4.1. 権限の2層構造



2層権限構造

- 両方の権限をグループに付与してシンプルに管理
- ロールごとに組み合わせを変えて細かい調整も可能

4.2. グループによる一元管理



グループ管理のメリット

1. 権限付与が簡単

グループに権限を付与すれば
メンバー全員に適用される

2. オンボーディングが即座

新メンバーをグループに追加するだけで
必要な権限がすべて付与される

3. 監査が容易

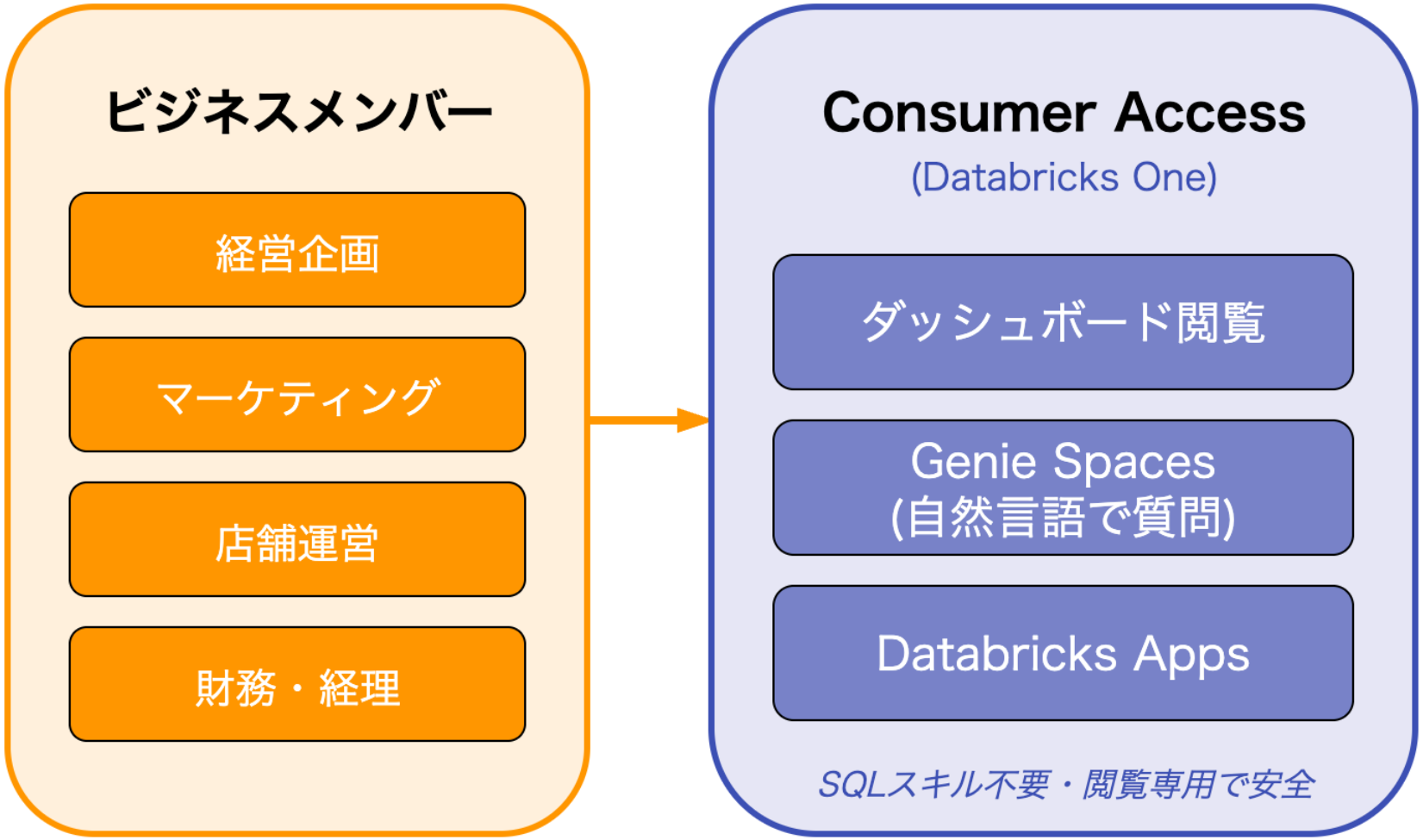
誰がどのグループに属しているか
グループにどの権限があるか一目瞭然

4. M&A対応がスムーズ

新会社用のグループを作成→テンプレート権限を付与

5. 活用シーン

5.1. ビジネスメンバー向け (Databricks One)

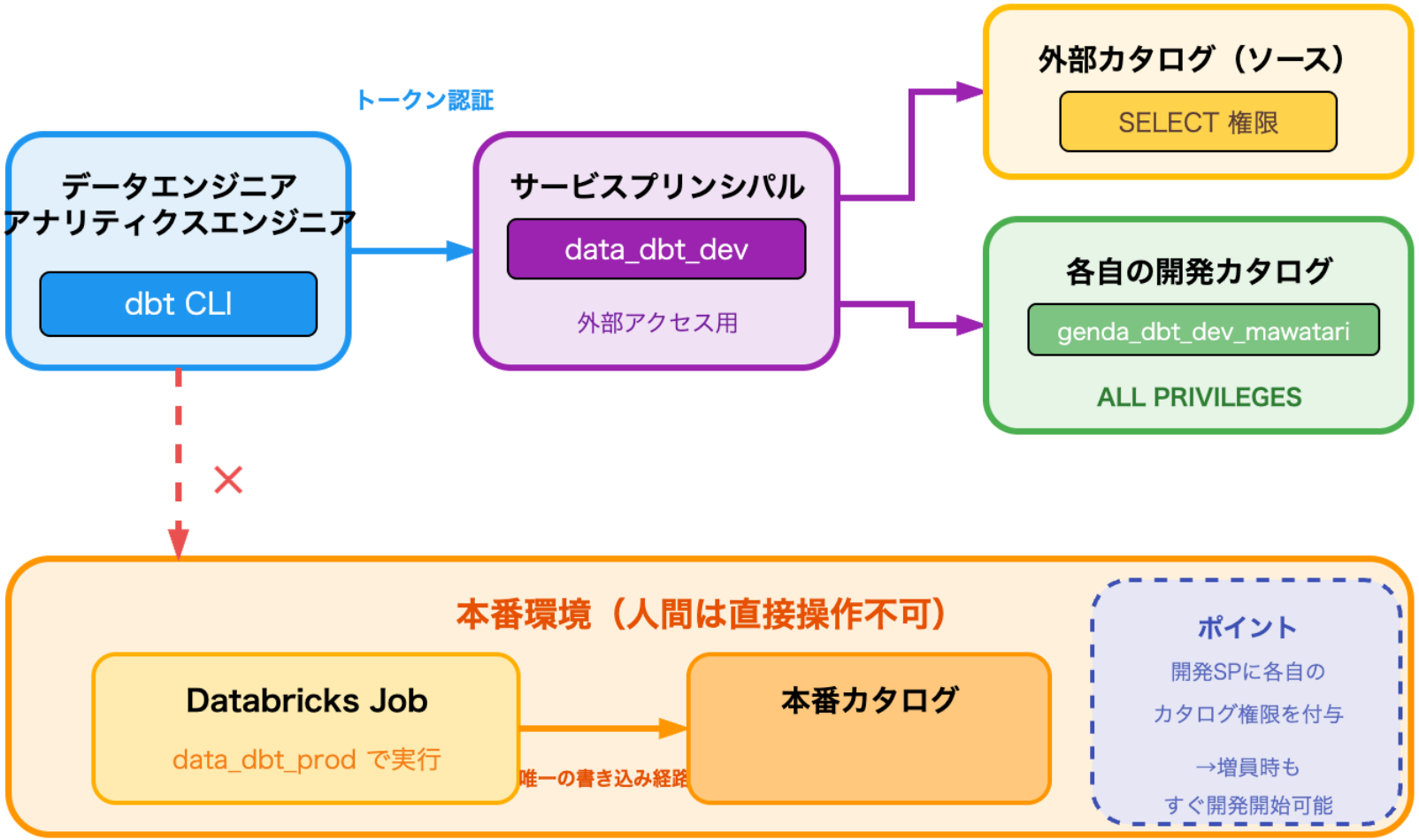


Databricks One / Consumer Access

5.2. Consumer Access のメリット

- シンプルなUI: 複雑な機能が非表示で迷わない
- SQLスキル不要で自然言語でデータ分析 (Genie)
- 閲覧専用で安全にデータ活用
- M&A後の新規メンバーも即座にオンボーディング可能

5.3. データエンジニア / アナリティクスエンジニア向け

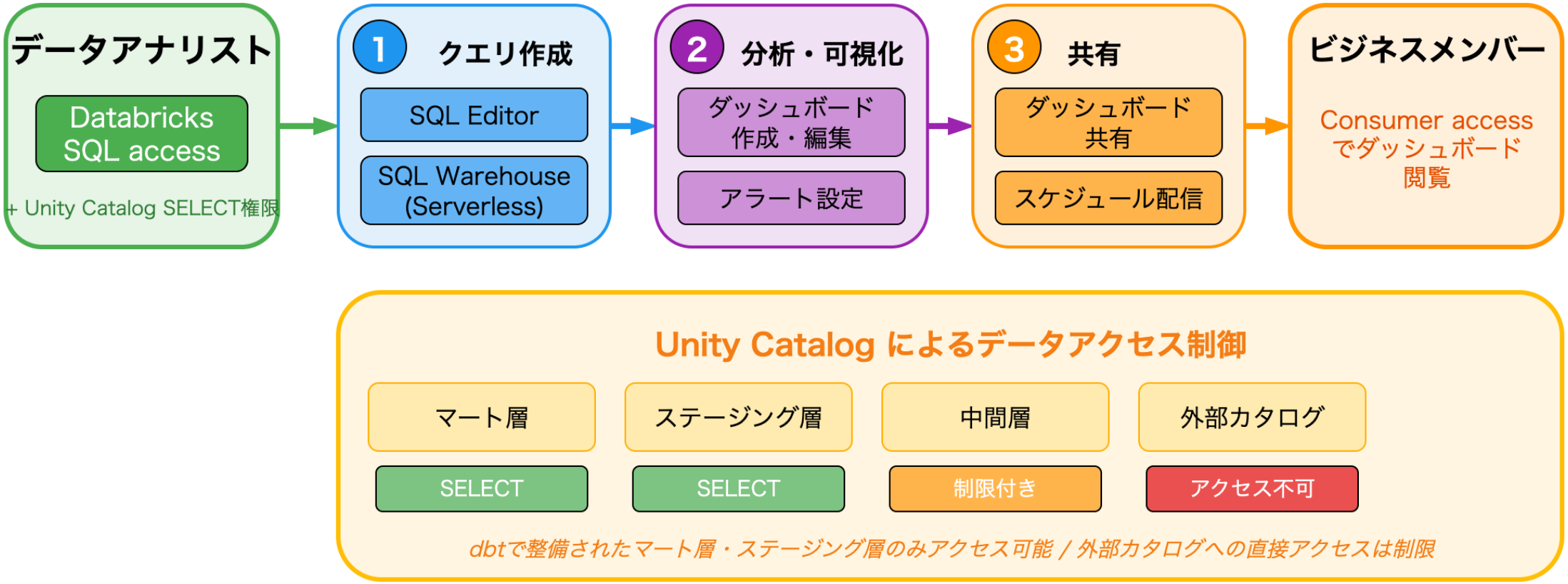


データエンジニアのワークフロー

5.4. データエンジニアの権限活用

- サービスプリンシパル経由でdbt開発
  - 開発SPに各自のカタログ権限を付与済み
  - トークン発行ですぐ開発開始
  - 本番反映はジョブ実行のみ（人間は直接操作不可）
- アナリティクスエンジニア増員に対応
  - 新メンバーにも同じ仕組みですぐオンボーディング
  - 開発環境は共有、本番は分離

5.5. データアナリスト向け



データアナリストのワークフロー

5.6. データアナリストの権限活用

- SQL access でクエリ作成・ダッシュボード作成
- Unity Catalog で必要なデータのみ SELECT 権限付与
- 行レベルセキュリティで顧客データを制限
- 列マスキングで機密情報を自動的にマスク
- 生データ (bronze) へのアクセスは制限し、加工済みデータ (gold) のみ提供

6. サービスプリンシパル管理

6.1. サービスプリンシパルとは

- 人間ではなくアプリケーション用のIDのこと
- ジョブやCI/CDパイプラインの実行主体として使用
- 個人アカウントと同様にグループに所属し権限を持てる
- トークンを発行して外部ツール（dbt, Terraform等）から認証

6.2. サービスプリンシパルのメリット

- ジョブ実行者を非人間にできる
  - 個人アカウントに依存しない本番運用
  - コスト管理が容易（誰の実行か明確に分離）
- 本番カタログ書き換え権限を集約
  - 人間は誰も本番環境を直接操作できない
  - 本番ジョブ実行でのみ書き換え可能
  - 事故防止とガバナンス強化
- トークン発行でdbt開発者がすぐ開発開始
  - 個別の権限設定なしでオンボーディング
  - 共有のサービスプリンシパルで開発環境統一

7. まとめ

7.1. 権限管理のポイント

- ロール別にワークスペース権限を使い分け
- Unity Catalogで細かいデータアクセス制御
- サービスプリンシパルは命名規則と運用ルールを統一
- セキュリティと利便性のバランスを実現

7.2. 今後の展望

- 新規グループ企業への展開を迅速化
- 権限テンプレートの整備
- 自動化の推進