

 Data Maturity Assessment — Spotify

Dimension	Niveau actuel (1-5)	Observations (Forces / Faiblesses)	Plan d'amélioration
Data Governance	3	Les rôles clés existent (CDO, DPO, Data Stewards). Des principes sont définis, mais non appliqués de manière uniforme entre les équipes produits, marketing et équipes techniques. Le modèle organisationnel basé sur des squads autonomes entraîne des silos de données (standards et pratiques hétérogènes).	La gouvernance actuelle est décentralisée et dépend des départements. Une gouvernance centralisée via un CoE est recommandée. Mettre en place un Data Governance Committee + définir des data owners/data stewards par domaine + pilot d'implémentation progressif (modèle CoE).

Data Quality	2	<p>Spotify est une organisation fortement data-driven, où les décisions produits, le ciblage marketing, et les modèles de recommandation reposent sur les données. Cela rend la qualité des données stratégique. Les variations de qualité proviennent de différences dans les outils d'ingestion, les équipes produit, et la localisation géographique (pays/régions), entraînant des formats, conventions et niveaux de complétude non uniformes. La qualité de données inégale affecte directement les recommandations et les décisions marketing.</p>	<p>Mettre en place un outil de data quality monitoring (ex: Ataccama (Data Quality, MDM, Catalog) / Informatica (ETL, Data Quality, Governance, MDM)) + audits de qualité trimestriels.</p>
---------------------	----------	--	--

Data Architecture	4	<p>Architecture moderne & scalable (cloud, microservices, data lake + warehouse). Forte expertise technique interne.</p> <p>Faiblesse : dépendance à plusieurs catalogues internes → manque de standardisation.</p>	<p>Centraliser la gestion des métadonnées via un data catalog entreprise (ex : Alation (Data Discovery, Collaboration) / Collibra (Data Governance, Privacy, Data Quality)).. Standardiser les pipelines ETL/ELT et les API via des templates et librairies communes.</p>
Compliance (RGPD, CCPA, PCI-DSS)	3	<p>Spotify respecte les règles majeures mais la gestion des consentements et du droit à l'oubli est complexe dans certains systèmes historiques. C'est-à-dire des plateformes ou bases de données mises en place avant l'introduction de ces réglementations.</p>	<p>Renforcer la gouvernance autour du data retention + automatisation des processus d'effacement utilisateur. Nécessite une harmonisation internationale et une automatisation des audits.</p>

Data Usage & Accessibility	4	<p>Forte culture de self-service analytics. Cependant, accès non harmonisé selon les équipes → risque de accumulation de privilèges / droits.</p>	<p>Implémenter un RBAC (Contrôle qui peut accéder à quelles données selon son rôle) + ABAC (Contrôle l'accès en fonction de conditions) unifié & audits d'accès réguliers.</p>
Data Security	4	<p>Très bon niveau de sécurité & SOC interne (Rapport d'audit de conformité). Les systèmes critiques sont protégés & chiffrés. Risque : complexité du shadow data dans les équipes produit.</p>	<p>Mettre en place des scans automatisés de data exposure & tagging automatique des données sensibles.</p>
Data Literacy	3	<p><u>Bonne culture data</u> dans les équipes tech, moins dans les équipes business & créatives.</p>	<p>Lancer un programme Data Literacy Academy interne (formations + référents locaux).</p>

Data Integration	3	Pipelines performants, mais multiplicité d'outils → manque d'uniformité des formats.	Définir des standards d'échange + gouvernance des API internes.
Analytics & BI	4	Forte maturité : dashboards produits, ML, personnalisation, recommandations. Limite : qualité & origine des données parfois opaques.	Coupler BI avec le catalog de métadonnées pour renforcer la traçabilité des données utilisées.



Résumé

Spotify possède une **maturité data avancée**, caractérisée par une infrastructure moderne, une forte capacité analytique et une culture data bien ancrée dans les équipes techniques. Cependant, cette maturité **n'est pas homogène** à l'échelle de l'entreprise, notamment sur :

- **La gouvernance transversale**
- **La qualité et la standardisation des données**
- **La gestion des consentements utilisateurs et des droits GDPR**
- **Le contrôle et la traçabilité de l'accès à la donnée**

Les enjeux principaux sont donc :

- **Harmoniser la gouvernance** entre pays, équipes & produits
- **Améliorer la qualité & la documentation** des données critiques
- **Renforcer la conformité & la transparence** vis-à-vis des utilisateurs
- **Simplifier l'accès sécurisé** aux données pour toutes les équipes

Spotify est **mûr pour un Data Governance Framework structuré**, en commençant par un **pilote** sur :

→ un domaine métier clé (ex : Marketing team).