

hymaia

**Data Mesh**  
**Tour d'horizon**



# Vos présentateurs



**Elhadi Cherifi**

**Lead Data Engineer**



**Yoann Benoit**

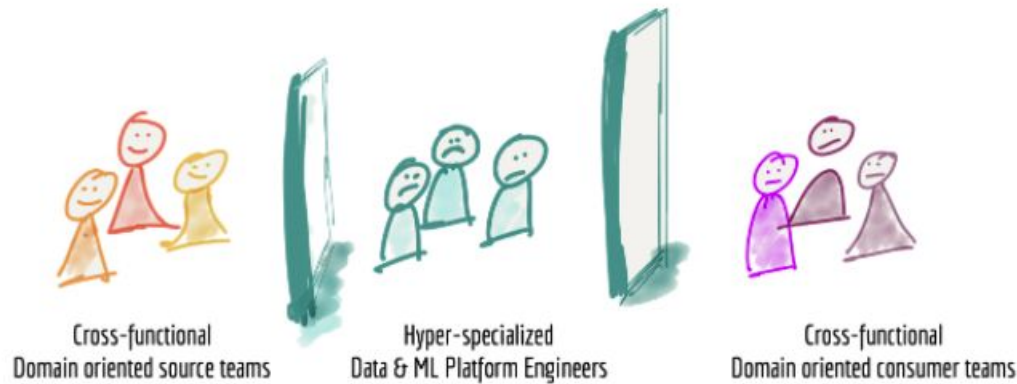
**Head of Data &  
Co-Founder**

# Le constat

# Le constat

Passage à l'échelle difficile pour les plateformes data construites autour d'une équipe dédiée

- Equipe data qui devient un goulot d'étranglement ;
- Pénurie des profils data;
- Data pipelines difficiles à maintenir ;
- Peu de connaissances métier spécifiques.



Source : Data Mesh - Introduction to Data Mesh - Zhamak Dehghani - [https://www.youtube.com/watch?v=\\_bmYXWCx\\_F\\_Q](https://www.youtube.com/watch?v=_bmYXWCx_F_Q)

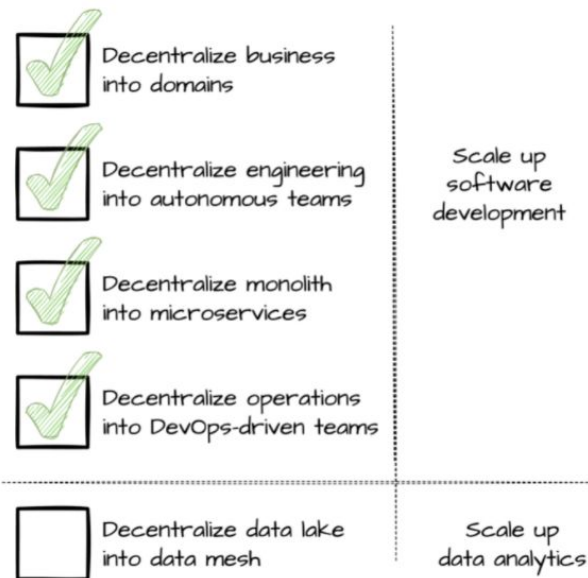
# **Data Mesh : Grands principes**

# Data Mesh & Domain-Driven Design

Le Data Mesh est l'application du **DDD (Domain-Driven Design)** au monde de la Data :

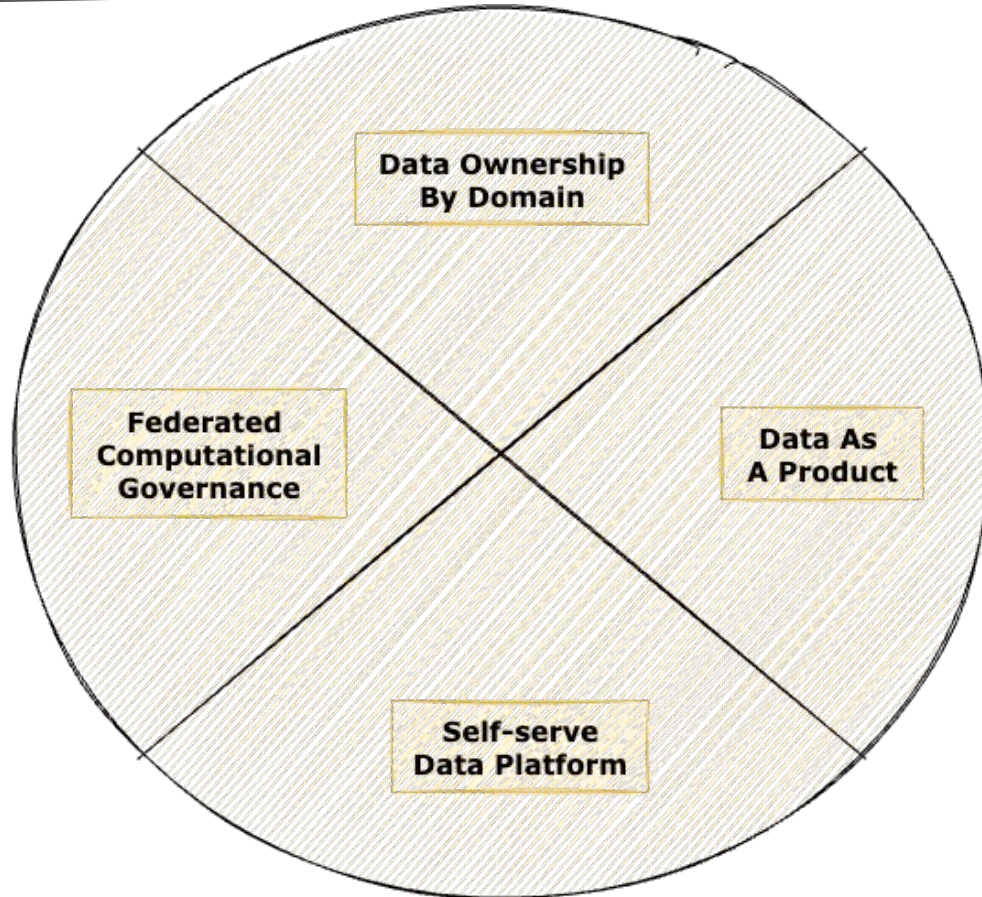
- Le DDD est une méthodologie de développement logiciel qui vise à **structurer les développements par domaine métier**
- Permet d'éviter de devoir gérer un monolithe difficilement compatible avec un passage à l'échelle

Les problématiques sont similaires avec les data plateformes centralisées.



Source : <https://www.datamesh-architecture.com/>

# Les 4 piliers du Data Mesh

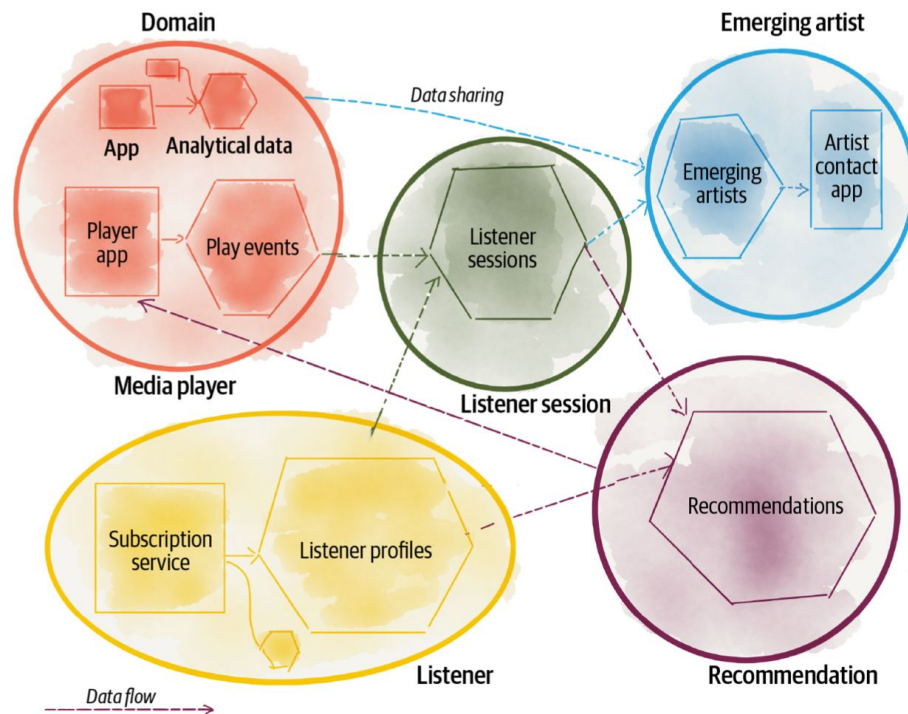


# Data Ownership By Domain

La donnée est **gérée et servie par chaque domaine métier**, qui en maîtrise le mieux les **aspects fonctionnels**.

Un domaine métier doit donc :

- **Produire** un ensemble de datasets ;
- Les **maintenir** et mettre à jour ;
- **Faciliter** leur accès aux différents consommateurs de l'entreprise.

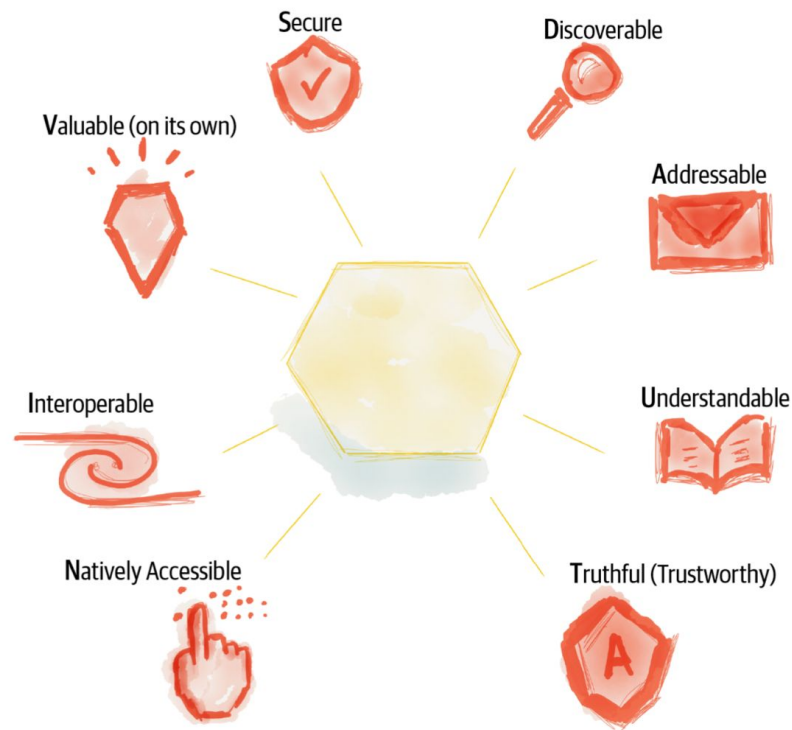


Source : Data Mesh - Delivering Data-Driven Value at Scale - Zhamak Dehghani - O'Reilly



# Data As a Product

- Pallier les problématiques d'accessibilité et d'harmonisation
- Éviter les silos de données
- Application des règles du **Product Thinking** pour considérer la data comme un **produit à part entière**.



Source : Data Mesh - Delivering Data-Driven Value at Scale - Zhamak Dehghani - O'Reilly

# Self-serve Data Platform : Principes

- Éviter la création de **données redondantes entre domaines** ;
- Une **infrastructure agnostique** aux domaines métiers ;
- A pour responsabilité de **fournir des services et ressources** pour faciliter la création de produits data.

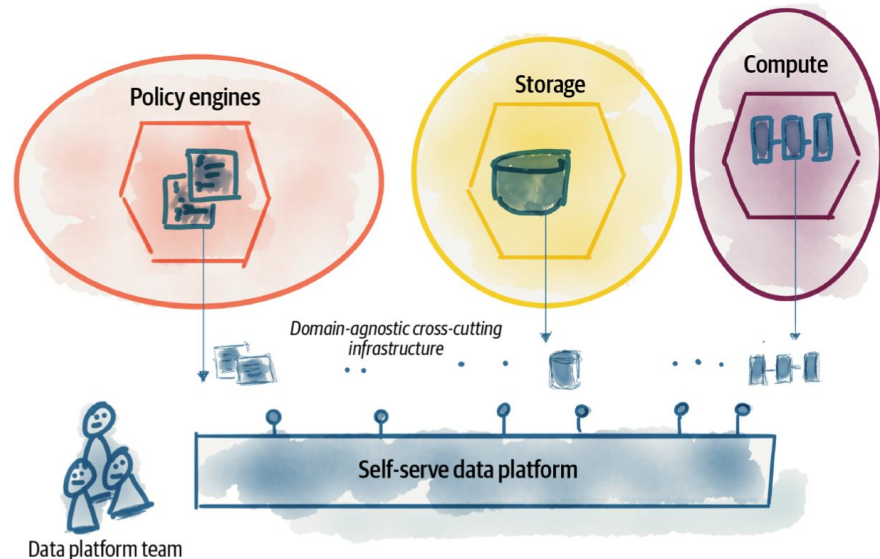
# Self-serve Data Platform : Services fournis

- **Stocker** la donnée de manière scalable ;
- Fournir un **catalogue de données** unifié ;
- Mesurer la **qualité** des données, **monitorer** et **alerter** les produits data;
- Fournir une couche applicative pour **standardiser** les données;
- Pouvoir gérer qui a **accès** à quelle donnée.

# Self-serve Data Platform

**Il est possible de s'appuyer sur sa plateforme data classique et de modifier ses usages:**

- Elle n'aura plus la responsabilité de créer des datasets;
- Elle fournira des services aux différents domaines pour qu'ils puissent créer des produits Data par eux-mêmes.



Source : Data Mesh - Delivering Data-Driven Value at Scale - Zhamak Dehghani - O'Reilly

# Federated Computational Governance

La gouvernance des données assure que nos produits data soient :

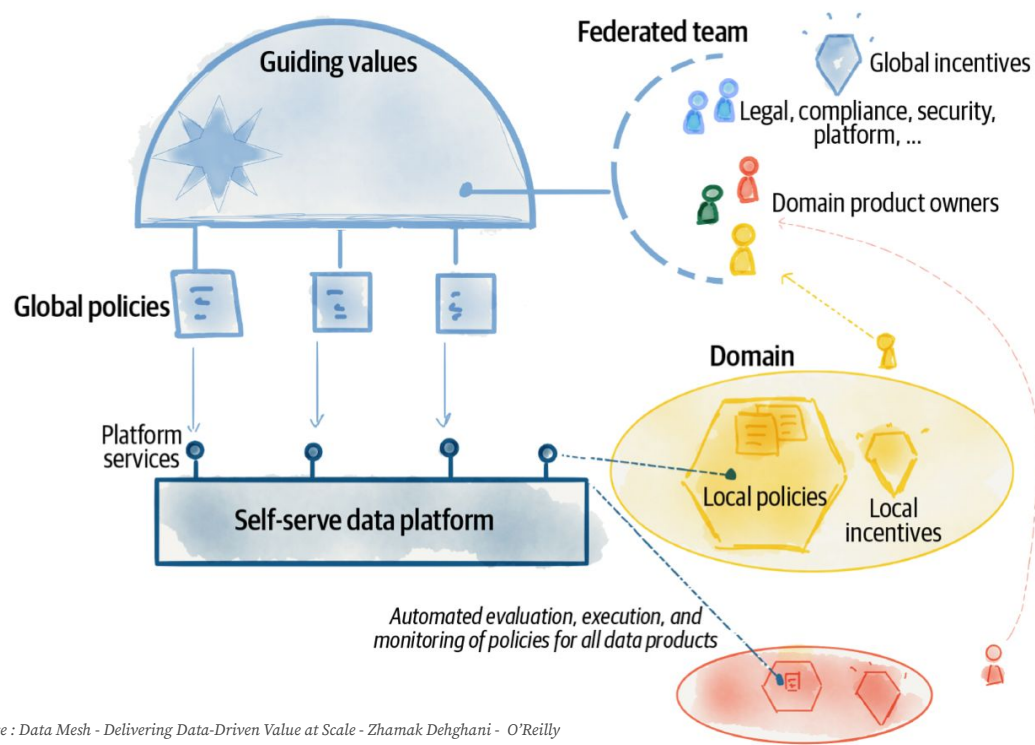
- **Sécurisés ;**
- **Fiables ;**
- **Délivrent de la valeur** dans notre organisation Data Mesh.

# Federated Computational Governance

Cette gouvernance doit :

- **Déléguer** la responsabilité de la qualité au producteur de la donnée ;
- **Automatiser** des vérifications sous forme de politiques pour s'assurer de la conformité des produits data

# Federated Computational Governance



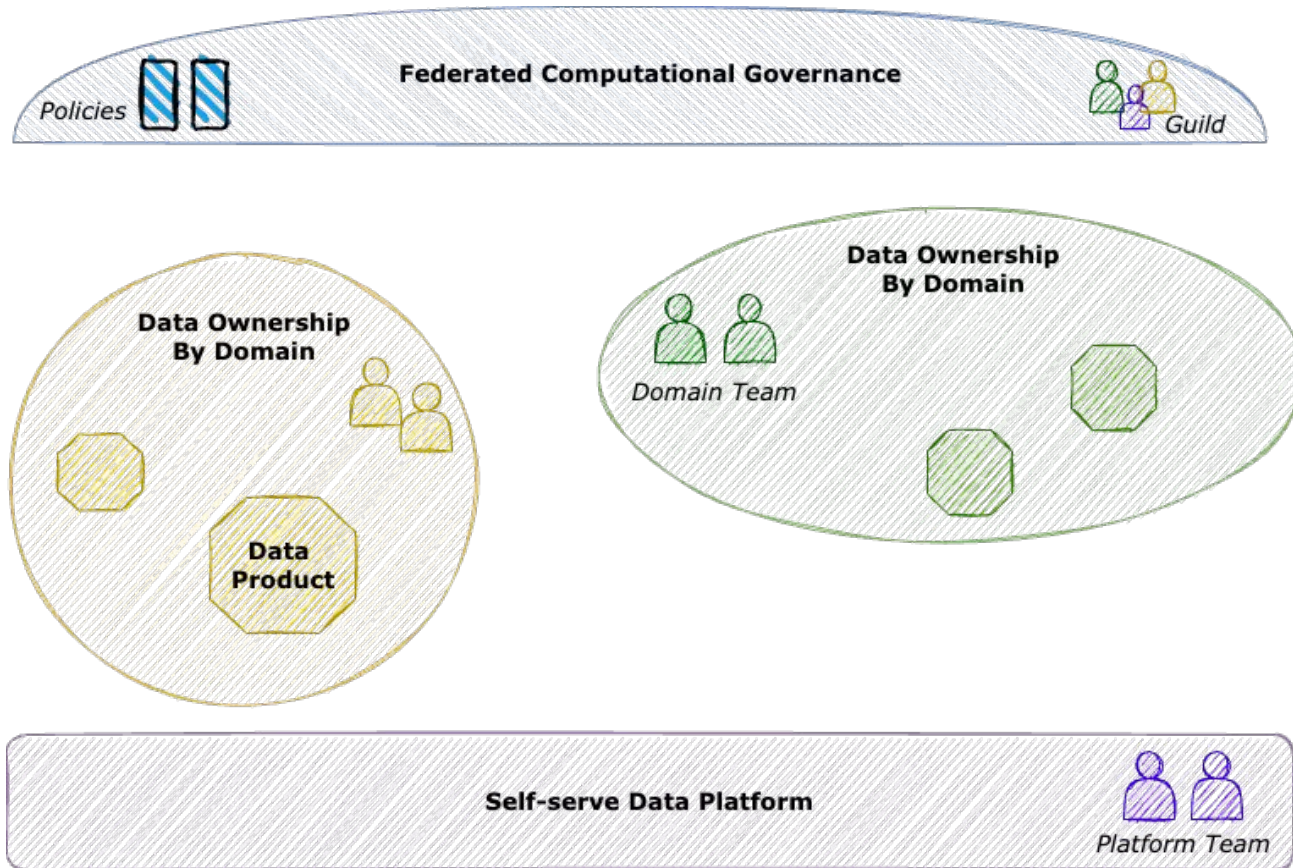
Source : Data Mesh - Delivering Data-Driven Value at Scale - Zhamak Dehghani - O'Reilly

**C'est un bon point  
de départ pour  
entamer une  
transformation vers  
le Data Mesh**

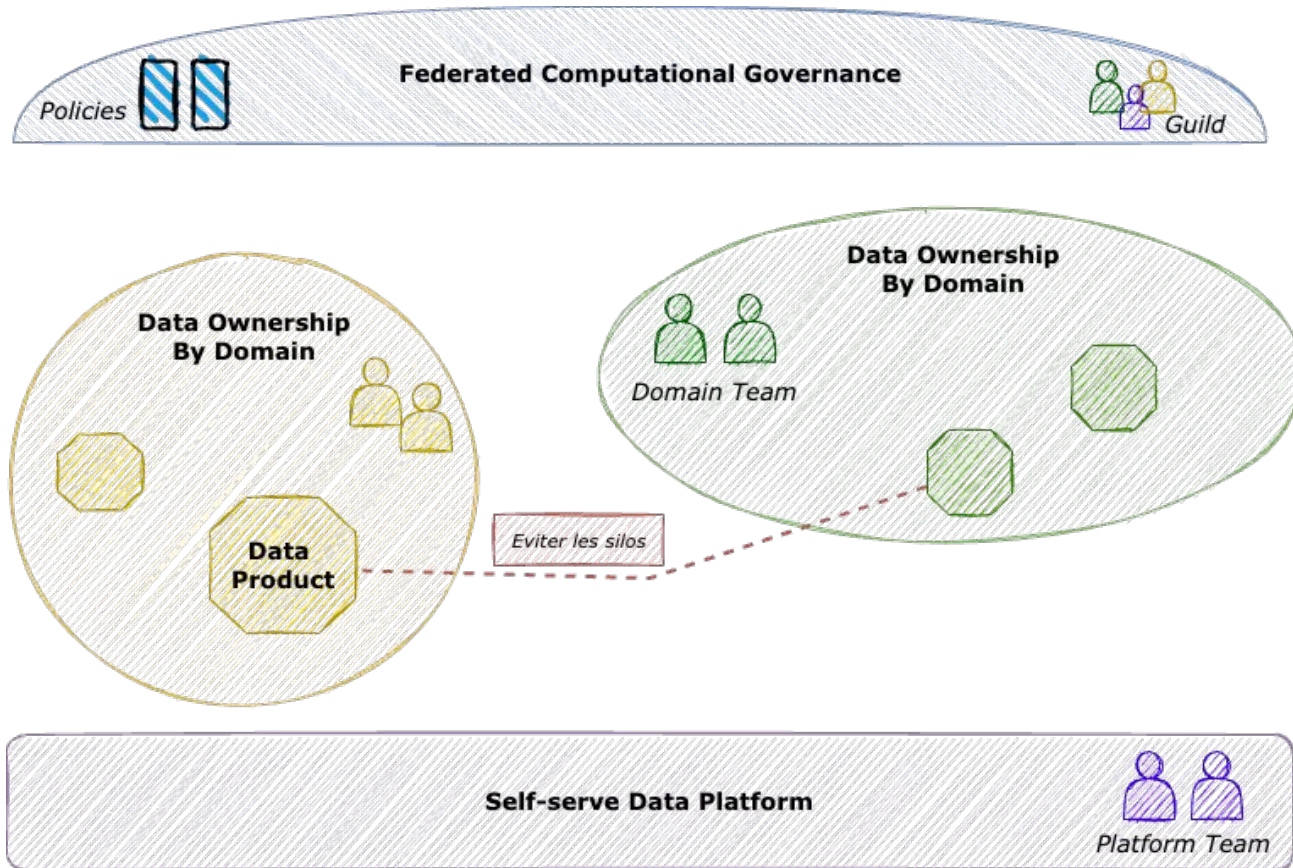
# **Interdépendances entre les principes du Data Mesh**



# L'interdépendance entre les principes du Data Mesh

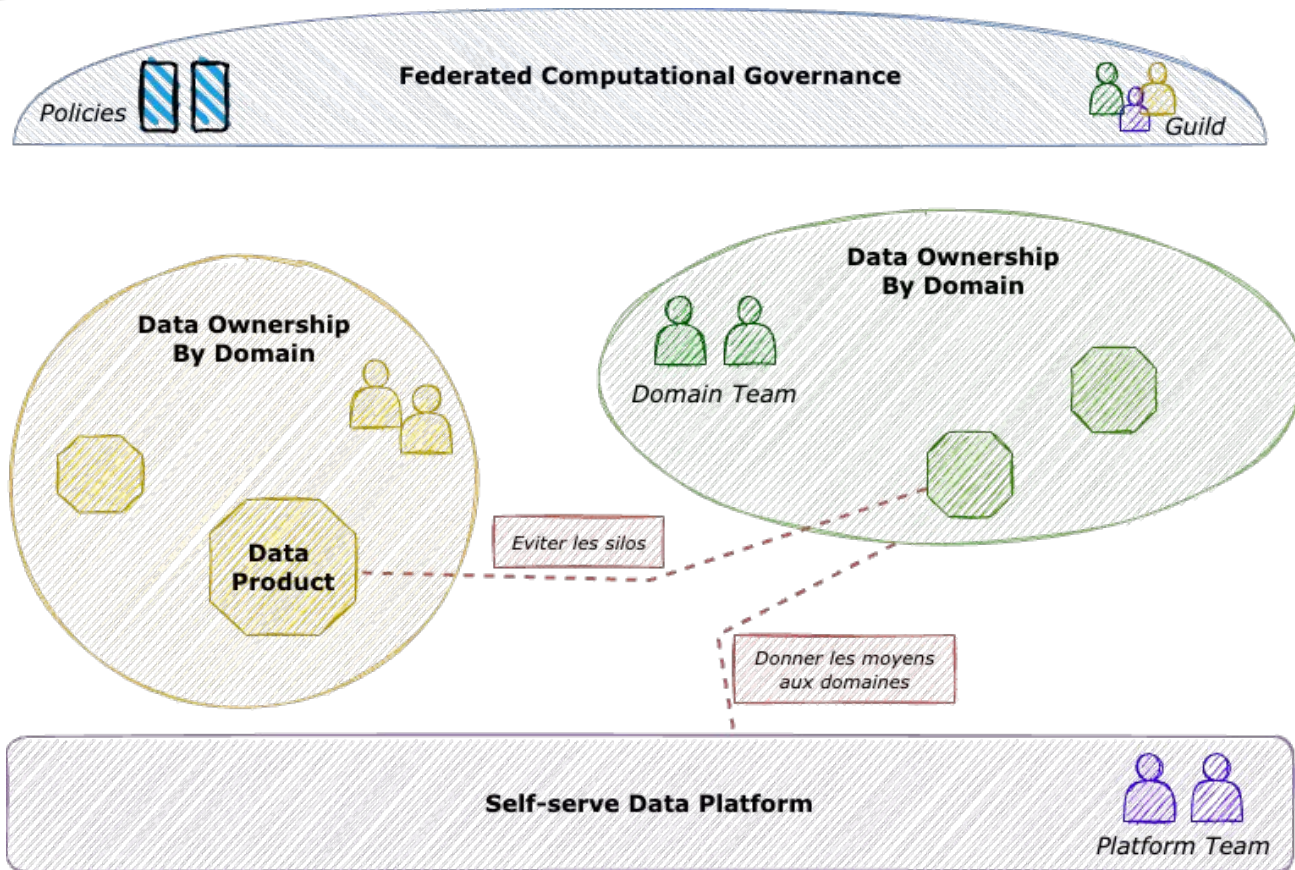


# L'interdépendance entre les principes du Data Mesh

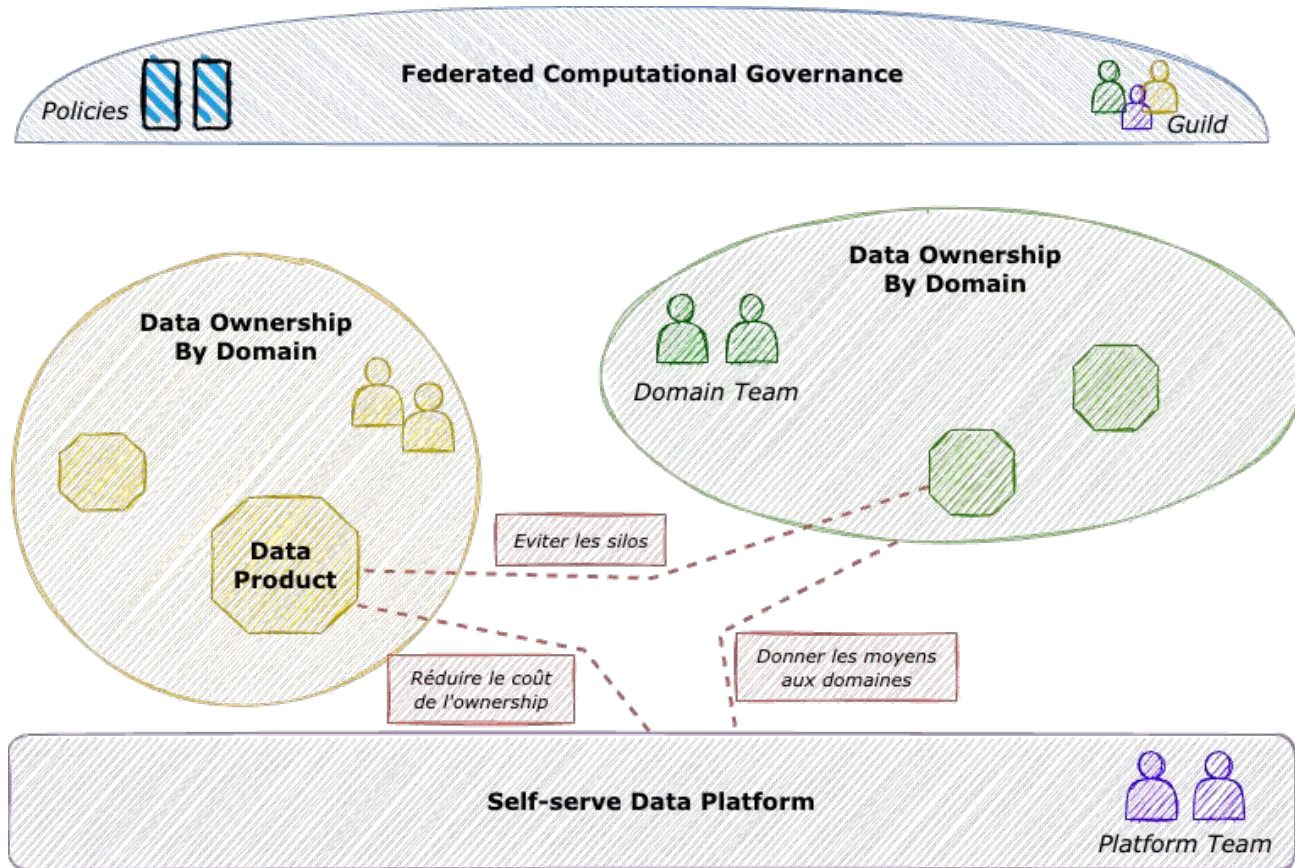




# L'interdépendance entre les principes du Data Mesh

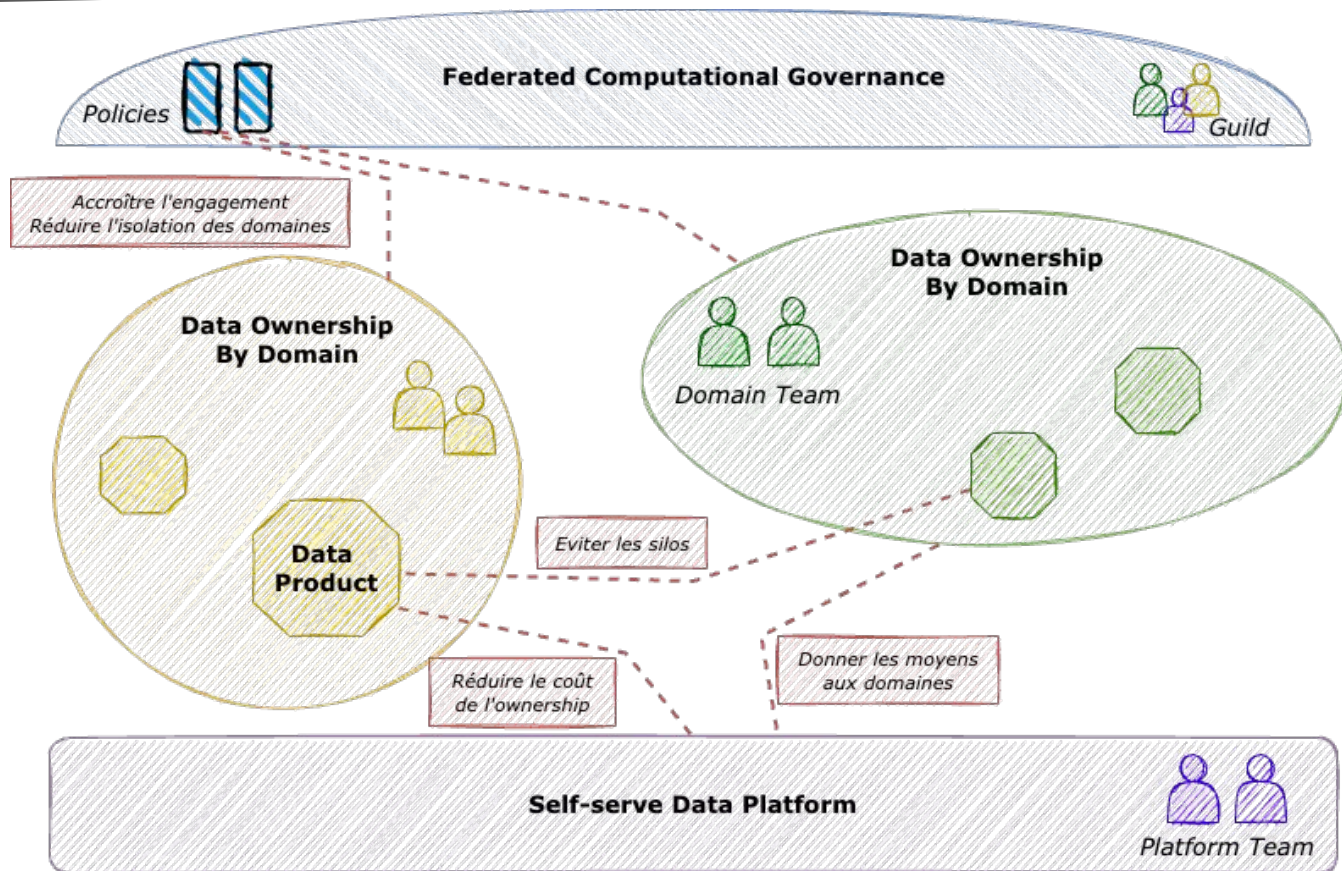


# L'interdépendance entre les principes du Data Mesh





# L'interdépendance entre les principes du Data Mesh



# **Data Mesh : un breaking change ?**

# Une évolution naturelle de la data

Le paradigme du data mesh fait évoluer le monde de la data par des méthodologies déjà prouvées par le passé:

- **L'application du DDD** pour découper le monolithe créé par les data plate-forme classiques.
- **S'inspirer du mouvement DevOps** pour casser les silos entre les équipes data isolées et le reste de l'entreprise.

# Des rôles qui évoluent



## Data Engineer

Construction et exposition des Data Products  
Acquisition d'une compréhension métier  
Placé au coeur technique d'une orga Data Mesh  
N'est plus isolé dans une équipe Data centrale



## Data Product Owner

Décisions concernant la vision des Data Products  
Créer des KPIs mesurables pour assurer le niveau de qualité requis  
Réculte de feedback des consommateurs de la donnée



# Le Data Mesh favorise l'interdépendance entre les profils

**L'organisation en Data Mesh favorise les échanges  
des bonnes pratiques et des connaissances**

Data  
Engineers



Software  
Engineers

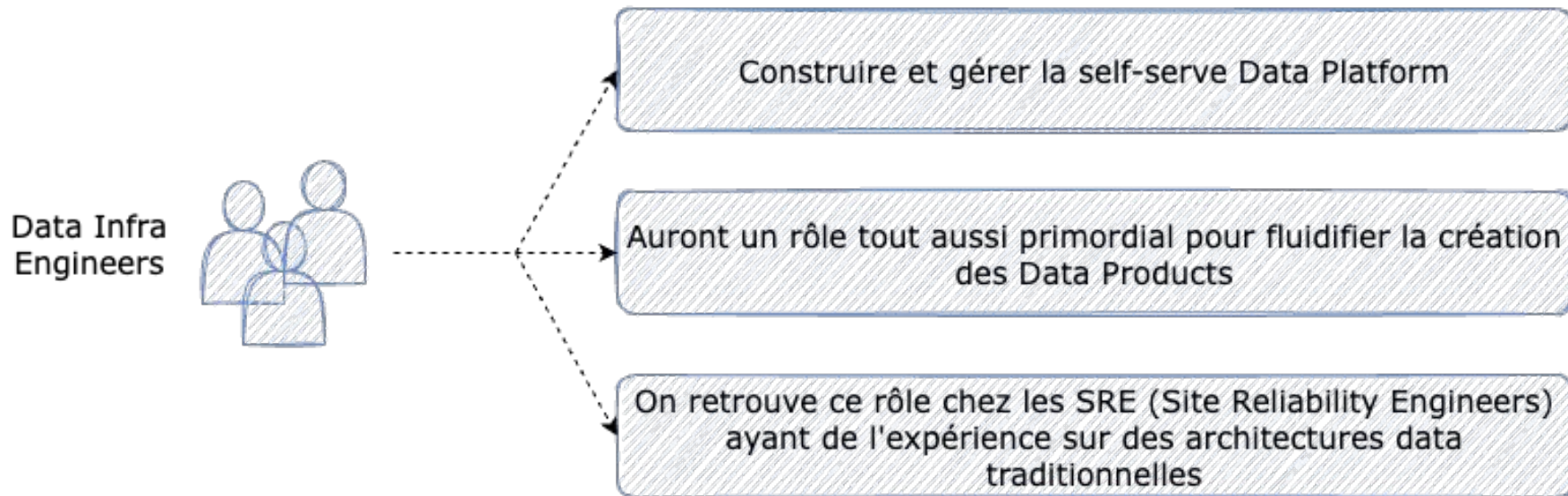


Un Software Engineer peut travailler sur la construction de Produits Data

La production de datasets doit être faite en respectant les normes de développement du software engineering (CI/CD, tests automatisés ...)

Des formations doivent être proposées pour préparer les Software Engineers mais aussi les Ops à travailler avec des outils data

# De nouveaux rôles émergent



# Conclusion

# Conclusion

- Une transformation vers les principes du Data Mesh va **bien au-delà d'un shift technologique** ;
- Définit par Zhamak Dehghani comme un paradigme “**socio-technique**”, mettant en avant le caractère humain primordial à cette transformation ;
- C'est un changement de paradigme complet dans notre approche de la data en entreprise.

**Merci !**  
**Questions ?**