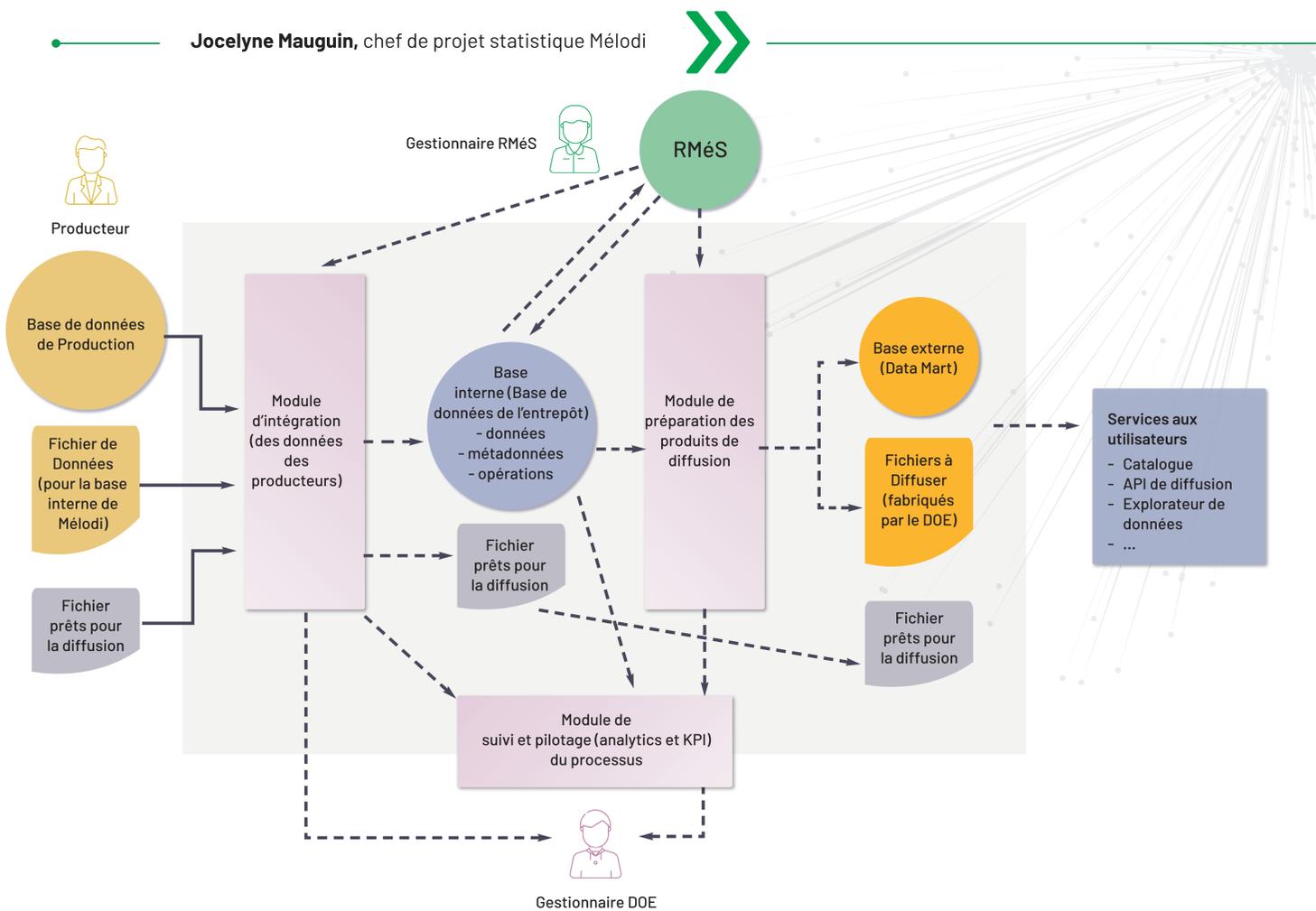


La diffusion des données avec Mélodi et RMÉS

Le projet Mélodi vise à améliorer l'offre de données de l'Insee. Les producteurs livreront leurs données statistiques à Mélodi qui les conditionnera pour les mettre à disposition des utilisateurs via différents services. Pour décrire les données, Mélodi s'appuiera sur RMÉS. C'est une organisation essentielle pour permettre aux utilisateurs d'avoir des données cohérentes entre les différents services et des métadonnées cohérentes entre sources, standardisées et complètement accessibles.

Jocelyne Mauguin, chef de projet statistique Mélodi



Le projet Mélodi vise deux objectifs principaux :

- En externe, améliorer le service rendu aux utilisateurs de nos données en proposant de nouveaux services en Open Data et en améliorant la qualité des données diffusées.
- En interne, rationaliser la diffusion des données statistiques pour atteindre l'objectif externe.

Pour l'ensemble des utilisateurs :

- le catalogue qui est un produit phare de l'offre et permet de porter à connaissance des utilisateurs l'existence des jeux de données de diffusion, en améliorant au passage le référencement naturel dans les moteurs de recherche ;
- l'explorateur de données pour permettre à l'utilisateur d'extraire l'échantillon de données de son choix, afin d'éviter la sur-consommation de données et permettre l'exploration de données de très gros fichiers ;
- le statbot (dont le projet prévoit seulement une expérimentation) renvoyant des réponses automatiques aux questions d'utilisateurs portant sur des demandes de chiffres-clés, afin de rendre l'accès à ces informations plus immédiat.

Pour les utilisateurs les plus techniques, en visant la réutilisation des données, et afin de positionner l'Insee sur des diffusions modernisées :

- la généralisation des formats d'open data pour faciliter l'exploitation de ces fichiers et progresser dans les classements internationaux d'ouverture des données (formats xlsx et csv à la place de xls) ;
- un outil de téléchargement de masse, ce qui évite à l'utilisateur de faire des clics répétés pour visiter le site et qui réduit la charge sur le site en diminuant le temps de recherche sur le site ;
- Une API de diffusion logée sur le portail des APIs qui favorise la réutilisation des données par un dialogue machine-machine (rediffusion ou intégration dans d'autres SI) ;
- enfin la mise à disposition de certaines données en Linked Open Data pour les utilisateurs experts en ce domaine et pour positionner l'Insee sur ce créneau.

Et pour alimenter ces services, un entrepôt de données où les opérations sont pilotées par la métadonnée depuis RMÉS.

Aujourd'hui la diffusion de données, c'est...

5 000

fichiers téléchargeables sur insee.fr chaque année

+ 1 000

fichiers envoyés à Eurostat chaque année

150 000

séries chronologiques dans la BDM

RMÉS

Référentiel de Métadonnées Statistiques