

L'Insee fait évoluer son offre de données

Le projet Mélodi – Mon espace de Livraison d'Open Data à l'Insee



BUREAU DU CNIS - 10 JUIN 2021

# **1 CADRE DE LA DIFFUSION DES DONNÉES À L'INSEE**

## **2 STRATÉGIE D'OFFRE DE DONNÉES**

### **3 PERSPECTIVES**

# 01

## CADRE DE LA DIFFUSION DES DONNÉES À L'INSEE

---

## CONTEXTE

- une demande de transparence de la société civile vis à vis des décideurs politiques ou d'entreprises
  - une société et une économie à l'ère du numérique
  - une évolution de la demande et de l'offre de données
  - des besoins et des usages des utilisateurs des données qui évoluent
- 
- Des représentations plus visuelles des informations : infographies, illustrations interactives
  - Une réutilisation croissante des données : des formats adaptés avec des métadonnées
  - Des accès en temps réel aux dernières informations : des web services et API
  - Des accès sur différents supports : ordinateur, tablette, smartphone et papier

## STRATÉGIE DE L'INSEE

- multiplier les modes d'accès à l'information (site internet, réseaux sociaux, appli smartphone)
- sur [insee.fr](https://www.insee.fr), déploiement :
  - d'une stratégie d'offre de données
  - d'une stratégie de datavisualisation

## STRATÉGIE DE DATAVISUALISATION :

- Renforcement des modalités de diffusion des données dans les publications par des infographies, des illustrations dynamiques basées sur de nouvelles fonctionnalités, des fichiers Excel systématiquement associés
- Renforcement de l'offre d'outils interactifs pour une meilleure appropriation de données démographiques, économiques, locales...

avec l'objectif d' **aller au-devant de tous les publics**

## LE SUIVI DES UTILISATEURS

- **Statistiques de consultation**
  - tableau de bord ou par produit
- **Enquêtes auprès des internautes**
  - enquête mensuelle sur la visite du jour, enquêtes ponctuelles sur les formats de fichiers, sur une collection
- **Tests utilisateurs sur des produits**
- **Focus groups**

# 02

## STRATÉGIE D'OFFRE DE DONNÉES

**CONNAISSANCE DES UTILISATEURS ET DE LEURS  
USAGES DES DONNEES**

**UN PROJET DE REFONTE DE LA DIFFUSION**

**UNE PREMIÈRE RÉALISATION**

## ETUDE SUR LES UTILISATEURS DE DONNÉES (PRESTATAIRE EXTERNE, RECHERCHES INTERNET, TRAVAUX EUROPÉENS)

- Classification des types d'utilisateurs de données (fréquence et intensité)
- Personae réalisées pour les principaux utilisateurs : économiste, data analyste, data journaliste, enseignant en économie, développeur web, data ingénieur...
- Classification des besoins des utilisateurs : découvrir, visualiser, explorer, analyser, ré-utiliser

Type d'utilisateurs	Statut
Gros utilisateurs	<ul style="list-style-type: none"> <li>• Rediffuseurs de données</li> <li>• Web développeurs</li> <li>• Chercheurs (universités, entreprises)</li> <li>• Data journalistes</li> <li>• Data scientists / consultants / chargés d'études</li> </ul>
Utilisateurs intermédiaires	<ul style="list-style-type: none"> <li>• Autres instituts nationaux de statistiques, et organismes internationaux</li> <li>• Observatoires et Administrations</li> <li>• Enseignants et étudiants</li> <li>• Médias, data-journalistes</li> <li>• Influenceurs (« social media multipliers »)</li> <li>• Entreprises</li> </ul>
Utilisateurs légers	<ul style="list-style-type: none"> <li>• Particuliers, entreprises, médias</li> </ul>



**Marie**

28 ans, chargée d'études

Secteur d'activité : public  
Ses outils :



Mode de diffusion des données

- API ○○○○○○
- Spreadsheet ●●●●●●
- CSV / TXT ●●●●●●
- RDF ○○○○○○

## Biographie –

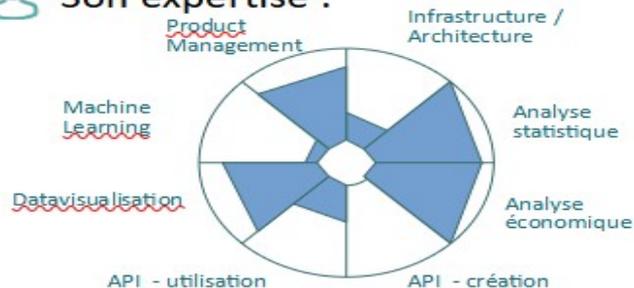
Après sa thèse, Marie a été embauchée comme chargée d'études à l'observatoire de la précarité et de la pauvreté de Gironde. Elle est passionnée par ces réflexions à long terme : chaque chiffre permet d'affiner son constat et ses prospectives.

Elle se base sur les fichiers du recensement de la population et des fichiers détails localisés pour réaliser ses propres traitements avec Access ou le logiciel R qu'elle a beaucoup utilisé pendant sa thèse.

## Ses objectifs :

- Récupérer les données détaillées à exploiter, les mettre en forme pour ses traitements
- Être informée des mises à jour, corrections et de la vie des séries en général
- Comprendre la fabrication (de ce qu'elle télécharge, et comment elle peut s'en servir)
- Répondre à une problématique

## Son expertise :



“

Chaque chiffre permet d'affiner mes analyses et projections

### Types d'usage données Insee



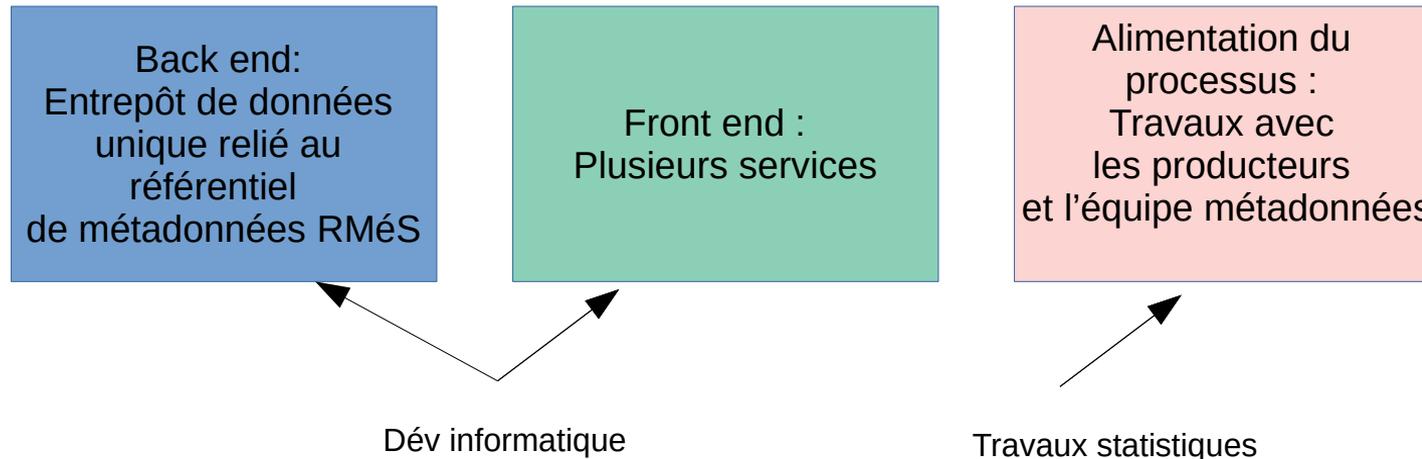
### Ses critères de jugement

- Pouvoir sélectionner rapidement les données dont elle a besoin
- Être guidée dans sa recherche de sources de données
- Avoir une vue sur les sources disponibles permettant de répondre à sa problématique

Le projet Mélodi « **Mon Espace de Livraison d'Open Data à l'Insee** » vise à moderniser le processus de diffusion des données statistiques à l'Insee.

Deux objectifs :

- rationalisation du processus de diffusion des données et gains de temps pour les producteurs
- nouveaux services et amélioration de la qualité pour l'utilisateur





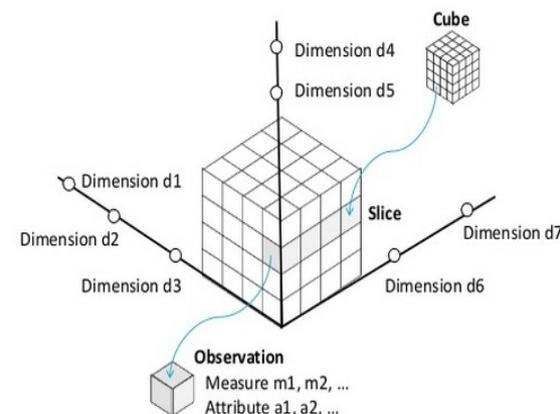
(de plus en plus complexe  
à développer)

Utilisateurs de plus en plus  
techniques

La métadonnée est centralisée dans un unique référentiel (Rmés).

Il y a différents types de métadonnées et elles devront toutes être interprétables par une machine :

- Métadonnées de base : variables et modalités
- Métadonnées de structure : description des cubes
- Métadonnées de catalogage
- Métadonnées de calcul
- Paradonnées



Permet de gagner en qualité grâce à une meilleure cohérence des métadonnées entre sources.

## UNE VÉRITABLE APPLICATION EN PRODUCTION RELIÉE À MÉLODI

Outil de **datavisualisation interactive** sur les salaires dans le secteur privé

Généralités Dimensions Mesures Slices

## Dimensions

Actions	Code	Concept
	ACTIVITY	ACTIVITY - Activité
	REF_AREA	REF_AREA - Géographie
	AGE	AGE - Age
	SEX	SEX - Sexe
	CONDITION_EMPLOI	CONDITION_EMPLOI - Condition d'emploi

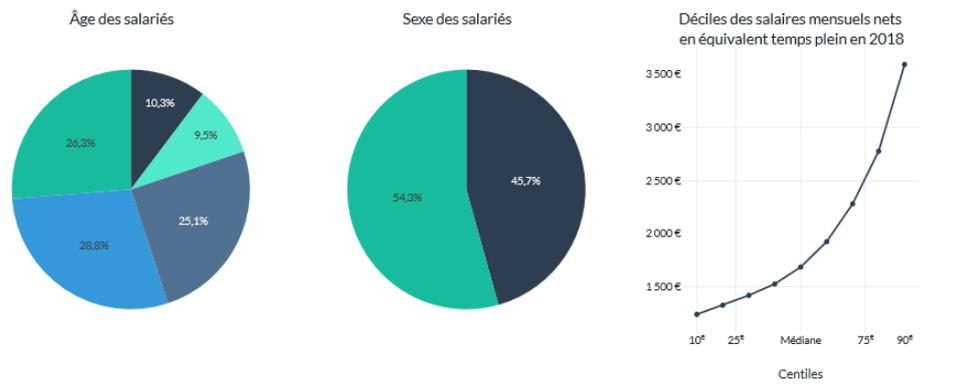


PCS	ESE	SEX	ACTIVITE	TAILLE	ETABLISSEMENT	CONDITION_EMPLOI	AGE	MEASURE	OBS	VALUE	STAT	STAT	MEASURE	UNIT	MEASURE	DD	FREQ	OPESTAT	STAT	CHAMP	R
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	806	216	CENTILE_1	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1061	78	CENTILE_2	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1159	608	CENTILE_3	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1185	737	CENTILE_4	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1209	834	CENTILE_5	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1221	971	CENTILE_6	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1253	964	CENTILE_7	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1273	777	CENTILE_8	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1289	906	CENTILE_9	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1304	658	CENTILE_10	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1300	684	CENTILE_11	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1336	9	CENTILE_12	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1351	241	CENTILE_13	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1366	22	CENTILE_14	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1380	58	CENTILE_15	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1393	808	CENTILE_16	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1405	861	CENTILE_17	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1417	492	CENTILE_18	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1428	684	CENTILE_19	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1439	999	CENTILE_20	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1451	249	CENTILE_21	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1463	14	CENTILE_22	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1474	363	CENTILE_23	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1485	864	CENTILE_24	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1497	343	CENTILE_25	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1509	174	CENTILE_26	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1520	258	CENTILE_27	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1531	767	CENTILE_28	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1543	648	CENTILE_29	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1555	382	CENTILE_30	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1567	197	CENTILE_31	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1578	709	CENTILE_32	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1590	623	CENTILE_33	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1602	336	CENTILE_34	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1614	267	CENTILE_35	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1626	777	CENTILE_36	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1639	667	CENTILE_37	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1652	47	CENTILE_38	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						
TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	TOTAL	sal aire_net_ eqpp_m ensuel_min	1665	225	CENTILE_39	sal aire_net_ eqpp_m ensuel_min	2018A	DADS	T						

## Professeurs d'art (hors établissements scolaires)

Professionnels, salariés ou non, enseignant une discipline artistique en dehors d'un établissement d'enseignement général. L'exercice de l'activité peut être pratiqué au domicile du professeur, dans un local particulier ou au sein d'un établissement spécialisé dans la délivrance des différents enseignements artistiques.

**Professions les plus typiques :** correcteur / correctrice a domicile, directeur d'enseignement artistique, directeur d'école de musique, directeur d'école nationale (supérieure) des (beaux) arts, directeur d'établissement artistique, directeur de conservatoire, enseignement artistique, formateur (formation et accompagnement artistique), musicien intervenant / musicienne intervenante, professeur / professeure de batterie, professeur / professeure de chant, professeur / professeure de clarinette, professeur / professeure de danse, professeur / professeure de flute, professeur / professeure de formation musicale, professeur / professeure de guitare, professeur / professeure de percussions, professeur / professeure de piano, professeur / professeure de saxophone, professeur / professeure de theatre, professeur / professeure de trompette, professeur / professeure de violon, professeur / professeure de violoncelle, professeur d'art (sauf établ. d'enseignement général ou technique), professeur d'art dramatique, professeur d'école nationale (supérieure) des (beaux) arts, professeur de conservatoire, professeur de danse (sauf établ. d'enseignement général ou technique), professeur de dessin (sauf établ. d'enseignement général ou technique), professeur de musique (sauf établ. d'enseignement général ou technique), professeur de peinture (sauf établ. d'enseignement général ou technique)



Close

## UN CONTENU DIVERS

### – Des outils interactifs

- Se positionner par rapport aux salaires
- Se positionner par rapport aux professions

### – Des publications augmentées

- Sur l'analyse des salaires
- Sur les séries longues

### – Des focus

- Inégalités salariales femmes-hommes
- Disparités géographiques

# 03

## PERSPECTIVES

---

## LA CIBLE (SELON TIM BERNERS-LEE) : LE LINKED OPEN DATA

### – Faire interagir différents silos

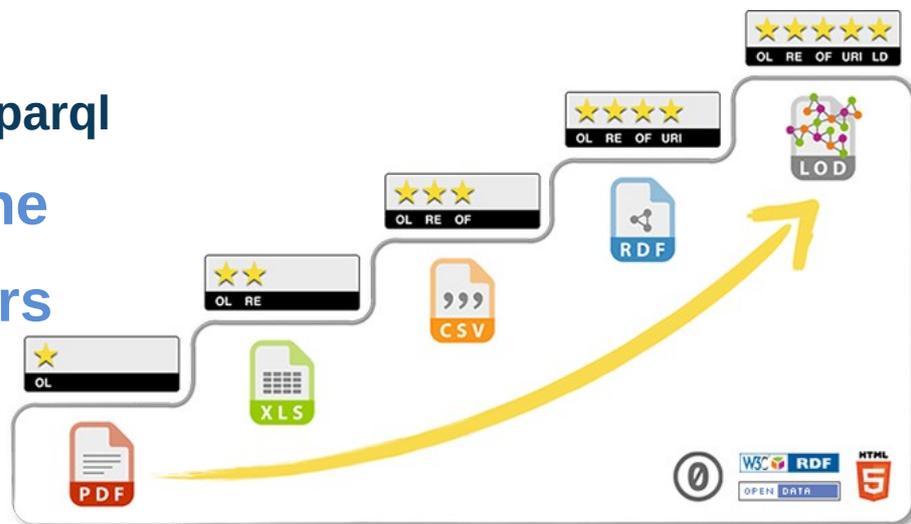
- Au niveau national (SSM)
- Au niveau international

ex : <https://statistics.gov.scot/sparql>

### – Par un dialogue machine-machine

### – Pour élargir l'offre aux utilisateurs

- Par de nouveaux outils
- Grâce à l'intelligence artificielle



## UN CHEMIN TECHNIQUE

- **Développement des API accessibles depuis le portail [api.insee.fr](https://api.insee.fr)**
  - Mode natif avec Melodi
- **Une « usine » pour fabriquer les produits de diffusion**
  - Gain de temps pour les producteurs, de délai pour les utilisateurs
  - Cohérence intrinsèque des données
- **Une modélisation technique des données**
  - Appui sur le format SDMX, promu par les institutions internationales majeures
  - Modélisation des métadonnées selon le standard datacube
- **Des outils parfois complexes à appréhender**

## UN CHEMIN SÉMANTIQUE

- Finaliser l'intégration de toutes les sources dans le référentiel de métadonnées statistiques (RMÉS)
  - Y compris potentiellement les jeux de données SSM
- Avec une description homogène et exploitable du contenu des jeux de données (standards de modélisation)
- Pour proposer aux utilisateurs un catalogue complet
- Afin d'explorer toutes les facettes d'un sujet
  - Quels jeux de données comprennent la dimension âge ?
- Et aussi relier plus systématiquement les études aux données
  - Identification des jeux de données

# 04

## POUR CONCLURE

---



TABLEAU DE BORD DE L'ÉCONOMIE FRANÇAISE

FR - EN

THÈMES > MARCHÉ DU TRAVAIL - SALAIRES Emploi - Population active

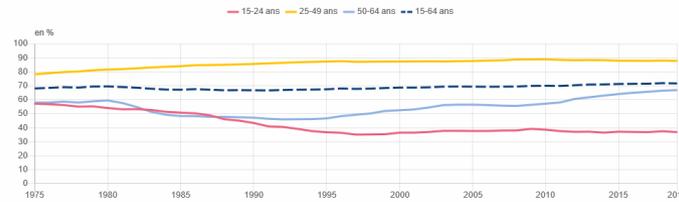
FRANCE EUROPE

En 2019, en France, 71,7 % des personnes âgées de 15 à 64 ans sont actives, c'est-à-dire en emploi ou au chômage. Ce taux d'activité moyen diffère selon l'âge : 36,9 % pour les 15 à 24 ans ; 87,8 % pour les 25 à 49 ans et 66,9 % chez les 50 à 64 ans.

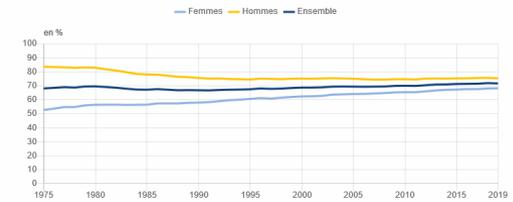
Le taux d'activité des jeunes a baissé depuis 1975, en lien avec l'allongement de la durée des études. Celui des seniors (50 à 64 ans) s'est redressé depuis le milieu des années 90, sous l'effet du recul de l'âge de départ à la retraite. En 40 ans, le taux d'activité des femmes s'est rapproché de celui des hommes.

[Lire la suite](#)

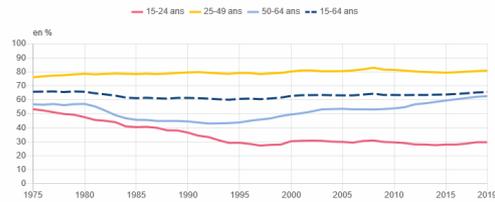
Taux d'activité selon l'âge



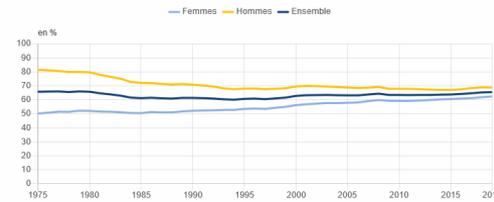
Taux d'activité selon le sexe



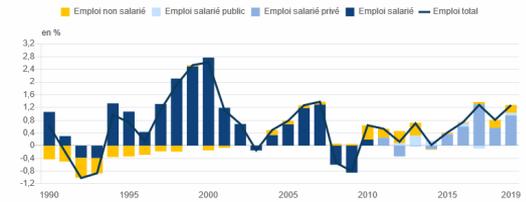
Taux d'emploi selon l'âge



Taux d'emploi selon le sexe



Évolution de l'emploi



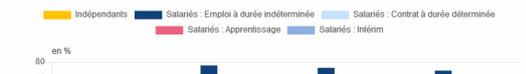
Évolution de l'emploi salarié



Emploi par catégorie socioprofessionnelle selon l'âge en 2019



Satut d'emploi et type de contrat selon l'âge en 2019



Collection Insee Références  
Édition 2020



266 pages

