

Glossaire

L'analyse fédérée : l'état de la science

Série d'apprentissage collectif

Dernière mise à jour : 10 janvier 2024

Le présent glossaire vise à établir une compréhension commune des termes et des concepts présents dans la série d'apprentissage collectif *L'analyse fédérée : l'état de la science*.

Principaux termes	Définition provisoire
Modèle de données commun	<p>Un modèle de données commun permet l'analyse systématique de diverses bases de données d'observation de différentes autorités compétentes en transformant l'information contenue dans les bases de données sources en un ensemble convenu d'éléments et de tableaux de données normalisés. Les modèles de données communs présentent deux grands avantages : ils raccourcissent le délai nécessaire pour traiter les requêtes provenant de multiples centres de données et ils permettent un contrôle de la qualité des résultats (1, 2).</p> <p>Exemples de modèles de données communs existants :</p> <ul style="list-style-type: none">• Modèle de données commun du Canadian Network for Observational Drug Effect Studies (3), une adaptation du modèle Sentinel de la Food and Drug Administration (4).• Modèle de données commun du Observational Medical Outcomes Partnership (5), adopté à différents degrés dans de nombreux pays, y compris les États-Unis, le Royaume-Uni, l'Australie et 22 pays d'Europe.
Données distribuées	Données stockées dans de multiples organismes, établissements ou centres de données (6).
Données fédérées	Données distribuées qui peuvent être analysées conjointement sans être regroupées (6, 7, 8).
Analyse fédérée	Analyse de données provenant de multiples fichiers de données et qui ne sont pas stockés au même endroit (données distribuées), réalisée rapidement et de façon sécuritaire (9).
Apprentissage fédéré	Apprentissage automatique appliqué à des données décentralisées. Comme pour l'analyse fédérée, les modèles sont appliqués localement dans chaque établissement



	participant et seules les caractéristiques des modèles (paramètres, gradients, etc.) sont transférées (6 , 7 , 8 , 10).
Fédération horizontale	Partitions comprenant les mêmes caractéristiques ou mesures, mais qui portent sur des populations différentes (p. ex. différentes provinces, données couramment définies pour utilisation dans le contexte des soins de santé) (11).
Fédération verticale	Partitions qui portent sur les mêmes populations, mais qui comprennent des caractéristiques différentes (p. ex. fichiers de données différents, l'un qui porte sur l'utilisation des soins de santé, l'autre qui comprend les données d'un sondage) (11). <i>Remarque : Même s'il n'est pas nécessaire qu'un lien clair soit établi, on présume que les populations sur lesquelles portent les fichiers de données fédérés verticalement sont en totalité ou en très grande partie les mêmes.</i>
Analyse parallèle	Méthode empirique utilisée pour déterminer le nombre de composantes dans une analyse en composantes principales (ACP) ou de facteurs dans une analyse en facteurs communs (AFC) (12).
Méta-analyse	Méthode statistique qui consiste à regrouper les résultats d'études indépendantes (13).
Analyse groupée	Analyse de données individuelles regroupées provenant de différents emplacements ou de différentes sources.
Données synthétiques	Données générées par des algorithmes dans le cadre de simulations par ordinateur, utilisées en remplacement de données du monde réel. Les données synthétiques peuvent servir à entraîner l'apprentissage fédéré (exemple d'article ici) (14).
Environnements de confiance	Environnements informatiques hautement sécurisés qui permettent un accès à distance aux données sur la santé (15 , 16).

Références

1. [Canadian Network for Observational Drug Effect Studies \(CNODES\) CDM](#)
2. [The Book of OHDSI: Chapter 4 The Common Data Model](#)
3. [CNODES](#)
4. [FDA Sentinel CDM](#)
5. [OMOP](#)
6. [From distributed machine learning to federated learning: A survey.](#)



7. [Federated learning: Challenges, methods, and future directions.](#)
8. [Federated learning for healthcare informatics.](#)
9. [Federated networks for distributed analysis of health data.](#)
10. [The future of digital health with federated learning.](#)
11. [Federated machine learning: Concept and applications.](#)
12. [A rationale and test for the number of factors in factor analysis.](#)
13. [What is meta-analysis?](#)
14. [Synthetic data example](#)
15. [Trusted Research Environments.](#)
16. [Trusted research environments \(TRE\) green paper.](#)

