RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

Secrétariat général pour l'investissement

FRANCE 2030

PROGRAMME DE RECHERCHE
DIGITALISATION ET DECARBONATION DES MOBILITES

Mob Sci-Dat Factory – Project Cible 3

Partage d'outils de traitement et d'analyse des données de mobilité

Porteur : Aline CARNEIRO VIANA, INRIA

Co-porteur : Alexandre CHASSE, IFPEN

Partenaires publics et ses associés : CEREMA, IGN, UGE, INSA-Lyon, CNRS, Sorbonne Paris Nord, ENPT

Inria · IFP Energies nouvelles · Université Gustave Eiffel · IGN · Cerema

Sommaire

1. **Positionnement au sein du PEPR**
   1. Contexte scientifique
   2. Objectifs
2. **Description du projet**
   1. Stratégie scientifique
   2. Structuration
   3. Organisation du partenariat
3. **Contributions aux PEPR**
   1. Les délivrables
   2. Les contributions du PEPR pouvant enrichir le projet
   3. Les possibilités de contributions et de mise en œuvre dans les territoires

Positionnement au sein du PEPR

1. Contexte Scientifique

Mob Sci-Dat Factory



Contexte Scientifique

Priorité EU

*« La mobilité future des **personnes et des biens** »*

L'échange, la gestion et l'infrastructure numérique des données de mobilité sont des besoins de recherche essentiels pour aborder la transformation énergétique et numérique qui change la nature de la mobilité



Les questions soulevées



Les données à mobiliser

- leur disponibilité, leur accessibilité, leur qualité et leur représentativité

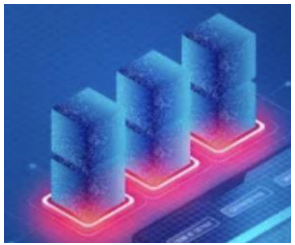


Les méthodes et les outils

- pour traiter, analyser et extraire des connaissances
- pour la compréhension et l'exploitation des données

La plateforme d'aide à la décision

- pour le partage des outils et des données



Les questions soulevées

Les données à mobiliser

- leur disponibilité
- ur représentativité
  - Le
- aire des connaissances
- et l'exploitation des données

La plate à la décision

- pou tage des outils et des données

Répondre conjointement à ces questions est une tâche difficile !

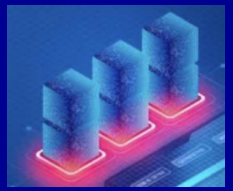
L'objectif de Mob Sci-Dat Factory



Les Défis







Mob Sci-Dat Factory



Les défis : les données à mobiliser

1. a-L'accessibilité et l'acquisition

- accès **difficile et restreint** (sauf acheteur et collaborations étroites)
- manque de sources de données hétérogènes

Conséquences

- Solutions et décisions limitées en

réalisme | précision | validité | généralité



- Recherche

non reproductible | sans réplicabilité | non exhaustive





Les défis : les données à mobiliser

1. b-La collecte

- un « **combat** » qui demande :
  - Le recrutement de bénévoles
  - Le déploiement d'une infrastructure de collecte
  - Des efforts logistique et administratif considérables…



Conséquences

- Echantillon homogène d'utilisateurs
- Vision incomplète des mesures
- Manque d'évolutivité



Les défis : les données à mobiliser

1. c-La qualité et représentativité

   - fortement liées au contexte et dépendent de la collecte
   - nécessite un enregistrement transparent de leur traçabilité

Conséquences

- Des campagnes de collecte (en général) indépendante
- Manque de traçabilité

Les défis : les méthodes et outils

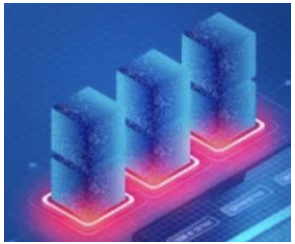


2. Le traitement, l'amélioration et l'exploitation de données
   - d'initiatives indépendantes et fragmentées
   - manquant de : unification, modularité, flexibilité, documentation

- Traitement long et couteux (en ressources et en effort)
- Ralentit la recherche et les avancements socio-économique
- Limite la reproductibilité
- Difficulte le calibrage et d'exploitation des données multi-sources

Conséquences

Les défis : la plateforme d'aide à la décision



3. Les infrastructures de partage et de traitement des données
   - d'initiatives dispersées et fragmentées
   - manque d'accès et d'interopérabilité
   - régies par des entités indépendantes

Conséquences

- Rend difficile leur utilisation
- Limite la reproductibilité et le répétabilité de solutions

Positionnement au sein du PEPR

2. Objectifs



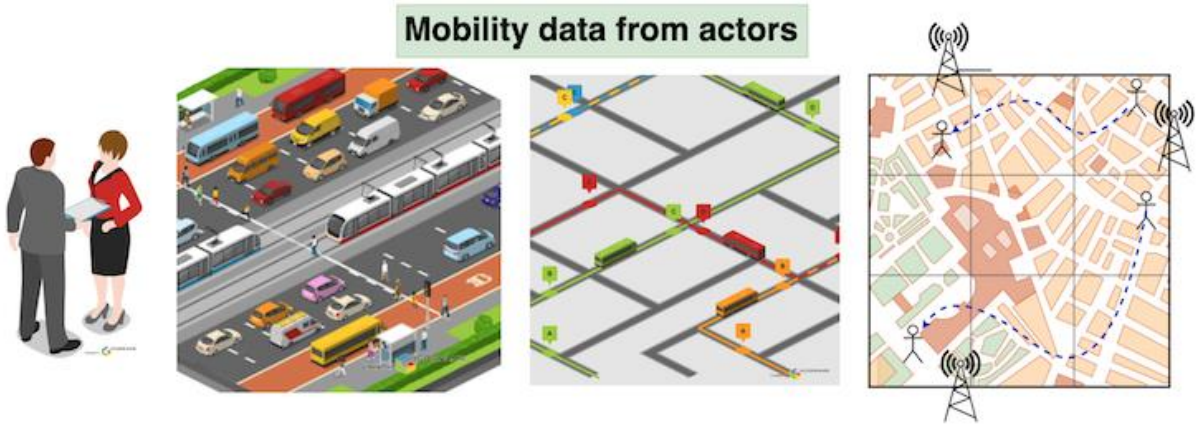
Filling the gaps!

Mob Sci-Dat Factory

Des données de mobilités

…. inventoriées, collectées et acquises, respectueux de la vie privée,
… avec leur écosystème, leur cycle de vie et les modifications subies
➡ *assurant leur traçabilité*

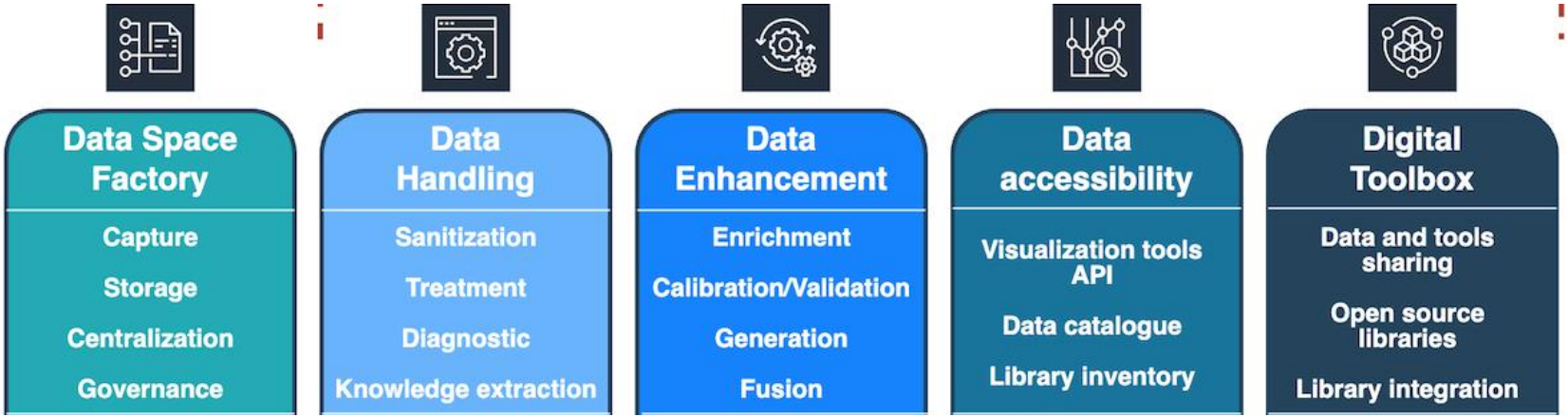
Des sources homogènes et hétérogènes
à grande échelle



Des methodes, bibliothèques et outils ….

… *open-source, unifiées, modulaires et interopérables dès leur conception*

➡ *Digital ToolBox*



Mob Sci-Dat Factory

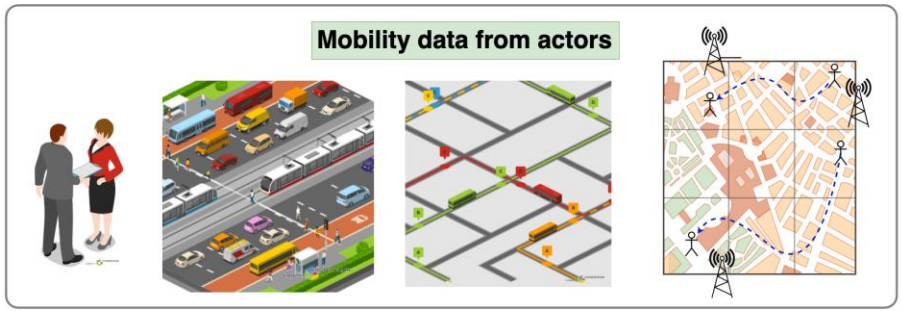
Une infrastructure cloud

… de partage (données et outils) et de traitement évolutifs et sécurisée

… avec une gouvernance technologique et des politiques d'accès et de traitement respectueuses de la vie privée



Mob Sci-Dat Factory



Use-cases requirements and specifities

Contributions by the scientific community

PC 3

PC 2

PC 1

Data Space Factory	Data Handling	Data Enhancement	Data accessibility	Digital Toolbox
Capture	Sanitization	Enrichment	Visualization tools API	Data and tools sharing
Storage	Treatment	Calibration/Validation	Data catalogue	Open source libraries
Centralization	Diagnostic	Generation	Library inventory	Library integration
Governance	Knowledge extraction	Fusion		
Storage and sharing gouvernance	Processed, privacy compliant, and exploitable Data	Improved, representative, and calibrated Data	Dashboards, websites, data and tool information	Ressource management sharing and processing platform

PC3 -- Mob Sci-Dat Factory

Sharing with the scientific community

Description du projet

1. Stratégie Scientifique et
2. Structuration

Mob Sci-Dat Factory



Stratégie scientifique

Pour le **benefice collectif** de la compréhension de la mobilité”

Assurer une gouvernance de partage et de traitement **adaptée et respectueux** de la vie privée

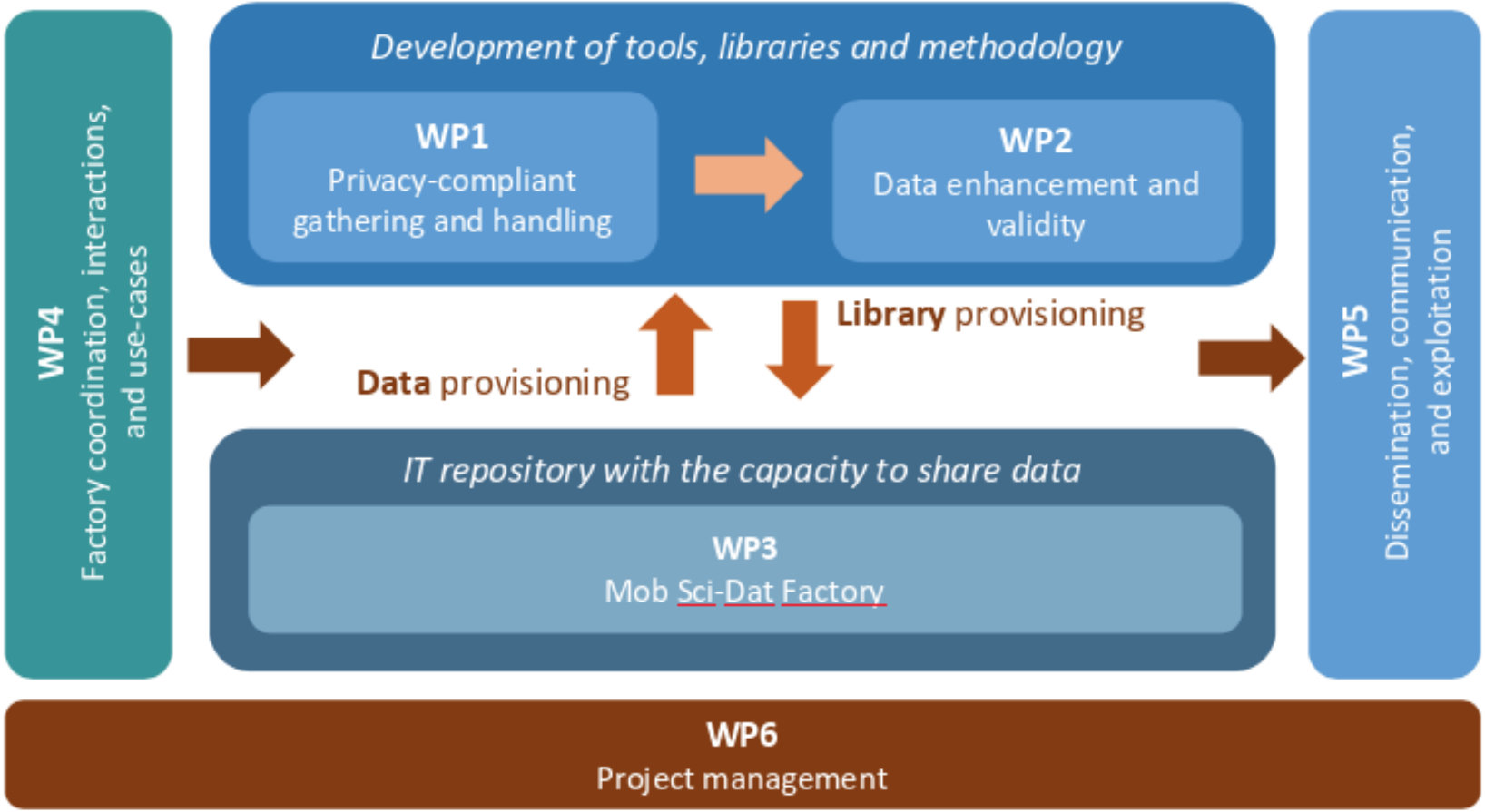
Faciliter l'adoption d’une **pratique commune** de partage, de stockage et de traitement

Mob Sci-Dat Factory



Assurer une utilisation **transparente et continue** de la plateforme par de nombreux acteurs

Structuration



Structuration

Technique

WP1 : Collecte et traitement respectueux la vie privé

T1.1 : Spécificités des données, traçabilité et traitement

T1.2 : Conformité à la vie privée

T1.3 : Analyse de données, diagnostic et extraction de connaissances

Technique

WP2 : Amélioration et validité des données

T2.1 : Augmentation, fusion et reconstruction des données

T2.2 : Validité et représentativité des données

T2.3 : Réalisme, calibrage et amélioration des *data-to-model*

Technique

WP3 : Mob Sci-Data Factory

T3.1 : *Data space factory*

T3.2 : *Resource factory* et gestion

T3.3 : Accessibilité des données et des outils

WP1 - WP3 assurent les tâches scientifiques et pratiques requises pour la conception du **Digital Toolbox** et la mise en place de la plateforme **Mob Sci-Dat Factory**

Opérationnel

WP4 : Coordination, interaction et cas d'usage du *Factory*

T4.1 : **Architecture** : spécification, intégration, interaction des éléments

T4.2 : Identifier les **cas d'usage**, ses exigences et données associées

T2.3 : Actions **d'appel ouvert**

Gestion

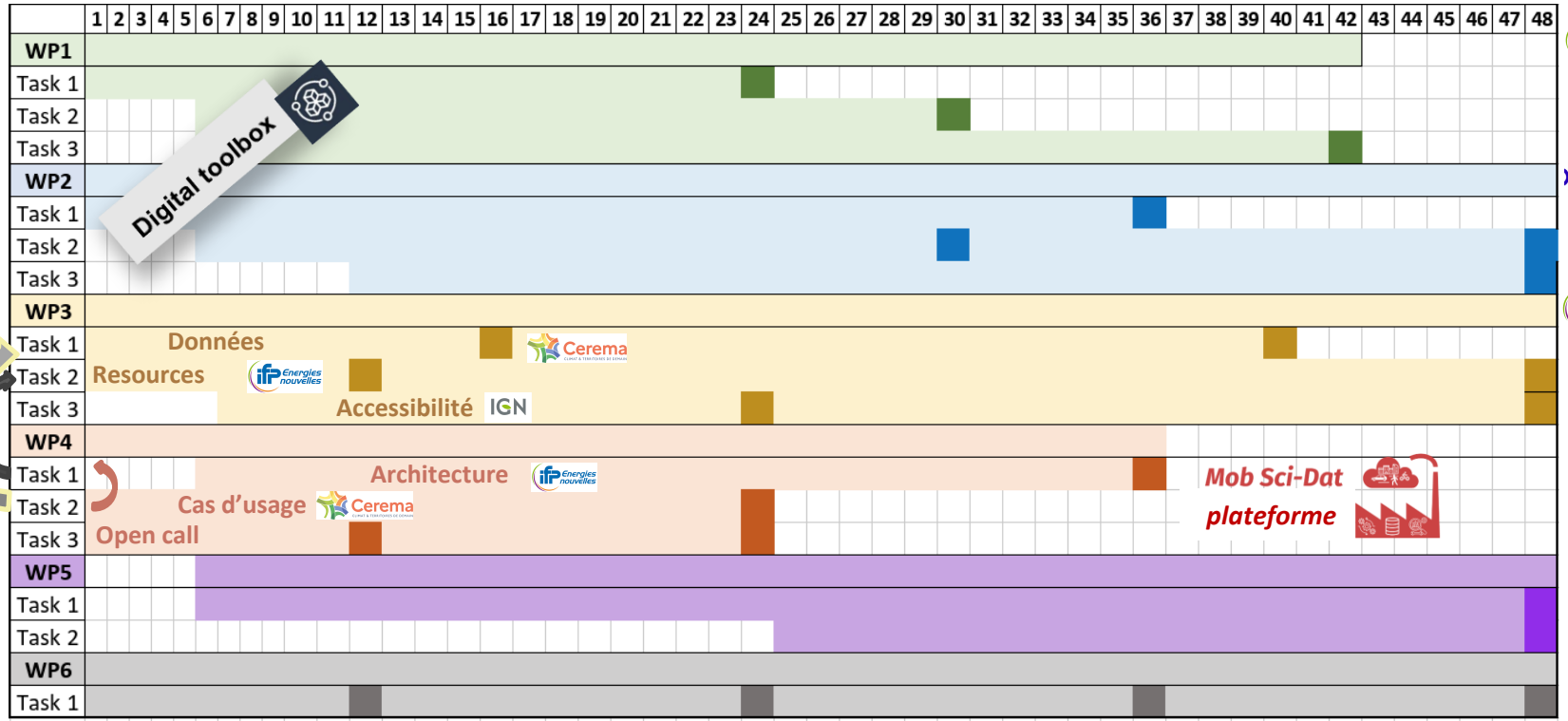
WP5 : Diffusion, communication et exploitation

T4.1 : Dissémination and communication (interne + externe)

T4.5 : Exploitation (pédagogiques, territoriales, publiques,…)

WP6 : Gestion de projet

Leaders



Mob Sci-Dat Factory

Description du projet

3. Organisation du partenariat

Mob Sci-Dat Factory



Le partenariat : multidisciplinaire et complémentaire

… **autour** de l'informatique, de l'ingénierie informatique, de la science des données, des transports et des acteurs sociétaux

… des **compétences** dans : les technologies de mobilité, de transport, d'énergie, de données et de cloud

Expertise	INRIA	IFPEN	UGE	CEREMA	IGN
Data measurement	x			x	
Data handling and analysis	x	x	x	x	x
Profiling behaviors in mobility	x	x			
Traceability and lineage			x		x
Transport modelling and profiling		x		x	
Data sanitization	x		x		x
Anomaly detection and diagnosis	x				
Reconstruction, modeling, enrichment	x	x	x		x
Data fusion and calibration			x	x	x
Validity and representativeness			x		
Simulation and data generator	x		x		x
Digital platforms	x	x		x	x
Accessibility: visualization, API, etc		x		x	x

… des acteurs publics territoriaux, de la recherche technologique et du transport

Activities and Interaction	Inria	IFP Energies nouvelles	Université Gustave Eiffel	Cerema	IGN
Territorial and public actors		x		x	x
Technological research actors	x	x	x		
Transport actors				x	x

L'organisation

INSA-Lyon
CNRS
Sorbonne Paris Nord

AGORA, ASCII, COATI, FUN, TRiBE

* Coordinator and leader of WP4, WP5, WP6, T1.2 and T4.3

Aline Carneiro Viana

Copilot and leader of WP1, WP3, T1.3, T3.2, T4.1, T5.1 and T5.2

Alexandre Chasse

Leader of WP2, T1.1, T2.2 and T2.3

Olivier Bonin

Leader of T2.1 and T3.3

Arnaud Le Guilcher

* Leader of T3.1 and T4.2

Christophe Damas

* Partenaires du PC1 et non-pilotes

Contributions du et au PEPR

Les délivrables et ses *milestones*

Digital toolbox

D1.1 - M24 : Intégration d'algorithmes de traitement de données brutes en bibliothèques
D1.2 - M30 : Rapport recommandant les bonnes pratiques en matière d'anonymisation des données
D1.3 - M42 : Modèle et son instanciation avec des données réelles

D2.1 - M36 : Méthodes et algorithmes associés pour la fusion de données
D2.2 - M30 : Outils logiciels avec cartes pré-calculées
D2.3 - M48 : Outils et méthodes pour calibrer le modèle à partir de données dynamiques

Mob Sci-Dat Factory

D3.1a - M16 : Documentation de la structuration des données
D3.1b - M40 : Outil de génération de données
D3.2 - M12 et M48: Infrastructure informatique fonctionnelle et alimentée de la première source de données
D3.3a - M12 : Première version de l'interface permettant l'accès aux données et bibliothèques
D3.3b - M48 : V. finale de l'interface intégrant les données produites et l'accès aux bibliothèques produites

D4.1 - M36 : Diagramme schématique et documentation des éléments constitutifs de *Data and Toolbox Factory*
D4.2 - M24 : Ensemble des besoins exprimés par les décideurs dans un document « besoin utilisateur »
D4.3a - M12 : Liste des thèmes identifiés pour le premier appel ouvert
D4.3b - M24 : Liste des thèmes identifiés pour le deuxième appel ouvert

D5.1 - M48 : Publications du projet, rapports techniques et description des actions de diffusion
D5.2 - M48 : Description de l'exploitation des sorties de Mob Sci-Dat Factory par des acteurs externes

T6.1 – M12, M24, M36 et M48 : Rapports annuels du projet

KPIs – contributions

- Infrastructure cloud de partage et le traitement sécurisé
  - *Digital ToolBox* : bibliothèques, outils, simulateurs…
    - Jeux de données anonymisées, traitées
  - Gouvernance d'accès et outils de visualisation de données

- Création d'une base de traitement commune, afin…
  - d'accélérer les travaux de recherche sur la mobilité
  - d'éviter la maîtrise de la chaîne de traitement et les biais

- Concevoir une base de développement open source adaptée « d'intérêt général »

KPIs – contributions

- Communication et exploitation
  1. Publications et documentations (outils et jeux de données)
     - accessibles depuis l'interface web de Factory
  2. Interventions (nationaux et/ou internationaux)
     - Conférences, GTs, acteurs territoriaux, publiques, EU …
  3. Autres interactions :
     - EONA-X/GAIA-X,
     - *EU Mobility Data Space initiatives (CSA PrepDSpace4Mobility)*
     - EU CARA pour des solutions de transport
     - *Autres PEPRs : Cloud, IA, Réseaux du Futur, …*

Mob Sci-Dat Factory: Une plateforme destinée à durée

- 1ere (à notre connaissance) plateforme scientifique de partage et traitement de données et d'outils, accessible à des acteurs publiques et territoriaux

- Une preuve de concept pour la conception d'un *Mobility Data Hub* en France

Contributions aux PEPR

3. Les contributions et de mise en œuvre dans les territoires

Mob Sci-Dat Factory



Les interventions

L'objectif du WP3 et WP4 et avec le savoir-faire auprès des :

- … des autorités locales, des acteurs territoriaux et nationaux
- … de l'association l'ATEC ITS (Les Acteurs de la Mobilité Intelligente)
- … des GTs du *Comité Scientifique de France Mobilités*
- … de l'Agence d'Innovation pour le Transport (AIT)
- … de la Direction Interministérielle du Numérique (DINUM)
- *… du projet Belgrand-GEBS (Grand Equipment de Base de Données)*

Les contributions

- Montrer l'utilité des données (quand bien traité) à répondre aux questions des cas d'usage identifiés (WP4.2)
- Orienter les acteurs à collaborer au partage des données
- Montrer que les outils peuvent être exploité par les territoires
- Permettre aux acteurs publics de tester et d'évaluer les politiques publiques en faveur d'une mobilité décarbonée


FRANCE 2030
PROGRAMME DE RECHERCHE
DIGITALISATION ET DECARBONATION DES MOBILITES
Merci pour votre attention !