

Quelle boîte à outils pour observer, comprendre, agir ?

J.P. Orfeuil

Professeur émérite, Université Paris Est,
IVM, Paris

Plan

- **La situation**
- **Avantages et inconvénients des données d'enquête et des données d'exploitants**
- **Avantages et inconvénients des données massives**
- **Exercices**

La situation

Un contexte factuel:

1 Les enquêtes « déplacements » ont beaucoup d'intérêt mais aussi des limites

2 Le big data est une opportunité à saisir

Nb de données générées/jour / véhicule connecté >> nb de données d'une EMD / EMC tous les 5-10 ans.

3 Les moyens de déplacement se sont diversifiés, il y a beaucoup de « petits modes » populaires dans les médias dont la somme peut être significative

4 Des territoires sont très mal décrits, par exemple « le » périurbain, et pas la mosaïque des périurbains

La situation

Une croyance assez répandue dans de nombreux domaines à l'égard des données massives:

Avec cette abondance de données, plus besoin d'hypothèses, de problématiques et de modèles, le data mining, les réseaux de neurones, le deep learning, etc. vont permettre d'extraire tout le « jus » des données et de révéler les structures sous jacentes, sans questionnement ou modèle a priori

On a déjà connu ça dans les années 70 avec l'analyse factorielle des correspondances

Les données massives: une nouvelle révolution scientifique?

Source: the fourth paradigm, data intensive scientific discovery

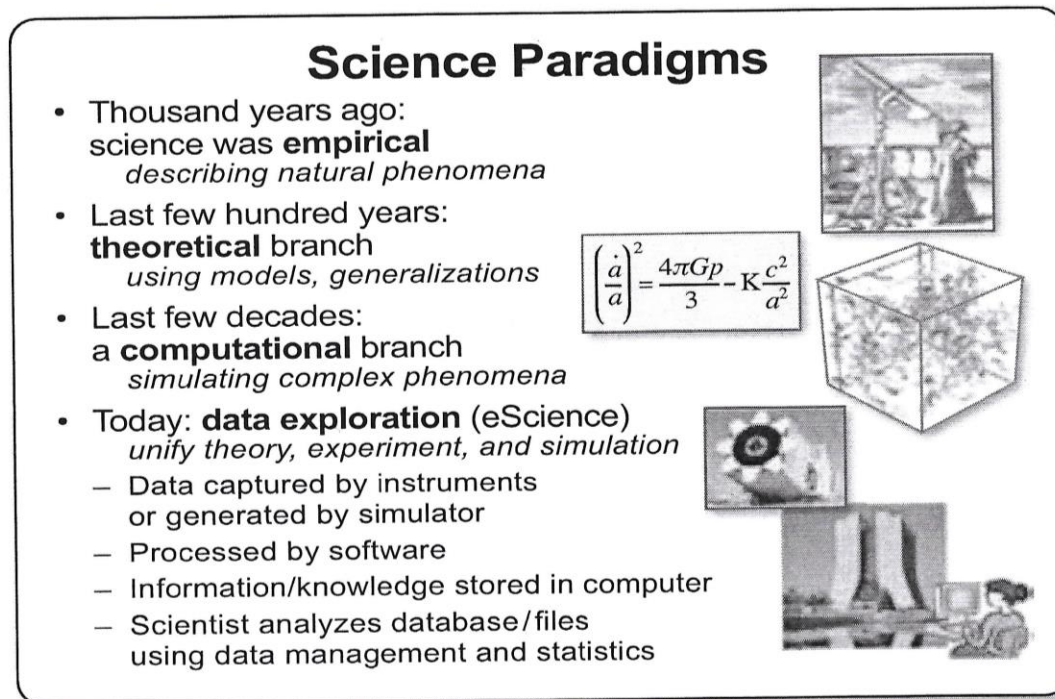


FIGURE 1

Avantages et inconvénients des enquêtes et des données d'exploitation

Les EMD / EMC ont des qualités, mais aussi des limites

Les avantages

**On a pensé les informations (#données) à recueillir dans une
optique de connaissance et de moyen terme**

**C'est l'individu (et même parfois les membres du ménage) qui
est observé. Il n'est pas « saucissonné » en modes.**

**Les comparaisons dans le temps et l'espace sont possibles
grâce à un corpus de données standardisées**

(France et monde)

**Des exploitations sont potentiellement systématisables
(Deed, budget logement+ transport...)**

Les EMD / EMC ont des qualités, mais aussi des limites

Les inconvénients

Coûteuses

Trop peu fréquentes / rythme des évolutions (Quoi que?)

Trop peu étendues dans le temps (1 journée observée / individu en standard) pour représenter les programmes d'activité des individus et les possibilités de vivre sans voiture personnelle

Pour certains objectifs (CO2 en particulier), la limitation à un territoire est un problème

Des effectifs faibles dès qu'on veut mettre le focus sur un « petit » territoire ou sur un « petit » mode

Le gros inconvénient perçu des EMD/EMC

- **Les résultats des exploitations standard sont pauvres et peu problématisés:**

ils répondent à des questions qu'on ne se pose pas, et ne répondent pas à des questions qu'on se pose

- **Les enquêtes sont à la base de nombreux travaux intéressants, mais très peu connus.**

**Il y aurait intérêt à rassembler les travaux menés à partir des EMD/EMC/
EGT/ENT**

Les enquêtes « quali » sont souvent intéressantes, mais les effectifs sont limités, la généralisation problématique et les traitements longs

Les différentes techniques d'IA et de traitement automatique du langage naturel peuvent sans doute aider à « débroussailler » une question

Les données administratives et les données d'exploitants sont souvent riches, mais le plus souvent modales

Avantages et inconvénients des données massives

Les données massives / EMD, c'est la possibilité du microscope

Exemples

**Connaissance très fine des vitesses à la minute
et en tout point**

**Connaissance d'éléments sur la mobilité des
habitants de telle ou telle petite commune**

**Possibilité de savoir rapidement si la croissance
d'un mode « marginal » est due à sa diffusion
dans la population ou à un usage plus intensif
par les « early adopters »**

Les données massives/ EMD, c'est la quasi immédiateté des diagnostics

Exemple

L'appréciation très rapide de la variation des temps de parcours consécutive à une politique

Mais aussi

Le risque d'oublier les transformations structurelles de moyen / long terme

L'usage des données massives risque d'inciter au court-termisme?

Mais l'exploitation des données massives « à la manière des EMD » a (aura) aussi des limites

- **Les data « transport » sont recueillies mode par mode, pas individu par individu. La notion de ménage est absente. Ni l'accessibilité publique, ni la représentativité ne sont acquises**
- **Les « data » ne sont pas toujours des « data primaires », et il est parfois compliqué de le savoir**
- **Les data de type suivi « smartphone » sont une mine très intéressante, mais**

**Leur usage est (et sera?) probablement (et légitimement)
fortement régulé.**

**Les motifs, les caractéristiques socio-éco des personnes seront
imprécis ou limités**

Exercice 1: 11 dimensions de la mobilité

Sur quelles dimensions pourrait-on progresser?

- **En mobilisant mieux les données existantes**
 - **En améliorant les EMD/C**
 - **En recourant au big data?**

11 dimensions de « la » mobilité pour comprendre les individus et évaluer des politiques(à court terme)

- **Conditions** générales de la mobilité (je suis à 30 minutes de...,) déclinées par mode (et par période horaire)
- **Injonctions** de mobilité (un emploi à 1 heure est un emploi acceptable, les premiers commerces sont à 5 km, il faut ramener mon enfant à 17h...)
- **Aptitudes et compétences personnelles** de mobilité: permis, pratique du vélo...
- **Pratiques**: les déplacements effectués sur le territoire observé, leur coût(argent, temps)
- **Pratiques de sortie de ce territoire: courts séjours, vacances, leur coût**
- **Marché des déplacements**: parts modales, types de liaisons, parts de chacun dans le marché
- **Enveloppe** des pratiques, en temps, argent, fatigue: la limite supérieure de ce qu'on peut accepter. Elle détermine ce à quoi on a accès

11 dimensions de « la » mobilité pour comprendre les individus et évaluer des politiques (à court terme)

Le vécu des pratiques assis ou collé-serré dans les TC,
congestion, sentiment de danger

La mobilité déléguée: livraisons, visites reçues, une
nounou pour ramener mon enfant

La mobilité virtuelle: gestion / loisir de chez soi, du bureau

Représentations collectives: le bien, le mal, les croyances

Les dimensions de la mobilité (à moyen/ long terme)

- Le suivi des lieux de résidence, d'emploi, d'études entre deux enquêtes
- Le suivi des changements de niveaux d'équipement

Dans ce domaine des approches fondées sur des panels seraient utiles

Exercice 2

**Quelles sont les « meilleures » données
pour répondre à des questions simples**

Quelques questions simples

- A qui profite le tarif unique dans les transports urbains, départementaux, régionaux?
- De bonnes compétences de mobilité améliorent-elles le niveau de vie des personnes?
- Quel est le bassin de vie de telle ou telle commune de banlieue lointaine?
- Quel est le potentiel de substitution du VAE à la voiture?
- Quelle part des déplacements ou des km VP est massifiable?
- La surmobilité est-elle le fait d'un petit nombre qui fait un très grand nombre de km, ou de tous qui en font « un peu trop » les uns les autres?

Merci de votre attention

Ce qu'on suit bien, ce qu'on peut améliorer avec les données massives, ce qui restera hors champ

Ce qu'on suit bien et qu'on peut beaucoup améliorer

Les pratiques de mobilité

Le marché des déplacements

Ce qu'on peut connaître aisément, mais qu'on utilise peu

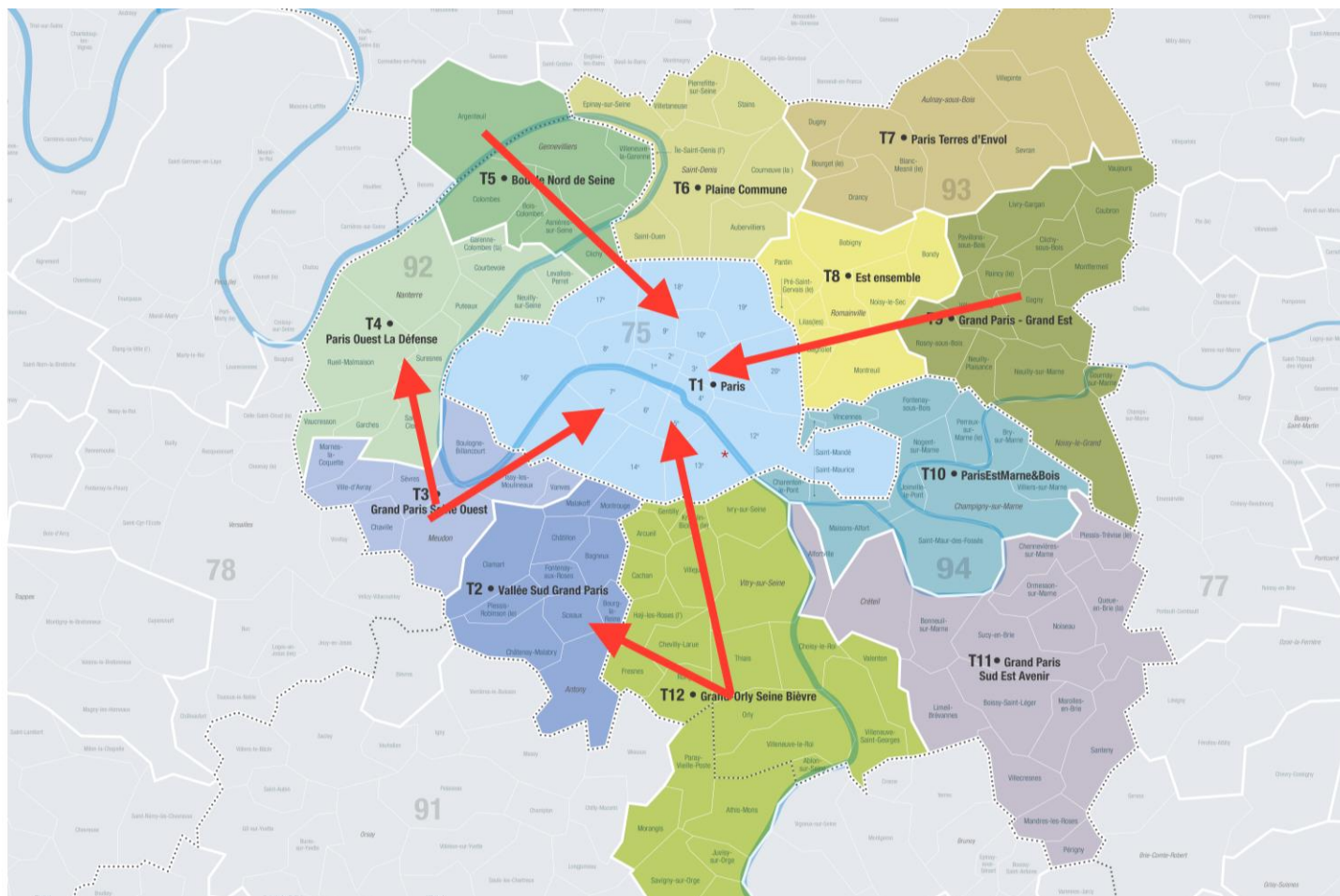
Les conditions générales de la mobilité

Les enveloppes des pratiques la mobilité

Ce qu'on suit mal, et qu'on continuera à mal connaître sans efforts supplémentaires

Les injonctions

Les aptitudes et compétences

[illegible]

Ou mesurer la dispersion des flux?

L'exemple des déplacements entre domicile et travail des actifs domiciliés en Seine et Marne

500 communes en Seine et Marne et environ 1300 en IdF

En fait environ **36000** relations pour **611.800** actifs au total, soit **17** actifs *en moyenne* par relation

126.500 actifs travaillent dans leur commune de résidence, soit **181** / commune *en moyenne*

485.300 changent de commune,

137.800 sont sur des flux >100 actifs

(répartis sur 700 relations, soit environ 200 actifs / relation)

Il reste **347.500 (57 %)** sur des relations à <100 actifs, avec en moyenne **10** actifs par relation

Une absence de permis pénalisante sur le marché du travail

Taux de chômage selon le diplôme (%)

	Scolarité obligatoire	Brevet	Bac	> Bac
Centre, permis	20	11	12	6
Centre, sans permis	36	25	15	15
Banlieue permis	13	8	6	4
Banlieue sans permis	29	20	24	14
Hors des grandes villes permis	15	6	6	4
Hors des grandes villes sans permis	36	29	36	9

L'exploitation en temps réel des données massives rend possibles et peu coûteuses des politiques d'exploitation innovantes

Par exemple

Un péage urbain rendu possible par une forte baisse des coûts de mise en œuvre.

Une tarification en temps réel et à l'échelle de l'îlot du stationnement sur voirie

L'annonce de la charge du prochain bus qui inciterait ceux qui n'ont qu'un petit déplacement à faire de le faire à pied