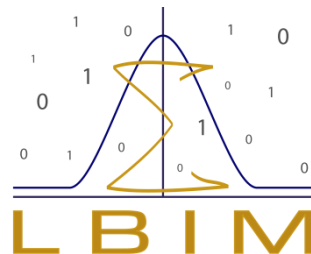


Big Data : c'est quoi ?

Erik A. Sauleau
ea.sauleau@unistra.fr

Pôle de Santé Publique - Santé au Travail - Hygiène hospitalière
Hôpitaux Universitaires de Strasbourg

Laboratoire de Biostatistique – Université de Strasbourg
ICube UMR CNRS 7357



En résumé

- **Définition** à finaliser
- **Enjeux** majeurs
 - Scientifiques
 - Sociétaux

Introduction

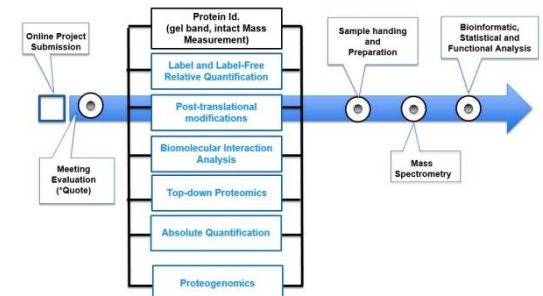
- « Bigdata : c'est quoi ? » → je ne sais pas
- Traduction de la *Commission générale de terminologie et de néologie*

mégadonnées

données structurées ou non dont le très grand volume requiert des outils d'analyse adaptés

- Retient deux caractéristiques
 - Axée sur leurs analyses
- Acception *de facto* différentes
 - Radiologie ou génomique (pipeline)
 - Entrepôt de données

Pipeline to work with the proteomics Platform



Définition : 5 V historiques

- De 3 à 7 V
- Définition *managériale*
 - Valeur → données sensibles de l'entreprise
 - Volume* → techniques particulières
 - Variété* → types différents
 - Vélocité* → rapidité d'acquisition et/ou traitement
 - Véracité → contrôles qualité
- Méthodes de gestion des données et d'analyse

Et dans les données de santé ?

- Définition **inclusive** des données de santé
 - Dossier du patient
 - Dossier médical hospitalier
 - Examens complémentaires avec leur résultat
 - ...
 - Dossiers partagés
 - Objets connectés
 - Blogues
 - ...

Et dans les données de santé ?

- Valeur
 - Monétaire (SIGAPS)
 - Données personnelles de sujets
 - Enjeux éthiques et réglementaires
- Volume
 - Fouille de données, intelligence artificielle
- Variété
 - Texte structuré ou non
 - Images et comptes-rendus
 - Objets connectés
 - ...

Et dans les données de santé ?

- Valeur
- Volume
- Variété
- *Vélocité* (à venir)
 - Bases de données pour le soin
 - Base de données pour la recherche (extraction)
- Véracité
 - Contrôles qualité et processus d'assurance qualité
 - Normalisation et certification des automates, appareils médicaux, ...

Et dans les données de santé ?

- Valeur
- Volume
- Variété
- *Vélocité*
- Véracité : donnée par rapport à la mesure
- Véridicité : mesure par rapport au concept
- **Mobilité** (se mouVoir)
 - Situations d'analyse
 - Ré-utilisation (vs recueil *ad hoc*)

Big data : nouvelle définition ?

- Poïkiloméga données (ποικίλος=varié)
- Plus simple

méga données (hétérogènes)

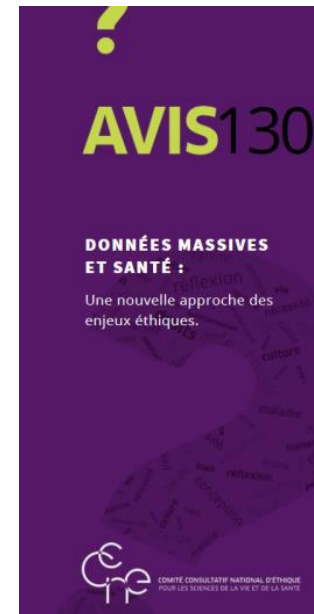
données structurées ou non dont le très grand volume et la grande variété de types requiert des outils de gestion et d'analyse adaptés



gestion technique mais aussi éthique, réglementaire ...

Enjeux éthiques

- Éthique comme partie intégrante de la démarche scientifique
- Retrait de Tay
 - Outil de conversation pour 18-24 ans
 - Après 96.000 tweets (8 heures) → propos sexistes, révisionnistes
- Avis 130 du Comité Consultatif National d'Éthique *Données massives en santé* (29/05/2019)
 - Donnée de santé = prise en charge, objets connectés des patients, données des téléphones mobiles, ...
 - Vigilance sur l'égalité d'accès aux avancées, absence de discrimination
 - Garantie humaine à chaque étape de l'analyse des données
 - Secret médical, responsabilité médicale, relation personnelle soignant-soigné
 - Importance de la **gouvernance de la donnée**



Enjeux du futur (demain)

- Enjeux **éthiques** et **réglementaires**
- **Dommages sociaux**
 - Vulnérabilité, dépendance
 - Données sensibles
 - Individu → RGPD
 - Groupes
- **Conservatisme**
 - **Ré-utilisation** des données
 - Nécessité d'historiser le contexte
 - La donnée n'est **pas la représentation du réel** mais d'un certain réel → conceptualisation relationnelle



Enjeux du futur (demain)

- **Entrepôts de données**
 - Réalisation, gestion
 - Interopérabilité
- Processus **décisionnel**
 - Numérique (sur les données)
ou
 - Symbolique (systèmes experts, exemple : **Mycin**)
 - Décision « en direct » → sur le recueil de soin
 - **Intelligence artificielle**

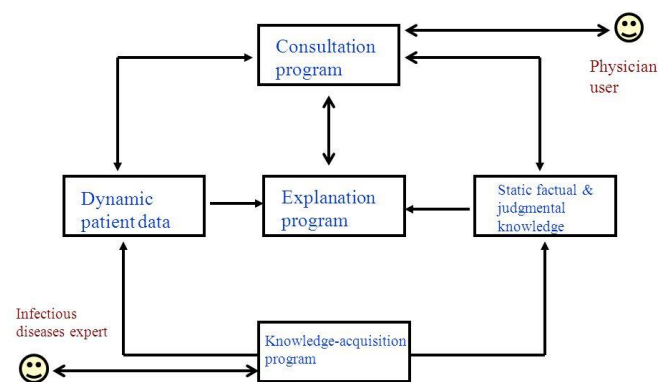
Enjeux éthiques et réglementaires

- Révolution **épistémologique**
 - De la déduction à l'induction



Kuhn T. *La structure des révolutions scientifiques*. Flammarion, Paris, 1983.

The MYCIN Architecture



Merci de votre attention

