



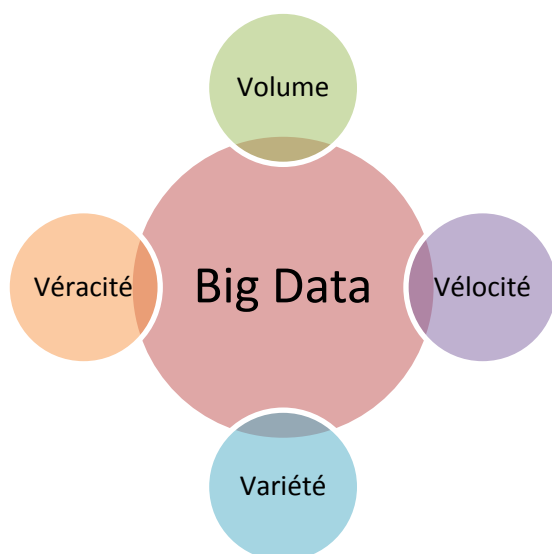
LE BIG DATA EN SANTE

En santé comme dans de nombreux autres domaines, les progrès technologiques ont fait exploser la quantité d'informations recueillies à chaque instant. Ces avancées technologiques ont ainsi fait augmenter de manière exponentielle le volume de données disponibles. Cela représente une aubaine pour la recherche en santé puisque les données massives (Big Data) sont une source presque inépuisable de nouvelles connaissances qui sont de plus en plus indispensables à l'innovation et donc aux progrès médicaux.

Quelle définition pour le Big Data ?

Littéralement le Big Data signifie grosses données ou méga données ou encore données massives. Si aucune définition précise n'est donnée au Big Data, ce terme désigne un ensemble très volumineux de données qu'il va falloir stocker en vue de les analyser. Mais au-delà de cette notion de volume, le Big Data est une réelle opportunité d'obtenir des connaissances sur des nouvelles données et sur des nouveaux contenus permettant ainsi par exemple d'ouvrir la porte à des innovations médicales.

Les 4 V du Big Data



Volume : Evolution exponentielle (OMIC, Objets connectés, images)

Vélocité : Production au fil de l'eau et réutilisation en temps réels pour une exploitation au plus près des utilisateurs

Variété : Données extrêmement hétérogènes (soin, recherche, médico-administratif)

Véracité : Qualité très variable selon les sources et les processus de recueil des données. La vraie richesse du Big data vient de la complémentarité des sources d'information



Les défis du Big Data

Les gros volumes de données que nous avons à notre disposition donnent lieu à de nombreux défis en termes de stockage, de capacité d'exploitation et d'analyse. La principale problématique est l'hétérogénéité des données collectées qui sont de plus en plus diverses.

- ❖ Premièrement diverses de par leur format (images 2D ou 3D, valeurs numériques, valeurs textuelles...)
- ❖ Mais aussi diverses de par leur nature (physiologique, biologique, sociale...)
- ❖ Et enfin diverses de par leur provenance de différents systèmes d'information

Pour faciliter leur traitement, ces informations doivent être acquises de manière structurée avant d'être intégrées dans des bases de données ou des entrepôts de données. Grâce à des standards, certains hôpitaux sont équipés pour compiler toutes les données collectées dans des entrepôts de données biomédicales interrogeables par les chercheurs.

Les perspectives du Big Data

Les données collectées à long terme sur des grandes populations pourront permettre d'identifier des facteurs de risque pour certaines maladies (cancers, diabète, maladies neuro dégénératives...). Ces facteurs pourront ensuite servir à la construction de messages de prévention et mettre en place des programmes à destination des populations à risque. Le Big Data pourrait donc permettre le développement de systèmes d'aide au diagnostic et d'outils permettant l'adaptation des traitements.

Le contexte de notre projet interrégional

Le but est de rationaliser et d'organiser au sein des établissements hospitaliers **l'exploitation des données des patients pour la recherche clinique** :

- ❖ Etablir le cadre organisationnel et réglementaire (aspect éthique, déontologique et juridique, CNIL)
- ❖ Déployer et mettre en œuvre des entrepôts de données biomédicales
- ❖ Mutualiser les compétences et les expertises sur l'exploitation des données numériques en santé
- ❖ Réaliser des études de faisabilité et de recrutement multicentriques



Les objectifs du projet et les actions à mettre en œuvre

L'objectif est d'étendre le déploiement et de partager :

- ❖ Une gouvernance, des procédures, une méthodologie d'exploitation
- ❖ Des outils pour un écosystème cohérent

Pour cela il est nécessaire de coordonner et d'animer un réseau structurant pour :

- ❖ Partager les compétences
- ❖ Réaliser une exploitation multicentrique
- ❖ Intégrer à d'autres réseaux et sources de données
- ❖ Utiliser des Registres, données de soin primaire, SNIIR-AM, Objets connectés

Les actions à mettre en œuvre :

- ❖ Déployer l'infrastructure des outils
- ❖ S'appuyer sur un industriel reconnu
- ❖ Mutualiser les acquis techniques
- ❖ Organiser/participer à des actions de formation aux nouveaux métiers des CDC et du Big Data en santé