

# BigData, intelligence artificielle, blockchain: guide pratique

Pr CHRISTIAN LOVIS<sup>a,b</sup>, CHRISTOPHE GAUDET-BLAVIGNAC<sup>a,b</sup>, RAPHAËL CHEVRIER<sup>a,b</sup>,  
ARNAUD ROBERT<sup>a,b</sup>, DAVID ISSOM<sup>a,b</sup> et Dr VASILIKI FOUFI<sup>a,b</sup>

Rev Med Suisse 2018; 14: 1559-63

La société en général, la médecine en particulier, sont emportées par la vague de la digitalisation. Ce phénomène s'appuie sur une production d'immenses quantités de données, parfois du fait de la dématérialisation de processus, comme l'écriture ou la photographie, parfois du fait de l'acquisition de nouvelles données, comme la géolocalisation. Ceci nécessite de nouveaux instruments pour le transport, le stockage et le traitement de l'information. Cet article présente quelques enjeux et instruments utilisés, telles les techniques de traitement du langage naturel, de l'intelligence artificielle et des bases de données en graphes. Enfin, nous décrivons brièvement la technologie de la *blockchain*, qui est de plus en plus proposée en médecine pour des processus non monétaires, tels que l'échange de données ou la gestion du consentement.

## Bigdata, artificial intelligence and blockchain for dummies

*Digitalization is transforming every aspect of life, it is also transforming deeply medicine. The digitalization era is characterized by a large production of new data streams while existing processes are progressively migrated, such as writing or imaging. The very large and fast-growing amount of data available requires new storage, transport and analytical tools. This paper presents some of them, such as natural language processing, artificial intelligence, and graph databases. A short introduction to the technology of blockchain is also provided, as it is increasingly used in some non-monetary transaction in medicine, such as data exchanges and consent management.*

## INTRODUCTION

Données massives, intelligence artificielle et *blockchain*... voilà des mots-clés qui font régulièrement la une. Et il n'y a rien d'étonnant à cela. La digitalisation rapide et massive de toutes les activités humaines, la mesure de plus en plus fine de toutes les caractéristiques de l'environnement dans lequel les humains vivent, tout cela nous rapproche de plus en plus de cette pensée si bien exprimée par Pierre Simon Laplace dans son *Essai philosophique sur les probabilités*: «Une intelligence qui, à un instant donné, connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était suffisamment vaste pour soumettre ces données à l'analyse, embrasserait dans la

même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome; rien ne serait incertain pour elle, et l'avenir, comme le passé, seraient présents à ses yeux». C'est le vieux rêve, ou peut-être le cauchemar, du déterminisme, de la capacité à comprendre, mais aussi à prédire. Pour cela toutefois, il faut arriver à lire, et à comprendre. Lire les chiffres, lire les images, mais surtout lire les textes. Car les données structurées, comme une température, un poids ou encore une glycémie, ne représentent qu'une petite île à la surface de l'immense océan du savoir et de l'information écrite. Dans cet article, nous ferons un très rapide survol des techniques de traitement du langage naturel, puis de quelques aspects liés à l'intelligence artificielle. Les bases de données en graphes, qui ont la puissance nécessaire pour représenter la complexité multidimensionnelle des données que nous traitons seront également présentées. Nous terminerons avec une brève description des technologies de la *blockchain*, que nous utilisons, notamment pour implémenter un consentement centré sur le citoyen, partagé et dynamique.<sup>1,2</sup>

## TRAITEMENT AUTOMATIQUE DU LANGAGE NATUREL

Le Traitement Automatique du Langage Naturel (TALN) est un domaine multidisciplinaire impliquant la linguistique, l'informatique et l'intelligence artificielle. Le langage étant présent partout, écrit ou oral, les applications du TALN sont multiples et variées (correction orthographique, traduction automatique, extraction d'information, détection d'opinion, filtrage de l'information, synthèse de la parole, reconnaissance vocale).

Dans le passé, tous les systèmes de TALN étaient basés sur des méthodes symboliques. Celles-ci se basent sur des grammaires et des ressources lexico-sémantiques construites manuellement et utilisent tous les différents niveaux d'analyse comme la phonétique, la syntaxe ou la sémantique. Les systèmes basés sur des méthodes symboliques offrent une large couverture des phénomènes linguistiques étudiés et peuvent garantir une grande précision dans la détection des structures. En particulier, les systèmes de reconnaissance qui sont basés sur les automates à états finis sont très efficaces.<sup>3</sup> De l'autre côté, il est vrai que la construction de ressources linguistiques est une tâche coûteuse, mais une fois construites ces ressources sont gérables et réutilisables par les utilisateurs.

Depuis le début des années 1990, les applications pratiques se sont multipliées, conduisant à une utilisation plus large des méthodes statistiques et probabilistes (regroupées parfois sous le terme de méthodes stochastiques), produisant des

<sup>a</sup> Service des sciences de l'information médicale, HUG, 1211 Genève 14,

<sup>b</sup> Université de Genève, 1211 Genève 4

christian.lovis@unige.ch | christophe.gaudet-blavignac@unige.ch

raphael.chevrier@hcuge.ch | arnaud.robert@unige.ch

david.issom@unige.ch | vasiliki.foufi@unige.ch

résultats plus rapidement. En 1993, les deux numéros spéciaux de *Computational Linguistics* ont été consacrés aux grands corpus et la phrase qui a dominé était: «Le vent souffle désormais au quantitatif».<sup>4</sup>

Le principe de ces méthodes est de faire de la langue un objet mathématique et d'appliquer différentes formules probabilistes et statistiques aux données de la langue (mots, phrase...). Une telle application pourrait constituer la fréquence d'apparition des mots dans un texte pour en déduire les thèmes principaux de celui-ci.<sup>5</sup> L'avantage de ces méthodes est qu'elles sont rapides à implémenter et les résultats sont très satisfaisants. L'inconvénient majeur est qu'elles ne sont efficaces que sur des corpus de données de très grandes tailles. De plus, elles peuvent nécessiter l'annotation manuelle de gros corpus pour donner des résultats satisfaisants.

Les deux méthodes peuvent être employées pour le traitement de données textuelles, soit de façon indépendante ou hybride. Dans le domaine médical en particulier, le TALN est essentiel pour extraire l'information précieuse stockée sous forme de texte libre et la transformer en données structurées.<sup>6</sup> Des outils de TALN ont déjà été appliqués en médecine pour la désidentification de documents, la détection d'effets indésirables, la prise de décision, et autres.

## INTELLIGENCE ARTIFICIELLE

Afin de pouvoir analyser la masse grandissante de données médicales, il faut se tourner vers l'intelligence artificielle (appellation populaire pour désigner un panel d'outils à la frontière des statistiques et de l'informatique).

L'apprentissage automatique (une sous-classe de l'intelligence artificielle) permet de créer des algorithmes sans qu'ils aient été explicitement codés par un humain. Il suffit pour cela de montrer à l'ordinateur une série d'exemples (x, y) où x sont les données et y les cibles. Un exemple concret en médecine serait: x une radiographie et y le diagnostic associé.

La variété d'algorithmes existants propose différentes façons d'aborder ce problème. Le point commun de toutes ces méthodes est qu'elles essaient d'approximer une fonction inconnue qui décrit la relation entre x et y. Plus il y a d'exemples à disposition plus l'approximation est proche de la véritable relation, jusqu'à éventuellement atteindre une limite inhérente au modèle. Le nombre de données requises pour une performance raisonnable a tendance à augmenter avec la complexité de la tâche (la complexité de la fonction à approximer).

Les applications de l'apprentissage automatique dans le domaine médical se divisent en deux catégories:<sup>7</sup>

- Les tâches que les médecins (experts) font de façon performante comme l'analyse d'une radiographie.<sup>8</sup> Dans ce cas de figure, les algorithmes auront tendance à essayer d'approximer la performance des experts. L'intérêt étant que sans délai et à moindre coût un avis médical de qualité peut être généré.
- Les tâches qui dépassent même les compétences des experts. Il s'agit principalement de tâches de prédiction ou d'évaluation de risques. Par exemple, prédire le risque

qu'une personne présente un problème cardiaque sur une période donnée. Dans ce contexte précis, l'apprentissage automatique proposera des manières inédites de soigner, prévenir ou détecter certaines maladies.

Les défis à venir pour l'apprentissage automatique en médecine sont présents dans plusieurs contextes. Il y a des barrières légales construites autour des données médicales qui limitent le partage des données et donc la construction de bases de données suffisamment complètes pour permettre un apprentissage de qualité.

Du point de vue de l'implémentation de ces techniques dans la pratique médicale, pour une collaboration efficace entre l'algorithme et le médecin, il serait préférable que les algorithmes arrivent à justifier leurs décisions. Pour les modèles les plus simples, la raison pour laquelle un algorithme a fait un choix spécifique peut être révélée par l'analyse. Cependant, pour des modèles plus complexes, ces raisons peuvent ne pas être flagrantes.

Cette notion de feedback de l'algorithme est cruciale à la collaboration. Les algorithmes qui fournissent un feedback avec leurs prédictions sont dits bayésiens. Ceux-ci produisent une prédiction accompagnée de la probabilité que leur prédiction soit correcte. Cette probabilité peut être vue comme la confiance que l'algorithme a en sa prédiction.

Finalement, le domaine fait également face à une «crise de la reproductibilité». Dans une publication récente dans *Science*, ce problème est abordé. Moins de 10 % des publications fournissent les algorithmes, et donc la possibilité de reproduire les résultats, pour 30 % qui fournissent les données.<sup>9</sup>

## PROTECTION DES DONNÉES

La croissance de la technologie de l'information a offert aux chercheurs la capacité d'exploiter de grands volumes de données médicales. L'utilisation secondaire des données est vue par les chercheurs comme un catalyseur important de la recherche biomédicale.

Des initiatives fédérales comme le SPHN<sup>10</sup> (Swiss Personalized Health Network) ou le PHRT<sup>11</sup> (Personalized Health and Related Technologies) montrent qu'il y a une conscience forte que l'avenir de la recherche médicale passera par le partage à grande échelle de données liées à la santé telles que des données génomiques, métabolomiques et phénotypiques. Néanmoins, il a été démontré dans d'autres pays que de telles collaborations sont plus rares et difficiles à établir que ce que l'on pourrait attendre.<sup>12</sup> Une des raisons qui freine de tels partages est la possibilité de réidentifier les patients à partir de leurs données.

### Anonymisation

Le terme rencontré le plus souvent dans la littérature pour décrire le traitement des données afin de rendre la réidentification impossible est l'«anonymisation». Mais ce terme n'est pas aussi clair qu'il y paraît. En effet, si l'on consulte la littérature sur le sujet, on constate que la définition de ce terme

est variable. Les juristes, les chercheurs biomédicaux ou les ingénieurs en informatique ne sont pas d'accord sur ce qu'implique cette anonymisation.

S'il est clair pour la plupart que l'anonymisation consiste à rendre impossible ou très difficile la réidentification des patients à partir des données, il n'y a pas de consensus réel sur la méthode ni sur le seuil qui permet de dire que des données sont anonymes. Pourtant, des critères mathématiques existent. Fréquemment utilisés par les chercheurs issus de l'informatique, ils permettent de définir pour un set de données, un niveau d'anonymat d'un patient de ce set.<sup>13</sup> Mais ce genre de concepts, difficiles à mettre en pratique par des non-spécialistes, apparaît rarement dans des publications émanant du domaine biomédical ou juridique.

De plus, concernant spécifiquement les données de santé, de récentes études ont montré qu'elles ont ceci de particulier qu'un nombre très restreint de valeurs permet la réidentification univoque d'un patient sur un groupe très grand, à moins que l'on modifie ces données, diminuant par là même leur utilité.

Il est ainsi urgent qu'une terminologie commune soit définie pour parler d'anonymisation, afin de permettre aux différents acteurs de se comprendre les uns les autres. Une meilleure compréhension du domaine pourrait à elle seule améliorer la qualité des collaborations, la sécurité des données et la productivité de la recherche.

### Désidentification

Dans ce contexte, des réglementations ont été définies afin d'assurer la protection des données personnelles. Le défi de la gestion de données de santé a été lancé en 1996 par les règles HIPAA (US Health Insurance Portability and Accountability Act) et ses révisions. En Europe, le règlement général sur la protection des données (GDPR) a été approuvé en avril 2016 et implémenté en mai 2018. Ces textes fournissent un cadre légal et en même temps un guide pour l'usage et la divulgation des données personnelles.

En pratique, l'approche proposée par HIPAA est la désidentification des textes médicaux. Cela consiste à retirer des documents certaines informations identifiables (18 catégories en tout). Comme il a été prouvé que la désidentification manuelle de dossiers de patients est coûteuse,<sup>14</sup> l'automatisation de cette tâche avec des outils du TALN est fortement recommandée. En TALN, la désidentification est conçue comme une tâche de reconnaissance des entités nommées (noms de personnes, noms d'organisations, etc.).

La nécessité de techniques de désidentification est reflétée par le grand nombre de systèmes qui ont été développés dans ce but ces dernières années. Ils sont basés sur des règles formelles,<sup>15,16</sup> des modèles entraînés pour le traitement de textes (Conditional Random Field)<sup>17</sup> ou des approches hybrides.<sup>18,19</sup>

### BASES DE DONNÉES EN GRAPHES

Les systèmes de gestion de bases de données de types relationnels dominant le monde de l'informatique depuis des

décennies se sont révélés inadaptés au traitement des données générées en masse par les grandes entreprises du web telles que Amazon ou Facebook. Ces compagnies ayant amené à d'inédites fonctionnalités (par exemple, gestion de relations par nature flexibles entre personnes fortement connectées, systèmes de recommandations d'achats), de nouvelles problématiques sont apparues telles que le stockage de gros volumes de données distribués sur plusieurs serveurs et la visualisation rapide de ces données.

Un utilisateur de Facebook typique a un réseau d'amis, qui peut varier, s'étendre, se contracter. Chaque membre de ce réseau a lui-même un réseau de relations dans lequel chaque individu a la possibilité, par exemple d'ajouter des photos et de «taguer» dessus une ou plusieurs relations «amicales» ou d'«aimer» un post. Dans un tel cas, très courant dans un réseau social populaire tel que Facebook, les informations sur les relations autour d'un individu s'accumulent très rapidement. Il se trouve que les systèmes de gestion de bases de données relationnelles, de par leur nature nécessitant d'avoir un index et des tables non liées entre elles, sont incapables de fournir un accès efficace à de tels volumes d'informations. L'exploitation des données sur ces systèmes traditionnels s'effectue avec le «Structured Query Language», ou SQL, un langage de requêtes structurées qui ne permet pas de gérer de manière performante ces bases de données.

Les limites intrinsèques de ces dernières ont amené à la mise en place d'autres visions de la base de données, telles que les bases de données «clé-valeurs» ou les bases de données dites en graphe.<sup>20</sup> La **figure 1** montre une comparaison entre bases de données de type relationnel et de type graphe.

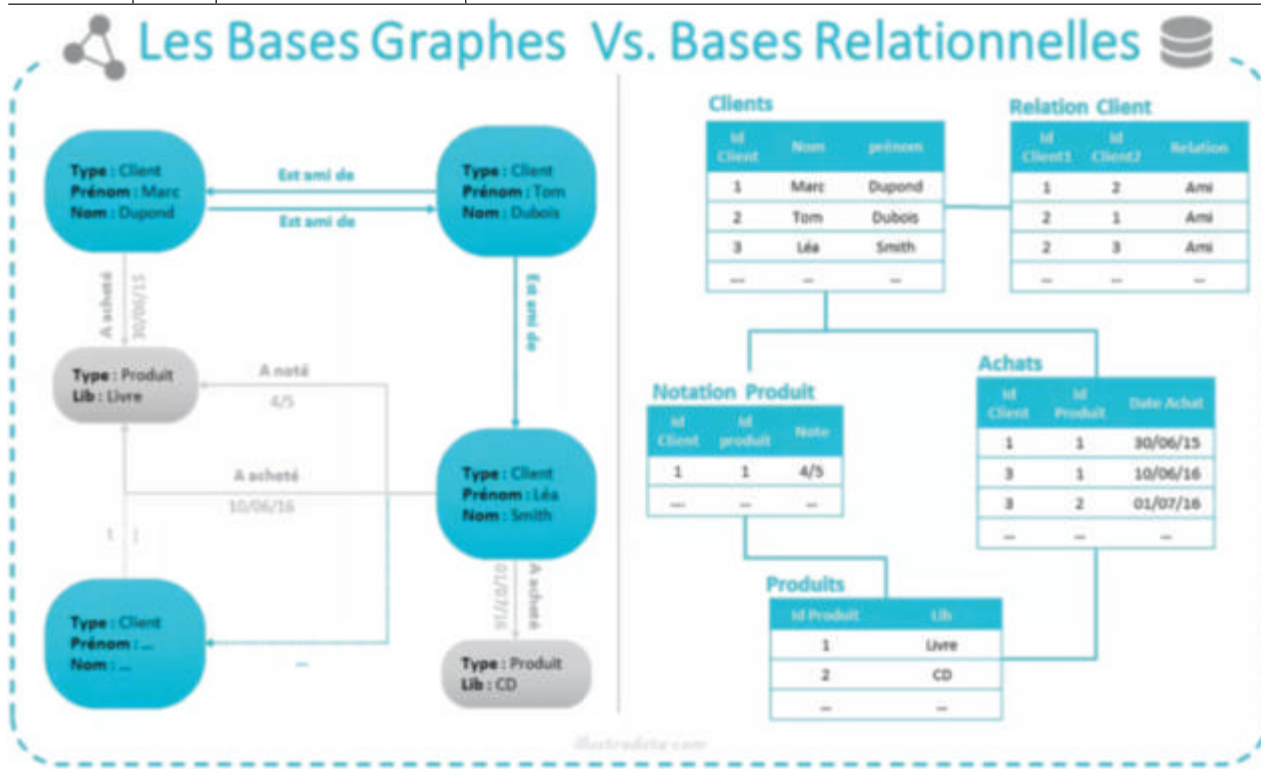
Les entreprises suscitées ont donc chacune créé leurs propres systèmes basés sur le paradigme du graphe, ce qui a amené à la popularisation du terme NoSQL. Le terme NoSQL représente donc les systèmes de stockage et de gestion de données en graphe ayant pour particularité de ne pas utiliser le langage SQL. Cependant, certains systèmes de gestion de bases de données (par exemple, CosmosDB) permettent de gérer les trois types de bases de données (clé-valeurs, tables, graphes).

Dans le domaine des bases de données médicales,<sup>21</sup> où la vision centrée sur le patient devient la norme et les informations liées à un patient sont flexibles et de types multiples, on peut facilement imaginer comment des systèmes en graphes pourraient faciliter le parcours des différentes relations (nœuds) qu'un patient possède. Se basant sur la théorie mathématique des graphes, ces systèmes permettent d'analyser rapidement les multiples facettes d'une relation thérapeutique.

### LA BLOCKCHAIN

L'histoire débute en 2008 lorsque le pseudonyme Satoshi Nakamoto, s'il existe vraiment, publie un article sur Internet dans lequel il décrit l'idée de la *blockchain*: une base de données publique et distribuée, synchronisée à intervalle régulier sur de nombreux ordinateurs, accessible à tous, crypté. Chaque modification est une transaction qui sera encodée sous forme d'un bloc, crypté et validé sur le réseau des autres nœuds avec les blocs précédemment vérifiés, avant de pouvoir

FIG 1 Bases en graphes<sup>27</sup>



être ajouté à la chaîne. Ainsi se construit la «chaîne de blocs», qui recense l'ensemble des transactions, vérifiable et inaltérable. La technologie de la *blockchain* peut être utilisée pour de nombreuses situations, notamment toutes les situations qui requièrent d'avoir un registre vérifiable et inaltérable des transactions: finance, actes notariés, contrats, etc. En 2009, Satoshi Nakamoto va utiliser cette technologie pour créer une nouvelle monnaie, le bitcoin. Cela donnera naissance aux cryptomonnaies. Aujourd'hui, la technologie de la *blockchain* commence à être utilisée également en santé, par exemple pour la sécurisation de certains échanges de données ou encore la gestion du consentement.<sup>22-24</sup> Dans ce dernier cas, l'implémentation de contrats intelligents (*smart contracts*)<sup>25,26</sup> dans la *blockchain* permet de changer complètement le paradigme actuel du consentement. Ainsi, les contrats intelligents permettent de créer des consentements centrés autour de la personne plutôt que d'un état: «être patient». Il devient entièrement dynamique, chaque personne peut suivre les consentements qui la concernent, les modifier en tout temps, être notifiées lorsqu'ils sont utilisés. Il y a de nombreuses autres propriétés, comme la création d'une autorité distribuée de confiance.

## CONCLUSION

La digitalisation entraîne la production de grandes quantités de données issues de l'activité des humains et de l'environnement dans lequel nous vivons. Elle entraîne également la production croissante d'informations et de connaissances, de publications. La quantité, la complexité, la vitesse de production ne sont que quelques-unes des raisons qui rendent ces

données et ces informations hors de portée de la capacité humaine, nécessitant ainsi de nouveaux instruments, dont la plupart restent à inventer. Le TALN et l'intelligence artificielle sont les deux instruments clés qui doivent accompagner cette révolution digitale, afin que tout cela puisse être utilisé par l'humanité. Les bases de données en graphes sont nécessaires pour répondre à la complexité croissante de la représentation de la sémantique de ces données. La *blockchain* est une technologie qui est aussi devenue indispensable. Non pas tant pour les cryptomonnaies que pour assurer la traçabilité et l'authenticité des transactions dans un monde où, tout étant sous forme digitale, tout peut se falsifier facilement. Ce type de technologie est donc nécessaire afin de créer des chaînes de confiance dans les cycles de production-analyse-interprétation de l'information en assurant l'intégrité irréversible de chaque étape.

**Conflit d'intérêts:** Les auteurs n'ont déclaré aucun conflit d'intérêt en relation avec cet article.

Les auteurs bénéficient de financements de recherche du fonds national (407540\_167223), des projets de recherche Langage et Communication de l'université de Genève (LC2016-MicMac), de la Fondation Leenaards (SantéPerSo) et Swiss Personalized Health Network (C3-Study, DeID, L4CHLAB).

- 1 Laplace PS. Essai philosophique sur les probabilités [Internet]. Cambridge: Cambridge University Press; 2009 [cited 2016 Jan 8] ; accessible à : <http://dx.doi.org/10.1017/CBO9780511693182>
- 2 Groupe Médecine&Hygiène. Consentement centré sur le citoyen : partagé, transparent et dynamique [Internet]. Accessible à : [www.santeperso.ch/Projets/Consentement-centre-sur-le-citoyen-partage-transparent-et-dynamique](http://www.santeperso.ch/Projets/Consentement-centre-sur-le-citoyen-partage-transparent-et-dynamique)
- 3 Gross M. The use of finite automata in the lexical representation of natural language [Internet]. In: Gross M, Perrin D, editors. Electronic dictionaries and automata in computational linguistics. Berlin, Heidelberg : Springer Berlin Heidelberg; 1989 [cited 2018 Jul 23]. p. 34–50. Available from: [http://link.springer.com/10.1007/3-540-51465-1\\_3](http://link.springer.com/10.1007/3-540-51465-1_3).
- 4 Cori M. Des méthodes de traitement automatique aux linguistiques fondées sur les corpus. *Langages* 2008;171:95-110.
- 5 Qu'est-ce que le TAL ? [Internet]. Inbenta. 2014 [cited 2018 Jun 24] ; accessible à : [www.inbenta.com/fr/blog/quest-ce-que-le-tal/](http://www.inbenta.com/fr/blog/quest-ce-que-le-tal/)
- 6 Friedman C, Rindflesch TC, Corn M. Natural language processing: state of the art and prospects for significant progress, a workshop sponsored by the National Library of Medicine. *J Biomed Inform* 2013;46:765-73.
- 7 \* Deo RC. Machine Learning in Medicine. *Circulation* 2015;132:1920-30.
- 8 Rajpurkar P, Irvin J, Zhu K, et al. CheXNet : radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv :171105225 [cs, stat]* [Internet] 2017 [cited 2018 Jun 22] ; accessible à : <http://arxiv.org/abs/1711.05225>
- 9 \* Hutson M. Artificial intelligence faces reproducibility crisis. *Science* 2018;359:725-6.
- 10 \* SPHN - Home [Internet]. [cited 2018 Jun 24]; accessible à : [www.sphn.ch/en.html](http://www.sphn.ch/en.html)
- 11 sfa-phrt | What is PHRT? [Internet]. sfa-phrt. [consulté le 24 juin 2018] ; accessible à : [www.sfa-phrt.ch/what-is-phrt](http://www.sfa-phrt.ch/what-is-phrt)
- 12 Bierer BE, Li R, Barnes M, Sim I. A global, neutral platform for sharing trial data. *N Engl J Med* 2016;374:2411-3.
- 13 El Emam K, Dankar FK, Issa R, et al. A globally optimal k-anonymity method for the De-identification of health data. *J Am Med Inform Assoc* 2009;16:670-82.
- 14 Douglass M, Clifford GD, Reisner A, Moody GB, Mark RG. Computer-assisted de-identification of free text in the MIMIC II database. *Comput Cardiol* 2004;341-4.
- 15 Jacimovic J, Krstev C, Jelovac D. A rule-based system for automatic de-identification of medical narrative texts. *Informatica* 2015;39:43.
- 16 Foufi V, Gaudet-Blavignac C, Chevrier R, Lovis C. De-Identification of Medical Narrative Data. *Stud Health Technol Inform* 2017;244:23-7.
- 17 Stubbs A, Uzuner Ö. Annotating longitudinal clinical narratives for de-identification: The 2014 i2b2/UTHealth corpus. *J Biomed Inform* 2015;58:S20-9.
- 18 Ferrández O, South BR, Shen S, Friedlin FJ, Samore MH, Meystre SM. BoB, a best-of-breed automated text de-identification system for VHA clinical documents. *J Am Med Inform Assoc* 2013;20:77-83.
- 19 Deleger L, Molnar K, Savova G, et al. Large-scale evaluation of automated clinical note de-identification and its impact on information extraction. *J Am Med Inform Assoc* 2013;20:84-94.
- 20 Angles R. A comparison of current graph database models. In: *IEEE 28th International Conference on Data Engineering Workshops* 2012;171-7.
- 21 Park Y, Shankar M, Park BH, Ghosh J. Graph databases for large-scale health-care systems : a framework for efficient data management and data services. In: *IEEE 30th International Conference on Data Engineering Workshops* 2014;12-9.
- 22 Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System. :9. [consulté le 22 juin 2018] ; accessible à : <https://bitcoin.org/bitcoin.pdf>
- 23 \* Bitcoin expliqué par son inventeur [Internet]. Bitcoin.fr. 2013 [consulté le 22 juin 2018 ] ; accessible à : <https://bitcoin.fr/bitcoin-explique-par-son-inventeur/>
- 24 Satoshi Nakamoto (Français) [Internet]. Wikipédia. 2018 [consulté le 22 juin 2018 ] ; accessible à : [https://fr.wikipedia.org/w/index.php?title=Satoshi\\_Nakamoto&oldid=149279025](https://fr.wikipedia.org/w/index.php?title=Satoshi_Nakamoto&oldid=149279025)
- 25 Smart Contracts are self-executing contractual states, stored on the blockchain. [Internet]. [consulté le 24 juin 2018 ] ; accessible à : [www.smartcontract.com/](http://www.smartcontract.com/)
- 26 \* Contrat intelligent [Internet]. Wikipédia. 2018 [consulté le 24 juin 2018] ; accessible à : [https://fr.wikipedia.org/w/index.php?title=Contrat\\_intelligent&oldid=148443897](https://fr.wikipedia.org/w/index.php?title=Contrat_intelligent&oldid=148443897)
- 27 illustradata. Que sont les bases NoSQL orientées graphes et à quoi servent-elles ? [Internet]. Illustradata. 2016 [cited 2018 Jul 23]; Available from : [www.illustradata.com/bases-nosql-orientees-graphes-a-quoi-servent/](http://www.illustradata.com/bases-nosql-orientees-graphes-a-quoi-servent/)

\* à lire  
\*\* à lire absolument