

Communiqué de presse Académie des technologies
A Paris, le 22 octobre 2020

L'ADN, une solution pour l'archivage des mégadonnées à l'horizon 2040 ?

L'Académie des technologies établit un état des lieux des progrès et enjeux du stockage de données dans l'ADN et recommande un soutien aux initiatives françaises dans le domaine

L'Académie des technologies dévoile son rapport sur les technologies de stockage moléculaire et en particulier le stockage dans l'ADN. Elle dresse un état des lieux de l'avancée de ces technologies, des barrières actuelles, et des efforts de recherche et développement académiques et industriels engagés en France et à l'étranger. Elle recommande un soutien à cette technologie prometteuse dans l'optique d'une montée en puissance progressive entre 2025 et 2040.

La sphère globale des données (SGD) créées par l'humanité était estimée en 2018 à trente-trois mille milliards de milliards de caractères (« octets »), du même ordre que le nombre estimé de grains de sable sur terre. Ces données proviennent de la recherche et de l'industrie, mais aussi de nos connexions amicales et professionnelles, livres, vidéos et photos, informations médicales ; sans oublier dans le futur voitures autonomes, capteurs, télésurveillance, réalité virtuelle, diagnostic et chirurgie à distance. La SGD augmente d'un facteur environ mille tous les vingt ans. Une part majoritaire de ces données est ensuite stockée dans plusieurs millions de centres de données, en utilisant des technologies de stockage qui sont sujettes à l'obsolescence et posent aussi des problèmes croissants d'approvisionnement en ressources.

Une alternative prometteuse est offerte par les supports moléculaires porteurs d'information, tel que l'ADN utilisé ici comme agent chimique en dehors du vivant. Potentiellement, l'ADN permet des densités informationnelles dix millions de fois supérieures aux mémoires traditionnelles : toute la SGD actuelle tiendrait ainsi dans une fourgonnette. L'ADN est stable à température ordinaire durant plusieurs millénaires, sans consommation énergétique. Il peut être aisément multiplié ou détruit à volonté. Enfin, sa technologie ne deviendra pas obsolète car il constitue notre matériel héréditaire.

Pour archiver et retrouver des données dans l'ADN, il convient d'enchaîner cinq étapes : coder le fichier de données binaires dans l'alphabet de l'ADN qui possède quatre lettres, puis écrire, stocker, lire l'ADN et, enfin, décoder l'information lue. Un prototype réalisant ces opérations fonctionne depuis mars 2019 chez Microsoft aux États-Unis. Actuellement, plusieurs ordres de grandeur manquent pour atteindre la viabilité économique de cette approche : un facteur environ mille pour le coût et la vitesse de lecture, et cent millions pour ceux d'écriture. Ces facteurs peuvent sembler colossaux, mais sont à mettre en regard des progrès des technologies ADN, proches d'un facteur mille tous les cinq ans, donc bien plus rapides que dans les domaines électronique et informatique.



Le rapport de l'Académie aboutit à la conclusion que la technologie moléculaire d'archivage des mégadonnées a le potentiel de devenir économiquement viable entre 2025 et 2040, progressant de marchés de niche vers des marchés plus globaux. Dans le futur proche, son usage se cantonnera à l'archivage de données nécessitant d'être conservées longtemps, où ses avantages sont évidents, en compétition ou complémentarité avec l'actuelle solution, la bande magnétique.

Plusieurs laboratoires académiques, des jeunes pousses et quelques grandes entreprises se sont explicitement positionnés sur ce défi, aux États-Unis, au Royaume-Uni, en Allemagne, Irlande, Suisse, et probablement Chine. En France, au moins un laboratoire et deux petites entreprises, dont l'une vient d'être dotée par une agence des États-Unis, y ont des positions originales et fortes dans les segments-clé que sont les polymères non-ADN, la synthèse enzymatique de l'ADN et son stockage de très longue durée. Au-delà, il existe en France un gisement de compétences pertinentes en biologie, chimie, informatique et sciences de l'ingénieur, qui pourraient être mobilisées dans une nécessaire synergie entre secteurs public et privé.

L'archivage moléculaire de données constitue un enjeu majeur et stratégique à horizon proche. Le rapport recommande de lancer une action concertée au plan national et de proposer une programmation européenne sur cette thématique.

Accéder à l'avis de l'Académie des technologies <https://www.academie-technologies.fr/blog/categories/publications-de-l-academie/posts/archiver-les-megadonnees-au-dela-de-2040-la-piste-de-l-adn>

Le thème du stockage fait l'objet d'un colloque « ADN, polymères et big data » organisé par le CNRS et les Académies des technologies et de médecine le 26 octobre

<https://www.academie-technologies.fr/blog/categories/actualites/posts/colloque-adn-polymeres-et-big-data-save-the-date>

Contact Académie des technologies

Catherine Côme catherine.come@academie-technologies.fr

Contacts presse

Jean-François Kitten jf@licencek.com +33 (0)6 11 29 30 28

Stéphane Laurain s.laurain@licencek.com +33 (0)6 98 58 38 3

Valentine Fricoteaux v.fricoteaux@licencek.com +33 (0)1 88 33 41 07