

ATELIER 2020-2021 de la PLATEFORME ETHIQUE ET BIOSCIENCES

USAGES ETHIQUES DES BIG DATA EN BIOSCIENCES

La Plateforme « Ethique et Biosciences » de la Génopole® Toulouse, Occitanie, organise chaque année, depuis 2006, un atelier thématique de réflexion éthique, destiné en priorité à la communauté scientifique et ouvert à toute personne intéressée. Il s'agit de donner la parole aux chercheurs, à l'ensemble des professionnels, et au grand public, afin de leur offrir l'opportunité d'interagir avec des spécialistes ayant une expérience dans le domaine. Ces ateliers se déroulent en 3 volets d'une après-midi chacun et analysent les différentes facettes d'un thème choisi. Ils font une large place à la discussion et sont animés par un membre de la plateforme. Chaque atelier forme un tout cohérent ; cependant la participation à l'un de ces volets n'est pas conditionnée par une présence à chacun d'entre eux.

PRESENTATION DE L'ATELIER

Les Big Data - ou données massives - font parler d'elles depuis plusieurs années, compte tenu des nouvelles capacités d'explorations qu'elles permettent et des perspectives d'usages qu'elles ouvrent pour l'ensemble des secteurs d'activités : la santé (ex. CCNE Avis 130: <https://www.ccne-ethique.fr/fr/publications/donnees-massives-et-sante-etat-des-lieux-prospective-et-nouvelles-questions-ethiques>), la sécurité, le marketing, etc.

La science, qui produit et utilise de plus en plus de données, se retrouve au premier rang des acteurs du Big Data. Cette production exponentielle est due aux capacités technologiques contemporaines (informatiques, techniques) permettant de générer des données massives et d'en assurer le stockage. Quant aux usages, ils sont poussés par la « Science de la donnée » postulant que les questions de recherche peuvent être directement formulées à partir de données existantes ; renversant le paradigme prévalant jusqu'à lors de l'acquisition des données *a posteriori* de la formulation des hypothèses de recherche. Un défi majeur persiste pour assurer le traitement de ces données massives en toute fiabilité et sécurité, et de nouveaux outils devront être élaborés pour permettre leur usage et leur traitement de manière responsable et transparente (notamment grâce à l'intelligence artificielle). Ces enjeux techniques se doublent d'une politique de plus en plus fréquente de partage de ces

PLATEFORME ETHIQUE ET BIOSCIENCES

Génopole® Toulouse Occitanie
Unité mixte Inserm - Université Toulouse III Paul Sabatier
UMR 1295 - CERPOP
Faculté de Médecine, 37 allées Jules Guesde
31000 Toulouse - Tél. + 33 (0) 5 61 14 59 38
<http://societal.genotoul.fr/>

Contacts :

Emmanuelle RIAL-SEBBAG (Responsable scientifique)
emmanuelle.rial@univ-tlse3.fr
Gauthier CHASSANG (Responsable opérationnel)
gauthier.chassang@gmail.com
Lucie SERRES (Coordinatrice)
serres.lucie@gmail.com
Anne CAMBON-THOMSEN (Fondatrice de la Plateforme)
anne.cambon-thomsen@univ-tlse3.fr

données, de leur mise à disposition et donc de leur ouverture à l'ensemble de la communauté scientifique et de leur réutilisation. Ainsi, du fait notamment du volume et de la variété des données en présence, plusieurs questions éthiques et pratiques se posent pour collecter, stocker, utiliser ces données dans le respect des droits de l'homme, dans l'intérêt général et pour le bien commun.

Les ateliers de la Plateforme visent à aborder les questions éthiques, juridiques et pratiques des Big Data en tant que ressources indispensables au développement des biosciences¹ modernes, depuis la définition du concept des Big Data à la réalité de la recherche de terrain, au travers d'exemples de projets utilisant ces ressources et des débats qui suivront. Qu'appelle-t-on Big Data ? Quelles sont les opportunités d'usages et les enjeux éthiques liés aux Big Data ? Comment assurer un accès et une utilisation éthique des Big Data dans le domaine des biosciences ? Quels défis restent-ils à relever pour que les Big Data soient une ressource accessible et fiable ?

Ces réflexions s'articulent autour de **3 volets** :

Volet 1 : « Définition des Big Data : mise en commun et partage »

Judi **12 mars 2020** de 13h30 à 17h à l'Université Fédérale Toulouse Midi-Pyrénées, 41 allées Jules Guesde, Toulouse - Salle du Conseil et Salle de réception.

Volet 2 : « Big Data et innovations thérapeutiques »

Judi **4 mars 2021** de 13h30 à 17h à Toulouse – Salle à déterminer.

Volet 3 : « Enjeux socio-économiques des Big Data »

Judi **20 mai 2021** de 13h30 à 17h à Toulouse – Salle à déterminer.

[Le format sera adapté selon le contexte sanitaire (en présentiel ou en visioconférence)]

Inscription gratuite mais recommandée par envoi d'une [fiche d'inscription](#) à Lucie Serres (serres.lucie@gmail.com).

¹ Biosciences, ou sciences de la vie, c'est à dire l'ensemble des études et disciplines scientifiques qui s'intéressent à la vie, qu'elle soit animale, végétale, humaine (ex: la zoologie, la botanique, la génétique, l'agronomie, la paléontologie, la recherche médicale, l'écologie...) ou qui utilisent des êtres vivants dans différentes applications (ex : biotechnologies).

Volet 1 : « Définition des Big Data : mise en commun et partage »

Animateur : **Gauthier Chassang**, Juriste, UMR 1027 INSERM - Université Paul Sabatier, Co-Responsable de la Plateforme Ethique et Biosciences (Genotoul Societal).

Intervenants :

- **Nicolas Savy**, Maître de conférences – Institut de Mathématiques de Toulouse.
- **Jérôme Béranger**, Fondateur de la société ADEL (*Algorithm Data Ethics Label*), Chercheur (PhD) associé à l'UMR 1027 Inserm - Université Paul Sabatier, Auteur de plusieurs ouvrages dont : « Le code éthique algorithmique » (2018, ISTE Editions).

Date : le jeudi 12 mars 2020 de 13h30 à 17h.

Lieu : Université Fédérale Toulouse Midi-Pyrénées, 41 allées Jules Guesde, Toulouse - Salle du Conseil et Salle de réception.

PRESENTATION DU VOLET

L'exploitation des données pour la recherche, l'innovation et le développement technologique est aujourd'hui au cœur des stratégies étatiques². Il n'est pas rare d'entendre ici et là que de nombreux secteurs économiques se fondent désormais sur des analyses complexes de données massives dites « Big data » pour faire évoluer les connaissances mais aussi les services, au bénéfice de tous. Mais le concept de Big data reste assez hermétique pour les citoyens non-initiés en ce qu'il englobe une variété de sources et de données, qu'il s'agisse de données géographiques, climatiques, statistiques, liées au transport, à la santé, aux finances et d'usages. Qu'appelle-t-on « Big data » ? Comment sont-elles constituées ? Quels apports attend-t-on de ces nouvelles ressources pour les domaines des biosciences ? Quels enjeux éthiques suscitent les Big data pour les biosciences, mais aussi pour les décideurs politiques et pour les citoyens ?

Afin de clarifier les concepts de référence, les pratiques actuelles à l'œuvre derrière ces concepts et les défis, présents et à venir des Big data en biosciences, nous aborderons ces différents aspects au

² Communication de la Commission Européenne : Vers une économie de la donnée prospère ([COM/2014/0442 final](#)), et document de travail accompagnant la communication ([SWD\(2014\)214 final](#))

PLATEFORME ETHIQUE ET BIOSCIENCES

Génopole® Toulouse Occitanie
Unité mixte Inserm - Université Toulouse III Paul Sabatier
UMR 1295 - CERPOP
Faculté de Médecine, 37 allées Jules Guesde
31000 Toulouse - Tél. + 33 (0) 5 61 14 59 38
<http://societal.genotoul.fr/>

Contacts :

Emmanuelle RIAL-SEBBAG (Responsable scientifique)
emmanuelle.rial@univ-tlse3.fr
Gauthier CHASSANG (Responsable opérationnel)
gauthier.chassang@gmail.com
Lucie SERRES (Coordinatrice)
serres.lucie@gmail.com
Anne CAMBON-THOMSEN (Fondatrice de la Plateforme)
anne.cambon-thomsen@univ-tlse3.fr

travers de deux présentations d'experts du domaine, d'exemples concrets, et durant les débats qui suivront avec les participants.

Résumé de Nicolas Savy – « Big Data en Santé : la place de la donnée dans la recherche médicale de demain »

Au cours de cet exposé, nous discuterons de plusieurs thématiques au cœur desquelles se pose la question suivante : comment en sommes-nous arrivés là ?

Ce sera l'occasion de parcourir les grandes étapes de l'évolution du rôle de la donnée dans la recherche - notamment médicale - et des conséquences sur la conception de la preuve scientifique et ce jusqu'au modèle disruptif de « Big Data ». Nous discuterons les spécificités du « Big Data » et les conséquences de ses spécificités sur le traitement de ces données, avec la montée en puissance du *machine learning* et de l'Intelligence Artificielle (IA). Quelques généralités sur le *machine learning* nous permettront de mettre en avant les forces, les faiblesses et les écueils méthodologiques de ces approches en biosciences.

Après cet état des lieux du matériel et des techniques disponibles, nous serons en mesure de présenter quelques exemples d'innovations ou de pistes d'innovation afin d'illustrer le rôle de la donnée dans la recherche médicale de demain.

Résumé de Jérôme Béranger – « L'éthique des Big Data en santé »

Face au développement constant de l'e-santé, de la télémédecine et des Big data médicaux, des changements juridiques qui s'accroissent, des ruptures économiques qui s'amplifient, une sensation de flou survient autour de la donnée personnelle en santé. La manipulation et l'exploitation de l'information médicale se trouvent alors en perpétuelle équilibre entre confidentialité et transparence. Ces Nouvelles technologies de l'information et de la communication (NTIC) perturbent sensiblement le microcosme médical allant jusqu'à remettre en question les paradigmes de la Médecine d'Hippocrate telle que nous la connaissons auparavant. Les innovations à la clé sont considérables, susceptibles à la fois de réduire les coûts de santé et d'ouvrir la voie vers la médecine personnalisée. Il est incontestable que la gestion de ces données massives est un important levier pour une meilleure compréhension des maladies, du développement de médicaments et du traitement des

PLATEFORME ETHIQUE ET BIOSCIENCES

Génopole® Toulouse Occitanie
Unité mixte Inserm - Université Toulouse III Paul Sabatier
UMR 1295 - CERPOP
Faculté de Médecine, 37 allées Jules Guesde
31000 Toulouse - Tél. + 33 (0) 5 61 14 59 38
<http://societal.genotoul.fr/>

Contacts :

Emmanuelle RIAL-SEBBAG (Responsable scientifique)
emmanuelle.rial@univ-tlse3.fr
Gauthier CHASSANG (Responsable opérationnel)
gauthier.chassang@gmail.com
Lucie SERRES (Coordinatrice)
serres.lucie@gmail.com
Anne CAMBON-THOMSEN (Fondatrice de la Plateforme)
anne.cambon-thomsen@univ-tlse3.fr

patients. Pour autant, les avantages attendus pour le pilotage de la santé et la mise en œuvre d'une médecine de plus en plus précise et ciblée n'en soulèvent pas moins un certain nombre de questions. Par exemple : Comment déterminer qui peut ou doit avoir accès à ces données et dans quelles conditions ? Comment conjuguer intérêt collectif et protection des personnes ? Comment se prémunir contre toute utilisation commerciale, voire frauduleuse de ces données ? Quelles instances doivent les gérer, en autoriser l'accès, assurer une efficacité et une sécurité optimales tout en garantissant le caractère démocratique et transparent de cette gestion ?

Autant de questions qui méritent d'être examinées et débattues pour que ces enjeux déterminants pour l'avenir de la santé publique fassent l'objet de choix éclairés. Dès lors, une réflexion sur l'étude des risques associée aux enjeux éthiques autour de la donnée de santé à caractère personnel s'impose à nous. L'enjeu est de donner du sens à la réalisation, la mise en place et l'utilisation des Big Data afin de mieux les contrôler. Nous proposons d'aligner les concepts d'une ontologie des risques à ceux d'une ontologie des exigences d'objectifs éthiques sur la base des meilleures pratiques dans le domaine de la sécurité des données. En plus du mécanisme de guidage, nous détaillerons les démarches utilisées pour la construction et l'alignement des deux ontologies.

PLATEFORME ETHIQUE ET BIOSCIENCES

Génopole® Toulouse Occitanie
Unité mixte Inserm - Université Toulouse III Paul Sabatier
UMR 1295 - CERPOP
Faculté de Médecine, 37 allées Jules Guesde
31000 Toulouse - Tél. + 33 (0) 5 61 14 59 38
<http://societal.genotoul.fr/>

Contacts :

Emmanuelle RIAL-SEBBAG (Responsable scientifique)
emmanuelle.rial@univ-tlse3.fr
Gauthier CHASSANG (Responsable opérationnel)
gauthier.chassang@gmail.com
Lucie SERRES (Coordinatrice)
serres.lucie@gmail.com
Anne CAMBON-THOMSEN (Fondatrice de la Plateforme)
anne.cambon-thomsen@univ-tlse3.fr

PROGRAMME

- 13h30 *Accueil des participants et remise des documents*
- 13h40 **Présentation du volet 1 de l'atelier par Gauthier Chassang**
- 13h50 Tour de table
- 14h00 **Nicolas Savy**, Maître de conférences – Institut de Mathématiques de Toulouse
« Big Data en santé : la place de la donnée dans la recherche médicale de demain »
- 14h30 Discussions
- 14h45 **Jérôme Béranger**, Fondateur de la société ADEL (Algorithm Data Ethics Label), Chercheur (PhD) associé à l'UMR 1027 Inserm - Université Paul Sabatier, Auteur de plusieurs ouvrages dont : « Le code éthique algorithmique » (2018, ISTE Editions)
« L'éthique des Big Data en santé »
- 15h15 Discussions
- 15h30 **Identification des questions à débattre**
- 15h45 *Pause*
- 16h00 **Atelier de réflexion**
- 17h00 *Fin de l'atelier*

PLATEFORME ETHIQUE ET BIOSCIENCES

Génopole® Toulouse Occitanie
Unité mixte Inserm - Université Toulouse III Paul Sabatier
UMR 1295 - CERPOP
Faculté de Médecine, 37 allées Jules Guesde
31000 Toulouse - Tél. + 33 (0) 5 61 14 59 38
<http://societal.genotoul.fr/>

Contacts :

Emmanuelle RIAL-SEBBAG (Responsable scientifique)
emmanuelle.rial@univ-tlse3.fr
Gauthier CHASSANG (Responsable opérationnel)
gauthier.chassang@gmail.com
Lucie SERRES (Coordinatrice)
serres.lucie@gmail.com
Anne CAMBON-THOMSEN (Fondatrice de la Plateforme)
anne.cambon-thomsen@univ-tlse3.fr