HAL open science

Les trois cultures de la psychiatrie computationnelle

Christophe Gauld, Guillaume Dumas, Éric Fakra, Jérémie Mattout, Jean-Arthur Micoulaud-Franchi

▶ **To cite this version:**

Christophe Gauld, Guillaume Dumas, Éric Fakra, Jérémie Mattout, Jean-Arthur Micoulaud-Franchi. Les trois cultures de la psychiatrie computationnelle. Annales Médico-Psychologiques, Revue Psychiatrique, 2021, 179 (1), pp.63-71. 10.1016/j.amp.2020.11.011 . hal-03956938

HAL Id: hal-03956938

https://hal.science/hal-03956938v1

Submitted on 22 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S0003448720303371
Manuscript_aa6ad451f02b08fb695b0be3f7eeb39d

Les trois cultures de la psychiatrie computationnelle

The three cultures of computational psychiatry

Christophe Gauld[1,2*], Guillaume Dumas [3,4], Éric Fakra[5,6], Jérémie Mattout[7,8], Jean-Arthur Micoulaud-Franchi[9,10]

[1] Institut d'Histoire et de Philosophie des Sciences et des Techniques, 13 rue du Four, 75006 Paris, France

[2] Centre Hospitalier Universitaire Grenoble-Alpes, Département de psychiatrie, Avenue du maquis de Grésivaudan, 38 000 Grenoble, France.

[3] Precision Psychiatry and Social Physiology Laboratory, CHU Sainte-Justine Research Center, Department of Psychiatry, University of Montreal, Quebec, Canada

[4] Mila – Quebec Artificial Intelligence Institute, University of Montreal, Quebec, Canada

[5] Department of Psychiatry, University Hospital of Saint-Étienne, Saint-Étienne, France.

[6] INSERM, U1028; CNRS, UMR5292; Lyon Neuroscience Research Center, Psychiatric Disorders: from Resistance to Response, PSYR2 Team, Lyon, France.

[7] University Lyon 1, Lyon, F-69000, France.

[8] Lyon Neuroscience Research Center, CRNL; INSERM, U1028; CNRS, UMR5292; Brain Dynamics and Cognition Team, Lyon, F-69000, France.

[9] Service Universitaire de Médecine du Sommeil (SUMS), CHU de Bordeaux, Place Amélie Raba-Léon, 33 076 Bordeaux, France.

[10] USR CNRS 3413 SANPSY, CHU Pellegrin, Université de Bordeaux, France.

* Auteur correspondant :

Christophe Gauld, Centre Hospitalier Universitaire Grenoble-Alpes, Département de psychiatrie, Avenue du maquis de Grésivaudan, 38 000 Grenoble, France.

Courriel : gauldchristophe@gmail.com

Résumé

Le champ de la psychiatrie computationnelle prend de l'ampleur depuis quelques années. Dans cet article, nous proposons de distinguer trois champs de la psychiatrie computationnelle. Le premier champ correspond à la *Digital Psychiatry*, ou psychiatrie utilisant les outils numériques, qui peut être définie comme un champ de la psychiatrie qui utilise des outils connectés pour recueillir des données numériques. Le deuxième champ correspond à la *Big Psychiatry*, ou psychiatrie fondée sur les données massives, qui traite de grandes quantités de données, en particulier par des méthodes

© 2020 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

d'apprentissage automatique, une des branches de l'intelligence artificielle. Le troisième champ correspond à la *Psychiatry Modeling*, ou modélisation en psychiatrie utilisant notamment les neurosciences computationnelles, avec le développement et l'utilisation de modèles formels (mathématiques) du fonctionnement (et dysfonctionnement) cérébral et cognitif, afin de caractériser les mécanismes à l'origine des symptômes observés en clinique. Ces trois champs se complètent et sont fortement dépendants les uns des autres. Ainsi, l'accès aux données est fourni par la *Digital Psychiatry* ; le traitement des données est opéré par la *Big Psychiatry* ; et la formalisation des hypothèses est offerte par la *Psychiatry Modeling*. Cette triple organisation de la psychiatrie computationnelle offre un cadre de réflexion robuste pour appréhender la psychiatrie personnalisée et de précision, articulée autour de méthodologies statistiques et mathématiques, axée sur la prédiction et l'explication et utilisant des données qualitativement et quantitativement variées – tout en s'adressant nécessairement à un sujet commun, le patient de la clinique psychiatrique.

Mots-clés : Apprentissage automatique ; Données massives ; Intelligence artificielle ; Neurosciences ; Outils numériques ; Objets connectés ; Physiopathologie ; Psychiatrie computationnelle ;

Abstract

Introduction. – Whether on the social, economic or scientific level, the digital sciences tend to change the conception of health. Computational Psychiatry, in the sense of a psychiatry based on "numbers" and information flow, has evolved rapidly.

Methods. – In this article, we propose the distinction between three fields of Computational Psychiatry. A first field corresponds to "Digital Psychiatry", i.e. a field using digital, connected, tools in the main goal to collect digital data. A second field corresponds to "Big Psychiatry", or Big Data Psychiatry, which deals with large amounts of data, e.g. through recent methodologies in machine learning or artificial intelligence. A third field corresponds to "Psychiatry Modeling", which corresponds to the utilization of formal hypothesis (i.e. mathematical models) about brain and behavior (and their dysfunctions) in line with computational neurosciences.

Results. – The collection of digital data fits into methodologies of assessments and interventions in daily life, named Ecological Momentary Assessment. Of course, these digital data, which differ quantitatively and qualitatively from what psychiatry has been able to collect in its history, raise numerous epistemological and ethical questions. In the field of Big Psychiatry, most machine learning techniques provide predictions rather than pathophysiological mechanisms, and these machine learning techniques makes it possible to propose new delineations of disorders in a logic of stratified medicine. Lastly, resulting from studies in computational neurosciences, explanatory

modeling of the brain (often called “Generative Modeling”) proposes a number of theories to understand the functioning of the brain in psychiatric disorder (e.g. predictive coding, reinforcement learning, decision making theories, but also dynamical systems theories and graph and network theory).

Discussion and Conclusion. – This field could offer a framework to characterize the origin of the psychiatric symptoms. Obviously, these three fields are highly mutually dependent, with for instance a data access provided by Digital Psychiatry (with Digital Tools), a data processing operated by Big Psychiatry (with Machine Learning) and a formalization of hypotheses offered by Generative Modeling of the brain from Psychiatry Modeling. This triple organization of Computational Psychiatry offers a robust framework for personalized and precision psychiatry, articulated around statistical and mathematical methodologies, focused on prediction and explanation and using qualitatively and quantitatively varied data. However, such a framework is necessarily geared to a common subject: the patient of the psychiatric clinic.

Key words: artificial intelligence; big data; computational psychiatry; digital tools; machine learning; neuroscience; pathophysiology

« J'ai dit que la réprobation attachée autrefois au péché de sorcellerie s'attache maintenant dans les esprits humains aux spéculations de la cybernétique moderne. »

« Ainsi l'un des grands problèmes futurs que nous devons résoudre est celui de la relation entre l'homme et la machine, des fonctions qui devraient être correctement attribuées à ces deux organes »

Norbert Wiener, 1964

God and Golem, On Cybernetics

1. Introduction

La psychiatrie computationnelle est en pleine expansion et soulève des enjeux pour cette spécialité médicale. Ces enjeux mêlent philosophie et éthique, pratique clinique, mais aussi

technologies numériques,[1] nouvelles méthodes d'analyses de données (ex. de l'intelligence artificielle), et modélisation de la cognition et du comportement en lien avec les phénomènes neurophysiologiques sous-jacents [24]. Il s'agit donc de connaître les technologies utilisées, de comprendre pourquoi et comment elles s'adaptent à des contextes cliniques donnés, en sachant interpréter les résultats et expliquer les principes de base de ces nouvelles méthodes aux patients et aux autres professionnels de santé [22]. En l'absence d'un tel socle de connaissances, le risque est grand de voir les psychiatres appliquer des techniques en mutation et en constante évolution sans en saisir le sens.

La psychiatrie fondée sur des données numériques et leurs analyses, qui ont inspiré la psychiatrie computationnelle, peut être retrouvée dès son origine. Par exemple, Esquirol est connu pour ses tables de statistiques. Pinel et la génération de cliniciens du début du XIXe siècle (Esquirol, Leuret et Lisle notamment) ont su aussi s'intéresser très tôt aux sciences adjacentes à la psychiatrie, permettant de proposer une clinique prudente, raisonnable et fondée sur la reconnaissance de régularités chiffrées. Les psychiatres de cette époque ont été appelés « numéristes » [27], inscrits dans une véritable « révolution probabiliste » [28] et ils étaient organisés autour du « héros méconnu des cliniciens », comme Pierre-Charles Alexandre Louis qui déploya la méthode numérique en médecine, et est considéré à ce titre comme le père spirituel de la médecine fondée sur les données probantes (*Evidence Based Medicine*, EBM) [4], à l'instar du père de la médecine expérimentale, Claude Bernard, qui a fondé les bases de l'explication physiologique moderne [7]. Mais depuis quelques années, la psychiatrie computationnelle s'oriente vers de nouvelles manières de recueillir des données chiffrées, de les analyser et de les utiliser.

Que ce soit sur le plan social, économique ou scientifique, les sciences du numérique ont tendance à changer en effet la conception de la santé. La psychiatrie n'échappe pas à cette influence. Alors que l'épidémiologie et les grandes cohortes de patients ont permis de fournir de nouvelles données pour comprendre les troubles mentaux, et que les découvertes de la physiologie et de l'imagerie morphologique et fonctionnelle de la seconde moitié du XXe siècle ont permis d'appréhender les troubles mentaux dans un cadre de réflexion centré autour des neurosciences, au XXIe siècle la psychiatrie s'organise désormais aussi autour d'un troisième cadre de réflexion, celui des sciences computationnelles. L'approche computationnelle renvoie à la fois à l'acquisition et au traitement des données numériques, faisant appel aux technologies numériques, aux systèmes connectés et à l'informatique. La psychiatrie computationnelle pourrait ainsi constituer un champ disciplinaire traitant des données fournies sous forme de nombres et qui peuvent être maniées selon des règles de calcul (ou algorithmes).

[1] En français, le terme de « numérique » est la traduction du terme anglais de « *digital* ».

L'intérêt croissant pour la psychiatrie computationnelle a été mis en avant dans une publication récente du *Lancet* avançant que les grands défis de la psychiatrie à venir (à l'image des problèmes de Hilbert décrits en mathématiques) dépendraient largement des sciences mathématiques (modélisation) et des neurosciences computationnelles [32, 33]. Une brève revue des publications issues de la littérature portant sur la psychiatrie computationnelle permet de montrer l'intérêt pour cette approche, augmentant de manière exponentielle (**Figure 1).**

[INSÉRER FIGURE 1]

Nous proposons ainsi dans cet article un cadre d'appréhension qui distingue trois « champs » de la psychiatrie computationnelle. Le premier correspond à la psychiatrie utilisant les outils de santé numérique (*Digital Psychiatry*), compris comme les interfaces connectées, telles les applications sur smartphones. Le deuxième champ appartient au champ de l'application des données massives (*Big Data*) dans le domaine de la santé mentale (*Big Psychiatry*). Celle-ci traite de grandes quantités de données, par exemple, par des méthodes d'intelligence artificielle et notamment d'apprentissage automatique. Le troisième champ correspond aux développements de modèles biologiquement plausibles, génératifs et explicatifs (*Psychiatry Modeling*). Cette démarche est née des théories et approches en neurosciences computationnelles, visant à relier les manifestations cliniques des troubles mentaux aux processus neurobiologiques, au travers d'une modélisation mathématique formelle.

[INSÉRER FIGURE 2]

2. Digital psychiatry

2.1. Recueil des données numériques

La *Digital Psychiatry* (ou psychiatrie numérique) peut être définie comme un champ de la psychiatrie qui gère des données numériques en s'appuyant sur les nouvelles technologies (« Outils Numériques » – **Figure 2**).

Le recueil de ces données numériques, au niveau individuel, s'intègre en particulier dans des méthodologies d'évaluations et d'interventions en vie quotidienne appelées *Ecological Momentary Assessment* (EMA) pour évaluation momentanée écologique [31]. Pour cela, des capteurs comportementaux (ex. activité physique ou sommeil), physiologiques (ex. fréquence cardiaque), de

voix (ex. prosodie, champ lexical) ou environnementaux (ex. pollution atmosphérique) peuvent être disposés sur une technologie mobile et connectée, comme un smartphone. Toute trace numérique laissée lors de l'usage de ces objets connectés peut aussi constituer une donnée numérique d'intérêt (appelée « empreinte » ou *footprint*).

Le développement récent et en augmentation croissante de l'Internet des objets (*Internet of Things – IoT*) permettra de passer du recueil issu d'une technologie spécifique à un recueil issu d'une technologie générique et ubiquitaire, implantée dans les objets du quotidien. En effet, de nombreux objets du quotidien pourront posséder les trois qualités qui permettent leur classification dans l'IoT : ils seront instrumentés (par une collecte et une intégration des données en temps réel), interconnectés (par un partage entre systèmes d'information) et intelligents (par un enrichissement des analyses pour les usagers et une aide à la prise de décision, permise notamment par l'analyse des données massives recueillies).

Enfin, le recueil des données numériques peut se faire au niveau de la population. Le développement des *smartcities* en est un exemple. De plus en plus de villes possèdent les qualités de l'IoT (instrumentation, interconnexion, intelligence), permettant de fournir des données sur les objets publics, les stocks, les flux ou les tendances citoyennes, dont une grande partie pourrait renseigner sur la santé mentale des citoyens. Par exemple, une étude française corrèle actuellement les données météorologiques à l'humeur des patients souffrant de troubles bipolaires [12]. La fouille de données textuelles sur les réseaux sociaux peut constituer un autre exemple d'extraction des données à l'échelle de la population générale [11].

2.2. Digital Psychiatry et RDoC

L'ensemble des données recueillies par les technologies numériques peuvent être potentiellement intégrées dans le cadre de la matrice neuroscientifique présentée en 2009 par le National Institute of Mental Health (NIMH) et désignée par le terme de « Research Domain Criteria » (RDoC) [16]. Les RDoC sont structurés autour d'une matrice dont les lignes témoignent de différentes fonctions cérébrales et les colonnes de différentes échelles du vivant. Les données classiquement intégrées à la matrice sont issues d'expérimentations neuroscientifiques. Toutefois, une littérature croissante propose d'y insérer des données numériques issues des technologies digitales à toutes les échelles du vivant [23, 39]. L'intérêt d'aborder la matrice des RDoC au travers de la *Digital Psychiatry* serait que les données numériques ainsi recueillies comportent des caractéristiques que ne possèdent pas les données classiques de la psychiatrie ou des neurosciences.

2.3. Caractéristiques spatio-temporelles des données

Les nouvelles technologies permettent en effet de réduire l'écart spatio-temporel entre les vécu et comportement du patient et le recueil des informations pendant la consultation psychiatrique. Par ailleurs, elles permettent d'apporter des données écologiques (c'est-à-dire « en situation réelle ») sur des comportements de la vie quotidienne. En effet, les données numériques sont potentiellement continues dans le temps et ubiquitaires dans l'espace. De ce fait, en fonction de la question posée, la résolution temporelle ou spatiale peut être modulée. Cette résolution permet d'obtenir un effet de « *zoom-in* » sur certaines données recueillies, comme sur les plaintes d'insomnie, qui peuvent alors être distinguées en « insomnie d'endormissement », « insomnie de maintien » ou « insomnie de réveil précoce ». L'ajustement de la résolution temporelle ou spatiale peut également permettre de fournir un effet de « *zoom-out* », notamment lorsqu'il s'agit d'intégrer de nombreux facteurs environnementaux pouvant interférer avec, par exemple, une tristesse de l'humeur [3]. Ces effets de *zoom* permettent d'adapter la « granularité » des données numériques et de servir à fournir des données très précises pour une évaluation psychiatrique. Les données permettraient ainsi d'affiner les évaluations psychiatriques de par leur objectivité et la pluralité de leur nature. Enfin, plus généralement, de telles données numériques s'avéraient plus économiques par rapport à d'autres marqueurs paracliniques classiques (biologiques, imagerie, EEG) et pourraient être plus aisément utilisées en recherche [17].

Ces données numériques qui diffèrent quantitativement et qualitativement de ce que la psychiatrie a pu recueillir dans son histoire offrent des opportunités nouvelles de recherche. Par ailleurs, au côté des données numériques pouvant provenir des outils connectés de la vie quotidienne, la *Big Psychiatry* ouvre la possibilité d'extériorisation du laboratoire, en permettant de constituer des échantillons dont la taille et la nature offrent des perspectives inédites (par exemple, applications ou réseaux sociaux avec des expériences en ligne ou des questionnaires). Cependant, ces données numériques posent de nombreuses questions pratiques et éthiques. Il s'agira notamment d'interroger à qui elles doivent appartenir, leur utilisation secondaire et leur possible détournement, leur validité et leur fiabilité ou leur utilité en pratique clinique et de recherche. Malgré ces défis, l'enjeu principal des données numériques est de constituer une nouvelle masse d'informations potentiellement intéressantes car, pour la première fois à notre connaissance dans l'histoire de la psychiatrie, elles offrent un aperçu des conditions de vie du patient.

3. *Big psychiatry*

3.1. Définition de la Big Psychiatry

La *Big Psychiatry*, ou psychiatrie fondée sur les données massives, peut être définie comme un sous-domaine de la médecine qui utilise des banques de données numériques de très grande envergure et de tous types, nommées « Big Data ». Il s'agit alors d'effectuer du « minage de données » (ou *data mining*), correspondant à analyser des banques de données avec ou sans hypothèses sur leurs causes [25]. Ces grandes quantités de données peuvent provenir des technologies numériques citées précédemment (de la *Digital Psychiatry*), mais elles peuvent également provenir de sources très diverses, notamment des études cliniques plus traditionnelles qui ont tendance à être d'ampleur toujours plus importante (données multimodales, études multicentriques, échantillons plus grands) et des Omics. Les Omics sont des ensembles de données spécifiques structurés à un niveau d'organisation du vivant depuis la génomique jusqu'à l'exposomique [7, 41]. Ces deux sources de données sont largement complémentaires, pouvant provenir d'objets connectés de la vie quotidienne ou de l'extériorisation du laboratoire (celui-ci tendant à « intégrer » la vie quotidienne).

3.2. Apprentissage automatique

La nouveauté apportée à la psychiatrie en termes de masse de données, d'une part, et le fait de toucher une population nouvelle, plus générale, d'autre part, appelle des méthodes d'analyse adaptée. Toutes ces méthodes ont été largement développées au sein d'autres disciplines médicales. Ainsi, la pertinence de grandes quantités de données dépend intrinsèquement du traitement auxquelles elles sont soumises. Actuellement, ces traitements statistiques sont principalement rassemblés sous le terme d'apprentissage automatique (*Machine Learning* ou « Apprentissage machine » – **Figure 2**), la principale méthode développée en intelligence artificielle et qui se confond généralement avec elle.

L'intelligence artificielle (IA) est une discipline scientifique qui cherche à reproduire certaines fonctions de l'intelligence humaine en basant ses théories et ses méthodes sur l'utilisation de machines. L'apprentissage automatique est une méthode d'IA fondée notamment sur des algorithmes statistiques (par exemple, des machines à vecteur de support ou des réseaux de neurones) et caractérisé de manière schématique par trois axes non hiérarchisés : la formulation d'hypothèses théoriques *a priori* (description de l'algorithme), la liberté fournie aux données (entraînement de l'algorithme à reconnaître ces données) et la capacité d'exploitation des connaissances structurées à partir de ces données (généralisation de l'utilisation de l'algorithme par l'utilisation d'un jeu de données différent). Les objectifs de l'apprentissage automatique sont de classer (ranger les données dans des classes prédéterminées, on parle alors d'apprentissage supervisé), de classifier (ranger des données dans des classes non prédéterminées, on parle alors

d'apprentissage non supervisé), de suivre des évolutions et des dynamiques (on parle de régression, souvent entre plusieurs variables). L'une des méthodes pour atteindre ces différents objectifs est par exemple d'utiliser une méthode de renforcement d'un comportement (on parle alors d'apprentissage par renforcement).

3.3. Prédiction et classifications

L'apprentissage automatique fournit des prédictions [14]. La prédiction ne nécessite pas de préciser la nature du mécanisme qui conduit à l'apparition du phénomène [2]. Pour autant, la prédiction peut être d'une grande utilité clinique. Au quotidien, la *Big Psychiatry* pourrait accompagner le clinicien pour mieux prédire le diagnostic (nouvelles bases diagnostiques), l'évolution, le pronostic et la réponse prédictive à la thérapeutique. Pour autant, si la *Big Psychiatry* peut augmenter la capacité de prédiction de la clinique et fournir par exemple des *patterns* de comportement (c'est-à-dire des ensembles d'observations) (**Figure 2**), elle n'en augmente pas pour autant sa compréhension et sa portée explicative.

On retrouve donc dans la *Big Psychiatry* le débat des années 1980, suite à la parution du DSM-III, où les catégories diagnostiques étaient définies selon un enjeu de fiabilité, mettant l'accent sur une communication efficace entre cliniciens, plutôt que sur un enjeu de validité, qui cherche à cerner le type naturel d'un trouble psychiatrique (c'est-à-dire ce qu'il est censé représenter dans la réalité). Autrement dit, la physiopathologie des catégories du DSM est encore mal définie. Étant donné ces difficultés et les apports potentiels de l'apprentissage automatique en termes de prédiction et de classification, les espoirs fournis par cette approche semblent majeurs. Pour autant, les difficultés persistent de manière importante. Le fait que la psychiatrie fondée sur les données massives s'appuie sur la prédiction plutôt que sur les mécanismes physiopathologiques permet certes de proposer de nouvelles délimitations des troubles (catégorielles ou dimensionnelles) dans une logique de médecine stratifiée [10]. Mais reste à définir le lien de ces nouvelles délimitations au regard d'une compréhension mécanistique physiologique. Pour autant, les données numériques traitées par la *Big Psychiatry* pouvant être des mesures comportementales reliées à des mécanismes physiologiques, il est donc possible que l'apprentissage automatique conduise à des catégories également pertinentes physiopathologiquement, et non pas uniquement d'un point de vue clinique. Cela soulève cependant l'enjeu du degré de contrainte dans le cadre d'un apprentissage supervisé ou non.

Les espoirs d'une telle psychiatrie fondée sur de grandes quantités de données doivent pour autant être tempérés. Notamment, la reproductibilité n'est pas systématiquement garantie entre lieux géographiques, techniques (notamment en ce qui concerne les méthodes d'apprentissage

automatique) et populations (par exemple, pour un même trouble psychiatrique avec des comorbidités différentes). Enfin, la gestion des données demande une organisation en consortium dont la gestion n'est pas évidente. On retrouve cependant de telles initiatives, qui s'appuient sur des infrastructures importantes de recueil de données (nommées *data lake* pour « lacs de données »), par exemple avec l'*Autism Brain Imaging Data Exchange* (ABIDE) [5], le *Human Connectome Project* (HCP) [40] ou le consortium ENIGMA [35]. Enfin, l'acquisition de données longitudinales reste un défi organisationnel pour la recherche en lien avec la clinique.

4. *Psychiatry modeling*

4.1. Modélisation psychiatrique et neurosciences computationnelles

La *Psychiatry Modeling* peut être définie comme un axe de recherche consistant au développement de la modélisation en psychiatrie utilisant notamment les neurosciences computationnelles. Au sein des neurosciences computationnelles, une partie des méthodes utilise les mêmes outils que la *Big Psychiatry* (notamment l'apprentissage automatique). Cependant, l'originalité des neurosciences computationnelles est de proposer également des modèles explicatifs et génératifs du fonctionnement cérébral et cognitif (*Generative Modeling /* « Modèles génératifs » – **Figure 2**), qui peuvent enrichir la compréhension des troubles psychiatriques [1]. En effet, les neurosciences computationnelles correspondent à une discipline scientifique utilisant des modèles formels (mathématiques), qui rendent explicites les fonctions (et dysfonctions) cérébrales et cognitives et permettent de caractériser ainsi l'origine des symptômes observés en clinique [8]. La recherche en psychiatrie utilisant les neurosciences computationnelles vise à redéfinir les troubles psychiatriques en termes de dysfonctions cérébrales utilisant les modèles mathématiques [1]. Pour cela, elle se fonde sur des modèles « explicites et formels » empiriquement testables, c'est-à-dire quantifiables et falsifiables, avec des algorithmes qui « génèrent, à partir d'une information externe (stimulus), un comportement observable » [1]. Les neurosciences computationnelles tentent de relever le défi du « fossé explicatif » (*explanatory gap*) qui sépare le cerveau et les comportements [19].

Les modèles computationnels utilisés possèdent plusieurs propriétés : il doit être mathématiquement explicite (l'algorithme décrit des processus physiopathologiques) [20], génératif (l'algorithme imite les opérations cérébrales et cognitives et décrit comment les données observées, comme l'activité cérébrale ou les symptômes cliniques, ont été générées par différents mécanismes), biologiquement plausible (l'algorithme doit pouvoir être réalisé par le cerveau) et il doit comporter des paramètres libres (des variables ajustables selon les comportements individuels).

Pessiglione et al. [26] ont ainsi proposé la notion de « phénotype computationnel », qui correspond à l'ensemble des paramètres computationnels d'un individu donné.

4.2. Les théories en neurosciences computationnelles

Nous pouvons citer cinq principales théories retrouvées dans le champ des neurosciences computationnelles :

- *Le codage prédictif (predictive coding) répond au principe de l'énergie libre (free energy principle).* Dans ces modèles, les calculs neuronaux seraient conformes aux principes bayésiens : les représentations sont mises à jour dans le but de produire des croyances ou des attentes en s'appuyant sur la combinaison d'attentes antérieures et de données sensorielles, selon la loi de Bayes. Lorsque le comportement corrige l'erreur de prédiction en agissant sur l'environnement, on parle d'inférence active [9, 18].
- *L'apprentissage par renforcement*. Dans ce cadre, les attentes antérieures représentent la récompense. Une erreur de prédiction accompagnée d'une récompense négative va signaler à l'agent une perte de confiance (ou de précision) dans ses attentes [6].
- *Les théories de la décision optimale (decision-making)*, avec par exemple la théorie des jeux et les modèles de prise de décision. L'entraînement modifie les distributions de probabilité qu'un agent va avoir dans son environnement futur : il pourra ainsi s'attendre à avoir peu d'influence dans ses décisions, c'est-à-dire peu de contrôlabilité, ou l'inverse. Ainsi, une croyance antérieure impliquant que l'environnement est peu susceptible d'offrir une contrôlabilité substantielle découragera l'exploration. De ce fait, cette approche explique que les comportements dépendent de la contrôlabilité passée. Cette interprétation computationnelle de « l'incontrôlabilité » offre une nouvelle manière de comprendre les interactions entre un agent et son environnement. Des altérations dans ce modèle peuvent expliquer de nombreux troubles psychiatriques, ainsi que des symptômes tels que l'apathie, l'impulsivité ou les compulsions.
- *Les théories des systèmes dynamiques*. La caractérisation de la dynamique cérébrale à l'aide de concepts issus de la théorie des systèmes dynamiques, comme la criticité, les lois du chaos, les bifurcations et la notion d'émergence permettent de fournir de nombreuses hypothèses pouvant être utilisées pour expliquer les troubles psychiatriques. Les modèles les plus élémentaires utilisés sont des modèles

biophysiques, qui ont plus ou moins de plausibilité biologique, mais qui permettent le maniement d'ensemble de données (à la différence de modèles mécanistes qui visent plutôt à saisir les causes ou les processus), ce qui conduit à d'excellentes prédictions [15, 34].

- *La théorie des graphes et des réseaux*. La science des réseaux et son application à la psychiatrie sont basées sur une branche des mathématiques appelée « théorie des graphes » [36] et permettent des applications dans le cadre d'une « médecine des réseaux ». La théorie des graphes permet d'analyser les connexions et les voies spécifiques entre variables d'un réseau (ou graphe) et les propriétés globales de ce réseau. Les connexions (au niveau « micro ») fournissent des informations sur les voies mécanistiques, tandis que la structure (au niveau « macro ») fournit des informations sur l'organisation structurelle du système, qui pourrait illustrer le fonctionnement de nombreux troubles psychiatriques, conçus comme des systèmes adaptatifs complexes possédant des vulnérabilités [30].

Le champ des neurosciences computationnelles et donc de la *Psychiatry Modeling* comprend deux perspectives concernant les modèles formels utilisés : la première correspond à des modèles formels traitant principalement des symboles (c'est-à-dire avec une sémantique et une syntaxe) et sont issus notamment du courant cognitiviste (et les notions de représentation, attente ou croyance), et la seconde correspond à des modèles formels fondés principalement sur la syntaxe et qui sont issus notamment du courant connexionniste (et les notions de dynamique et de réseau).

4.3. Les défis de la modélisation en psychiatrie computationnelle

Les principaux défis que devra relever la modélisation en psychiatrie utilisant les outils des neurosciences computationnelles tiennent à des défis techniques, théoriques et épistémologiques. En termes techniques, la modélisation des comportements complexes ne peut intégrer tous les paramètres dans les équations, les simulations ou les protocoles expérimentaux. En termes théoriques, aucune donnée ni information ne sont théoriquement neutres, conduisant à rattacher des présupposés issus de la psychiatrie ou de la neurobiologie aux modélisations formelles. En termes épistémologiques, les ambitions larges des théories (par exemple du « cerveau bayésien ») contrastent avec l'implémentation relativement restreinte au niveau du codage neuronal. De plus, le fait que certains modèles et concepts soient appliqués largement et de manière ubiquitaire (comme c'est par exemple le cas avec le principe d'énergie libre) ne permet pas de les rendre « potentiellement réfutables ». Leur valeur scientifique devrait donc être discutée.

5. Discussion

5.1. Intrications de la Digital Psychiatry, Big Psychiatry *et* Psychiatry Modeling

Les relations entre ces trois champs de la psychiatrie computationnelle sont essentielles (**Figure 2**). En effet, la *Digital Psychiatry* utilise de nombreux outils numériques qui apparaissent nécessaires pour collecter de grandes quantités de données [38, 37, 39]. La *Big Psychiatry* peut ainsi s'appuyer sur les données recueillies pour développer ses modèles prédictifs ou utiliser ses algorithmes pour traiter ces données (*Machine learning*). Enfin, la modélisation en psychiatrie (*Psychiatry Modeling*) utilisant les neurosciences computationnelles fournit des modèles (*Generative modeling*) qui permettent d'interpréter les grandes quantités de données suivant des contraintes explicatives, et en retour, ces grandes quantités de données peuvent être utilisées pour formaliser des hypothèses et développer, vérifier, réfuter ou confirmer les modèles explicatifs et génératifs du fonctionnement cérébral et cognitif, tout en prenant en compte les contraintes prédictives liées à la clinique et à la thérapeutique. Plus précisément, la psychiatrie computationnelle va pouvoir prendre tout son sens en étant déployée à large échelle, en utilisant des grandes quantités de données que les méthodes mathématiques permettent de manier.

5.2. Phénotypes numériques versus phénotypes computationnels

L'enjeu pourrait alors être de confronter des « phénotypes numériques » (issus du *Machine Learning*) à des « phénotypes computationnels » (issus du *Generative Modeling*). En effet, certaines des données numériques recueillies peuvent devenir des « biomarqueurs numériques » permettant d'identifier un phénotype numérique, grâce aux outils de la *Big Psychiatry* notamment. Ce phénotype numérique est constitué sur la base de données numériques, contrairement à un phénotype clinique constitué de données sémiologiques. Il serait alors nécessaire que le biomarqueur numérique possède une capacité prédictive pour la clinique, associée à une mesure fiable mais aussi valide. Ce biomarqueur pourrait refléter un processus biologique en lien avec un mécanisme physiopathologique identifié ou supposé, pouvant correspondre au phénotype computationnel afin d'asseoir la validité du biomarqueur numérique, ce qui soulève un enjeu de recherche majeur.

La psychiatrie computationnelle paraît disruptive par rapport à ce qui était développé auparavant au sein de la discipline, et si elle n'opère pas de réelle rupture de paradigme (c'est-à-dire de changement radical de théorie qui contredit les découvertes scientifiques passées), il existe cependant un changement de « style de raisonnement scientifique ». Un style de raisonnement

scientifique est une « façon de découvrir », c'est-à-dire une certaine méthode scientifique, un style de pensée, une méthode d'enquête. Ian Hacking, dès 1992 et dans ses cours de 2002–2003 du Collège de France, a élaboré un vaste projet visant à comprendre quels styles de raisonnement scientifique sont utilisés dans les sciences, en se référant aux travaux de l'historien des sciences Alistair Cameron Crombie (1915–1996) [13]. Hacking présente ainsi six styles de raisonnement scientifique qui sont voués à résumer l'intégralité de la méthode scientifique : le style mathématique (géométrique et combinatoire), le style de laboratoire (des instruments et de la création de phénomènes), le style galiléen (de la modélisation), le style taxinomique, le style statistique et des probabilités et enfin le style historico-génétique.

5.3. Styles de raisonnement de la psychiatrie computationnelle

Toutes les disciplines scientifiques, quelles que soient leur origine et leur nature, peuvent être intégrées au sein d'un de ces styles. Chaque style a une histoire propre. Par exemple, le style du laboratoire, né il y a quatre siècles, a bouleversé la science puisqu'autrefois le chercheur observait et spéculait sur les phénomènes, alors qu'en adoptant ce style il s'est mis à fabriquer les phénomènes. Chaque style de raisonnement scientifique a développé un ensemble de techniques qui assure sa stabilité. Par exemple, le style mathématique s'appuie sur des théorèmes mathématiques. De plus, selon Hacking [13] ou Ruphy [29], chacun de ces styles crée une nouvelle classe d'objet. Par exemple, le style statistique (que l'on pourrait faire naître à la suite des travaux de John Graunt qui a estimé la population de Londres en 1662 en s'aidant des registres paroissiaux), a permis de créer un champ disciplinaire et des « groupes statistiques », objets de science qui n'existaient pas auparavant.

La psychiatrie computationnelle provoque un franc changement de style de raisonnement scientifique. Au lieu de s'appuyer sur le style taxonomique (comme c'est le cas pour le *Manuel Diagnostique et Statistique des troubles mentaux*), sur le style historico-génétique fondé sur le jugement d'experts (comme c'est le cas de la psychodynamie) ou sur le style de laboratoire (comme c'est le cas avec une grande partie de la psychiatrie neuroscientifique), cette forme de psychiatrie manie des données chiffrées (provenant de mesures physiologiques, environnementales, cliniques et issues de comportements de la vie réelle). Elle ordonne ces données selon des modèles algorithmiques dont l'objectif est de prédire (notamment classer, classifier, montrer l'évolution) [21]. Elle crée également des modèles mathématiques servant à expliquer le fonctionnement du cerveau, modèles qui sont par la suite confirmés ou réfutés par les données. La psychiatrie computationnelle s'inscrit donc dans un style tantôt mathématique, tantôt statistique.

6. Conclusion

Dans ces interactions entre les trois champs de la psychiatrie computationnelle (**Figure 2**), de nombreux enjeux éthiques sont encore à prendre en compte. En effet, dans le cadre d'un important recueil de données (par exemple au sein de « lacs de données »), quelle valeur auront les données qui ne sont pas médicalement utiles ? De plus, le traitement par des modèles prédictifs d'apprentissage automatique pourra-t-il permettre de déterminer des biomarqueurs numériques à la fois stables, intégrables à des modèles, correspondant à des strates pertinentes de patients et actionnables pour la clinique ? Et enfin, les neurosciences computationnelles sauront-elles utiliser leurs modèles pour aider à la supervision de ce traitement massif de données ?

Conflit d'intérêt : Les auteurs déclarent n'avoir pas d'intérêt en lien avec cet article.

Références

[1] Beaumont S, Dhôte J, Vinckier F, Palminteri S, Wyart V, Jardri R, Domenech P. Introduction à la psychiatrie computationnelle. EMC - Psychiatrie; 2020. https://doi.org/10.1016/S0246-1072(20)69595-5

[2] Bzdok D, Varoquaux G, Steyerberg EW. Prediction, Not Association, Paves the Road to Precision Medicine. JAMA Psychiatry; 2020. https://doi.org/10.1001/jamapsychiatry.2020.2549

[3] Cohen AS, Schwartz E, Le T, Cowan T, Cox C, Tucker R, et al. Validating digital phenotyping technologies for clinical use: The critical importance of "resolution." World Psychiatry 2020;19:114–5. https://doi.org/10.1002/wps.20703

[4] Demazeux S. (). L'Éclipse du symptôme. L'observation clinique en psychiatrie, 1800-1950. Ithaque; 2019.

[5] Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, et al. The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular Psychiatry 2014;19:659–67. https://doi.org/10.1038/mp.2013.78

[6] Dickinson A. Contemporary Animal Learning Theory. CUP Archive; 1980.

[7] Dumas G, Cermolacce M, Batail JM, Quiles C, Micoulaud-Franchi JA. Vers une approche physiologique de la sémiologie en psychiatrie. Partie 2: Perspectives offertes par la biologie systémique. Ann Méd Psychol 2019;177:289–94. https://doi.org/10.1016/j.amp.2019.01.002

[8] Friston KJ, Stephan KE, Montague R, Dolan RJ. Computational psychiatry: The brain as a phantastic organ. The Lancet. Psychiatry 2014;1:148–58.

[9] https://doi.org/10.1016/S2215-0366(14)70275-5

[10] Friston K, Mattout J, Kilner J. Action understanding and active inference. Biological Cybernetics 2011;104:137–60. https://doi.org/10.1007/s00422-011-0424-z

[11] Gauld C, Dumas G, Darrason M, Salles N, Desvergnes P, Philip P, Micoulaud-Franchi JA. Médecine du sommeil personnalisée et syndrome d'apnées hypopnées obstructives du sommeil: Entre précision et stratification, une proposition de clarification. Médecine du Sommeil 2020. https://doi.org/10.1016/j.msom.2020.08.003

[12] Gauld C, Maquet J, Micoulaud-Franchi JA, Dumas G. Mining big data about representations of autism spectrum disorder: A comparison from Twitter to PubMed, a TwiMed proof-of-concept. (in prep.).

[13] Geoffroy PA. Étude portant sur les patients bipolaires et les conditions météorologiques. (in prep.).

[14] Hacking I. 'Style' for historians and philosophers. Studies in History and Philosophy of Science Part A 1992;23:1–20. https://doi.org/10.1016/0039-3681(92)90024-Z

[15] Hempel C. Aspects of Scientific Explanation. Free Press 1965.

[16] Hilborn R, Mangel M. The Ecological Detective: Confronting Models with Data (1st edition). Princeton University Press; 1997.

[17] Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Sanislow C, Wang P. Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. Am J Psychiatry 2010;167:748–51.

https://doi.org/10.1176/appi.ajp.2010.09091379

[18] Insel TR. Digital phenotyping: A global tool for psychiatry. World Psychiatry 2018;17:276–7. https://doi.org/10.1002/wps.20550

[19] Knill DC, Pouget A. The Bayesian brain: The role of uncertainty in neural coding and computation. Trends in Neurosciences 2004;27:712–9. https://doi.org/10.1016/j.tins.2004.10.007

[20] Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron 2017;93:480–90. https://doi.org/10.1016/j.neuron.2016.12.041

[21] Love BC. The Algorithmic Level Is the Bridge Between Computation and Brain. Topics in Cognitive Science 2015;7:230–42. https://doi.org/10.1111/tops.12131

[22] Marr D. Vision: A computational investigation into the human representation and processing of visual information. MIT Press; 2010.

[23] McCoy LG, Nagaraj S, Morgado F, Harish V, Das S, Celi LA. What do medical students actually need to know about artificial intelligence? Npj Digital Medicine 2020;3:1–3. https://doi.org/10.1038/s41746-020-0294-7

[24] Micoulaud-Franchi JA, Quiles C, Batail JM, Daudet C, Cermolacce M, Dumas G. Vers une approche physiologique de la sémiologie en psychiatrie. Partie 1: Approches RDC, DSM, RDoC et HiTOP.Ann Méd Psychol 2019;177:282–8. https://doi.org/10.1016/j.amp.2019.01.001

[25] Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. Trends in Cognitive Sciences 2012;16:72–80. https://doi.org/10.1016/j.tics.2011.11.018

[26] Monteith S, Glenn T, Geddes J, Bauer M. Big data are coming to psychiatry: A general introduction. International Journal of Bipolar Disorders 2015;3:21. https://doi.org/10.1186/s40345-015-0038-9

[27] Pessiglione M, Vinckier F, Bouret S, Daunizeau J, Le Bouc R. Why not try harder?

Computational approach to motivation deficits in neuro-psychiatric diseases. Brain: A Journal of Neurology 2018;141:629–50. https://doi.org/10.1093/brain/awx278

[28] Pinel. (1807). Résultats d'observations " Mdm. Sci. Math. Et phys."

[29] Porter TM. The Rise of Statistical Thinking. 1986.

[30] Ruphy S. Scientific Pluralism Reconsidered: A New Approach to the (Dis)Unity of Science. University of Pittsburgh Press; 2017.

[31] Saxe GN.Network Psychiatry: Computational Methods to Understand the Complexity of Psychiatric Disorders. Journal of the American Academy of Child and Adolescent Psychiatry 2017;56:639–41. https://doi.org/10.1016/j.jaac.2017.06.001

[32] Shiffman S, Stone AA, Hufford MR. Ecological Momentary Assessment. Annual Review of Clinical Psychology 2008;4:1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

[33] Stephan KE, Bach DR, Fletcher PC, Flint J, Frank MJ, Friston KJ. Charting the landscape of priority problems in psychiatry, part 1: Classification and diagnosis. The Lancet Psychiatry 2016;3:77–83. https://doi.org/10.1016/S2215-0366(15)00361-2

[34] Stephan KE, Binder EB, Breakspear M, Dayan P, Johnstone EC, Meyer-Lindenberg A. Charting the landscape of priority problems in psychiatry, part 2: Pathogenesis and aetiology. The Lancet Psychiatry 2016;3:84–90.

[35] https://doi.org/10.1016/S2215-0366(15)00360-0

[36] Strogatz SH. Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering (1st edition). CRC Press; 2000.

[37] the Alzheimer's Disease Neuroimaging Initiative, EPIGEN Consortium, IMAGEN Consortium, Saguenay Youth Study (SYS) Group, Thompson, P. M., Stein, J. L., Medland, S. E., Hibar, D. P., Vasquez, A. A., Renteria, M. E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., Wright, M. J., Martin, N. G., Agartz, I., Alda, M., Alhusaini, S., Almasy, L., Almeida, J., Alpert, K., … Drevets, W. (2014). The ENIGMA Consortium: Large-scale

collaborative analyses of neuroimaging and genetic data. Brain Imaging and Behavior, 8(2), 153–182. https://doi.org/10.1007/s11682-013-9269-5

[38] The Structure and Dynamics of Networks 2006. https://press.princeton.edu/books/paperback/9780691113579/the-structure-and-dynamics-of-networks

[39] Torous J, Hsin H. Empowering the digital therapeutic relationship: Virtual clinics for digital health interventions. Npj Digital Medicine 2018;1:1–3. https://doi.org/10.1038/s41746-018-0028-2

[40] Torous J, Kiang MV, Lorme J, Onnela JP. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. JMIR Mental Health 2016;3:e16. https://doi.org/10.2196/mental.5165

[41] Torous J, Stern AP, Padmanabhan JL, Keshavan MS, Perez DL.A proposed solution to integrating cognitive-affective neuroscience and neuropsychiatry in psychiatry residency training: The time is now. Asian Journal of Psychiatry 2015;1:116–21. https://doi.org/10.1016/j.ajp.2015.05.007

[42] Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K. The WU-Minn Human Connectome Project: An overview. NeuroImage 2013;80:62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

[43] Yadav SP. The Wholeness in Suffix -omics, -omes, and the Word Om. Journal of Biomolecular Techniques : JBT 2007;18:277.

FIGURES

Figure 1. Nombre de publications portant sur la psychiatrie computationnelle (*Digital Psychiatry*, *Big Psychiatry* and *Psychiatry Modeling*).

Source : Pubmed ([psychiatry OR psychology OR psychopathology] AND "Computing Methodologies"[Mesh] OR "Artificial Intelligence"[Mesh] OR "Mathematical Concepts "[Mesh] OR "Computational"[Mesh] OR "Digital"[Mesh])). À noter que les données n'ont pas été normalisées en fonction de l'évolution du terme « psychiatrie », « psychologie » ou « psychopathologie » dans la base PubMed par année.

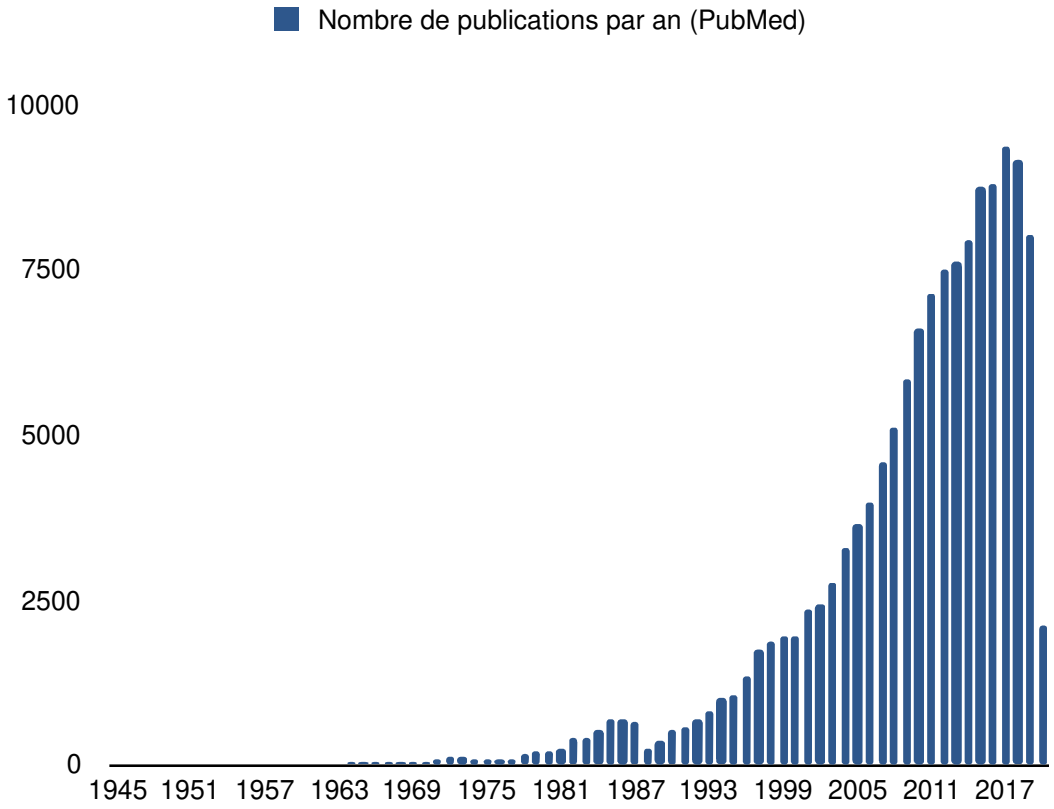
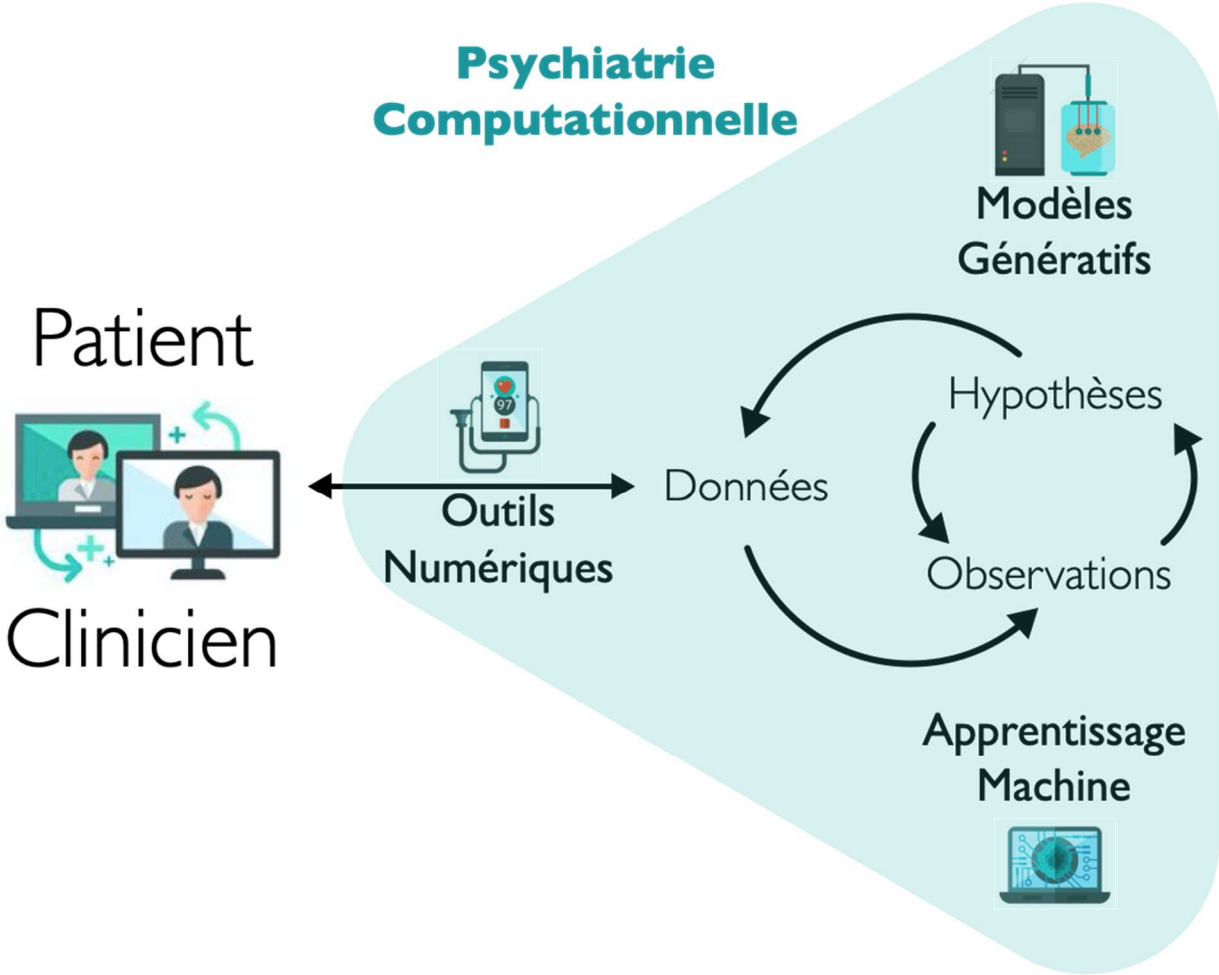
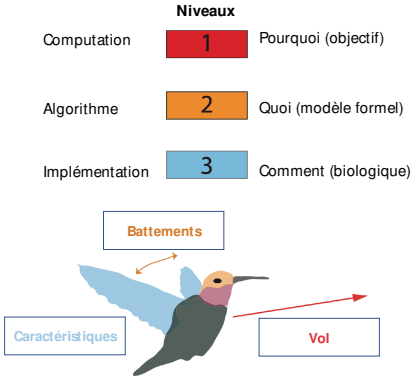


Figure 2. Trois champs principaux de la psychiatrie computationnelle : *Digital Psychiatry*, *Big Psychiatry* and *Psychiatry Modeling* avec respectivement les « Outils Numériques » (*Digital Tools*), l'« Apprentissage Automatique » / « Apprentissage Machine » (*Machine Learning*) et les « Modèles Génératifs » (*Generative Modeling*).



ENCADRÉ



Le cadre de travail le plus connu pour formaliser le lien entre le cerveau et le comportement correspond aux trois niveaux de David Marr (conçu à l'origine comme une critique du travail réductionniste en neurophysiologie dans le cadre des systèmes complexes) (Marr, 1982/2010).

Le Niveau 1 correspond au niveau physique ou biologique, le Niveau 2 aux modèles formels et le Niveau 3 à l'objectif. Zenon Pylyshyn (1984) appelle ces niveaux, « Physique » (Niveau 1), « Syntaxique » (Niveau 2) et « Sémantique » (Niveau 3). D'autres auteurs comme Glass, Holyoak et Santa avaient pu les nommer dès 1979 « Support » (Niveau 1), « Forme » (Niveau 2) et « Contenu » (Niveau 3).

« Tenter de comprendre la perception en comprenant les neurones, c'est comme essayer de comprendre le vol d'un oiseau en étudiant uniquement les plumes. Cela n'est tout simplement pas possible » (Marr, 1982/2010). L'intuition de Marr est qu'il est beaucoup plus difficile de déduire à partir du matériel neutre (c'est-à-dire de « l'implémentation », le Niveau 3) quel algorithme (le Niveau 2) le système nerveux utilise pour analyser un problème (le Niveau 1).

Dans la figure, un oiseau tente de voler (objectif) en battant des ailes (modèle formel), et son aérodynamique dépend des caractéristiques de ses plumes (physique ou biologique). Les trois niveaux sont intimement liés. En pratique, il faut d'abord observer le comportement et il est ensuite possible de tester des règles simples qui conduiront à la reproduction du comportement (par exemple, par simulation). Autrement dit, il convient de demander d'abord pourquoi le cerveau exécute un tel comportement, puis de vérifier la manière dont ce comportement est mis en œuvre.

Une autre analogie est la compréhension du jeu d'échec : comprendre le jeu d'échec ne dépend pas de la manière dont le plateau ou les pièces sont fabriqués (Niveau 1, physique ou biologique) (Marr suggère que les composants du système nerveux pourraient même n'avoir aucune importance pour comprendre les comportements). La compréhension des échecs dépend des règles (Niveau 2, les modèles formels) et de l'objectif (Niveau 3, computationnel).

GLOSSAIRE

Apprentissage automatique (ou Apprentissage machine) : Principale méthode d'Intelligence Artificielle qui permet de faire apprendre à la machine. L'apprentissage automatique peut être classé en quatre types : classique, renforcement, réseaux de neurones, des ensembles. Il peut également être classé suivant ses objectifs : apprentissage supervisé, non supervisé, par renforcement ou semi-supervisé.

Apprentissage par renforcement : Une branche de l'apprentissage automatique inspirée de la psychologie comportementale qui modélise le comportement comme une maximisation de la récompense attendue.

Biomarqueur : Indicateur mesurable, fiable et valide, qui reflète un processus biologique en lien avec un mécanisme physiopathologique identifié ou supposé.

Biomarqueur numérique : Biomarqueur collecté au moyen d'un outil numérique.

Computation : Opération de calcul qui permet de répondre à un problème en utilisant des données et des algorithmes, c'est-à-dire des règles.

Computationnel : On associe au terme de sciences du numérique (ou sciences du digital) les sciences dites « computationnelles ». La notion de computation vient du latin *computatio*, qui renvoie au calcul (ou opération). Selon Turing, les sciences computationnelles correspondent de manière large à toute science du calcul. Par la suite, la dénomination de science computationnelle a été l'un des trois niveaux décrits par Marr (1982) : voir **Encadré**. Actuellement, la science computationnelle intègre une conception connexionniste, syntaxique et formelle. Elle peut donc être comprise comme une méthodologie de traitement des données numériques servant à résoudre un problème de la psychiatrie de manière mathématique et informatique au sein d'un modèle donné (par exemple, une équation).

Connexionnisme : Il se définit par référence à l'utilisation de réseaux de neurones formels, c'est-à-dire la représentation mathématique et informatique d'un ensemble de neurones biologiques. Le principe de base du connexionnisme est que les phénomènes mentaux peuvent être décrits à l'aide de réseaux d'unités interconnectées. Il s'agit historiquement d'un traitement de l'information qui est

1) séquentiel, 2) distribué (pas d'unité centrale), 3) gérant des poids de connexion et 4) il traite des données continues.

Déduction / Induction / Abduction : L'épistémologie propose des outils pour expliquer comment améliorer des modèles, c'est-à-dire des représentations ordonnées de la réalité. Il existe deux principales conceptions épistémologiques des modèles. Une première conception épistémologique propose d'émettre des hypothèses, qui seront ensuite confirmées par les données (ou des « faits » dans la tradition sociologique). Il s'agit de la conception épistémologique dite « déductive-nomologique », dans laquelle une déduction est faite à partir d'hypothèses et de lois. Une seconde conception épistémologique propose d'accumuler des données qui vont ensuite confirmer des hypothèses. Il s'agit d'une conception dite « inductive-statistique », dans laquelle la relation entre des statistiques et une hypothèse préconçue est réalisée. Historiquement, dans cette conception et en médecine, les données sont principalement apportées par les technologies. Dans le cadre du paradigme numérique, il est fait mention d'approche « *top-down* » pour la conception hypothético-déductive (en lien avec la *Psychiatry Modeling*) et d'approche « *bottom-up* » pour la conception inductive-statistique (en lien avec la *Big Psychiatry*). Enfin, des auteurs comme Peirce (1931 – 1958) ont introduit la notion d'abduction, qui permet de passer d'une observation élémentaire à une hypothèse. L'abduction est donc « *bottom-up* », car il s'agit de partir des données (comme l'induction), mais à la différence de l'induction, elle peut faire émerger une hypothèse (comme la déduction), comme proposé dans la **Figure 2**, montrant les apports mutuels des trois cultures de la psychiatrie computationnelle.

Donnée : Toute entité élémentaire recueillie, comme par exemple un symptôme ou une donnée numérique, à l'aide d'un modèle donné et par une technologie particulière.

Intelligence artificielle (IA) : Discipline scientifique, au même titre que la chimie ou la médecine, faite de sciences des machines (sciences computationnelles), d'ingénierie, de robotique, etc., qui utilise comme technologie principale des machines (ordinateurs).

Loi de Bayes : Règle de mise à jour des croyances fondée sur le fait que la probabilité que deux événements se produisent ensemble est la probabilité du premier événement, en supposant que le second s'est produit, multipliée par la probabilité que le deuxième événement se produise.

Modèle : Représentation de la réalité qui permet d'ordonner des données, comme par exemple une classification. Un modèle est une représentation de la réalité. Un modèle rassemble des données

dans un cadre commun. Il peut chercher à ressembler la réalité (comme la carte du métro de Paris qui est fidèle à son territoire géographique) ou à être utile en termes de calcul (comme la carte du métro de Londres qui n'est pas fidèle par rapport à son territoire géographique mais permet une meilleure représentation du trajet).
Un modèle est une structure qui permet d'agencer les données, selon l'image de briques (les données, en l'occurrence des biomarqueurs numériques) rassemblées pour construire une maison (le modèle). Le modèle de classification le plus utilisé correspond au DSM (modèle descriptif fondé sur des données sémiologiques). Mais il existe d'autres modèles de classification, comme les *Research Domain Criteria* (modèle fondé sur des données neuroscientifiques), la *Hierarchical Taxonomy of Psychopathology* (modèle fondé sur des données statistiques), les modèles de stadification (modèles dynamiques prenant en compte l'évolution longitudinale des troubles) ou des réseaux de symptômes (modèle fondé sur l'hypothèse d'une influence mutuelle entre symptômes).

Paradigme : Un paradigme est défini comme un ensemble de points de vue et des règles permettant de répondre aux problèmes circonscrits dans un champ scientifique. Pour constituer un paradigme, il faut qu'un ensemble cohérent « de lois, de théories, d'applications et de dispositifs expérimentaux » puisse fournir « des modèles qui donnent naissance à des traditions particulières de recherche » (Kuhn, 1962). Autrement dit, un paradigme repose sur l'idée d'améliorer l'unification des idées. Par exemple, concernant le paradigme neuroscientifique, un de ses projet est de décrire et unifier les différents niveaux du vivant et de l'environnement. Cette unification passe par une connaissance accrue des faits concernant le fonctionnement cérébral et par une amélioration de la précision des prédictions. En neuroscience, un modèle peut être, par exemple, l'explication du mécanisme de la peur, ou le circuit de la mémoire ou de la récompense).

Pattern : Un ensemble de données modélisées, ou un ensemble d'observations.

Phénotype : Ensemble de traits mesurables au sein d'un organisme.

Phénotype computationnel : Phénotype comportemental ou neuronal mesurable, défini en termes d'un modèle computationnel. Par analogie avec d'autres phénotypes, un phénotype computationnel devrait montrer une variation entre les individus et la sélection naturelle pourrait agir sur cette variation. Le phénotypage computationnel à grande échelle chez l'homme n'a pas été réalisé ; par conséquent, l'utilité ultime du concept n'a pas été rigoureusement testée.

Phénotype numérique : Ensemble de biomarqueurs numériques déterminant un phénotype par des mesures des comportements en vie réelle, à l'aide d'appareils numériques.

Rupture de paradigme : La psychiatrie computationnelle modifie la perception de la psychiatrie sans pour autant entraîner de véritable rupture paradigmatique. La rupture d'un paradigme scientifique est liée à la perception d'une « curiosité scientifique » qui serait difficilement explicable par le paradigme en cours. Dans l'histoire de la science, cette curiosité prend de l'ampleur et peut aller jusqu'à provoquer une crise obligeant à changer de paradigme, suivant le modèle de Kuhn (1962). La véritable rupture paradigmatique reste rare. Le plus fréquemment, on assiste en science à des ajustements des théories existantes aux anomalies constatées[2]. Dans le cas de la psychiatrie computationnelle, les technologies numériques apportent de nouvelles données, ce qui permet la création de nouveaux modèles explicatifs et prédictifs pour la psychiatrie. Ces nouveaux modèles peuvent certes modifier la pratique clinique, mais ils ne contredisent pas l'intégralité des découvertes scientifiques gravitant autour de la psychiatrie. La psychiatrie computationnelle n'opère donc pas de rupture de paradigme. De plus, selon Kuhn (1962), il est particulièrement important que les nouvelles théories puissent conserver l'essentiel des résultats passés.

Sémantique : Renvoie couramment au sens des mots ou plus généralement des entités élémentaires d'un système. Dans le cadre de cet article, il renvoie aux composants d'un système porteur d'une représentation, c'est-à-dire un système intentionnel qui porte un sens, un symbole, un contenu. Il représente le Niveau 3 selon Marr (voir **Encadré**).

Syntaxique : Renvoie couramment à la grammaire et définit la manière dont les entités élémentaires doivent s'agencer entre eux. Dans le cadre de cet article, il renvoie aux composants d'un système qui ne porte pas de représentation (ou contenu), mais gère la structure du système. Il représente le Niveau 2 selon Marr (voir **Encadré**).

Théorie des jeux : L'étude des modèles mathématiques d'interactions entre agents rationnels.

---

[2] En général, un paradigme scientifique (par exemple, le paradigme neuroscientifique) écarte les problèmes qui ne se posent pas en termes compatibles avec le paradigme : on dit que la science normale garantit l′ existence d′ une solution.