

# IA@ONERA

## Intelligence artificielle

Quels enjeux pour l'ASD ?

Quels objectifs pour l'ONERA ?



**L'intelligence artificielle**  
une impulsion de recherche et d'innovation  
pour tout le domaine  
**aéronautique-spatial-défense**

## Introduction

Les évolutions techniques de l'intelligence artificielle (l'IA) seront très certainement au cœur d'un nombre important d'innovations de la filière industrielle aéronautique-spatial-défense dans les toutes prochaines années. L'IA s'adresse en effet déjà de manière importante aux techniques de conception, de simulation, de commande et d'autonomie des systèmes aérospatiaux ainsi qu'aux systèmes de reconnaissance, de surveillance, de robotisation et plus largement de traitement de l'information pour les besoins de la défense. D'une manière plus large, l'IA aura certainement un impact à terme sur la pratique scientifique et l'ingénierie.

L'ONERA doit anticiper et accompagner ce mouvement par une recherche appropriée au domaine ASD et par un développement des applications dont le potentiel est tout à fait important.

Dans ce livre blanc de l'IA à l'ONERA, à partir d'une analyse de certaines opportunités pour la recherche à l'ONERA – analyse orientée par notre potentiel et les recherches actuelles –, nous proposons un ensemble d'axes de recherche et de développement. Pour un certain nombre, ceux-ci ne font que prolonger voire amplifier des travaux actuels. D'autres cherchent à couvrir plus largement les applications potentielles de l'IA dans le domaine ASD ou à répondre à des défis scientifiques comme ceux de la robustesse, de l'explicabilité, du passage à l'échelle des méthodes d'apprentissage auquel nous pensons pouvoir contribuer de manière particulière compte-tenu de l'expertise singulière de l'ONERA.

L'intégration des activités de l'IA à l'ONERA dans le contexte national et international, les indispensables coopérations avec le monde académique et industriel sont discutées dans la dernière partie du document.

# *Sommaire*

<b>INTRODUCTION.....</b>	<b>5</b>
<b>1. L'IA, UNE NOUVELLE DYNAMIQUE .....</b>	<b>9</b>
1.1 Dès 1976, à l'ONERA .....	9
1.2 Aujourd'hui à l'ONERA.....	9
1.3 L'IA dans le monde et en France (2018) .....	11
<b>2. L'IA ET SES DOMAINES SCIENTIFIQUES .....</b>	<b>13</b>
2.1 L'IA : de quoi parle-t-on ? .....	13
2.2 L'IA : un peu d'histoire ... ..	13
2.3 L'IA, ses techniques et ses applications aujourd'hui.....	19
2.3.1 Apprentissage artificiel .....	19
2.3.2 Planification et décision .....	26
2.3.3 Autonomie et robotique.....	30
2.3.4 Interactions et cognitive .....	33
2.4 L'IA et les calculateurs .....	36
2.5 L'IA et les données.....	37
2.7 L'IA et les mathématiques .....	38
<b>3. L'IA, QUELS ENJEUX POUR LE DOMAINE ASD.....</b>	<b>39</b>
3.1 Aéronautique (civile et militaire).....	39
3.1.1 Modèles hybrides pour la simulation numérique.....	40
3.1.2 Réduction de modèles et représentation des connaissances expertes pour l'aide à la conception .....	41
3.1.3 Smart cockpit – Smart skies .....	42
3.1.4 Systèmes à forte autonomie, leur sûreté et leur sécurité .....	43
3.1.5 Systèmes de production agile .....	43
3.1.6 Maintenance individualisée – Pronostic et diagnostic .....	44

3.2	Spatial .....	45
3.2.1	Le renseignement géospatial .....	46
3.2.2	Satellites agiles – Autonomie des missions spatiales .....	47
3.2.3	Surveillance de l'espace.....	47
3.3	Défense/Sécurité .....	49
3.1.1	Système de combat aérien du futur .....	50
3.1.2	Surveillance de l'environnement opérationnel .....	51
3.1.3	La perception active et distribuée .....	51
3.1.4	Systèmes de défense multi-vecteurs .....	52
3.1.5	Prévention des risques et sécurité civile.....	53

#### **4. IA@ONERA : QUELS AXES STRATÉGIQUES ?..... 54**

4.1	Recherches concernant le « cœur » de l'IA .....	54
4.1.1	L'intégration de représentations symboliques et de méthodes formelles aux modèles avec un apprentissage automatique ....	54
4.1.2	La modélisation hybride robuste, combinant des modèles physiques avec modèles d'apprentissage automatique .....	55
4.1.3	La distribution de modèles d'apprentissage automatique dans des systèmes multi-agents ou dans des grands systèmes.....	55
4.2	Méthodes de l'IA pour les applications dans le domaine ASD .....	56
4.2.1	Perception augmentée pour la compréhension de scènes dynamiques .....	56
4.2.2	Robots, drones et systèmes intelligents pour la sécurité et la défense.....	56
4.2.3	Sûreté et sécurité des systèmes autonomes .....	57
4.2.4	Perception artificielle cognitive & perception distribuée.....	57
4.2.5	Apprentissage automatique pour le renseignement d'origine électromagnétique .....	57
4.2.6	Application de l'apprentissage par renforcement à la commande et synthèse de comportements .....	58
4.2.7	Modèles d'intégration d'informations multi-sources .....	58
4.2.8	Simulation physique hybride, prédiction du comportement et aide à la compréhension de phénomènes.....	58

4.2.9	Planification des missions spatiales .....	59
4.3	IA@ONERA : insertion dans l'environnement académique et le domaine ASD .....	59
4.3.1	Quels objectifs .....	59
4.3.2	Collaboration nationale autour de grands projets.....	61
4.3.3	Collaborations internationales .....	62
<b>ABRÉVIATIONS .....</b>		<b>64</b>

## 1. L'IA, une nouvelle dynamique

### 1.1 Dès 1976, à l'ONERA

Les premiers travaux de recherche en intelligence artificielle (IA) à l'ONERA ont débuté en 1976 dans le cadre d'une collaboration entre les départements de recherche en automatique et en informatique (DERA et DERI). Ceux-ci portaient alors principalement sur les méthodes de raisonnement pour la génération de plans d'actions en robotique.

En 1987, la décision est prise de créer une équipe en intelligence artificielle rassemblant les automaticiens et informaticiens impliqués dans les activités en IA, équipe intégrée à la direction des moyens informatiques et placée sous la responsabilité de Jean Erceau (à droite sur la photo).



#### – Création du groupe d'intelligence artificielle

Quelques mois plus tard, le président de l'ONERA officialise l'existence du groupe d'intelligence artificielle (GIA), qui est constitué à sa création de 33 personnes (18 ingénieurs et 15 doctorants), issues de différents départements.

Les thématiques placées au cœur des activités du GIA étaient alors : la modélisation des connaissances et le raisonnement symbolique, la recherche opérationnelle pour les problèmes à forte combinatoire, l'apprentissage symbolique, les systèmes experts, les systèmes multi-agents, la classification en reconnaissance de formes, les interfaces homme-machine, etc. Les domaines d'application considérés étaient la robotique mobile, la robotique industrielle, les missions spatiales, etc. Une action de formation fut également mise en place pour former des ingénieurs « cogniticiens » à l'ONERA.

Très vite, des collaborations se sont développées avec des laboratoires universitaires des universités Paris 6, Paris 5, Paul Sabatier (Toulouse) et plus largement avec les membres du PRC IA (programme de recherches coordonnées) créé par le ministère de la Recherche (sur le Fonds de la recherche technologique) en 1985. En 1987, ce programme fut transformé en un groupement de recherche (GdR) par le CNRS. D'autres collaborations ont vu le jour par la suite notamment avec le CNES, l'Aérospatiale, PSA et des PME comme Verilog, Techlog, etc.

#### – Les premiers travaux sur les réseaux de neurones artificiels

Le traitement des images puis des flux vidéo (la vision par ordinateur) est devenu très tôt un domaine d'intérêt majeur pour l'Intelligence Artificielle à l'ONERA. La reconnaissance des formes par appariement puis l'interprétation des images a été l'un des cadres d'expression du potentiel de premiers réseaux de neurones artificiels. Les techniques de régression non-linéaires, fondées sur les *perceptrons* multicouches à rétro-propagation, ont été développées à l'ONERA dès le début des années 90 pour la classification des images.

### 1.2 Aujourd'hui à l'ONERA

Aujourd'hui, les techniques relevant de l'intelligence artificielle sont développées à l'ONERA principalement au département Traitement de l'information et systèmes (DTIS) et sont exploitées dans l'ensemble des départements scientifiques.

Après la réorganisation de 2018, le DTIS a défini « l'intelligence artificielle et la décision » comme l'une de ses huit thématiques scientifiques. Une trentaine de permanents et une vingtaine de doctorants développent des travaux au sein de cette thématique dans cinq domaines techniques.

### Domaines techniques de l'IA dans lesquels l'ONERA est plus particulièrement impliqué :

- 1) **la perception artificielle**, qui constitue un véritable moteur pour le développement des techniques d'apprentissage statistique destinées à la classification ;
- 2) **l'autonomie des systèmes et la robotique** où les techniques de planification d'actions fondées sur l'apprentissage symbolique et de navigation par localisation visuelle sont particulièrement importantes ;
- 3) **l'apprentissage de modèles** de systèmes dynamiques (modèles non-linéaires) ou de comportements (modèles hybrides) pour la commande des systèmes aéronautiques et robotiques ou la simulation d'humains virtuels ;
- 4) **les modèles d'approximation** (régression) à partir de résultats d'expériences ou de calcul dans différents domaines scientifiques de l'ONERA pour les besoins de l'optimisation en conception, de la prédiction de phénomènes, ou de la simulation hybride ;
- 5) **la logique et la décision** : la représentation des connaissances et des préférences, les modèles de décision distribuée et la décision en présence de points de vue multiples et/ou dans l'incertain.

Au-delà du DTIS, les départements Optique et techniques associées (DOTA) et Électromagnétisme et radar (DEMR) font usage des techniques d'apprentissage supervisé pour la classification des images satellitaires ou aéroportées dans des finalités de modélisation de l'environnement et d'observation de l'espace. Le département d'Aérodynamique, acoustique, aéroélasticité (DAAA) utilise des modèles réduits et de l'estimation par apprentissage par renforcement et le DMAS (département Matériaux et structures) des régressions multilinéaires par réseaux de neurones pour prédire les propriétés d'alliages en fonction de leur composition chimique. Au département Multi-physique pour l'énergétique (DMPE), on exploite l'apprentissage pour l'identification des propriétés acoustiques des matériaux. Le DPHY (département Physique, instrumentation, environnement, espace) travaille sur l'utilisation de réseaux de neurones pour l'analyse des séries temporelles.

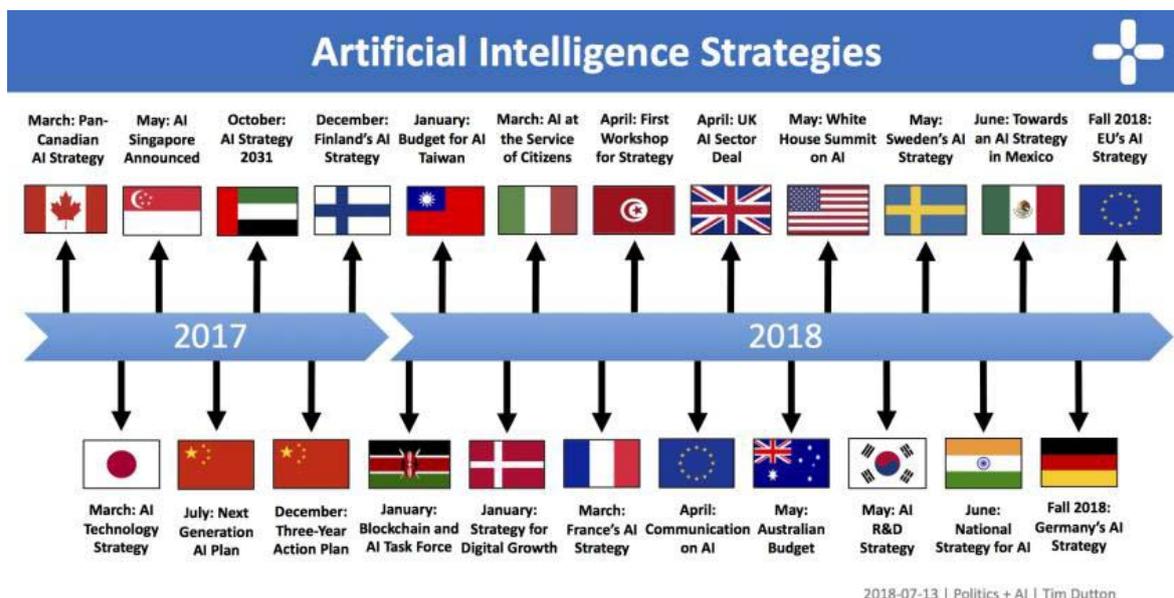
À noter également que depuis plusieurs années, des formations internes sont proposées aux chercheurs, ingénieurs et doctorants de l'ONERA sur l'apprentissage et les modèles stochastiques pour anticiper et accompagner l'exploitation des techniques de l'IA.

Une structure interne d'animation scientifique a été mise en place en 2018 pour favoriser la diffusion de ces techniques dans des domaines autres que le domaine du Traitement de l'information et des systèmes mais aussi pour mener des investigations sur des technologies émergentes et structurer la diffusion des ressources logicielles de l'ONERA en IA (<https://www.onera.fr/IA>).

### Chiffres clés de l'IA@ONERA

- 1987 : création du groupe Intelligence artificielle et premiers travaux
- 35 chercheurs développant aujourd'hui les techniques de l'IA
- 12 thèses en cours sur les techniques de l'IA
- 8 projets internes impliquant l'IA : DELTA, ADAC, R2D2, MOSKITO, SUPER, MEDUSE, GUIMAUVE, CADENCE
- 6 projets collaboratifs : ANR-JCJC, H2020, partenariat recherche industrie (SNCF)
- 4 plates-formes logicielles de capitalisation des développements (DeepLab, InCell, GENETIC, MEDUSE)
- 2 instituts 3IA dans lesquels l'ONERA est impliqué : ANITI et 3IA Côte d'Azur

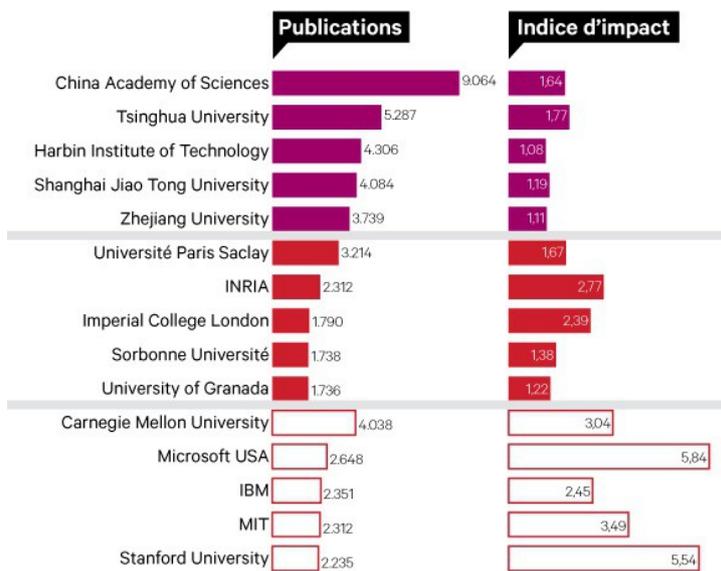
## 1.3 L'IA dans le monde et en France (2018)



### Publications scientifiques sur l'IA dans le monde

Top 5 des établissements en Chine, en Europe et aux Etats-Unis pour la période 2013-2017

■ Chine ■ Europe □ Etats-Unis



\* LES ECHOS \* / SOURCE : ELSEVIER ; SCIVAL

Tous les grands pays industrialisés se sont dotés d'un programme national en IA ces dernières années. Parmi les pays les plus ambitieux on compte la Chine, la Corée du Sud, le Japon, les États-Unis<sup>1</sup> (pour plus de détails lire<sup>2</sup>). L'Europe a également mis en place en 2018 un plan IA<sup>3</sup>. Pour bonne part, ces plans visent à soutenir le développement de la recherche à moyen et long termes, recherche qui dans certains domaines de l'IA connaît une croissance spectaculaire ces dix dernières années : les publications sur les réseaux de neurones ont eu un taux de croissance annuel d'environ 3% entre 2010 et 2014 ; ce taux est passé à 37% entre 2014 et 2017 ; et de 2010 à 2017 la part des publications en « apprentissage automatique et raisonnement probabiliste » a doublé par rapport à l'ensemble de la production scientifique en IA.

L'Europe arrive en tête devant la Chine et les États-Unis en termes de publications scientifiques sur l'intelligence artificielle. Selon une étude publiée par l'éditeur Elsevier<sup>4</sup>, les établissements français Université Paris-Saclay, INRIA et

<sup>1</sup> [https://www.nitrd.gov/news/national\\_ai\\_rd\\_strategic\\_plan.aspx](https://www.nitrd.gov/news/national_ai_rd_strategic_plan.aspx)

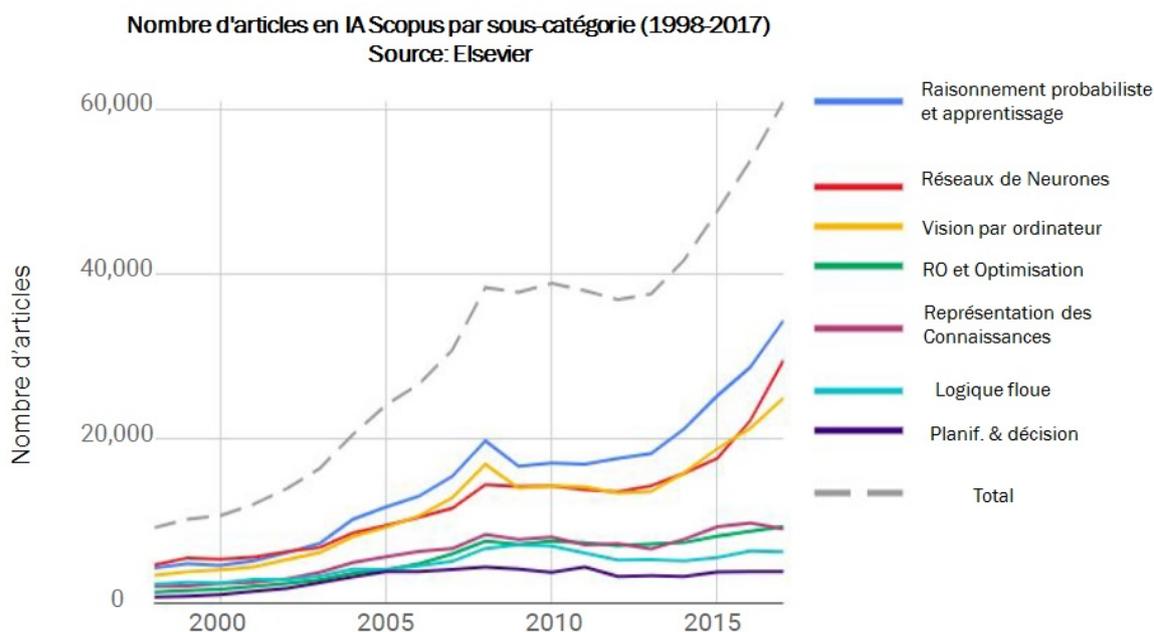
<sup>2</sup> <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>

<sup>3</sup> <https://fpf.org/2018/07/19/policy-brief-european-commissions-strategy-for-ai-explained/>

<sup>4</sup> *Artificial Intelligence : How Knowledge is Created, Transferred, and Used*, Elsevier Artificial Intelligence resource centre

Sorbonne Université apparaissent en bonne position au niveau européen puisqu'ils se situent respectivement en première, deuxième et quatrième place en nombre de publications. L'étude relève que, contrairement à d'autres domaines, les publications se font essentiellement dans des actes de conférences.

Pour soutenir la dynamique française autour de l'IA, le secrétaire d'État chargé de l'Industrie, du Numérique et de l'Innovation et le secrétaire d'État chargé de l'Enseignement supérieur et de la Recherche ont lancé, le 20 janvier 2017, la démarche dite « France IA » (voir rapport<sup>5</sup>). Cette démarche de mobilisation de la communauté française de l'intelligence artificielle a pour objectif principal l'élaboration d'une stratégie nationale en s'appuyant en particulier sur la recherche publique. En IA, la France est historiquement engagée de manière significative dans une grande diversité de domaines relevant de l'IA.



Évolution des articles de recherche consacrés à l'intelligence artificielle au cours des 20 dernières années, par sous-catégories, dans le monde selon les données collectées par l'IA Index 2018<sup>6</sup>.

En septembre 2017, Cédric Villani, député de l'Essonne, a été chargé de conduire une mission sur la définition d'une politique nationale en IA. Le rapport produit – *IA for Humanity*<sup>7</sup> – a été présenté en mars 2018 au Président de la République. Il aborde successivement différentes facettes de l'IA : politique économique, recherche, emploi, éthique, cohésion sociale. Cinq annexes identifient des domaines d'intérêt particulier : éducation, santé, agriculture, transport, défense et sécurité.

Ce rapport constitue le socle de la stratégie nationale en IA dont le volet « moyens de la recherche », chiffré à 1 milliard d'euros, a été dévoilé le 28 novembre 2018 (665 millions d'euros de dépenses du budget de l'État jusqu'en 2022, auxquels s'ajoutent des investissements du secteur privé). Six grandes orientations ont été prévues pour : « faire en sorte, avec l'Allemagne et l'Union européenne, que la France soit capable de rivaliser avec la Chine et les États-Unis ».

<sup>5</sup> <https://www.economie.gouv.fr/France-IA-intelligence-artificielle>

<sup>6</sup> <https://aiindex.org/>

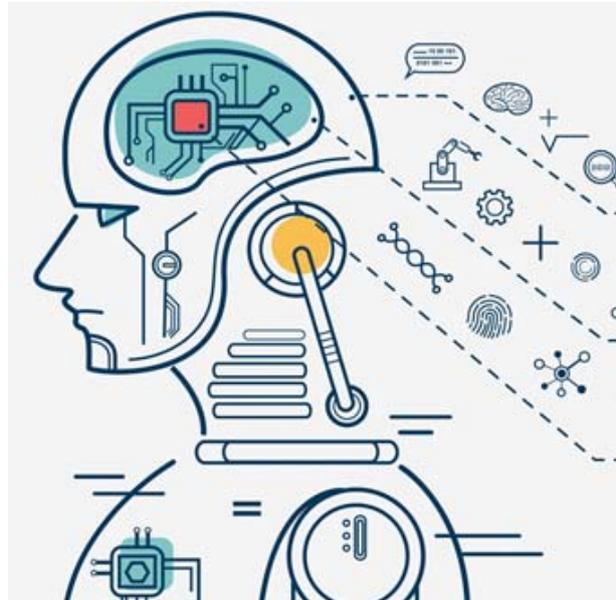
<sup>7</sup> <https://www.aiforhumanity.fr>

## 2. L'IA et ses domaines scientifiques

### 2.1 L'IA : de quoi parle-t-on ?

Définir ce qui est sous-tendu par le concept d'intelligence artificielle est toujours sujet à discussion voire à polémique. Il est toutefois assez clair que dans son acception la plus large il s'agit de l'ensemble des algorithmes (des technologies computationnelles) qui cherchent à reproduire, sous différentes formes, des fonctions du système nerveux central (CNS) qui, chez l'homme ou chez l'animal, servent à percevoir, apprendre, raisonner et agir.

L'IA vise d'une manière générale la conception de machines intelligentes intégrant des fonctions sensibles (informations proprio et extéroceptives), des fonctions motrices, les fonctions d'intégration cognitive des informations et des activités dans des schémas de planification. Précisons que la planification s'appuie généralement sur une représentation unifiée des fonctions de perception et d'action et exploite les fonctions du raisonnement et l'interaction (sociale notamment, comme le langage).



Les domaines de recherche en IA s'alignent traditionnellement sur les fonctions du système nerveux central. Une liste typique des sous-domaines de l'IA pourrait être : vision par ordinateur, traitement du langage naturel, robotique, interactions homme-système, recherche opérationnelle et planification, systèmes multi-agents, analyse des médias sociaux, représentation des connaissances, raisonnement et décision.

### 2.2 L'IA : un peu d'histoire ...

La mise sous forme d'algorithmes pour construire des machines intelligentes capables de rivaliser avec l'homme, dans différents registres réputés requérir de l'intelligence, a été imaginée dès les années 50 par Alan Turing<sup>8</sup> dans le cadre du paradigme du jeu d'échecs.

Dès cette époque, l'accent est mis sur le besoin de l'apprentissage dans les heuristiques de résolution de problèmes. Différentes notions furent avancées pour définir ce domaine émergent – cybernétique, théorie des automates, traitement complexe de l'information, etc. – traduisant la diversité des idées sur ce que seraient ces machines. Les considérations sur la capacité des calculateurs nécessaires au déploiement d'une intelligence artificielle ne sont pas absentes du débat à cette époque. À défaut de disposer d'un ordinateur puissant, Alan Turing s'est contenté de le simuler, le calcul de chaque coup prenant environ une demi-heure !

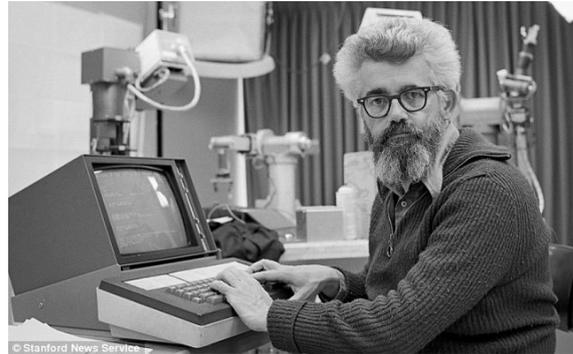
Le terme même d'intelligence artificielle a été proposé en 1955 par John McCarthy, alors jeune professeur de mathématiques au *Dartmouth College* (USA). Il avait décidé d'organiser un groupe de travail pour clarifier et développer des idées sur « les machines à penser ». L'expression « intelligence artificielle » a été choisie pour ce nouveau champ en partie pour sa neutralité, évitant de se concentrer sur la théorie étroite des automates, et évitant la cybernétique qui était fortement axée sur l'analogie avec l'humain.

<sup>8</sup> <https://www.csee.umbc.edu/courses/471/papers/turing.pdf>

### – Les pères de l'IA symbolique...

Dans les années 60, avec les progrès sur le raisonnement non monotone, les méthodes formelles connaissent un fort développement.

John McCarthy a été un contributeur majeur à la logique formelle pour le raisonnement en IA et est à l'origine de beaucoup d'idées clés en AI. Il cherche à « imiter » le raisonnement du cerveau en manipulant des symboles à l'aide de règles par une logique mathématique formelle du premier ordre, pour réaliser des tâches élémentaires « cognitives », introduisant le partage du temps, le raisonnement par le bon sens, l'algorithme d'élagage alpha-bêta, etc.

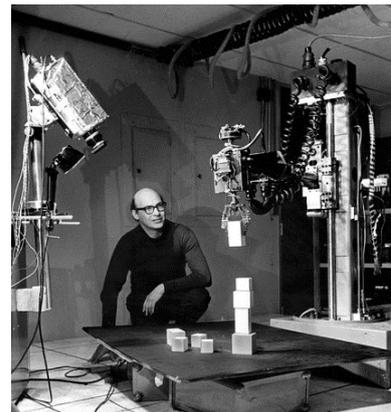


John McCarthy est l'un des pionniers de l'IA dite symbolique qui s'opposait alors à la vision connexionniste émergente fondée sur les principes des réseaux neuro-mimétiques dotés d'une capacité d'apprentissage avec le célèbre *perceptron* de Rosenblatt (1962), un algorithme de régression linéaire basé sur la méthode de « rétro-propagation du gradient de l'erreur ». Il est par ailleurs l'inventeur du langage de programmation déclarative et fonctionnelle LISP (*LISt Processing*) qui fut le premier langage de cette intelligence artificielle.

### – ...du connexionnisme

Marvin Minsky, qui cofonda en 1959 le groupe d'intelligence artificielle du *Massachusetts Institute of Technology* (MIT) avec John McCarty, a inspiré une approche qui considère le système cognitif comme une société de micro-agents susceptibles de résoudre des problèmes localement – une approche hiérarchique, modulaire et distribuée (bio-inspirée) favorisant l'émergence de comportements. Cette approche comportementale a été reprise dans les années 90 par Rodney Brooks aujourd'hui professeur émérite au MIT. Pour les connexionnistes, il faut aller au-delà des opérations symboliques, vers ce qu'ils appellent le niveau sub-symbolique.

Ce concept d'architecture hiérarchique a été utilisé pour la première fois en vision par ordinateur dans laquelle différentes sources de connaissances interagissaient les unes avec les autres, de manière opportuniste, pour résoudre le problème d'interprétation des images. Marvin Minsky conçoit en 1951 la première machine à réseau neuronal stochastique, le SNARC, considéré par beaucoup comme la première machine d'auto-apprentissage artificiel.



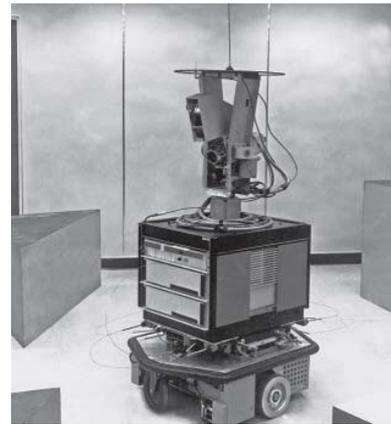
### – ...de la robotique

McCarthy initie également, comme professeur à l'Université de Stanford, la vision robotique (vision active) à travers le projet main-œil<sup>9</sup>, un système dans lequel une caméra (dont les images étaient traitées par ordinateur PDP-10) pouvait voir de vrais blocs 3D et commander un bras robotique pour réaliser des exercices d'empilement et d'arrangement.

<sup>9</sup> <https://www.ijcai.org/Proceedings/69/Papers/046A.pdf>

En janvier 1965, Charles A. Rosen, lance le projet Shakey au Stanford Research Institute (SRI International), certainement l'un des premiers projets de robotique mobile « autonome » utilisant une reconnaissance de l'environnement.

Ce projet expérimental cherchait à intégrer les sous-champs de l'intelligence artificielle et voulait intégrer dans un système de représentation et de raisonnement : la planification, l'apprentissage automatique, la vision par ordinateur, la compréhension du langage naturel. L'objectif était que Shakey puisse planifier et réaliser une séquence de déplacements pour atteindre un lieu donné dans un environnement non connu a priori de manière autonome.



#### – ...de la planification

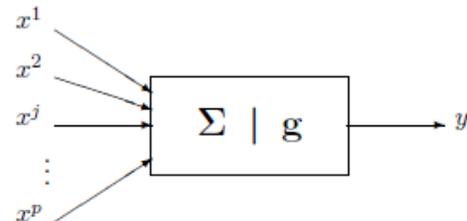
Le langage STRIPS<sup>10</sup> (*Stanford Research Institute Problem Solver*, de Richard Fikes et Nils Nilsson 1971, tous deux membres du projet Shakey) fût l'un des premiers langages pour modéliser et résoudre des problèmes de planification. STRIPS est né de la combinaison de deux grandes idées : la première était la stratégie de planification basée sur un moyen d'analyse pour atteindre le but ; la seconde d'utiliser le calcul des prédicats (logique du premier ordre) pour décrire de manière expressive l'état du monde. Peu de temps après la conception de STRIPS, l'équipe de Stanford a trouvé une façon de généraliser un plan STRIPS en remplaçant les constantes dans le plan par des variables. Ils ont également inventé une donnée structurée appelée « table de triangle » qui représente les dépendances internes d'un plan généralisé.

#### – L'émergence du connexionnisme et de l'apprentissage artificiel

Le connexionnisme désigne une famille de modèles formels développés pour la description et la simulation des capacités d'apprentissage, de reconnaissance et d'évocation de mémoires artificielles ou naturelles. Ces mémoires sont décrites comme des réseaux (souvent de grande taille) d'unités de traitement en interconnexion totale ou partielle.

Le connexionnisme est particulièrement approprié à l'étude du fonctionnement cérébral et des activités cognitives. Il est fondé sur les principes des réseaux neuro-mimétiques dotés d'une capacité d'apprentissage de tâches cognitives « de base ».

Un réseau neuronal est l'association, en un graphe de neurones formels qui de manière élémentaire se caractérisent par un état interne, des signaux d'entrée et une fonction d'activation qui opère une transformation d'une combinaison affine des signaux d'entrée (fonction seuil, linéaire, sigmoïde, stochastique, etc.). L'un des principaux intérêts des réseaux connexionnistes réside dans leurs capacités d'apprentissage.



Les unités de sortie d'un réseau dépendent non seulement des valeurs des unités d'entrée, mais aussi des valeurs des poids des connexions et des seuils des unités. On peut modifier la correspondance entre entrées et sorties d'un réseau en changeant ces poids et ces seuils. Le processus d'apprentissage consiste à effectuer de telles modifications à partir d'une suite de données construite par un tirage aléatoire avec remise dans un ensemble d'exemples. Ce processus d'entraînement peut être vu comme la résolution d'un problème d'optimisation par des méthodes de descente de gradient, du gradient conjugué, etc.

Les principaux réseaux se distinguent par leur architecture – l'organisation du graphe en couches possiblement multiples –, par la présence ou non de boucles de rétroaction dans le réseau, par le nombre ou le type de neurones, par leurs fonctions de transition ou d'activation.

<sup>10</sup> "STRIPS, A New Approach to the Application of Theorem Proving to Problem Solving," R. E. Fikes and N. J. Nilsson, in *Proceedings of the Second International Joint Conference on Artificial Intelligence (IJCAI)*, 1971.

La structure la plus simple, celle du *perceptron* multicouche dite statique (*MLP : Multi-Layer Perceptron*), mise au point en 1986, ne présente pas de boucle de rétroaction. Des réseaux dynamiques, avec boucles de rétroaction, des réseaux profonds, des réseaux récurrents à mémoire court-terme et long terme, des cartes de Kohonen ou cartes auto-organisatrices, qui verront le jour plus tard, offrent un potentiel bien plus important pour la modélisation de systèmes non linéaires.

L'histoire des réseaux de neurones artificiels remonte aux années 1950 et aux efforts de psychologues comme Franck Rosenblatt (1928-1971), accomplis dans les laboratoires aéronautiques de Cornell pour comprendre le cerveau humain. Initialement, ils sont conçus dans le but de modéliser mathématiquement le traitement de l'information par les réseaux de neurones biologiques du cortex des mammifères.

Le *perceptron* trouve ses origines dans des travaux de deux autres psychologues nord-américains – Warren McCulloch (1898-1969) et Donald Hebb (1904-1985). Le premier publie, avec le logicien Walter Pitts (1923-1969), en 1943, un article dans lequel il propose des neurones artificiels inspirés de leurs équivalents biologiques, en les dotant de propriétés mathématiques. Le second, en 1949, fournit les règles pour que ces neurones formels apprennent, comme le cerveau apprend par essais et erreurs.



L'enthousiasme créé autour des premiers réseaux de neurones artificiels retombe cependant assez rapidement. En 1969, le livre *Perceptrons*, par Seymour Papert et Marvin Minsky (réédition *MIT Press*, 2017) critique le *perceptron* (approximant des fonctions linéaires) du fait qu'il ne peut résoudre que des problèmes simples.

#### – Un regain d'intérêt pour les systèmes experts dans les années 80

Apparu dans le milieu des années 60, les systèmes experts, contrairement aux techniques connexionnistes, sont des systèmes à base de connaissances. Ils prennent leur essor dans les années 1970 et sont à l'origine du regain d'intérêt pour cette IA formelle dans les années 80.

Les systèmes experts sont des programmes qui utilisent des stratégies heuristiques, préalablement acquises à travers des entretiens avec des experts puis converties en règles d'action. La résolution des problèmes se fait par des méthodes dites déductives.

Le raisonnement effectué par un système expert peut être rendu compréhensible. Il fournit en effet l'information sur l'état de la résolution du problème et des explications sur les choix et les décisions. Le moteur d'inférence est le cœur du système expert. Ce dernier applique une stratégie de résolution en utilisant les connaissances, ceci pour en dériver de nouvelles informations.

Il existe de nombreux types de moteurs capables de traiter différentes formes de règles logiques pour déduire de nouveaux faits à partir de la base de connaissances. Les plus simples des systèmes experts s'appuient sur la logique des propositions « *logique d'ordre 0* ». D'autres systèmes s'appuient sur la logique des prédicats du premier ordre (dite aussi « *logique d'ordre 1* ») comme le langage PROLOG ou encore sur des inférences statistiques, probabilistes (ou possibilistes), etc.

– Après le marasme de l'IA des années 90, tout s'accélère



Durant les années 1990, l'IA en tant que domaine académique se trouvait indiscutablement dans un certain marasme. Bien que les algorithmes aient régulièrement progressé, les freins au progrès résidaient principalement du côté du matériel qui limitait la mise en œuvre des algorithmes.

En 1997, le programme Deep Blue d'IBM battait aux échecs le champion Gary Kasparov – un tournant dans l'histoire de l'informatique – sur la base de la « force brute » (600.000 parties jouées mémorisées dont celles de Kasparov et 256 microprocesseurs pour analyser 200 millions de positions par seconde, soit jusqu'à 100 milliards de situations à chaque coup).

Vingt ans plus tard, en 2016, Google AlphaGo bat le champion du monde de jeu de go Lee Sedol en utilisant la technologie des réseaux de neurones profonds (*DNN : Deep Neural Network*) à base de neurones convolutionnels et de parcours de graphes. AlphaZero surpasse aujourd'hui largement, par auto-apprentissage, le niveau de tous les joueurs humains aux échecs, au go, etc. et AlphaFold réalise des estimations en biologie structurale 3D jamais réalisées par les scientifiques.

Les techniques de base des réseaux neuronaux profonds, les réseaux de neurones convolutionnels et la rétro-propagation, étaient connues dès les années 1980. Cependant, les réseaux profonds n'ont été véritablement exploités en IA qu'à partir de 2010. Ce retard de plus de 20 ans dans l'application des réseaux profonds aux principaux problèmes d'IA est dû à l'absence de deux éléments clés : les données pour l'apprentissage et le matériel insuffisamment puissant.

Pour la classification d'image, d'importants ensembles de données d'apprentissage (par exemple ImageNet) étaient disponibles en 2005, mais ce ne fut qu'en 2012 que les réseaux profonds ont remporté le concours de reconnaissance visuelle ImageNet grâce à leur mise en œuvre sur des GPU.

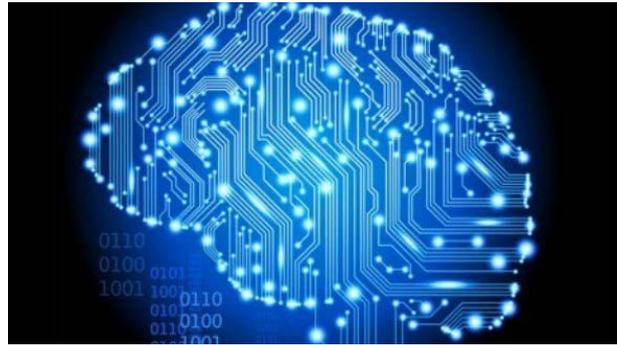
Avec la disponibilité des GPU, il est devenu possible d'entraîner de grands réseaux sur des ensembles de données de grande taille. Aujourd'hui encore, la taille d'un réseau et la taille de l'ensemble de données sur lequel il est formé restent limitées par le matériel disponible.

L'impact des réseaux profonds sur plusieurs domaines de l'IA a été très spectaculaire au cours des cinq dernières années. En utilisant les réseaux profonds, entre 2011 et 2015, le taux d'erreur pour le sous-titrage des images par ordinateur est tombé de 25% à environ 3%, mieux que le chiffre accepté pour la performance humaine qui est d'environ 5%. Les réseaux profonds ont dépassé les performances humaines dans de nombreuses tâches, y compris la reconnaissance faciale, la reconnaissance d'objets et surtout la compréhension du langage naturel dont le taux d'erreur de reconnaissance est tombé de 23% à 8% entre 2013 et 2015.

– L'influence des neurosciences computationnelles sur l'IA

Les neurosciences computationnelles visent à développer des modèles de calcul pour mieux comprendre les relations complexes entre la structure générale du système nerveux et ses grandes fonctions (perception, représentation, apprentissage, mémoire, raisonnement, planification).

Outre une meilleure connaissance de la cognition et de ses dysfonctionnements, les progrès réalisés en matière de modélisation permettent aujourd'hui de proposer des nouveaux modèles connexionnistes de traitement de l'information.

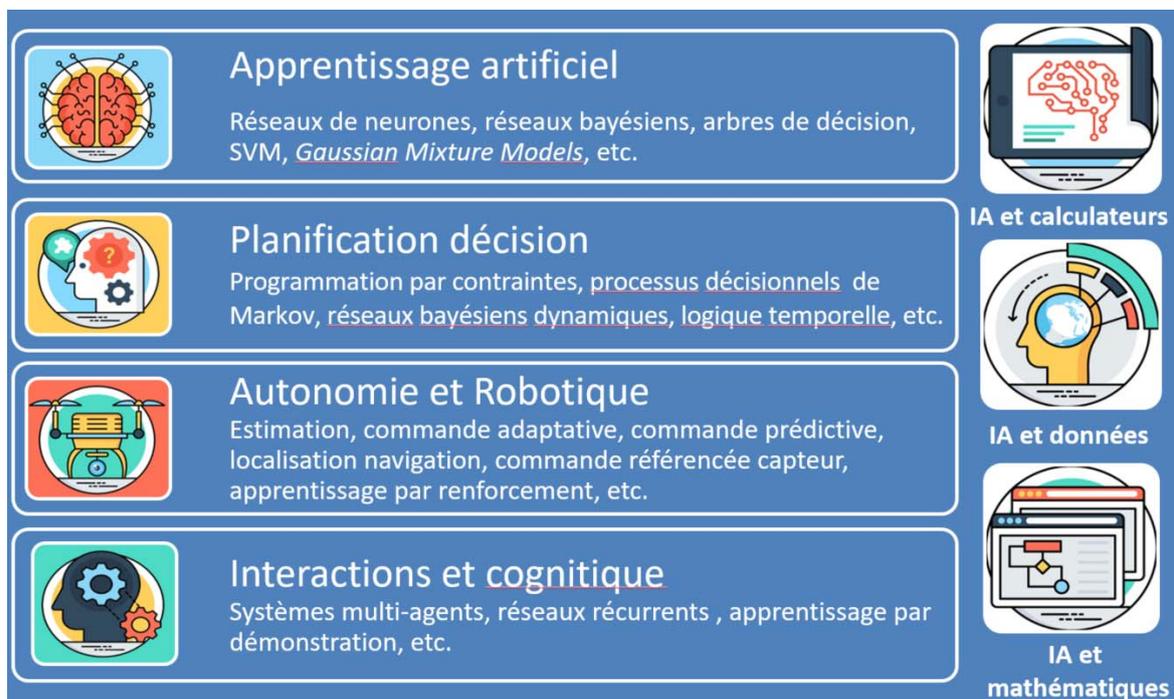


Les modèles connexionnistes connaissent aujourd'hui une influence considérable des travaux de recherche en neurosciences computationnelles et plus largement en sciences cognitives :

- ils ont largement contribué à rapprocher l'étude de l'architecture de calculateurs massivement parallèles de celle du software « cognitif ». Une nouvelle génération de puces dites « neuro-morphiques » s'appuyant sur une électronique mixte, analogique et digitale, a été proposée. Ces nouvelles architectures s'inspirent du fonctionnement dynamique du cerveau (les dynamiques non-linéaires à retard notamment) et de sa biophysique interne.
- ils ont replacé le concept de mémoire au centre des débats théoriques, que les courants théoriques dominants du cognitivisme avaient eu tendance à sous-estimer, voire à éliminer.
- ils ont permis l'élaboration de formalismes qui permettent non seulement de modéliser, mais aussi de simuler le fonctionnement de mémoires, naturelles ou artificielles, sur des calculateurs.

## 2.3 L'IA, ses techniques et ses applications aujourd'hui

Les algorithmes dits d'intelligence artificielle touchent aujourd'hui un très grand nombre de fonctions « cognitives » et « exécutives » pour des systèmes très variés et avec différents niveaux d'intégration. Ils ouvrent des champs de recherche dans des domaines comme ceux de la représentation des connaissances, du raisonnement en présence d'informations incomplètes et/ou incertaines, de l'apprentissage machine, de la robotique et des interactions homme-système.

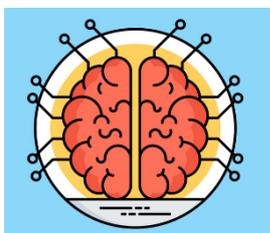


On reviendra rapidement dans ce qui suit sur ces technologies logicielles dont une forme de classification est proposée ci-dessus (celle que nous utiliserons).

On s'arrêtera en particulier sur les techniques qui font l'objet de développements à l'ONERA et pour lesquelles sont organisés des processus de capitalisation sous la forme d'atelier logiciel ou de bibliothèques logicielles auxquels nous ferons référence dans le domaine concerné.

Les questions des données, de l'évolution des calculateurs pour l'IA, ainsi que de l'apport des mathématiques appliquées seront évoqués.

### 2.3.1 Apprentissage artificiel



*L'apprentissage est un « processus de modification des connaissances ou de modification du comportement au cours des interactions d'un système avec son environnement ». Chez l'homme il mobilise différentes structures du cortex et de l'hippocampe. Le fait d'apprendre induit des modifications physiques, par la plasticité cérébrale, de la structure des réseaux neuronaux du système nerveux central (CNS) et conduit par conséquent à modifier, adapter notre façon de percevoir, penser, planifier et exécuter. Apprentissage et mémoire sont souvent confondus. L'apprentissage concerne toute modification du comportement ou des activités psychologiques attribuables à l'expérience du sujet tandis que la mémoire est une fonction plus restrictive qui inclut l'emploi d'un apprentissage.*

L'apprentissage « artificiel », ou « machine » ou encore « automatique » fait l'objet d'un ensemble de techniques algorithmiques visant à élaborer ces fonctions d'acquisition de connaissances sous une forme artificielle et à différents niveaux d'intégration. L'apprentissage machine constitue aujourd'hui une discipline scientifique

à part entière, à la frontière entre les mathématiques et l'informatique, visant d'une manière générale « la construction de règles d'inférence et de décision pour le traitement automatique des informations ».

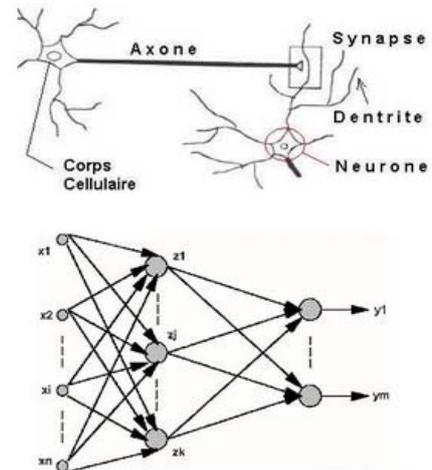
Les processus d'acquisition et de structuration des informations disponibles dépendent de la nature même des interactions du système avec l'environnement. Les informations et les règles qui les structurent peuvent être fournies directement par des experts sous des formes plus moins précises. On parle alors d'apprentissage symbolique (systèmes experts) consistant à déduire des données un système à base de règles (diagnostic, arbres de décision...). Il s'agit d'une forme dite déductive de l'apprentissage.

L'explosion combinatoire du traitement des problèmes réels a rendu nécessaire la mise en œuvre de « prétraitements statistiques ». L'intégration de ces prétraitements dans le système d'apprentissage a conduit à considérer que l'apprentissage automatique relève d'une méthodologie statistique. Dans bon nombre de problèmes, les connaissances restent à structurer, ou sont difficiles à structurer et reposent simplement sur des données formant un échantillon d'exemples (observations préalables) représentatifs de la relation recherchée (fonction aléatoire) entre les données d'entrée et celles de sortie (une base d'apprentissage).

Les fondements même du processus d'apprentissage reposent alors sur un traitement statistique des données. On parle alors d'apprentissage inductif ou mixte inductif-déductif pour les réseaux profonds. Ce traitement va dépendre des données mais aussi du problème à traiter. Plusieurs techniques algorithmiques sont exploitées pour construire un modèle « d'approximation » à partir d'apprentissage déductif exploitant des données sans une connaissance explicite du domaine dont elles sont issues. Les plus utilisées sont les réseaux de neurones (RN) formels et les arbres de décision.

Les réseaux de neurones connaissent un succès croissant face aux autres méthodes de traitement statistique en raison de leur puissance, leur polyvalence et simplicité d'utilisation. Ils ont une remarquable faculté à donner un sens, extraire des règles et des tendances à partir de données compliquées, bruitées et imprécises. Ils peuvent être vus comme des architectures connexionnistes bio-inspirées sans toutefois prétendre à une analogie avec les neurones biologiques.

L'architecture donnée à ces RN (type de réseau et fonction de transition et fonction d'activation) va dépendre du domaine (données d'entrée) et de l'objectif visé (la tâche d'apprentissage) qui est généralement un problème de classification (variables catégorielles) ou de régression (prédire une quantité).



#### – Apprentissage supervisé

Bon nombre d'applications d'apprentissage machine aujourd'hui touchent à la classification des images, la détection des objets, la reconnaissance de la parole et la compréhension du langage naturel, etc. Ces applications utilisent des techniques dites d'apprentissage supervisé. Lors de la phase d'apprentissage, on exploite une base d'apprentissage qui est formé d'un ensemble de données annotées (statiques ou dynamiques). Ces données contiennent des informations (attributs) sur les  $n$  variables explicatives  $x_i$  et sur la variable expliquée  $y$  associée. Ces « observations » (des couples  $(x_{1k}, x_{2k}, \dots, x_{nk}), y_k$ ) sont utilisées par un algorithme d'apprentissage pour rechercher le modèle de la fonction entrée/sortie  $y=f(x)$  (les poids) qui permettra de prédire la valeur de la sortie  $y$  (une classe par exemple) à partir d'entrées non présentes dans l'ensemble d'apprentissage. Le vecteur d'entrée  $x$  est généralement un vecteur de données aléatoires régies par une distribution marginale de probabilité. L'apprentissage consiste alors à ajuster les paramètres du réseau pour une famille de fonctions. Le modèle prédictif créé constitue une forme de modélisation décisionnelle à partir de données.

#### – Apprentissage non supervisé

Dans l'apprentissage statistique non supervisé (aussi appelé « analyse de données » en statistique ou « fouille de données »), on dispose de données qui ne sont pas annotées. L'objectif d'un algorithme d'apprentissage non supervisé est d'organiser les données pour en décrire la structure cachée. Cela peut signifier un regroupement en classes homogènes ou la recherche de différentes formes de visualisation des données complexes afin d'en simplifier l'affichage ou de les visualiser plus efficacement (*visual analytics*).

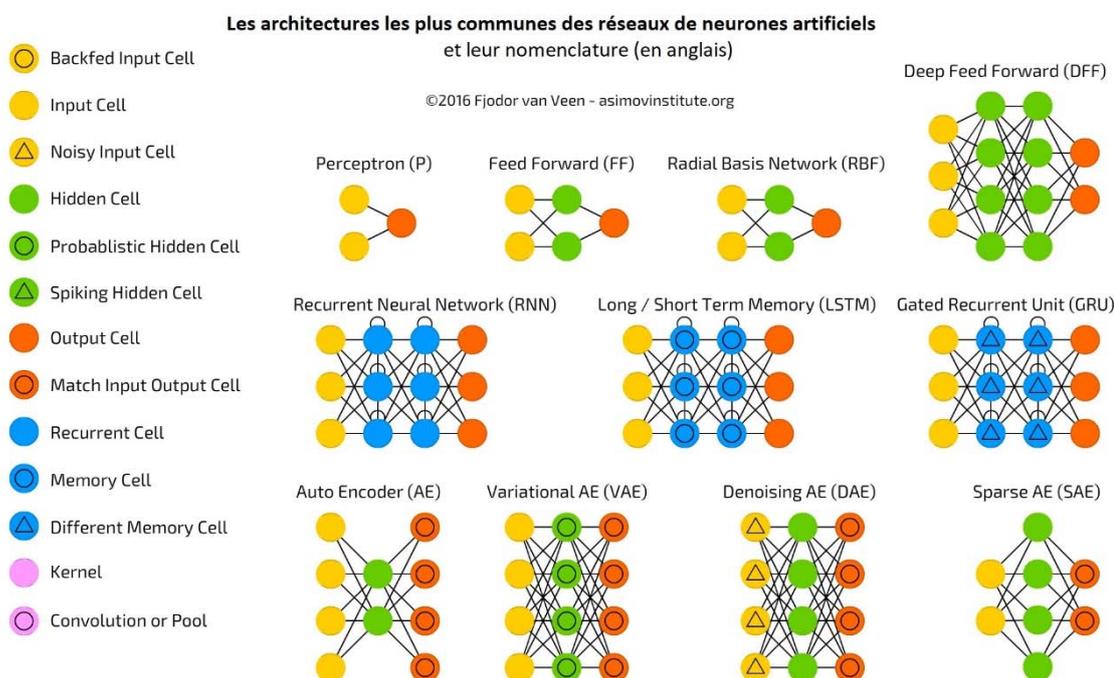
## – Apprentissage semi-supervisé

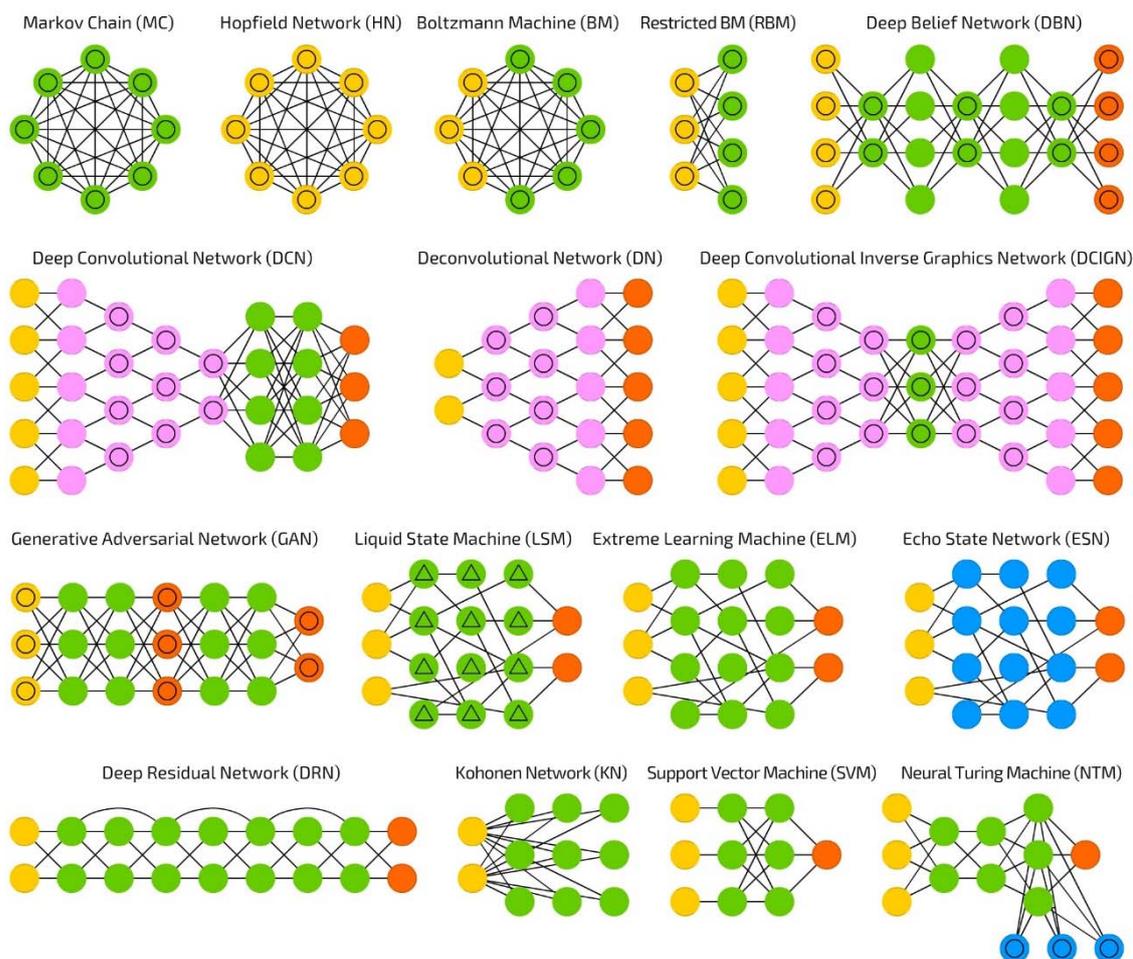
Le principe d'apprentissage semi-supervisé consiste à exploiter à la fois des données étiquetées et non étiquetées. Ce principe permet de pallier le manque de données annotées mais aussi d'améliorer la capacité des classificateurs, en accélérant la vitesse de convergence par de l'auto-apprentissage ou sa robustesse par un apprentissage actif. Pour réaliser un auto-apprentissage, on entraîne le classificateur sur des données étiquetées puis on utilise ce dernier pour étiqueter les données incomplètes, cela sous une forme itérative. L'apprentissage actif consiste à introduire un opérateur dans ce processus d'apprentissage hiérarchique pour annoter les données sélectionnées comme apportant le plus d'informations ou inversement celles dont on est le moins sûr en s'appuyant sur des estimateurs d'erreur de généralisation. L'idée de l'apprentissage semi-supervisé est d'une manière générale d'adapter le modèle de classification à la structure du problème.

## – Apprentissage par renforcement

L'apprentissage par renforcement diffère fondamentalement des apprentissages supervisé et non supervisé. Il correspond au cas où l'algorithme apprend de manière itérative un comportement qu'il optimise à partir d'observations. L'action de l'algorithme sur l'environnement produit en effet une valeur de retour qui guide l'algorithme d'apprentissage. Dans l'apprentissage par renforcement, l'algorithme choisit une action en réponse à chaque point de données. L'algorithme d'apprentissage reçoit en retour un signal de récompense (*reward*) qui indique la qualité de la décision. Il s'agit d'un problème de décision séquentiel dans lequel l'objectif est la récompense totale à long terme. On cherche une politique optimale qui maximise la récompense totale, les transitions entre états définissant les récompenses élémentaires. L'apprentissage par renforcement est un cadre pour la résolution de problèmes dans des espaces discrets et continus.

Les architectures des réseaux connexionnistes utilisées aujourd'hui dépassent très largement les limites des réseaux classiques à couches multiples de neurones. Les architectures les plus communes sont regroupées dans la table ci-dessous.





Dans ce qui suit, plusieurs types d'algorithmes d'apprentissage exploités dans les travaux en IA à l'ONERA sont décrits dans leurs principes.

#### – Réseaux de neurones convolutifs

Les réseaux de neurones profonds (*DNN : Deep Neural Network*) combinés à des données de grande taille donnent l'apprentissage profond (*DL : Deep Learning*). Les DNN permettent de modéliser avec un haut niveau d'abstraction des données grâce à des architectures à grand nombre de couches articulant différentes transformations non-linéaires. La plupart de ces réseaux exploitent aujourd'hui les réseaux de neurones convolutifs (ou réseaux de neurones convolutifs) (*CNN : Convolutional Neural Network*). Ces derniers sont généralement constitués de quatre types de couches : la couche convolutionnelle, la couche de *pooling*, la couche de correction ReLU et la couche *fully-connected*. La couche convolutionnelle est constituée d'un noyau de convolution qui est appliqué sur l'ensemble des données d'entrée par balayage. La profondeur d'un réseau de neurones convolutif compte le nombre de couches de convolution.

Les DNN peuvent être entraînés pour classer les images, détecter des objets, identifier des personnes à partir de visages, générer du texte à partir de la parole, traduire des langues naturelles, et beaucoup d'autres tâches. Pour beaucoup de ces tâches, les DNN ont atteint des performances dépassant ce que les humains font habituellement.

#### – Réseaux récurrents

Les réseaux récurrents (*RNN : Recurrent Neural Networks*) sont des réseaux de neurones dans lesquels l'information peut se propager dans les deux sens, y compris des couches profondes aux premières couches. Ces réseaux possèdent des connexions récurrentes au sens où elles conservent des informations en mémoire : ils

peuvent prendre en compte à un instant un certain nombre d'états passés. Pour cette raison, les RNN sont particulièrement adaptés aux applications faisant intervenir le contexte, et plus particulièrement au traitement des séquences temporelles, c'est-à-dire quand les données forment une suite et ne sont pas indépendantes les unes des autres. Néanmoins, pour les applications faisant intervenir de longs écarts temporels (typiquement la classification de séquences vidéo), cette « mémoire à court-terme » n'est pas suffisante. En effet, les RNN « classiques » (réseaux de neurones récurrents simples) ne sont capables de mémoriser que le passé dit proche, et commencent à oublier au bout d'une cinquantaine d'itérations environ. Ce transfert d'information à double sens rend leur entraînement beaucoup plus compliqué, et ce n'est que récemment que des méthodes efficaces ont été mises au point comme les LSTM (*Long Short Term Memory*). Ces réseaux à large « mémoire court-terme » ont notamment révolutionné la reconnaissance du langage naturel ou de textes.

#### – Réseaux auto-encodeurs

Un auto-encodeur est un type de réseau neuronal artificiel qui permet de construire une nouvelle représentation d'un jeu de données compact qui présente moins de descripteurs, ce qui permet de réduire la dimensionnalité. L'objet d'un auto-encodeur est d'apprendre une représentation (encodage) « efficace », une représentation aussi proche que possible de son entrée d'origine, d'où son nom. L'architecture d'un auto-encodeur est constituée de deux parties : l'encodeur et le décodeur. L'encodeur est constitué par un ensemble de couches de neurones qui traitent les données afin de construire de nouvelles représentations dites « encodées ». À leur tour, les couches de neurones du décodeur reçoivent ces représentations et les traitent afin d'essayer de reconstruire les données de départ. Les différences entre les données reconstruites et les données initiales permettent de mesurer l'erreur commise par l'auto-encodeur. L'entraînement consiste à modifier les paramètres de l'auto-encodeur afin de réduire l'erreur de reconstruction mesurée sur les différents exemples du jeu de données. Un exemple typique d'usage est auto-encodeur est la reconstruction d'images dans lequel il va apprendre à identifier et à représenter les caractéristiques les plus significatives, les plus petits détails étant ignorés ou perdus et le bruit atténué.

#### – Réseaux antagonistes génératifs

La technologie des réseaux antagonistes génératifs (*GAN : Generative Adversarial Networks*) est une approche de programmation pour l'élaboration de modèles génératifs, c'est-à-dire capables de produire eux-mêmes des données. Les GAN reposent sur l'entraînement non supervisé de deux réseaux de neurones artificiels appelés « générateur » et « discriminateur ». Ces deux réseaux s'entraînent l'un l'autre dans le cadre d'une relation contradictoire : le générateur est en charge de créer des échantillons ; le discriminateur reçoit des échantillons provenant (i) du générateur et (ii) d'une base de données constituée d'échantillons réels et se charge d'en identifier la source.

Des algorithmes d'apprentissage sont par ailleurs à choisir pour optimiser la réponse du réseau pour une tâche donnée. Nous évoquons ci-dessous certaines des méthodes sur lesquelles reposent des algorithmes récemment proposés.

#### – Apprentissage de bout en bout

L'apprentissage de bout en bout consiste à former un système d'apprentissage travaillant directement sur des données brutes. Une même architecture de réseau réalise l'extraction automatique des caractéristiques (des informations pertinentes) et la prédiction (classification par exemple). Ces architectures d'apprentissage de bout en bout sont conçues à partir de modules, souvent organisés en couches et groupes de couches avec des rôles spécifiques. Des exemples marquants sont des séquences alternées de convolution et de regroupement des couches, des couches entièrement connectées, des encodeurs automatiques (couches à goulot d'étranglement) et des couches LSTM.

L'apprentissage de bout en bout bénéficie d'un degré élevé d'automatisation, mais il ne fonctionne toutefois pas de manière totalement autonome : la conception et la configuration d'un système d'apprentissage adapté à une situation donnée est nécessaire. Cette tâche nécessite un certain niveau d'expérience et de connaissances en machine.

### – Apprentissage par adaptation de domaine

L'hypothèse faite dans l'application des méthodes traditionnelles d'apprentissage automatique est que les données d'apprentissage et les données de test sont issues du même domaine, de sorte que l'espace des caractéristiques en entrée et la distribution des données soient les mêmes. Cependant, dans beaucoup de scénarios d'apprentissage, les données annotées nécessaires à l'apprentissage ne sont pas disponibles ou pas de façon simple. On peut chercher alors à réutiliser les connaissances acquises dans un domaine (une distribution de probabilités) voire de plusieurs domaines pour les transférer à une autre tâche ou un autre domaine partageant des similitudes pour améliorer un taux de classification par exemple. L'intérêt est notamment d'utiliser des réseaux pré-entraînés disponibles.

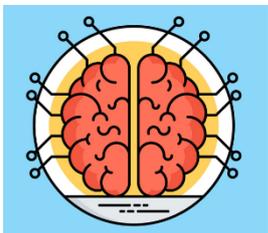
Des méthodes d'apprentissage par adaptation de domaine ont été proposées pour cela. Elles peuvent permettre aujourd'hui de considérer un apprentissage par transfert « homogène ». La majorité des solutions d'apprentissage par transfert homogène utilisent des stratégies qui consistent à corriger, entre les données source et cible, la différence de distribution marginale, la différence de distribution conditionnelle, ou à la fois les différences de distribution marginale et conditionnelle. Sinon, on parle d'apprentissage par transfert hétérogène qui vise l'alignement des espaces d'entrée des domaines source et cible, avec comme hypothèse que les distributions de domaines restent les mêmes.

### – Arbres de décision

Les arbres de décision forment une classe particulière d'algorithmes d'apprentissage se basant sur la représentation des choix sous la forme graphique d'un arbre avec les différentes décisions de classification placées dans les feuilles. C'est une technique d'apprentissage supervisé dans lequel on utilise un ensemble de données pour lesquelles on connaît la valeur de la variable-cible afin de construire de manière récursive l'arbre (données dites étiquetées). On génère pour cela une séquence hiérarchique de tests, aussi courte que possible, qui divise successivement l'ensemble des données d'apprentissage en sous-ensembles disjoints, afin que des sous-groupes de cas appartenant à la même classe soient rapidement détectés.

Un arbre de décision peut être utilisé pour représenter de manière explicite les décisions réalisées et les processus qui les amènent. Ils peuvent être utilisés à la classification ou à la prédiction. C'est la raison pour laquelle les méthodes d'arbres de décision sont particulièrement bien adaptées aux tâches de *data mining*, dans lesquelles nous ne disposons généralement que de connaissances a priori limitées, et où nous ne disposons ni de théories, ni de prévisions cohérentes quant aux variables qui sont corrélées, ni à la manière dont elles le sont. Dans ce type d'analyses de données, les méthodes d'arbres de décision permettent souvent de mettre en évidence des relations simples entre un nombre restreint de variables, alors que ces relations peuvent parfaitement être occultées par d'autres techniques analytiques. Un arbre de décision décrit les données mais pas les décisions elles-mêmes, l'arbre serait utilisé comme point de départ au processus de décision.

## L'apprentissage artificiel à l'ONERA



*L'apprentissage artificiel (ou automatique) est développé historiquement à l'ONERA principalement dans des usages classiques du traitement des images issues de caméras 2D et 3D comme celles de divers imageurs optroniques (lidar, stéréo, RGB-D, IR, etc.) et radar (SAR). Nous en donnons un exemple d'application ci-dessous avec l'analyse de scènes.*

*Les techniques de l'apprentissage machine diffusent aujourd'hui dans l'ensemble des domaines scientifiques couverts par l'ONERA. Elles sont également exploitées pour des mesures de champs en soufflerie ou de simulation, des données de capteurs biométriques (ECG, EMG, oculomètres), ou encore des données de vol des aéronefs, d'essais de vibrations, etc. Elles font l'objet d'un processus de capitalisation organisé dans le cadre du DeepLab DELTA.*

## Le DeepLab DELTA



Le projet de recherche DELTA (*Deep learning for aerospace applications*) a été mis en place à l'ONERA en 2017 dans l'objectif d'étendre le champ d'utilisation des technologies d'apprentissage automatique à différents domaines de la recherche aérospatiale. Il intéresse notamment des applications en mécanique des fluides, mécanique des matériaux, électromagnétisme, optique et en robotique.

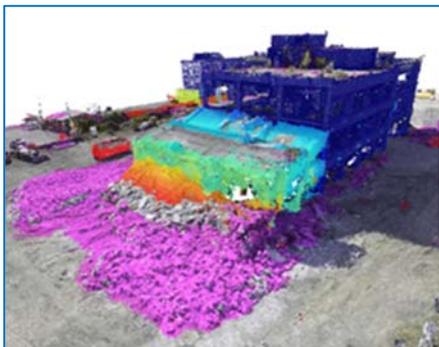
DELTA propose un environnement logiciel dédié d'une part à l'expérimentation des techniques d'apprentissage profond – réseaux profonds, récurrents, adverses, *Long Short-Term Memory* – et d'autre part à la génération de données adaptées et d'architectures à différents niveaux de complexité.

DELTA est développé en s'appuyant sur des *frameworks* : Tensorflow ([tensorflow.org](https://www.tensorflow.org)), PyTorch ([pytorch.org](https://pytorch.org)), Torch ([torch.ch](https://torch.ch)), Caffe ([caffe.berkeleyvision.org](https://caffe.berkeleyvision.org)).

Ce « DeepLab » est basé matériellement sur des GPU (cartes graphiques) et s'appuie sur un parc d'un millier de machines dont des serveurs avec des GPU Nvidia de nouvelle génération (Titan6, Titan Xp, GeForce GTX). Voir site GitHub : <https://delta-onera.github.io>

## L'analyse de scènes

L'un des grands enjeux dans l'exploitation des images (aériennes et satellitaires) est de comprendre des « scènes » en s'appuyant sur l'analyse de changement sur différentes échelles de temps pour, par exemple, suivre l'occupation du sursol, détecter des véhicules, identifier des déplacements, etc.

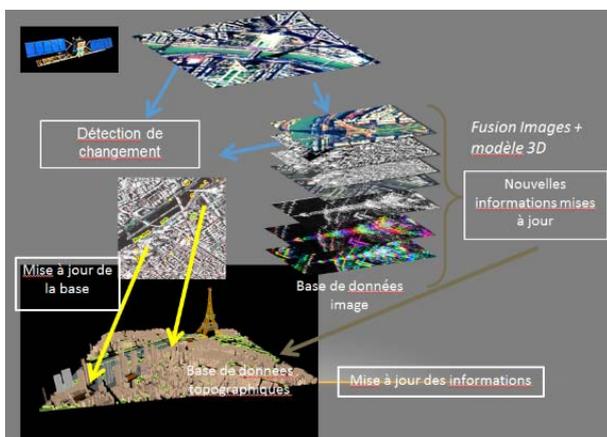


L'ONERA a développé et mise en œuvre de nouveaux outils de classification, en s'appuyant notamment sur les méthodes d'apprentissage adaptées au cadre des données de télé-détection. Ceci nous a conduit à rechercher des structures de réseau de neurones capables d'apprendre à la fois la classification et la segmentation, c'est-à-dire à réaliser une segmentation sémantique.

En particulier, l'architecture SegNet conçue sur le principe des réseaux de neurones entièrement convolutifs a été mise au point grâce à une architecture de type « encodeur/décodeur » qui parvient à replacer les représentations abstraites de haut niveau aux points de saillance de l'image originale. Nous avons également proposé des architectures de réseaux neuronaux convolutifs pour la détection de changements. Il s'agit d'extensions siamoises d'architectures entièrement convolutives. Ces réseaux sont capables d'apprendre à détecter des changements en utilisant des paires annotées en termes de changement, sans intervention humaine ni post-traitement. Une illustration de résultat est donnée sur la figure qui met en évidence des modifications sur un bâtiment suite à une catastrophe naturelle.

## Traitement *big data* des données

Les données géospatiales issues de capteurs SAR et optroniques permettent de constituer des représentations multi-échelles de l'environnement. Cette numérisation du monde et de sa sémantisation peuvent servir à l'analyse de la variabilité spatio-temporelle et la compréhension de scènes dynamiques au sens large, l'anticipation de son évolution prévisible, la détection d'événements imprévus ou anormaux.



Les traitements des données satellitaires réalisés par l'ONERA visent en particulier à répondre à différents domaines d'application :

- le suivi de l'évolution de l'occupation des sols
- la surveillance du trafic routier ou maritime
- la surveillance des déformations des sols et des déformations des structures 3D
- la classification des îlots de chaleur urbains, etc.

Les apports des réseaux de neurones en télédétection dans différentes approches (*end-to-end*, basée région, fusion d'images hétérogènes, segmentation sémantique) pour la classification et la détection ont fait l'objet d'investigations importantes en particulier pour parvenir à maîtriser la robustesse des stratégies d'apprentissage supervisé à l'aide de réseaux de neurones convolutifs (CNN) et entièrement convolutifs (FCN, *Fully-Convolutional Networks*).

Nous avons proposé récemment Berunda Net, une architecture de FCN multitâche, pour l'entraînement semi-supervisé avec un auto-encodage qui agit comme une régularisation de l'apprentissage et nous avons mis en évidence la capacité de ce modèle à bénéficier des images non-annotées, notamment pour les classes ou caractéristiques rares.

### 2.3.2 Planification et décision

*La décision est liée au raisonnement (pour les décisions raisonnées) ainsi que bien évidemment aux connaissances sur lesquelles elle s'appuie. Les prises de décision sont considérées comme des fonctions cognitives évoluées. Elles font intervenir le cortex préfrontal, une structure cérébrale parmi les plus développées et connue pour assurer les processus décisionnels.*



Les méthodes d'apprentissage peuvent servir à construire une prédiction en vue d'une décision. L'apprentissage statistique, qu'il soit réalisé dans une perspective de classification ou de régression, intègre des « formes » de décision (outils d'inférence statistique et de mesure de la qualité de l'apprentissage). La plupart des méthodes de classification issues de l'apprentissage numérique apportent toutefois rarement de notion de confiance dans la décision. Par ailleurs, les méthodes d'apprentissage numérique nécessitent des adaptations parfois complexes pour permettre un apprentissage prenant en compte des observations arrivant de manière continue. Qu'il s'agisse de problèmes de classification, de diagnostic, etc., l'état du « monde » est considéré comme statique. Dès lors où le monde évolue dans le temps, le problème de décision devient alors combinatoire.

## – Le problème général de la décision

Le problème de la décision est généralement posé en considérant un agent immergé dans un environnement. Ce dernier reçoit des observations  $E$  générées par cet environnement, appelées aussi perceptions. En utilisant ces perceptions, l'agent infère un état discret  $S$  probable de l'environnement. Il doit décider d'une action  $A$  à réaliser, qui va l'amener vers un état  $S_i(A;S)$  plus désirable. Dans un processus déterministe, la solution est un plan. Lorsque l'état du monde est partiellement observé (ou seulement estimé) par les capteurs, quand les actions produisent des incertitudes, le processus d'observation devient stochastique. On élabore alors une politique optimale basée sur la maximisation d'une utilité, considérant l'impossibilité de représenter explicitement toutes les contingences (combinatoire) pour produire un plan conditionnel.

Une solution s'établit sous forme d'une séquence d'actions permettant de passer d'un état initial vers un état final (« but ») c'est-à-dire un chemin dans l'espace des états (l'espace de recherche peut être représenté par un graphe orienté). Résoudre un problème de planification consiste à trouver un ou plusieurs chemins possibles. Une fonction de coût des chemins permet de sélectionner une solution préférée parmi l'ensemble des solutions.

On peut tenir compte directement de l'incertain et faire prendre par un agent « intelligent » les meilleures décisions possibles par l'utilisation de la théorie des probabilités, pour apprendre automatiquement, raisonner et prendre des décisions dans un domaine incertain à partir de perceptions. La théorie des probabilités aide à intégrer les incertitudes (le degré de croyance de l'agent) sur l'état du monde par des formalismes tels que ceux des réseaux bayésiens dynamiques, les chaînes de Markov (automates à états finis stochastiques : monde totalement observable), les modèles de Markov cachés, etc.

La décision la plus appropriée, pour qu'un agent « intelligent » atteigne un but, est celle qui maximise la fonction d'utilité, maximum qui peut être estimé sous la forme d'un maximum de vraisemblance ou à l'aide d'un classifieur naïf de Bayes.

## – Les processus décisionnels de Markov

Les processus décisionnels de Markov (PDM) et ceux partiellement observables (PDMP0) permettent de représenter et de résoudre des problèmes de décision séquentielle dans l'incertain. Les décisions sont fondées sur une estimation de l'état, prenant la forme d'une distribution de probabilité définie sur l'espace des états du système. Cette distribution qui forme l'état de croyance est mise à jour à chaque étape du processus. La mise à jour, effectuée à l'aide de la règle de Bayes, nécessite une connaissance parfaite de l'état de croyance initial de l'agent, ainsi que des distributions de probabilité de transition et d'observation du processus.

De nombreux algorithmes sont utilisables à la résolution de problèmes de planification : STRIP, recherche heuristique ( $A^*$ ), par résolution de contraintes (*PSP-Partial Satisfaction Planning*, *OSP-Over Subscription Planning*, *POP-Partial Order Planning*, etc.) et leurs versions stochastiques fondées sur la théorie des probabilités et le raisonnement bayésien.

Dès lors où les problèmes comportent de fortes incertitudes, ou qu'ils sont de grande taille ou indécidables, on procède par décision séquentielle. L'introduction de la séquentialité dans la décision permet de tenir compte des opportunités d'information future et introduit de la flexibilité dans la prise de décision. Le problème de la prise de décision séquentielle (*SDM-Shared Decision Making*) est toutefois plus complexe car les séquences de décision optimales ne sont généralement pas formées en sélectionnant la meilleure décision individuelle à chaque étape, les décisions qui composent la séquence optimale n'étant pas indépendantes.

## – La capacité d'apprentissage incrémental

La capacité d'apprentissage incrémental est essentielle pour prendre en compte l'évolution des modèles du monde et donc faire dépendre le modèle du temps. L'apprentissage incrémental est une réponse possible à ce problème et en particulier l'apprentissage bayésien qui consiste à calculer les probabilités *a posteriori* (après l'observation d'événements) à partir de vraisemblance des événements connaissant le phénomène qui est susceptible de les avoir générés et la probabilité a priori sur ces événements. L'apprentissage de réseaux bayésiens revient à rechercher parmi un ensemble de distributions, la distribution la plus proche possible, en un certain sens, de la distribution représentée par les données. En limitant l'ensemble de recherche, on peut éviter le problème de sur-apprentissage, qui revient dans ce cas à calquer exactement la distribution représentée par les exemples.

L'essentiel des algorithmes développés pour l'inférence et l'apprentissage dans les réseaux bayésiens et les outils disponibles pour mettre en œuvre ces algorithmes utilisent des variables discrètes. La machinerie des algorithmes d'inférence est essentiellement fondée sur une algèbre de tables de probabilités. De même, les algorithmes d'apprentissage modélisent en général les distributions de probabilité des paramètres contenus dans les tables du réseau, c'est-à-dire de probabilités discrètes. Tout un travail reste à faire pour pouvoir mélanger des variables discrètes et continues afin de considérer simultanément des variables logiques (comme celles issues d'une carte d'occupation d'un environnement) avec des variables continues (des trajectoires pour un mobile par exemple).

## Planification et décision à l'ONERA

*Dans le domaine de la décision, les travaux de l'ONERA visent en particulier la planification de missions d'engins autonomes (drones, robots terrestres et sous-marins, satellites) pour disposer à terme des systèmes capables de réaliser à haut niveau d'autonomie des missions complexes avec une forte résilience aux aléas et perturbations.*



*On se fonde pour cela principalement sur des modèles formels représentant les systèmes considérés, les évolutions possibles de l'environnement et les objectifs à atteindre. Ceux-ci sont capitalisés dans la bibliothèque InCell (voir ci-dessous).*

*On cherche d'une manière générale à produire diverses stratégies (des plans) permettant d'optimiser les actions réalisées par le système et rendre ce dernier résilient. Ces stratégies peuvent être réactives (les plans étant reconstruits – adaptés – en cas de déviation entre le modèle et la réalité) ou pro-actives (prenant la forme*

*de politiques de décision qui spécifient les actions à réaliser pour plusieurs scénarios envisageables) et sont combinées à différents niveaux hiérarchiques. Trois illustrations en sont données ci-dessous.*

@ONERA

## La bibliothèque InCell



L'ONERA a développé un ensemble de techniques de planification reposant sur le formalisme de description PDDL (*Planning Domain Definition Language*). Celles-ci reposent sur des méthodes déterministes ou probabilistes. Elles permettent la prise en compte de contraintes spatio-temporelles liées à l'évolution de l'environnement et à l'état des systèmes dans la réalisation des plans d'actions.

Celles-ci sont regroupées dans la bibliothèque InCell. Cette dernière s'appuie sur le cadre des CBLS (*Constraint-Based Local Search*) qui combine (1) des modèles formels définis par des variables de décision et par des contraintes et critères portant sur ces variables, et (2) des techniques de recherche locale qui appliquent une succession de modifications locales sur un plan courant.

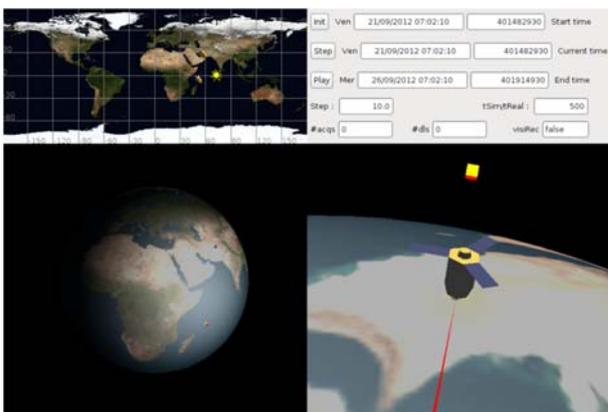
Un travail particulier été effectué concernant la gestion des aspects temporels (durée des tâches), travail qui a mené à la définition des TSTN (*Time-Dependent Simple Temporal Networks*). La bibliothèque InCell est étendue actuellement au contexte dynamique et incertain de la génération de plans robustes (réparation de plans et planification sur un horizon glissant).

## Planification des missions de systèmes robotiques autonomes



Une architecture décisionnelle, développée à l'ONERA, permet d'atteindre un haut niveau de résilience dans le déploiement de systèmes multi-robots autonomes. Celle-ci intègre planification, supervision de l'exécution et réparation de plans d'actions. Elle s'articule autour du planificateur hybride HiPOP (*Hierarchical Partial Order Planning*) qui organise les actions sous une forme hiérarchique. Ce planificateur gère les contraintes spatio-temporelles portant sur l'enchaînement des diverses activités (synchronisation, rendez-vous), des contraintes sur les ressources (mémoire, énergie...), et des objectifs d'optimisation multicritères. La structuration hiérarchique des plans est exploitée pour faciliter l'ajout dynamique d'actions et la réparation efficace des plans suite à un aléa. Elle permet de limiter la recherche à effectuer et de produire un plan proche du plan initial par ajout itératif d'actions, le nouveau plan contenant ainsi les actions passées.

## Observation spatiale



Plusieurs missions correspondant à des problèmes de planification en ligne à bord de satellites autonomes sont traitées à l'aide des techniques de planification développées à l'ONERA :

- la mission HotSpot, consistant à planifier les activités d'acquisition et de vidage de données vers des stations sol pour un satellite capable de détecter des « points chauds » à la surface du globe ;
- la mission Agile correspondant à la mission HotSpot étendue à des satellites agiles équipés d'actionneurs gyroscopiques et d'un système de détection en avant de la couverture nuageuse.

Dans ce cadre, les problèmes d'optimisation pour l'affectation et l'ordonnancement des ressources non partageables sous contraintes de précedence, de durées minimum dépendantes du temps, etc sont traités à l'aide de la librairie générique InCELL. Complémentairement, des algorithmes de recherche gloutonne et locale ont été développés pour produire des plans de vidage, compte-tenu du temps et des ressources de calcul limités disponibles à bord des satellites.

### 2.3.3 Autonomie et robotique



*Chez l'homme, les fonctions associant perception et action impliquent différentes régions du CNS (cortex moteur, cortex visuel, cortex auditif, cervelet, hippocampe, etc.) et mettent en œuvre différents schémas d'intégration. L'intégration multi-sensorielle permet l'élaboration de représentations internes de l'environnement spécifiques de l'appareil sensoriel considéré, adaptées à l'action. Cela peut par exemple s'illustrer par la spécialisation de neurones de l'hippocampe pour la représentation de la localisation, sur la base de données allocentriques (visuelles, tactiles, olfactives) et égocentriques (proprioception).*

*Les structures de contrôle moteur exploitent différents substrats neuronaux centraux dans l'apprentissage (cortex frontal, ganglions de la base, formation hippocampique, cervelet). Les fonctions emploient des architectures multi-boucles utilisant des circuits sous-corticaux rapides pour les actions réflexes. Au niveau cortical, les analyses sensorielles sont plus riches mais également plus lentes ; les aires associatives et frontales se chargent de la fusion des données, des changements de référentiels, de la construction et de la mémorisation de représentations internes en vue de leur exploitation pour une action plus délibérative, guidée vers un but.*

La recherche en robotique pose des questions de nature plus « intégrative » à l'intelligence artificielle. Comment construire une représentation unifiée de l'environnement à partir des données multi-sensorielles ? Comment exploiter cette représentation pour naviguer ? Comment apprendre différentes tâches en parallèle ? Comment planifier et exécuter des actions de manière autonome avec une grande capacité d'adaptation ? Ces interrogations (parmi d'autres) induisent des travaux sur les représentations de l'information perceptuelle pour la navigation notamment, sur des mécanismes d'apprentissage incrémentaux et stables à long terme pour les schémas de commande mais aussi sur les méthodes d'apprentissage pour la planification d'actions et les interactions avec des humains.

#### – La localisation et la navigation sémantique multimodale

L'autonomie de déplacement implique la capacité pour les systèmes à se localiser et à naviguer dans des environnements non connus *a priori* et souvent dynamiques. Ce type de fonction fait appel à des techniques dites de SLAM (*Simultaneous Localisation and Mapping*) dont le développement repose généralement sur une représentation géométrique de l'environnement extraite par des capteurs embarqués. Un objectif général à terme est de développer des approches sémantiques reposant sur des représentations hiérarchiques d'éléments fonctionnels de l'environnement utiles à la navigation – mur, porte, fenêtre, meuble, objets, etc.

La navigation « sémantique » suppose des développements sur les méthodes de reconnaissance de ces différents éléments qui pourront bénéficier des nombreux travaux en vision par ordinateur utilisant les réseaux de neurones convolutifs profonds pour la reconnaissance et la classification des images. Elle suggère également des développements sur des aspects plus spécifiques à la robotique tels que la fusion de données télémétrie/vision et l'exploration active de l'environnement (perception active).

#### – L'apprentissage en ligne

Les méthodes d'apprentissage incrémentales, qui sont bien adaptées à l'apprentissage en ligne, peuvent être également envisagées dans un cadre d'apprentissage interactif ou d'apprentissage actif (basé sur la motivation). La répartition spatiale adaptative des apprentissages pour focaliser des zones d'intérêt (détectées par exemple par catégorisation non supervisées des éléments stables et récurrents de l'environnement) ou apprendre à structurer progressivement l'environnement en combinant différentes représentations dans la mémoire à court et long terme sont également des questions pour la recherche en IA.

#### – L'apprentissage pour l'estimation de modèles

Les méthodes de régression non linéaire sont classiquement utilisées pour l'estimation de modèles de systèmes dynamiques ou de lois de commande pour des systèmes complexes, non linéaires et non stationnaires.

Dans le cadre de la modélisation et de la commande de processus, l'utilisation de réseaux de neurones (de type MLP) a tout d'abord été réalisée pour l'estimation des coefficients des correcteurs simples (type PID) pour la commande de systèmes présentant des non-linéarités (frottement, saturations, etc.) ou encore pour la détermination d'une structure optimale de régulation. Les capacités d'approximation de fonctions des réseaux de

neurones ont également été utilisées pour obtenir la dynamique inverse des systèmes non linéaires. L'utilisation du modèle neuronal inverse ainsi obtenu constitue une structure de commande neuronale directe d'un système non linéaire. Des structures hybrides, combinant en parallèle un contrôleur simple et un contrôleur neuronal, ont été proposées pour assurer le rejet des perturbations et le suivi par la sortie d'un signal de référence en présence d'erreurs de modèle.

Ces capacités d'apprentissage rendent les tâches plus robustes. Elles peuvent permettre d'assurer le maintien de performances de manière autonome mais aussi de synthétiser des contrôleurs, en ligne, par des résultats d'expériences. Les progrès dans la compréhension des méthodes d'apprentissage profond et de leur relation aux méthodes de régression plus classiques permettent d'envisager l'application de ces méthodes à l'identification en ligne pour la commande adaptative comme celle par modèle de référence. Le contrôle adaptatif par modèle de référence (MRAC) est un exemple de ce qui est utilisé dans de nombreuses applications aéronautiques pour améliorer les performances du système en présence d'incertitudes ou quand sa dynamique évolue au cours du temps.

#### – L'apprentissage pour la synthèse de comportement

La synthèse de commandes dites « orientées tâche » est envisageable également par apprentissage en ligne. Ces commandes permettent l'adaptation du comportement aux variations de l'environnement par un asservissement direct sur des informations extraites de capteurs embarqués. L'apprentissage profond, et en particulier l'apprentissage par renforcement profond, a permis de montrer qu'il est possible d'apprendre des comportements en utilisant simplement des images de la scène en entrée sans passer par des espaces d'état bien définis. Bien que permettant de résoudre des tâches de manière efficace, ces algorithmes conduisent à des contrôleurs qui ne peuvent pas en général être simplement réutilisés pour d'autres tâches car ils mélangent de manière intriquée la perception des informations pertinentes et la prédiction de la performance des actions. Toutefois, les techniques d'apprentissage par transfert ouvrent une voie pour aborder ces problèmes de manière plus générique.

#### – La supervision et la planification de mission des systèmes multi-robots

La robustesse dans le déploiement de systèmes multi-robots autonomes passe par des capacités de supervision et de planification dynamique des activités du système et d'adaptation par la commande pour faire face aux évolutions des tâches, de l'environnement et de l'état des entités.

Certaines tâches nécessitent d'avoir recours à une coordination des trajectoires entre des vecteurs de même nature ou hétérogènes. Plusieurs paradigmes ont été développés pour coordonner les trajectoires sous un certain nombre de contraintes sur les variables d'état des entités. Cependant leur champ d'application demeure aujourd'hui restreint en raison de la formulation du problème de coordination. La coordination doit pouvoir en effet être induite par l'objectif de la mission et tirer profit de méthodes de décision pour la planification de la mission. La navigation coordonnée pour l'exploration visuelle distribuée d'un environnement et sa reconstruction par fusion des cartes locales peut s'envisager par des modèles probabilistes bayésiens hiérarchiques. Cette représentation commune peut servir à la localisation mutuelle des engins embarquant les capteurs et à la détermination de stratégies efficaces d'exploration minimisant par exemple les taux de revisites.

Pour de tels systèmes, la gestion des missions par l'homme d'une supervision de haut niveau est souvent nécessaire. Cette gestion doit être reliée à la fonction d'interprétation des données issues des systèmes de capteurs, des trajectoires des vecteurs supports des capteurs et des éléments de contexte. Les techniques de réalité virtuelle exploitant les informations géométriques et sémantiques extraites de la scène ainsi que des représentations par cartes cognitives dynamiques sont envisageables pour constituer des interfaces homme-système d'aide à la décision.

## Autonomie et robotique à l'ONERA



La robotique et l'autonomie constituent un domaine d'activité important de l'ONERA. Les applications concernent les domaines de la défense et la sécurité (pour la surveillance de zone notamment), de l'aéronautique (les drones en particulier) et du spatial (On-Orbit-Servicing et exploration planétaire).

Les fonctions de navigation coordonnée et de perception distribuée constituent un champ d'application pour les techniques de l'IA : l'un des exemples pour de tels systèmes est la patrouille de drones partageant une mission commune de surveillance ou de protection d'une zone critique. Pour de tels systèmes, il faut concevoir et architecturer une fonction de perception globale, distribuée sur tous les acteurs. Celle-ci doit être reconfigurable dynamiquement en fonction des aléas ou des capacités de chaque drone à aller chercher par lui-même l'information au bon endroit.

Une partie des développements réalisés dans ce cadre fait l'objet de la plate-forme MAUVE (<https://gitlab.com/MAUVE>). Les applications de l'IA à la commande de systèmes sont également un champ d'investigation qui fait l'objet d'une capitalisation dans la bibliothèque GENETIC (<https://github.com/OneraHub/GENETIC>)

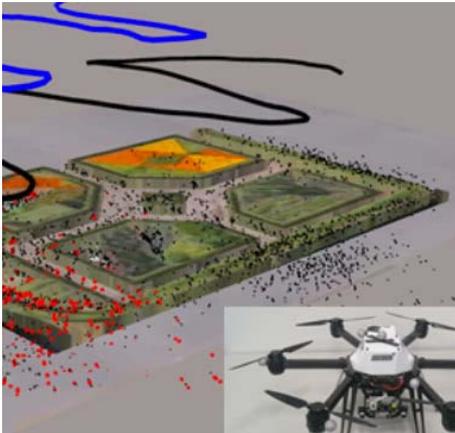
@ONERA

## Identification, commande neuronale



Les techniques d'identification et de commande de systèmes développées à l'ONERA regroupées dans la bibliothèque logicielle GENETIC intègrent notamment des techniques d'apprentissage profond et d'apprentissage par renforcement. Celles-ci peuvent être insérées dans des schémas d'identification et de commande de systèmes, les capacités d'apprentissage rendant les tâches réalisées plus robustes aux perturbations et plus autonomes si on intègre le niveau décisionnel. Des techniques dites de commande neuronales ont été développées sur des bras robotisés. Celles-ci permettent d'aborder des problèmes d'identification en ligne des paramètres d'un modèle de systèmes dynamiques et la mise en œuvre de lois de commande adaptative pour prendre en compte **des phénomènes non linéaires et non stationnaires tels que le frottement ou la flexibilité**. Des techniques de régression pour fournir des approximations de modèles, ont également été développées pour la commande par inversion de modèles.

## Navigation visuelle



On cherche aujourd'hui à constituer des systèmes intelligents embarquant des capteurs doués d'une capacité élevée d'analyse et de décision, à la fois pour comprendre l'environnement dans lequel ils évoluent mais aussi pour commander leur mouvement, accéder à des informations de manière plus active et adapter les traitements aux finalités.

Ces problèmes d'intégration sensorimotrice s'illustrent par la navigation à partir d'informations visuelles d'un engin dans un milieu intérieur ou extérieur non connu a priori voire dynamique (obstacles mobiles par exemple). Nous avons développé des techniques de SLAM distribué afin d'améliorer les capacités de localisation et la robustesse du guidage visuel. Ces travaux exploitent des méthodes d'apprentissage pour apporter de l'adaptabilité à la localisation, à la cartographie ou à la reconnaissance d'objets.

Nous développons également des techniques d'apprentissage par renforcement profond qui permettent d'apprendre à naviguer dans une scène pour atteindre une cible visuelle donnée, ceci en anticipant les éventuelles collisions avec l'environnement. Techniquement, le modèle sans carte suit une méthode d'apprentissage par renforcement par acteur-critique où la fonction de récompense a été conçue pour être consciente des collisions.

La navigation sémantique est également un domaine dans lequel nous développons des travaux importants. La compréhension visuelle de la scène est une capacité clé pour permettre aux robots d'évoluer sur la base d'une reconnaissance de leur environnement. Nous avons mis au point notamment des techniques d'étiquetage sémantique de la scène perçue par un robot par des approches multi-vues exploitant une meilleure intégration des images et de la géométrie. Sur ces principes, nous avons réalisé des reconstructions de structures 3D à partir de la sémantique 2D sur les bases de test SUNRGBD et NYUDv2, conduisant à des avancées significatives de l'état de l'art.

### 2.3.4 Interactions et cognitive



*Les interactions de l'homme avec son environnement humain et matériel reposent sur des processus cognitifs qui, d'une manière générale, concernent la perception ou les activités sensorimotrices et la cognition sociale. Ces fonctions résident dans des régions particulières du cerveau (cortex moteur pour les neurones miroir, amygdale pour les émotions et la confiance en soi, lobe préfrontal pour la « perception », etc.). Celles-ci sont influencées par l'environnement humain et physique qui induit des adaptations du comportement notamment à travers des processus psychologiques engageant les émotions, les codes des relations interpersonnelles, etc.*

Les interactions homme-système et la conception des moyens qui les structurent, que ce soit pour la mise en œuvre des boucles de pilotage ou pour des fonctions de plus haut niveau (supervision, décision, etc.), sont à envisager dans une perspective qui intègre ces dimensions, ceci dès lors que l'homme et la machine partagent des tâches, des décisions ou l'espace. Que ces interactions concernent des entités humaines ou des systèmes « numériques », les activités partagées doivent être mutuellement intelligibles voire prédictibles et fondées sur une même représentation de l'activité, de la connaissance et du raisonnement. Elles doivent être en outre susceptibles d'évoluer par apprentissage et de s'adapter au contexte.

### – Analyse et prédiction de comportement

De la connaissance et de la bonne maîtrise des interactions entre l'homme et ces systèmes dépend la qualité des activités induites : la performance, la sécurité et la sûreté des opérations. Ces interactions engagent des fonctions cognitives et sensorimotrices qui sont souvent multimodales (plusieurs modalités perceptives – par exemple visuelles, auditives, haptiques –, ou plusieurs modalités d'action – par exemple gestuelles, vocales, oculomotrices, tactiles –, sont souvent exploitées conjointement) et qui s'inscrivent dans une structure dynamique dépendante des objectifs et du contexte.

On peut accéder à certaines données relatives aux interactions par le biais de divers instruments de mesure et de modélisation permettant de capturer des paramètres physiologiques (EEG, EMG, NIRSI, etc.) et biomécaniques (MOCAP, dynamique, accélérométrie, etc.) pour les intégrer dans des simulations physiques (humains virtuels) afin de reconstruire et d'analyser des comportements, des activités motrices ou sensorielles à l'aide de méthodes d'apprentissage statistique. Le traitement (analyse multivariée) de ces longues séries temporelles peut permettre de détecter des précurseurs de la fatigue, de la perte d'attention, de la désorientation spatiale, etc. Des variantes temporelles des modèles de Bayes naïfs augmentés et des réseaux de nœuds temporels bayésiens (TNBN), ou encore des réseaux de type LSTM peuvent permettre d'envisager la détection ou la prédiction d'évènements.

### – Décision collective

Les problèmes d'aide à la décision collective incluant des notions d'apprentissage peuvent également être abordés par les techniques de l'IA. Lorsque cela s'avère nécessaire du fait de la complexité de l'objectif à atteindre, les agents « intelligents » sont intégrés dans des systèmes distribués dénommés systèmes multi-agents. Ceux-ci sont constitués d'un ensemble d'agents autonomes reliés et collaborant entre eux. On va alors chercher à maîtriser l'intelligence collective, ou une intelligence partagée émergeant de la collaboration ou de la compétition entre plusieurs individus.

Dans ce cadre, l'apprentissage collaboratif de l'intelligence artificielle distribuée (IAD) peut être envisagé de façon décentralisée. À l'ONERA, nous abordons tous types de problèmes de décision, qui vont de la décision individuelle à la décision de groupe, notamment par le biais des processus décisionnels de Markov partiellement observables qui permettent de formaliser la notion d'apprentissage par renforcement. En théorie ces modèles formels sont aussi applicables aux systèmes collectifs composés de plusieurs agents.

## Interaction et cognitive à l'ONERA



*Les travaux réalisés à l'ONERA en IA en matière d'interactions homme-systèmes concernent en particulier l'analyse de l'état des opérateurs – stress, engagement, attention, etc. – à partir de signaux physiologiques et, à travers des modèles, celle de performances et de robustesse aux perturbations de systèmes interactifs. Ceux-ci sont plus spécifiquement des systèmes aéronautiques (avions, hélicoptères, drones), des systèmes de défense (tous types d'engins, systèmes de surveillance, etc.) mais aussi plus largement des systèmes d'information et de gestion de données (tenue de situation, simulations, expérimentations à grande échelle, etc.).*

## Détection et prédiction des pertes d'attention

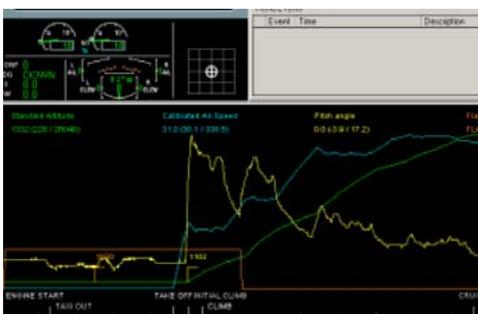


Les processus attentionnels sont essentiels dans la perception, la prise de décision, la sélection et l'exécution de l'action. Les activités des contrôleurs aériens par exemple reposent sur leurs capacités à déplacer et diviser leur attention spatiale sur les interfaces homme-machine de contrôle pour identifier d'éventuels risques de collision.

L'ONERA développe des méthodes de détection et de prédiction de la perte d'attention pour proposer des adaptations des interfaces à l'état cognitif des opérateurs. Ces méthodes exploitent des signaux physiologiques (notamment les ondes alpha de relaxation profonde des EEG qui renseignent sur la charge attentionnelle d'un individu) ainsi que des signaux de l'activité oculomotrice et motrice.

Le décodage et la classification des signaux physiologiques associés à l'attention spatiale et à la charge attentionnelle sont réalisés à l'aide d'algorithmes d'apprentissage supervisé classique tels que l'analyse discriminante linéaire (LDA), l'analyse discriminante quadratique (QDA), le *Support Vector Machine* (SVM), la classification probabiliste naïve bayésienne et aussi les modèles d'apprentissage par renforcement. Des modèles de classification capables de préserver et analyser la dépendance temporelle des signaux physiologiques à base de LSTN sont également utilisés pour classifier en ligne la charge attentionnelle et pour comprendre comment s'organisent les différentes tâches de contrôle ainsi que l'évolution des performances comportementales des opérateurs.

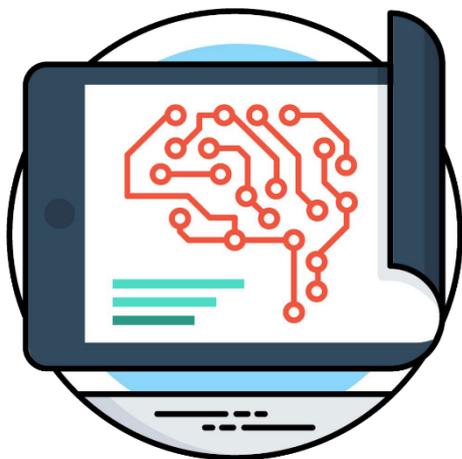
## Détection d'anomalies dans les données FDR



Les données des enregistreurs FDR (*Flight Data Recorder*) des vols commerciaux sont exploitables à la détection et à la caractérisation d'anomalies dans les opérations pouvant avoir une incidence sur la sécurité des vols ou encore sur les besoins de maintenance. Les données concernent plusieurs centaines de paramètres relatifs à des grandeurs physiques très variées, enregistrées à des fréquences de 1 à 4 Hz. L'analyse de ces données vise en particulier à détecter des écarts par rapport aux opérations normales, des dépassements qui caractérisent des événements de sécurité et la recherche de précurseurs.

L'ONERA développe pour cela des méthodes d'analyse dans un contexte *Big Data* (des données portant sur plus de 1 million de vols) pour caractériser des événements par phase de vol. Les algorithmes de détection adaptés à des séries temporelles de données hétérogènes, discrètes et continues, basés sur des modèles d'opération ont été développés (*MKAD-Multiple Kernel Anomaly Detection*). Ils sélectionnent et combinent automatiquement les noyaux les plus pertinents dans un ensemble donné. Ces techniques ont été appliquées notamment à l'analyse des conditions conduisant à la remise de gaz lors des phases d'approche finale.

## 2.4 L'IA et les calculateurs



Les progrès que nous observons aujourd'hui dans le domaine de l'apprentissage artificiel tiennent à la convergence de trois domaines clés : la donnée, l'infrastructure pour la capter, gérer, traiter ainsi que la puissance de calcul des puces pour la digérer, soit en termes techniques : *Big Data*, *Cloud/Deep Machine Learning*, CPU/GPU.

Les architectures et les capacités des calculateurs ont connu des évolutions très rapides. De 1971 à aujourd'hui, nous sommes ainsi passés des 2300 transistors du premier microprocesseur d'Intel, à des puces embarquant plusieurs milliards de transistors. Des calculateurs accessibles à tous (à base de GPU par exemple) permettent aujourd'hui la mise en œuvre de nombreuses innovations du *Deep Learning* – CNN (*Convolutional Neural Network*) et RNN (*Recurrent Neural Network*) et la prochaine vague de progrès proviendra des réseaux générateurs d'opposition (GAN) et de l'apprentissage par renforcement.

La course à l'informatique pour l'IA concerne également d'autres architectures de calculateurs pour le calcul haute performance (HPC), le calcul neuromorphique (NC) et le calcul quantique (QC). D'autre part, des fabricants de puces comme Intel parient sur l'informatique probabiliste comme un composant majeur de l'IA qui permettrait de faire des calculs avec les incertitudes inhérentes aux données naturelles.

Tandis qu'Intel, Nvidia et d'autres fabricants de puces traditionnels s'efforcent de tirer parti de la nouvelle demande pour les GPU, d'autres comme Google et Microsoft développent leurs propres puces propriétaires qui accélèrent les traitements à l'aide de leurs propres plates-formes d'apprentissage profond.

Les algorithmes d'IA présentent de vraies spécificités dans leurs exigences techniques. Là où un modèle de simulation requiert une grande précision dans les calculs en virgule flottante, la mise en œuvre d'un réseau de neurones demande une précision bien moindre, mais un très grand nombre de cœurs de calcul interconnectés à très haute vitesse. De fait, les processeurs de cartes graphiques (GPU) sont bien adaptés, avec un ratio performance/consommation électrique très intéressant.

Les calculateurs modernes voient ainsi leur nombre de puces graphiques s'accroître en raison de la montée en puissance de l'IA. L'Institut de technologie de Tokyo (Tokyo Tech) a récemment annoncé la construction de Tsubame 3.0, un ordinateur dédié à l'IA. Son architecture hybride compte 540 nœuds de calcul, avec deux processeurs Intel Xeon E5 pour chacun d'eux et 4 GPU Tesla P100 de Nvidia, soit un total de 2 160 GPU. La capacité de cette machine est estimée à 12,2 pétaflops (soit 12,2 millions de milliards d'opérations par seconde) en calcul double précision, mais à 47,2 pétaflops en "demi-précision", c'est-à-dire pour les applications d'IA.

Une nouvelle vague de processeurs conçus pour l'IA arrive sur le marché. IBM a fourni au Lawrence Livermore National Laboratory sa puce neurosynaptique TrueNorth afin de réaliser de l'analyse de données de type *machine learning*. Ce composant, issu d'un programme de la Darpa, l'agence du département de la défense des États-Unis spécialisée dans l'innovation de rupture, exploite un réseau de 16 millions de neurones et 4 milliards de synapses pour une consommation limitée à 2,5 watts. Intel a acquis la start-up Nervana pour disposer d'une technologie comparable. Et Google en est à la deuxième génération de sa puce dédiée à l'IA, le TPU (*Tensor Processing Unit*). Ce dernier a conçu un ordinateur spécialisé, équipé de 64 de ces puces, qui affiche une puissance de 11,5 pétaflops. Un modèle de traduction automatique qui nécessitait une journée d'apprentissage sur une machine dotée de 64 des plus puissants GPU ne nécessite plus qu'une demi-journée de travail en exploitant 8 TPU.

## 2.5 L'IA et les données



Les fondements même de l'apprentissage machine reposent sur les données et une partie des enjeux de l'IA portent sur l'exploitation de flux de données hétérogènes, de très grande taille, parfois sous des contraintes temporelles, ceci pour extraire des connaissances : des comportements, des événements rares notamment. Il s'agit là du contexte dit du *Big Data* qui repose sur les cinq éléments clés : volume, vitesse, variété, véracité et valeur (les 5V). Le *Big Data* et l'intelligence artificielle sont deux technologies inextricablement liées, au point que l'on peut parler d'une *Big Data Intelligence*.

La science des données a évolué rapidement ces dernières années et suit l'évolution des moyens de calcul et des volumes de données disponibles. Les bases de données se déstructurent et exploitent le *cloud*, les moyens de calculs se groupent (*cluster*).

Aux techniques statistiques usuelles sont venues s'ajouter des méthodes d'apprentissage non supervisées ou supervisées pour l'analyse des données non structurées, semi-structurées et structurées, des données temporelles, spatio-temporelles, les flux de données provenant de toutes sortes de terminaux (capteurs, assistants personnels, applications du web, etc.)

L'essor des mégadonnées a entraîné de nouvelles demandes en matière de systèmes d'apprentissage automatique, pour apprendre des modèles complexes, dotés de millions de paramètres, en mesure de digérer d'énormes jeux de données (pour une part dans des flux) et offrant une capacité d'analyse prédictive (caractéristiques latentes de grande dimension, représentations intermédiaires et fonctions de décision).

Pour exécuter des algorithmes d'apprentissage machine à de telles échelles, on est amené à distribuer les calculs sur des *clusters* comportant des dizaines à des milliers de machines. Des efforts importants en ingénierie des logiciels sont alors nécessaires pour optimiser les temps de traitement sur des bases de données massivement parallèles. Des *frameworks* tels qu'Hadoop et Spark <sup>11</sup> sont disponibles pour cela – *frameworks* conçus initialement pour le traitement distribué des données auxquels sont ajoutés des bibliothèques d'apprentissage automatique.

Pour des tâches complexes d'apprentissage machine, et en particulier pour la mise en œuvre des réseaux de neurones profonds, les modèles à parallélisme de données sont abandonnés au profit d'un parallélisme de modèles. Ceci permet de distribuer l'apprentissage et l'inférence dans les réseaux profonds. Plusieurs instances du modèle d'apprentissage sont distribuées pour résoudre un même problème d'optimisation ou des problèmes différents (approches centralisées, décentralisées ou concurrentes de l'apprentissage distribué).

Ce problème de passage à l'échelle des algorithmes d'apprentissage profond est un domaine de recherche relativement récent – les premières propositions comme celle de *Google DistBelief*, devenu *TensorFlow*, reposant sur une approche centralisée de descente de gradient stochastique asynchrone datent de 2012 – et particulièrement actif compte-tenu des enjeux relatifs à la prédiction et des connaissances extraites sur des modèles de très grande taille, domaine du *data analytics*.

---

<sup>11</sup> Big Data: Principles and Paradigm publié par Rajkumar Buyya, Rodrigo N. Calheiros, Amir Vahid Dastjerdi

## 2.7 L'IA et les mathématiques



Bien que les réseaux de neurones aient démontré leur capacité à apprendre à partir d'informations complexes très variées, y compris maintenant avec les réseaux profonds convolutionnels sans extraction préalable de caractéristiques, on ignore encore pourquoi, d'un point de vue mathématique, ils offrent de telles capacités de généralisation, ou encore de reconnaissance (lié à l'apprentissage de dépendances lointaines, dans le temps et dans l'espace). Leur structure est généralement conçue de manière heuristique et leur comportement reste pour la plupart une boîte noire avec fonctionnement régi par un ensemble d'hyper-paramètres.

### – Modélisation et analyse des réseaux de neurones

Une compréhension mathématique plus profonde des réseaux de neurones peut améliorer la maîtrise de leurs performances et bénéficier à la recherche de nouvelles architectures. Elle peut également donner des clés au problème de « l'explicabilité » du comportement des réseaux de neurones.

Les mathématiques se sont penchées sur le fonctionnement des réseaux de neurones dès les années 80. Les premiers travaux traitaient de plusieurs sujets liés aux propriétés des réseaux à rétro-propagation : minima locaux, unicité des poids, capacité de classification, etc.

En 1989, il a été démontré que dans certaines conditions de régularité (théorème d'approximation universelle), un réseau de neurones à une seule couche cachée et une fonction d'activation sigmoïdale peut permettre d'approximer toute fonction continue lorsque des paramètres appropriés leur sont attribués sans expliquer la capacité d'apprentissage algorithmique de ces paramètres.

De plus, il n'existe pas de cadre algébrique standard pour la modélisation et l'analyse notamment des erreurs des réseaux de neurones. Des travaux récents cherchent à fournir des moyens génériques de descriptions mathématiques des réseaux de neurones pour manipuler des cartes génériques (définissant les entrées réseau, ou les paramètres associés à chaque couche) à valeurs vectorielles dans l'hypothèse de la dimensionnalité finie. Ce cadre mathématique pour les réseaux de neurones devrait évoluer pour aller vers les dimensions entrées et sorties stochastiques et infinies afin de pouvoir augmenter leur expressivité et l'extraction automatique de caractéristiques.

### – Robustesse des extracteurs

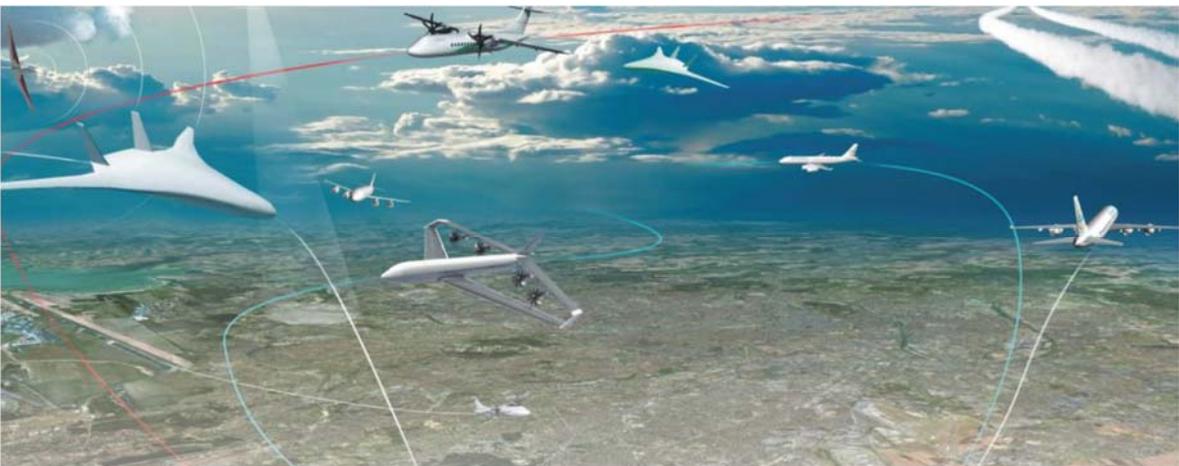
L'analyse mathématique des extracteurs de caractéristiques générés par les réseaux de neurones convolutionnels visant à établir les invariants (structures géométriques), dans la traduction et leur stabilité devant certaines déformations non linéaires, a été initiée à l'aide de méthodes de décomposition multi-échelles. Toutefois, la nature exacte de cette invariance face aux déformations ainsi que des caractéristiques des réseaux de neurones pour contrôler les variations de prédiction (telles que l'étiquette de classe dans une tâche de classification) restent encore à expliquer. Obtenir des preuves de stabilité aux petites déformations permettra d'atteindre des modèles plus robustes pouvant exploiter ces invariances, pour construire la base d'apprentissage sur des représentations stables.

### 3. L'IA, quels enjeux pour le domaine ASD



Les techniques de l'Intelligence Artificielle font l'objet d'un usage important dans le domaine aéronautique-spatial-défense. Les développements récents de ces techniques, induits par les nouvelles architectures de calculateurs, le partage des bases de données d'apprentissage et des algorithmes, ouvrent de nouvelles perspectives. Certaines d'entre-elles sont discutées dans ce chapitre sur la base des enjeux de l'aéronautique civile et militaire, du *New Space* et des technologies à logiciel prépondérant pour la défense et la sécurité. Nous nous arrêterons plus spécifiquement sur celles auxquelles l'ONERA peut, du fait de son potentiel de recherche, apporter des contributions significatives en combinant les expertises des chercheurs en IA à celles des autres domaines des STIC, de la physique au sens large et de la simulation numérique.

#### 3.1 Aéronautique (civile et militaire)

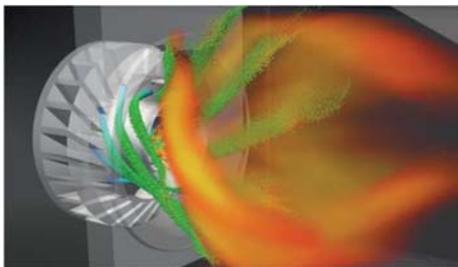


La transformation numérique du secteur aéronautique civil et militaire, comme dans beaucoup d'autres filières industrielles, est un enjeu de toute première importance. Elle porte sur les aéronefs eux-mêmes, qui intègrent des réseaux de calculateurs (plus d'une centaine sur les A320). Elle touche aussi les logiciels embarqués (dont le volume double tous les 10 ans, de l'ordre de 20 MB sur un A320) pour gérer l'ensemble des fonctions de vol et des équipements de bord ainsi que pour assurer l'enregistrement des données sur l'état de l'appareil (un téraoctet de données pour un vol d'une heure).

Cette transformation numérique concerne également les moyens de production qui doivent s'adapter à l'accroissement des cadences, ceux nécessaires à l'analyse de données afin d'améliorer les opérations, la maintenance, etc. et bien évidemment ceux exploités pour l'ingénierie des aéronefs : les moyens de maquettage numérique, de simulation pour la conception, de certification, de recherche de nouvelles architectures, etc.

L'IA ouvre beaucoup de perspectives à l'aéronautique civile et militaire, nous en évoquons un certain nombre dans cette section.

### 3.1.1 Modèles hybrides pour la simulation numérique



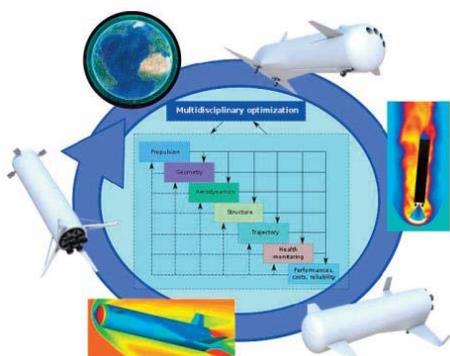
Les modèles de phénomènes ou de systèmes complexes, dans bon nombre de domaines des sciences et de l'ingénierie, peuvent être envisagés sous une forme hybride combinant des modèles physiques à des modèles basés données. Par ailleurs, l'inclusion de techniques inductives puissantes, telles que celles fournies par l'apprentissage automatique dans les moyens de simulation, offre la possibilité d'améliorer les prédictions obtenues à partir d'approches déductives, voire de favoriser l'émergence de nouvelles connaissances.

L'utilisation des empirismes issus de l'apprentissage automatique dans un modèle basé sur la physique aboutit à un modèle dit « hybride ». Ces modèles hybrides permettent de combiner les prédictions de modèles physiques avec d'autres sources de données correspondant à des phénomènes « non modélisables ». Pour ces derniers, on cherche des relations entre les variables d'état du système, sans disposer d'une connaissance explicite (ou seulement partielle) du comportement physique du système, et sachant que ces modèles doivent souvent être réduits à un nombre relativement petit de variables pour un calcul pratique. Par ailleurs, il peut être difficile d'identifier et de paramétrer les caractéristiques qui doivent être incorporées dans le modèle.

Des approches statistiques puissantes deviennent disponibles pour cela. Elles sont basées sur l'exploitation de grands volumes de données par l'apprentissage automatique, prenant en compte l'évolution temporelle du phénomène (comme les réseaux de neurones récurrents et plus spécifiquement les LSTM). Par ailleurs, dans certains cas, la quantité de données disponibles peut ne pas être suffisante pour que le réseau neuronal apprenne un modèle avec la précision requise. D'autres approches, telles que l'augmentation des données (génération de données bruitées) par des méthodes classiques telles que le krigeage ou les forêts aléatoires ou encore des réseaux de neurones génératifs, peuvent être utilisées pour définir un modèle plus robuste. L'apprentissage peut également chercher à inclure des contraintes physiques directement dans les réseaux de neurones.

La dynamique des fluides et celle des structures peuvent tirer parti d'autres méthodes d'apprentissage automatique. Un réseau de neurones DNN auto-encodeur peut être utilisé par exemple pour apprendre un modèle réduit d'un système dynamique. Les méthodes de réduction d'ordre des modèles comme la décomposition orthogonale appropriée (POD) ou SOD (*Smooth Orthogonal Decomposition*) peuvent être combinées avec des DNN (susceptibles de capturer des phénomènes transitoires, intermittents ou multi-échelles non sensibles aux invariances dues à la translation, rotation ou mise à l'échelle) dans le but de créer un cadre d'apprentissage des caractéristiques d'un système dynamique (multi-échelle spatiale et temporelle), de grande dimension avec des structures intermittentes rotationnelles et translationnelles. Au lieu d'apprendre une carte à partir des entrées des coefficients des représentations POD, la réduction de modèle basée sur les données cherche alors à apprendre les opérateurs de modèles réduits. L'utilisation des méthodes hybrides développées pour « l'explicitabilité » des réseaux peut prétendre à donner accès à de nouvelles connaissances dans le domaine de la simulation en structurant les relations entre différents phénomènes.

### 3.1.2 Réduction de modèles et représentation des connaissances expertes pour l'aide à la conception



Les méthodes avancées de conception d'engins aéronautiques reposent sur des techniques d'optimisation portant généralement sur un très grand nombre de variables issues de modèles relevant de plusieurs « disciplines ». Ces méthodes dites de *MDO (Multi-Disciplinary Optimisation)* exploitent des formulations du problème d'optimisation qui décomposent ou pas le problème d'optimisation en sous-problèmes « disciplinaires ». Cette optimisation peut être mono ou multi-niveau. Les couplages, parfois adaptatifs, entre les disciplines sont définis en cohérence avec une méthode d'optimisation et la gestion des incertitudes (décrites à l'aide de différents formalismes, épistémiques des modèles). Les méthodes d'optimisation font appel à des modèles de substitution (encore dit méta-modèles) intégrant le plus souvent les incertitudes dans les modèles de substitution et qui permettent de réduire drastiquement les temps de calcul.

Les réseaux de neurones classiques MLP associés à des méthodes d'optimisation de la topologie des réseaux (par plans d'expérience adaptatifs) sont très souvent utilisés pour fournir ces métamodèles. La réduction de modèles paramétriques ainsi que la réduction de modèle des systèmes dynamiques non paramétriques, comme pour les besoins de la simulation hybride évoqués avant, peuvent chercher à exploiter des modèles d'approximation à différents niveaux de fidélité qui seraient issus de différentes architectures de DNN. Un réseau de codeurs et de décodeurs basé sur un réseau de neurones à convolution (CNN) peut être formé à l'aide du jeu de données d'apprentissage généré à basse résolution. Ensuite, pour le raffinement en deux étapes par exemple, un réseau des réseaux conditionnels génératifs adverses (CGAN) est formé à partir des structures optimisées appariées aux résolutions basse et haute, et est connecté au réseau d'encodeurs et de décodeurs basé sur le CNN.

Les modèles à variables aléatoires mixtes, continues, discrètes ou catégorielles (nombre et type de propulseurs par exemple) peuvent également être abordés par apprentissage profond pour des problèmes à grande dimension à l'aide de processus gaussiens à noyaux non stationnaires pour interpoler les points des bases d'apprentissage.

L'exploration de l'espace de conception (nécessaire à la construction des modèles réduits) et l'optimisation de la conception sont restés des processus manuels relativement fastidieux. Les ingénieurs ont besoin d'une expertise considérable dans les différents outils et les méthodes utilisés. Les outils eux-mêmes restent quelque peu fragmentés. Ils considèrent séparément la conception de plans d'expériences, les différents types d'optimisation – structurelle, multidisciplinaire, multi-objectifs, stochastique, etc. Beaucoup d'interprétations et de corrélations sont nécessaires de la part des experts pour capturer toutes ces connaissances spécialisées et explorer l'espace de conception potentiel, non seulement de son point de vue disciplinaire, mais également avec une visibilité sur tous les autres phénomènes, et à n'importe quel stade du processus de conception.

Dans ce cadre, l'un des enjeux est de parvenir à mettre en place des formulations et des méthodes de conception qui reposent sur un ensemble d'expertises et qui par ailleurs conduisent à des « traçables et reproductibles ». Accompagner l'activité de conception d'un moyen d'interaction, de synthèse de séquences de décision (d'une politique) et d'aide à la décision est un objectif pour lequel certaines techniques d'IA peuvent être mises à profit.

L'interaction avec le processus de conception passe par l'analyse et la visualisation des données (*data analytics*) à l'aide des techniques d'exploration et de traitement statistique des données appropriées (classification, *clustering*, analyse statistique, analyse de sensibilité, etc.) pour révéler des informations nécessaires aux processus décisionnels. L'aide à l'automatisation des processus de simulation, à tous les niveaux de fidélité du modèle, sont la clé pour l'efficacité et la traçabilité des résultats.

La mise en place de systèmes experts (basés sur la connaissance) comme aide à l'utilisateur est également à promouvoir, pour formaliser la connaissance acquise au cours du processus itératif de la mise en œuvre des algorithmes sous forme d'une base de faits et de règles heuristiques, et exploiter cette base de connaissances par un moteur d'inférence qui peut exploiter simplement des déductions logiques mais aussi intégrer des modèles de décision séquentiels dans l'incertain. Les diagrammes d'influence sont aussi utilisables tout comme les processus de décision de Markov pour représenter les liens causaux dans le processus de décision. Apprendre le modèle de décision qui répond à une tâche précise peut conduire à exploiter les réseaux bayésiens.

Ces modèles sont bien étudiés en ce qui concerne l'inférence causale mais l'apprentissage de leur structure reste à approfondir.

### 3.1.3 Smart cockpit – Smart skies



Les systèmes de gestion du trafic aérien vont très probablement connaître dans les années qui viennent de profondes transformations en raison de l'impact des technologies du numérique en général (communications, GNSS, capteurs embarqués, fonctions d'autonomie, etc.).

La gestion du trafic aérien devra également s'adapter au trafic d'aéronefs sans pilote (*UTM, Unmanned Traffic Management*) et faire face à diverses contraintes (nouvelles organisations, réglementations, évolution du trafic) pour une diversité d'objets aériens beaucoup plus grande qu'aujourd'hui.

Cette future gestion de l'espace aérien et du trafic associé devra être suffisamment souple et dynamique pour pouvoir être appliquée de manière cohérente dans tous les types d'espace aérien. La prédiction, la détection d'anomalies, la planification réactive à grande échelle des trajectoires, leur adaptation locale pour s'adapter à des phénomènes météorologiques, etc. sont des questions auxquelles les techniques de l'IA peuvent apporter des réponses en intégrant bien évidemment l'échelle du système concerné.

Les futurs systèmes d'ATM (*Air Traffic Management*) pourront s'appuyer sur des données temps réel du trafic comme celles fournies par l'*Automatic Dependent Surveillance-Broadcast ADS-B* issues du réseau mondial actuel de 20.000 capteurs. Ce réseau fournit les données sur les trajectoires des avions ainsi que tout un ensemble d'informations (identité, etc.) avec une fréquence de l'ordre du hertz. L'exploitation de ces données de plusieurs téraoctets nécessite des traitements parallèles s'exécutant sur des dizaines, des centaines, voire des milliers de serveurs, afin de répondre aux attentes en matière de contrôle aérien. Une nouvelle classe de technologies *Big Data* a émergé récemment et peut permettre d'envisager des traitements distribués sur des bases de données telles que NoSQL, Hadoop et MapReduce.

L'analyse des données de vol en vue de la détection d'anomalies sur les trajectoires de vol comme la prédiction des trajectoires (conflits et retards, en intégrant les données météo) est abordée par des réseaux neuronaux notamment à l'aide des LSTM sur des fenêtres glissantes ainsi qu'à l'aide de réseaux bayésiens. L'analyse de la variabilité dans les opérations aériennes a été réalisée avec des méthodes de *data mining*. L'application des méthodes d'apprentissage profond est à considérer pour l'ensemble des problèmes de classification pour l'identification des zones à risque de collision entre aéronefs, pour la détection des trajectoires atypiques et pour la prédiction des choix de pistes et de *runways*.

L'IA propose également des éléments pouvant être intégrés aux cockpits du futur, à travers des moyens d'analyse et des aides techniques, pour assurer la sécurité des vols et la gestion de la complexité de certaines situations – conditions météorologiques changeantes, incidents avec des moteurs, avec le train d'atterrissage – dans lesquelles doivent être prises des décisions critiques.

Les techniques de classification et de prédiction utilisant notamment des réseaux de neurones récurrents (RNN) profonds ou des architectures de mémoire à court terme (LSTM), peuvent permettre de suivre l'activité des pilotes et de détecter des motifs dans des longues séries temporelles, ou des séquences d'événements traduisant divers phénomènes comme la fatigue, le stress, la perte de contrôle ou d'attention.

Des interfaces d'assistance au pilote pourraient permettre d'effectuer plus généralement une évaluation en temps réel des états cognitifs du pilote et fournir des alertes utiles et opportunes basées sur les prévisions des niveaux de performance du pilote. Les interfaces peuvent pour cela donner accès à des signaux physiologiques mais aussi des signaux verbaux ou non verbaux. Les interactions en langage naturel avec un assistant personnel comme Amazon Alexa, Microsoft Cortana, Siri d'Apple, Watson d'IBM, s'appuient pour cela sur des mécanismes d'attention mettant en œuvre des TDNN (*Time Delay Neural Network*, très utilisés en reconnaissance vocale pour classifier les phonèmes par exemple, et qui permettent d'ignorer les décalages temporels dans les schémas à détecter) ainsi que sur un système de combinaison de réseaux de confusion (CNS) couplé avec un système de recherche du meilleur sous-ensemble pour trancher parmi les propositions des réseaux et ne garder que la meilleure.

### 3.1.4 Systèmes à forte autonomie, leur sûreté et leur sécurité



Les véhicules autonomes ont le potentiel d'améliorer la qualité des systèmes de transport, leur sécurité et leur efficacité. La mise au point de ces nouveaux moyens pour les mobilités terrestre, aérienne, maritime, nécessite des progrès dans de nombreux aspects de l'autonomie des véhicules, allant de la conception du véhicule au contrôle, à la perception, la planification, la coordination et les interactions humaines.

Les véhicules autonomes évoluant dans des environnements dynamiques complexes nécessitent des moyens d'élaboration de « comportements » qui doivent être en mesure de faire face à des situations imprévisibles et/ou complexes, de raisonner en temps opportun afin d'atteindre le niveau fiabilité et de sécurité au moins égal à celui qui pourrait être atteint par un opérateur humain.

La fonction de pilotage autonome (comme l'AutoPilot de niveau 3 de Nvidia par exemple) repose sur un système de perception de l'environnement qui reconstruit les informations utiles au pilotage (détection d'obstacles, localisation, cartographie, calcul de trajectoires, mémorisation des parcours, etc.) et qui satisfait des exigences en matière de sécurité imposées par la réglementation. L'homologation de ces autopilotes, qui intègrent des composants logiciels d'apprentissage et de décision, constitue une exigence à laquelle il n'existe pas à ce jour de réponse. Cela pose des questions sur les bases d'apprentissage dont dépend la capacité de généralisation des réseaux ainsi que sur les méthodes d'apprentissage et d'inférence pour garantir la stabilité des classificateurs. La représentation des structures des réseaux doit aussi permettre de rendre leur fonctionnement interprétable et servir à l'explicabilité des décisions. Pour des fonctions élémentaires (comme la détection de collisions), l'utilisation de méthodes de vérification basées sur des solveurs SMT (Satisfaisabilité Modulo-Théories) est envisageable pour la vérification de propriétés comme l'existence de solutions.

La commande pour le pilotage de systèmes dynamiques a été réalisée à l'aide de techniques de commande prédictive MPC (*Model Predictive Control*) dans lesquelles les fonctions de prédiction sont apprises en ligne.

Les techniques d'apprentissage profond pourraient intégrer perception-commande-planification et générer une entrée de contrôle directement à partir d'informations sensorielles (visuelle par exemple) et d'une connaissance du contexte. Des techniques d'apprentissage par renforcement utilisant des réseaux profonds pour apprendre des politiques Q-learning ont été proposées pour des systèmes discrets. Pour le contrôle de systèmes à variables d'état continues, des algorithmes de type « acteur-critique » ont été expérimentés ainsi que des algorithmes Q-learning et d'apprentissage par renforcement ou par démonstration.

Notons que les questions de sûreté et de sécurité des systèmes sont déterminantes dans ce contexte pour l'intégration des composants logiciels exploitant de l'IA aux systèmes critiques.

### 3.1.5 Systèmes de production agile



Les constructeurs aéronautiques anticipent des augmentations importantes de la production de certaines gammes de leurs avions. Airbus se prépare ainsi à porter le rythme de production de ses avions monocouloirs A320 à 75 appareils par mois pour la prochaine décennie alors qu'elle était de 40 en 2015.

Pour suivre la hausse des cadences, les lignes de production doivent adopter des technologies agiles et flexibles ainsi que des moyens de planification « dynamique » de la chaîne de production (et de la *supply chain*). Les procédés de fabrication et d'assemblage pour la production d'avions, d'hélicoptères, de moteurs et de fusées, s'adressent en effet à de petites séries (quelques dizaines par mois), ce qui est très différent de l'industrie automobile. Les opérations de maintenance sont multiples et parfois longues : démonter un réacteur sur un avion est une opération complexe qui prend six heures en moyenne.

Planifier de manière sûre et optimale les activités est tout à fait fondamental pour la maîtrise de la production. Le résultat incertain des actions et les observations bruitées et/ou partielles rendent toutefois la planification difficile.

Dans de nombreux contextes de ce type, les processus de décision de Markov partiellement observables (POMDP) définissant des actions optimales pour un seul agent et un POMDP décentralisé (DEC-POMDP) pour plusieurs agents coopératifs sont envisageables. Les POMDP et DEC-POMDP ou des alternatives comme les *Stochastic Safest and Shortest Path* (S3P) sont des modèles expressifs mais exigeants sur le plan de la résolution (définition des probabilités, existence d'une politique qui atteint les buts avec probabilité garantie, etc.). Les techniques par satisfaction de contraintes comme les *Partial Satisfaction Planning* (PSP) ou les *Over-Subscription Planning* (OSP) associées à des recherches heuristiques ont monté leur capacité à traiter des problèmes de planification stochastique représentatifs des problèmes de gestion de production dans le domaine aéronautique.

### 3.1.6 Maintenance individualisée – Pronostic et diagnostic



Le ministère des Armées a placé le MCO/MRO 4.0<sup>12</sup> aéronautique et la logistique qui l'accompagne au cœur des besoins de modernisation. L'industrie aéronautique civile a également inscrit la MRO/MCO 4.0 parmi les grands domaines dans lesquels un effort d'innovation est fortement encouragé. La MRO 4.0 concerne la transformation numérique des moyens de maintenance pour l'aéronautique civile. Dans le domaine militaire, on parle de M2CO 4.0 (maintenance pour le maintien en condition opérationnelle 4.0). MRO/MCO est un domaine d'intérêt civil et militaire pour lequel les enjeux économiques sont très importants (marché de 80 milliards d'euros en 2017, croissant à un rythme de près de 10% par an depuis 10 ans).

La détection d'anomalies (le diagnostic) dans le domaine aéronautique concerne tous les sous-ensembles des aéronefs : cellule, propulseurs, commandes de vol, équipements de navigation, avionique, etc. Elle s'inscrit dans un cadre réglementaire (normes) très précis sur l'exploitation des aéronefs. Des données sont collectées pour cela lors des différentes phases de maintenance (quatre types d'opération de maintenance sont organisées périodiquement, auxquelles s'ajoutent des inspections routinières) et pendant les opérations (événements opérationnels, données de vol : conditions extérieures, état moteur, etc.). La détection et l'identification des anomalies exploitent parfois des moyens de mesure et des traitements qui dépendent du domaine concerné (structure, réseau électrique, logiciel, équipements, etc.).

Le diagnostic est l'un des domaines d'usage privilégié des algorithmes d'apprentissage. Des réseaux de neurones et des systèmes experts ont été utilisés dès les années 90 dans le domaine aéronautique pour la détection et l'aide au diagnostic. De nombreuses méthodes d'apprentissage non supervisé sont exploitées aujourd'hui pour la classification (SVM, *k-means*, Kohonen, etc.) et pour la détection de données anormales. Des méthodes d'apprentissage supervisé (perception multi-couches, forêts aléatoires, arbres de décision) utilisent des modèles de défaillance ou des anomalies souvent simulées car on ne dispose que de peu de bases de données d'anomalies labellisées.

Les anomalies recherchées sont souvent multidimensionnelles. Les vecteurs d'entrée des réseaux d'apprentissage sont généralement de grande taille et multi-domaines. Les données sont présentes sur de longues séries temporelles qui nécessitent d'être normalisées pour les rendre indépendantes du contexte notamment. Dans ce contexte, les réseaux convolutifs bi-directionnels (pour l'extraction automatique de caractéristiques des séries temporelles pour former des *Time-Series Bag of Features*), combinés avec des réseaux de neurones récurrents comme les LSTM (bien adaptés au traitement de séquences dont la dynamique s'étale sur un long terme) ont été expérimentés à partir de données brutes. L'apprentissage de la métrique de classification par des réseaux de neurones siamois pour construire une mesure de similarité est également beaucoup utilisé pour l'inférence des classes.

Les réseaux de neurones récurrents peuvent être également utilisés à la prédiction séquentielle du comportement (pronostic) des systèmes dynamiques. Ils ont montré leur efficacité dans des domaines comme l'interprétation et la génération du langage naturel. Ces modèles génératifs à vecteurs latents capturent bien les dynamiques temporelles et permettent de former des modèles non linéaires de prédiction. Ils nécessitent toutefois une quantité de données importante pour atteindre des bonnes performances de prédiction. Le pronostic prévoit les performances futures d'un composant en évaluant la déviation du comportement par rapport aux

<sup>12</sup> MCO : Maintien en Condition Opérationnelle / MRO : Maintenance, Repair, Operations / dans le cadre de l'industrie 4.0

conditions normales de fonctionnement. Il est donc nécessaire de disposer d'informations initiales sur les éventuelles défaillances (y compris le site, le mode, la cause et le mécanisme). Une tendance actuelle est de produire des données par des simulations avec des jumeaux numériques pour constituer des bases d'apprentissage maîtrisées incorporant les phénomènes de dégradation et des aléas. Les utilisations potentielles pour le pronostic sont dans la maintenance conditionnelle.

En plus de la détection des pannes, des diagnostics et des pronostics, l'aide à la décision des actions à entreprendre est souvent nécessaire.

## 3.2 Spatial



Les applications au domaine spatial de l'intelligence artificielle sont très nombreuses et déjà à l'origine d'avancées importantes que ce soit pour les besoins de l'observation de la Terre, de l'exploration ou de l'exploitation de l'espace.

Depuis 2000, la surveillance de la Terre par l'imagerie satellitaire a conduit à accumuler plus de 100 pétaoctets d'images et développe chaque jour plus de 3 000 000 km<sup>2</sup> de vues de la Terre. Identifier les objets dans une image donnée, détecter des changements et les catégoriser sur différentes échelles spatiales et temporelles sont des problèmes où l'apprentissage automatique a montré son efficacité tout comme pour le traitement des images issues des télescopes terrestres ou spatiaux.

La gestion autonome des données recueillies par des algorithmes d'apprentissage intégrés aux traitements à bord des satellites pour supprimer les données superflues et planifier les téléchargements afin d'optimiser la transmission de données, etc. est à envisager, pour atteindre des moyens « intelligents », à l'image du rover martien Curiosity de la NASA qui dispose à son bord d'un logiciel de sélection des objectifs pour sa caméra *ChemCam Chemistry* (ChemCam).

Des fonctions plus complexes peuvent être envisagées pour les télécommunications par exemple. Une radio intelligente ou radio « cognitive » pourrait apprendre à se protéger temporairement des radiations lors d'événements météorologiques extrêmes. Cela suppose des moyens de détection et de classification des perturbations et d'adaptation des paramètres : modulation, puissance d'émission et bandes de fréquence mais aussi des trajets par réseau radio des données vers le sol.

Un autre domaine dans lequel l'apprentissage automatique peut améliorer la technologie actuelle est celui du contrôle relatif du mouvement des engins spatiaux et des satellites. Des capacités reposant sur l'apprentissage automatique augmenteraient les capacités d'ajustement de l'orbite, de navigation autonome et d'amarrage de véhicules spatiaux ou leur atterrissage.

Également pour la surveillance de l'environnement spatial, l'apprentissage sous ses diverses formes permettrait de découvrir des anomalies, reconnaître des structures, prédire des événements potentiellement dangereux pour les satellites et les astronautes, etc.

### 3.2.1 Le renseignement géospatial



Le GEOINT dans sa définition la plus large (production d'informations pertinentes et analyse géospatiale à l'intention des décideurs) connaît actuellement une révolution, dont le signe le plus évident est le succès des méthodes d'apprentissage profond dans le domaine de la vision artificielle et de l'interaction homme-machine.

Ces nouveaux traitements sont rendus possibles grâce à l'accès aux bases de données d'imagerie satellitaire et à des moyens de calcul massifs. Les investissements réalisés par les acteurs du domaine ont conduit à la création de bases de données ouvertes, en service gratuit ou payant, à l'échelle mondiale et, dans le même temps, à la diffusion d'outils de support et de traitement pour l'exploitation de ces données (Google Earth Engine, Amazon AWS, etc.).

Toutefois, l'interprétation automatique des données issues des capteurs embarqués, le passage du signal à une information sémantique exploitable en maîtrisant les contraintes de temps, de volume, de ressources et de qualité de l'information produite reste un défi. Les performances des algorithmes d'interprétation de données par l'utilisation de techniques d'apprentissage connaissent des évolutions rapides. Des descriptions multi-label (descriptions de comportements et d'actions) s'accroissent de manière importante et bénéficient pour cela des apports des travaux dans le domaine du *Big Data*.

Dans ce contexte, la perception augmentée de l'environnement revêt un enjeu stratégique qui est celui de la compréhension d'une scène dynamique au sens large, l'anticipation de son évolution, la détection d'événements imprévus ou anormaux.

Il peut s'agir de la surveillance globale à l'échelle planétaire d'activités humaines et/ou de leurs conséquences environnementales, de la surveillance côtière, d'un site urbain ou industriel ou encore de moyens de renseignement mis en place sur un théâtre d'opération donné en soutien à des interventions. Les besoins associés concernent aussi bien les domaines de la défense et la sécurité, des enjeux économiques, ou encore des problématiques écologiques et environnementales.

Les techniques de traitement doivent répondre à des problèmes d'une complexité extrême, demandant à maîtriser en même temps les dimensions : physique (caractérisation des objets et des scènes), temporelle (détection d'évolution, modélisation de comportement), multimodale (informations hétérogènes et dépendantes des conditions d'environnement), sociotechnique (modélisation de l'activité humaine, d'un contexte opérationnel).

La chaîne de perception joue alors un rôle crucial pour construire l'information souhaitée au bon niveau de performance. Elle doit être la plus automatisée possible pour absorber les flux, réduire les délais et assister un opérateur. Elle doit être capable de prendre en compte des données distribuées, asynchrones, multi-physiques. Cela veut dire maîtriser le référencement géographique des données, savoir extraire l'information pertinente dans la masse de données, pouvoir les fusionner pour exploiter au mieux les différentes sources d'informations, et enfin concevoir les produits de valeur ajoutée pour l'utilisateur.

A la révolution de l'accès à la donnée et de son caractère massif, il faut aussi associer la performance technologique croissante des capteurs (résolutions spatiale et spectrale) dans toutes les bandes de fréquences (radar, du visible à l'infrarouge thermique, lidar, etc.) offrant ainsi l'accès à des données multiphysiques qui augurent d'un bond qualitatif pour la caractérisation du contenu des scènes. Il s'agit là d'un cadre pour le développement de capteurs « cognitifs », capteurs de vision (pour des imageurs multi-spectraux comme les caméras *Return Beam Vidicon* (RBV) et capteurs radars (de type SAR et SWIR) pour adapter la fonction de perception aux objectifs de l'observation.

### 3.2.2 Satellites agiles – Autonomie des missions spatiales



Les moyens futurs d'observation de la Terre à partir de l'espace comprendront un mélange de multiples constellations de satellites hétérogènes et homogènes (commerciaux ou non commerciaux), pouvant impliquer diverses configurations et tailles de géométrie (petites et moyennes tailles, par exemple), ainsi qu'un large spectre de modalité de capteur et de résolution.

L'agilité des plates-formes – c'est-à-dire la manœuvrabilité dans des directions spécifiques pour la souplesse d'observation – sera également renforcée.

Il est envisagé que la charge utile de chaque plate-forme soit constituée d'un seul capteur – optique (par exemple, panchromatique, hyperspectral, multispectral, infrarouge) ou radar (par exemple, radar à synthèse d'ouverture SAR). Les capacités de traitement par les calculateurs de bord seraient également renforcées pour s'affranchir des problèmes de communication (visibilité des engins par la station sol), de modifications de l'environnement (couverture nuageuse au-dessus d'une zone à imager), modifications de l'état de l'engin (défaillances matérielles).

Dans ce contexte, la planification de la collecte d'images multi-satellite constitue un problème particulièrement complexe. Celle-ci consiste à affecter aux ressources d'observation, sur un horizon prédéterminé, un ensemble de demandes d'informations (zones à observer) sur des objectifs possiblement multiples. Dans ce problème, tout un ensemble de contraintes sont à prendre en compte. Elles peuvent concerner des considérations relatives à la mission, à la tâche, aux opérations, au collecteur, aux ressources d'appui, à la communication, à la capacité, au temps, à l'itinéraire et aux coûts. L'ordonnement initial des missions multi-satellite vise à déterminer des plans d'allocation ressource-tâche de collecte afin d'optimiser un ou plusieurs objectifs/coût/utilité attendus sur un horizon temporel donné.

Les tâches typiques des missions d'observation, à l'appui de la connaissance permanente de la situation par exemple, comprennent : la surveillance, le suivi, la recherche, la détection, la classification, l'identification de zones d'intérêt dans des images.

Aujourd'hui, la planification des missions spatiales est réalisée hors ligne ou de manière continue par mise à jour d'un plan initial. Le problème de planification est classiquement abordé comme un problème de satisfaction de contraintes maximisant une fonction objective. Celle utilisée pour la planification des tâches de capture est une mesure de performance qui représente différentes dimensions des missions et la diversité des missions considérées. Elle combine typiquement : le niveau de service de l'information (débit de la tâche/satisfaction de la demande, équité vis-à-vis de la demande, retard et âge de l'information), la consommation globale de ressources et d'énergie, l'exposition au risque, la couverture attendue (tâche d'enquête), la covariance d'erreur cinématique, l'erreur quadratique moyenne ou entropie (tâche de suivi), le gain d'information attendu (entropie-incertitude) ou la divergence d'information (estimation de l'état pour la tâche de classification-identification).

### 3.2.3 Surveillance de l'espace



Le spectre électromagnétique des appareils de télécommunications terrestres et spatiales est de plus en plus complexe. Il intègre les ondes radio (AM et FM), les télévisions (VHF), des téléphones sans fils, les communications satellitaires et surtout les radars qui utilisent des fréquences qui s'échelonnent entre 0.5 et 24000 mégahertz.

Ce spectre électromagnétique constitue une source d'information essentielle pour le renseignement stratégique et tactique sur lequel se fonde le renseignement d'origine électromagnétique ou ROEM (*Signals Intelligence - SIGINT*).

Qu'il s'agisse de systèmes de communication (COMINT) ou de signaux électroniques provenant de signaux autres que de communication (intelligence électronique – ELINT), le SIGINT a connu un développement rapide au cours des dernières années.

En plus de détecter la présence d'une source, les signaux peuvent être utilisés pour obtenir des informations sur les équipements (émetteurs et/ou récepteurs) et leurs caractéristiques (type de rayonnement, modulation,

fréquence, forme d'onde, fréquence radio, etc.). Les radars peuvent être identifiés et comparés aux types de signaux connus. Ces informations peuvent ensuite être utilisées pour déterminer sa portée, son utilisation, sa capacité et sa vulnérabilité aux contre-mesures.

La surveillance de l'utilisation des émetteurs est tout à fait stratégique quant à l'intention des forces opposées. L'incapacité à identifier les signaux de manière unique dans un environnement (en les différenciant du fond) crée un risque opérationnel en raison du manque de connaissance de la situation, de l'impossibilité de cibler les menaces et de la vulnérabilité des communications face aux actions malveillantes.

La collecte du renseignement électromagnétique est un outil stratégique. La France a confirmé par la dernière loi de programmation militaire le lancement en 2020 de CERES, un système composé de trois satellites volant en formation et destiné à la détection et la localisation par triangulation des émissions électromagnétiques au sol. Au-delà des applications purement militaires, le programme, conçu pour pouvoir analyser les différents types d'émissions, permettra en plus de déterminer l'architecture de systèmes en réseaux, tel des réseaux de communication, de suivre, en fonction des caractéristiques des signaux, l'évolution du niveau des performances des matériels, etc.

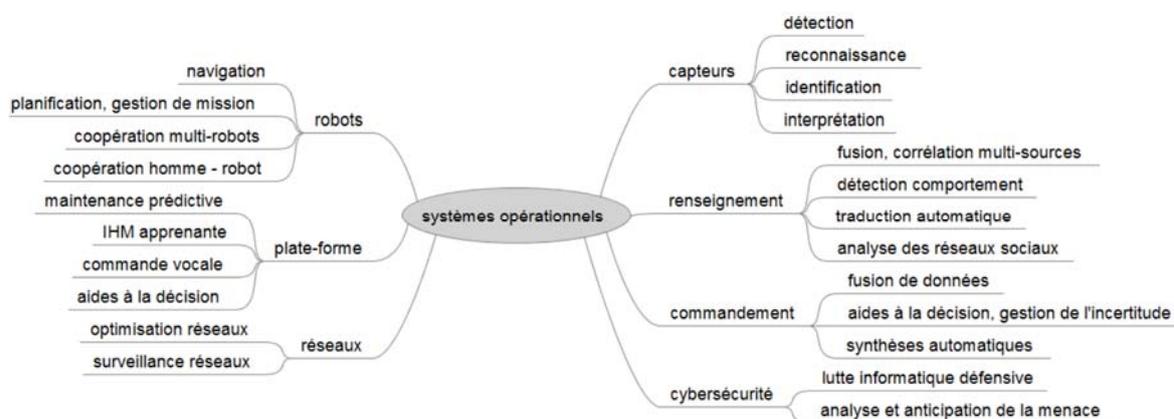
L'application des techniques de l'apprentissage machine au REOM, comme le *MLRF (Multi-Label Random Forest)*, à des fins de classification robuste est un domaine relativement émergent mais prometteur. L'apprentissage automatique devrait donner des capacités beaucoup plus agiles et polyvalentes : un système MLRF, avec une base d'entraînement suffisamment riche, devrait être en mesure d'identifier une vaste gamme de données connues et inédites et de formes d'onde.

### 3.3 Défense/Sécurité



La défense a fait de l'intelligence artificielle un enjeu stratégique et de souveraineté nationale. L'IA recèle une grande variété d'applications potentielles comme l'illustre le graphique ci-après issu de la feuille de route IA 2018 de la DGA.

Pour l'Agence de l'innovation de défense (créée au 1<sup>er</sup> septembre 2018), ce domaine est l'un de ses principaux axes de développement. Plusieurs grands programmes centrés sur l'IA ont été lancés par la DGA.



L'IA est en particulier au cœur des développements liés à l'autonomie des systèmes robotiques envisagés dans des missions de surveillance de l'environnement. Ces missions pouvant être réalisées par des systèmes multi-robots pour lesquels les questions de planification, de supervision sont centrales.

Les interactions entre l'homme et les systèmes à forte autonomie (UAV, UGV, USV, ROV), qu'elles soient directes ou à travers des interfaces de type C4ISR (*Command, Control, Communications, Computers, Intelligence, Surveillance, Reconnaissance*) constituent également des questions qui intéressent l'IA.

Celles-ci sont placées dans le cadre du combat aéroterrestre de demain et d'une « bulle opérationnelle aéroterrestre » (BOA) où divers types d'unités seront simultanément engagés sur un même théâtre d'opérations (avec une forte présence d'engins terrestres, navals, ou aériens, d'autonomie diversifiée et pilotés à distance), connectés entre eux et au C4ISR par des liaisons satellitaires.

Pour des besoins à plus long terme, un programme d'étude amont de la DGA sur le *Man Machine Teaming*<sup>13</sup> a été initié en 2018 avec comme thématiques principales : la gestion des missions des systèmes robotisés, les interactions homme-système et les capteurs intelligents dans l'aéronautique du futur.

À ce programme s'ajoutent des développements importants sur les vecteurs robotiques à haut niveau d'autonomie (programme Scorpion de l'Armée de terre), sur les drones de surveillance et de combat UCAV (*Unmanned Combat Air Vehicle*) ainsi que sur la surveillance de l'espace.

Dans tous ces développements, les techniques de l'IA prennent une part importante : problèmes de classification, prédiction, décision collective, dialogue en langage naturel, représentation des connaissances, analyse de scènes dynamiques, etc.

Les besoins spécifiques des armées en matière de traitement des données pour la logistique, le maintien en condition opérationnelle, de renseignement, de cyber-sécurité, d'analyse prédictive et d'aide à la décision font par ailleurs l'objet d'un grand programme de partenariat d'innovation sur 15 ans : ARTEMIS (Architecture de Traitement et d'Exploitation de l'Information Multi-Source).

Le domaine de la sécurité – qui couvre l'ensemble des champs de la prévention et de la protection face aux risques, qu'ils soient naturels ou technologiques pour les populations ou les infrastructures – partage une partie de ces enjeux. La mise en œuvre des fonctions concourant à la sécurité repose en effet sur tout un ensemble de technologies logicielles pour « l'intelligence des données » à laquelle concourt l'IA notamment pour :

- le traitement automatisé de données hétérogènes extraites de capteurs variés (caméras de surveillance, drones, satellites, objets connectés) ou encore des réseaux sociaux (traitement des données textuelles, analyse de fiabilité des sources, modèles de diffusion et dynamique d'opinion, etc.) ;
- la modélisation et la simulation comme outil dans la prévention des risques : systèmes multi-agents, modélisation multi-échelle des environnements, modèles spatio-temporels des activités (dont les systèmes de transport), prédiction et analyse des risques, décision distribuée, méthodes et outils d'analyse de fiabilité, résilience des plateformes et des systèmes distribués, reconfiguration et adaptation des architectures.

### 3.1.1 Système de combat aérien du futur



Pour répondre à ces futurs enjeux, l'Armée de l'air conduit des réflexions sur le développement du SCAF (système de combat aérien futur). Ce dernier doit lui permettre de mettre en œuvre ses différentes plates-formes aéronautiques, systèmes de commandement et systèmes d'armes au sein d'un même réseau, incluant des systèmes habités et des systèmes non-habités. L'objectif est de pouvoir produire à tout moment l'effet final recherché, grâce à la meilleure combinaison des capacités aériennes, combinées ou intégrées dans un cadre national ou multinational, face à toutes les menaces (conventionnelles, asymétriques).

Dans le cadre du SCAF, le projet *Man Machine Teaming* explore la possibilité de développer un système aérien cognitif de combat. Le principe de ce dernier est de doter les différents systèmes-machines de davantage d'autonomie et d'intelligence artificielle au service d'une relation homme-machine élargie et repensée. Dans cette perspective, les systèmes intelligents ne se limiteraient plus seulement à la simple exécution des actions demandées par un opérateur. Ils permettraient un travail collaboratif qui rendrait les actions et décisions des opérateurs plus efficaces et performantes tout en économisant les ressources mentales et physiques de ces derniers.

Pour ce faire, ces systèmes seraient dotés d'une connaissance accrue des situations à l'aide notamment de différents moyens de perception et d'analyse (état des opérateurs, interactions, prédiction des intentions des acteurs, situations tactiques de combat, etc.). Cette capacité permettrait aux systèmes d'apprendre des situations rencontrées, de s'adapter en conséquence et de partager les informations pertinentes afin d'apporter une aide à la prise de décision et à la planification des opérations. Pour garantir un haut niveau de performance,

<sup>13</sup> <https://man-machine-teaming.com/>

gage du succès des missions, ce système aérien cognitif intégrerait également de nouvelles modalités d'interactions plus naturelles et adaptées aux situations rencontrées par les opérateurs.

### 3.1.2 Surveillance de l'environnement opérationnel



L'interprétation automatique des données issues de capteurs, le passage du signal à une information sémantique exploitable en maîtrisant conjointement les contraintes de temps, de volume, de ressources et de qualité de l'information produite reste un défi. L'automatisation complète ou partielle de l'interprétation peut être déclinée selon plusieurs axes, relevant : du traitement interactif des données, de l'allocation dynamique de ressources, de la décision séquentielle, de la planification dans l'incertain, de la vision active, de la qualité de service des systèmes perceptifs, etc.

Pour être utilisable en pratique, l'interprétation automatique doit donc être polyvalente pour répondre simultanément à une diversité de besoins. Elle doit s'adapter à l'évolution temporelle des données et gérer les interruptions, les défauts de qualité des données source ou les modifications de mission. Elle doit exploiter efficacement des ressources potentiellement limitées (car souvent embarquées).

La temporalité de l'interprétation est aussi une problématique centrale. Elle intervient au niveau des données elles-mêmes (flux vidéo) sur lesquelles reposent des descriptions spatio-temporelles du contenu de la scène (détection-pistage-reconnaissance d'objets ou de personnes, description de comportement...). Les modes d'interprétation peuvent être séquentiels ou itératifs incluant très souvent l'humain.

La mise en rapport des données extraites des systèmes de perception avec d'autres types de « données » à portée locale (données obtenues par des moyens capteurs *in situ* comme des vidéos de surveillance, des photos, des informations textuelles) suppose notamment de travailler sur un référentiel géographique tridimensionnel partagé. L'intégration des données dans un référentiel 3D nécessite de disposer de données à la fois structurées et normalisées, suffisamment précises, complètes et actuelles, neutres et homogènes, sur une couverture géographique étendue.

### 3.1.3 La perception active et distribuée



L'analyse de grandes scènes dynamiques ne se réduit pas à la simple collection de données. Elle intègre en effet des dimensions propres à la multi-modalité (plusieurs capteurs et plusieurs traitements associés sont exploités simultanément), à l'intégration spatio-temporelle des informations multi-sensorielles dans des modèles de représentation sur différentes échelles de temps ainsi qu'à l'interprétation pour relier la perception réalisée à une prise de décision.

Les systèmes de perception sont souvent des systèmes contraints par différents facteurs (puissance des calculateurs, environnement, communications, etc.). On cherche aujourd'hui à en faire des systèmes intelligents, doués d'une capacité élevée d'analyse et de décision, à la fois pour comprendre l'environnement dans lequel ils évoluent mais aussi pour commander le mouvement des vecteurs qui les intègrent. On accède ainsi à des informations de manière plus active et les traitements peuvent être orientés par les finalités.

Le processus de perception dans son ensemble doit intégrer les besoins d'extraction de l'information, de robustesse des traitements notamment ceux exploitant de l'apprentissage statistique et leur contrôle par l'explicitabilité des décisions algorithmiques.

Les représentations pour l'intégration « cognitive » des informations doivent permettre d'analyser des variations temporelles : détection de changement, caractérisation des phénomènes et des comportements par l'exploitation

de séries de signaux temporels et une analyse prédictive (exemple : prédiction de conditions atmosphériques, ionosphériques et d'environnement ionisant spatial), ceci en mettant à profit des architectures d'apprentissage.

Les interactions de l'opérateur avec le système de perception influencent également beaucoup la dynamique de l'acquisition des informations et leur structuration, pour des besoins d'apprentissage (interactif) et de validation, mais aussi pour l'interrogation en langage naturel, la fouille de données, l'aide à la décision, et la visualisation informative des résultats (*visual analytics*).

### 3.1.4 Systèmes de défense multi-vecteurs



Le déploiement de ces systèmes robotiques en environnement naturel nécessite la mise en œuvre de moyens logiciels de planification des actions nécessaires à la réalisation des missions à partir de connaissances a priori acquises sur l'environnement.

Les fonctions de surveillance, à des fins de sécurité notamment, reposent sur des systèmes de capteurs qui peuvent être mobiles (embarqués sur des plates-formes mobiles disposant d'une « certaine » auto-nomie, de nature différente et organisés de manière dynamique).

Les moyens logiciels permettent d'organiser l'enchaînement des actions pour tenir compte des objectifs de la mission et des contraintes à satisfaire. Cette planification s'effectue en ligne pour adapter le comportement aux variations de l'environnement ou à des aléas. Elle doit intégrer la détermination des mouvements continus souvent contraints des engins qui est réalisée par des méthodes probabilistes de recherche de chemins ou des méthodes de diffusion. D'une manière générale, le problème de la planification référence à tout ce qui permet de définir des plans d'action pour les acteurs du dispositif à partir d'objectifs de haut niveau : la modélisation des problèmes ainsi qu'aux techniques de résolution permettant d'optimiser des plans de tâche multi-acteurs prenant en compte de multiples contraintes que ce soit la prise de risque, l'efficacité, la stabilité, la robustesse des plans.

Dans des missions de surveillance, la perception est souvent distribuée. L'exploration de l'espace, l'acquisition d'informations, la représentation des connaissances à différents niveaux sémantiques et leur interprétation peuvent être réalisées sous une forme orientée par les objectifs, ces derniers variant souvent dans l'espace et dans le temps.

La fonction de perception peut s'envisager à travers un modèle intégrant les processus de cognition associés (répartis, travaillant sur des représentations internes et externes, exploitant des connaissances liées à l'environnement). Les questions sous-tendues de perception mutuelle, d'égo-localisation distribuée et robuste, de mobilité coordonnée sous les contraintes physiques liées à l'environnement, de traitement embarqué, de communication et d'inter-visibilité sont également à traiter dans ce cadre.

La coordination et les interactions de ces systèmes robotiques entre eux et/ou avec l'opérateur constituent également des dimensions essentielles dans leur déploiement. Elles posent des problèmes de communication, de distribution des fonctions, d'autonomie ajustable, de commande coordonnée ainsi que d'ergonomie cognitive dans la conception des interfaces homme-système.

Enfin, il conviendrait aussi de s'intéresser à la justification des prises de décision du système artificiel lors des calculs de plan, aux processus de validation, ainsi qu'à l'embarquabilité de tels algorithmes qui devront s'exécuter en temps réel directement au sein des acteurs concernés.

### 3.1.5 Prévention des risques et sécurité civile



Les données issues des services de police et de gendarmerie, de la justice, de la sécurité civile et de la sécurité routière regroupées avec d'autres sources de données géolocalisées (caméras de surveillance, informations géospatiales et météorologiques, trafic, etc.) pourraient permettre de générer des alertes, prédire des risques, procéder à des analyses sur différentes échelles de temps et d'espace et ainsi permettre aux forces de l'ordre ou de la sécurité civile d'intervenir de manière efficace.

Dans un tel contexte *Big Data*, plusieurs types de traitement de données par les techniques de l'IA pourraient être utilisés. Parmi celles-ci, les techniques de classification par des modèles d'apprentissage profond distribué pour aider à prédire par exemple des points chauds (lieux où la probabilité d'évènement est élevée). La criminologie prédictive qui étudie les façons dont les groupes de données interagissent pour comprendre des « scènes » dynamiques constitue également un cadre d'usage pour les réseaux LSTM.

L'apprentissage propose d'une manière générale une base d'outils non seulement pour ce qui concerne le traitement des données importantes, mais aussi dans le processus d'extraction de connaissances à partir de ces données. Cette extraction repose sur des solutions basées sur la reconnaissance des modèles comportementaux (un comportement malveillant par exemple) et sur l'analyse de grands volumes de données en temps réel pour la détection d'anomalies. Détecter un comportement malicieux dans le contexte de la cybersécurité revient à analyser des séries temporelles multivariées, autrement dit, des données provenant de plusieurs sources et ordonnées dans le temps, ceci afin de mettre en évidence des motifs d'attaques.

Une première approche concerne les méthodes d'apprentissage supervisé. Ces méthodes requièrent des flux de données d'entraînement annotées. Ces méthodes d'apprentissage montrent généralement une grande capacité prédictive : elles sont capables de généraliser à partir des motifs des attaques connues et de reconnaître des attaques similaires aux menaces « vues » pendant la phase d'entraînement. Toutefois, ces approches requièrent la collecte et l'étiquetage manuel des échantillons sur un volume représentatif et donc conséquent de données, et de disposer d'une base de motifs pour les phénomènes observés.

Une autre approche viserait à regrouper les flux de données analysés par similarité (une similarité à plusieurs dimensions). Cette approche met en évidence la formation de plusieurs groupes de comportements. Les groupes sous-représentés (avec peu d'éléments ou une faible densité) désignent un ensemble de comportements statistiquement anormaux, se basant sur l'hypothèse que ces comportements interviennent rarement. Afin de préciser la sémantique propre à chaque groupe de comportements anormaux (illégal, suspicieux, malveillant, etc.), il est nécessaire d'intégrer des informations contextuelles. Dans ce cadre des réseaux RNN pourraient être envisagés.

## 4. IA@ONERA : quels axes stratégiques ?

L'identification d'un certain nombre d'axes stratégiques propres à la recherche en IA, dans sa diversité, est nécessaire à la lisibilité des objectifs de recherche de l'ONERA.

Dans ce qui suit, plusieurs axes de recherche structurants sont proposés. Ils concernent à la fois le développement des connaissances dans des domaines « cœur » de l'IA ainsi que le développement des méthodes pour des applications dans lesquelles l'IA joue un rôle fondamental.

Ceux-ci sont définis en tenant compte d'une part de la dynamique de recherche actuelle ainsi que celle envisagée dans le cadre notamment des feuilles de route de l'ONERA et d'autre part du potentiel de chercheurs mobilisable. Ne sont évoqués que les axes qui sont les plus pertinents au regard de certains enjeux du domaine ASD, en particulier ceux cités dans la partie 3 de ce document.

Le contexte national et international dans lequel ils seront développés est précisé afin de situer les perspectives de soutien financier et de collaboration.

### 4.1 Recherches concernant le « cœur » de l'IA

Les applications des techniques de l'IA ouvrent des perspectives et des possibilités tout à fait nouvelles pour les sciences et ses applications au domaine ASD, considérant que ces dernières s'appuieront désormais autant sur l'analyse de données que sur la modélisation mathématique ou la simulation.

L'ONERA a engagé des travaux qui portent sur des aspects fondamentaux des techniques de l'IA. Une structuration et une amplification de ces activités, notamment à travers des partenariats, sur des thématiques définies ci-dessous sont jugées stratégiques compte-tenu des enjeux relatifs au domaine ASD. Ces thématiques sont relatives à trois aspects de l'apprentissage machine : (1) la maîtrise des composants logiciels à base d'apprentissage automatique, (2) la modélisation hybride physique-données, (3) l'apprentissage distribué.

#### 4.1.1 L'intégration de représentations symboliques et de méthodes formelles aux modèles avec un apprentissage automatique

La valeur d'un algorithme d'apprentissage machine dépend avant tout de sa robustesse, de sa faculté à supporter des données du monde réel (incohérentes, éparées) et sa capacité à passer à l'échelle. Tout un ensemble de métriques sont utilisables pour la mesure de la robustesse de la fonction de classification ou de régression. Cette question de la robustesse des algorithmes et la stabilité de la prédiction est abordée notamment par l'agrégation d'estimateurs sous forme d'une moyenne pondérée des estimations (en régression) ou d'un vote à la majorité (en discrimination). Ces principes sont à la base des méthodes de *boosting*, de *bagging* et des forêts aléatoires. Cependant, malgré les performances montrées par les algorithmes d'apprentissage machine, ceux-ci se contentent de détecter des corrélations sans les « expliquer ».

Au-delà de la performance même des algorithmes, un enjeu majeur dans le développement des architectures d'apprentissage est de produire des algorithmes dont le fonctionnement soit interprétable et au-delà explicable. L'interprétation peut porter sur : (1) les données (fournir une trace d'exécution et des analyses de robustesse face aux exemples contradictoires), (2) l'architecture du réseau (la représentation des données dans le réseau) ou (3) la décision produite par le réseau.

Le but de l'interprétation est de décrire les éléments internes d'un système et de son fonctionnement d'une manière compréhensible par l'homme (à commencer par un expert). Une explication peut être jugée « plus complète » quand elle permet d'anticiper le comportement du système dans le plus de situations possibles. L'analyse des structures des réseaux de neurones se fonde par exemple sur des méthodes comme : *Local Interpretable Model-Agnostic Explanations* (LIME), *Layer-Wise Relevance Propagation* (LRP) ou des analyses de sensibilité. L'interprétation peut également reposer sur des méthodes de visualisation comme celles d'analyse des perturbations (occlusions). Cette interprétation du fonctionnement des réseaux de neurones peut aussi mettre à profit des techniques d'interaction homme-système en langage naturel.

A partir de ces différentes interprétations, on peut envisager de produire des explications sur la base d'une représentation du fonctionnement du réseau sous une forme symbolique hiérarchique et de méthodes de raisonnement exploitant les méthodes formelles, domaine dans lequel l'ONERA dispose d'une expertise.

Les travaux engagés à l'ONERA dans le domaine de l'explicabilité de l'IA s'adressent à des enjeux portant sur la maîtrise des traitements à base d'apprentissage notamment dans le cadre de leur utilisation dans des systèmes critiques.

#### 4.1.2 La modélisation hybride robuste, combinant des modèles physiques avec modèles d'apprentissage automatique

La modélisation basée sur les données peut être considérée comme une approche de la modélisation axée sur l'utilisation de l'apprentissage dans la construction de modèles physiques pour compléter ou remplacer les modèles « basés sur la connaissance ». On apprend alors la physique par l'expérience ! Les modèles par apprentissage permettent également d'atteindre des modèles de prédiction qui sont généralement plus rapides que la résolution des équations décrivant un modèle physique. L'apprentissage peut également servir à la « découverte » des structures d'un modèle.

L'un des principaux problèmes dans ce cadre porte sur la « maîtrise » des données. La construction des ensembles de données et de descripteurs robustes est essentielle à la maîtrise des performances d'un algorithme de prédiction ou de régression basé sur des réseaux de neurones ou des réseaux bayésiens. L'apprentissage, dans ce cadre, peut s'envisager sous différentes formes : par l'assimilation des données plus ou moins partielles et/ou bruitées, pour en inférer des variables d'intérêt des phénomènes observés. On peut enrichir les techniques d'apprentissage en intégrant des contraintes physiques de manière à ce que certaines lois/principes de conservation physiques soient naturellement encodées dans la méthode. Le processus d'apprentissage peut aussi être guidé par le transfert de politique. Les réseaux *end-to-end* sont possiblement en mesure de capturer des phénomènes intermittents et multi-échelles tout en respectant d'importants invariants du système, etc.

Les données restent souvent une denrée rare en physique, ceci est encore plus vrai pour des modèles à grand nombre de variables. Des simulations à différents niveaux de fidélité peuvent bien évidemment permettre de produire des données utilisables dans un apprentissage par renforcement et/ou par transfert.

La « modélisation basée données » ouvre des perspectives importantes que nous souhaitons explorer. On cherchera dans le développement de nos travaux plus particulièrement à exploiter des modèles basés données pour produire des modèles hybrides (physique-données) mais aussi à extraire des concepts physiques simples à partir de données expérimentales sans disposer de connaissances préalables supplémentaires (le réseau identifie les paramètres physiquement pertinents, exploite les lois de conservation pour faire des prédictions !). On peut envisager également l'exploitation des principes de l'interprétabilité des architectures d'apprentissage pour inférer une structuration de modèles de phénomènes physiques et pour obtenir des informations d'ordre conceptuel sur des phénomènes complexes.

#### 4.1.3 La distribution de modèles d'apprentissage automatique dans des systèmes multi-agents ou dans des grands systèmes

L'apprentissage automatique distribué fait référence à des algorithmes d'apprentissage multi-nœuds généralement conçus pour améliorer la robustesse d'une prédiction, augmenter la rapidité des traitements, s'adapter à l'échelle des données d'entrée, à la précision d'une approximation, etc. Pour les grandes tailles de données d'entrée, beaucoup d'algorithmes peuvent réduire considérablement l'erreur d'apprentissage. L'apprentissage automatique distribué conduit généralement à des décisions possiblement plus robustes par agrégation ou par répartition de données par classes plus homogènes.

La structuration des architectures de systèmes d'apprentissage profond distribués sur la base de modèles mettant en œuvre différents schémas de communication reste toutefois une question de recherche ouverte. Aujourd'hui, les activités de l'ONERA dans ce domaine restent limitées à l'exploitation de plates-formes logicielles telles que Spark MLlib.

Nous devons à l'avenir amplifier l'effort de recherche sur les modèles d'apprentissage distribué pour répondre aux enjeux du traitement *Big Data* du domaine ASD. Notons que cette recherche pourra bénéficier de l'expertise importante que nous avons sur les architectures des machines (*multi et manycore*) et sur le calcul réseau.

Un autre aspect voisin de l'apprentissage artificiel distribué est à considérer : celui concernant les systèmes multi-agents autonomes. Ce dernier est représentatif de nombreux problèmes de robotique collaborative ou de

collaboration homme(s)-système(s). Dans de tels systèmes, les agents, dotés de comportements « initiaux », doivent souvent apprendre de nouveaux comportements en ligne de sorte que les performances du système multi-agents dans son ensemble s'améliorent ou s'adaptent progressivement.

L'apprentissage par renforcement distribué, dans lequel les agents peuvent agir de manière coopérative, concurrentielle ou présenter des comportements neutres, est une voie à explorer. Pour gérer des systèmes de grande dimension, les techniques s'appuyant sur la théorie des jeux et de la décision distribuée peuvent être utiles pour modéliser et optimiser le comportement global du système. L'apprentissage distribué constitue un axe de développement pour nos travaux dans le domaine de la décision.

## 4.2 Méthodes de l'IA pour les applications dans le domaine ASD

Les axes de développement définis ci-après s'inscrivent pour la plupart dans des dynamiques existantes mises en place à travers des projets de différentes natures (internes, collaboratifs, etc.). Ils reposent sur un potentiel important en termes d'acquis et de nombre de contributeurs. Les activités actuelles sur ces différents axes peuvent être considérées comme des activités de premier plan au niveau international reposant sur des moyens logiciels et expérimentaux relativement uniques.

Ces axes offrent par ailleurs des perspectives de recherche importantes et s'adressent à des enjeux identifiés et largement partagés entre les acteurs du domaine ASD. Leur définition fait et continue à faire l'objet d'échanges internes à l'ONERA ainsi qu'avec les tutelles, les partenaires de recherche et les acteurs industriels.

Notons que certains font l'objet de grands projets structurants définis dans le cadre des feuilles de route de l'ONERA, comme pour ce qui concerne la « perception augmentée » et les « systèmes robotiques » ou encore la « sécurité des systèmes autonomes ». D'autres font l'objet ou feront l'objet de projets de recherche internes.

Le développement de ces projets sera réalisé pour une part dans le cadre de projets collaboratifs selon les opportunités de soutien qui se proposeront.

### 4.2.1 Perception augmentée pour la compréhension de scènes dynamiques

Les traitements par les technologies de l'IA d'images multimodales et plus largement de sources multimodales géo-référencées constituent une feuille de route à part entière de l'ONERA (Axe 4 « Surveillance de l'environnement opérationnel en défense & sécurité et de l'environnement naturel » – Thème 2).

Celle-ci s'adresse d'une manière générale aux besoins de la compréhension de scènes (au sens générique du terme) dans des finalités duales (surveillance de l'environnement et renseignement). Elle vise en particulier à exploiter et développer le potentiel des techniques de l'IA dans le traitement automatique des données de l'environnement : fusionner et interpréter à différents niveaux sémantiques des données hétérogènes multivariées, détecter et analyser des changements, des événements, etc. Elle se matérialisera par une plate-forme logicielle qui permettra d'explorer des scénarios de traitement dans le cadre de la surveillance globale à l'échelle planétaire d'activités humaines et/ou de leurs conséquences environnementales.

### 4.2.2 Robots, drones et systèmes intelligents pour la sécurité et la défense

Le mise en œuvre de systèmes autonomes (terrestres, aériens, maritimes) dans des missions de surveillance de l'environnement constitue également le cadre d'une feuille de route de l'ONERA (Axe 5 « Robots, drones et systèmes intelligents pour la défense et la sécurité » – Thème 1)

Dans cette feuille de route, les problèmes de première importance sont : (1) la planification des activités (mouvement, perception, transmission de données, etc.), (2) la coordination des robots à différentes échelles spatiales et temporelles pour réaliser une mission de manière robuste, (3) la construction de modèles de l'environnement et la navigation coordonnée, ainsi que (4) la supervision voire les interactions avec des humains.

Les techniques de l'IA qui seront plus spécifiquement développées dans ce cadre sont relatives à la navigation individuelle et coordonnée des engins et à la planification robuste de leur mission. Parmi les techniques considérées, celles de navigation et de localisation simultanée (SLAM) à haut niveau « sémantique » reposant sur l'exploration et la segmentation dynamique de l'environnement feront l'objet de développements importants.

La coordination dynamique des activités des systèmes multi-robots par la prise de décision séquentielle coopérative dans des conditions d'incertitude par des politiques apprises sur des Dec-POMDP par exemple feront

l'objet de travaux qui viendront compléter ceux sur les méthodes plus formelles de planification réactive, prédictive et répartie sous contraintes temporelles et les problèmes de décision séquentielle dans l'incertain.

### 4.2.3 Sûreté et sécurité des systèmes autonomes

Le thème de la sécurité et de la sûreté des systèmes autonomes à base d'IA fera l'objet d'un investissement particulier à l'ONERA. Il constituera le cadre principal pour nos travaux sur la maîtrise des composants logiciels IA critiques pour la mobilité et la perception, qui fait l'objet d'un thème central pour la feuille de route ONERA de l'axe 3 « Efficacité, sécurité et certification des systèmes de transport aérien ».

Certains des développements relatifs à ce thème sont envisagés dans le cadre de l'Institut 3IA ANITI et du projet DEEL de l'IRT Saint Exupéry (voir section 4.3 de ce document). D'une manière générale, il s'agira de fournir un ensemble de méthodes visant l'interprétabilité des modèles d'apprentissage automatique (incluant ou pas des phases de renforcement) et l'explicabilité des décisions locales et globales. Celles-ci s'appuieront notamment sur des représentations symboliques ou graphiques ou encore sur des techniques de dialogue homme-système en langage naturel. Leur potentiel sera qualifié par un ensemble de métriques de la robustesse des algorithmes de classification.

### 4.2.4 Perception artificielle cognitive & perception distribuée

La perception artificielle mono ou multi-capteurs peut être envisagée d'une manière générale sous la forme d'un ensemble de processus « cognitifs » réalisant les fonctions d'acquisition d'informations, d'attention, de mémorisation dans une structure de connaissances sur laquelle s'opèrent différents raisonnements visant à produire une action. La perception devient alors active, les fonctions cognitives travaillant conjointement pour intégrer la connaissance et créer une interprétation du « monde » à partir d'une exploration de l'espace.

La fonction de perception est souvent réalisée par un ensemble de capteurs. L'acquisition des informations, leur mémorisation, la représentation des connaissances à différents niveaux sémantiques et leur interprétation peut être placée dans le cadre de la « perception distribuée ».

L'objectif de la recherche sera de considérer la fonction de perception à travers les processus de cognition distribués (répartis, travaillant sur des représentations internes et externes, exploitant des connaissances liées à l'environnement). Les questions sous-tendues de perception mutuelle, d'égo-localisation distribuée et robuste, de mobilité coordonnée sous les contraintes physiques liées à l'environnement, de traitement embarqué, de communication et d'inter-visibilité seront à traiter dans ce cadre.

Cette recherche s'inscrit dans le prolongement de projets en cours sur la navigation visuelle multi-drones (projet de recherche GUIMAUVE). Ils sont inscrits également dans la feuille de route ONERA de l'axe 5 : « Robots, drones et systèmes intelligents pour la défense et la sécurité ») dont il a été fait mention section 4.2.2. Ils présentent également un intérêt pour le renseignement d'origine électromagnétique dont il est question juste après (section 4.2.5).

### 4.2.5 Apprentissage automatique pour le renseignement d'origine électromagnétique

Le projet, qui serait proposé uniquement comme un projet interne à l'ONERA, porterait sur l'application des techniques de l'apprentissage machine au REOM à des fins de classification robuste. Il intégrerait notamment plusieurs dimensions de la « perception active » : attention artificielle pour la détection dite de la saillance, la compréhension de scènes, l'attention axée sur les objectifs de détection, ainsi que la question de la configuration du système pour améliorer les performances de détection du spectre.

Le système d'apprentissage cherchera d'une part à savoir comment optimiser et configurer ses ressources matérielles afin de maximiser le nombre de signaux importants découverts dans des environnements RF difficiles et d'autre part à adapter automatiquement la réceptivité des signaux aux caractéristiques de signal pour discerner et caractériser les signaux dans les spectres encombrés. Il visera également à apprendre à reconnaître un émetteur spécifique sur la base de l'empreinte RF unique conférée naturellement par les imperfections matérielles de cet émetteur.

#### 4.2.6 Application de l'apprentissage par renforcement à la commande et synthèse de comportements

Ce projet vise à développer des techniques de commande fondées sur l'apprentissage par renforcement. Ces techniques, similaires à celle de la commande optimale, offrent un potentiel important pour la commande de systèmes continus et hybrides à comportement complexe et/ou évoluant dans des environnements dynamiques.

Nous considérerons en particulier une nouvelle classe d'asservissement visuel dit direct dans le cadre de la navigation autonome d'engins robotisés par caméra embarquée. Ce type d'asservissement consiste à travailler sur l'ensemble de l'image en utilisant directement l'information photométrique et non plus des primitives visuelles particulières. Il peut potentiellement conduire à une nette augmentation de la précision de positionnement, des algorithmes d'apprentissage profond pouvant être là utilisés pour estimer la configuration du mobile.

Ce type de commande présente un domaine de convergence plus local et une faible robustesse aux modifications de l'environnement (ex : une variation de l'intensité lumineuse). Elle ne fonctionne correctement que si le réseau est entraîné avec des images (pouvant être simulées) proches de la scène considérée. Pour aller plus loin, on abordera les questions du transfert d'apprentissage (TA) dans les algorithmes d'apprentissage par renforcement (AR) pour faire face à des environnements complexes, et d'apprentissage hiérarchiques, où les politiques sont organisées dans une hiérarchie de sous-tâches. L'objectif du TA dans ce cadre est de construire automatiquement une hiérarchie des compétences qui puisse être réutilisée pour résoudre plusieurs tâches.

Ces recherches pourraient faire l'objet d'un projet de recherche interne en 2020. Il pourrait être inscrit aux collaborations envisagées dans le cadre du 3IA ANITI.

#### 4.2.7 Modèles d'intégration d'informations multi-sources

Les travaux envisagés concernent la modélisation des processus psycho-cognitifs impliqués dans le pilotage des aéronefs (hélicoptères notamment) dans le cadre de missions exigeantes pour parvenir à la définition d'un « pilote virtuel ».

Le modèle explicatif visé aurait capacité à représenter les processus d'intégration observés chez l'homme par des informations multimodales et multi-sources en provenance de la « scène » pour répondre par des actions adaptées à cette situation. On cherchera à construire une architecture combinant différents types de réseaux de neurones permettant de reproduire notamment des mécanismes attentionnels ainsi que ceux de la prédiction.

Le « pilote virtuel » serait éprouvé en modifiant certains paramètres d'une situation donnée. Une phase de renforcement dans l'apprentissage permettrait d'individualiser les comportements et rapprocher les comportements prédits par le modèle de celui d'un pilote réel.

Cette recherche sera initiée dans le cadre d'un projet ANR JCJC en cours et proposée au financement de la DGAC. Nous envisagerons également une collaboration avec l'Institut 3IA de Sophia sur ce thème.

#### 4.2.8 Simulation physique hybride, prédiction du comportement et aide à la compréhension de phénomènes

L'ambition est ici de développer des méthodes avancées de pronostic et de diagnostic destinées à la maintenance des aéronefs en exploitant des moyens de simulation physique hybride (une forme de jumeau numérique) et des techniques d'IA pour la prédiction du comportement. La simulation physique par le jumeau numérique servirait à la mise au point et qualification en partie dans leur robustesse des algorithmes de détection par la maîtrise des jeux de données synthétiques issues des plans d'expérience réalisés.

Pour des machines fonctionnant dans un environnement dynamique, les données de surveillance recueillies sont soumises à de hauts niveaux d'incertitude et d'imprévisibilité. Par ailleurs, les connaissances et la compréhension sur les processus de détérioration sont limitées. Prédire le comportement d'une machine en service (comme un moteur d'avion) est un défi compliqué tout comme la détection robuste et fiable d'anomalies.

La détection d'anomalies utilise généralement des techniques d'apprentissage statistique non supervisées et non paramétriques (SVM et forêts aléatoires notamment). Nous explorons comment l'apprentissage en profondeur peut être utilisé pour résoudre certains problèmes importants de l'analyse de données volumineuses en flux, notamment pour l'extraction de schémas complexes et l'indexation sémantique.

Nous considérerons également l'usage des réseaux bayésiens statiques et dynamiques dans le domaine de l'analyse des dégradations pouvant affecter un système dynamique, les réseaux bayésiens dynamiques étant

utilisés pour tenir compte de la dynamique du système et prédire son comportement en fonction de son état actuel et d'autres variables ou contraintes exogènes.

Ce projet sera inscrit dans le cadre de la coopération envisagée avec les laboratoires CNRS (I2M et IMS) de l'Université de Bordeaux.

#### 4.2.9 Planification des missions spatiales

La planification de missions liées à l'observation spatiale doit s'appuyer sur des méthodes formelles permettant d'appréhender la complexité du problème (fonctions du système, contraintes, missions) et sur des techniques de décision performantes applicables à des problèmes de grande dimension.

L'objectif de ce projet sera de développer et d'éprouver sur la base de scénarios réalistes un ensemble de techniques de planification sous contraintes pour l'optimisation dans l'espace et dans le temps de l'exploitation des ressources (mesures, mouvements, communications, etc.). Techniquement, on explorera : la planification hiérarchique, celles de ressources discrètes et continues, les méthodes de recherche locales incrémentales rapides, les stratégies d'optimisation globale, l'explicabilité des résultats et des incohérences. Les aspects décision dans des contextes dynamiques incertains (évolution des problèmes et des données) portant sur des systèmes à grande voire très grande échelle, et de prédiction robuste seront également traités.

Le cadre privilégié pour ces travaux est celui des applications à l'observation spatiale par des constellations de satellites agiles. À ce titre, des collaborations avec le CNES (dans le cadre d'un PIC - programme d'intérêt commun) ainsi qu'avec l'Institut 3IA AMITI sont envisagées. D'autres domaines d'application seront concernés par ces travaux : ceux de la planification des missions multi-robots, de la production aéronautique et de sa logistique.

### 4.3 IA@ONERA : insertion dans l'environnement académique et le domaine ASD

#### 4.3.1 Quels objectifs

La France a mis en place en 2018 un plan national sur l'intelligence artificielle qui a été décliné du rapport Villani : <https://www.numerique.gouv.fr/ai-for-humanity>. Ce plan identifie quatre grands domaines d'application privilégiés pour l'IA en France : défense & sécurité, transport et mobilités, environnement, santé. Il concerne donc à différents titres les activités de l'ONERA et prévoit notamment :

- L'organisation de la recherche envisagée dans ce plan – sur le territoire national et en Europe –, la mise en place de hubs de recherche en IA à partir d'un programme national coordonné par l'INRIA en lien avec les universités et les organismes de recherche. L'objectif est de faire émerger « un réseau d'instituts dédiés localisés dans quatre ou cinq endroits en France », accompagné par un programme de chaires individuelles. L'un des éléments clés du programme réside dans la constitution d'un réseau de 3IA (instituts interdisciplinaires pour l'intelligence artificielle).

Quatre projets ont été présélectionnés par un jury international, le 6 novembre 2018 : les quatre 3IA présélectionnés sont « MIAI@Grenoble-Alpes » à Grenoble, « 3IA Côte d'Azur » à Nice, « PRAIRIE » à Paris et « ANITI » à Toulouse.

L'État a réservé une enveloppe de 100 M€ pour soutenir les 3IA jusqu'en 2022, enveloppe provenant du PIA et du budget du ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. S'y ajouteront 100 M€ apportés par les établissements publics qui portent ces instituts, ainsi que 100 M€ apportés par les partenaires industriels.

- Un renforcement du soutien de l'ANR aux projets de recherche en IA est également prévu. En 2018, ce sont 61 projets qui ont été soutenus à hauteur de 27 M€. L'objectif visé est de doubler ce nombre d'ici à 2022 en fléchissant spécifiquement un renforcement de 100 M€ de la dotation de l'agence.
- Les acteurs institutionnels et industriels seront sollicités pour s'engager dans une politique de partage des données notamment dans les transports, la santé, l'environnement. Cette politique est déjà engagée dans le secteur de la santé. Elle doit être accélérée au bénéfice des besoins en matière de cartographie – et plus

largement de modèles d'environnement pour le transport « intelligent » et autonome –, de l'agriculture, de l'aménagement du territoire, etc. Cette politique de partage pourrait être développée dans le cadre d'un partenariat public-privé.

- Une stratégie de R&D nationale et européenne dans les quatre grands domaines d'application privilégiés de l'IA.
- L'État prévoit une enveloppe de 65 M€ pour « mobiliser tous les leviers » disponibles, au-delà des 3IA : 20 M€ pour renforcer les chaires industrielles et les LabCom, laboratoires communs entre la recherche publique et des PME ou ETI, l'objectif étant de porter leur nombre à « au moins 50 » dans le domaine de l'IA ; 10 M€ de plus pour les Instituts Carnot ; et 35 M€ pour soutenir l'initiative « Engage IA » lancée par les IRT (instituts de recherche technologique).
- Des « grands défis » en IA financés par le fonds pour l'innovation et l'industrie. Deux défis ont déjà été lancés en juillet par le Conseil de l'innovation, qui seront dotés chacun d'environ 30 M€ sur trois ans.
- Le développement de la coopération franco-allemande dans le domaine de l'IA. L'axe franco-allemand a été clairement mis en avant dans le plan national. Ce cadre sera privilégié pour les questions de recherche à commencer par la coopération bilatérale franco-allemande. Les deux pays vont mettre en place un « programme bilatéral de promotion de la recherche fondamentale », dont l'un des appels à projets annuels portera sur l'IA et sera soutenu à hauteur de 3 M€ par an. Le couple franco-allemand prévoit également une mise en réseau de ses centres de compétences et de ses plates-formes technologiques, de porter en commun des projets de recherche à l'Europe et de « faire de l'IA un des points focaux de l'EIC » (Conseil européen de l'innovation).
- L'implication de l'ONERA dans les projets 3IA.

La mise en place d'instituts interdisciplinaires en d'intelligence artificielle (3IA) constitue un élément important du plan national pour l'intelligence artificielle. L'ONERA s'est impliqué dans deux des quatre propositions acceptées suite à la manifestation d'intérêt émise par l'ANR en juin 2018. Les quatre projets sélectionnés sont :

- Grenoble – « MIAI@Grenoble-Alpes » avec pour applications privilégiées : la santé, l'environnement et l'énergie
- Nice – « 3IA Côte d'Azur » pour des applications en biologie (traitement des données et modèles), et pour la ville et les territoires « intelligents »,
- Paris – « PRAIRIE », pour la santé, les transports et l'environnement
- Toulouse – « ANITI » avec pour applications privilégiées le transport, l'environnement et la santé.

L'ONERA est impliqué principalement dans le projet ANITI porté par l'Université fédérale de Toulouse et plus marginalement dans le projet « 3IA Côte d'Azur ». Ces instituts ont vocation à constituer une masse critique pour la formation, la recherche et l'innovation en intelligence artificielle et dans ses secteurs d'application. Ils visent à apporter des moyens propres et des apports industriels, deux tiers du financement des activités seront consacrés à des chaires.

Les développements de recherche proposés par l'ONERA dans le projet ANITI portent sur :

- (1) la robotique (génération de mouvement) – la perception du robot et l'intégration d'informations provenant de diverses sources pour diverses tâches complexes qui bénéficieront des techniques d'apprentissage.
- (2) la télédétection qui exploitera notamment les méthodes d'apprentissage automatique en combinant le traitement de différents types de données physiques pour réduire les incertitudes et améliorer la précision et l'exactitude des modèles. La maîtrise pour la certification des composants logiciels à base d'apprentissage pour l'autonomie des systèmes de transport constitue un autre axe de travail proposé.
- (3) les techniques de réduction de modèles avec incertitudes pour la conception multidomaines.
- (4) la maîtrise en vue de la certification des composants logiciels à base d'apprentissage pour l'autonomie des systèmes de transport.

Le 3IA Côte d'Azur privilégie des travaux sur les modèles bio-inspirés et le développement des territoires. Des discussions ont été engagées avec les porteurs du projet pour envisager la participation des unités de Salon de Provence à des projets qui pourraient concerner la télédétection pour la surveillance du littoral ainsi que la modélisation et l'analyse des activités cognitives des opérateurs de systèmes aéronautiques.

### 4.3.2 Collaboration nationale autour de grands projets

Un ensemble d'opérations de recherche dans des différents domaines de l'ASD se dessine autour de sujets scientifiques liés à l'apprentissage machine et la décision. Ces opérations de recherche se développeraient dans le cadre de collaborations précisées après.

#### – IRT Saint Exupéry

L'objet de la collaboration avec l'IRT Saint Exupéry porterait sur les problèmes d'interprétabilité et d'explicabilité des composants logiciels pour les systèmes critiques mettant en œuvre les techniques d'apprentissage statistique. Ces travaux s'inscriraient dans le cadre d'un projet baptisé DEEL, un projet de recherche collaboratif d'envergure qui implique des acteurs académiques membres de l'Université fédérale de Toulouse et industriels : Airbus, Thales Avionics, Thales Alenia Space, Continental, Safran et Renault. La contribution principale de l'ONERA concernerait l'IA hybride symbolique-statistique. À noter qu'une partie de ces travaux serait réalisée dans le cadre de l'institut 3IA ANITI.

#### – CNES

Le CNES et l'ONERA ont ambition de démontrer les apports potentiels des techniques de l'IA dans le développement de services exploitant l'imagerie haute résolution et les données géospatiales. Ce projet donnerait lieu à une plateforme hébergeant données et traitements. Elle permettrait le passage à l'échelle d'algorithmes d'IA appliqués à des données d'observation de la Terre, visant en particulier la production automatique de données 3D qualifiées. Les contributions de l'ONERA dans ce cadre seraient relatives au développement et au déploiement d'algorithmes d'apprentissage profond dans un contexte *Big Data*.

#### – DGA-ARTEMIS

L'ONERA inscrira une partie de ses développements sur les traitements *Big Data* dans l'initiative ARTEMIS de la DGA dont les infrastructures sont supportées par les sociétés Atos-Bull/Capgemini et le groupement Thales/Sopra-Steria. Ces infrastructures permettront de confronter les algorithmes de traitement à des données maîtrisées par la DGA pour faire mûrir jusqu'à des solutions utilisables par les forces une gamme complète de produits mettant en œuvre des avancées dans le domaine de l'IA.

L'ONERA pourrait en particulier mener des expérimentations sur des données de vol des avions et des drones en vue de la maintenance prédictive ainsi que sur des images satellitaires (radar/optroniques) pour la visualisation et l'aide à l'interprétation sémantique des images.

#### – CNRS-industrie aéronautique

L'ONERA souhaite développer des activités dans les secteurs de la maintenance et de la production des domaines aéronautique et spatial. Plusieurs projets de collaboration avec des laboratoires CNRS sont envisagés, en particulier avec les unités mixtes CNRS-Université de Bordeaux-ENSAM I2M (Institut de mécanique et d'ingénierie) et IMS (Intégration du matériau au système) ainsi qu'avec l'IRIT (Institut de recherche en informatique de Toulouse).

Dans ces différents cadres de collaboration, nous développerons des opérations de recherche portant pour une part sur l'exploitation des techniques d'apprentissage à la classification des défauts observés par des procédés de contrôle non-destructif et d'inspection visuelle. D'autre part, des travaux seront engagés sur les techniques de pronostic par apprentissage profond distribué, exploitant des jumeaux numériques.

Avec l'IRIT, nous conduirons des travaux en planification/conception conjointe de systèmes de fabrication et de logistique en considérant plus spécifiquement les problèmes liés à l'industrie aéronautique et au spatial.

### 4.3.3 Collaborations internationales

L'Europe a lancé au début 2019 un grand projet pour l'IA (*AI4EU – Artificial Intelligence for European Union*) et l'ONERA en est l'un des 79 membres. Ce projet sur 3 ans, piloté par Thales, est doté de 20 MEuros. Il a vocation à animer la communauté européenne en IA et développer des recherches dans des domaines comme l'explicabilité des algorithmes et de la certificabilité des systèmes à base d'IA.

Plusieurs axes de coopération internationale sont envisageables dans le domaine de l'intelligence artificielle.

#### – Allemagne

En premier lieu, l'axe franco-allemand. Ce dernier a été identifié comme stratégique dans le plan national. Comme la France, l'Allemagne a réservé une enveloppe conséquente (environ 30 MEuros) pour soutenir « de manière significative et durable » la recherche et le développement dans l'intelligence artificielle (IA)<sup>14</sup>.

Les deux pays se sont engagés en mai à « mettre en place un centre de recherche franco-allemand sur l'intelligence artificielle ». Au niveau national, la collaboration franco-allemande, soutenue par le fonds pour l'innovation et l'industrie qui sera à terme doté de 10 milliards d'euros, pourrait contribuer à des « projets importants d'intérêt européen commun ».

Dans ce contexte, une coopération avec le DLR et la TUM (Université technique de Munich) sur des thématiques comme les jumeaux numériques pour la prédiction au service de l'optimisation et de l'automatisation des processus de maintenance sera considérée.

La surveillance et l'analyse de l'activité des pilotes est également un thème sur lequel un projet de recherche s'appuyant le projet de « modèle d'intégration cognitive multi-source » est envisageable. L'opportunité de proposer un *Digital Innovation Hub*<sup>15</sup> sur les moyens expérimentaux du DLR, du NLR et de l'ONERA, destinés à l'analyse des interactions homme-système est une possibilité à explorer.

D'autres domaines pourraient être considérés dans ce cadre de collaborations avec les instituts Fraunhofer notamment ceux en relation avec la sécurité des systèmes cyber-physiques (thème identifié comme un axe stratégique de collaboration avec l'Allemagne<sup>16</sup>).

#### – Canada

La volonté d'un développement de coopérations entre la France et le Canada dans le domaine de l'IA s'est traduite par la création d'un G2IA. Le G2IA (Groupe international d'experts en intelligence artificielle) est une alliance qui vise à promouvoir une IA éthique et inclusive. En marge du prochain sommet du G7, qui se tiendra en France à l'été 2019, la France et le Canada organiseront à Paris une grande conférence mondiale sur l'IA,

<sup>14</sup> [https://www.ip.mpg.de/fileadmin/ipmpg/content/aktuelles/Outline\\_for\\_a\\_German\\_Artificial\\_Intelligence\\_Strategy.pdf](https://www.ip.mpg.de/fileadmin/ipmpg/content/aktuelles/Outline_for_a_German_Artificial_Intelligence_Strategy.pdf)

<sup>15</sup> <https://ec.europa.eu/digital-single-market/en/digital-innovation-hubs>

<sup>16</sup> [http://www.embedded-france.org/wp-content/uploads/2018/02/Rapport\\_CPS\\_Version\\_Publique\\_2017.pdf](http://www.embedded-france.org/wp-content/uploads/2018/02/Rapport_CPS_Version_Publique_2017.pdf)

qui visera à « établir un premier consensus scientifique sur les enjeux de l'intelligence artificielle » et à ouvrir des axes de coopération.

L'importance de l'activité aéronautique au Canada devrait nous inviter à considérer des sujets de coopération dans des domaines liés au développement des systèmes à forte autonomie comme les UAV ou les systèmes à autonomie renforcée. Des discussions ont été engagées avec le National Research Council sur le monitoring des pilotes et télé-pilotes d'UAV.

#### – USA

Les coopérations déjà actives avec la NASA pourraient être intensifiées. Celles avec NASA-Ames (*Intelligent Systems Division*) concernant l'IA portent sur la modélisation et l'analyse des interactions homme-système pour la conception des aides au pilotage des aéronefs. Elles exploitent les travaux que nous menons sur les réseaux dynamiques bayésiens pour la détection des risques de perte de contrôle des systèmes « intelligents ».

Les activités de l'ONERA sur les architectures pour la commande et la planification de missions multi-engins pourraient faire l'objet de coopérations avec la NASA/JPL. En particulier, les travaux réalisés sur les environnements de développement pour les systèmes embarqués critiques (basés sur des méthodes formelles) et la planification distribuée (supporté par la bibliothèque InCell) pourraient trouver dans cette coopération un espace de confrontation sur des missions spatiales telles que celles des programmes d'exploration planétaire. Une coopération sur ce même objet pourrait être envisagée avec l'Agence Spatiale Canadienne.

La participation de l'ONERA aux activités en IA pour les sciences spatiales dans le cadre du Frontier Development Lab<sup>17</sup> de NASA-Ames et de son réseau de partenaires internationaux serait également envisageable.

#### – Singapour

Singapour s'est doté d'un plan national en IA (\$150 million sur 5 ans) baptisé AI Singapore<sup>18</sup>. Ce plan vise en particulier à favoriser le développement des technologies de l'IA pour l'industrie et pour les services. Il comporte un volet recherche ambitieux qui sera mis en œuvre par la National Research Foundation (NRF). Nos collaborations avec Singapour en IA portent sur la planification d'activités pour les systèmes multi-robots. Elles s'inscrivent dans un projet commun entre la DGA et l'agence de R&D de la défense de Singapour (DSO). Elles pourraient être étendues à des problèmes de surveillance par des réseaux de capteurs dans le cadre du DSO AI Hub<sup>19</sup>.

---

<sup>17</sup> <https://frontierdevelopmentlab.org>

<sup>18</sup> <https://www.nrf.gov.sg/programmes/artificial-intelligence-r-d-programme>

<sup>19</sup> <https://www.straitstimes.com/singapore/room-to-work-on-robotics-and-ai-in-new-dso-complex>

## Abréviations

3IA	Instituts interdisciplinaires en intelligence artificielle
ADS-B	Automatic Dependent Surveillance-Broadcast
AM	Amplitude Modulation
AR	Apprentissage par renforcement
ASD	Aéronautique-Spatial-Défense
ATM	Air Traffic Management
BOA	Bulle opérationnelle aéroterrestre
C4ISR	Command, Control, Communications, Computers, Intelligence, Surveillance, Reconnaissance
CBLS	Constraint-Based Local Search
CGAN	Conditional Generative Adversarial Network
COMINT	Communications Intelligence
CNN	Convolutional Neural Network
CNS	Central Nervous System
	<i>ou</i> Combinaison de réseaux de confusion
CPU	Central Processing Unit
DEC-POMDP	Decentralized Partially Observable Markov Decision Process
DL	Deep Learning
DNN	Deep Neural Network
ECG	Électrocardiogramme
EEG	Électroencéphalogramme
ELINT	Electronic Intelligence
EMG	Électromyogramme
FDR	Flight Data Recorder
FM	Frequency Modulation
GAN	Generative Adversarial Network
GEOINT	Geospatial Intelligence
GIA	Groupe Intelligence artificielle (ONERA)
GNSS	Global Navigation Satellite System
GPU	Graphics Processing Unit
HiPOP	Hierarchical Partial Order Planning
HPC	High Performance Computing
IA	Intelligence artificielle
IAD	Intelligence artificielle distribuée
IR	Infrarouge
LDA	Analyse discriminante linéaire
LIME	Local Interpretable Model-Agnostic Explanations
LRP	Layer-Wise Relevance Propagation
LSTM	Long Short-Term Memory
M2CO	Maintenance pour le maintien en condition opérationnelle
MCO	Maintien en condition opérationnelle
MDO	Multi-Disciplinary Optimisation
MLRF	Multi-Label Random Forest
MLP	Multi Layer Perceptron
MKAD	Multiple Kernel Anomaly Detection
MOCAP	Motion Capture
MPC	Model Predictive Control
MRAC	Model Reference Adaptive Controller
MRO	Maintenance Repair and Overhaul
NC	Neuromorphic Computing
NIRSI	Near Infrared Spectroscopic Imaging
OSP	Over Subscription Planning

PDDL	Planning Domain Definition Language
PDM	Processus décisionnel de Markov
PID	Proportionnel, intégral, dérivé
POD	Proper Orthogonal Decomposition
POMDP	Partially Observable Markov Decision Process
POP	Partial Order Planning
PSP	Partial Satisfaction Planning
QC	Quantum Computing
QDA	Analyse discriminante quadratique
RBV	Return Beam Vidicon
REOM	Renseignement d'origine électromagnétique
RF	Radio-fréquence
RN	Réseau de neurones
RNN	Recurrent Neural Network
ROV	Remotely Operated Underwater Vehicle
S3P	Stochastic Safestand Shortest Path
SAR	Synthetic Aperture Radar
SCAF	Système de combat aérien futur
SDM	Shared Decision Making
SIGINT	Signals Intelligence
SLAM	Simultaneous Localisation And Mapping
SMT	Satisfaisabilité modulo-théories
SNARC	Stochastic Neural Analog Reinforcement Calculator
SOD	Smooth Orthogonal Decomposition
SRI	Stanford Research Institute
STIC	Sciences et technologies de l'information et de la communication
STRIPS	Stanford Research Institute Problem Solver
SVM	Support Vector Machine
SWIR	Short Wave Infrared Radar
TA	Transfert d'apprentissage
TDNN	Time Delay Neural Network
TNBN	Temporal Node Bayesian Network
TPU	Tensor Processing Unit
TSTN	Time Dependent Simple Temporal Networks
UAV	Unmanned Aerial Vehicle
UCAV	Unmanned Combat Air Vehicle
UGV	Unmanned Ground Vehicle
USV	Unmanned Surface Vehicle
UTM	Unmanned Traffic Management
VHF	Very High Frequency (radiodiffusion, télédiffusion ...)

---

Document élaboré à la Direction scientifique générale de l'ONERA, publié en mars 2020.