

Cartographie des ressources expertes en IA présentes au sein des laboratoires de recherche ligériens

Pierre Cointe

16 février, révision 4 juin 2018

<http://atlantisc2020.fr/>

1. Préambule

« *Artificial Intelligence : the field of research concerned with making machine do things that people consider to require intelligence. There is no clear boundary between psychology and AI because the brain itself is a kind of machine* » M. Minsky [1]

Contrairement aux Mathématiques et à la Physique, l'Informatique est une discipline scientifique jeune. Mais comme ses sœurs aînées, l'Informatique est constituée de sous disciplines parmi lesquelles l'Intelligence Artificielle (IA) et le Génie Logiciel (GL). Le premier laboratoire d'IA s'est constitué au MIT en 1959 sous la direction des professeurs [Marvin Minsky](#)¹ (l'auteur du best seller « The Society of Mind ») et [John McCarthy](#) (le père du langage Lisp). Leurs idées ont d'abord essaimé à Stanford (McCarthy) et Carnegie Mellon (Allen Newell and Herbert Simon) puis dans les principales universités américaines, européennes et japonaises. Elles ont également données lieu à la création de conférences prestigieuses comme IJCAI en 1969 et sa déclinaison européenne ECAI en 1974.

Dans les années 80, les travaux en IA ont suscité de nombreuses avancées dans les domaines des langages de programmation, des machines dédiées à un langage (Symbolics/Lisp, Xerox/Smalltalk, ERCIM/Prolog), des machines parallèles (la Connection Machine, le programme japonais de 5^{ème} génération FGCS), de la représentation des connaissances et des systèmes experts, du traitement de la langage naturel, de l'apprentissage automatique et des réseaux de neurones... Néanmoins à la fin du 20^{ème} siècle, principalement du fait du manque de puissance des processeurs matériels, les résultats obtenus n'ont pas été à la hauteur des espoirs soulevés par les zéloteurs de l'IA à tel point que de nombreux laboratoires de recherche ont renoncé à se revendiquer de cette sous discipline informatique « passée de mode ».

Cependant dans les années 2010, conformément à la conjecture de Moore, la puissance de calcul des processeurs n'a cessé de progresser alors que l'avènement des réseaux sociaux et de l'internet des objets a produit les grandes masses de données qu'il est devenu possible/nécessaire d'analyser relançant par la même les travaux sur l'apprentissage (Deep Learning), la représentation des connaissances (Web sémantique), l'algorithmique et la

¹ « In 1959, John McCarthy came to MIT from Dartmouth, and we started the MIT Artificial Intelligence Project. We agreed that the most critical problem was how minds do common sense reasoning. McCarthy was more concerned with establishing logical and mathematical foundations for reasoning, while I was more involved with theories of how we actually reason using pattern recognition and analogy » (Page 323 The Society of Mind [1]).

programmation distribués, le traitement automatique des langues, la robotique, l'informatique (représentation) musicale, ...

Par un effet de yoyo, l'IA revient aujourd'hui sur le devant des scènes médiatique et politique² alors même que la révolution numérique suscite de nombreuses transformations sociétales difficiles à appréhender. Très récemment, différentes études ont donné lieu à un état des compétences nationales en IA dont le rapport de synthèse France Intelligence Artificielle [2], le livre blanc Inria [3] mais également le rapport de l'association pas@aline [4]. Dans un même temps, l'institut INS2I du CNRS prépare la création d'un GDR IA [5] et le mathématicien/député Cédric Villani lance une consultation publique sur l'intelligence artificielle.

Le propos de ce texte est de dresser une cartographie des compétences en IA des laboratoires ligériens et très précisément des laboratoires en sciences du numérique définissant le premier cercle du programme régional Atlanstic 2020.

Pour réaliser cette carte, nous avons choisi le référentiel proposé dans le livre blanc Inria, référentiel déjà repris par de nombreux acteurs dont l'association pas@line et l'Institut Mines Télécom. Dans une première partie, nous analysons de manière quantitative « les forces recherche en présence » d'abord à partir des travaux des équipes de recherche (30 recensées appartenant aux 4 laboratoires LS2N, LERIA, LARIS et LIUM) ensuite à partir des lauréats aux appels à projet recherche Atlanstic 2020 (31 projets lauréats en 2016 et 2017 sur 80 soumis). Dans une seconde partie qualitative, nous mettons en exergue les travaux/chercheurs de notoriété internationale pouvant servir de référents ligériens. En conclusion nous livrons quelques pistes sur les actions à mener pour mobiliser une force IA Pays de Loire autour de deux à trois projets d'envergure.

² On parle même des intelligences artificielles !

2. Vision quantitative

Nous utilisons donc la taxinomie Inria qui propose les 8 thématiques suivantes : 1) Connaissances, 2) Apprentissage automatique, 3) Traitement du langage naturel, 4) Traitement des signaux, 5) Robotique, 6) Neurosciences & Sciences cognitives, 7) Algorithmique de l'IA et 8) Aide à la décision. Chacune de ces sous thématiques pouvant elle-même se décomposer en sous champs thématiques

3.1 Crible des 30 équipes de recherche LS2N, LARIS, LERIA et LIUM

Cette première grille de synthèse place en abscisse les 8 thématiques principales et en ordonnée les 30 équipes des 4 laboratoires étudiés. Nous utilisons l'acronyme des ces équipes, leurs noms complets sont développés en annexe1. A noter que dans le cas du « gros » LS2N l'ensemble des équipes est lui même organisé en 5 pôles.

Enseignants-chercheurs & chercheurs permanents		Labos du premier cercle Atlanstic 2020																											274			
		LS2N															LERIA	LARIS	LIUM													
		SDD			SIEL			RPC			SLS			CSS			20	43	26													
		51					34					28					35					37			20	43	26	274				
		19	7	15	8	2	10	11	8	6	9	9	5	6	7	6	5	7	6	4	8	12	5	13	9	11	15	15	13	13	13	274
Taxinomie Inria Livre blanc IA (pages 20-21)		Duke	TASC	SLP	ComBi	Méforbio	TALN	IPI	SIMS	PACCE	RoMas	ReV	ARMEN	OGRE	GDD	Galinette	NaoMod	STACK	AeLos	RIO	PSI	STR	IS3P	Commande	MOA / MO	ICLN / RR	SDO	ISISV	SFD	LST	IEIAH	
Connaissances		50%	10%			25%									25%	IDM										20%						
Apprentissage automatique		25%		10%	20%			10%	30%	20%																30%		25%	25%			
Traitement du langage naturel							50%																			20%			50%			
Traitement des signaux							30%	30%					50%															50%	25%			
Robotique											20%	50%	50%														25%					
Neurosciences, Sciences cognitives							10%	10%	10%																							
Algorithmique de l'IA			55%	50%	10%										50%	AD	30%	AC	20%	MA	10%			100%	30%							
Aide à la décision				15%	10%																								SF			
Pourcentage E/C et C du domaine de l'IA		75%	110%	60%	10%	30%	100%	60%	60%	10%	20%	50%	100%	50%	25%	30%	0%	0%	20%	0%	30%	40%	50%	0%	100%	100%	25%	50%	25%	100%	0%	
		14	8	9	1	1	10	7	5	1	2	4	5	3	2	2	0	0	1	0	2	5	3	0	9	11	4	8	3	13	0	
		82															20	15	13	130												
		44%															100%	34%	50%													

La partie supérieure de la grille recense, pour chaque équipe, le nombre d'enseignants-chercheurs et chercheurs permanents à l'exclusion des ingénieurs (permanents ou non permanents), des doctorants³ et des post-doctorants.

Pour chaque équipe nous indiquons le pourcentage de son implication dans chacune des thématiques (0% par défaut). La couleur vert foncé indique une forte visibilité de cette activité. La couleur bleue (sans pourcentage) indique une activité dans un domaine comme l'ingénierie des modèles (IDM), les multi-agents (MA), l'algorithmique distribuée (AD), sûreté de fonctionnement (SF) qui n'apparaît pas dans la taxinomie Inria mais qui aurait pu/du figurer.

La partie inférieure de la grille, livre les conclusions de l'analyse, c'est à dire, pour chaque équipe son pourcentage d'activité relevant de l'IA au sens du livre blanc Inria. On observera en particulier les résultats suivants :

- **130** des E/C permanents soit **47%** mènent des travaux dans le domaine de l'IA ;
- 6 équipes (indiquées par un fond gris) revendiquent de mener leurs travaux exclusivement dans le domaine de l'IA ce qui est également le cas du laboratoire LERIA à Angers. Il s'agit respectivement des équipes LS2N/TALN et LIUM/LST dans le domaine du traitement de la langue, des équipes LS2N/TASC, LERIA/MO et LS2N/Méforbio dans celui de l'algorithmique, de l'équipe LS2N/PSI dans celui des connaissances et de l'équipe LERIA/RR sur une grande partie du spectre ;
- le taux de « couverture IA » serait donc de 100% pour le LERIA, de 50% pour le LIUM, de 44% pour le LS2N et de 34% pour le LARIS.

³ Il est généralement admis que le nombre de non-permanents (doctorants, post-doctorants et ingénieurs sur contrat) est égal à celui des E/C et C, le taux d'encadrement étant de 1 non-permanent par permanent dans un laboratoire des sciences du numérique.

Cette grille de synthèse est obtenue à partir de la grille déroulée ci-dessous qui explicite pour chacune des 8 thématiques principales ses (éventuelles) sous thématiques. Elle détaille donc le champ des travaux de chacune des 30 équipes.

		Labos du premier cercle Atlanctic 2020																															
		LS2N												LERIA	LARIS	LIUM																	
		SDD				SIEL				RPC				SLS				CSS															
		51				34				28				35				37				20	43	26									
		19	7	15	8	2	10	11	8	6	9	9	5	6	7	6	5	7	6	4	8	12	5	13	9	11	15	15	13	13	13		
		Duke	TASC	SLP	ComBi	Méforbio	TALN	IPI	SIMS	PACCE	RoMas	ReV	ARMEN	OGRE	GDD	Galiette	NaoMod	STACK	AeLos	RIO	PSI	STR	IS3P	Commande	MOA / MO	ICLN/RR	SDO	ISISV	SFD	LST	IEIAH		
Taxinomie Inria Livre blanc IA (pages 20-21)																																	
Connaissances																																	
Base de connaissances																																	
Extraction & nettoyage de connaissances																																	
Inférence																																	
Web sémantique																																	
Ontologies																																	
Apprentissage automatique																																	
Apprentissage supervisé																																	
Apprentissage (partiellement) non-supervisé																																	
Apprentissage séquentiel et par renforcement																																	
Optimisation pour l'apprentissage																																	
Méthodes bayésiennes																																	
Réseaux de neurones ou neuronaux																																	
Méthode à noyau																																	
Apprentissage profond																																	
Fouille de données																																	
Analyse de données massives																																	
Traitement du langage naturel																																	
Traitement des signaux																																	
Parole																																	
Vision reconnaissance d'objets																																	
Vision reconnaissance d'activités																																	
Vision recherche dans les banques d'images/vidéos																																	
Vision reconstruction 3D et spatio-temporelle																																	
Vision suivi d'objets et analyse de mouvements																																	
Vision localisation d'objets																																	
Vision asservissement visuel																																	
Robotique																																	
Conception																																	
Perception																																	
Décision																																	
Action																																	
Interactions avec les robots																																	
Flotte de robots																																	
Apprentissage des robots																																	
Cognition pour la robotique et les systèmes																																	
Neurosciences, Sciences cognitives																																	
Compréhension et stimulation du cerveau																																	
Sciences cognitives																																	
Algorithmique de l'IA																																	
Programmation logique et ASP																																	
Dédution, preuve																																	
Théorie SAT																																	
Raisonnement causal, temporel, incertain																																	
Programmation par contraintes																																	
Recherche heuristique																																	
Planification et ordonnancement																																	
Aide à la décision																																	

3.2 Cible des 31 premiers projets Atlantic2020 cofinancés par la Région Pays de Loire

Une autre analyse possible porte sur les projets portés par Atlanctic2020 et résultant des appels à projets recherche réalisés en 2016 et 2017. Avec une première vague de 31 projets lauréats nous disposons d'une base intéressante pour observer si ceux-ci relèvent du champ de l'IA et dans l'affirmative les projeter sur les 8 thématiques Inria.

La figure suivante reprend donc la grille précédente en remplaçant en ordonnée les équipes de recherche par les projets lauréats⁴. Les porteurs des projets sont indiqués en regard de la (sous)thématique principale du projet. Le X indiquant un projet contribuant au développement

⁴ Nous donnons ici l'acronyme du projet, son intitulé développé figure en annexe2. Le projet scientifique correspondant est disponible sur le Wiki Atlanctic2020 dans sa version longue privée et le site web [Atlanctic 2020](http://atlanctic2020.univ-nantes.fr) dans sa version courte publique. Il s'agit respectivement des url :

- <http://atlanctic2020.fr/resultats-appels-a-projets-2017/>
- https://rfiatlanctic2020.univ-nantes.fr/doku.php?id=consciens:appelprojet:appels_a_projets2017

porteur d'une nouvelle thématique pour leur établissement et leur laboratoire. Deux dossiers sont en fin d'instruction pour accompagner la venue au LS2N en septembre 2018 de deux chercheuses :

1. [Isabelle Fantoni](#) (DR2 CNRS, chargée mission Robotique INS2I, mobilité UTC) porteuse du projet RAPID⁵ au sein de l'équipe LS2N/ARMEN dans le domaine de la Robotique et plus précisément de la commande robuste de flotte de drones donc de flotte de robots selon la taxinomie Inria ;
2. [Assia Mahboubi](#) (CR1 Inria, mobilité Orsay) porteuse du projet VerCoMa⁶ au sein de la nouvelle équipe projet Galinette. Il s'agit de proposer de nouveaux outils Coq pour écrire, vérifier et archiver des preuves mathématiques par informatique. Ce projet relève de la déduction et de la preuve dans la taxinomie Inria.

3.3 Chercheurs clefs

Nous présentons ici les chercheurs distingués à l'ERC ou l'IUF ainsi que ceux impliqués dans des instances nationales d'animation de la recherche en IA.

IUF : [Jin-Kao Hao](#) (PRE Université d'Angers) membre senior de l'Institut Universitaire de France et de l'équipe LERIA/MO. Sa recherche sur les « [metaheuristics](#) and nature-inspired intelligent search methods for solving large-scale complex combinatorial problems » relèvent de l'algorithmique de l'IA.

ERC : [Nicolas Tabareau](#) (CR1 Inria) porteur de l'ERC Starting Grant CoqHoTT responsable de l'équipe Galinette commune au LS2N et au centre Inria Rennes Bretagne Atlantique qui accueille deux autres chercheurs Inria dont Assia Mahboubi. Les travaux [Galinette](#) visent à définir une nouvelle génération d'assistant de preuves dans la lignée de Coq. Il relève de l'algorithmique de l'IA.

CNRS GDR IA [5] : Frédéric Saubion (PR Université d'Angers, LERIA/MO) et Philippe Leray (PR Université de Nantes, LS2N/DUKe) sont membres du comité d'animation de ce nouveau GDR consacré aux aspects formels et algorithmiques de l'intelligence artificielle. Frédéric Saubion est également vice président du CNU 27. Florian Richoux (MdC Université de Nantes, LS2N/TASC) est co-responsable du groupe de travail Jeux.

CNRS pre-GDR TAL (Traitement Automatique du Langage naturel) : Béatrice Daille (PRE Université de Nantes, LS2N/TALN).

3.3 Projets européens H2020 en cours relevant de l'IA

- [ICT X5gon](#)⁷ Colin de la Higuera (PRE Université de Nantes, LS2N/TALN) : utilisation des techniques d'IA pour naviguer dans un ensemble de ressources éducatives. Liaison avec la chaire UNESCO ;
- [chist-era M2CR](#)⁸ Loïc Barrault (MdC Université du Maine, LIUM/LST) : création d'une architecture neuronale unifiée pour la compréhension multimodale et multilingue du langage humain. Lauréat du prix européen Innovation Radar 2016 ;
- [chist-era ALLIES](#)⁹ Anthony Larcher (MCF Le Mans Université LIUM/LST) : développement de systèmes autonomes selon un scénario d'apprentissage donné ;
- à compléter ...

⁵ Formation de Robots Aériens hétérogènes pour l'exPloration avec des Dynamiques variables.

⁶ Verified Computer Mathematics.

⁷ Cross Modal, Cross Cultural, Cross Lingual, Cross Domain, and Cross Site Global OER Network

⁸ Multilingual Multimodal Continuous Representation for Human Language Understanding

⁹ Autonomous Lifelong Learning IntelligEnt Systems (ALLIES)

4. Conclusions et recommandations

Cette cartographie mériterait d'être complétée en intégrant les laboratoires de Mathématiques impliqués dans le programme DéfiMaths piloté par LMJL mais également les laboratoires du 2^{ème} cercle ayant comme tutelles l'INSERM, l'INRA, l'école nationale supérieure d'architecture de Nantes, ONIRIS, IFSTAR, ...

PRO : à ce stade nous observons une activité soutenue des équipes de recherche dans le domaine de l'IA avec des compétences bien réparties sur le territoire en traitement du langage naturel et en algorithmique de l'IA. A cet égard les projets régionaux structurants DEPART et LIGERO pré RFI ont eu un effet structurant certain. Nous constatons également une masse critique importante de 130 E/C permanents menant des travaux et/ou des applications dans le domaine de l'IA. A cette masse critique il convient d'ajouter les startups scientifiques issues des laboratoires comme Dictanova (F. Poulard, LS2N), COSLING (J.-G. Fages et T. Lapègue, LS2N), MyScript (P.-M. Lallican), CogniTalk (S. Knerr), Case Law Analytics (J. Lévy Vehel, LMJL), IAdvize (J. Hervouet), Voxolab (V. Jousse, LIUM), Allo Média (V. Jousse et A. Rousseau, LIUM)

CONS : à l'exception de quelques pépites nous ne disposons pas de leader international dans ce domaine. A titre d'exemple, aucun ligérien ne participe au conseil scientifique du GDR CNRS IA au titre des [EurAI fellows](#) Français [8]. Nous ne disposons pas non plus, à l'exception de la robotique, de réelles plateformes d'expérimentation/démonstration.

Notre recommandation serait de « challenger » la masse critique ligérienne en IA au travers de quelques défis sociétaux contribuant aux transformations numériques en cours. L'objectif pourrait être de lancer quelques projets fédérateurs de construction de vrais démonstrateurs de l'IA en action, en mobilisant par exemple notre expertise de pointe en robotique, algorithmique et traitement de la langue.

Le LS2N porte le projet de « Cluster » DELPHI déposé à l'appel NExT dans le but de développer un réseau local interdisciplinaire sur les dernières avancées de l'IA pour les mettre au service de la santé et de l'industrie du futur. Une autre initiative pourrait porter sur une approche algorithmique et logicielle dédiée au développement de l'industrie du futur incluant une dimension preuve/vérification dans le prolongement de l'atelier [aLIFE](#). Enfin l'apprentissage machine, pierre angulaire de beaucoup d'études en IA, demande des ressources de calcul sur GPU de plus en plus importantes. Une piste serait d'aider à étendre les méso centres de calcul ligériens dans ce sens : l'institut du calcul intensif ICI et le CCIPL.

Remerciements

Remerciements aux directeurs de laboratoires et au collègue recherche du programme Atlanstic 2020 pour leur contribution à cette cartographie et à ce texte de synthèse.

Références

- [1] The society of Mind, Marvin Minsky, Simon and Schuster editors, New York 1986
- [2] France IA [Rapport de Synthèse](#) (350 pages) et [Synthèse](#) (36 pages)
- [3] Inria [Livre blanc IA](#) 2016
- [4] [Note](#) sur les compétences recherchées liées à l'IA. Pasc@line 2017
- [5] [GDR IA](#) CNRS sur les aspects formels et algorithmiques de l'IA
- [6] Voyage au cœur de l'IA. [Hors série Libération](#)
- [7] [Rapport Villani](#) 2018 : donner un sens à l'IA
https://www.aiforhumanity.fr/pdfs/9782111457089_Rapport_Villani_accessible.pdf
- [8] https://www.eurai.org/awards_and_grants/fellows
- [9] https://afia.asso.fr/formulaire-de-recherche-afia/?obj_t=research_group

Annexe 1 : Acronymes et noms développés des équipes de recherche

A1.1 LS2N <https://www.ls2n.fr/les-equipes/>

1. DUKe = Data User Knowledge
2. TASC = Théorie, Algorithmes et Systèmes en Contraintes
3. SLP = Systèmes Logistiques et de Production
4. ComBi = Combinatoire et Bioinformatique
5. Méforbio = Méthodes Formelles pour la Bioinformatique
6. TALN = Traitement Automatique du Langage Naturel
7. IPI = Image Perception Interaction
8. SIMS = Signal, Image et Son
9. PACCE = Perception, Action, Cognition pour la Conception et l'Ergonomie
10. RoMas = Robots and Machines for Manufacturing, Society and Services
11. ReV = Robotique Et Vivant
12. ARMEN = Autonomie des Robots et Maîtrise des interactions avec l'Environnement
13. OGRE = Optimisation Globale et Résolution Ensembliste
14. GDD = Gestion de Données Distribuées
15. Gallinette
16. NaoMod = Nantes Software Modeling Group
17. STACK = Software *Stack* for Massively Geo-Distributed Infrastructures
18. AeLoS = Architectures et Logiciels Sûrs
19. RIO = Réseaux pour l'Internet des Objets
20. PSI = Pilotage des Systèmes Industriels
21. STR = Systèmes Temps Réels
22. IS3P = Ingénierie des Systèmes : Produits, Performances, Perceptions
23. Commande

A1.2 LARIS <http://laris.univ-angers.fr/fr/index.html>

1. SDO = Systèmes Dynamiques et Optimisation
2. ISIV = Information, Signal, Image et Sciences du Vivant
3. SFD = Sûreté de fonctionnement et aide à la Décision

A1.3 LERIA <http://www.info.univ-angers.fr/leria/>

1. MO = Métaheuristiques et Optimisation
2. RR = Représentation des connaissances et Raisonnement

A1.4 LIUM <https://lium.univ-lemans.fr/>

1. LST = Language and Speech Technology
2. IEIAH = Ingénierie des Environnements Informatiques pour l'Apprentissage Humain

Annexe 2 – Acronymes et noms développés des projets de recherche A2020 du domaine de l'IA

1. COPUL = Constrained Optimization Problems with Uncertainty Learning
2. KIMI = Knowledge sharing in Medical Imaging
3. Capillaroscope sans contact : preuve de concept et démonstrateur
4. Robot POODLE = PrOtotype de Désherbage LasEr
5. NéO : observatoire de la néologie
6. SIMPAEX = Segmentation et identification de segments de parole expressive dans un contexte multi-locuteur
7. ACS = Analyse Cognitive de Savoirs
8. CREATOR = Robot parallèle à câbles ayant un grand espace de travail en translation et en orientation

9. ADLACI = Aide à la décision pour la conception de lignes d'assemblage en contexte incertain
10. VISURBA = Visualisation de données urbaines 3D spatio-temporelles et multicouches
11. FABULA = Federated Query Processing for Big Linked Data
12. VAM+ = Exploring peripheral and central vision for visual attention modeling and applications in ICT
13. COMOOG = Conception de modèles graphiques métiers sous contraintes
14. RAPACE = Réseaux de neurones profonds pour le traitement de la langue orale et écrite
15. PROMPT = Performances de la Robotique Manufacturière en Perception de Tâche
16. ROBIBIO = Robot humanoïde BI-articulaires BIO-inspirée
17. ProbioSTIC
18. WebOfBrowsers = WeBrowse
19. REVISTIXX = Réalité virtuelle et stimulation cérébrale, une approche expérientielle
20. AI-GU = AI-based assistant for molecular QUantum chemistry
21. MASCOT = Medical Attendant Shift Conception and OpTimisation
22. EVENT-AI= Visualisation et recommandation contextuelle d'évènements culturels pour des utilisateurs mobiles dans la ville intelligente
23. ENGRAIS = Fusion de données pour la NaviGation (autonome) d'un Robot AgrIcole Symétrique
24. OPTISCAN = Apprentissage automatique de modèles statistiques pour la numérisation d'objets 3D à partir d'un ensemble de points de mesure optimaux