



HAL
open science

Intelligence artificielle et réseau scientifique et technique

Pierre-Cyril Aubin-Frankowski

► **To cite this version:**

Pierre-Cyril Aubin-Frankowski. Intelligence artificielle et réseau scientifique et technique. Sciences de l'information et de la communication. 2018. hal-01876156

HAL Id: hal-01876156

<https://enpc.hal.science/hal-01876156>

Submitted on 18 Sep 2018

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

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



Intelligence artificielle et réseau scientifique et technique

Mémoire de thèse professionnelle pour le Mastère
spécialisé PAPDD

Pierre-Cyril Aubin, année universitaire 2017/2018

Encadré par Messieurs Nicolas Hautiere et Luc Mathis

« Ecole des Ponts ParisTech, AgroParisTech et la Direction de la Recherche et de l'Innovation du Commissariat Général au Développement Durable n'entendent donner aucune approbation ni improbation aux thèses et opinions émises dans ce rapport ; celles-ci doivent être considérées comme propres à leur auteur. »

« J'atteste que ce mémoire est le résultat de mon travail personnel, qu'il cite entre guillemets et référence toutes les sources utilisées et qu'il ne contient pas de passages ayant déjà été utilisés intégralement dans un travail similaire. »

Remerciements

Je tiens en premier lieu à remercier de tout cœur Luc Mathis qui a rédigé, puis suivi, la commande de la DRI et m'a fait confiance dans les différentes phases d'élaboration de ce rapport. Avec mon encadrant académique, Nicolas Hautière, ils m'ont prodigué conseils et soutien à mesure que je m'éloignais de mon thème de prédilection, l'apprentissage automatique, pour pénétrer dans les territoires inconnus de sa déclinaison pour l'action publique.

Je salue chaleureusement mes collègues pour un temps de l'AST, Vincent Letrouit avec qui nous avons si souvent disserté des caractéristiques du scientifique artiste, Laurent Belanger et Claire Sallenave qui m'ont fait l'honneur de suivre la progression de la rédaction de ce manuscrit, Béatrice Yoffo et Sylvain Pradelle qui accompagnent Luc Mathis dans cette volonté d'irriguer le RST par le numérique. Je remercie vivement Serge Bossini et Thierry Courtine pour m'avoir donné la chance de mener cette mission.

Je me tourne enfin vers tous ceux qui, au cours des entretiens, m'ont offert de leur temps, de leur compétence et de leur vision. Ces discussions avaient le don de me motiver et je repense avec plaisir à celles qui m'ont éclairé sur le sens de ma mission. A ceux que je n'ai pu rencontrer dans les temps, en particulier parce que leur fonction mentionnait le mot d'intelligence artificielle, j'adresse mes plus cordiales salutations, en espérant que la lecture de ce rapport les intéressera.

Personnes rencontrées : Dominique Mignaux, Michèle Guilbot, Frédéric Bourquin, Jean-Patrick Lebacque, Bérengère Lebental, Serge Piperno, Jacques Roudier, Pierre-Alain Roche, Louis Fernique, Marion Gust, Céline Bonhomme, Fabien Leurent, Stéphane Trainel, Benoît David, Laurence Matringe, Thierry Coquil, Samuel Goldszmidt, Eric Premat, Marc Pontaud, Frank Fuchs, Nicolas Paparoditis, François Hissel, Hervé Brulé, Vincent Denamur, Pierre Lagarde, Vincent Labbé, Bruno Sportisse, Sandrine Lefebvre-Guillot, André Barkat, Belkacem Laimouche, Georges Uzbelger, Christine Duval, Pierre Moller, Pâris Mouratoglou, David Chupin, Van Bui Tran, Lionel Janin, Patrice Aknin, Loic Landrieu, Jean-Yves Chatelier, Olivier Rivière.

Table des sigles

| | |
|--------|---|
| AESM | Agence européenne pour la sécurité maritime |
| AFB | Agence française pour la biodiversité |
| ANR | Agence nationale de la recherche |
| AST | Animation scientifique et technique |
| BRGM | Bureau de recherches géologiques et minières |
| CEA | Commissariat à l'énergie atomique |
| Cerema | Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement |
| CETU | Centre d'études des tunnels |
| CGDD | Commissariat général au développement durable |
| CNES | Centre national d'études spatiales |
| CSTB | Centre scientifique et technique du bâtiment |
| DAM | Direction des affaires maritimes |
| DGITM | Direction générale des infrastructures, des transports et de la mer |
| DGPR | Direction générale de la prévention des risques |
| DIR | Direction interdépartementale des routes |
| DREAL | Direction régionale de l'environnement, de l'aménagement et du logement |
| DRI | Direction de la recherche et de l'innovation |
| ECN | Ecole Centrale de Nantes |
| EIG | Entrepreneur d'intérêt général |
| ENSM | École nationale supérieure maritime |
| ENSTA | École nationale supérieure de techniques avancées |
| IA | Intelligence artificielle |

IFPEN Institut français du pétrole Énergies nouvelles

IFSTTAR Institut français des sciences et technologies des transports, de l'aménagement et des réseaux

IGN Institut national de l'information géographique et forestière

INERIS Institut national de l'environnement industriel et des risques

IPEF Ingénieur(e) des ponts, des eaux et des forêts

IRSTEA Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture

IRT Institut de recherche technologique

ITE Institut pour la transition énergétique

ITPE Ingénieur(e) des travaux publics de l'État

MESRI Ministère de l'enseignement supérieur, de la recherche et de l'innovation

MTES Ministère de la transition écologique et solidaire

RST Réseau scientifique et technique

STAC Service technique de l'aviation civile

Sommaire

| | | |
|----------|---|-----------|
| 1 | Contextualisation institutionnelle du RST | 3 |
| 1.1 | Le RST : un concept du ministère de l’environnement | 3 |
| 1.2 | Le RST : un ensemble hétérogène d’organismes soumis à diverses tutelles | 6 |
| 1.3 | Le RST : un réseau et des institutions dans un contexte en évolution rapide | 7 |
| 2 | Conceptualisation de l’apprentissage automatique | 9 |
| 2.1 | Modèles et aide à la décision | 9 |
| 2.2 | Survol technique de l’apprentissage automatique | 17 |
| 2.3 | Décloisonnement scientifique et questions statistiques | 21 |
| 3 | Contextualisation technique de l’apprentissage automatique | 25 |
| 3.1 | Chaîne de production et de valorisation des données | 25 |
| 3.2 | Cultures métiers et interaction avec l’apprentissage | 28 |
| 3.3 | Se positionner vis-à-vis de l’apprentissage | 30 |
| 4 | Enjeux pour le RST en matière d’apprentissage automatique | 33 |
| 4.1 | Quelles incitations à agir ? | 33 |
| 4.2 | Quelles finalités pour le RST à s’approprier cette technique ? | 34 |
| 4.3 | Quels prérequis et quelles approches pour développer la technique au sein du RST ? | 44 |
| 4.4 | Trois scénarios d’évolution des institutions du RST | 47 |

Résumé

L'avenir des institutions du réseau scientifique et technique du ministère de l'environnement dépend-il des dernières formes d'intelligence artificielle ? L'apprentissage automatique n'est pas une solution miracle du monde numérique transposable instantanément dans une institution technique. L'apprentissage automatique n'est qu'un outil, extension des techniques statistiques. Sa conceptualisation augure de nouveaux moyens de modélisation et d'aide à la décision. Mais son développement a de nombreux prérequis et nécessite de penser son intégration. Avant de gagner en compétence, en efficience, en visibilité ou en innovation, chaque institution du réseau doit choisir son public de destination.

Abstract

Does the future of the institutions of the technical and scientific network of the ministry of environment depend on the latest creations in artificial intelligence ? Machine learning is not a silver bullet from the digital world that could instantaneously penetrate a technical institution. Machine learning is but a tool, an extension of statistical methods. Through conceptualizing this tool we expect new means of modelization and decision-making. But developing machine learning has many prerequisites, most of all to think about its integration in an existing data science process. Before gaining competence, efficiency and visibility as well as innovating, each institution of the network has to choose its audience.

Introduction

L'intelligence artificielle (IA) est la dernière mode. Après le Big Data, l'Internet des objets, la transformation numérique, voici l'IA sur le devant de la scène, auréolée en France du prestige de la mission Villani et d'un discours présidentiel. Derrière chacune des notions, les différents rapports ont souligné les mutations en cours, les vastes potentialités et les enjeux sociétaux à considérer. Cependant un rapport est marqué par son destinataire : les choix sont effectués en fonction du degré de connaissance de l'objet, du champ auquel le rapport s'applique. Cette étude vise pour sa part un public avec une formation scientifique, pour qui le battage médiatique rend le concept opérationnel plus difficile à cerner. Démystifier l'IA est une nécessité pour que l'objet ne soit pas perçu comme aliénant ou trop étranger aux pratiques. Ce point a déjà fait l'objet d'un excellent rapport par l'Académie des Technologies [AdT, 2018]. Recontextualiser les fondamentaux des techniques d'IA dans leurs questions scientifiques me paraît l'étape suivante pour dépasser le regard d'utilisateur et s'appropriier les enjeux de la discipline afin de rompre les barrières culturelles.

Mais de quelle discipline s'agit-il ? Offrir un regard cohérent sur l'ensemble des techniques qui se revendiquent ou non de l'IA est une gageure. Le domaine comme nous le verrons est un creuset où s'associent de très nombreux courants de pensée, ce qui favorise un décloisonnement des écoles scientifiques. Leurs fondamentaux s'entremêlent autour d'un socle commun : les données. Dans ce rapport, je fais le choix de restreindre l'IA à l'apprentissage automatique, aussi connu sous le nom d'apprentissage statistique ou de *machine learning*. Il s'agit en effet du domaine le plus susceptible d'être en interaction avec les autres disciplines pratiquées dans des institutions scientifiques et techniques. Pour discuter des conséquences potentielles sur l'activité dans ce secteur spécifique, il est difficile d'évoquer en même temps : les systèmes experts ou la recherche opérationnelle et leur contribution à l'industrie ; l'interaction avec des clients dans le secteur tertiaire ; ce qui pourrait participer d'une activité technique dans des métiers scientifiques. L'apprentissage automatique (le *machine learning*) est la science de la modélisation et de l'aide à la décision par des méthodes inductives. Il ne s'agit pas d'un thème nouveau, mais la plupart des méthodes d'intelligence artificielle, comme les *support vector machine* ou les réseaux

de neurones, qui depuis 20 ans ont attiré l'attention s'y rattachent. Le domaine a été un grand bénéficiaire de l'expansion du Big Data et de l'*Open data*, de la mise à disposition en *Open source* des briques logicielles, de la démultiplication des capteurs de l'Internet des objets.

L'apprentissage automatique est un outil dont le réseau scientifique et technique (RST) du ministère de l'environnement (MTES) considère l'opportunité de s'emparer. Avant d'évoquer les enjeux et d'envisager des scénarios d'évolution, il s'agit de contextualiser l'outil et son destinataire. Ces derniers étant protéiformes, il faut procéder à des éclairages multiples pour les saisir. C'est pourquoi ce rapport est écrit sous le signe du balancement. La première partie discute successivement du RST comme concept pour le MTES et de sa manifestation dans un ensemble hétérogène d'organismes partageant des problématiques transverses. La seconde partie tient d'un essai sur l'apprentissage automatique, entre formalisation mathématique et interaction avec les questions statistiques. Une fois posés les deux objets d'étude, les deux dernières parties s'attachent à rendre tangibles les effets de l'association de l'apprentissage automatique et du RST. Toutefois, l'angle théorique reste privilégié, car il ne s'agit pas de préconiser un outil particulier à un établissement donné. Ce rapport a pour objectif de montrer comment la montée en généralité, illustrée par des exemples tirés des entretiens, permet de faire émerger une problématique transverse aux établissements. Tout d'abord, il faut enraciner le bloc algorithmique dans une chaîne de valorisation des données, ce qui est l'objet de la troisième partie. Enfin, en s'appuyant sur la concaténation des chaînes, la quatrième partie développera les enjeux, les incitations, les finalités et les prérequis pour le RST afin de s'appropriier l'apprentissage automatique. Ceci conduira à ébaucher trois scénarios pouvant être suivis individuellement ou collectivement par les institutions du RST. Dans tout le rapport, le prisme de l'apprentissage automatique permet de souligner en filigrane les objectifs et contraintes partagés par les institutions du RST. Le rapport s'appuie sur 36 entretiens et sur la participation à différentes réunions du RST, avec ses dirigeants, le 20 juin 2018 à l'ENPC, et ses responsables de la valorisation, au cours d'un groupe de travail.

1 Contextualisation institutionnelle du RST

Dans cette partie exclusivement, on distinguera la notion de réseau des organismes qui le constituent.

1.1 Le RST : un concept du ministère de l'environnement

Le réseau scientifique et technique (RST) est un regroupement d'organismes associés aux thématiques du ministère de l'environnement (MTES). Il est chapeauté par la sous-direction de l'animation scientifique et technique (AST) de la direction de la recherche et de l'innovation (DRI) au sein du commissariat général au développement durable (CGDD). A ce titre, le RST est en soi un concept, comme l'est le CGDD. C'est un concept propre au MTES, hérité de son histoire d'ingénierie publique et un marqueur de son évolution.

Tentons de dégager le concept par une image miroir. Le concept de RST existant au MTES est par exemple inconnu du ministère de l'enseignement supérieur (MESRI) qui pour sa part a une approche plus universitaire de la notion de réseau. Ceci conduit le MESRI à un suivi très éloigné par thématiques, l'ANR ayant le véritable contrôle des financements. Le MESRI s'adapte à une culture très individualiste où la notion de gestion des ressources humaines reste l'apanage des grands organismes et corps de recherche qui lui sont rattachés (INSERM, CNRS, CEA, universités,...). L'autonomie obtenue par le statut d'agence, le pouvoir d'influence ou des lois comme celle relative aux libertés et responsabilités des universités (loi LRU de 2007), rend très difficile l'exercice d'une tutelle fine. Cette tutelle peut même être inversée car à l'autonomie importante s'ajoute la cohérence, la taille et la pérennité des établissements rattachés au MESRI. Le MESRI s'est ainsi reporté sur la stratégie nationale de recherche pour conserver un rôle, dans l'espoir d'une vision plus globale et transverse.

Au contraire, la tradition d'ingénierie d'un ministère plus opérationnel amène le MTES à une interaction plutôt avec des institutions que de manière directe avec des agents qui jouiraient de la liberté propre aux chercheurs. La proximité de la DRI avec les directions centrales du ministère, comme la DGITM ou la DGPR, la rapproche des décideurs publics et institutionnels qui déterminent ou approuvent les choix stra-

tégiques des organismes. Le paysage de l'ingénierie publique est paradoxalement bien plus éclaté que celui de la recherche publique. La spécialisation extrême des sujets a requis la constitution d'équipes de taille modeste, de l'ordre d'une centaine de personnes, suffisamment cohérentes pour être efficaces sur les politiques et thématiques qui les sollicitent. Les grands remembrements¹ ne font qu'occulter cet état de fait. Les directions centrales ont donc eu pour rôle naturel de fournir une cohérence générale à cet ensemble très diversifié. Il n'existait pas alors d'établissement parmi les organismes directement rattachés au MTES qui puisse choisir de s'abstraire de sa tutelle et de sa relation avec le réseau ministériel sans perdre à la fois ses missions, son financement et sa cohérence. Pour une part importante, le personnel lui-même est dépendant des corps techniques du ministère (ITPE, IPEF,...), ce qui contraint l'autonomie des institutions. Ce lien étroit a entraîné la pérennité des structures aussi longtemps que le ministère associé disposait d'une cohérence. C'est pourquoi la proximité même des organismes à leur tutelle les a confrontés à des bouleversements quand leurs tutelles, aussi bien corps d'Etat que ministères, évoluaient.

En effet l'observation du RST souligne que la stabilité des organismes dépendant de façon prépondérante de l'industrie (BRGM, IFPEN) ou de l'enseignement supérieur (CEA, CNES, MNHN) va de pair avec l'instabilité chronique des établissements liés à l'environnement (IRSTEA, IFSTTAR, Cerema, IGN). Cette instabilité est à mettre en regard des "batailles homériques ou guerres microcholines" [Lascoumes et al, 2014] qui ont déterminé la ligne du MTES à sa fondation. La fondation du ministère est un point central car la naissance du concept de RST sous sa forme actuelle provient de la création du CGDD qui reçoit en 2008 dans ses attributions la définition des "orientations du réseau des organismes scientifiques et techniques du ministère et [...] leur mise en œuvre ainsi [que] la diffusion des produits de la recherche et de l'innovation technique"². L'idée de RST porte les mêmes ambitions de transversalité et de mutualisation que le CGDD. La sous-direction AST

1. A titre d'exemple, le Cerema comporte douze directions, dont celle principale, réparties sur 32 sites.

2. https://www.legifrance.gouv.fr/affichTexteArticle.do;jsessionid=36966840399107F4900391357A128118.tplgfr32s_1?cidTexte=JORFTEXT000019146939&idArticle=LEGIARTI000036445670&dateTexte=20171231&categorieLien=id#LEGIARTI000036445670

correspondante tire son existence de la présence d'un RST, qui lui-même renforce la légitimité du CGDD face aux autres directions centrales par la maîtrise de la dimension recherche des thèmes du MTES. La cohérence institutionnelle du nouveau ministère et la recherche d'une culture commune ont ainsi pris racine dans la DRI, qui succède à la Direction de la recherche et des affaires scientifiques et techniques (DRAST) qui existait dans les formes antérieures du MTES.

IPEF, RST et transition

La relation du MTES au RST témoigne de la fondation du ministère et de l'identité en construction du corps des IPEF pour lequel le RST est un débouché fréquent. En particulier, pour le corps de IPEF, l'évolution du RST en agences plus autonomes traduit ce délitement du lien avec l'Etat central. Ce délitement est la conséquence avant tout de la décentralisation et de la réduction du périmètre de l'Etat, qui s'est exprimée le plus complètement dans la révision générale des politiques publiques (RGPP de 2007) et ses avatars, modernisation de l'action publique (loi MAP de 2012) et aujourd'hui action publique 2022 (AP22 lancé en 2017). Le récit qui est proposé aux promotions postérieures à la fusion de 2009 entre les ingénieurs des Ponts et Chaussées (IPC) et ceux du Génie Rural des Eaux et des Forêts (GREF) s'en ressent. La dimension sociale de gestion du personnel existant et la relation aux collectivités apparaissent comme l'activité prédominante du corps des IPEF à mesure que les privatisations et l'agencification éloignent les organismes du giron de l'Etat.

Le délitement du lien à l'État est à replacer dans un contexte de dévalorisation de la parole scientifique dans la sphère publique. Il sera illustré plus loin par le risque réel de dépréciation progressive des actifs techniques étatiques.

1.2 Le RST : un ensemble hétérogène d'organismes soumis à diverses tutelles



FIGURE 1 – La diversité du RST illustrée par ses logos (Source : site web du RST)

Le RST est composé de 32 organismes (dont les logos sont représentés sur la Figure 1) :

- 6 services à compétence nationale ou services techniques centraux ou équivalent : CETU, SCHAPI, SNIA, STAC, LCPP, STRMTG
- 10 établissements publics à caractère administratif : AFB, ANSES, Cerema, ENAC, ENGEES, EPSF, IGN, Météo-France, ONCFS, ONF
- 2 établissements publics à caractère scientifique et technologique : IFSTTAR, IRSTEA
- 5 établissements publics à caractère scientifique, culturel et professionnel : Ecole Centrale de Nantes, ENPC, ENSM, ENTPE, MNHN
- 9 établissements publics à caractère industriel et commercial : Ademe, BRGM, CEA, CNES, CSTB, IFPEN, IFREMER, INERIS, IRSN

Cette énumération de sigles masque l'hétérogénéité des organismes, depuis les 90 agents du CETU jusqu'aux 16000 du CEA. Les sigles n'indiquent pas non plus les tutelles multiples de nombreux établissements, partagés entre la santé, le travail, l'économie, l'agriculture, l'enseignement supérieur,... La DRI n'exerce de tutelle directe que sur quatre d'entre eux (IGN, Cerema, IFSTTAR et Météo-France). Son rapport aux autres établissements est plus ou moins distant. La sous-direction de l'animation scientifique et technique (AST) se partage donc entre un suivi spécifique et un appui à ses établissements proches et les questions transversales, souvent abordées par des thématiques, comme valorisation et modèles économiques, ou données

et numérique, et par des outils, comme l'intelligence artificielle. La participation au réseau repose sur le volontariat, et donc sur l'attrait des sujets abordés et la qualité des moyens proposés (masterclass, datathon, groupe de travail,...).

1.3 Le RST : un réseau et des institutions dans un contexte en évolution rapide

Pour les institutions qui proposent des prestations, la concurrence s'est accrue. La montée en compétences du privé ou du monde associatif est notable, aussi bien par la maîtrise des données, comme OpenStreetMap ou GoogleMaps en données géographiques, que pour la maîtrise technique, comme Setec ou Eren en inspection des infrastructures, ou même en recherche et innovation, comme Plume Labs ou Breezometer sur la qualité de l'air. L'origine publique (Caisse des Dépôts, SCNF, RATP,...) ou historiquement privée de ces concurrents importe peu car ces organismes se transforment rapidement. Des spin-offs, comme Altametris (ex-pôle drones de SNCF Réseau) ou Spygen³ (issu du laboratoire d'Ecologie Alpine), éloignent progressivement la partie innovante des organismes de leur attache publique, à mesure que ces organismes s'autonomisent eux-mêmes. L'heure est aux agences d'agences, comme l'ITE Efficacity ou l'IRT SystemX vis-à-vis de l'ANR, qui bénéficient de plus d'autonomie pour s'adresser à un marché dont le référentiel n'est plus exclusif à la puissance publique, mais dont les clients restent largement les collectivités.

Cet écosystème plus riche peut être une chance, car il permet de repenser les missions des établissements du RST, de se focaliser sur l'innovation en laissant le soin des capacités organisationnelles et de passage à l'échelle à d'autres acteurs. Mais la transformation des acteurs a lieu au rythme du numérique, en particulier dans la dernière décennie. En témoigne la reconversion des constructeurs automobiles pour qui l'électronique devient la source principale de valeur, par les services. Métiers et missions sont réévalués, ce qui engendre des frictions, car les institutions du RST sont souvent les têtes de leurs propres réseaux opérationnels (STAC, CETU) ou en incorporent une partie (AFB, Cerema, Météo-France).

Le Programme d'investissements d'avenir (PIA) lancé en 2010 accompagne cette transformation générale. Le numérique, qui n'y était pas une priorité, gagne en visi-

3. méthodes d'inventaires de la biodiversité basées sur l'étude de l'ADN environnemental

bilité après la pénétration exceptionnelle du smartphone depuis 2012. Sa traduction budgétaire dans le fonds pour la transformation de l'action publique (FTAP) offre une marge de manœuvre pour adapter les institutions au nouveau contexte (Météo-France a remporté un appel à projet de 2,9M d'euros). Par ailleurs, au niveau des ministères, des réseaux (Supervision générale des données au CGDD, qui s'est dotée d'une doctrine en juin 2018) et des pôles (DINSIC-Etalab) se constituent et s'affirment. Les collectivités ont les leurs, comme Polis ou l'IDDRIM, et la dimension européenne ou internationale se développe.

Voici donc le contexte dans lequel la DRI et les organismes du RST s'interrogent sur ce que l'IA pourrait leur apporter.

2 Conceptualisation de l'apprentissage automatique

L'apprentissage automatique (le *machine learning*) est la science de la modélisation et de l'aide à la décision par des méthodes inductives. Il ne s'agit pas d'un thème nouveau⁴, mais la plupart des méthodes d'intelligence artificielle, comme les *support vector machine* ou les réseaux de neurones, qui depuis 20 ans ont attiré l'attention s'y rattachent. Cette section est à la fois descriptive et technique, elle propose des typologies pour l'aide à la décision et l'apprentissage automatique. Ce cadre conceptuel est associé ensuite à la dynamique de décloisonnement des sciences et aux questions qui traversent ses fondements statistiques.

2.1 Modèles et aide à la décision

2.1.1 Qu'est-ce qu'un modèle ?

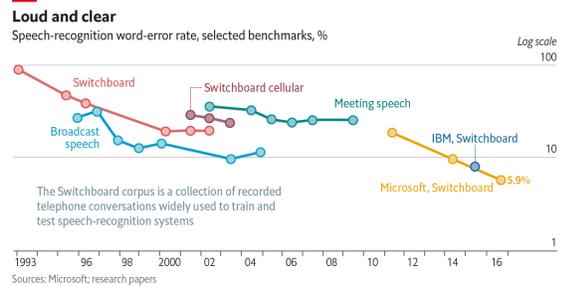
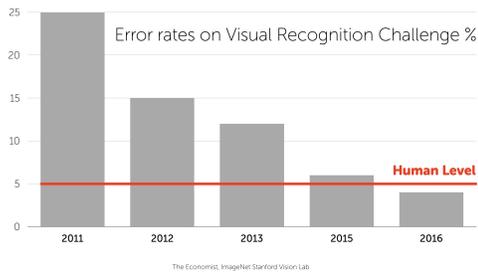
« Pour un dictionnaire pas trop récent, le substantif "modèle" est un emprunt à l'italien *modello* du XVIème siècle, représentation miniature de ce qui sera construit en grand (modèle réduit). Le terme prend ensuite le sens d'objet référent, puis d'idéal à imiter. Il faut attendre le XXème siècle pour voir apparaître le sens technico-scientifique sur lequel porte notre propos. [...] On peut être tenté d'en conclure ou de laisser entendre qu'un modèle mathématique possède par essence la perfection ultime liée aux deux mots qui le définissent. Certains ne s'en privent pas, souvent par ignorance, dans les médias ou chez les décideurs. Il faut consulter un dictionnaire plus récent pour trouver le sens de modèle comme "représentation simplifiée d'un processus, d'un système", et celui de "modèle mathématique" comme "traduction d'une observation dans le but de lui appliquer les outils, techniques et théories mathématiques et, en sens inverse, la traduction des résultats mathématiques obtenus en prédictions ou opérations dans le monde réel". » [Mawhin, 2017, p11]

Un modèle mathématique est la conséquence des choix de son concepteur. Loin d'une idéalisation de l'objet, John von Neumann lui attribue trois qualités : la conformité aux observations, la faculté de généralisation à d'autres situations et l'esthétique du dépouillement.

4. L'apprentissage statistique remonterait au moins aux travaux de Vladimir Vapnik de 1964.

« *La justification d'une telle construction mathématique est uniquement et précisément qu'on s'attend à ce qu'elle fonctionne, c'est-à-dire décrire correctement les phénomènes dans un domaine raisonnablement large. En outre, elle doit satisfaire certains critères esthétiques, c'est-à-dire, qu'en fonction de la portée de ce qu'elle décrit, elle doit être assez simple. Je pense qu'il vaut la peine d'insister sur ces termes vagues, par exemple sur l'emploi du mot 'assez'. On ne peut pas dire exactement combien 'simple' est simple.* » [von Neumann, 1955, p157-164], d'après [Mawhin, 2017, p89]

Qu'est-ce qu'un bon modèle ? Comment le juge-t-on acceptable scientifiquement ? L'énoncé de von Neumann n'offre pas de cadre au modélisateur autre que sa subjectivité artistique et sa faculté à représenter le réel. Ces facultés étant limitées, choisir un bon modèle s'est ramené à sélectionner une classe de modèles et en extraire le meilleur en certain sens. La théorie de l'apprentissage statistique a, depuis von Neumann, fourni un cadre conceptuel au critère de comparaison. Mais il restait une autre barrière à franchir. Pour qu'un modèle se substitue à l'être humain dans certaines tâches, il fallait pouvoir le dépasser, en vitesse comme en précision. En mécanique, la confrontation a eu lieu avec le cheval, ce qui a donné le cheval-vapeur. Aujourd'hui avec la percée de l'apprentissage dans les tâches de cognition et de perception, l'on assiste à une nouvelle comparaison, celle avec l'être humain. La comparaison s'est faite d'abord dans des cadres très spécifiques, comme les échecs (victoire de *Deep Blue* sur Gary Kasparov en 1996) ou le go (victoire d'*Alpha Go* sur Lee Seedol en 2016). La comparaison s'est ensuite étendue à l'être humain moyen, comme en témoigne la Figure 2.



Speak easy

Human scorers' rating* of Google Translate and human translation

Translation method | Phrase-based[†] | Neural-network[†] | Human



Source: Google

*0=completely nonsense translation, 6=perfect translation †Machine translation

FIGURE 2 – Exemples de performance de modèles d'apprentissage automatique confrontés à une moyenne humaine sur trois tâches classiques : reconnaissance d'objet, reconnaissance de la parole et traduction de texte. (Source : The Economist)

Apprentissage profond

L'apprentissage profond (*deep learning*) est, parmi les techniques d'apprentissage automatique, celle qui excite le plus de fantasmes. Cette extension des réseaux de neurones aux couches multiples est la méthode phare qui attire les spécialistes de tous bords au chevet de sa compréhension. Les meilleurs résultats des trois cas d'études ^a de la Figure 2 sont obtenus par cette technique. Avant de conclure abusivement que la supériorité sur l'humain est entérinée, il faut rappeler qu'il s'agit, pour la reconnaissance visuelle et vocale, d'un cadre très standardisé (données ImageNet et Switchboard) pour permettre la comparaison des techniques. Ce cadre est cependant couramment généralisé par l'utilisation d'assistants vocaux (Siri, Cortana, Alexa,...) et de méthodes d'annotation d'images sur les réseaux sociaux. Concernant les capacités cognitives à attendre de tels modèles, je me bornerai à rappeler que les réseaux de neurones proviennent du perceptron de Rosenblatt (1957) et ont connu déjà deux "hivers de l'IA". On peut noter que l'apprentissage profond contrevient sans nul doute à l'importance esthétique que Von Neumann accordait aux modèles, excepté pour la forme de ses empilements. Enfin, comme l'écrit Jean Mawhin dans [Mawhin, 2017, p100] :

« Tout comme une église est, dit-on, une secte qui a réussi, une théorie scientifique n'est peut-être qu'un modèle qui a (temporairement) réussi. Et par "réussi" il faut probablement entendre — au sens de l'économie de marché — "réussi à éliminer les autres", en attendant d'être lui-même remplacé. Seule peut-être une "théorie du tout" à laquelle d'aucuns rêvent encore, mériterait encore le nom de théorie. On peut se demander si sa venue sur Terre ne signifierait pas que la science est devenue une religion. »

^a. Pour la reconnaissance visuelle : <https://www.economist.com/special-report/2016/06/25/from-not-working-to-neural-networking> Pour la reconnaissance vocale ou la traduction : <https://www.economist.com/technology-quarterly/2017-05-01/language>

2.1.2 Quelle aide à la décision ?

Un modèle sert à appréhender la complexité. Ce besoin survient dès qu'un problème est difficile d'accès à la perception humaine. Cette perplexité peut avoir de multiples origines : la dimension des variables peut dépasser la compréhension d'un être humain ou sa faculté à se représenter le problème ; les relations entre les variables peuvent être médiées de façon complexe ; enfin le problème peut se répéter trop souvent pour qu'un opérateur y consacre l'attention requise.

Que propose un modèle pour accompagner une décision ? Je propose ici une typologie des sorties d'un modèle :

- **descriptive** : le modèle renvoie des valeurs qui représentent l'objet suivant des modalités objectives, liées à des caractéristiques statistiques, comme la forme de la distribution sous-jacente aux données ou ses composantes principales. Ces descripteurs sont le fruit d'hypothèses statistiques caduques hors de leur domaine de validité.
- **prédictive** : le modèle assigne des valeurs qui peuvent être comparées à des caractéristiques objectives de l'objet. On peut penser à l'appartenance à une classe ou à une valeur qui sera observée à une date ultérieure sans avoir été perturbée par l'utilisation du modèle.
- **prescriptive** : le modèle infère des causes et propose des actions qui, une fois effectuées, changeront l'état du système.

Pourquoi un décideur s'en remet-il à un modèle avant de choisir ? En effet, il pourrait faire valoir son autorité et en assumer l'arbitraire. Cependant il mettrait alors en jeu à la fois sa responsabilité face aux conséquences et sa crédibilité pour les choix futurs. Dans un monde où la figure de l'expert est aisément contestée⁵, le modèle a pour vertu d'atténuer la responsabilité d'un décideur comme justification d'une certaine rationalité. Il dénote aussi une perte de confiance dans l'à-propos de mettre en jeu une responsabilité dans un contexte complexe ou incertain.

La dimension descriptive est le rôle dévolu aux statistiques depuis qu'elles furent

5. Au point de faire l'objet d'un travail spécifique de France Stratégie sur les causes de cette défiance : <http://www.strategie.gouv.fr/travaux/chantiers/paroles-dexperts-quoi-faire>

présentées aux décideurs. En créant un cadre de discussion, elles définissent des références partagées dans leur langage. La dimension prédictive correspond à l'oracle, terme courant en modélisation, dont on peut vérifier la véracité car sa prédiction, même probabiliste, est associée à une fatalité pythique. Cette fatalité est le propre d'une situation qui tient plus de l'observation que de l'action. La dimension prescriptive est celle qui a le plus d'affinités avec les nécessités d'un décideur. Elle le décharge d'une large part de la réflexion pour lui fournir un éventail réduit de décisions. Cette dernière approche est-elle si rare ? Sauf à être qualifié d'inhumain, nul ne peut prendre des décisions ayant des effets considérables sur un grand nombre de personnes sans pouvoir s'appuyer sur une règle. C'est là peut-être une des origines du code juridique qui standardise les choix à effectuer afin de limiter le poids de la décision sur le magistrat. De même qu'un magistrat confronté au doute a néanmoins l'obligation de juger, un modèle proposera toujours une réponse, même inadéquate.

Illustrons les différences entre ces notions autour d'un thème récurrent au sein du RST⁶ : les activités de police⁷.

- descriptive : A New-York, une équipe municipale d'analystes⁸ trouve une corrélation entre des données d'égouts et des restaurants rejetant illégalement leurs déchets. Le modèle n'a pas d'information sur l'illégalité des restaurants, ils sont indiqués comme des anomalies d'une distribution statistique.
- prédictive : Il est courant aux États-Unis d'essayer de prédire le taux de récidive d'un condamné⁹. Le modèle fournit une information qui peut être confrontée à une réalité¹⁰.

6. On peut citer le contrôle des pesticides pour l'AFB, le contrôle des émissions des cheminées à l'INERIS,...

7. Pour des lecteurs à l'approche plus industrielle, les trois situations se déclinent aisément au cadre de la maintenance prédictive.

8. Pour une description de l'équipe et de ses motivations : http://ccnmtl.columbia.edu/projects/caseconsortium/casestudies/127/casestudy/www/layout/case_id_127_id_889.html

9. Pour une tentative de justification de cette activité de recherche : <http://theconversation.com/new-models-to-predict-recidivism-could-provide-better-way-to-deter-repeat-crime-44165>

10. Pour un exemple d'évaluation d'un modèle de prédiction de la récidive : <https://futurism.com/algorithms-no-better-predicting-repeat-offenders-inexperienced-humans/>

- prescriptive : Les modèles de police prédictive¹¹ suggèrent aux policiers de faire leur ronde en des lieux spécifiques. Suivre la prescription change l'information qui pourra être récoltée¹².

Ces trois exemples soulignent que le degré d'aide à la décision dépend d'un choix volontaire du décideur et de la temporalité associée. Dans le premier cas, l'analyse est rétrospective. Dans le second cas, l'analyse a vocation prospective, mais n'a d'incidence que si le décideur en tient compte pour modifier la peine infligée aux condamnés, ce qui revient à biaiser *a posteriori* le modèle. De plus, l'analyse est effectuée sur des données passées, antérieures à l'existence du modèle. Dans le troisième cas, l'analyse est prospective et a un effet *a priori* sur les observations possibles, le décideur s'en remet au modèle pour agir.

2.1.3 Les décideurs aux prises avec les modèles

Quelle différence entre un décideur et un simple acteur ? Le décideur bénéficie d'un "point de vue de surplomb" [Desrosières, 1993, p66]. Ses décisions l'engagent moins personnellement que celles d'un acteur, au même titre que le casino peut s'autoriser plus de pertes que le simple joueur et a par conséquent accès à d'autres outils que des martingales illusoire. Entre les deux acteurs, il y a une dissymétrie des rôles et de l'information. Alors que le joueur n'a accès qu'à des probabilités de croyance, depuis son point de vue subjectif, le décideur aspire à l'objectivité. Ceci rappelle la longue querelle statistique entre bayésiens et fréquentistes, sans cesse remise au goût du jour, comme dans [François et Frezal, 2016] pour le secteur de l'assurance.

Le vocabulaire autour de l'IA sous-entend fréquemment un risque de domination de l'être humain. Si une réponse classique est qu'un programme informatique est aujourd'hui incapable de créativité et de dépasser sa grammaire, au-delà de la crainte millénariste, un modèle peut-il dominer un décideur ? Il n'y a pas besoin

11. Pour une présentation des méthodes : <https://www.theguardian.com/cities/2014/jun/25/predicting-crime-lapd-los-angeles-police-data-analysis-algorithm-minority-report>

12. Sur la difficulté d'évaluer de tels modèles : <http://www.internetactu.net/2015/06/23/predpol-la-prediction-des-banalites/>

d'invoquer une intelligence supérieure pour obtenir un tel cas de figure. Un modèle peut se contenter de devenir une référence partagée qui par son formalisme fonde et contraint une communauté. Dans [François et Frezal, 2016], les auteurs montrent comment la directive Solvabilité II s'est imposée en façonnant les cadres de pensée des compagnies d'assurance. "Les modèles et les hypothèses implicites qui les fondent jouent en effet, dans la sphère financière, le rôle d'une institution : ils constituent des cadres a priori, normatifs et partagés." Ici il s'agit du vocabulaire stochastique pour la gestion du risque, la norme renvoyant à une contrainte sur le risque de défaut d'une entreprise. L'acceptation de ces outils se fait par la démission volontaire des responsabilités humaines : "[Les statistiques] permettent, dès lors, d'éviter la mise en critique et de la nécessité de se justifier ex post." Des indicateurs comme le classement de Shanghai dans l'enseignement supérieur n'agissent pas différemment.

Proposer un modèle permet par ailleurs de discuter du biais d'expert. En décrivant un objet par d'autres représentations, il peut y avoir une confrontation avec l'expertise antérieure, qui ne vise plus seulement à imiter mais à compléter ou à souligner les limites, du modèle comme de l'expert.

Ces réflexions sur les modèles ne sont pas exclusives au phénomène de l'intelligence artificielle. Quelle est donc la conséquence spécifique à la sophistication des modèles ? Le monde statistique usuel est surtout empreint de linéarité. Les notions de moyenne ou de variance en témoignent. Une large part des méthodes d'apprentissage correspondent à l'extension au cadre non-linéaire. Or l'être humain ne conçoit aisément que les dépendances linéaires ("la croissance d'une valeur induit la croissance d'une autre"). Le non-linéaire ne se prête pas à des raisonnements aussi sommaires. Il faut accepter pour se servir de tels outils de se départir des limites du langage vernaculaire pour embrasser l'abstraction mathématique. Cette remarque renvoie au caractère abusif de "boîte noire" assigné à des modèles. Tout modèle commence par être une "boîte noire" avant qu'il ne soit compris, mais la montée en abstraction fait qu'il ne le sera que par un public restreint.

Comme le rappelait John Von Neumann ([von Neumann, 1955, p157-164], d'après [Mawhin, 2017, p89]) : « Certaines des théories que nous avons adoptées, certains des modèles qui nous rendent très heureux et dont nous sommes très fiers ne paraîtraient probablement pas particulièrement simples à quelqu'un à qui on les expose pour la

première fois. »

La fin de la linéarité a des conséquences considérables car elle remet en cause l'idée et la légitimité de la moyenne ordinaire. Ceci ouvre la porte à l'individualisation. Un internaute aujourd'hui n'est plus soumis à une publicité généraliste s'adressant à un hypothétique auditeur moyen. Il est traité dans sa singularité, la sortie du modèle est adaptée à son historique d'achat ou de navigation¹³. Dans l'assurance, les entreprises cherchent à mieux segmenter le public afin d'adapter le produit. Nous assistons progressivement à la fin de la mutualisation, phénomène rendu possible par la complexité des modèles utilisés. Il fallait pour cela une classe de modèles suffisamment riche pour traiter efficacement la diversité, car une solution sur-mesure non automatique, si elle était envisageable, n'était pas économiquement viable. Comme l'écrit Olivier Fliche de l'Autorité de contrôle prudentiel et de résolution : "Dans le secteur financier, en particulier en assurance, le jeu de la concurrence incite à la segmentation et, d'une certaine façon, à une démutualisation des risques qui peut aller jusqu'à une exclusion, explicite ou par les prix, d'une partie de la population." [Enjeux Numériques, 2018a, Fliche, p30]

2.2 Survol technique de l'apprentissage automatique

Cette section cherche à souligner les principes communs aux nombreuses méthodes, sur un plan formel ainsi que sur un plan mathématique. Elle ne se prétend pas exhaustive dans la présentation des techniques. Je recommande la lecture du chapitre 1 du rapport sur l'IA de l'Académie des technologies [AdT, 2018] pour une présentation très accessible des différentes familles d'outils. Par ailleurs, l'aide à la décision n'est bien entendu pas exclusive à l'apprentissage automatique. La théorie du contrôle, la recherche opérationnelle proposent aussi une aide à la décision, souvent au sens d'optimalité d'un critère. Néanmoins leurs modèles sont définis indépendamment des données.

L'apprentissage automatique ou statistique (le *machine learning*) est l'étude systématisée et automatisée à partir d'observations, dans une logique inductive, de la

13. [Enjeux Numériques, 2018a, Niccoli et Le Ouay, p59]

relation entre une entrée x , un modèle f et une sortie y

- Pour x et y connus, il s'agit de modélisation où l'on cherche un modèle correspondant aux données observées.
- Pour x et f connus, il s'agit d'une aide à la décision, sous des modalités descriptives, prédictives ou prescriptives telles que discutées ci-dessus.
- Pour f et y connus, il s'agit d'un problème inverse ou de traitement du signal.

On distingue par habitude les problèmes de classification lorsque la sortie y est discrète, correspondant souvent à un label, des problèmes de régression où la sortie y est continue. Un problème de classification est fréquemment le résultat d'un seuillage sur une régression.

On peut définir trois grandes classes d'apprentissage :

- **supervisé** : l'apprentissage est effectué sur un ensemble de couples $(x_i, y_i)_i$. L'erreur commise par le modèle est connue *a priori* par l'algorithme.
- **non supervisé** : dans les cas où la sortie y est difficile à décrire, l'algorithme en génère une en cherchant une homogénéité dans ses données. L'erreur commise n'a pas nécessairement de sens numérique.
- **par renforcement** : l'algorithme génère son propre ensemble de couples $(x_i, y_i)_i$. L'erreur commise par le modèle est connue *a posteriori* par l'algorithme.

Illustrons ces idées dans leur cadre mathématique. Dans la pratique, bien des problèmes de modélisation peuvent s'écrire comme un problème de minimisation sur une classe \mathcal{M} de modèles f où l'on associe une fonction de perte L , qui décrit l'erreur par rapport aux données observées $(x_i, y_i)_i \in \mathcal{X} \times \mathcal{Y}$, à une fonction de régularisation R , qui limite les variations du modèle f .

$$\min_{f \in \mathcal{M}} L(\{x_i\}_i, \{y_i\}_i, f) + R(f)$$

Un premier exemple naturel d'un tel problème est la régression linéaire quadratique. La fonction de perte L relève d'une distance entre les prédictions du modèle et les observations. Supposons maintenant qu'il existe un modèle $f \in \mathcal{M}$ qui soit capable d'assigner simultanément à chaque x_i la valeur y_i . Ce modèle serait parfait au regard des observations, mais si ces observations étaient bruitées ou si elles étaient très in-

complètes, un tel modèle serait souvent erroné pour d'autres données. Ce phénomène est connu sous le nom d'*overfitting*. Ceci pose la question de la généralisation d'un modèle, au sens de la confrontation à des données qui n'ont pas servi à son apprentissage.

Que faire lorsque la classe de modèles est riche ? Si la classe \mathcal{M} peut approcher de nombreuses observations et que les sorties $f(x_i) \in \mathcal{Y}$ de ses modèles peuvent varier considérablement entre deux points x pourtant proches (au sens d'une distance sur \mathcal{X}), il est judicieux d'appliquer un principe d'homogénéité. Celui-ci se traduit par une fonction de régularisation R sur le modèle f . Une telle régularisation est inutile en basse dimension pour des classes de modèles pauvres comme les modèles linéaires. Le compromis à trouver entre complexité et robustesse des modèles a été largement étudié et est connu sous le nom de dilemme biais-variance.

Le problème a-t-il une solution ? On choisit souvent les fonctions L et R de façon à ce qu'une fois rassemblées le problème soit convexe en f , car ceci garantit dans la plupart des cas l'unicité de la solution et donne accès à des algorithmes déterministes pour trouver le modèle optimal f^* , comme la méthode de Newton ou la descente du gradient. Parfois le cadre de modélisation n'assure pas la convexité, c'est le cas pour les paramètres des réseaux de neurones, les algorithmes deviennent alors stochastiques et reposent sur de nombreuses heuristiques.

L'apprentissage non supervisé a été mentionné dans de nombreux rapports comme étant l'avenir des méthodes d'apprentissage. Il faut tempérer cette affirmation qui laisserait accroire que dans ce cadre la machine ferait preuve d'une créativité débridée. Au contraire, l'absence de données d'entraînement signifie qu'elles doivent être compensées par des hypothèses plus nombreuses et des heuristiques plutôt que des garanties. En pratique, l'apprentissage non supervisé est très proche des tâches de visualisation, où l'on cherche à décrire les observations dans un espace de plus petite dimension. L'analyse par composantes principales est ainsi une réduction de dimension qui permet, en apprenant les composantes, de ramener un problème de très grande dimension, comme les quelques 20000 gènes humains, à un espace perceptible par un opérateur, bi ou tridimensionnel. L'apprentissage non supervisé est en particulier adapté aux données peu structurées, comme des graphes, pour lesquelles

la qualité de la prédiction est en soi un objet difficile à évaluer. La visualisation par un être humain permet de s'assurer de l'à-propos de la sortie de l'algorithme.

L'apprentissage par renforcement présuppose pour sa part la possibilité de pouvoir générer de nombreux exemples et de connaître *a posteriori* l'erreur commise. Si ces éléments semblent moins contraignants qu'un ensemble d'apprentissage annoté, ces méthodes sont le mieux adaptées pour un environnement contraint de jeux, comme les bandits manchots ou les jeux de plateau. Les principes de l'apprentissage par renforcement découlent d'une logique de théorie des jeux. La machine explore des possibilités avant d'exploiter les connaissances acquises sur son environnement. L'apprentissage y est séquentiel, à la recherche de stratégies gagnantes, sans s'appuyer une mémorisation complète des situations.

En somme le volume des données et le degré de compréhension dont nous disposons détermine le choix des méthodes. Si les données disponibles couvrent de façon assez complète les cas envisagés et que l'évaluation de la qualité des prédictions laisse peu d'ambiguïtés, l'apprentissage supervisé a ses avantages. Si les données observées ne sont pas annotées mais qu'une généralisation du modèle n'est pas nécessaire et que le problème est difficile à décrire ou évaluer, le problème se décrit plus naturellement par l'apprentissage non supervisé. Enfin si les règles de génération de données sont bien comprises et dans un espace restreint et que l'évaluation est bien formulée, on reconnaît un cadre d'apprentissage par renforcement.

Bréviaire du modélisateur

- Est-il **pertinent** de définir une certaine fonction objectif à maximiser ou minimiser au vu des besoins ? Cela dépend du type d'erreur que l'on souhaite pénaliser afin de satisfaire l'utilisateur.
- Le choix de la classe de modèles est-il **cohérent** avec la fonction objectif ? Une classe de modèles trop riche ou trop pauvre peut

être inadaptée pour une fonction objectif si l'objectif n'a pas d'effet sur la sélection du modèle (l'objectif est inatteignable) ou trop d'effets (l'objectif est toujours atteint).

- L'algorithme sélectionné est-il **efficace** en optimisant effectivement la fonction objectif pour la classe de modèles ? Il faut que l'algorithme converge, qui plus est vers la bonne valeur.
- Le modèle sélectionné est-il **efficient** en étant généralisable à des cas nouveaux sans requérir plus de paramétrage ou plus de données ? Ceci correspond à un modèle parcimonieux en termes d'entrées d'apprentissage.
- Les sorties du modèles sont-elles **utiles** vis-à-vis des besoins ? Cela requiert un indicateur mesurant la performance des sorties du modèle au vu des attentes.
- Est-il **cohérent** d'optimiser une fonction objectif donnée s'il existe plusieurs objectifs ? Parfois un objectif n'est atteignable qu'au prix d'être sous-optimal pour d'autres objectifs.

2.3 Décloisonnement scientifique et questions statistiques

Une conséquence majeure de la publicité faite autour des thèmes de l'intelligence artificielle a été d'attirer des chercheurs de nombreuses écoles de pensée mathématique à s'y intéresser. Dans les formations dédiées¹⁴, comme le master Mathématiques-Vision-Apprentissage de l'ENS Paris-Saclay, de nouveaux cours s'ajoutent chaque année, issus de champs de recherche aussi divers que la topologie algébrique ou les équations aux dérivées partielles. Dans sa dimension académique, l'apprentissage entremêle trois profils rendus distincts par leurs moyens de publication scientifique : les mathématiciens, les analystes de données et les statisticiens. La théorie de l'optimisation est un pilier commun, car elle est un moyen puissant de sélection et donc un préalable à la décision. La confrontation à la notion d'incertitude ou d'erreur a pour sa part légitimé la description probabiliste. Cette interaction des métiers et des disciplines produit un retournement, car la spécialisation continue des thématiques

14. Pour une liste récente qui permet indirectement d'identifier les pôles d'enseignement et les spécialités : <https://www.actuia.com/actualite/etudes-en-intelligence-artificielle-le-grand-panel-des-formations-dossier/>

fait place à un décloisonnement progressif.

Cependant la dimension conceptuelle des écoles est largement occultée par la pratique. Comme décrit par [François et Frezal, 2016, p12], depuis les prémices de la formalisation probabiliste par Borel et Kolmogorov, les mathématiciens se sont éloignés des débats philosophiques sur leur objet. "La division du travail, au sein du monde intellectuel, est telle que la vigilance que les mathématiciens abandonnent à d'autres n'est en fait prise en charge par personne."

En effet une large part des techniques d'apprentissage automatique se situe dans la continuité de questionnements statistiques. Prenons l'exemple de la classification et d'un algorithme très courant d'apprentissage non supervisé : les k-moyennes. Il s'agit de subdiviser un ensemble en k classes correspondant à k représentants, les centroïdes, objets fictifs qui sont placés par l'algorithme de façon à couvrir au mieux les données suivant une logique de proximité. La proximité a ici un sens de distance, le plus souvent euclidienne, qui favorise les distributions normales autour des centroïdes, qui font office d'espérance statistique pour les données de leur classe. Le centroïde est un "homme moyen" de Quételet, généralisé à un contexte de mélange de populations (mixture de gaussiennes).

L'apprentissage automatique se situe ainsi à la confluence entre points de vue objectiviste, où l'algorithme "dévoile et approche un réel caché", et constructiviste, où l'algorithme "construit des objets en les dotant d'une relative stabilité" [Desrosières, 1993, p74]. Cette question se retrouve dans la formalisation mathématique qui sépare espérance issue d'une distribution statistique sous-jacente et moyenne empirique calculée à partir des seules données observées. Un tel raisonnement ne porte pas à conséquence quand les classes n'ont d'intérêt qu'en vue d'une étude. En revanche, les classes deviennent cruciales quand elles s'imposent à leurs éléments. Les règles de codification d'une population telles que celles pratiquées par l'administration fiscale sont héritées d'un long débat démocratique, ou du moins administratif, sur l'application de l'impôt. Les segmentations automatiques de populations d'internautes pour leur trouver une similarité à des fins publicitaires ou pour leur assigner un score de crédit peuvent avoir des conséquences tout aussi importantes, comme raconté dans [O'Neil, 2016], livre devenu incontournable sur les effets de la classification abusive.

A cet égard, la statistique a été de longue date confrontée à ses effets sociaux, ce qui a amené à considérer avec circonspection la définition de nouvelles classes d'objets. L'apprentissage automatique pour sa part produit avec grande facilité aussi bien des objets, au sens de nouvelles représentations, que des "formes pour décrire les relations entre les objets construits" [Desrosières, 1993, p79]. Dans son processus de codification, l'apprentissage automatique s'approche d'une position nominaliste : les classes sont arbitraires, susceptibles d'évoluer par ajout de variables. Elles ont une vocation essentiellement pratique. Cette position, héritière des "méthodes de Buffon", s'oppose à une classification "réaliste inspirée des systèmes de Linné" [Desrosières, 1993, p292-297], qui s'appuierait sur un nombre restreint de critères théoriques.

Cette absence de codification garantie par la théorie, fondée sur des données de qualité très variable mais en grand nombre, est étrangère à une approche systémique de statistiques. En particulier, si l'objet d'étude est sujet à un environnement variable dans le temps, comme pour beaucoup de grandeurs socioéconomiques, il faut une définition axiomatisée d'une classe ou d'un indicateur pour suivre son évolution. L'évolution d'un apprentissage dans le temps est complexe, elle s'appuie sur des fenêtres qui font office de temps caractéristiques des phénomènes et est souvent coûteuse en matière de calculs. Néanmoins, à la décharge de l'apprentissage, l'objet statistique et les modèles figés sont encore plus vulnérables au moment d'une transition. Si la structure de l'échantillon varie fortement (pensons par exemple au portefeuille de clients d'EDF en période de concurrence) les règles d'hier n'ont plus nécessairement cours. L'apprentissage permet de recalibrer les modèles et sert de moyen transitoire pour s'adapter.

Quelle histoire de l'IA ?

Pour qui souhaiterait reconstituer l'histoire de l'intelligence artificielle, la plupart des références renvoient à la conférence de Dartmouth de 1956, où le nom aurait été inventé, ou bien au perceptron de Rosenblatt. Ce récit semi-mythologique embrasse ensuite les "hivers de l'IA" et l'opposition entre réseaux de neurones connexionnistes et systèmes experts. Ayant été écrit majoritairement par des cybernéticiens et des informaticiens, ce récit fait la part belle aux prédictions sur l'IA forte, en décalage complet avec les réalités, passée et présente.

Au contraire, l'histoire de l'apprentissage automatique et de ses débouchés actuels est plus le fait des statisticiens. Leur discours s'attache au volume croissant d'informations et à l'autonomie progressive d'un pôle "science des données". Leo Breiman en 2001 opposait la tradition de l'inférence formalisée de causes à l'émergence de la prédiction agnostique de valeurs. En 2015, David Donoho soulignait combien le cadre des compétitions de données avait forgé une nouvelle culture.

Les mathématiciens pour leur part ne racontent pas d'histoires, étant plus attachés aux récits d'individus héroïques qu'à la dimension collective d'une communauté.

Pour qui souhaiterait approfondir ces points, je recommande :

- le billet ^a de Lionel Janin pour un excellent résumé des enjeux techniques et socioéconomiques ;
- l'article "De statisticien à Data Scientist" [Besse, 2015] et l'introduction ^b sur la transition de l'enseignement des statistiques vers la science des données ;
- la réflexion "50 Years of Data Science" [Donoho, 2017] sur les problématiques présentes et futures de la discipline.

a. http://www.strategie.gouv.fr/point-de-vue/lintelligence-artificielle-mythes-realites#_ftn4

b. https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-lm-Intro-Stat_SD.pdf

3 Contextualisation technique de l'apprentissage automatique

3.1 Chaîne de production et de valorisation des données

L'apprentissage automatique est un bloc algorithmique qui s'insère dans une chaîne d'activités autour des données. Il n'a pas de valeur propre, excepté comme objet d'étude pour les chercheurs. Je propose de décrire ainsi cette chaîne de production et de valorisation des données :

- acquisition : le recueil de l'information,
- traitement : le formatage numérique de l'information, par la sélection des variables et des échantillons,
- standardisation : l'élaboration d'un référentiel commun d'étude pour s'assurer d'une représentativité,
- stockage et calcul : l'architecture informatique pour conserver et manipuler l'information,
- modélisation : la représentation de l'information sous d'autres formes,
- aide à la décision : l'application de la modélisation à la production d'autres informations,
- utilisation : la finalité de l'information recueillie, une fois transformée par les autres maillons.

Comme discuté à la partie précédente, l'apprentissage automatique vient théoriser les maillons de modélisation et d'aide à la décision d'après des données. Il complète ou supplante des méthodes antérieures pour limiter l'intervention humaine dans le choix des paramètres. Si un élément est modifié dans un tel assemblage, quelles en sont les répercussions possibles sur les autres niveaux de la chaîne ?

A la différence de l'expérimentation qui court-circuite les différentes phases pour se concentrer sur un résultat, au contraire de la méthodologie scientifique qui revient à explorer intensément un maillon unique, l'industrialisation d'un procédé d'étude de la donnée requiert de concevoir la chaîne de valorisation dans son ensemble. L'ex-

haustivité du rapport Villani [Villani, 2018] montrait bien que, pour discuter d'intelligence artificielle, il était possible de consacrer des parties entières aux données ou à la dimension éthique des études poursuivies. En revanche, discuter de la chaîne est une attitude plus propre aux industriels qu'aux chercheurs. Le thème de l'intelligence artificielle touchant à l'immatériel, la matérialité de sa concrétisation se trouvait plutôt dans les études sur le Big Data. En effet, l'apprentissage automatique a déjà fait l'objet de rapports sous le nom des "algorithmes du Big Data" [AdT, 2015, p17-24].

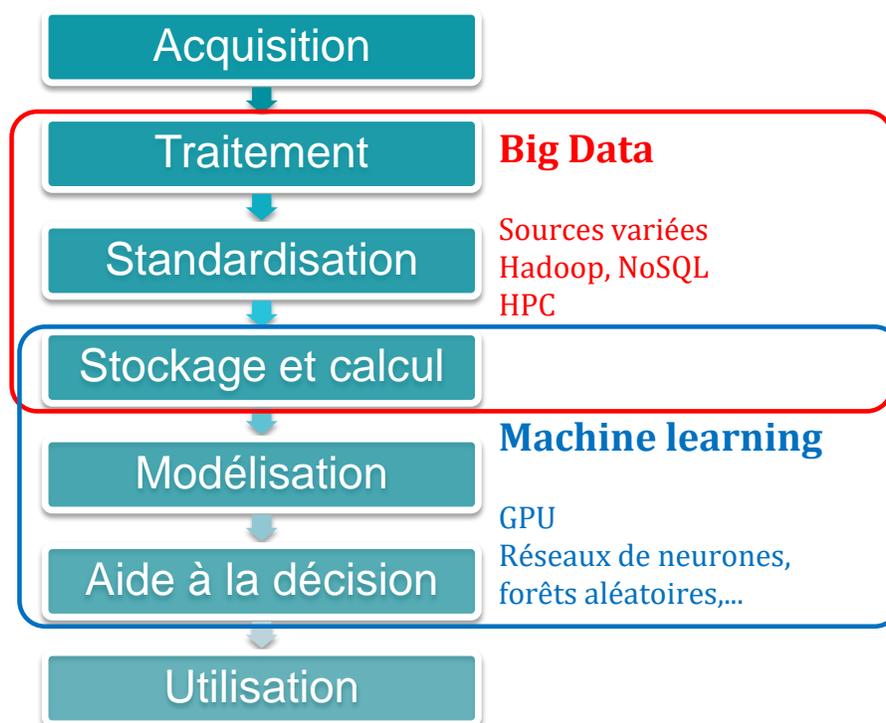


FIGURE 3 – Chaîne de production et de valorisation des données ainsi qu'effets principaux du Big Data et de l'apprentissage automatique (*machine learning*)

La Figure 3 représente l'interaction entre la chaîne de production et de valorisation des données et des thématiques, comme le Big Data ou l'apprentissage automatique. La notion de Big Data préfigurait la discussion sur l'utilisation de l'apprentissage en s'attachant aux caractéristiques des données¹⁵ et à leur conservation sous des architectures distribuées comme Hadoop. De nouveaux principes informatiques

15. Ce sont les fameux "3V" : volume, vitesse et variété.

étaient discutés comme la possibilité de s'éloigner des langages de requête structurée (SQL), en remettant en cause les fondements des bases relationnelles. Le Big Data était par conséquent un sujet investi préférentiellement par des industriels et des informaticiens. L'accent mis sur l'intelligence artificielle a dégagé la discussion de sa matérialité pour toucher de nouveaux cercles, en statistique et même en sciences sociales. Seule la notion de moyen de calcul a été conservée dans les derniers rapports [Villani, 2018, p58-59], les processeurs graphiques (GPU) étant devenus le nouveau support de prédilection. Ces derniers sont mieux adaptés aux opérations linéaires et parallélisées que les calculateurs à haute performance (HPC).

Au-delà de la question des infrastructures physiques, il faut se poser la question des personnes qui participent à la chaîne. Les différentes tâches de la chaîne peuvent être effectuées par une personne unique qui adopterait plusieurs postures, ou bien par différentes personnes, chacune étant spécialisée. Le second cas de figure est propre au passage de l'expérimentation au déploiement d'un procédé. Le déploiement repose sur une division des tâches motivée par une rationalisation taylorienne. La spécialisation et la technicité croissante des tâches requièrent pour leur part une compréhension des principes généraux de chaque activité, même en l'absence de pratique. La compréhension entraîne la confiance vis-à-vis du produit des autres tâches. En conséquence, l'introduction d'un nouvel élément vient perturber des pratiques établies, ce qui érode la confiance, l'objet méconnu étant taxé de "boîte noire".

La conduite du changement ne se borne donc pas à la conception des nouvelles pratiques, elle requiert une intégration dans un environnement préexistant. La difficulté tient à devoir réadapter les autres maillons pour aboutir à une cohérence. Dans le cas de l'apprentissage, l'Académie des Technologies [AdT, 2018, p54] cite le verrou que constitue l'acquisition d'une base de données adaptée et l'établissement d'un protocole de validation de l'apprentissage. Il n'est pas lieu de commenter ici les questions plus générales d'innovation ou de transformation numérique, comme par exemple les huit étapes du changement de Kotter¹⁶. En revanche, je me concentrerai sur le contexte métier dans lequel l'arrivée du maillon d'apprentissage automatique intervient.

16. <https://www.kotterinc.com/8-steps-process-for-leading-change/>

3.2 Cultures métiers et interaction avec l'apprentissage

Passons en revue, de façon archétypale, les corps de métiers de la chaîne présentée précédemment. Les métiers et attitudes correspondants (opérateur - métrologue - statisticien - informaticien - modélisateur - décideur - utilisateur) ont des logiques distinctes, mais pour parvenir à un fort niveau de complexité dans le service et dans les modèles, il faut que la chaîne repose sur la confiance lors des traitements successifs. Chaque processus peut ensuite être affiné pour tenir compte de la spécialisation croissante des métiers. L'amélioration peut se faire au détriment de la robustesse de l'analyse globale, en faveur de la finesse du modèle, ou bien en rendant moins intelligibles les autres étapes pour un métier donné.

La crainte majeure pour un métier confronté à l'automatisation de ses tâches est

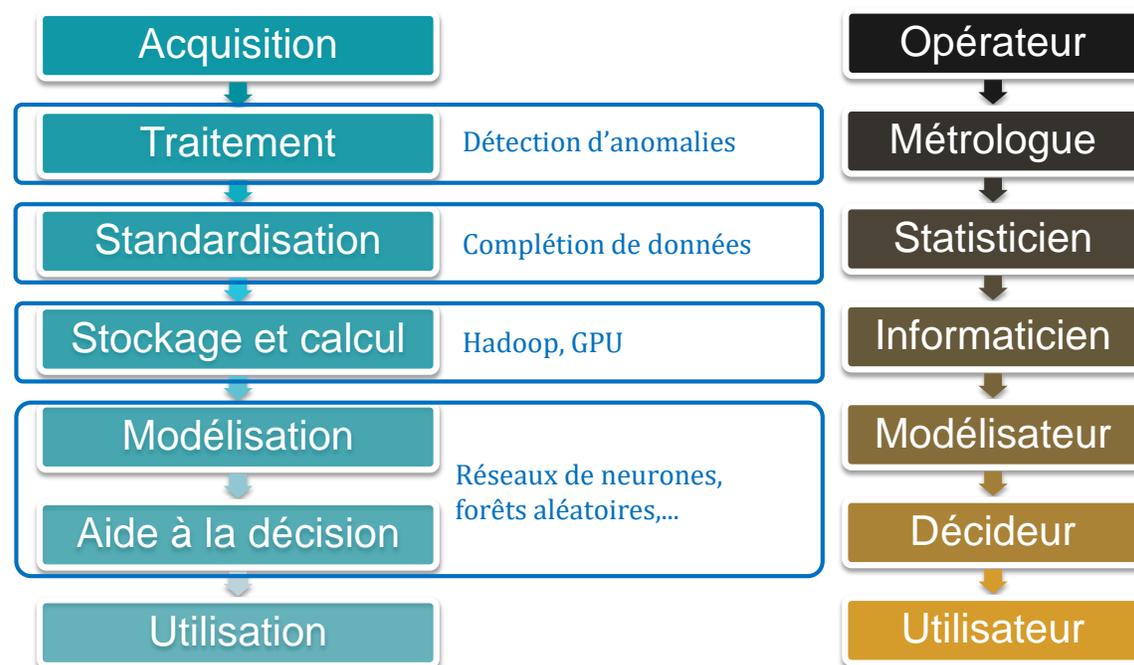


FIGURE 4 – Chaîne de production et de valorisation des données mise en regard des métiers et attitudes associés, ainsi qu'effets principaux de l'apprentissage automatique (*machine learning*)

de voir actée sa disparition. La "décision par algorithmes" qui laisserait envisager un "monde sans experts" [Vayatis, 2016] frappe aussi bien les traders, les publicitaires que les bibliothécaires. En effet, les potentialités de l'apprentissage automatique ne

se cantonnent pas au bloc que nous avons grossièrement dégrossi à la Figure 3. La Figure 4 complète ces étapes et souligne que la plupart des métiers et attitudes sont affectés. Dans la plupart des cas, l'opinion de l'agent concerné oscille entre curiosité et crainte de perdre le contrôle. Schématiquement, en voici les réactions :

- **opérateur** : l'acquisition de données est au cœur de l'apprentissage. Cependant, la numérisation requise signifie que la compétence humaine est moins sollicitée, surtout s'il faut une simultanéité avec l'analyse. Patrice Aknin, ancien directeur de l'innovation de la SNCF, témoigne des difficultés à passer à l'inspection par drones, sujet qui concerne aussi le Cerema. A l'AFB, François Hissel remarque que si l'automatisation des mesures d'hydromorphologie des berges est acceptée, l'automatisation de l'estimation des populations de poissons le serait bien moins.
- **météorologue** : l'inquiétude porte ici sur la qualité de la donnée quand elle provient de sources variées, problématique inhérente au Big Data. L'apprentissage automatique est perçu plus positivement car utile pour le calibrage, surtout d'instruments à bas coûts. Bérengère Lebental de l'IFSTTAR mentionne qu'en matière de calibrage, "avant 2010, il était anathème de parler d'autre chose que de modèles physiques, compléter un modèle par des statistiques était aberrant. Aujourd'hui le nerf de la guerre est d'arriver à exprimer la question comme un problème d'apprentissage".
- **statisticien** : la statistique publique associe une appétence pour l'hybridation à une défiance quant à la portée de ces évolutions. "Ces statistiques expérimentales pourront apporter des informations inédites sur ces domaines émergents ou à mieux couvrir, mais sans pouvoir être immédiatement mises sur le même plan que des productions régulières éprouvées sur la longue période." [Enjeux Numériques, 2018b, Blanchet et Givort, p84]
- **informaticien** : l'architecture informatique a été plus exposée dans le cadre du Big data ou de la politique d'Open data. L'apprentissage automatique pour sa part s'est consolidé plutôt autour des processeurs graphiques comme moyens de calcul. En cela, même les établissements habitués à la donnée, comme l'IGN ou Météo-France, ont été pris en défaut et leurs systèmes d'information doivent s'adapter.
- **modélisateur** : les modèles des experts scientifiques sont mis en concurrence

avec des modèles automatiques qui ont pour propriété de passer à l'échelle. Ainsi, Vincent Labbé du BRGM note que les modèles empiriques et locaux des hydrogéologues pourraient être systématisés afin de mailler le territoire tout en tenant compte des spécificités locales.

- **décideur** : le décideur est normalement le premier bénéficiaire puisqu'il dispose d'une nouvelle méthode pour l'aider et ainsi est amené à se consacrer aux cas les plus difficiles ou les plus extrêmes. Cependant, par la complexité des techniques à l'œuvre, il peut aussi être cantonné à un rôle de validation. La transition actuelle des prévisionnistes Météo-France en aval de la chaîne de décision en témoigne (voir Section 4.2.2).
- **utilisateur** : par ses retours d'expérience l'utilisateur est un élément clé du dispositif, mais le manque de compréhension de la provenance de l'information et de sa réutilisation dans le cadre de l'apprentissage peut engendrer une défiance. Le compteur électrique communicant Linky d'Enedis en est un exemple emblématique.

L'accélération en cours pour atteindre la vitesse du numérique est à elle seule un facteur perturbateur aggravant pour les métiers. L'inertie de chaînes de valorisation des données fermement enracinées, parfois tout à fait efficaces, en est contestée. A ce titre, l'apprentissage automatique s'inscrit dans la suite de la transformation numérique et a résolument pris la direction de l'*Open source*, là où l'*Open data* ne s'installe que progressivement, et très peu dans le monde privé. L'accessibilité de la programmation informatique et la standardisation par les bibliothèques de programmes permettent de disséminer les techniques plus vite aux différents niveaux. Après la baisse drastique des coûts de diffusion, puis des coûts de stockage, c'est au tour des coûts de traitement de l'information de s'effondrer.

3.3 Se positionner vis-à-vis de l'apprentissage

La plupart des entretiens conduits pour ce rapport soulignaient le besoin de positionner clairement les organismes du RST dans la chaîne de données. Chaque organisme présente en effet toutes les facettes métiers, mais les implique différemment sur ses projets. Savoir implémenter un réseau de neurones est certainement d'intérêt

pour qui produit des algorithmes, mais savoir évaluer la qualité d'un modèle est plus utile pour qui doit qualifier des solutions. Maîtriser au moins une technique permet de définir une référence (*baseline*) dans une compétition, mais pour qui doit définir le problème il faut surtout connaître les principes de l'apprentissage automatique.

Une position intéressante dans ce contexte est celle de concepteur de jeux de données pour compétition, car elle permet d'engager très rapidement le dialogue avec la communauté et, dans certains cas, d'avoir un point d'observation privilégié des dernières techniques. Sur la reconnaissance de panneaux de signalisation, les jeux de données allemands¹⁷ ont joué un tel rôle. S'agit-il de sous-traitance ou de sur-traitance? Dans un cas, cela signifie que l'organisme conserve le contrôle global sur la chaîne, dans l'autre cas, qu'il se positionne en plateforme pour accompagner le développement. Abandonne-t-on de la valeur à des extérieurs? Peut-être, s'il y a concurrence avec une production interne, mais tous ont à gagner à la résolution de la tâche fixée. Il faudra toujours un jeu de données, même minimal, pour déterminer la faisabilité d'un problème avant de rechercher la solution la plus performante. En outre, définir un problème analogue permet d'écarter les questions de confidentialité.

Les deux exemples qui suivent illustrent cette question du positionnement et de la relation aux chercheurs et data scientists. Dans les deux cas, il s'agit d'un problème d'aide à la décision et d'alertes pour des opérateurs.

Dans le domaine maritime, la DAM cherche à desserrer les liens entre administration et administrés et passe d'un contrôle systématique des bateaux a priori à un contrôle a posteriori. Un outil de police prédictive est envisagé pour être moins intrusif mais les données collectées ne sont pas interopérables. Il faut d'abord refonder la chaîne amont de la collecte jusqu'au stockage.

La DAM sert par ailleurs de point de transit pour la remontée des données de géolocalisation des navires. Les agents s'acculturent à un principe d'alerte à partir de scénarios préétablis de navigation. Ces scénarios définis par l'agence européenne pour la sécurité maritime (AESM) font ressentir le besoin d'une adaptation de l'algorithme à l'agent-décideur. En termes culturels, la DAM est confrontée à une tradition d'actions de police plutôt que de production de connais-

17. <http://benchmark.ini.rub.de/>

sances. L'échange international est limité, pour des raisons de souveraineté ou de diversité des situations. L'Association internationale de signalisation maritime (AISM), dont les travaux techniques sont nombreux et incontournables, n'a pas encore abordé de manière significative les sujets inhérents à l'intelligence artificielle.

Dans l'aviation civile, le STAC en est au début de sa réflexion concernant l'IA. Les deux pôles auditionnés (sûreté à Bonneuil et navigation aérienne à Toulouse) ont des approches et des contextes très différents au-delà de l'enjeu commun d'évaluer l'utilité des techniques algorithmiques.

Sur l'inspection des bagages les plus difficiles à analyser, la division sûreté pourrait constituer rapidement une base importante, annotée par les prestataires de sûreté. Elle souhaite découvrir les technologies et les rendre robustes au masquage ou aux leurres. En effet, la division sûreté du STAC produirait ou certifierait les solutions sans en être utilisateur direct. Un contact infructueux a été pris avec les INSA (Rennes et Lyon) qui proposaient une thèse, et donc un horizon en décalage avec les besoins du STAC qui pour 2019 doit être en mesure d'évaluer les produits en laboratoire et développer une spécification technique.

A la navigation aérienne, l'horizon est plus lointain, 2023 pour améliorer le système d'anticipation des collisions, 2030-40 pour l'intégration dans les systèmes techniques de contrôle aérien. Les collisions potentielles sont aujourd'hui des alertes fondées sur la distance entre avions. La division navigation aérienne considère l'opportunité de l'apprentissage automatique pour les anticiper de plusieurs minutes. Les prestataires IT sont décisionnaires sur ce point, la DGAC ne travaillerait que sur la certification, puis serait récipiendaire de ces logiciels pour ses contrôleurs. Un autre sujet pourrait être l'identification automatique des avions. Dans tous les cas, la communication européenne primerait sur l'effort national.

4 Enjeux pour le RST en matière d'apprentissage automatique

4.1 Quelles incitations à agir ?

Quelle urgence à agir en matière d'intelligence artificielle ? La notion est exhibée comme un "enjeu d'avenir" par des démarches gouvernementales comme #FranceIA, affirmation reprise en chœur par les revues spécialisées. Il faut pourtant de fermes incitations pour vaincre l'inertie d'institutions qui ne manquent pas d'activités et qui sont aux prises avec une diminution constante de postes. Deux principes émergent en la matière : "Nécessité fait loi" et "gouverner c'est prévoir". Le premier principe a vocation interne, le second renvoie à la gestion collective des enjeux.

"Nécessité fait loi." L'apprentissage automatique relève d'un changement de pratiques. En interne, il se développe essentiellement suite à deux contraintes, la saturation de l'emploi existant et la complexité croissante des tâches. La saturation de l'emploi est une contrainte partagée par la vaste majorité des administrations dans une période prolongée de rationalisation budgétaire. Elle est certes répartie inégalement entre et au sein des institutions, mais rares sont celles à échapper à la baisse des plafonds d'emploi. Pour s'adapter au besoin de faire autant ou mieux avec moins, les organismes considèrent l'automatisation de certaines tâches. Des tâches individuelles, ce sont des postes entiers qui peuvent être transformés. L'automatisation devient par ailleurs une compétence à acquérir, l'apprentissage automatique en étant un avatar. La complexité croissante des tâches demandées amène à repenser la place de l'être humain dans les dispositifs. Elle conduit à un effort de systématisation et de pré-traitement de l'information. L'adage "Nécessité fait loi" traduit aussi le fait qu'en l'absence de besoin pressant, l'inertie prendra le pas. C'est pourquoi les entretiens ont montré que la réflexion la plus développée sur la transformation numérique se trouvait dans les établissements les plus exposés à la concurrence, comme l'IGN et Météo-France.

"Gouverner c'est prévoir." Cette maxime s'adresse aux coordonnateurs que représentent les directions des institutions du RST, ainsi qu'à la DRI qui les accom-

pagne. Le rôle de ces groupes est d'anticiper les virages pour ne pas être pris au dépourvu, afin de ne pas réagir dans la fébrilité lorsque la contrainte se présente. Néanmoins, il faut raison garder et ne pas s'emparer des sujets avec cette même fébrilité que l'on souhaite s'éviter. Ce serait oublier la finalité pour ne se consacrer qu'à l'objet. Toute la difficulté pour une direction pour qui penser est une part de l'action est justement que le sujet de l'intelligence artificielle est pensé en même temps que les techniques s'inventent. Édifier un plan d'action immédiate est très complexe au vu des ramifications vis-à-vis des secteurs et des métiers, ainsi que de l'évolution rapide du sujet. En revanche, l'identification des finalités d'un déploiement de l'apprentissage automatique et des prérequis à rassembler est un objectif commun à toutes les directions. Je reprends à mon compte la préconisation du rapport de l'Académie des Technologies [AdT, 2018, p12] :

« L'IA est un outil, pas une fin en soi, mais la maîtrise de cet outil est bien un enjeu stratégique. Cette maîtrise doit s'appuyer sur des compétences dont il faut favoriser l'émergence, ce qui signifie que cela prend du temps et qu'il n'y a pas de relation directe de cause à effet, donc que le volontarisme n'est pas suffisant. »

Plus indirectement, des thématiques comme l'intelligence artificielle peuvent servir des fins détournées pour pousser à la transformation numérique. Le mot à la mode n'est alors que l'éperon du changement et il est d'autant plus efficace qu'il est hors de portée. La technique devient alors sa propre finalité et sa poursuite en est l'incitation intrinsèque. L'injonction répétée à évoluer risque cependant d'entraîner une insensibilisation progressive, une mithridatisation. La partie qui suit discutera par conséquent des finalités concrètes pour un établissement du RST à s'approprier l'apprentissage automatique.

4.2 Quelles finalités pour le RST à s'approprier cette technique ?

Le prisme de l'intelligence artificielle permet de souligner les objectifs et contraintes partagées par les établissements. La place que pourrait prendre l'apprentissage automatique en dépend. Nous verrons que pour servir une logique d'interaction avec un public, externe ou interne à l'établissement, il manque un facteur d'échelle dans les

interactions actuelles. Néanmoins, l'apprentissage dont nous parlons découle d'une automatisation de la modélisation ou de la prise de décision. Toute activité qui se rattache à ces thèmes peut donc opportunément envisager la possibilité d'incorporer des briques algorithmiques ou au moins de s'y comparer. Du point de vue des établissements, le thème de l'intelligence artificielle peut répondre à différents enjeux : **disposer d'une compétence supplémentaire, gagner en efficience, élargir le public concerné, produire de nouveaux services**. Détaillons ces différents aspects :

- **Disposer d'une compétence à faire valoir** : Cyniquement, l'image d'un établissement du RST peut bénéficier à peu de frais du battage médiatique sur l'intelligence artificielle¹⁸. Plus utilement, cela permet de répondre aux questions de décideurs, des ministères ou des collectivités, qui veulent déployer de "l'intelligence" dans leurs missions. De telles aspirations, avec un degré de préparation variable, ont été rencontrées par Stéphane Trainel du CGDD au cours des datathons organisés autour des données mises à disposition. En revanche, le milieu de l'apprentissage automatique étant très actif et changeant, il faut une pratique régulière et donc une association à un métier de l'établissement pour maintenir les compétences dans le temps.
- **Gagner en efficience** : Une loi historique rappelée par [France Stratégie, 2017] stipule que les tâches laborieuses et à faible valeur ajoutée sont les premières automatisées. Bien des métiers, même à forte valeur ajoutée, comportent de telles tâches. Il faut cependant une répétition fréquente, individuellement ou collectivement, pour qu'automatiser soit opportun. L'apprentissage excelle aujourd'hui dans les traitements automatiques du texte, de la parole ou l'image. Ces tâches peuvent être présentes au sein du RST, ou de ses interlocuteurs, en particulier plus opérationnels. On peut penser aux DREAL pour l'AFB, aux DIR pour le CETU. Ces acteurs peuvent eux aussi bénéficier de conseils pour l'automatisation de leurs activités.
- **Élargir le public d'utilisation** : Si la saturation des services entraîne une limite à l'activité, l'automatisation de l'analyse des données permet son ex-

18. Il faut à ce titre noter que pratiquer la novlangue autour de l'IA requiert un manque de déontologie qui n'est pas, par bonheur, si répandu dans la fonction publique.

tension à un public plus large, ou à d'autres publics. Un bloc d'apprentissage automatique adéquatement placé et calibré peut transformer la sortie d'un modèle principal, laissé pour sa part inchangé. Ce post-traitement permet d'envisager une information taillée sur mesure pour un destinataire qui ne retire pas actuellement d'utilité de l'information généraliste. Météo-France affine ainsi par des techniques de réseaux de neurones ses zones de prévision de pluie sans modifier son modèle de prévision.

- **Produire de nouveaux services** : Expérimenter des méthodes d'apprentissage permet d'imaginer d'autres offres pour un établissement à partir de ses données actuelles, et surtout sur de nouveaux jeux de données. Ce point rejoint la stratégie d'innovation et de valorisation d'un établissement. Un risque en la matière est de sous-estimer le taux naturel d'échec, car inventer un produit est aussi difficile dans un établissement qu'en startup. Formuler le besoin d'un public, en particulier pour des institutions qui n'échangent qu'avec un public restreint, ne fait pas nécessairement partie de la culture des établissements.

4.2.1 Disposer d'une compétence à faire valoir

Expérimenter va de pair avec l'objectif de compétences, objectif partagé de façon indiscriminée par tout le RST. Il serait par conséquent évident de préconiser à toutes les institutions de se doter de nouvelles compétences, comme l'apprentissage automatique. La mode de l'intelligence artificielle s'étant installée, les institutions pourraient faire valoir leurs capacités d'innovation à bon compte auprès de leurs financeurs. S'il est possible de suivre la tendance à peu de frais en requalifiant des métiers déjà présents, il est difficile de soulever d'autre objection que la tromperie d'un public déjà abusé par bien des publicités.

Il est cependant illusoire d'imaginer séparer l'apprentissage automatique de ses cas d'études. C'est pourquoi l'expérimentation n'est qu'un volet qui doit préfigurer une activité régulière si la compétence est jugée nécessaire d'être conservée. En effet, il faut une pratique constante, une mise en compétition et une capacité d'adaptation pour affirmer que la compétence de sciences des données est bien présente. Le milieu de l'apprentissage automatique évolue vite. Considérer en avoir fait le tour

pour le mieux fossiliser accroît le risque d'obsolescence. L'apprentissage s'appuie sur des données sans cesse renouvelées et d'origine variée. Cette compétence ne peut être appréhendée au même titre qu'une compétence métier qui jamais ne se dévaluerait. Il ne s'agit pas de conserver une alternative à d'autres modèles dans une logique autarcique et régaliennne. Il ne s'agit pas de dissuasion nucléaire ou d'abris anti-atomiques. Si elle n'est pas confrontée régulièrement à la pratique, et donc à un usage et à un destinataire, cette technique dépérit.

L'IFSTTAR aborde le thème de l'apprentissage statistique comme une facette de la "raquette de compétences" ^a à maintenir pour répondre aux attentes du ministère. L'absence de bases de données en transport qui soient centralisées, structurées et accessibles, est un frein majeur pour dépasser la simple étude et passer à l'échelle. Des chercheurs ont individuellement analysé les données Vélib à Paris ou les données de mobilité billettique à Rennes. Mais les données des cartes Navigo n'ont par exemple pas été relevées à des fins d'exploitation.

^a. Entretien avec Serge Piperno

L'IGN s'appuie sur l'apprentissage pour annoter (sémantiser) les images 2D ou 3D qu'elle recueille. En 2D, la classification automatique des surfaces agricoles sera utilisée pour l'automatisation des déclarations PAC si le consortium auquel l'IGN participe remporte ce projet Horizon Europe. En 3D, dans le cadre des projets iTowns et Stereopolis, il s'agit plus de développer la compétence que de procéder à une production de données systématique, ceci afin d'anticiper la standardisation de la remontée de données et d'exporter la technologie de collecte hors de l'IGN. L'apprentissage par les exemples avait déjà provoqué à partir de 2005 un "véritable changement culturel" ^a en supplantant la grammaire figée des types de bâti. Un algorithme de l'IGN est pour le moment en tête de la compétition internationale Semantic3D.

^a. Entretien avec Nicolas Papanoditis

4.2.2 Gagner en efficience

L'automatisation n'est pas le propre de l'apprentissage statistique. C'est pourquoi il n'est pas rare d'associer intelligence artificielle et robotisation, comme une déclinaison concrète. L'intérêt du point de vue statistique réside surtout dans l'accès à une vision d'ensemble, qui permet de constituer une équivalence entre des activités locales et de dépasser les contingences géographiques. L'efficience globale provient de façon prédominante d'un renouvellement des techniques de mesure. L'apprentissage automatique accompagne ces nouvelles techniques qui induisent une discontinuité des pratiques. Par exemple, au Cerema, l'image qualité des routes nationales (IQRN), collectée en trois ans par la machine Aigle RN pour la couverture complète, le sera en une année seulement par Aigle 3D grâce à la télédétection laser (lidar). Cette même machine permettra au CETU de passer d'une inspection visuelle en deux nuits avec fermeture du tunnel à une simple traversée en une heure. L'outil technique permet à lui seul le gain de temps sur le terrain, mais si le traitement de l'image est effectué par un opérateur, comme cela a été proposé au CETU par un bureau d'études, le gain n'est que marginal.

Dans un institut de recherche, le gain peut se trouver dans la réduction du nombre d'expériences, physiques ou numériques, nécessaires pour explorer des situations. Dans un cadre opérationnel, il peut s'agir du pré-traitement d'une image, d'un texte ou d'un son pour accélérer le travail d'expertise. Bien que le RST soit de moins en moins associé à des tâches directement opérationnelles, il est souvent requis pour des activités de conseils. L'AFB, le MNHN et l'IRSTEA accompagnent ainsi les DREAL au sujet des algues diatomées, dénombrées au microscope dans les DREAL sans reconnaissance automatique. Les DIR en pleine réforme sollicitent de façon croissante le CETU, qui choisit ses cas d'études. Une doctrine sur l'automatisation peut donc servir à d'autres organismes.

A Météo-France, le gain d'efficience par l'apprentissage est élevé au rang de projet d'établissement et a reçu 2,9M d'euros dans le cadre du fonds pour la transformation de l'action publique (FTAP). Le métier de prévisionniste y est exposé à une information toujours plus importante suite aux prévisions d'ensemble à un rythme horaire. Les modifications qu'il apportait aux sorties des modèles de prévision ne sont plus possibles ou nécessaires, les bulletins qu'il rédigeait peuvent être désormais

générés automatiquement avec une qualité suffisante.

Exemples ^a des progrès obtenus dans la génération automatique de textes à partir de prévisions :

« 1) *Avant* : *Nuages masquant le soleil.*

1) *Après* : *En début d'après-midi, les nuages sont bien présents, mais quelques éclaircies sont encore possibles. Ensuite le ciel se couvre complètement jusqu'à la fin de journée.*

2) *Avant* : *Beau temps en dépit de quelques passages nuageux.*

2) *Après* : *Des Alpilles au golfe de Fos, les nuages sont nombreux le matin, mais bien vite le soleil l'emporte. Ailleurs, il brille sans partage toute la journée. »*

Ceci induit une "mutation de l'établissement" qui touche à un des ses "cœurs du métier" ^b. Les prévisionnistes d'hier, physiciens de l'atmosphère, deviennent des ingénieurs-conseil devant interpréter des distributions statistiques. Si la décision finale en matière de vigilance reposera toujours sur une décision humaine, ce métier se déplace en aval de la chaîne d'aide à la décision pour se tourner vers l'utilisateur.

^a. Gracieusement fournis par Olivier Rivière

^b. Entretien avec Marc Pontaud

La division Sûreté du STAC souhaite fournir une aide à la décision au contrôle des bagages à partir des informations renvoyées par les équipements de sûreté. Ces informations incluent, pour les appareils de détection automatique d'explosifs (EDS), l'interprétation d'alarmes et d'images radioscopiques (2D et/ou 3D). Il ne s'agit a priori pas d'automatiser les cas simples ou répétitifs. Le véritable gain serait pour le traitement de situations complexes, comme les masquages, qui peuvent aller jusqu'à requérir des équipes de déminage. Des essais avec des arbres de décision sont à l'étude en interne. Pour l'apprentissage profond, le STAC pourrait se tourner vers l'extérieur. Les techniques seraient ensuite analysées à l'aune de l'expertise des meilleurs opérateurs. A cause de la nature sensible du sujet, techniquement et socialement, le STAC produirait ou certifierait les solutions, bien que s'en remettant aux prestataires de sûreté pour l'utilisation.

4.2.3 Apprendre du public, penser au destinataire

Nous avons vu que l'apprentissage automatique s'inscrit dans la continuité de nombreuses techniques, parfois en se contentant d'englober l'existant. Expérimenter ces méthodes, dont le coût d'accès et de programmation est si faible aujourd'hui, est envisageable par bien des institutions. Pour autant quelles sont les raisons pour pousser l'expérimentation au sein de leurs services au-delà d'une activité scientifiquement récréative ? La pratique peut bien entendu générer des idées de nouvelles applications, mais c'est en disposant d'une finalité que l'objet reçoit sa pleine dimension et que l'on détermine si les prérequis pour sa mise en pratique sont remplis. Une technique d'apprentissage automatique intéresse surtout par ses retours, à qui sont-ils faits ? De nombreux retours indirects sont possibles. Ils peuvent être proposés à l'opérateur pour accélérer l'acquisition des données, au responsable de la base de données pour en améliorer la qualité, au modélisateur pour comparer sa pratique. Cependant c'est à l'utilisateur qui consulte l'oracle qu'une telle technique est la mieux destinée.

En effet la plupart des exemples d'apprentissage automatique mis en avant dans la presse sont ceux d'interaction avec le public. Que ce soit dans la banque, dans le transport ou en médecine, pour reprendre les thèmes choisis par [France Stratégie, 2017], mais aussi dans la publicité ou l'assurance, nous retrouvons des requêtes nombreuses d'un public qui accepte, et même parfois souhaite, une relation avec une machine. L'apprentissage allant dans tous ces cas de pair avec le Big Data, la présence du public permet d'accroître la taille de la base, par ses traces Internet, par son utilisation d'un service. Cette configuration rend possible l'apprentissage par le retour des utilisateurs, apprentissage de nature agnostique, sans nécessité d'un cadre trop imposé. En effet la quantité de données et leur origine numérique, et donc sûre, à défaut d'être complète, compense le peu d'hypothèses de travail originellement disponibles.

Or les institutions du RST ne disposent pas de ce public massifié. Elles ont des interlocuteurs privilégiés, et au premier titre : l'Etat. A qui sont destinées les productions du RST ? Bien entendu il n'est pas question d'apporter une réponse unique, ce sont aussi bien des collectivités, des entreprises, jeunes pousses ou non, d'autres services de l'Etat, opérationnels ou non. En revanche, si le citoyen est toujours servi

dans le cadre des missions de service public, il est rarement le destinataire direct d'un produit¹⁹. Celui-ci lui sera transmis par les autres utilisateurs, ce sont donc aux objectifs de ces derniers que le RST participe. La conséquence est de ne pouvoir transposer aisément des modèles d'innovation qui doivent leur succès à une standardisation du rapport à l'utilisateur. S'il est aisé de se représenter l'automatisation au sein des services d'imposition des Finances Publiques [Villani, 2018, p65], ou à Pôle Emploi [Enjeux Numériques, 2018a, Chapuis, p38-43], il est bien plus difficile de la penser au Cerema dans la constitution des études de transport pour les collectivités. Alors que d'un côté, l'utilisateur est diffus, virtuel, dans l'autre cas, le bénéficiaire est commanditaire et a des attentes spécifiques. Ceci vaut pour l'utilisateur externe. Il est aussi possible de penser un service pour l'utilisateur interne comme une recherche dans une banque d'informations, projet par exemple mentionné par le CETU. Mais là encore, il manque le facteur d'échelle. Aucun établissement du RST ne peut se prévaloir d'un personnel pour qui une solution commune soit profitable, comme pour les 20.000 conseillers du Crédit Mutuel²⁰ et leurs 300.000 mails quotidiens. Il faut donc chercher d'autres moyens qu'un échange croissant avec le public actuel pour justifier d'une procédure d'automatisation, ou bien diversifier le public.

Tout public requiert une information adaptée à son cadre de travail : une échelle départementale n'est pas adaptée à des projets municipaux, une moyenne ou une tendance annuelle répondent plus aux besoins d'un gestionnaire qu'à ceux d'un acteur local. Trop souvent, les outils du RST n'ont été pensés que pour un décideur public, services municipaux ou préfectoraux, dans un contexte antérieur aux technologies de l'information. Si le choix est fait d'élargir le public auquel les informations sont diffusées, il faut assumer d'abandonner la médiation par des notables. L'apprentissage automatique, suivant les différentes modalités de l'aide à la décision (descriptive, prédictive ou prescriptive), conduit à produire une information à partir d'exemples. Une approximation ou une interpolation pour raffiner une carte permet, sans changer les instruments de mesure, de fournir une information adaptée à des besoins plus

19. A titre d'exception, on peut mentionner les applications d'information comme celle de Météo-France, souvent des déclinaisons du cœur de métier, sans bénéficier du même degré d'excellence.

20. Le Crédit Mutuel collabore avec IBM pour développer une solution Watson analysant les mails, les annotant et pré-remplissant des réponses [Enjeux Numériques, 2018a, Rublé, p22-26].

locaux. L'idée d'étendre certaines informations existantes à un public élargi froisse la culture du secret, de maîtrise de l'information et de son utilisation. Ceci amène à la question de la transparence.

Transparence, contrôle et ouverture

On peut faire remonter la culture française du secret statistique à l'Ancien Régime [Desrosières, 1993, p40-41]. Plus récemment, un défenseur emblématique de l'Open data notait que "[les redevances de réutilisation traduisent], dans certains cas, une réelle inquiétude de l'administration de ne pas être en mesure de satisfaire les réutilisateurs, en termes de qualité des données et du service, mais aussi une crainte que les informations ainsi communiquées ne soient réutilisées pour critiquer le service public." [Trojette, 2013, p4]

La politique d'ouverture met à mal la volonté d'institutions-forteresses qui souhaiteraient obstinément garder le contrôle sur la diffusion et la réutilisation de leurs productions. Sauf à défendre une impérite radicale des autres organismes de l'Etat, il est judicieux de limiter l'ambition de contrôle, en particulier sur les conséquences politiques. Si l'information fournie amène à systématiser les bilans, cela permet de mieux responsabiliser (*accountability*) les décideurs publics.

A mesure que ses modèles s'affinent, Météo-France peut toucher des publics plus locaux. Par le post-traitement de la prévision, il est possible de définir les zones de pluie mieux qu'un expert. Météo-France est par ailleurs responsable de l'élaboration et de la diffusion de la vigilance météorologique, aujourd'hui au niveau départemental. Dans le cadre du contrat d'objectif performance 2017-2021, il s'agit de toucher l'infra-départemental tout en proposant une information pertinente. Sans remettre en question la technique de prédiction, la descente en échelle conduit à considérer de nouveaux besoins, pour la grande distribution, le secteur de l'énergie ou l'organisation des grands prix de Formule 1. Elle amène à se détacher de l'organisation administrative des services étatiques.

4.2.4 Produire de nouveaux services

Dans son livre-manifeste de 2013, Marianna Mazzucato défendait un rôle historique de l'Etat pour développer l'innovation, sans se restreindre à corriger les imperfections de marché : "Il faut considérer le problème depuis la représentation de l'Etat à travers ses agences (en les encourageant à être entreprenantes, ce qui inclut d'accepter plutôt que de craindre les échecs), jusqu'aux relations de l'Etat aux autres acteurs de l'innovation (en assumant un rôle plus actif)." [Mazzucato, 2013, p12] Au vu de la complexité de la chaîne de valorisation des données, l'échec est en effet un risque à accepter. De nombreux éléments peuvent manquer au moment de la réalisation, mais la méthode agile qui caractérise ces approches amène à apprendre de ses erreurs.

Produire de nouveaux services signifie ici plus que d'adapter ou de faire mieux avec l'existant. Il s'agit de dépasser les attentes pour créer un objet qui n'aurait que peu de parenté avec les sujets principaux d'une institution. L'enjeu central est ici de croiser les expertises et les données. Force est de constater qu'au cours des entretiens de nombreux projets ne sont ressortis que comme des potentialités. Le département Transport Santé Sécurité de l'IFSTTAR espère croiser ses données à celles du système national des données de santé (SNDS). Météo-France pourrait associer ses modèles météorologique et climatique aux modèles agronomiques. Mais d'autres projets, comme INSPIRE entre l'AFB et le BRGM autour de l'infrastructure de données Hub'eau, ont effectivement vu le jour.

Les projets les plus efficaces requièrent un porteur de projet à plein temps, situation parfois difficile à mener en interne pour les établissements. C'est pourquoi les porteurs peuvent aussi être extérieurs ou externalisés, temporairement ou définitivement. L'institution délègue alors l'innovation en se bornant à un rôle d'accompagnement. Loin d'être dégradant, ce dernier permet la montée en compétences des équipes associées au projet et de développer la culture d'innovation. Ce choix a été fait aussi bien à IGNFab, l'accélérateur de l'IGN, qu'à la Fabrique numérique, le pôle intrapreneuriat du MTES. A IGNFab, des projets comme In Sun We Trust associent startup, organisme et laboratoires de recherche, autour d'un porteur de projet extérieur. A la Fabrique numérique, un programme d'intrapreneuriat reconduit de manière semestrielle permet à des personnels de l'administration de se placer à la

tête d'une "Startup d'État". Dans les deux cas, la démarche a une forte composante d'acculturation, pour établir ce prérequis qu'est l'état d'esprit d'entreprendre.

4.3 Quels prérequis et quelles approches pour développer la technique au sein du RST ?

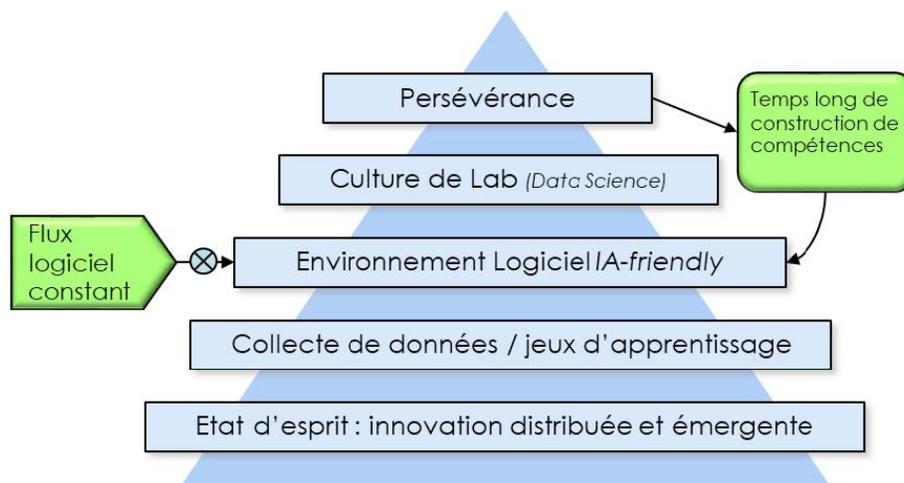


FIGURE 5 – Pyramide hiérarchique des conditions de mise en place d'une démarche IA [AdT, 2018, p71] (Source : Académie des technologies)

La culture métier est le point bloquant le plus souvent cité au cours des entretiens. Si l'on se réfère à la pyramide hiérarchique de la Figure 5, reprise de [AdT, 2018, p71], la dimension culturelle n'est pourtant qu'une étape tardive. Il ne faut pas confondre culture et état d'esprit. Ce dernier s'obtient par le décroissement là où la culture de sciences des données provient de la pratique. Il faut changer de milieu pour changer d'état d'esprit, avoir accès à d'autres profils auxquels on fait confiance et dont on ne mésestime pas les compétences. Une large section de ce rapport a été consacrée à recontextualiser scientifiquement le domaine de l'apprentissage automatique, pour rappeler qu'il n'était pas qu'une collection de solutions toutes faites, directement extraites d'une bibliothèque de programmes²¹, mais aussi une source de questionnements.

21. L'une des bibliothèques les plus emblématiques, scikit-learn, est d'ailleurs française, créé au sein de l'INRIA en 2010 <https://www.inria.fr/centre/saclay/actualites/scikit-learn-donne-de-l-intelligence-a-nos-systemes>

Points de blocages d'après [Villani, 2018, p48] :

« – une limite organisationnelle : les administrations ne sont pas structurées de façon à accueillir de l'IA, celle-ci étant par nature transverse à leurs missions ;

– un héritage historique : l'appropriation du sujet se heurte souvent à une culture et des modes de fonctionnement défavorables au développement de l'IA, en particulier pour ce qui relève des processus, des achats, des pratiques en matière de systèmes d'information, de l'exploitation, de l'acquisition et de l'ouverture des données ;

– un changement de paradigme : l'IA met notamment en défaut les logiques classiques d'expression du besoin et de spécification dans un contexte où les besoins émergent parfois de pair avec les solutions ;

– un effet de silo : le manque de réflexion transversale et prospective sur des usages futurs conduit à privilégier des systèmes conçus en vase clos, incompatible des développements futurs de l'IA. Ce manque s'accompagne souvent de la crainte de perdre la maîtrise sur ses données, une crainte qui entretient ces logiques de silos et limite grandement la circulation des données (y compris en interne) ;

– une absence matérielle de plateformes adaptées aux contraintes des secteurs, qui rassemblent les données d'intérêt pour l'IA, les moyens de calcul pour les exploiter et les piles logicielles nécessaires au développement d'applications expérimentales et opérationnelles ;

– les cadres réglementaires et juridiques qui peuvent sembler inadaptés aux besoins liés au développement de l'IA. »

L'état d'esprit est impossible à instaurer *ex nihilo* car il se heurte à de nombreux points de blocage recensés par la mission Villani pour les administrations, repris dans l'encadré 4.3. Il faut au préalable déstabiliser la structure en place. Ceci peut avoir lieu au cours de datathons, comme RSTAnalytics, ou des projets "bac à sable", comme à l'IFPEN. Le choix du mode d'animation de ces phases d'incitation est crucial car si la confiance vis-à-vis de l'équipe organisatrice fait défaut il est impossible de prolonger la démarche. Cependant, entre la confiance totale envers une démarche interne et le monde inconnu des initiatives extérieures, il ne suffit pas de toujours rester dans une zone de confort. L'organisation de sessions requiert un degré de professionnalisme qui ne s'acquiert que par leur répétition. L'improvisation serait le signe d'une obédience au dernier culte de transition permanente sans en rencontrer

les porteurs.

L'état d'esprit de l'innovation n'est pas si étranger aux milieux de recherche. Le degré d'exposition à la compétition internationale, la tenue ou la participation à des congrès, dénotent de l'ouverture à d'autres méthodes et de la volonté de rester à l'état de l'art. L'IGN organisera par exemple le congrès de 2020 de l'*International Society for Photogrammetry and Remote Sensing* (ISPRS). L'INERIS participe à une intercomparaison au niveau européen des prédictions de qualité de l'air. Ce jugement n'est à tempérer que pour des communautés scientifiques très repliées sur elles-mêmes pour qui la dimension culturelle, et parfois militante, dépasse la logique de problèmes communs. Pour ces dernières, le basculement provient quand suffisamment de membres se convertissent à de nouvelles techniques, au rythme de l'impact des publications. Les directions scientifiques de l'IGN et de Météo-France sont ainsi à l'œuvre pour décloisonner recherche et enseignement dans leurs écoles nationales, de météorologie (ENM) et de sciences géographiques (ENSG). Pour les communautés fermées, l'apprentissage automatique est aussi un moyen de décloisonnement et de remise en question. C'est une nouvelle donne pour les métiers liés aux données.

La phase de collecte des données requiert l'industrialisation du procédé et le passage à l'échelle. Il vaut mieux à ce titre acquérir de nouvelles données que d'espérer d'abord tirer profit de celles passées. Si les données n'ont pas été collectées de façon exhaustive, ou pensées pour une exploitation, le projet tient souvent de l'affichage. L'environnement logiciel pour sa part est affaire de direction des systèmes d'information qui doit s'adapter à une logique de cycles courts pour les programmes utilisés, de puissance de calcul spécifique, de stockage approprié, de requêtes nombreuses,...

La culture de laboratoire de sciences des données présuppose une sollicitation fréquente et une pratique régulière qui ne s'obtient que par la constitution d'équipes dédiées. Pour des raisons de confiance au sein de l'organisation, il est difficile de la créer hors sol, ce qui est d'ailleurs impossible au vu des trajectoires budgétaires et salariales des institutions du RST. Aussi longtemps que dure la montée des compétences en interne, il est envisageable d'externaliser cet aspect. Suivant la maturité du projet et les attentes associées, il existe différentes formes d'externalisation : stages, EIG, thèse CIFRE, consultants, intrapreneurs,... Le point fondamental est que s'agissant de méthodes évoluant rapidement, il ne faut pas en attendre une solution logicielle

figée. Sans accès au code informatique ou dans l'incapacité de l'analyser et de le modifier, la solution se trouvera rapidement obsolète. Un accompagnement extérieur suffisamment proche doit à terme permettre aux solutions de pénétrer dans la culture de certaines parties de l'institution, qui serviront de socle pour une équipe interne.

4.4 Trois scénarios d'évolution des institutions du RST

En somme, trois scénarios émergent de cette discussion sur les enjeux de l'apprentissage automatique pour le RST. Ces scénarios dépassent le strict cadre des aspects techniques car ils reflètent les orientations stratégiques prises individuellement par les institutions du RST. Par ailleurs les scénarios ne sont mutuellement exclusifs ni inter-établissements, au vu de l'hétérogénéité du RST, ni intra-établissement par la diversité des activités menées. Il revient aux fonctions de coordination d'arbitrer localement entre les stratégies des directions et aux actions transversales de pousser aux scénarios les plus positifs. Ces scénarios sont construits en fonction du destinataire final des stratégies.

Scénario 1 : l'institution tournée vers l'État de demain lui sert d'avant-garde dans l'automatisation des services techniques de la puissance publique.

Scénario 2 : l'institution tournée vers d'autres acteurs, citoyens ou entreprises, leur étend ses services actuels ou en propose de nouveaux.

Scénario 3 : l'institution tournée vers l'État d'hier préserve ses compétences et subit une lente dépréciation de ses actifs.

Dans le premier scénario, l'accent est mis sur la montée en compétences et le gain en efficience pour à terme proposer les deux aux services étatiques en interaction avec l'institution du RST. C'est le chemin le plus souhaitable quand l'organisme est fortement associé aux missions de l'Etat.

Dans le second scénario, l'autonomie est plus forte ou l'injonction de la tutelle à dégager des ressources propres plus marquée. L'institution se tourne vers d'autres acteurs auxquels il faut adapter les productions actuelles à coût réduit ou proposer

de nouveaux services.

Dans le troisième scénario, il n'est pas possible d'ajouter de nouvelles activités. Les efforts se focalisent sur la préservation des missions existantes soumises à la contrainte de faire autant avec moins. Certaines activités sont abandonnées au profit de concurrents plus performants, la dépréciation est lente, au rythme des réductions de postes.

Enfin aucun scénario ne stipule qu'il sera suivi à périmètre constant par l'institution. Il est possible que sa structure organisationnelle évolue, par scission, fusions, spin-offs,... Dans ces différents cas, l'entité créée est susceptible de s'éloigner encore plus du giron de l'État, ce qui se traduit par un délitement du lien qui les unissait et une érosion des capacités techniques pouvant être mobilisées par l'État.

Recommandations

Accompagner l'essor et la transition culturelle

Dans le cadre d'une formation spécifique aux institutions du RST, il faudrait :

- structurer le discours sur l'apprentissage pour mieux **définir** les problèmes afin d'établir le dialogue avec la communauté de science des données,
- être en mesure d'identifier les points bloquants pour **sélectionner** plus rapidement des projets à expérimenter,
- veiller à ne pas présenter la technique comme une boîte à outils sans cohérence méthodologique pour **résoudre** des problèmes,
- pousser à **participer** à des compétitions de données en externe pour s'approprier la démarche.

Définir le rôle des institutions dans l'écosystème

Une fois l'objet démystifié et la formation générale acquise, il s'agirait de :

- déterminer les ambitions et les limites à se fixer, pour donner la priorité à l'étude approfondie des méthodes correspondantes, en fonction du profil retenu (producteur de données ou de solutions, utilisateur, certificateur,...),
- proposer soi-même des compétitions pour bénéficier d'un retour d'expérience et d'une comparaison par rapport à un score obtenu en interne,
- définir le meilleur choix de porteur de projet en fonction du rôle et de la maturité de l'institution.

Passer à l'échelle par la division du travail

Au sein de la puissance publique ainsi qu'à l'extérieur, le RST peut :

- transmettre à ses partenaires la démarche pour une montée globale en compétence et développer l'écosystème institutionnel,
- construire la complémentarité avec d'autres réseaux (ex : supervision générale des données) pour conseiller sur les réutilisations sectorielles et les variables d'intérêt.

Conclusion

Au terme de ce rapport qui a cherché à embrasser sous divers angles le sujet de l'apprentissage automatique pour les institutions scientifiques et techniques du RST, que retenir ? Tout d'abord, loin des fantasmes sur une supposée intelligence, l'outil est une extension de méthodes statistiques ou algorithmiques. A ce titre, cette technique partage les mêmes problématiques que bien des solutions logicielles ou approches statistiques. L'apprentissage automatique pose la question du bien-fondé des modèles et de leurs modalités d'utilisation. L'outil requiert une intégration dans une chaîne de valorisation existante pour des institutions dont les missions sont pérennes. Par conséquent déployer de telles méthodes au-delà d'une simple expérimentation a une implication pour de nombreux corps de métiers. La conduite du changement doit anticiper ces effets et savoir sur quels groupes s'appuyer.

Pour le RST, le défi revient à se construire un discours spécifique car l'interaction privilégiée avec l'Etat ne permet pas de calquer un modèle issu du monde numérique. Si, en s'appropriant l'apprentissage automatique, le RST pourrait gagner en compétence, en efficacité, en visibilité ou en innovation, chaque institution doit déterminer la finalité d'un tel investissement. S'agit-il de mieux répondre aux attentes actuelles de l'Etat ou d'aborder de nouvelles questions et de nouveaux publics ? Le décloisonnement scientifique en cours permet de diversifier les compétences présentes et de transformer des cultures d'institutions. En définitive, l'outil par sa seule existence amène à questionner les axes stratégiques des institutions du RST et appelle à la réflexion sur leur périmètre d'action.

Références

- [AdT, 2018] Commission technologies de l'information et de la communication, *Renouveau de l'Intelligence artificielle et de l'apprentissage automatique*, Académie des technologies, Rapport, Mars 2018
- [AdT, 2015] Commission technologies de l'information et de la communication, *Big data : un changement de paradigme peut en cacher un autre*, Académie des technologies, Rapport, 2015
- [Besse, 2015] Philippe Besse et Béatrice Laurent, *De Statisticien à Data Scientist : Développements pédagogiques à l'INSA de Toulouse*, Statistique et Enseignement (ISSN 2108-6745), Société Française de Statistique, 2016, 7 (1), pp.75-93.
- [Donoho, 2017] David Donoho, *50 Years of Data Science*, Journal of Computational and Graphical Statistics, 26 :4, 745-766
- [François et Frezal, 2016] Pierre François et Sylvestre Frezal, *Pourquoi utilisons-nous des modèles « faux » ?*, PARI, Document de travail, Septembre 2016
- [Desrosières, 1993] Alain Desrosières, *La politique des grands nombres - Histoire de la raison statistique*, La Découverte, 2010 (première édition : 1993)
- [Enjeux Numériques, 2018a] *L'intelligence artificielle : un enjeu d'économie et de civilisation ?*, Annales des Mines (ed.), Enjeux numériques N° 1, Mars 2018
- [Enjeux Numériques, 2018b] *Big data : économie et régulation*, Annales des Mines (ed.), Enjeux numériques N° 2, Juin 2018
- [France Stratégie, 2017] France Stratégie, *Intelligence artificielle et travail*, Rapport aux ministres du travail et du numérique, Mars 2018
- [Inspection générale, 2014] Inspection générale interministérielle, *Évaluation de la politique de sécurité routière*, Rapport de diagnostic, Juillet 2014
- [Lascoumes et al, 2014] Pierre Lascoumes, Laure Bonnaud, Jean-Pierre Le Bourhis et Emmanuel Martinais, *Le développement durable. Une nouvelle affaire d'État*, Paris, Presses universitaires de France, coll. « L'écologie en questions », 2014
- [Mawhin, 2017] Jean Mawhin, *Les modèles mathématiques sont-ils des modèles à suivre ?*, Académie royale de Belgique (ed.), L'Académie en poche, 2017
- [Mazzucato, 2013] Marianna Mazzucato, *The Entrepreneurial State*, Anthem Press (ed.), 2013

- [O’Neil, 2016] Cathy O’Neil, *Weapons of Math Destruction*, Crown (ed.), 2016
- [Trojette, 2013] Mohammed Adnène Trojette, *Ouverture des données publiques - Les exceptions au principe de gratuité sont-elles toutes légitimes ?*, Rapport au Premier ministre, Juillet 2013
- [Vayatis, 2016] Nicolas Vayatis, *Big Data et politiques publiques dans les transports : La décision par algorithme* (Chapitre 2), Economica (ed.), A. de Palma et S. Dantan (coord.), 2016
- [Villani, 2018] Cédric Villani et al., *Donner un sens à l’intelligence artificielle : Pour une stratégie nationale et européenne*, Rapport de mission parlementaire, Mars 2018
- [von Neumann, 1955] John von Neumann, *Method in the Physical Sciences*, L. Leary (ed.), The Unity of Knowledge, 1955

Table des matières

| | | |
|----------|--|-----------|
| 1 | Contextualisation institutionnelle du RST | 3 |
| 1.1 | Le RST : un concept du ministère de l’environnement | 3 |
| 1.2 | Le RST : un ensemble hétérogène d’organismes soumis à diverses tutelles | 6 |
| 1.3 | Le RST : un réseau et des institutions dans un contexte en évolution rapide | 7 |
| 2 | Conceptualisation de l’apprentissage automatique | 9 |
| 2.1 | Modèles et aide à la décision | 9 |
| 2.1.1 | Qu’est-ce qu’un modèle ? | 9 |
| 2.1.2 | Quelle aide à la décision ? | 13 |
| 2.1.3 | Les décideurs aux prises avec les modèles | 15 |
| 2.2 | Survol technique de l’apprentissage automatique | 17 |
| 2.3 | Décloisonnement scientifique et questions statistiques | 21 |
| 3 | Contextualisation technique de l’apprentissage automatique | 25 |
| 3.1 | Chaîne de production et de valorisation des données | 25 |
| 3.2 | Cultures métiers et interaction avec l’apprentissage | 28 |
| 3.3 | Se positionner vis-à-vis de l’apprentissage | 30 |
| 4 | Enjeux pour le RST en matière d’apprentissage automatique | 33 |
| 4.1 | Quelles incitations à agir ? | 33 |
| 4.2 | Quelles finalités pour le RST à s’approprier cette technique ? | 34 |
| 4.2.1 | Disposer d’une compétence à faire valoir | 36 |
| 4.2.2 | Gagner en efficacité | 38 |
| 4.2.3 | Apprendre du public, penser au destinataire | 40 |
| 4.2.4 | Produire de nouveaux services | 43 |
| 4.3 | Quels prérequis et quelles approches pour développer la technique au sein du RST ? | 44 |
| 4.4 | Trois scénarios d’évolution des institutions du RST | 47 |