



Un interview de Jérôme PESENTI chef de division en recherche en intelligence artificielle chez Facebook

Propos recueillis par D. GAYET

D'après Wikipedia tu as remplacé Yann Lecun chez Facebook au poste de, je cite « chef de division en recherche en intelligence artificielle ». Peux-tu nous dire en quoi ça consiste ?

Je n'ai pas remplacé Yann. Son rôle c'est d'être directeur scientifique du laboratoire d'intelligence artificielle de l'entreprise. Moi je suis manager de l'équipe. Je m'occupe aussi des problèmes d'implémentations, des problèmes de recherche plus appliquée.

Ton équipe, c'est combien de personnes ?

C'est 500 personnes dont 150 chercheurs à plein temps.

Facebook est souvent décrié pour son manque d'intérêt pour les questions de données personnelles, comme l'a démontré par exemple le scandale de Cambridge Analytica. Est-ce que c'est un problème pour toi, moralement et professionnellement parlant ?

C'est un problème, bien sûr, moi je veux travailler pour une entreprise qui ait un impact positif sur le monde. Mais entre ce qui est présenté dans les médias et la réalité, il y a quand même une grande différence, et Facebook a pas mal évolué sur ce thème des données personnelles. L'une des raisons pour lesquelles je suis chez eux, c'est que je peux faire ce que je veux, au sens où chacun peut avoir sa propre opinion, l'exprimer de façon complètement ouverte et faire le mieux qu'on pense.

Peux-tu nous décrire ton parcours universitaire et professionnel ?

Je suis rentré en 1991 à l'École normale supérieure de Paris en mathématiques, j'ai passé l'agrégation

et fait une thèse en géométrie arithmétique à Orsay. En même temps que ma thèse j'ai suivi un cursus universitaire de philosophie puis de sciences cognitives. Celles-ci m'ont vraiment passionné. Après mon DEA de sciences cognitives, pour mon service militaire je suis parti faire ma coopération. J'ai alors soutenu ma thèse mais j'ai décidé d'abandonner les mathématiques à ce moment-là et de me lancer dans l'aventure de l'intelligence artificielle. J'ai fait un peu de recherche pendant un an et demi à Carnegie Mellon, puis j'y ai lancé une entreprise, Vivisimo, dans le domaine de l'IA, qui a employé jusqu'à 150 personnes. J'ai ensuite vendu la société à IBM, chez qui j'ai été embauché pour lancer la commercialisation de Watson¹. Puis je suis devenu co-CEO d'une start-up à Londres dans le domaine de l'IA associée aux découvertes de médicaments, et maintenant je suis chez Facebook.

En quoi les mathématiques t'ont-elles été utiles dans ton parcours professionnel ?

Je ne suis plus mathématicien ni même chercheur. De mon point de vue, les mathématiques permettent deux choses : d'une part, une facilité à comprendre et manipuler le formalisme, de plus en plus présent en IA. D'autre part, elles fournissent une sorte de rigueur de raisonnement qui aujourd'hui est assez important en IA, c'est assez difficile de comprendre la performance des systèmes, leur méthodologie, qui repose sur les sciences des données et les statistiques.

Comment vois-tu les liens entre informatique et mathématiques pour ce qui te concerne, pour la formation des chercheurs en IA, et pour l'IA en général ?

1. Watson est le programme phare d'IA d'IBM, célèbre pour avoir battu en 2011 les champions du jeu *Jeopardy!*

Comme je l'ai dit avant, les mathématiques fournissent une capacité à comprendre le formalisme. Mais plus directement, les probabilités et les statistiques sont vraiment au centre de l'IA, tout comme l'algèbre linéaire. Ce ne sont pas des mathématiques très complexes, mais ces systèmes vivent en très grandes dimensions, et sont en fait très difficiles à comprendre du point de vue mathématique.

Que recherche-t-on à Facebook ?

La raison pour laquelle l'IA est importante pour Facebook est la suivante : nous avons une plateforme avec une production énorme de contenus, qu'il faut modérer. Il faut que les gens ne se harcèlent pas, qu'ils ne publient pas d'images violentes, bref qu'ils respectent les règles de Facebook. Même si l'on a 20000 modérateurs, ça ne suffit pas et il faut des outils automatiques pour comprendre les contenus, les vidéos, les textes. Ça nécessite une IA assez raffinée. Par ailleurs nous essayons d'accroître la qualité de ce qui est proposé aux utilisateurs, donc il faut comprendre la qualité d'un contenu, son impact sur les gens, s'il les rend plus informés, plus heureux, etc. En résumé, l'IA nous aide à limiter le contenu négatif et à promouvoir les contenus positifs.

Et pour ce qui concerne la recherche proprement dite ?

Ça signifie qu'il faut comprendre un minimum le langage, ainsi que la vision. Et comprendre le langage, c'est comprendre ce que c'est que l'intelligence, en fait. Aujourd'hui on utilise des systèmes qui sont très statistiques, mais qui permettent de traduire de manière assez efficace. Ils comprennent donc un peu l'intention du langage. Mais ça ne suffit pas, par exemple pour différencier une blague d'un harcèlement. Côté images, il faut pouvoir détecter par exemple la violence dans une vidéo. Le faire de façon automatique ça n'est pas du tout évident.

Qu'est-ce qui distingue la recherche en IA chez Google, IBM et Facebook ?

Je dirais que Google et Facebook ont des recherches très proches, mais les applications sont différentes. Par exemple Google s'occupe de conduite automatique. IBM est plutôt tournée vers les applications pour les entreprises. Je dirais que Google essaie d'organiser l'information du monde, Facebook de connecter les humains.

Est-ce qu'il y a des liens entre Facebook et les instituts de recherche ?

Dans notre équipe de recherche, 70% des articles publiés par nos chercheurs ont un coauteur qui n'est pas de Facebook, ce qui se comprend puisque tous nos chercheurs ont fait une thèse à l'université. Par ailleurs toute notre recherche est publiée, et souvent de façon libre, donc il y a très peu de contraintes pour collaborer avec des chercheurs extérieurs.

Est-ce qu'il ya des liens avec d'autres domaines de recherche, comme la sociologie, la psychologie ou la linguistique ?

De plus en plus. Bien sûr, il existe des liens avec la linguistique, souvent conflictuels d'ailleurs. J'aime bien raconter cette anecdote : Frederick Jelinek, un chercheur d'IBM qui a initié tout le champ de la linguistique statistique pour résoudre les problèmes de reconnaissance vocale et traduction automatique, gérait une équipe de linguistes et d'ingénieurs. Un jour un linguiste part, Jelinek le remplace par un ingénieur parce qu'il ne trouve pas de linguiste, et se rend compte que le système s'est amélioré. Il convainc alors un autre linguiste de partir, et le système s'améliore encore, et ainsi de suite. Cela dit les linguistes disent que ces systèmes ne comprennent pas le langage, et finalement depuis deux ou trois ans c'est ce que je pense aussi. On a des systèmes vraiment très performants grâce à une force brute, statistique. Mais ces systèmes n'ont pas de sens commun. Ils donnent l'illusion de comprendre ce qui se passe, mais en fait ce sont des formes qu'ils reconnaissent, et pas du sens.

Concernant la sociologie, aujourd'hui les systèmes d'IA sont de plus en plus utilisés pour recommander des rencontres amoureuses, du contenu, ou aider au recrutement. Quand on fait ça il y a un problème de biais, d'équité (*fairness*). Est-ce qu'on peut accepter que ces systèmes reproduisent les biais existant dans la réalité ? Souvent on répond par la négative. Mais les gens font des choix qui leur ressemblent, alors jusqu'à quel point l'IA doit-elle s'abstenir d'agir ainsi ? Ce sont des problèmes très complexes d'éthique ou de sociologie, plus que des problèmes d'IA.

Quel est le profil des chercheurs embauchés chez Facebook ?

On embauche beaucoup de jeunes docteurs en science de l'apprentissage avec des applications autour de la vision, du langage, ou de l'audio. On recrute aussi des chercheurs qui peuvent avoir jusqu'à 15 ans d'expérience dans le domaine. Et puis on embauche aussi des physiciens, parce qu'on se

rend compte que pour comprendre un réseau de neurones, il faut presque en faire un modèle physique, savoir comment l'information se propage, etc. Et l'on a commencé aussi à recruter également des gens qui ont une formation en sciences cognitives, et en mathématiques.

[Est-ce que vous êtes sensibles à Facebook de l'image qu'a le grand public de l'IA ?](#)

Un de nos buts, c'est de communiquer sur l'IA, et de montrer qu'elle a un impact positif sur le monde. Les gens ont pensé que l'IA allait prendre nos emplois, par exemple. Mais ces systèmes ne peuvent effectuer que des tâches malgré tout très restreintes, donc cette menace n'est pas réelle. Les vrais problèmes de l'IA, c'est ce que je disais avant, ce sont les biais qu'ils peuvent renforcer. Si l'IA recommande pour les postes d'informaticiens uniquement des hommes, on perpétue un biais. Nous aimerions que ces problèmes éthiques soient discutés ouvertement. Est-ce que l'IA doit être utilisée pour l'armement, par exemple ? C'est une technologie très puissante, et comme à chaque fois elle va produire des changements importants dans la société. C'est important d'en discuter avec la société de façon ouverte.

[Peux-tu nous faire un court historique des réseaux de neurones \(RNs\), qui sont actuellement les piliers de l'IA ?](#)

Le modèle a été découvert dans les années 50. Ensuite des chercheurs ont décrit les limites théoriques des RNs. Les gens en avaient conclu que les RNs étaient très limités. Dans les années 80-90, des gens comme Yann Lecun ont commencé à montrer que les RN pouvaient résoudre des problèmes réels, mais bizarrement l'intérêt et les investissements ont diminué, c'est ce qu'on appelle l'hiver de l'IA. Dans les conférences on invitait un ou deux connexionnistes pour être gentil. À partir de 2009 les RNs se sont mis à battre, selon les tests (*benchmark*) classiques de l'IA, tous les autres types de systèmes. Depuis les progrès sont constants, ce qui n'était pas le cas avant.

[J'ai souvent entendu souvent dire qu'on ne comprenait pas pourquoi les RNs fonctionnaient aussi bien, peux-tu nous expliquer cette double affirmation ?](#)

Oui c'est un peu vrai. Ça marche très bien : dans la plupart des tâches classiques, les systèmes reposant sur les IAs sont les meilleurs. De façon un peu plus technique, les RNs ont énormément de paramètres, et semblent généraliser la tâche sans trop

surinterpréter (*overfit*), c'est-à-dire sans coller trop aux données déjà fournies. Plus les réseaux de neurones sont gros, plus ils sont efficaces. C'est un peu magique.

[Chez Facebook, y a-t-il des chercheurs qui essaient de comprendre pourquoi ce phénomène ?](#)

Oui, on fait de la recherche fondamentale en IA.

[Peux-tu nous parler des problèmes fondamentaux de l'IA non résolus, outre ceux dont tu as déjà un peu parlé ?](#)

Les systèmes actuels apprennent de façon scolaire, supervisée. Mais les enfants, pour une même tâche simple, n'apprennent pas de façon supervisée. Un enfant en bas âge apprend en interagissant et en observant le monde. On ne lui donne pas un milliard d'exemples en précisant lesquels sont bons et lesquels ne le sont pas. Le nombre de supervisions est donc très limité par rapport à la capacité qu'il acquiert. Un enfant peut même apprendre quelque chose en une seule fois (*one-shot learning*). Comme on l'a dit les systèmes n'ont pas non plus de sens commun, ils ne comprennent pas par exemple la causalité, contrairement à un humain qui détecte rapidement une cause à partir d'un effet. Cela étant dit, les enfants bénéficient d'un cerveau qui est le résultat de millions d'années d'apprentissage. Il faudrait de façon générale que les systèmes, qui actuellement fonctionnent de manière très statistique, apprennent plus par concept, et possèdent un modèle du monde qui fasse sens.

[Comment penses-tu que ça sera résolu ?](#)

Ça viendra à mon avis d'une combinaison. D'abord, même si l'on atteint un peu les limites de ce phénomène, on va continuer à accroître la taille des RNs, puisqu'ils fonctionnent mieux ainsi. Ensuite, et là c'est un débat entre les communautés, il faudrait associer les méthodes statistiques actuelles avec une approche plus symbolique. Les approches symboliques des années 80 n'ont pas porté beaucoup de fruit, elles étaient trop naïves. La question intéressante, c'est comment ancrer ce symbolisme dans une représentation statistique du monde. Un symbole ou un concept représente une combinaison d'expériences. Aujourd'hui on utilise des vecteurs en très grande dimension pour représenter la complexité des concepts. Ce qui nous manque, c'est une compréhension théorique, mathématique, de la dynamique des systèmes apprenants, de la façon dont ceux-ci généralisent ou mémorisent l'information. Il y a une focalisation de la communauté sur les

tests, avec des articles qui améliorent d'un ou deux pourcents une performance donnée, alors qu'il est plus intéressant de comprendre pourquoi le système est performant. Par exemple il y a des expériences qui retirent certaines informations au RN, et l'on se rend compte que le réseau reste beaucoup plus performant qu'on ne l'aurait pensé.

[Il semble que les RNs fonctionnent par corrélations. Dans un article récent du Monde le journaliste donnait l'exemple d'un loup que les RNs distinguaient d'un chien uniquement par le fond de l'image, toujours naturel dans le premier cas. On imagine bien les problèmes quand il s'agit de détecter une maladie ou de décider si une voiture s'engage ou pas sur une route. Que penses-tu de ce problème?](#)

C'est un gros problème. Les systèmes apprennent de données historiques et de manière statistique. On ne peut pas faire face correctement à une situation nouvelle à moins d'avoir une compréhension conceptuelle du problème. Concernant le loup, c'est vrai que le système se base beaucoup sur le fond, même de façon locale. Mais on peut faire des RNs qui font des choses plus fines, par exemple qui comprennent les silhouettes. Le problème c'est que l'on crée des tests où le RN va utiliser le décor plutôt qu'une compréhension plus fine. Les tests ne favorisent pas les systèmes qui ont une approche conceptuelle. Le problème de l'intelligence, c'est de généraliser des situations qu'on n'a rencontrées qu'une seule fois, de transférer ce qu'on connaît d'un domaine dans un domaine. On essaie de créer des tests plus interactifs et moins statiques. À Facebook, nous construisons des environnements qui interagissent avec les gens, plutôt que de réaliser une tâche fixe avec un test qui est toujours le même. Par exemple, on est en train de réaliser un guide dans les rues de NYC avec une IA.

[L'IA pourra-t-elle un jour aider les mathématiciens dans leur recherche? L'IA pourrait-elle parcourir l'espace abstrait des mathématiques comme AlphaGo explore l'espace des parties de go?](#)

On en est vraiment très, très loin. L'espace des parties de Go possède une certaine régularité, et on peut évaluer la valeur d'une position dans une partie. Dans une preuve mathématique, savoir si on est sur le chemin d'une bonne preuve ou pas, ça semble très difficile. Mais mon équipe vient tout de même de dé-

marrer un projet ambitieux sur ce sujet en utilisant des techniques similaires.

[Va-t-on être tous transformés en trombones un jour?](#)²

D'un côté c'est une bonne question, de l'autre non. C'est une mauvaise question parce que tous les systèmes d'IA aujourd'hui se comportent dans des limites bien définies, et il n'y a aucune raison de penser que ça va changer. Le problème intéressant et même crucial, c'est par contre celui de la fonction d'objectif (*objectif function*). Trouver un bon objectif à un système d'IA est assez difficile, parce que ce sont souvent des objectifs à long terme. Ce qui est difficile c'est de savoir si l'on se trouve bien sur le chemin de la tâche finale. On trouve toujours des *proxys* pour cette tâche, c'est-à-dire des paramètres censés indiquer si l'on est proche ou non de l'objectif. Mais souvent on crée des systèmes qui essaient d'optimiser les *proxys*, mais pas la tâche finale. Concernant Facebook, si l'objectif est que les gens soient heureux mais que le proxy choisi est le temps passé sur Facebook, il se peut que le système pousse les gens à perdre leur temps. Un autre exemple est le fameux *bot* de Microsoft, qui avait été créé pour maximiser les interactions sur Twitter. Ce que le *bot* a trouvé très rapidement, c'est que s'il insultait les gens ou s'il faisait des commentaires choquants, les gens réagissaient beaucoup, ce qui encourageait le bot dans cette stratégie.

[Quelles sont selon toi les bornes morales pour les décisions autonomes des IA, l'exemple extrême étant les robots tueurs?](#)

Dans l'armement, c'est quand même très délicat que des armes prennent des décisions offensives, en particulier tuer. La plupart des chercheurs sont inconfortables avec cette idée. Cela dit on pourrait par exemple créer des armements intelligents qui les rendraient moins létaux. Et si l'on utilise une IA qui permette de ne jamais tuer des civils, est-ce que c'est une mauvaise chose? Dans tous les cas l'IA a un système de valeurs quand il prend des décisions, et il faut que ça se fasse dans une certaine transparence, notamment parce qu'ils ont tendance à reproduire les biais des données.

[Est-ce que tu regrettes le monde des mathématiques?](#)

2. En 2003, le philosophe Nick Bostrom a illustré les risques posés par l'IA en décrivant un système à qui on aurait demandé de produire autant que possible des trombones. Le système pourrait alors se débarrasser des humains qui pourraient chercher à l'arrêter dans sa quête de maximiser la production de trombones.

Très bonne question! Oui parfois je ressens une certaine nostalgie. Est-ce que je regrette mes choix, non, parce qu'aujourd'hui je peux avoir énormément d'impact positif sur le monde. Mais ce qu'il y a d'amusant c'est que pas mal de gens que j'ai connus en mathématiques ont obtenu des prix, y compris la médaille Fields. Bien sûr je n'aurais pas eu la médaille Fields, je n'étais pas aussi brillant qu'eux, mais je me

retrouve à lire des articles sur leur domaine, d'essayer de comprendre ce qu'ils font. C'est vrai que cette recherche de connaissances pures me manque. Découvrir quelque chose de nouveau, de vraiment nouveau, que personne d'autre n'avait trouvé avant, c'est quand même vraiment exaltant, et je ne peux plus le faire maintenant.



Depuis 20 ans, **Jérôme Pesenti** travaille dans le domaine de l'intelligence artificielle, le traitement du langage naturel, les moteurs de recherche, les données massives et l'apprentissage automatique. Il dirige actuellement les équipes de recherche et développement en intelligence artificielle de Facebook.

ETH zürich

Professor of Mathematics

→ The Department of Mathematics (www.math.ethz.ch) at ETH Zurich invites applications for the above-mentioned position.

→ Successful candidates should demonstrate an outstanding research record and a proven ability to direct research work of high quality. Willingness to participate in collaborative work both within and outside the school is expected. The new professor will be responsible, together with other members of the department, for teaching undergraduate (German or English) and graduate level courses (English) for students of mathematics, natural sciences, and engineering.

→ **Please apply online: www.facultyaffairs.ethz.ch**

→ Applications should include a curriculum vitae, a list of publications, a statement of future research and teaching interests, and a description of the three most important achievements. The letter of application should be addressed **to the President of ETH Zurich, Prof. Dr. Joël Mesot. The closing date for applications is 15 November 2019.** ETH Zurich is an equal opportunity and family friendly employer, strives to increase the number of women professors, and is responsive to the needs of dual career couples.