Groupe de travail interdépartemental IA de la Confédération

Contribution préliminaire de l'EPFL du 16 avril 2019

au thème Cybersécurité et Politique de sécurité (GT 12 – IDAG KI)

Lausanne, le 16 avril 2019

Contribution préparée à la demande de M. Maurice Eglin (SG-DDPS), Président du GT 12 – IDAG KI, par

- ***Professeur Pierre Vandergheinst**, Vice-Président de l'EPFL pour l'éducation, directeur du laboratoire LTS II (modern challenges in data processing)*
- ***Stéphane Decoutère**, Délégué du Président de l'EPFL aux affaires gouvernementales*

Rassemblée et rédigée avec le concours des contributions des professeurs et chercheurs affiliés

- ***aux Centres de l'EPFL liés à la digitalisation** – 1/ Center for Learning Sciences **LEARN,** 2/Center for Digital Trust **C4DT,** 3/ Center for Intelligent Systems **CIS***
- ***et au Centre conjoint EPFL-ETHZ** Swiss Data Science Center **SDSC***

1ère Partie : Concrétisation du domaine thématique

Développement de l'IA

L'intelligence artificielle (IA) a déjà démontré sa capacité à disrupter les industries établies (Industrie 4.0). Elle nous obligera bientôt à revisiter les hypothèses fondamentales de notre société, y compris les notions de base de souveraineté, de confiance, d'égalité, de responsabilité, d'éthique, de normes et de droit international.

Le cycle actuel est principalement marqué par les progrès réalisés par l'apprentissage automatique profond ("*Deep Learning*"). Cette technique d'apprentissage automatique (« Machine Learning ») est très souvent utilisée derrière l'appellation générique IA. Ce sont ces applications qui permettent d'atteindre des performances humaines ou surhumaines dans la reconnaissance d'images et le traitement du langage naturel, ce dont attestent par exemple les logiciels de traduction.

Dans les cas les plus récents, aucune expertise humaine n'est utilisée dans le processus d'apprentissage, mis à part un modèle paramétrique liant entrée et sortie (apprentissage dit "de bout en bout"). Les algorithmes informatiques appelés "réseaux neuronaux" utilisés par ces mécanismes d'apprentissage en profondeur sont ajustés dans un processus automatisé d'entraînement itératif qui utilise de grandes quantités de données annotées.

Autrement dit : la machine apprend par elle-même à reproduire des tâches complexes indépendamment de l'homme qui l'a conçue et qui l'opère. Ce qui rend la compréhension et le contrôle des algorithmes d'autant plus important. Dans ce domaine, les récentes méthodes de « reinforcement learning » utilisent des environnements de simulation, souvent très fidèles, plutôt que des données du monde réel et permettent par exemple à des agents intelligents d'apprendre à se mouvoir.

Utilisations de l'IA aujourd'hui

Les utilisations les plus évidentes de l'IA et du *Deep Learning* touchent actuellement 4 domaines :

1. Cybersécurité, cyberdéfense et confiance digitale

   Les implications de l'IA en matière de souveraineté et de sécurité publique sont particulièrement sensibles, par exemple lorsque les informations véhiculées par les réseaux sociaux mondiaux et les algorithmes qui les opèrent peuvent être facilement manipulés par des tiers. De même, l'économie de l'information, et en particulier les secteurs de l'industrie financière et de l'industrie manufacturière, est particulièrement vulnérable aux violations de données, aux intrusions dans les réseaux et au vol de la propriété intellectuelle des logiciels et du matériel. D'où l'urgence prise par la cybersécurité, la cyberdéfense et plus globalement la question de la confiance digitale dans les stratégies des Etats et des sociétés privées.

2. Formation continue aux savoir-faire digitaux nécessaires aux emplois 4.0 utilisant l'IA

   De manière plus générale, l'avenir de l'emploi et le rôle de la formation continue tout au long de la vie sont au cœur des interrogations autour de la digitalisation en général et de l'IA en particulier. Elles sont particulièrement d'actualité pour des institutions internationales établies à Genève comme l'Organisation internationale du travail (OIT) ou l'Organisation mondiale de la propriété intellectuelle (OMPI). Comme l'a écrit l'historien Yuval Noah Harari dans son dernier ouvrage (« 21 leçons pour le 21ème siècle ») : "*Le marché de l'emploi de 2050 pourrait ainsi se caractériser par une coopération hommes-IA plutôt que par une concurrence. Dans divers domaines, de la police à la banque, des équipes hommes-IA pourraient surpasser à la fois les hommes et les ordinateurs*". A contrario, une déficience dans cette collaboration homme-machine dans des contexte d'automatisation, qu'elle soit appuyée ou non par l'IA, peut générer des catastrophes comme en témoigne, de l'avis de plusieurs experts, l'accident récent des deux Boing 737 Max 8, en Indonésie et en Ethiopie. Cet état de fait nous force à repenser l'éducation pour permettre aux populations actuelles et futures d'acquérir et de mettre à jour leur savoir-faire par rapport à cette nouvelle réalité de l'intelligence artificielle dont dépendent les emplois 4.0. Il s'agit en particulier de doter toutes les populations de notre pays d'une compréhension des mécanismes de base de l'IA, de son potentiel, de ses limites, et de ses failles spécifiques de qualité ou de sécurité.

3. Applied Machine Learning, algorithmes et santé

   L'un des domaines dans lequel la coopération hommes-IA se développe rapidement est celui de la médecine. Le secteur de la santé est particulièrement intéressant pour les applications d'IA, étant donné la numérisation en cours de tous les types d'informations de santé à travers notamment les objets connectés qui nous accompagnent au quotidien. L'IA pour la santé offre aussi potentiellement de nouveaux moyens pour remédier à la pénurie de professionnels de la santé ou pour accélérer la couverture médicale dans des parties peu accessibles du monde. L'IA peut enfin potentiellement améliorer de manière significative les processus de décision en matière de diagnostics médicaux et de traitements basés sur des données numériques.

4. Systèmes intelligents

   Une autre tendance transformationnelle stimulant actuellement l'innovation et la recherche en robotique et en IA est la "fusion" des mondes virtuel et physique par le développement de systèmes autonomes ou intelligents, de technologies portables, de collègues robotiques ou intelligents (« digital twins »), d'assistants intelligents, d'appareils intelligents, de technologies de drone, de réalité augmentée, de maisons intelligentes et de nombreuses autres applications d'intelligence artificielle.

L'autonomie ou l'intelligence est un élément clé de bon nombre de ces technologies. Une fois disponibles, ces systèmes intelligents présenteront au moins quatre types de défis :

i. Ils auront des répercussions radicales dans de nombreux domaines, notamment la fabrication, le transport, le commerce, l'emploi, les soins de santé, le gouvernement, le droit, la sécurité, la protection de la vie privée et aussi l'éducation. Le corps professoral, les étudiants, les entreprises, les autorités politiques, les administrations qui leurs sont subordonnées et l'ensemble de la population devront prendre conscience de ces implications.

ii. Ces systèmes exigeront de nouveaux paradigmes de conception qui reconnaissent que les machines intelligentes et les humains doivent collaborer pour atteindre un but qui peut affecter un grand nombre d'individus, et que les humains et les systèmes peuvent partager le même espace physique. Les questions liées à la sécurité, à la fiabilité, au contrôle, à l'explicabilité, à la responsabilité et à la réactivité deviendront primordiales.

iii. Une nouvelle couche de systèmes interconnectés et intelligents verra le jour. De plus, un nouvel ensemble de questions scientifiques et techniques devra être abordé afin de permettre la conception de systèmes interconnectés complexes avec un "esprit" qui leur est propre. Nous en voyons déjà des éléments rudimentaires dans les approches utilisées pour l'étude des systèmes cyberphysiques (CPS). Cependant, il faudra beaucoup plus pour tenir compte de l'élément d'"intelligence" qui deviendra omniprésent.

iv. Enfin, le niveau accru d'interdépendance et d'autonomie pose un certain nombre de nouvelles questions éthiques et juridiques qui exigent une réflexion approfondie.

2ème Partie : Enjeux spécifiques de l'IA dans le domaine de la sécurité

Challenges généraux posés par l'autonomie des systèmes d'Intelligence artificielle

Contrairement aux systèmes automatisés, qui sont programmés pour exécuter des tâches répétitives dans des conditions limitées et des lois de contrôle bien définies, les systèmes autonomes ou intelligents peuvent apprendre de l'expérience et des humains, rechercher activement des informations et en générer de nouvelles, planifier et prendre des décisions complexes dans des environnements dynamiques, quitter des zones géographiquement limitées, se réparer eux-mêmes et se maintenir, et éventuellement travailler avec des humains et partager leur milieu.

- On s'attend à ce que les futurs systèmes de fabrication, de soins de santé, d'énergie et de transport ainsi que les installations scientifiques à grande échelle s'appuient fortement sur des entités autonomes capables d'opérer dans des environnements complexes, en présence d'humains, en présence de contraintes évolutives, et même dans de nouvelles frontières (comme l'exploration spatiale).

- On s'attend aussi à ce que les systèmes autonomes aient un impact considérable sur notre façon de vivre et de travailler, et leur présence aura des répercussions sociétales profondes sur l'emploi, l'éducation, la sécurité et la vie privée, en plus des questions éthiques et juridiques.

Challenges spécifiques de l'IA du point de vue sécuritaire

Du point de vue sécuritaire, plusieurs débats agitent en ce moment la communauté scientifique internationale spécialisée dans l'IA :

- Débat no 1 : Comment rendre un algorithme résistant et robuste ? (voir par ex. le Programme de recherche sur l'interprétabilité de l'IA de l'US National Science Foundation).
- Débat no 2 : Comment donner confiance à ceux qui donnent accès à leurs données pour entraîner l'IA ?
- Débat no 3 : Comment assurer une diffusion sure d'algorithmes effectifs dans des domaines d'applications spécifiques ?
- Débat no 4 : Quelles sont les conditions cadres à réunir pour tirer parti des opportunités et minimiser les risques inhérents à l'IA d'un point de vue économique et sociétal ?

Nous distinguons par conséquent dans les lignes qui suivent quatre types de problèmes à résoudre sur lesquels l'EPFL travaille activement.

- Les deux premiers problèmes concernent la science des données (mécanismes fondamentaux de l'IA tels que reinforcement learning via la simulation, IA en réseau, exemples adversariaux, etc.).
- Les deux suivants concernent les sciences digitales (développement de l'utilisation de l'IA dans tous les domaines d'applications possibles, santé, finance, e-voting, défense nationale, éducation, etc.).

1. Challenge sécuritaire no 1
Maîtriser l'entraînement des algorithmes

i. Les systèmes IA sont entraînés par les données qui lui sont fournies. Cet entraînement est fragile. Une petite modification d'une entrée peut influencer fortement la décision prise par l'IA. C'est ce que montrent les systèmes dit adversariaux (par ex : signalisation routière où une petite perturbation introduite par ex. sur le coin d'un panneau de signalisation est de nature à tromper l'algorithme). Il est ainsi possible de placer un overlay (invisible à l'œil nu) sur la photo d'une personne faisant que la machine identifie cette photo comme l'image d'un tigre, par exemple. Nous voyons la photo d'une personne, tandis que l'ordinateur « voit » un tigre. Ce principe s'apparente à une « illusion d'optique » pour les machines.

ii. Les « adversarial examples » sont d'abord dangereux puisqu'ils sont susceptibles de duper les machines autonomes. Pensez par exemple à une voiture autonome qui interpréterait, de cette manière, un signal routier de façon erronée. Les chercheurs qui se concentrent sur le machine learning évaluent comment sécuriser et renforcer leurs systèmes d'intelligence artificielle et comment identifier les « adversarial examples ».

iii. Sur le plan de la cyberdéfense ou de la cybersécurité, l'ennemi ou plus simplement les personnes voulant imposer la diffusion d'un certain type d'informations ou qui veulent influencer l'opinion et pour cela provoquer un certain type de réponses de la part des algorithmes, peuvent tester eux-aussi nos systèmes et tenter d'en influencer les réponses (aux Etats-Unis, les recherches ont montré comment l'envoi de certaines données médicales permettait de prévoir les diagnostics qui seraient faits par l'algorithme concerné).

iv. La contre mesure consiste à développer des systèmes d'entrainement robustes permettant d'analyser les systèmes IA et de protéger l'intégrité des systèmes. Cela demande des connaissances pointues en science des données et en mathématiques fondamentales. Ce d'autant que nous n'échapperons pas à devoir nous pencher sur la problématique dite de la maîtrise par l'homme de ces superintelligences artificielles capables d'apprentissage autonome, problématique illustrée dès 2014 par le Professeur Nick Bostrom, directeur du « Future Of Humanity Institute » de l'Université d'Oxford dans son livre « Superintelligence – Paths, Dangers, Strategies ».

2. Challenge sécuritaire no 2
Sécuriser les données nécessaires à entraîner ces algorithmes

i. La sécurité des données – privacy, intégrité – est le second enjeu sécuritaire posé par l'IA.

ii. Pour pouvoir entraîner l'IA, il faut pouvoir garantir aux clients mettant à disposition leurs données que celles qui sont utilisées ne peuvent pas être compromises, soit en sécurisant l'infrastructure d'apprentissage soit en effectuant l'apprentissage directement sur des données cryptées. Cela vaut aussi bien pour les données réelles que pour les données simulées.

iii. Pour y parvenir, le développement d'outils systématiques permettant de construire des technologies défendant l'intégrité des données contre des attaques par des outils de machine learning est nécessaire.

iv. Une approche développée à l'EPFL (Prof. Troncoso) s'appuie sur la théorie des graphes (codage de toutes les possibilités d'interprétation et de leur coût). De telles techniques demandent aussi une connaissance pointue en science des données et donc en mathématique fondamentale.

3. Challenge sécuritaire no 3
Certifier les algorithmes de manière à pouvoir en assurer une diffusion appropriée

i. L'IA peut potentiellement améliorer de manière significative les processus de décision dans toute sortes de matière, par exemple le diagnostic médical et de traitement, basés sur des données numériques.

ii. Cependant, et en admettant qu'elle soit sure du point de vue de l'interprétabilité (challenge 1) et de l'intégrité des données (challenge 2), l'IA est encore rarement déployée dans la pratique à l'échelle mondiale, dans la santé par exemple, en raison de contraintes juridiques, commerciales, techniques ou financières.

iii. C'est la raison pour laquelle l'UIT et l'OMS ont mis récemment sur pied à Genève un groupe de travail conjoint chargé d'établir un cadre d'évaluation normalisé comportant des points de repère ouverts pour l'évaluation des méthodes de santé basées sur l'IA, telles que les décisions en matière de diagnostic, de triage ou de traitement. Le professeur de l'EPFL Marcel Salathé en assure la vice-présidence.

iv. Dans cet exemple lié à la santé globale, trois enjeux sont particulièrement importants à relever :
   a) Le benchmarking et la certification par des instances reconnues d'algorithmes mis sur le marché de la santé ;
   b) leur diffusion de manière à en faire bénéficier le plus grand nombre possible de personnes, indépendamment de leur situation économique et de leur localisation;
   c) leur utilisation dans la prévention et le traitement des pandémies.

v. Par extension, le challenge est de développer des mécanismes de certification et de benchmarking des algorithmes. Nous y voyons un potentiel certain pour la Suisse, ses hautes écoles et ses entreprises de même que pour la Genève internationale, y inclus le volet humanitaire posé au CICR par l'utilisation de l'IA sur des théâtres d'opérations militaires (killer robots par exemple).

4. Challenge sécuritaire no 4

Créer rapidement les conditions cadres permettant d'intégrer l'IA comme composante disruptive de la transition digitale (à l'exemple des leçons apprises du E-Voting)

i. L'IA est une composante « sur-disruptive » de la transition digitale qui, elle-même, disrupte déjà la société toute entière. Par conséquent, un développement sûr de l'IA dépend aussi de la mise en place de conditions cadres à la transition digitale dans son ensemble comme l'illustre l'exemple du e-voting en Suisse, le mécanisme de E-voting pouvant être profondément impacté par l'IA.

ii. Premier constat : en matière d'IA et de E-voting, la Suisse n'est pas protégée. La digitalisation de la société nécessite (en Suisse et ailleurs) une règlementation et un monitoring adéquat. Il paraît essentiel d'incorporer ces risques nouveaux dans toute spécification de vote (électronique ou non) afin de minimiser les risques de manipulations. Or aujourd'hui, la Suisse ne possède pas de mécanisme d'identification numérique commun susceptible d'être utilisé par toutes les administrations. Les divers services publics utilisent chacun leurs propres identifiants. De plus, ces identifiants utilisent une approche simpliste basée sur les mots de passe et non pas sur une authentification forte à facteurs multiples. En conséquence, toute proposition de vote électronique doit actuellement compenser le manque de cartes d'identités électroniques, avec de fortes conséquences sur la solution proposée. D'un point de vue purement technologique, la généralisation d'un système E-ID à authentification forte est une condition-cadre qui devrait donc précéder l'adoption massive du vote électronique, qui, lui ensuite, pourrait se baser dessus. Elle pourrait d'ailleurs également renforcer la sécurité du vote par correspondance.

iii. Deuxième constat : de manière encore plus problématique, nous anticipons une évolution rapide des systèmes de vote électronique dans le monde, basés sur des normes internationales et sur des innovations technologiques futures. Le processus démocratique est à risque de manière globale, en Suisse et dans le monde. Il s'agit, en partie, d'une conséquence directe de la digitalisation. L'histoire récente du Brexit, des élections américaines de 2016 et de certaines élections au Congrès de 2018 (p.ex. en Caroline du Nord) ont montré clairement que des intérêts étrangers ont utilisés des moyens digitaux et l'IA pour manipuler le processus démocratique (e.g., Cambridge Analytica). Plutôt que de se précipiter seul dans un chemin isolé, il paraît judicieux pour la Suisse d'inviter régulièrement la communauté internationale pour une revue externe de notre approche étape-par-étape prenant en compte l'évolution constante des technologies.

3ème Partie : Activités en cours dans les deux Ecoles polytechniques fédérales

Qu'il s'agisse de digitalisation ou d'IA, les deux EPFs font toutes deux partie des meilleures universités mondiales en matière de computer sciences et de sciences digitales, qu'il s'agisse de formation, de recherche ou d'innovation. Certains classements les situent mêmes toutes deux parmi les dix meilleures du monde. Et le capital-risque international investi en Suisse dans le domaine digital privilégie des start-ups établies sur ou à proximité de leurs campus.

Cette situation est le fruit des pré-investissements que les deux écoles ont fait à partir de leurs dotations de base au cours de ces dernières années pour anticiper la digitalisation de la société et préparer l'économie suisse de demain.

Ces pré-investissements des deux EPFs donnent à la Suisse un avantage momentané dans le domaine de la digitalisation et de l'IA. Ils sont à la fois conjoints ou complémentaies selon les accents spécifiques donnés à ces domaines par chacune des deux EPFs.

1. Exemples d'efforts conjoints EPFL-ETHZ dans le domaine digital

i. Swiss Data Science Center : les deux EPFs ont ouvert en janvier 2017 leur centre conjoint dénommé Swiss Data Science Center ou SDSC actif à la fois à Lausanne et à Zurich. Le centre réunissant déjà plus de 50 professionnels à plein temps propose son expertise en IA. Son but est d'accélérer l'adoption des data sciences par l'industrie et les universités parce que la non-collaboration entre experts de domaine (santé, banque, etc.) et les data scientists limite jusqu'ici l'impact de l'IA en Suisse. Pour y remédier, le SDSC met en premier lieu à disposition une équipe de talents en data science basée sur les deux campus qui conduit des projets de recherche et d'innovation basés sur les besoins des partenaires académiques et industriels, principalement dans les domaines de la banque, de l'advanced manufacturing et de la pharma. Ce faisant, comme il est difficile pour les entrepises suisses de recruter des experts de top niveau en data science en raison de la forte concurrence dans ce domaine, le SDSC met à disposition des talents de « niveau EPF » en data science pour tous les domaines de façon à éviter que certains domaines applicatifs des entreprises suisses engagées dans une concurrence internationale intense manquent de talents ou doivent recruter des collaborateurs de second choix (« the second crop »).

ii. Master conjoint EPFL - ETHZ en cybersécurité : le 19 mars 2019, les deux EPFs ont annoncé en présence de M. le Conseiller fédéral Parmelin, le début cet automne d'un master conjoint en cybersécurité né dans le cadre de leur coopération à la mise sur pied du Campus Cyberdéfense du DDPS dont les deux campus vont abriter une antenne dès cette année.

iii. Projet de Joint EPFL-ETHZ National Support Center for Cybersecurity : les deux EPFs proposent la mise en place à Berne d'un centre d'appui conjoint à la mise en œuvre de la stratégie nationale de cybersécurité. Les discussions à ce sujet sont en cours.

2. Exemples de centres complémentaires dans les deux EPFs pour accélérer la transition digitale

Dans le cadre des travaux menés conjointement par les deux EPFs dans le cadre des différentes stratégies numériques nationales, nous avons identifié, en plus du SDSC, l'existence de 8 centres à disposition des acteurs suisses, 4 dans chacune des EPF. Nous les citons ici pour mémoire :

i. Centres EPFL
   a) LEARN : Center for Learning Sciences
   b) C4DT : Center for Digital Trust
   c) CIS : Center for Intelligent Systems
   d) IRGC : International Risk Governance Council

ii. Centres ETHZ
   a) ZISC / SCION : Information Security Center
   b) CSS: Center for Security Studies
   c) ABZ : Center for Computer Science Teaching
   d) SNSC : Swiss National Supercomputing Center

3. Exemples d'actions spécifiques de l'EPFL en intelligence artificielle

Dans le domaine spécifique de l'IA, l'EPFL a pris trois initiatives majeures.

i. Formation en Computational Thinking

- Le computational thinking, c'est à dire la compréhension des mécanismes de fonctionnement des algorithmes est depuis la rentrée 2018 le troisième pilier de la formation polytechnique à l'EPFL. Ce pilier est obligatoire pour tous les étudiants et vient s'ajouter aux deux piliers traditionnels de l'enseignement polytechnique, à savoir les mathématiques et la physique.

- S'y ajoute un effort de l'EPFL en matière d'outreach pour tous les niveaux de formation en computational thinking en Suisse et pour tous les publics : formation primaire avec le robot Thymio, formation professionnelle avec le scale up de la Swiss Leading House en formation professionnelle existant à l'EPFL depuis 2007, formation continue avec l'Extension school et le centre for Digital Trust (par ex. formation des juristes du CICR aux rudiments du Machine Learning) y compris des formations tout public du type « 50 Things you should know about …)

ii. Applied Machine Learning Days (AMLD)

- L'EPFL organise chaque année depuis janvier 2016 les Applied Machine Learning Days (AMLD), aujourd'hui certainement le plus grand événement de Machine Learning et d'IA en Suisse et l'un des plus importants en Europe. L'objectif de la conférence, comme son nom l'indique, est de mettre l'accent sur l'application de l'apprentissage machine et de permettre aux acteurs concernés de travailler en réseau et d'échanger leur savoir-faire sur les derniers développements dans ce domaine, qui secoue l'industrie, la science, la politique et la société en général.

- L'un des points forts de l'AMLD est qu'il ne s'agit ni d'une conférence purement académique sur la technologie, ni d'une conférence purement commerciale. Il s'agit plutôt d'une manifestation intermédiaire, à cette intersection importante mais souvent difficile à trouver entre les deux, et qui se recoupe de plus en plus avec le secteur public, comme en témoignent nos co-parrains organisationnels que sont l'OMS et le CICR.

- Les AMLD attirent les meilleurs experts mondiaux dans ce domaine, tant sur le plan technologique que sur celui de l'industrie et nous souhaitons que sa quatrième édition (27-28 janvier 2020), accueille également des leaders politiques.

- Pour sa part, l'édition 2019 (du 26 au 29 janvier au Swisstech Convention Center de l'EPFL) a pu compter sur des conférenciers à fort impact, parmi lesquels Garry Kasparov, Zeynep Tufekci (techno-sociologue de l'University of North Carolina et de l'Université Harvard) et Jeff Dean, directeur de Google AI.

- Cette manifestation a réuni environ 2000 personnes tout au long de l'événement, avec des participants du milieu universitaire, de l'industrie et d'organisations gouvernementales et non gouvernementales. Des keynotes y ont alterné avec 16 ateliers parallèles, spécifiques à un domaine, allant de l'IA et la santé à l'IA et la société, ou l'IA et le transport.

iii. EPFL Center for Intelligent Systems (CIS)

- Le CIS est le dernier-né des centres de l'EPFL. Il est entièrement dédié à l'IA et à ses applications. Partant de l'idée qu'il restera difficile pour les universités, en particulier pour les universités européennes, de rivaliser avec les départements IA les plus puissants d'entreprises comme Google, Microsoft ou Facebook, l'EPFL a décidé de mettre sur pied un centre dédié à l'IA et aux systèmes intelligents.

- Son but d'attirer des chercheurs et des étudiants de haut niveau, de lancer des projets passionnants et visibles et de conduire la Suisse dans la prochaine avancée technique du 21e siècle, par exemple de la génération des futurs standards internationaux de l'IA, un domaine dans lequel la neutralité helvétique couplée à son excellence scientifique pourrait s'avérer un atout alors qu'il lui est plus difficile de rivaliser en termes économiques avec les géants digitaux, en tout cas actuellement, et tout à fait impossible de concurrencer le géopolitique des grandes puissances en matière d'IA, Etats-Unis, Chine, Russie, Japon ou Europe.

- Dans ce contexte, le CIS de l'EPFL créé fin 2018 veut :

    - Faciliter l'embauche de professeurs et de chercheurs dans ce domaine.
    - Fournir une infrastructure commune et partagée pour la recherche.
    - Élaborer un programme d'études solide appuyant le développement des systèmes intelligents dans nos écoles et dans notre pays.
    - Créer une riche dynamique d'activités pour rendre nos efforts visibles en Europe et dans le monde.
    - Développer des liens étroits et productifs avec l'industrie suisse pour l'aider à se positionner rapidement dans ces secteurs.
    - Collaborer étroitement avec les stratégies nationales numériques et avec la Genève internationale.

- La tâche est considérable : la construction de systèmes intelligents implique un grand nombre de disciplines de recherche parfois regroupées sous "AI" :

    - Science des données ;
    - Machine Learning ;
    - Vision ;
    - Traitement audio ;
    - Parole ;
    - Traitement du langage naturel ;
    - Systèmes intelligents intégratifs ;
    - Systèmes cyberphysiques ;
    - Collaboration homme-machine ;
    - IA en réseau.

- En tant que domaine dont les fondements sont encore balbutiants, les mathématiques et la théorie computationnelle fondamentale jouent un rôle essentiel.

- Un recrutement ciblé et des collaborations plus étroites avec les autorités fédérales, cantonales et avec l'économie doivent produire une présence beaucoup plus forte et visible de la Suisse dans les disciplines émergentes de l'application de l'intelligence informatique aux défis technologiques, techniques et sociétaux et à la construction de systèmes "centaure" (homme-machine), sachant :

    - que les disruptions scientifiques provenant de la recherche fondamentale et des innovations technologiques qui en découlent tirent aujourd'hui toute la chaîne de l'innovation ;

    - et que l'avantage pris par le premier acteur présent sur le marché est absolument déterminant (« First takes it all »), contrairement aux décennies précédentes où l'optimisation pouvait suffire à assurer la croissance de l'économie et les emplois de la population suisses.

4ème Partie : Propositions de mesures à prendre sur le plan fédéral

Cet ultime constat montre que si l'EPFL et l'ETHZ appuient déjà activement la transition numérique y inclus l'IA au niveau suisse, il est primordial que les conditions cadres soient améliorées rapidement si la Suisse veut réellement relever le challenge de l'IA et des systèmes intelligents d'une part, réussir et accélérer sa transition numérique d'autre part.

Aux yeux de l'EPFL, trois mesures entrant dans les compétences de proposition de l'IDAG KI paraissent nécessaires pour relever ce double challenge lié :

1. Intégrer rapidement les mesures permettant de relever les quatre challenges de l'IA dans les stratégies digitales en cours de préparation ou de mise en œuvre en Suisse (moyennant aussi le renforcement de leur coordination de ces stratégies entre elles) :

    i. **Stratégie nationale numérique « Suisse Numérique » ;**

    ii. **Stratégie nationale de cyberdéfense PACD ;**

    iii. **Stratégie nationale de cybersécurité NCS ;**

    iv. **Réseau national de sécurité Cantons – Confédération RNS.**

2. Intégrer en parallèle le développement de l'IA dans un **Programme national d'accélération de la transition numérique** avec les moyens financiers correspondants y compris mais pas seulement dans le futur Message FRI 2021-2024 incluant :

    i. **la formation** à tous les niveaux (computational thinking pour tous) y compris dans la formation continue et professionnelle ;

    ii. **la recherche** qu'il s'agisse de la recherche fondamentale en science des données (mécanismes fondamentaux de l'IA reposant sur les mathématiques de pointe) et de la recherche translationnelle disruptive en sciences digitales (développement de l'utilisation de l'IA dans tous les domaines d'applications possibles, santé, finance, e-voting, défense nationale, éducation, etc.) ;

    iii. **l'innovation** afin de renforcer l'écosystème suisse (Trust and AI Valley) dans ce domaine.

3. S'appuyer sur le savoir-faire **des deux écoles polytechniques fédérales** pour réussir la transition numérique et l'intégration de l'IA et des systèmes intelligents dans l'économie et la société suisses, écoles qui comme leur nom l'indique ont une mission nationale doublée d'une excellence reconnue sur le plan international dans le domaine digital, à charge pour elles d'impliquer les acteurs académiques, économiques, politiques ainsi que la population suisse dans cet effort de première priorité.

* * * *