



SIG et intelligence artificielle : quels développements et quel futur ?

Christian CAROLIN

AXES CONSEIL – CAROLIN@AXES.FR – HTTP://WWW.AXES.FR

L'Intelligence Artificielle, souvent dénommée par ses initiales IA, constitue un sujet à la mode. Les définitions de l'IA ne manquent pas. Elles varient selon les approches conceptuelles : cognitivisme (manipulation des symboles élémentaires par le vivant) ou connexionnisme (connexion de processus qui s'auto-organisent). Globalement, l'IA va de la reproduction/simulation de processus humains jusqu'à la construction autonome (sans intervention humaine) de ces processus.

GIS and artificial intelligence: what developments and what future?

Artificial Intelligence, often dubbed AI, is currently all the rage. Its definitions are many. They vary according to conceptual approaches: cognitivism (manipulation of elementary symbols by living beings) or connectionism (connection of self-organising processes). Roughly speaking, AI goes from reproduction/simulation of human processes all the way to fully autonomous construction (without human assistance) of those processes.



IA, mythes et réalités

Contrairement à ce qui est souvent supposé, l'IA n'est pas une technologie nouvelle, ni même récente : la machine de Turing a été conceptualisée dans les années 40, les théories de logique floue furent élaborées en 1965 et Mycin, l'un des premiers systèmes experts, date de 1973. Après diverses tentatives de reproduction/imitation du fonctionnement du cerveau humain, l'IA s'est résolument tournée vers les mathématiques et la mise au point d'algorithmes sophistiqués. Les logiciels restent essentiellement basés sur les réseaux de neurones artificiels, la statistique probabiliste et sa déclinaison que constituent les réseaux bayésiens, ainsi que les systèmes de raisonnement à base de règles et de cas.

Pourquoi maintenant ?

En clair, ce qui constitue un réel progrès et permet de présenter les succès dont la presse nous abreuve régulièrement, réside davantage dans la puissance croissante des machines que dans la progression algorithmique fondamentale. Ce qui n'était concevable que d'un point de vue théorique il y a quelques décennies, est désormais rendu possible par la loi de Moore, même empirique et déclinante : des processeurs surpuissants, les super-calculateurs, les architectures distribuées, en attendant les processeurs neuromorphiques et les ordinateurs quantiques. Les progrès paraissent spectaculaires, parce qu'ils sont illustrés par des applications du quotidien :

véhicules (plus ou moins) autonomes, chatbots sur les sites de vente en ligne, discussion avec son téléphone portable, via SIRI, Google Assistant, etc.

Quelques principes

L'IA sera de plus en plus présente dans la vie quotidienne, mais on constate déjà la disponibilité de nombreuses applications professionnelles. Les calculs de risque bancaire ou assurantiel, la prédiction dans le domaine marketing, par exemple, utilisent des fonctions d'intelligence artificielle depuis des années. S'il est difficile d'établir une approche ontologique de l'IA tant le champ de réflexion est vaste, il est possible de classer quelques technologies de base :

- Les systèmes supervisés : il s'agit de fournir au système un modèle d'apprentissage constitué d'un ensemble de cas significatifs, sur lesquels le moteur s'appuiera pour apprendre puis pour déterminer des actions (phase de test). Ces systèmes sont utilisés pour obtenir un résultat sur la base de critères prédéterminés ;
- Les systèmes non supervisés : base du *deep learning*, ces algorithmes puisent dans de vastes sources de données pour constituer des classifications, sans base de référence obligatoire permettant l'apprentissage. Ils permettent par exemple la mise en évidence autonome de corrélations, sans déterminisme préalable (exemple : les cartes auto-adaptatives, ou *maps de Kohonen*).

Désormais, on utilise une sémantique évoquant les notions d'IA forte et IA faible, correspondant

globalement, dans le premier cas, à la capacité de génération d'un raisonnement autonome (dit cognitif), là où l'IA faible (la plus développée à ce jour) raisonne sur des algorithmes déjà connus, maîtrisés par l'humain et ne progresse donc que dans les limites de la connaissance humaine (approche dite pragmatiste).

IA et SIG, « GeoIA »

Pour faire simple, on peut considérer que toute donnée géolocalisée, liée à une notion de déplacement et son corollaire, le temps, peut offrir une thématique géographique utilisable par l'IA. L'information statique est bien entendu également exploitable.

Fonctions IA intégrées ou non au SIG

Le SIG ne vit plus tout seul : il est interfacé. Il n'est plus de projet applicatif significatif sans une composante géographique et les données géolocalisées contribuent significativement à la masse d'informations stockées dans le contexte *big data*. Les sources de données peuvent être :

- Matérielles : caméra, radar, LiDAR... ;
- Les bases de données, les données non structurées ;
- Et de plus en plus, les données de source IoT (Internet des objets) et les réseaux sociaux.

À ce jour, les principaux SIG du marché n'intègrent pas de fonctions d'IA. D'une façon générale, on constate davantage la construction d'interfaces que du développement spécifique.

AI, myth and reality

Contrary to a widespread assumption, AI is neither a cutting-edge technology, nor even a recent one: the Turing Machine was formalised in the '40s, the first fuzzy logic theories appeared in 1965 and *Mycin*, one of the first expert systems, dates back to 1973. After trying to mimic or replicate the function of the human brain, AI development turned to mathematics and highly sophisticated algorithms. Software development still relies on neural networks, statistics and probability theory, together with their offspring known as "bayesian networks", and systems of reasoning based on rules and cases.

Why now?

To be honest, the real progress which leads to the success the press harps on is more likely due to the steady increase in computing power rather than any breakthrough in fundamental research. What was only dreamt of decades ago is now commonplace. Moore's Law looks like it will break down soon, but, in the meantime, it has given us ultra-powerful CPUs, supercomputers, distributed architectures, with neuromorphic CPUs and quantum processing in the offing. Progress appears outstanding, because it impacts everyday applications: (more or less) autonomous vehicles, chatbots used in various interactive websites, chat applications on mobile phones, *Apple Siri*, *Google Assistant*, and so on.

A quick AI primer

AI has yet to invade our daily life, but it has already given birth to numerous professional softwares. The fields of banking, insurance risk assessment and marketing prediction have been using AI engines for years. While an ontological approach to the vast field of AI can be tricky, some basic technologies can be listed:

- Supervised systems: the system is fed with a training model made up of a set of relevant cases, which the engine then uses to learn and later make decisions (test phase). These systems are used to get a result out of predetermined cases;
- Non-supervised systems: these algorithms are the essence of *deep learning*. They draw from huge data sources to create classifications, with no prior references given during a learning phase. They allow for autonomous discovery of correlations, without preexistent determinism (e.g. auto-adaptative maps, also known as *Kohonen's maps*).

Recently, a new taxonomy using the names "strong AI" or "weak AI" has emerged. The former concept corresponds to the ability to create an autonomous reasoning (dubbed "cognitive"). The latter operates only on known algorithms, designed by humans, and is thus bound to the limits of human knowledge (this approach is dubbed "pragmatic"). It is currently the most widespread type of AI.

AI and GIS, "GeoAI"

To keep it simple, we can assume that every georeferenced

or timestamped data can offer geographical information to an AI program. Statistical figures are obviously also processable.

AI function integrated in the GIS... or not!

No more is GIS a castaway: it communicates. There is currently no large data processing framework which does not include a tool to crunch geodata and where geodata do not represent a significant amount of stored information in a *big data* context. Data sources can be:

- Hardware: cameras, radars, lidars...;
- Databases, unstructured data;
- And more and more, IoT sources and social networks.

To date, the mainstream GIS softwares do not integrate any AI modules. Generally speaking, the trend is to offer interfacing rather than bespoke development. Several projects conflate GIS with AI engines. Their goals are:

- Improvement in analysis result quality, easing up decision making;
- Clearer interpretation of results, in a forecast context;
- Risk mitigation ahead of decision making.

A few outliers exist. The AI platform developed by the American company *Alteryx* offers functions that exploit geodata (geocoding, model building, mapping) and can also read or save geodata using most of the well-known formats (SHP, MIF/MID, MAP, KML...). The most common databases can be used as data sources.



Il existe de nombreux projets mixant les fonctionnalités natives SIG à des moteurs IA. Les objectifs du couple SIG/IA sont généralement :

- L'amélioration de la qualité des résultats d'analyse, permettant une meilleure aide à la décision ;
- La progression des interprétations, dans un contexte prédictif ;
- La limitation des risques en amont des prises de décision.

Il existe quelques exceptions, telles que la plateforme IA de l'éditeur américain *Alteryx*, qui intègre des fonctions d'exploitation de l'information géographique (géocodage, modélisation, restitution) et peut également produire des données SIG dans les formats principaux du marché (SHP, MIF/MID, MAP, KML...). Les principaux SGBDR peuvent être utilisés en source de données.

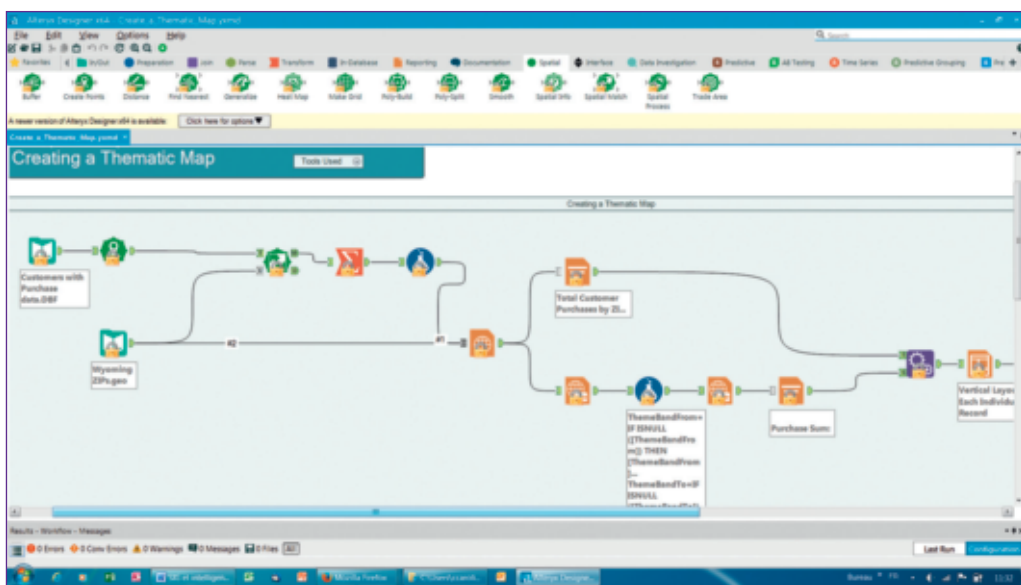
Applications thématiques

Les applications sont constituées d'association des fonctions natives des SIG avec des algorithmes IA combinés : prédictifs, classification, etc. Quelques exemples de cartes interprétatives, souvent générées à l'aide de l'inférence bayésienne :

- Simulation de l'évolution de l'occupation des sols ;
- Estimation de la variation de la végétation ;
- Simulation des catastrophes naturelles résultant des changements climatiques, pour l'aide à la gestion de crise ;
- Simulation d'accidents industriels en 3D, dans un cadre de prévention des risques ;
- Prédiction des mouvements de terrain : à titre d'exemple, l'université canadienne de

Sherbrooke a réalisé une étude portant sur l'utilisation de la télédétection, des SIG et de l'intelligence artificielle pour déterminer le niveau de susceptibilité aux mouvements de terrain, avec une application dans les Andes de la Bolivie (<http://savoirs.usherbrooke.ca/handle/11143/2709>) ;

- Il existe de nombreux projets de recherche dans le domaine de la prédiction de la déforestation, de la sécheresse ou *a contrario*, des inondations ;
- Utilisation de l'IA en géologie et pour la prospection minière et pétrolière, afin de déterminer des champs de prospection future ;
- La société canadienne *Globvision* propose un progiciel d'estimation de l'humidité et de la rugosité du sol, permettant des applications hydrologiques et agricoles. Sont cités : le suivi et prévision de la sécheresse, l'optimisation du rendement des cultures et la modélisation climatique. Le logiciel intègre une interprétation de données de télédétection pour l'estimation des paramètres de la surface du sol, à partir d'images satellite ;
- Quand la donnée n'existe pas : un éditeur SIG a réalisé une application permettant de prédire les conditions de trafic dans la ville de Koweït City, sachant qu'une surface significative de la ville ne dispose pas de données géographiques exploitables. L'IA permet, via des algorithmes appropriés, la prédiction et l'abstraction sur des zones cibles (même non intégralement couvertes par des données géographiques). Les solutions algorithmiques déployées permettent de déterminer les directions dans lesquelles la congestion de trafic s'étendra et quelles sont les mesures préventives à envisager.



Modélisation d'une carte thématique avec Alteryx.
Thematic map modelling with Alteryx.

Thematic applications

These applications combine native GIS functions with AI algorithms: prediction, classification... Here are some examples of interpretative maps. Most of them are the result of bayesian inference:

- Land use / land use evolution simulations;
- Vegetation variation assessment or estimation;
- Simulation of natural disasters caused by climate change, to aid in crisis management;
- 3D simulation of industrial disasters, in a risk mitigation context;
- Prediction of landslides: for instance, the Sherbrooke university has conducted a study over the use of remote sensing, GIS and AI to assess the landslide risk,

with an application in the Bolivian Andes (<http://savoirs.usherbrooke.ca/handle/11143/2709>);

- Several research projects target deforestation, drought or flooding prediction;
- AI is used in geology and in the mining industry to help locate future prospection fields;
- The Canadian company *Globvision* sells a software which estimates soil wetness and ruggedness, for hydrological or agricultural applications. Are cited: management and forecast of droughts, crop yield optimisation, and climatic modelling. The software includes interpretation of remote sensing data and satellite imagery to estimate soil surface parameters;
- When data is unavailable or non-existent: a GIS software company has created an application able to predict traffic conditions in Koweit City, with

the caveat that most of the city is uncharted. Using relevant algorithms, the AI engine allows prediction and abstraction over target areas (even in the absence of thorough geospatial data coverage). The software predicts how gridlocks expand and tells how to mitigate them.

Strengths and weaknesses: a case study

In the *en vogue* field of urban security, the city of Zurich has been recently using a German software named *Precobs* (*PRE Crime OBServation System*). The analysis criteria of this weak AI application (since the decisions to monitor or intervene are made by an officer) include: the type of offense, place, date, motive, and means used. The algorithms

The screenshot shows the Alteryx Designer x64 interface. On the left, a map of Sydney is displayed with several blue location markers and green drive-time regions. The main panel on the right contains a workflow diagram for the project 'ABS Census LGAB02 DriveTime v03.yxmd'. The workflow starts with 'Join ABS Local Government Area spatial data with Census attributes' (Source: ABS). This is followed by 'Use the Map Input tool to drop points right on to a map to create spatial data' and 'Specify your drive time in minutes or kms'. The output is 'Cookie cut whole Census regions with your drive times', which is then processed to 'Output to database, spatial files, Tableau, Qlik, SAS, SPSS etc'. The workflow is based on ABS data.

Understand characteristics of people near an existing / potential site
 Select locations on the map to identify sites of interest and specify a drive time in km or minutes around those sites.
 Use the drive time regions you have created to cookie cut Census demographic information and output the data to explore and analyse further.

Join ABS Local Government Area spatial data with Census attributes
 Source: ABS

Use the Map Input tool to drop points right on to a map to create spatial data

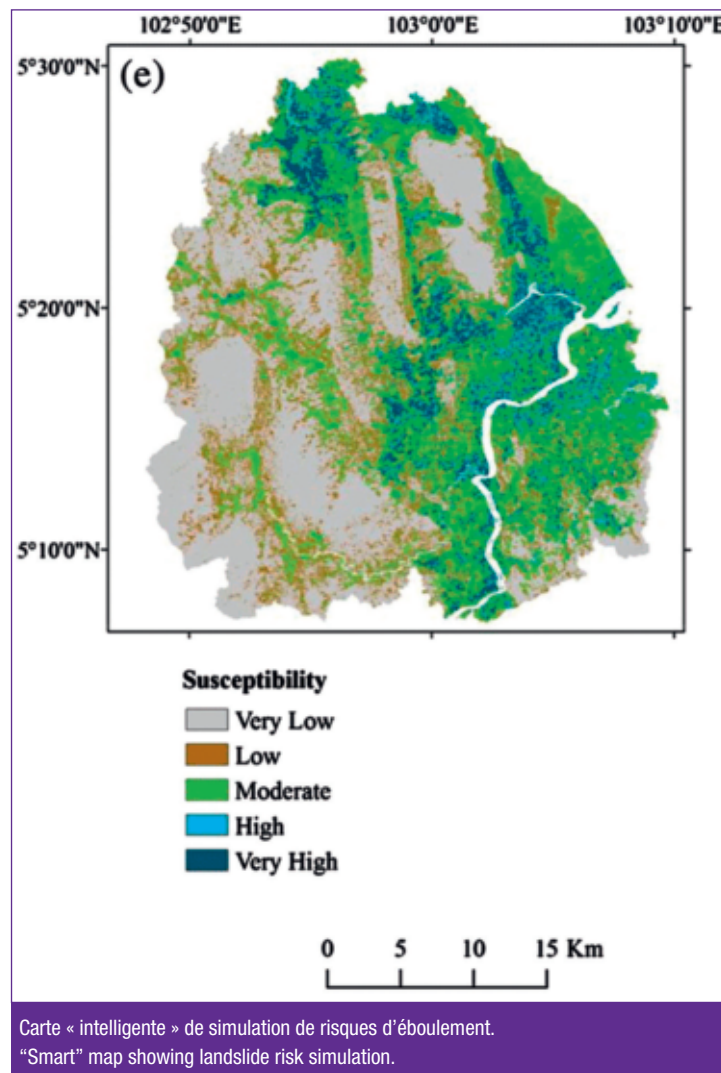
Specify your drive time in minutes or kms

Cookie cut whole Census regions with your drive times

Output to database, spatial files, Tableau, Qlik, SAS, SPSS etc
 Based on ABS data

Susan Day, MIP, 2015
 ABS Census LGAB02 DriveTime v03.yxmd

Carte générée.
 Generated map.



Points forts points faibles : étude de cas

Dans le domaine en vogue de la sécurité urbaine, la ville de Zurich utilise depuis peu un logiciel allemand nommé *Precobs* (*Pre Crime Observation System*). Les critères d'analyse de cette application IA faible (les décisions d'intervention ou de surveillance sont prises par un policier), sont notamment : le type d'infraction, lieu, date, but du délit, moyens utilisés. Les algorithmes utilisés sont de type

statistique (corrélation et régression) et intègrent également des caractéristiques topologiques et de psychologie criminelle. La probabilité des délits est évaluée sur un intervalle de temps allant jusqu'à sept jours.

La ville annonce une baisse de la criminalité de 30 %, cependant les résultats et l'efficacité du progiciel semblent être remis en question selon les interlocuteurs et usages :

- Insuffisance des sources de données (notamment, pas de

données IoT ou issues des réseaux sociaux) ;

- Lien causal entre la diminution du nombre de crimes et utilisation du progiciel non prouvé ;
- En phase de test du progiciel, la ville allemande de Nuremberg a vu sa criminalité diminuer puis augmenter selon la période de l'année, sur des bases algorithmiques identiques ;
- Limitation du fonctionnement du progiciel aux actes criminels professionnels.

Les limites de la « Geo IA »

Au-delà de l'exemple précédent d'IA faible, une future Géo IA forte pose naturellement le problème du comportement humain. La simulation des déplacements dans un espace géographique implique en effet que l'algorithme reproduise le raisonnement humain. Or, celui-ci est souvent arbitraire, chaque personne ayant sa propre perception de son mode de déplacement, de sa vitesse, de son environnement géographique. Les notions de vitesse ou de proximité sont certes des principes assimilables par la logique floue, liés à l'environnement de chacun. Mais l'apport de sciences et approches additionnelles, telles que les neurosciences, semblent indispensables pour élaborer des raisonnements exploitables.

Il convient également d'intégrer le fait que la valeur de l'IA réside très largement dans la qualité de la donnée source, au-delà de la performance algorithmique, que l'IA soit forte ou faible. Par exemple, le poids synaptique des liens d'un réseau neuronal formel dépend totalement des

used are primarily statistical (correlation and regression), but they also take into account topological details as well as elements of criminal psychology. Forecast data is available up to seven days in advance.

The city boasts a 30% decrease in criminal activity, yet the software's results and efficiency vary according to its users or usages:

- Scarcity or insufficiency of data sources (especially no data from IoT or social networks);
- The statistical correlation between the decrease in crime activity and use of the software is moot;
- During the test phase, the city of Nurnberg saw its crime activity fluctuate along the year, despite the algorithms being unchanged;
- The software is limited to "professional crime" activity.

GeoAI limits

Beyond the former example of weak AI, a future strong AI poses the problem of emulating human behaviour. Simulation of movements in geographical space entails reliable reproduction of human thinking. However, that thinking is often arbitrary, each person having their own perception of their trajectory, speed or surroundings. While it is true that speed or proximity can theoretically be modelled in fuzzy logic, further scientific support or more complex approaches, such as neurosciences, seem required to mimic reasoning.

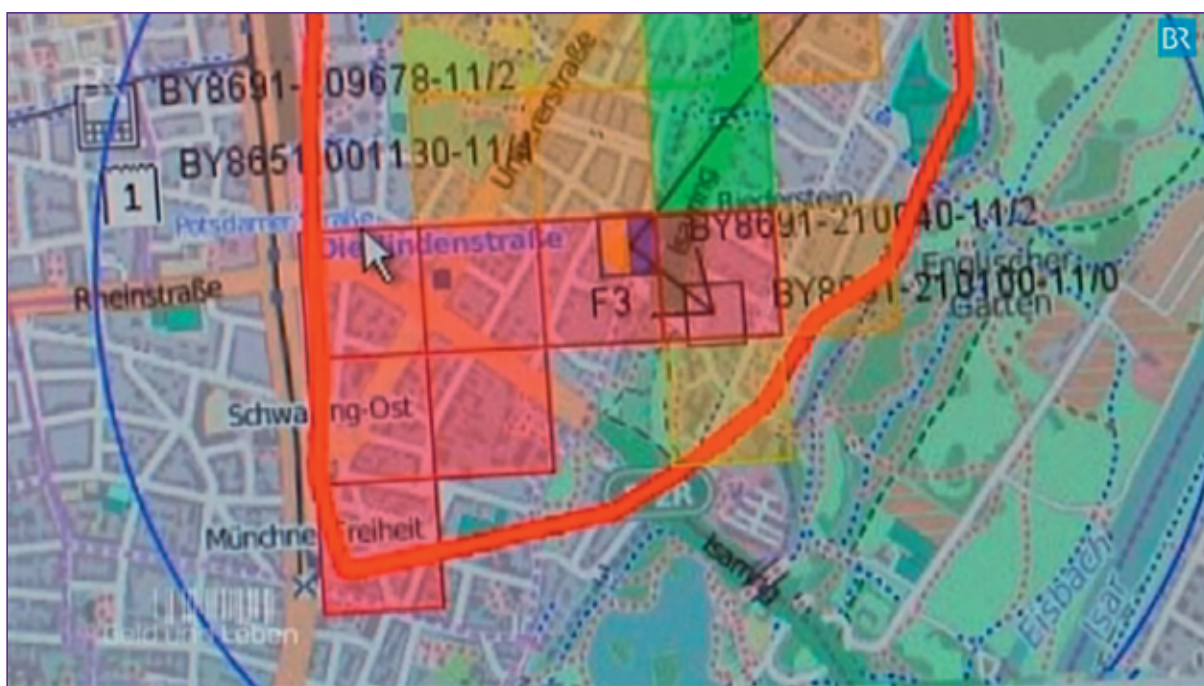
One should also consider that data quality is paramount to weak or strong AI's performance, irrespective of algorithm efficiency. For example, the synaptic weights of the connexions in a formal neural network depend

on, and only on, the data used and its quality. All these considerations usher us into a vast field of investigation, mostly unexplored as of today.

Into the future: is strong AI inevitable despite the risks?

The pair GIS/AI is still in its early stage. As aforementioned, many applications coupling GIS with weak AI are available. Strong AI is still years down the line.

It is difficult to conclude without mentioning or pondering on the risks associated with autonomous AI. This goes beyond technology and flirts with ethics and law. A deep political reflection should be ordered, and measures taken so that in all circumstances a human could stop



Sur la carte, les zones représentant un risque fort de délit (ici, les cambriolages) sont en rouge (grille de 250 mètres).
On this map, areas with the highest risk of burglary are coloured in red (250 m grid).



données utilisées et de leur qualité intrinsèque. Tous ces aspects mettent en évidence un champ d'investigation vaste, à ce jour embryonnaire.

Le futur : l'IA forte, inéluctable mais risquée ?

Le couple SIG/IA est balbutiant. Comme présenté précédemment, de nombreuses applications couplant SIG et IA faible sont déjà disponibles mais amenées à progresser et se développer. La « Géoinformatique IA forte » reste à venir.

Il est difficile de conclure ce propos sans évoquer et s'interroger sur les risques potentiels inhérents à l'IA autonome. Ceci dépasse la technologie et est du ressort de l'éthique et de la loi. Une réflexion politique approfondie devrait être menée pour qu'en toutes circonstances, l'humain puisse interrompre unilatéralement le fonctionnement autonome d'une IA. Restreint à notre champ de réflexion géomatique,

que penser d'un algorithme dédié au militaire qui déterminerait la trajectoire d'un missile, d'un avion, d'un drone, sans possibilité d'intervention d'un ingénieur ? Que penser d'une voiture autonome qui ne permettrait pas à l'homme de reprendre le contrôle du véhicule si le logiciel décidait d'une direction en contradiction avec la volonté du passager ? Le danger est d'ores et déjà largement annoncé par Elon Musk ou Stephen Hawking, personnalités parfois fantasques mais rarement fantaisistes. (<https://futureoflife.org/ai-open-letter>).

Il est étonnant de constater que la mesure de ce danger, même potentiel, ne semble faire l'objet d'aucune identification réelle, d'aucun groupe de travail dédié, dans le rapport national (https://www.economie.gouv.fr/files/files/PDF/2017/Rapport_synthese_France_IA_.pdf), qui a suivi la Journée de l'IA en mars 2017, conclue par le président de la République. La notion de risque y est évoquée mais majoritairement dédiée à la dimension économique. En France, pays du principe de précaution, l'auto-

rité politique aurait pourtant une occasion, pas si fréquente, de réfléchir en amont et peut-être de légiférer de façon proactive et non en réaction à des événements passés, souvent dramatiques. Il faudrait penser un texte de loi simple, s'inspirant partiellement, pourquoi pas, des lois d'Asimov et imposer qu'il n'existe aucun cas dans lequel l'humain ne peut reprendre le contrôle d'une intelligence artificielle, qu'une application IA forte ne peut jamais être diffusée, commercialisée et exploitée sans respecter cette règle.

Cependant, des experts certifient que le danger est faible, voire inexistant. Alors, que penser des performances et de la progression de *DeepMind*, du programme *Watson* ou du programme *Google Brain*, IA qui a créé son propre langage, incompréhensible par les humains ? La théorie de la Singularité est-elle totalement stupide ? Ray Kurzweil est-il fou ? Souhaitons donc ces experts soient plus clairvoyants que ceux qui prédisaient qu'*Internet* ne remplacerait jamais le *Minitel*... ■

the function of any AI. What of a military algorithm able to compute the trajectory of a drone, a plane or a missile without the supervision of an engineer? What of an autonomous vehicle where the passenger cannot override the steering mechanism, even if the onboard computer is driving it astray? That danger is already trumpeted by Elon Musk or Stephen Hawking, two eccentric but not flighty people (<https://futureoflife.org/ai-open-letter>).

Amazingly enough, this danger, however remote, is not mentioned in the recent national report ([https://www.economie.gouv.](https://www.economie.gouv.fr/files/files/PDF/2017/Rapport_synthese_France_IA_.pdf)

[fr/files/files/PDF/2017/Rapport_synthese_France_IA_.pdf](https://www.economie.gouv.fr/files/files/PDF/2017/Rapport_synthese_France_IA_.pdf), in French) of the “AI’s day”, an event held in March 2017 and capped off with an address from the *Président de la République*. Risk is touched upon, but the focus is clearly on economical threats. In France, the country that invented the precautionary principle, authorities would have a unique opportunity to ponder ahead of time and not act in response to a past, and maybe dramatic, incident. A bill that could refer to Asimov’s three laws of robotics should be passed, enforcing mandatory human control over any type of AI program. No

strong AI application should ever be written, sold, deployed or operated that would violate this law.

However, some experts claim that this danger is insignificant, even possibly fanciful. If this is the case, what should we think of the progress and performance of *DeepMind*, *Watson*, or of *Google Brain*, an AI software which created its own language, obscure to any human? Is Ray Kurzweil a kook? Let’s hope that these people are more insightful than those who predicted that the *Internet* would never overthrow the French *Minitel*... ■