

CARTE
BLANCHELes GAFA
pour réguler
Google Duplex ?

Par LAURENT ALEXANDRE

L'intelligence artificielle (IA) connaît une bulle médiatique et financière qui concerne surtout les « chatbots », capables de communiquer oralement avec des humains. Yann LeCun, qui dirige FAIR, le laboratoire d'IA de Facebook, s'est emporté contre les créateurs de Sophia, le robot de Hanson Robotics : « Elle n'a pas de sentiment, pas d'opinion, aucune compréhension. C'est une marionnette. » Choqué par cette « IA Potemkine », il a même accusé d'escroquerie les médias qui la promeuvent. Yann LeCun a raison de dire qu'un chatbot capable de mener une discussion généraliste ne sera pas sur le marché avant dix à quinze ans.

Pour des applications plus limitées, Google vient cependant de frapper l'opinion en présentant Google Duplex, capable de passer des coups de téléphone à notre place. Cette IA communique dans un langage naturel et peut prendre par exemple des rendez-vous chez un coiffeur, en discutant de l'horaire et du type de coupe. Google a obtenu une avancée spectaculaire : la conversation est fluide, avec un jeu varié d'intonations. Google Duplex est désormais en avance sur Apple, dont l'assistant Siri ne progresse que lentement.

Des « hum hmm » dans les discussions

Bien sûr, pour interagir avec un coiffeur ou avec un restaurateur, Google Duplex est entraîné pour reconnaître les questions et mots utilisés dans un contexte donné ; il n'est pas encore capable de discuter sur des sujets généraux. Duplex fonctionne en pratique avec un réseau de neurones éduqués à partir de communications téléphoniques, pour comprendre le sens des mots en fonction du contexte. Pour chaque domaine (coiffure, restauration, pédiatry...), Google a créé un modèle de compréhension ad hoc. Pour répondre, Duplex utilise plusieurs moteurs dits « TTS » extrêmement complexes qui permettent même de contrôler l'intonation de la voix en fonction des circonstances. Pour humaniser sa conversation, Google Duplex intègre des petits « hum hmm » dans les discussions.

Google prévoit un déploiement progressif pour proposer à terme une IA aidant les utilisateurs en toute circonstance. Nous allons devoir gérer les immenses possibilités mais aussi les risques de ces automates. Une IA comme Google Duplex, manipulée par des hackers, pourrait réserver un million de tables de restaurants en quelques instants ou encore tromper notre banquier ou nos collègues de travail : les canulars radiophoniques sembleront alors bien mièvres ! Il va falloir contrôler ces IA d'un genre nouveau et encadrer leur utilisation, par le biais de régulateurs qui garantiront que nous – humains – avons bien l'intention d'aller au restaurant ou demandons vraiment un virement. Pour cela il faudra un registre centralisé des IA de type Duplex et de leurs utilisateurs, qui pourrait d'ailleurs utiliser la blockchain.

Quelle que soit la technologie utilisée, seuls les géants du numérique – GAFA américains et BATX chinois – seront en mesure d'être ces tiers de confiance, tant la lutte contre les utilisations frauduleuses de ce type d'IA sera complexe. Il n'y a qu'une façon d'améliorer la sécurité de l'écosystème numérique mondial sur lequel notre société repose : encore et toujours plus d'IA. Il est évident que seule l'IA peut nous protéger contre des attaques ultrasophistiquées. Ainsi, chaque banque affronte des millions d'attaques par jour qu'aucune équipe humaine ne pourrait même compter, et Google a déjà montré, en juillet 2017, que son IA dépitait mieux que les humains les vidéos extrémistes postées sur sa filiale YouTube. La régulation des IA de type Duplex pose des problèmes encore plus redoutables : elles vont, à l'horizon 2030, renforcer le monopole des géants du numérique. ■

Laurent Alexandre
Chirurgien urologue,
président de DNAVision
l.alexandre@dnavision.be
PHOTO: JEAN CHISCANO

La protection du droit d'auteur ne doit pas
entraver la circulation des logiciels

TRIBUNE - Des informaticiens s'inquiètent des dangers du projet européen de réforme du droit d'auteur présenté à Bruxelles. Il constitue, selon eux, une menace pour les processus actuels de production « open source »

Les logiciels sont partout. Vous les utilisez au quotidien pour communiquer, travailler et vous divertir. Ils sont essentiels pour la gestion de nos entreprises, la recherche avancée, la création et la diffusion des connaissances et des arts. Nos industries, notre société, notre culture, voire nos propres vies dépendent des logiciels, qui font aujourd'hui partie intégrante du patrimoine de l'humanité.

Mais ces logiciels, véritable moteur de la transformation numérique, ne tombent pas du ciel : ils sont développés par des êtres humains qui les écrivent dans une forme qu'on appelle le code source, en utilisant des langages de programmation. Nous, qui développons ces logiciels, sommes donc bien des auteurs : les codes sources des logiciels que nous créons sont couverts par le même droit d'auteur qui protège la musique, les livres ou les films.

Par conséquent, les dispositions relatives au droit d'auteur nous concernent en premier lieu, et nous sommes très inquiets des propositions contenues dans le projet de directive européenne réformant le droit d'auteur actuellement examiné par le Parlement européen et le Conseil.

En particulier, l'article 13 de ce projet de directive introduit l'obligation, pour toute plate-forme permettant de partager des contenus, de mettre en place des filtres automatiques, du style de ceux qui bloquent sur YouTube les vidéos qui réutilisent des contenus protégés – l'objectif affirmé étant de

prévenir la diffusion d'œuvres sans l'autorisation des auteurs, et de garantir ainsi leur rémunération.

Dans la continuité de la lettre ouverte au Conseil du 26 avril rassemblant 147 organisations européennes, que nous avons soutenues, nous souhaitons aujourd'hui alerter les députés européens et les représentants des Etats membres sur les menaces spécifiques que ce projet de texte fait peser directement sur le logiciel libre et, par son biais, sur toute l'industrie du logiciel.

Aujourd'hui, la plupart des logiciels sont construits en réutilisant des composants préexistants, développés et distribués sur des plates-formes ouvertes de développement collaboratif. Tout comme Linux, qui est au cœur de plus de 80% des téléphones portables, il y a des millions de logiciels construits par des auteurs qui ont choisi d'en faire des logiciels libres, ce qui veut dire que tout le monde peut lire, étudier, modifier, faire modifier et redistribuer leurs codes sources, sans restriction ni autorisation particulière.

On estime que 80% à 90% d'une application informatique moderne sont issus de cette réutilisation, et la suppression de chacun de ces composants peut avoir des conséquences imprévisibles : on a pu le constater quand, en 2016, la disparition de onze lignes de code source a cassé des millions de sites Web. Imposer des filtres automatiques sur ces plates-formes ouvertes de développement collaboratif reviendrait donc à menacer les pro-

IMPOSER
DES FILTRES
AUTOMATIQUES
AURAIT DES IMPACTS
MAJEURS SUR
L'INNOVATION
DE NOS INDUSTRIES

cessus actuels de production de ces logiciels, et aurait des impacts majeurs sur l'innovation de nos industries et la compétitivité de nos économies.

Le logiciel libre, également appelé *open source*, est en effet un socle technologique indispensable qui permet de développer plus rapidement la plupart des logiciels dont a besoin notre société. Et il est aussi une filière économique dynamique, qui représente aujourd'hui en France un chiffre d'affaires de 4,5 milliards d'euros, avec plus de 500 entreprises et 50 000 emplois.

Si nous comprenons l'inquiétude de certains acteurs de l'industrie culturelle qui se sentent démunis face aux changements apportés par la révolution numérique, il convient de rappeler que le droit d'auteur concerne tout autant les auteurs de logiciels que les auteurs de l'industrie culturelle. Cette réforme doit donc être élaborée en concertation avec tous ces acteurs concernés par le droit d'auteur, et non uniquement ceux de l'industrie culturelle.

Nous tenons donc à alerter sur les menaces que l'actuel projet de directive comporte pour les acteurs du logiciel, et par là même pour la société tout entière : cela va des freins au développement de nouvelles technologies, à cause du blocage dans l'accès au *text and data mining* (art. 3), aux entraves graves au développement collaboratif et à la réutilisation des logiciels que nous avons analysés ici en détail (art. 13).

Une exclusion totale des logiciels des dispositions de l'article 13 et la levée de toute restriction sur le *text and data mining* (art. 3) nous apparaissent nécessaires pour ne pas créer d'effets collatéraux majeurs avec cette réforme, pensée avant tout pour les acteurs de l'industrie culturelle. ■

¶ Serge Abiteboul, informaticien, Inria et ENS Paris, Académie des sciences ; Pierre Baudracco, président du programme des Paris Open Source Summit 2017 et 2018 ; Laurent Baudart, délégué général de Syntec Numérique ; Roberto Di Cosmo, directeur de Software Heritage, professeur d'informatique, Inria et université Paris-Diderot ; Stéphane Fermigier, coprésident du Conseil national du logiciel libre (CNLL) ; Philippe Montargès, coprésident du CNLL ; Pierre Paradinas, président de la Société informatique de France, professeur au CNAM ; Emmanuelle Roux, directrice associée de SC21 et de LeChaudron.io ; Cédric Thomas, directeur général d'OW2.

Le supplément « Science & médecine » publie chaque semaine une tribune libre. Si vous souhaitez soumettre un texte, prière de l'adresser à sciences@lemonde.fr

COLLECTION GÉNIES DES MATHÉMATIQUES

Riemann, une hypothèse
de premier plan

Lorsqu'il s'éteignit à l'âge de 40 ans, en 1866, le mathématicien allemand Bernhard Riemann avait publié seulement six articles. Son œuvre complète, qui les rassemble au côté de textes posthumes, ne totalise même pas 500 pages. Cela semble bien peu, comparé aux dizaines de milliers de pages d'écrits mathématiques d'Euler ou Cauchy. Pourtant, ces textes font de Riemann l'un des plus grands mathématiciens de tous les temps. En effet, il modifia profondément la manière de concevoir certaines des notions les plus importantes dans ce domaine.

Plusieurs de ces notions remontent aux *Éléments* d'Euclide. Cet ouvrage fondateur du style de présentation des preuves mathématiques contient, entre autres, des théories au sujet des figures simples du plan et des nombres entiers. Les nombres premiers y jouent le rôle d'espèces atomiques, d'abord parce qu'ils sont indécomposables, n'étant divisibles que par un et par eux-mêmes, ensuite parce que tout nombre entier s'écrit de manière unique comme produit de nombres premiers. Mais, à la différence des espèces d'atomes chimiques, il y a une infinité de nombres premiers. La preuve qu'en donna Euclide n'a jamais été dépassée en simplicité.

Y a-t-il un lien entre le plan et les nombres premiers ? Personne n'en avait trouvé avant que Riemann ne publie en 1859 un article intitulé « Sur le nombre de nombres premiers inférieurs à

une grandeur donnée ». Par exemple, combien de nombres premiers sont inférieurs à 1859 ? Il est possible de les compter, mais cela ne dirait rien sur la réponse concernant 2018. Riemann proposa une formule approchée d'une excellente précision, valable pour toute valeur de la grandeur donnée. Cependant, il ne savait prouver cette formule que sous l'hypothèse que toutes les solutions d'une certaine équation dont l'inconnue est cherchée parmi les points du plan sont situées sur deux droites spéciales. Cette « hypothèse de Riemann » au sujet de la « fonction zêta de Riemann » – l'expression figurant dans l'équation – n'a toujours pas été confirmée ou infirmée, en dépit de grands efforts. Elle est fascinante par ce rapprochement entre les nombres premiers et le plan.

Dans l'article de Riemann, le plan représentait géométriquement les nombres dits « complexes ». Ces nombres semblaient paradoxaux lors de leur invention au XVI^e siècle, puisque en les élevant au carré on obtient parfois des nombres négatifs. Comme ils se montraient de plus en plus importants, ils furent progressivement apprivoisés. Au XIX^e siècle on s'était habitué à les voir simplement comme coordonnées des points dans un plan. On avait aussi découvert que bon nombre d'équations ou de fonctions, dont l'inconnue était d'abord cantonnée à la droite des nombres réels, pouvaient mieux s'étudier si l'inconnue était cherchée dans tout le

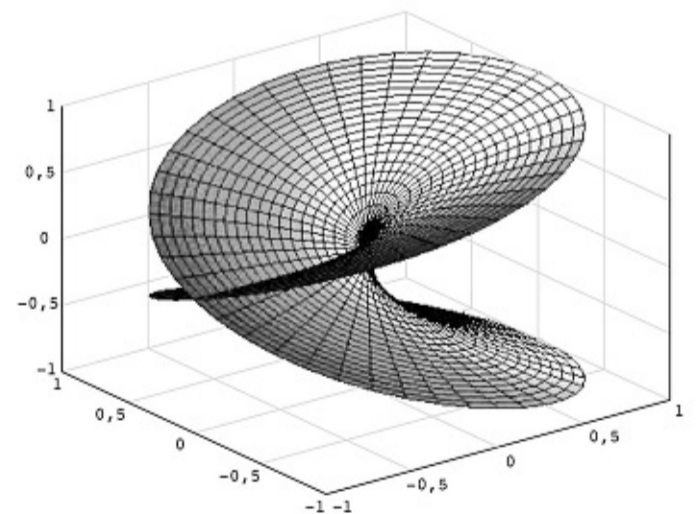


Image
détaillée
de la surface
de Riemann
de la fonction
racine carrée.
ARCHIVES RBA

La conjecture fondamentale
sur les nombres premiers
Riemann



9,99 €, en kiosque
le 24 mai.

plan des nombres complexes. Par exemple, la fonction zêta de Riemann avait déjà été considérée par Euler pour une inconnue réelle, mais c'est Riemann qui comprit l'avantage que l'on pouvait tirer du passage au plan.

Riemann a aussi imaginé une nouvelle manière de concevoir le plan et l'espace de la géométrie d'Euclide. Il les engloba sous une notion très générale d'espace courbe de dimension quelconque. Il pensait que cette notion pouvait servir à mieux comprendre la structure du cosmos : « Il faut donc, ou que la réalité sur laquelle est fondé l'espace forme une variété discrète, ou que le fondement des rapports métriques soit cherché en dehors de lui, dans les forces de liaison qui agissent en lui », écrivait-il en 1854. Cette vision eut une profonde influence sur la théorie de la relativité générale d'Einstein de 1915 et, par la suite, sur le développement de la cosmologie. ■

PATRICK POPESCU-PAMPU,
PROFESSEUR DE MATHÉMATIQUES
AU LABORATOIRE PAUL-PAINLEVÉ
DE L'UNIVERSITÉ DE LILLE, SPÉCIALISTE
EN THÉORIE DES SINGULARITÉS