



Data ontsluiten

Frank Robben



Het e-Healthplatform kan een massa gegevens die opgeslagen zitten in medische dossiers mee helpen beschikbaar maken voor secondary use. Dat zou onder meer kunnen dienen om AI-systemen te trainen. De rol van het platform is die van een trusted third party.

Frank Robben, administrateur-generaal van de Kruispuntbank van de Sociale Zekerheid en van het eHealth-platform (en gedelegeerd bestuurder van Smals vzw) uit zich wat gereserveerd over artificiële intelligentie. Hij gelooft in het nut van beslissingsondersteunende systemen voor de arts. Maar volgens hem worden de mogelijkheden op korte termijn van AI overschat.

"Wat vooral werkt is *natural language processing* en *supervised learning*. Bij dit laatste probeert het systeem te leren om met een bepaalde input van informatie tot een bepaalde output te komen. Die kennis gebruikt het vervolgens op nieuwe input. Het systeem stelt de beslissingsregels die het heeft berekend voor aan een 'supervisor', die beoordeelt of ze juist zijn. Blijkt uit de feedback dat een regel twijfelachtig is, dan krijgt die voortaan een lagere 'threshold value' en

zal de computer er minder gauw gebruik van maken. Is de regel juist dan wordt de drempel om hem te gebruiken lager gelegd. Zo leer je het systeem om tot het juiste resultaat te komen." *Unsupervised learning* of *reinforcement learning*, waarbij het systeem niet vooraf verteld wordt waar het moet uitkomen en er geen feedbackmechanisme is, levert nog niet zo een goede resultaten op. "Een AI-algoritme kan bijvoorbeeld nog moeilijk onderscheiden wat oorzaak en gevolg is. Het diagnosticeert dan bijvoorbeeld een longziekte aan de hand van de aanwezigheid van een port-à-cath op de röntgenfoto."

Frank Robben ontwikkelde als jonge jurist-informaticus beslissingsondersteunende systemen voor juristen – en

bestudeerde zogenaamde 'neurale netwerken', computersystemen die zijn ontworpen naar analogie met het menselijk brein. "Volgens mij hanteert men vandaag nog grotendeels dezelfde logische systemen. De methoden zijn in wezen niet zo geëvolueerd. Maar computers zijn ondertussen veel krachtiger geworden, ze kunnen met elkaar praten over performante netwerken, en er zijn een massa gegevens beschikbaar gekomen."

Vertrouwenspartner

Een goede toepassing voor AI zou volgens Robben kunnen zijn om te helpen bij het 'coderen' in niet-gestructureerde medische dossiers. Een 'robot' kan het dossier van de patiënt scannen

om ICD11-, ICPC3- en/of Snomed-codes te suggereren als die nog niet zijn toegepast. Waar het eHealth-platform kan bij helpen is het ontsluiten van gegevens die gedecentraliseerd zijn opgeslagen. Die zouden kunnen dienen om AI-systemen te trainen. "Maar het is niet de rol van het eHealth-platform om onderzoeksprojecten op te zetten", zegt Robben. "Wat de rol is die wij moeten spelen, is die van een *trusted third party* (TTP). We kunnen geen vertrouwenspartner zijn als we zelf als partij belang krijgen bij het ontsluiten van de gegevens. We moeten zorg dragen voor de informatieveiligheid: de gegevens beveiligen en de privacy van personen verzekeren."

Het eHealth-platform biedt services aan die, onder meer, de datasets voor het onderzoek anoniem maken. Eventueel door pseudonomisering, waarbij de informatie die een persoon identificeert, wordt vervangen door 'betekenisloze' informatie." Dat houdt in dat men wel nog weet welke gegevens bij één individu horen, zodat bijvoorbeeld de evolutie van een aandoening over gegevens verspreid in de tijd kan worden ingeschat, of medische data gekoppeld kunnen worden aan sociaal-economische data – maar zonder dat er nog gegevens worden meegegeven die de persoon herkenbaar maken. De medische gegevens over een persoon krijgt het eHealth-platform overigens zelf nooit te zien. De services van het platform versleutelen ze tijdens de transmissie. Dat is een stuk van de informatieveiligheid.

Hergebruik

Om het 'secundair' gebruik van gezondheidsgegevens te faciliteren, wordt binnen de schoot van de FOD Volksgezondheid een gezondheids(zorg) data-autoriteit (Health Data Authority of HDA) opgericht. "De HDA moet onderzoeken met welke methodologieën de kwaliteitsvolle *primary data* over gezondheidszorg kunnen worden hergebruikt voor *secondary use*. Hoe bijvoorbeeld big data-technologieën kunnen worden gebruikt op anonieme data voor wetenschappelijke research, beleidsondersteuning, de evaluatie van geneesmiddelen bij behandelingen,... Het moet dit soort van onderzoek op een veilige manier faciliteren." In de beheersorganen van de HDA zitten de

betrokken overheidsinstellingen, de onderzoeksWereld, de professionals, patiëntenvertegenwoordigers,... In het gebruikerscomité (dat advies geeft) is ook de farma-industrie vertegenwoordigd. "Maar het eHealthplatform blijft een TTP en gaat als dusdanig in het HDA zeker geen leidende rol nemen. Overigens heeft ook de EU een project in de steigers staan om het hergebruik van gezondheidsgegevens mogelijk te maken, de *European Health Data Space*. De HDA moet België hierop helpen voorbereiden."

Vandaag al gaat het onafhankelijke Informatieveiligheidscomité na of voorgestelde onderzoeken wel op een juiste manier gebruik maken van gegevens uit de gezondheidszorg (via het eHealth-platform) of de sociale zekerheid – dat is vanuit het oogpunt van veiligheid. Wie op de website van het IVC nagaat wie deze gegevens voor

'AI-toepassingen in de gezondheidszorg moeten een tool blijven voor de arts, die de ultieme controle uitoefent'

onderzoek wil gebruiken, ziet dat de vragen vooral komen van overheidsinstellingen (zoals het Riziv, het KCE, ...) en de academische wereld. Af en toe komt er ook al eens een vraag waar een farmabedrijf bij betrokken is – bijvoorbeeld om gegevens te verzamelen over een nieuw 'contractgeneesmiddel' dat een voorlopige toegang heeft gekregen tot de markt.

Covid

"Dat het gebruik van de gegevens van patiënten een goed gezondheidsbeleid kan ondersteunen, is mooi gebleken tijdens covid. De ziekenfonden en de huisartsen hebben aan de hand van de gegevens die ze beheren, de personen helpen selecteren die, toen de covidvacins geleidelijk ter beschikking kwamen, als eerste in aanmerking kwamen

voor vaccinatie. Wie prioriteit kreeg in de vaccinatiestrategie, maakte toen echt wel verschil uit."

Dat was weliswaar geen AI. "AI-toepassingen gebruiken we vandaag bijvoorbeeld al in de sociale zekerheid om fraude op het spoor te komen. Ook sommige softwaretoepassingen die tijdens de covidcrisis in de gezondheidszorg werden gebruikt, waren toch al behoorlijk intelligent. Naar welke laboratoria testcentra de monsters stuurdelen voor PCR, berustte initieel op historische gronden en verliep niet optimaal. Rekening houdend met het aantal verwachte tests per centrum, de tijd voor het transport, de kenmerken van het laboratorium is op korte tijd de kaart hertekend van wie tests aan welk lab moest leveren, zodat de tijd tussen afname en het meedelen van het resultaat opmerkelijk verkort kon worden. Voor de contact tracing in de scholen, hebben we gebruik gemaakt van graph databases. Die brengen het risico van blootstelling aan het virus dat iemand loopt, visueel en veel efficiënter in kaart."

"De software die we voor zo'n toepassingen gebruiken, wordt meestal ter beschikking gesteld door Smals vzw – dat overheidsbedrijf is eigenlijk een ledenorganisatie waar talrijke diensten van de overheid lid van zijn. Dankzij Smals kunnen we gespecialiseerde software snel (her)gebruiken in uiteenlopende situaties en kunnen we beschikken over mensen met de expertise om ermee te werken."

Instrument

AI-toepassingen in de gezondheidszorg moeten voor Frank Robben een tool blijven voor de arts, die de ultieme controle uitoefent. Hij stelt zich een robot voor die de chirurg helpt om een niertumor weg te halen. "Die kan de omtrek van de tumor met behulp van foto's duidelijk afbakenen, de incisie met precisie helpen uitvoeren, bepalen welke microvaatjes waar en wanneer moeten samengedrukt worden om de bloedtoevoer juist lang genoeg af te sluiten. Met zo'n robot kan de chirurg wellicht veel preciezer te werk gaan. Maar is dat een slimme machine? Dat een computer een grote toevloed van gegevens snel kan verwerken en veel sneller dan de mens berekeningen kan uitvoeren, weten we. Maar het blijft

vandaag nog bij deterministische, goed afgebakende processen – de robot bepaalt niet autonoom het doel. En als de robot iets stoms doet, kan de chirurg altijd ingrijpen." De computer zal de menselijke intelligentie nog niet zo snel vervangen, denkt Robben. En wanneer dat wel mogelijk wordt, moeten er eerst ethische vragen beantwoord worden. Robben zegt niet pessimistisch te zijn. "Ik ben voorstander van digitalisering – maar je moet een computer niet zomaar zijn gang laten gaan als je de risico's niet kunt beheren."

Leerdoelen

Na het lezen van dit artikel bent u vertrouwd met:

» 'Open ended' artificiële intelligentie levert nog geen voldoende betrouwbare resultaten op. AI-systemen worden getraind met situaties waarbij de uitkomst vastligt.

» Het eHealth-platform kan medische gegevens ontsluiten die gedecentraliseerd zijn opgeslagen maar belangrijk kunnen zijn voor onderzoek en ondersteuning van de gezondheidszorg.

» De Health Data Authority moet methodieken ontwikkelen voor het correcte hergebruik van medische gegevens.

» Ook farmabedrijven kunnen op deze gecontroleerde manier klinische gegevens gebruiken over de behandeling met een geneesmiddel na toelating ervan tot de markt.

» De rol van het eHealth-platform is om als trusted third party de informatieveiligheid te verzekeren.



Déverrouillage des données



Frank Robben

La plate-forme e-Health peut contribuer à rendre une masse de données stockées dans les dossiers médicaux disponibles pour un *secondary use*. Cela pourrait servir, entre autres, à former des systèmes d'IA. La plate-forme a le rôle de *trusted third party*.

Frank Robben, administrateur général de la Banque carrefour de la Sécurité sociale et de la plate-forme e-Health (et administrateur délégué de l'ASBL Smals), fait preuve d'une certaine réserve à l'égard de l'intelligence artificielle. Il est convaincu de l'utilité

des systèmes d'aide à la décision pour le médecin, mais pense qu'on surestime les possibilités de l'IA à court terme. « Ce sont surtout le natural language processing et le supervised learning qui fonctionnent. Dans ce dernier cas, le système essaie d'obtenir un certain output grâce à un input d'informations. Il utilise

ensuite cette connaissance pour un nouvel input. Le système présente les règles de décision qu'il a calculées à un supervisor, qui vérifie leur correction. Une règle que le feedback juge discutable se voit attribuer un threshold value inférieur et est moins susceptible d'être utilisée. Si la règle est correcte, le seuil d'utilisation est abaissé. C'est ainsi qu'on apprend au système à parvenir au bon résultat. » Le unsupervised learning ou reinforcement learning ne permet pas au système de prédire à quel résultat il doit aboutir et est dépourvu de mécanisme de feedback. Il ne produit pas encore d'aussi bons résultats. « Par exemple, un algorithme

d'IA a encore du mal à distinguer la cause de l'effet. Il va par exemple diagnostiquer une maladie pulmonaire en se basant sur la présence d'un port-à-cath sur la radiographie. »

Jeune juriste informaticien, Frank Robben a développé des systèmes d'aide à la décision pour les avocats et a étudié les « réseaux neuronaux », des systèmes informatiques conçus par analogie avec le cerveau humain. « Selon moi, on continue à utiliser majoritairement les mêmes systèmes logiques actuellement. Les méthodes n'ont pas beaucoup évolué mais les ordinateurs sont beaucoup plus puissants, ils peuvent communiquer

entre eux par le biais de réseaux très performants et nous disposons d'une masse de données.»

Un partenaire de confiance

Selon Robben, l'IA pourrait aider au «codage» des dossiers médicaux non structurés. Un «robot» peut scanner le dossier du patient pour suggérer des codes CIM11, ICPC3 et/ou Snomed s'ils n'ont pas encore été appliqués. Quant à la plate-forme e-Health, elle peut aider à déverrouiller les données stockées de manière décentralisée, qui pourraient ensuite servir à former des systèmes d'IA.

«Toutefois, la plate-forme e-Health n'a pas pour vocation le lancement de projets de recherche», précise Robben. «Nous occupons le rôle de trusted third party (TTP). Nous ne pouvons pas être un partenaire fiable si nous avons intérêt à déverrouiller les données. Nous devons veiller à la sécurité des informations, les protéger et assurer la privacy des personnes concernées.»

La plate-forme e-Health fournit des services qui rendent notamment les sets de données anonymes. Éventuellement par pseudonymisation, ce qui signifie que les informations pouvant identifier une personne sont remplacées par des informations «sans signification». On sait toujours quelles données correspondent à tel individu, notamment pour être en mesure de suivre l'évolution d'une maladie grâce aux données recueillies au fil du temps, ou pour pouvoir associer des données médicales à des données socio-économiques, mais sans transmettre d'éléments qui permettent d'identifier la personne. D'ailleurs, la plate-forme e-Health ne peut jamais accéder aux données médicales d'une personne. Ses services les verrouillent pendant leur transmission pour assurer la sécurité des informations.

La réutilisation

Afin de faciliter l'usage «secondaire» des données sanitaires, on a fondé au sein du SPF Santé publique une autorité en matière de données médicales, la *Health data authority* ou HDA. «La HDA est chargée d'examiner les méthodologies permettant la réutilisation de primary data de qualité pour un secondary use. Par exemple, comment utiliser les technologies big data sur des

données anonymes pour la recherche scientifique, le soutien à la gouvernance, l'évaluation des médicaments dans les traitements... La HDA doit faciliter ce type de recherche en toute sécurité.» Les organismes publics concernés, le monde de la recherche, les professionnels, les représentants des patients siègent dans les organes de gestion de la HDA.

L'industrie pharmaceutique est également représentée au sein du comité des utilisateurs (qui donne son avis). «Toutefois, la plate-forme e-Health demeure une TTP et n'endossera donc aucun rôle prépondérant au sein de la HDA. D'ailleurs, l'UE a également lancé un projet permettant la réutilisation des données sanitaires, l'European Health Data Space. La HDA doit aider la Belgique à s'y préparer.»

Le Comité de sécurité de l'information (CSI), indépendant, vérifie déjà si les études proposées utilisent correctement

«Dans le domaine de la santé, les applications d'IA doivent rester un outil pour le médecin, qui exerce le contrôle ultime.»

les données provenant des soins de santé (via la plate-forme e-Health) ou de la Sécurité sociale - au niveau de la sécurité. Un coup d'œil au site web du CSI révèle que ce sont surtout le milieu académique et les institutions publiques, comme l'ONSS, la BCSS, etc., qui demandent l'autorisation d'utiliser ces données pour des recherches. De temps à autre, une demande émane d'une firme pharmaceutique, par exemple pour rassembler des données sur un nouveau «médicament contrat» qui a été provisoirement autorisé sur le marché.

Le covid

«On a pu constater pendant le covid que l'utilisation des données des patients pouvait bel et bien favoriser une bonne

politique de la santé. À partir des données qu'ils gèrent, les caisses d'assurance maladie et les médecins généralistes nous ont aidés à sélectionner les personnes prioritaires durant la campagne de vaccination. Les priorités accordées dans la stratégie de vaccination ont vraiment fait une différence ce moment-là.»

Il ne s'agissait pas d'IA. «À l'heure actuelle, nous avons recours aux applications d'IA pour détecter les fraudes en matière de sécurité sociale. Certains logiciels employés par le secteur médical pendant la pandémie étaient déjà très intelligents. Initialement, les centres de test envoyoyaient les échantillons PCR à des laboratoires sur base de motifs historiques et cela ne se déroulait pas de manière optimale. En peu de temps, en tenant compte du nombre attendu de tests PCR par centre, du temps de transport et des caractéristiques des laboratoires, nous avons établi qui devait envoyer ses tests à quel labo, pour réduire drastiquement le laps de temps entre le test et la communication du résultat. Nous avons employé des bases de données graphiques pour le traçage des contacts dans les écoles. Elles permettaient de visualiser plus efficacement le risque d'exposition au virus des différentes personnes.»

«En général, c'est l'ASBL Smals qui met à notre disposition ce type de logiciel. Cette entreprise publique est en fait une association de membres, essentiellement issus des nombreux services publics. Grâce à Smals, nous pouvons (ré)utiliser rapidement le logiciel spécialisé dans toutes sortes de circonstances et nous disposons de personnes dotées de l'expertise requise pour s'en servir.»

Un instrument

Aux yeux de Frank Robben, les applications d'IA doivent rester un instrument pour le médecin, le seul habilité à procéder à l'ultime contrôle. Il imagine un robot qui aide le chirurgien à procéder à l'ablation d'une tumeur rénale. «Grâce à des photos, il peut délimiter clairement le pourtour de la tumeur et donc procéder à une incision précise, déterminer quels micro-vaisseaux comprimer et quand pour n'interrompre la circulation sanguine que le temps requis. Pareil robot aide le chirurgien à travailler avec beaucoup plus de précision. Mais est-ce une machine intelligente? Nous savons qu'un ordinateur peut traiter rapidement un grand nombre de données et

effectuer des calculs beaucoup plus vite que l'être humain mais à l'heure actuelle, il est cantonné à des processus bien définis. Le robot ne détermine pas l'objectif de manière autonome. Et si le robot commet une erreur, le chirurgien peut toujours intervenir.» L'ordinateur ne remplacera pas de sitôt l'intelligence humaine, selon Robben. Et quand ce sera possible, il faudra d'abord se pencher sur des questions éthiques. Il affirme pourtant ne pas être pessimiste. «Je suis partisan de la digitalisation mais il ne faut pas laisser l'ordinateur suivre son chemin si on ne peut pas gérer les risques.»

Objectifs d'apprentissage

La lecture de cet article vous aura familiarisé(e) avec :

- L'intelligence artificielle «Open ended» ne fournit pas encore de résultats suffisamment fiables. Les systèmes IA sont formés à partir de situations dont le dénouement est fixé;
- La plate-forme e-Health peut déverrouiller des données médicales stockées de manière décentralisée qui peuvent être importantes pour la recherche et le soutien sanitaire;
- La *Health Data Authority* doit mettre au point des méthodes de réutilisation correcte des données médicales;
- Les firmes pharmaceutiques peuvent également utiliser ces données cliniques contrôlées sur l'efficacité d'un médicament après son entrée sur le marché;
- En tant que trusted third party, la plate-forme e-Health doit assurer la sécurité des informations.