

Bias als buzzwoord: bijeenkomst KB over Artificiële Intelligentie en bibliotheek

Bart Janssen

16-07-2019

Het is tegenwoordig niet ongebruikelijk om in een willekeurige bibliotheek aan te lopen tegen een paar vrolijk beppende Peppertjes of ander robotisch vernuft, maar waarheen voert de weg der Artificiële Intelligentie (kortweg AI) de bibliotheek verder nog? Het was de vraag die centraal stond tijdens een bijeenkomst die de Koninklijke Bibliotheek op 2 juli organiseerde.

Vanuit de Verenigde Staten reflecteerde via Skype 'bibliotheekgoeroe' David Lankes als eerste op de relatie tussen de bibliotheek en AI. In zijn lezing, getiteld '*Decoding AI and Libraries*', gaat Lankes in op drie aspecten op het gebied van AI: (big) data, algoritmes en *machine* cq *deep learning*. 'If we, librarians do not seek to have positive impacts at each of these levels – I would argue we are not doing our job and putting our communities in danger,' aldus Lankes.

Bijna iedereen lijkt het als een gegeven te beschouwen dat aan online zijn kosten verbonden zijn in de vorm van het delen van data met de grote internetbedrijven. De rol die bibliotheken hier moeten opeisen is die van '*steward of public data*', zoals Lankes het noemt. 'Libraries have a VERY long history of data stewardship that includes respect for privacy and seeking equitable access to information. If we are going to allow our governments and our businesses to harvest data then we need to ensure our communities have a strong say in how that happens and trust in those that make the decisions. Right now, libraries have a stronger level of trust than Apple, Google, Facebook, and most elected governments,' aldus Lankes.



Ook met betrekking tot de toepassing van algoritmes ziet Lankes een rol voor bibliotheken. Verwijzend naar enkele boeken, zoals *The Power of Habit* van Charles Duhigg en *Weapons of math destruction* van Cathy O'Neil, noemt hij een aantal verbluffende voorbeelden van de gevolgen van *data mining* en het gebruik van algoritmes. Met name O'Neil wijst in haar boek op de misvatting dat vaak gedacht wordt dat algoritmes objectief en neutraal zijn, terwijl er in feite allerlei aannames onder liggen die de uitkomst beïnvloeden. Lankes volgt haar daarin en ziet hier een rol voor bibliotheken op het gebied van voorlichting en bewustwording en waar het gaat om regulering. Maar, zo stelt hij, ook bibliotheken moeten zich bewust zijn van de aard van hun eigen stem. Lankes introduceert op dit punt hét buzzwoord van de dag: '*bias*', (vaak niet-bewuste) vooringenomenheid, een term die in de lezing van de tweede internationale spreker van de dag, Kriti Sharma, een sleutelrol zal spelen. Lankes stelt: 'For too long librarians saw ourselves as neutral actors. We collected, described, and provided materials believing that these acts were either without bias, or that those biases were controlled.' Maar, zo voegt hij er aan toe, langzaam maar zeker worden bibliotheken zich er bewust van dat ook zij een product zijn van hun cultuur en context of in sterkere mate beïnvloed worden door de dominante stemmen binnen de gemeenschap. Ze zouden zich hier in ieder geval bewust van moeten zijn. 'So as a library we are not asking to be neutral arbiters of data collection and uses. We are seeking to improve society through data and algorithms – that means we have a point of view. We have a definition of what improve means. However, the biases we bring, or more precisely the principles, we bring to the Googles and Facebooks of the world is that of a strong voice that advocates for transparency, privacy, the common good, and a need for a durable memory is important. We recognize that bias exists even if we can't always identify it, and so we require diversity and inclusive voices in our work. In this act we are not simply advocates, we are activists. A missionary corps of professionals equipping our communities to fight for their interests.'

Machine/deep learning omschrijft Lankes als het creëren door AI van algoritmes en procedures zonder enige menselijke tussenkomst. Ook hier ligt het gevaar van '*bias*' op de loer, maar extra problematisch is dat dat plaatsvindt in een '*black box*', waarbij nauwelijks te doorgronden valt hoe de door AI gegenereerde uitkomsten tot stand komen vanwege het gebrek aan transparantie van het proces. Voor bibliotheken ziet Lankes een rol in het '*social*' maken van AI, zodat het in dienst staat van de gemeenschap. 'Our communities, and our society, needs a voice to ensure the data being used is representative of all of a community, not just the dominant voice, or

the most monetizable. Our communities need support, understanding, and organizing to ensure that the true societal costs of A.I. are evaluated, not simply the benefits. (...) Librarians need to become well versed in these technologies, and participate in their development, not simply dismiss them or hamper them. We must not only demonstrate flaws where they exist but be ready to offer up solutions. Solutions grounded in our values and in the communities we serve.'



Kriti Sharma, verbonden aan het Britse techbedrijf **Sage Group** als '*vice president of artificial intelligence and ethics*', bouwde op haar vijftiende haar eerste robot. Op dertigjarige leeftijd geeft ze advies aan het Britse Parlement, de UN Tech Labs, heeft ze meerdere TEDx talks op haar naam staan en staat ze in de Forbes 30 under 30-lijst. Sharma gaat in op de ethische aspecten van AI en zoals gezegd is de term '*bias*' (of '*Algorithmic bias*') een sleutelbegrip in haar lezing. Sharma neemt ons in een bevlogen en met humor doorspekt betoog mee langs een aantal voorbeelden van beslissingen gemaakt op basis van AI die feitelijk stoelen in '*bias*', hetzij raciaal (cq racistisch), seksistisch of anderszins van aard. In reactie daarop pleit Sharma voor een meer ethisch verantwoorde inzet van AI en algoritmes. AI kan, rekening houdend met alle '*bias*'kwesties die ermee gepaard gaan, volgens Sharma zo ook ingezet worden voor het realiseren van een betere, meer sociale, wereld. Als het om bibliotheken gaat kan AI volgens Sharma ingezet worden om de toegang tot informatie eerlijker en breder te maken, voor grotere groepen in de samenleving, maar wel op een ethisch verantwoorde en sociale manier zodat het iets bijdraagt aan de gemeenschap.

Sharma formuleert vijf uitgangspunten voor 'Social AI':

- 'AI should reflect the diversity of its users'
- 'AI has to be held to account - and so must its users (accountability).'¹ Er zijn processen nodig ter controle/ backup
- Transparantie: 'Reward AI for sharing its workings'
- 'AI should level the playing field, should be developed for the common good and the benefit of humanity.'
- 'AI will replace, but it must also create.'



Ook de laatste spreker van de dag, **Jan Willem van Wessel** van de KB, ging in zijn **lezing** (pdf) in op de ethische kanten van AI. Maar eerst ging hij in hoog tempo langs verschillende invalshoeken met betrekking tot AI. Als het gaat om de klant kan AI bijvoorbeeld veel betekenen op het gebied van '*search / discovery*', '*recommender*'-systemen of spraakinterfaces. Voor de bibliotheek zelf en haar werkprocessen valt te denken aan het automatisch toekennen van metadata en synopsissen of het verbeteren van OCR-technieken. Trends op dit gebied zijn het automatisch genereren van luisterboeken en automatisch samengestelde boeken cq automatische contentcreatie. Daarnaast zijn bibliothecarissen met hun ervaring op het gebied van curatie bij uitstek geschikt om data te leveren om de AI te 'trainen'. Ook Van Wessel besluit met ethische principes voor het gebruik van AI. De KB heeft er zeven geformuleerd:

- Toegankelijk: zoveel mogelijk informatie beschikbaar
- Inclusief: rekening houden met het risico van bias
- Onder toezicht: altijd onder toezicht van mensen

- Transparant: 'trainings'data voor het algoritme moeten bekend zijn
- Neutraal: niet willen manipuleren van gebruikers of voor eigen gewin
- Veilig: onder andere aspecten als privacy e.d.
- Compliant: voldoen aan lokale wet- en regelgeving

Wat specialistischer lezingen verzorgden [Thomas Smits](#), oud researcher-in-residence van de KB, en [Jacco van Ossenbruggen](#), onderzoeker bij het Centrum Wiskunde & Informatica (CWI) en tevens verbonden aan de Vrije Universiteit (VU), met presentaties over respectievelijk beeldherkenning middels AI en de zogeheten 'Cultural AI', het gebruik van AI om de menselijke cultuur te bestuderen ('digital humanities'). Smits maakt in zijn [lezing](#) (pdf) duidelijk hoe AI gebruikt kan worden bij het herkennen en classificeren van beeld en dat de technologie door (wetenschappelijke) bibliotheken onder andere ingezet kan worden voor het automatisch toevoegen van metadata aan beeldcollecties, het verbinden van beeldinformatie met informatie op het internet of voor het bieden van nieuwe toegangen tot visuele collecties voor de gebruikers, Van Ossenbruggen legt in zijn [lezing](#) (pdf) uit dat het bij 'Cultural AI' enerzijds gaat om machines te leren om menselijke normen en waarden beter te begrijpen en anderzijds om sociaalwetenschappelijke kennis (de humaniora) te gebruiken om AI beter te maken. Daartoe is het nodig dat wij beter begrijpen hoe AI werkt. Ook in dit verband speelt het begrip 'bias' - en alle nuances en vragen die erbij komen kijken - weer een rol. Van Ossenbruggen pleit tenslotte voor de oprichting van een 'nationaal Cultural AI Lab' om multidisciplinair verbindend onderzoek te doen. 'Dit is niet iets wat de beta's in hun eentje kunnen, dit is niet iets wat alfa's of gamma's in hun eentje kunnen, maar we moeten het onderwerp gezamenlijk oppakken. We moeten gezamenlijk niet alleen technologie ontwikkelen die geoptimaliseerd is voor onze culturele normen en waarden, maar die zich impliciet of expliciet ook bewust is van onze culturele normen en waarden. De publieke waarden van de openbare bibliotheek sluiten hier prima bij aan.'

Conclusie van de KB na de bijeenkomst: het onderwerp AI is belangrijk genoeg om nog vaker aandacht aan te besteden, reden waarom er vervolgsessies georganiseerd zullen gaan worden waarbij meer verdieping gezocht zal gaan worden rond verschillende specialismen en thema's.

In het kader van [Bieb-a-palooza](#) zijn na afloop van de bijeenkomst met de sprekers enkele circa tien minuten durende interviews gemaakt.

Registraties van de lezingen zijn te vinden op YouTube in de reeks [KB-colleges](#). De presentaties zelf zijn te vinden op de [website van de KB](#).

Tekst: Bart Janssen

Foto's: Jacqueline van der Kort, Beeldstudio KB

Reacties op dit artikel (0)

Er zijn nog geen reacties.