

Met wat voor kenmerkende verschillen tussen een raptekst en 'gewonere' teksten moet een automatische tekstgenerator rekening houden?

'Rap onderscheidt zich in veel opzichten van gewone tekst, zowel wat betreft inhoudelijke als stilistische aspecten. De stilistische eigenschappen, zoals rijm, metrum en flow, springen wellicht nog het meest in het oog. Dit zijn allemaal kenmerken waar "gewone" tekstgenerators niet direct mee om kunnen gaan, en waar we dus nieuwe oplossingen voor hebben moeten bedenken.'

Worden dergelijke kenmerken vooraf als regels in jullie programma ingevoerd, of leert het programma het helemaal op basis van voorbeelden?

'De meeste kenmerken worden gegenereerd op basis van machine learning. Daarbij moet wel aangetekend worden dat we niet zomaar een grote hoeveelheid teksten "voeren" aan een algoritme in de hoop dat het zelf kenmerken als rijm en ritme zal herkennen en vervolgens ook zal toepassen. We hebben speciale modules voor specifieke kenmerken gemaakt, die – bijvoorbeeld in het geval van rijm – de regel hebben dat er gerijmd moet worden, maar waarbij niet bepaald wordt op welk woord. Ons systeem past een vrij losse definitie van rijm toe, die zegt dat twee woorden rijmen als hun bekleemtoonde klinkers overeenkomen.'

Begin het programma met wat willekeurige woorden, of zijn er bepaalde aangeleverde zinnen of thema's waarop het voortbordurt?

'We hebben een reeks verschillende systemen geïmplementeerd die op verschillende taalniveaus tekst produceren. Zo hebben we een systeem gemaakt dat op letterniveau teksten maakt, maar ook systemen die op lettergreep en woordniveau opereren. Voor elk van deze systemen geldt dat ze beginnen met het selecteren van een willekeurige letter, lettergreep, of woord, afhankelijk van het gekozen systeem.'

Een rappende tekstgenerator

Deze zomer deden de makers van software die automatisch rapteksten genereert, op het Lowlands-festival een soort Turing-test: herkenden bezoekers of een raptekst door een computer of een mens gemaakt was? Op de VOGIN-IP-lezing komt Folgert Karsdorp (Meertensinstituut) daarover vertellen. IP stelde hem vooraf al wat vragen.

Hoeveel van de door jullie programma gegenereerde teksten werden bij het experiment tijdens Lowlands door de proefpersonen als 'echt' beoordeeld?

'Gemiddeld konden de proefpersonen in zestig procent van de gevallen het onderscheid maken tussen echte en gegenereerde teksten. De deelnemers kregen ongeveer evenveel authenticiteit als gegenereerde teksten te zien, dus in veertig procent van de gevallen werden gegenereerde teksten "foutief" als origineel beschouwd.'

Is er van jullie teksten ooit eentje echt door een rapper uitgevoerd?

'Nog niet!'

Wat hebben jullie van dit experiment geleerd?

'Allereerst hebben we geleerd dat het genereren van rapteksten een grote uitdaging is. Gegeven al de verschillende eigenschappen en kenmerken van raps, is het lang niet zo eenvoudig om overtuigende teksten te schrijven. Dat was op zichzelf opvallend, omdat in eerdere experimenten met literair materiaal relatief eenvoudige tekstgenerators vrij makkelijk overtuigende literaire zinnen konden maken. Diezelfde generators waren echter ontoereikend voor de rapteksten. Daarnaast heeft het experiment ons inzicht gegeven in de perceptie van mensen in gegenereerde teksten. Wat zijn de verwachtingen van mensen als ze een gegenereerde tekst zien? Een opvallende en leuke uitkomst was dat mensen zich blijkbaar niet kunnen voorstellen dat een automatisch systeem schuttingtaal zou gebruiken. Nou, reken maar! Tot slot, ik denk dat het belangrijk is dat er in veel verschillende domeinen (zoals literatuur, muziek, nieuws, enzovoort) geëxperimenteerd wordt met tekstgeneratiesystemen. Op die manier kunnen we beter zicht krijgen op de enorme hoeveelheid uitdagingen die tekstgeneratie nog heeft en ook op de grote complexiteit van creativiteit en taal.'



Lowlands-festival 2018: is die raptekst wel of niet door een mens geschreven?