

## LA CONVERSATION

### L'intelligence artificielle peut-elle créer une poésie d'un genre nouveau ?

18 novembre 2021

Thierry Poibeau, DR CNRS, École normale supérieure (ENS) – PSL



La boîte à poésie. RAFFARD ROUSSEL

Il existe des centaines de programmes de génération de poésie sur Internet – c'est-à-dire des systèmes capables de produire de la poésie automatiquement –, mais à quoi peuvent-ils servir ? Ces programmes ont-ils un intérêt, au-delà de celui de satisfaire leur concepteur ?

On a beaucoup entendu parler de GPT2 ou GPT3, ces énormes programmes informatiques capables de produire des textes très réalistes, et même de la poésie. GPT2 et GPT3 sont en fait des « modèles », des espèces de bases de connaissances, alimentés par des milliards de phrases et de textes glanés sur Internet, et « digérés » afin de pouvoir produire des textes nouveaux, inspirés des textes anciens, mais en même temps très différents.

#### Ajouter des contraintes

Pour produire de la poésie, il « suffit » d'y ajouter des contraintes : surveiller les rimes et la longueur des vers, respecter la structure globale, l'absence de répétition en position rimée, etc. Les systèmes de génération automatique connaissent un certain succès (on en trouve un nombre phénoménal sur Internet) car la tâche est amusante, ludique, mais aussi complexe si on veut produire des textes avec du sens (et encore plus si on veut contrôler ce qui est dit).

On a donc ici un cadre idéal pour expérimenter (souvent hors contrainte de financement) : la génération de poésie est souvent un loisir et un passe-temps, chez le chercheur comme chez l'amateur éclairé.

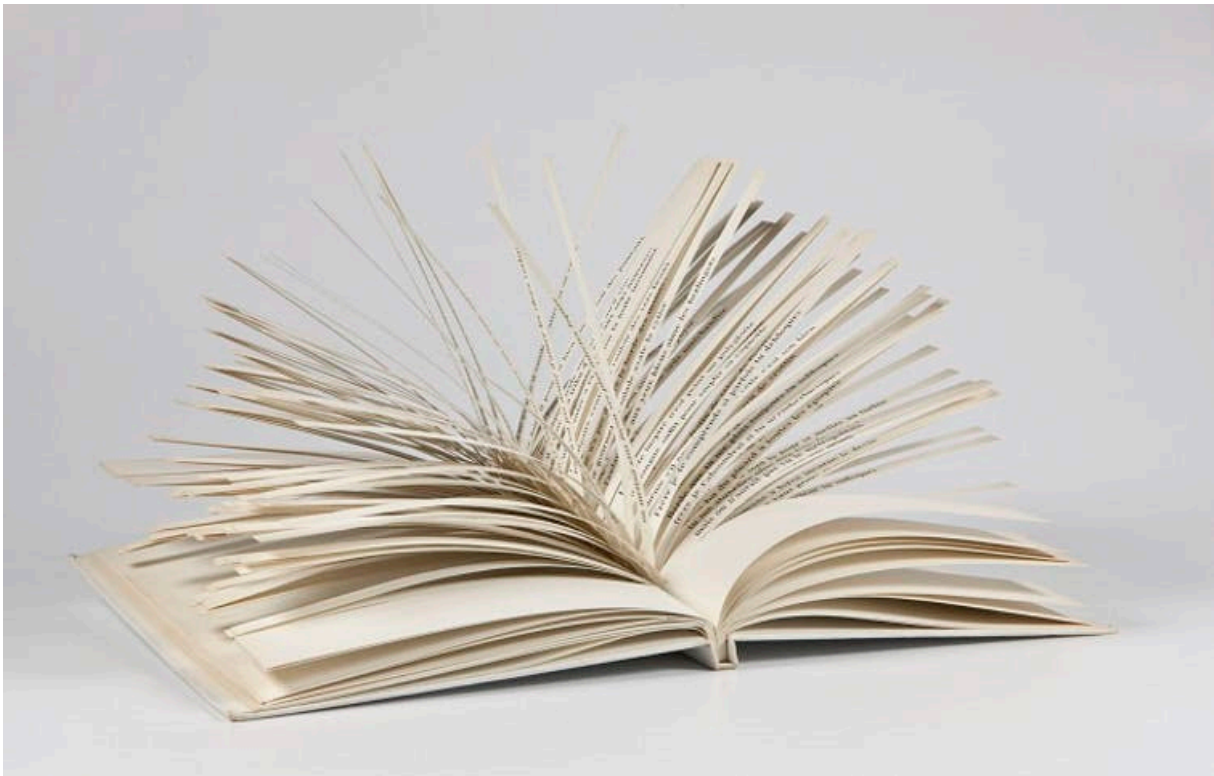
Une question se pose toutefois. Ces systèmes sont-ils dignes d'attention ?

Sur le plan littéraire, la plupart des systèmes sont encore, il faut bien l'avouer, assez rudimentaires et ont du mal à rivaliser avec du Baudelaire ou du Rimbaud. Les plus avancés sont toutefois bluffants, et c'est avant tout la base d'entraînement qui joue un rôle crucial (c'est-à-dire l'ensemble des textes qui ont permis au système d'apprendre. On peut en effet fournir à un système un jeu de données réduit, mais spécialisé (des œuvres de poètes du XIX<sup>e</sup> siècle par exemple) pour « spécialiser » un système, l'adapter à moindre coût. On peut alors obtenir des systèmes qui rédigent des paragraphes à la manière de Balzac, ou de la poésie à la manière de Baudelaire.

Il faut toutefois noter que les résultats apparaissant dans la presse (qu'il s'agisse de génération de prose ou de poésie) sont souvent le fruit de multiples essais, voire le fruit d'un travail de postédition de la part du journaliste.

### **Le projet Oupoco**

Le projet [Oupoco](#) (Ouvroir de poésie combinatoire) que nous avons développé avec une équipe du [laboratoire LATTICE](#), avait un but plus modeste. À l'image de l'expérience de [Queneau](#) dans *Cent mille milliards de poèmes*, notre ambition première était de produire des milliards de poèmes simplement en recombinant des vers issus d'un corpus poétique français représentatif.



*Cent mille milliards de poèmes*, Raymond Queneau, 1961. [Musée Médard](#)

À cette fin, nous avons assemblé une base de plus de 4000 sonnets d'auteurs allant du début du 19<sup>e</sup> jusqu'au début du XX<sup>e</sup> siècle. Alors que tous les vers de Queneau riment ensemble, il nous a fallu déterminer automatiquement la rime de chaque vers afin de pouvoir produire de la poésie avec rimes. Le projet était donc dès l'origine davantage un projet d'analyse qu'un projet de génération.

Ce projet peut paraître dommageable, en ce qu'il ferait passer la poésie pour du « n'importe quoi ». Mais le but est évidemment bien différent. Des expériences concrètes et la rencontre avec le public nous ont montré que cette crainte est en grande partie injustifiée. Le public (jeunes et vieux, femmes et hommes) est amusé, intrigué, veut en savoir plus. Un public d'ordinaire peu attiré par la poésie s'intéresse à ce qui est produit. Le public n'est pas naïf, même quand il s'agit d'enfants : il voit bien le caractère fabriqué, étrange et ludique de l'affaire. Il sait que derrière ce qui est produit se cachent d'autres textes et l'incongruité d'un vers hors norme pousse souvent à aller voir le contexte original, c'est-à-dire le poème d'origine.

Les vendangeurs lassés ayant rompu leurs lignes,  
La trirème d'argent blanchit le fleuve noir  
Sachant que notre cœur s'emporte et se résigne,  
**Étrangler** tes sanglots dans le fatal mouchoir.

C'est par un ciel pareil, tout blanc du vol des cygnes,  
**L'éclair laisse, en fuyant**, l'horizon triste et noir.  
Je hais le mouvement qui déplace les lignes,  
Et son verger fleuri n'est qu'un vaste encensoir.

Mais, le corps étendu, n'oublions pas que l'âme,  
**Ils** me rendent aveugle au jour qui te proclame  
Si proche et déjà loin de celui que j'aimais.

Haute théologie et solide morale,  
Et maintenant j'habite, hélas! et pour jamais,  
**S'envole, tinte et meurt** dans le ciel rose et pâle.

Un exemple de poésie (Oupoco) – en rouge, les passages présentant un problème de cohérence. Oupoco, Fourni par l'auteur

Le générateur de poésie, avec les dispositifs de diffusion qui vont avec (comme la [Boîte à poésie](#), une œuvre d'art conçue par [l'Atelier Raffard-Roussel](#) et permettant d'obtenir un objet portatif intégrant le générateur de poésie d'Oupoco), permettent à un public large de renouer avec la poésie, alors que c'est une forme souvent délaissée même par les lecteurs régulier.

Quant aux expériences en génération pure (où la poésie produite n'est pas composée à partir de vers préexistants, mais est réellement conçue par ordinateur), elles amènent à réfléchir à d'autres aspects. Sur le texte lui-même : quelle est la richesse du texte produit ? Qu'est-ce qui fait la valeur d'un texte poétique ? Si on est dans le cadre d'une génération « à la manière de » (de façon similaire à la production de musique « à la manière de »), on peut s'interroger sur la valeur du résultat, sur les caractéristiques d'un auteur, sur ce qui fait le style d'un auteur, finalement.

### Différents niveaux de créativité

Ces questions amènent enfin à s'interroger sur la notion de créativité elle-même. [Margaret Boden](#), une informaticienne anglaise ayant développé une théorie sur la question, distingue trois niveaux de créativité, chez les humains comme chez les ordinateurs : la « créativité exploratrice », qui consiste juste à étendre un peu ce qui existe déjà (écrire un poème à la façon

de Hugo) ; la « créativité combinatoire », qui consiste à combiner de façon originale des éléments existants autour de nous, mais de nature éloignée (les travaux de [l'Oulipo](#), mêlant littérature et contraintes mathématiques sont probablement de cet ordre). La troisième forme de créativité, qualifiée de « transformationnelle », est d'une autre nature, elle change radicalement la façon de voir la réalité et produit généralement toute une nouvelle lignée d'œuvres. Margaret Boden parle de l'invention du cubisme par Picasso ; on peut penser à l'abandon des codes du roman dans les années 1950, autour du nouveau roman, mais la notion de rupture en littérature serait un concept à discuter en lui-même.

Le système Oupoco, recombinaison juste des vers existants, est indéniablement de nature exploratrice, même si cette exploration est fondée sur la combinatoire. Le Graal de la créativité par ordinateur serait d'atteindre la créativité transformationnelle, au sens que Boden accorde à ce mot. Un ordinateur serait-il capable d'atteindre ce niveau ? On peut en douter, car ce niveau implique une certaine conscience de soi, une prise de recul par rapport au réel, pour imaginer des mécanismes complètement nouveaux. L'apprentissage artificiel, à la source de la plupart des développements récents et médiatiques en matière d'IA (intelligence artificielle), est très bon pour généraliser et recombinaison les milliards de données reçues en entrée, mais est incapable de « faire un pas de côté », pour réellement transformer le réel.

Notons enfin que les humains apprennent aussi à partir de stimuli et par imitation. La nature et la réalité de la créativité transformationnelle n'est pas complètement prouvée. Peut-être qu'à partir des milliards de perceptions reçues au cours de sa vie l'homme est capable de recombinaison de manière suffisamment libre pour donner l'impression d'une créativité transformationnelle. On est alors au cœur de la cognition !

---

*Ont participé à Oupoco les personnes suivantes : Claude Grunspan, Mylène Maignant, Clément Plançq, Frédérique Mélanie Becquet, Marco Naguib, Yann Raphalen, Mathilde Saurat, ainsi que l'Atelier Raffard-Roussel.*