

Grands modèles de langue (GPT, BERT) : enjeux pour la recherche et l'enseignement

Julien Boelaert (MCF - Université de Lille – CERAPS)

Organisé par la SFR GANESHs

2 décembre 2025, 9h30 à 12h30, Université Jean Monnet de Saint-Étienne

Le 2 décembre 2025, la structure fédérative de recherche GANESHs avait invité Julien Boelaert, maître de conférences en sciences politiques à l'université de Lille, pour une conférence traitant des nouveaux enjeux des grands modèles de langues en intelligence artificielle (IA) pour la recherche et l'enseignement. Julien Boelaert précise d'emblée préférer le terme « grands modèles de langues » à celui « d'intelligences artificielles », pour des questions de précision. Il a par exemple rappelé que les intelligences artificielles existent depuis plusieurs décennies (ne serait-ce que pour jouer à des jeux en ligne comme le backgammon ou les échecs). Il y a, dès lors, un enjeu de définition au cœur des IA. Les grands modèles de langues (ou *large language models* – LLMs) sont en réalité des modèles mathématiques qui associent une probabilité à un texte. Si un LLM est confronté à la phrase « longtemps le soir je me suis couché de bonne... », par exemple, il doit pouvoir prédire avec une forte probabilité que le prochain mot est « heure », pour compléter la citation de Proust. Ces modèles sont calculés à partir de très grandes bases de données et sont applicables aux textes autant qu'aux sons et aux images. Un nouveau modèle de LLM a émergé en 2017, proposant un saut qualitatif, avec les Transformers : pour Julien Boelaert, ils permettent de « s'augmenter en tant que chercheur », dans des tâches comme la classification ou encore la recherche et l'extraction d'informations. Ils permettent également une plus grande exhaustivité, ainsi qu'une remise en question de l'ancienne opposition entre internalisme (analyser le contenu du texte) et externalisme (en analyser les métadonnées : auteur, date...). Après une introduction et de nombreux rappels bienvenus sur les définitions, Julien Boelaert propose d'articuler sa réflexion en deux parties, en traitant séparément les LLMs avant et après l'apparition des Transformers.

Avant les Transformers, chaque texte est considéré, dans le cadre des LLMs, comme un ensemble de vecteurs. Là où les Transformers traitent les textes par blocs, les autres LLMs traitent les mots séparément : Michael Jackson, par exemple, est la somme du vecteur pour Michael et de celui pour Jackson, ce qui ne restitue pas toute la richesse sémantique de la

notion. Le mot n'est donc pas réellement caractérisé par ce qui l'entoure. Ce système, qui est également celui des matrices documents-termes (ou DTM), montre deux autres limites principales. La première, que Julien Boelaert qualifie de « sac de mots » : avec le système pré-Transformers, l'ordre des mots n'a que peu d'importance, ce qui peut donner lieu à des contresens (« le chat mange la souris » n'est pas synonyme de « la souris mange le chat ») ou à des erreurs liées à la polysémie des termes (on pense notamment à « chat », qui peut faire référence à l'animal comme au système de discussion en ligne). Le second écueil principal repose sur le fait que les termes nécessitent, pour ce système, un pré-traitement, pour conserver la casse et les majuscules (« la maison blanche » n'est pas la même chose que « la Maison Blanche »). Julien Boelaert conclut cette partie en rappelant le rôle prépondérant de l'École Française de lexicométrie dans le développement des LLMs, qui n'est donc pas seulement le fruit des initiatives américaines, portées par des grandes entreprises comme Google. Cette première partie très éclairante a ainsi permis de retracer les origines des LLMs et de les mettre en lien avec des outils bien connus dont on ignore parfois le fonctionnement. Notons également que le conférencier n'a pas hésité, à plusieurs reprises, à s'interrompre pour répondre aux questions du public, ce qui a permis de clarifier certains points avant de continuer la démonstration.

Dans un second temps, Julien Boelaert a proposé de traiter plus précisément des Transformers, qui sont des modèles de langage sur texte brut, pré-entraînés sur des corpus massifs et qui permettent de traiter des blocs entiers de texte. Ils fonctionnent également avec un mécanisme de requête et de clé, c'est-à-dire que pour chaque terme, on recherche quel autre mot dans le contexte informe le sens du terme initial, proposant donc un enrichissement par le contexte. Ces modèles sont également novateurs en ce qu'ils ne traitent pas réellement des mots mais plutôt des « tokens », c'est-à-dire des parties, plus ou moins longues, de mots (« longtemps », par exemple, est composé des tokens « long » et « temps »). Il existe de nombreux modèles de Transformers, qui ne sont pas tous égaux devant les tâches demandées : GPT, par exemple, est un transformer décodeur particulièrement à même de prédire le prochain token, tandis que BERT est un transformer encodeur non génératif performant pour ajouter des données. Julien Boelaert rappelle également que ces modèles évoluent très vite (dans le cas de GPT, il y a un nouveau modèle par an depuis 2018) et que ces évolutions sont spectaculaires : GPT2, sorti en 2019 seulement un an après GPT1, est dix fois plus performant et propose une génération de texte crédible, par exemple. GPT3, sorti en 2020, est mille fois plus performant que GPT1 et fonctionne en *few-shot learning*, c'est-à-dire presque sans

entraînement, à partir de seulement quelques exemples en contexte. Après ces quelques remarques de définition et de mise en contexte sur les Transformers, Julien Boelaert propose de définir quelques concepts clés importants pour le fonctionnement des LLM (LLM de fondation, LLM *fine tuned*, *instruct*...), ce qui l'amène à évoquer la question des limites de ces grands modèles de langage, notamment à travers la notion de passage à l'échelle (*scaling laws*) : alors que précédemment l'on considérait que la taille du modèle faisait sa qualité, l'on revient à présent sur cette idée en proposant davantage de mieux entraîner les modèles, plutôt que d'en créer de plus gros. Cela fait émerger deux nouvelles questions, que Julien Boelaert soumet au public : que faire lorsque toutes les données du web sont intégrées aux LLMs ? Comment procéder quand une partie du contenu du web est justement produit par les LLMs (à l'heure actuelle, c'est le cas pour environ 30% du nouveau contenu du web) ? Julien Boelaert évoque ensuite un autre problème, lié à l'alignement : aligner, c'est contraindre les réponses d'un LLM, afin de ne pas produire de texte nuisible, selon les principes de *truthfulness*, *helpfulness* et *non-offensiveness*. Pour cela, l'on utilise soit un prompt éliminatoire qui indique au modèle ce qu'il ne doit pas dire ; soit du *fine tuning*, en entraînant le modèle avec des exemples ; soit une censure par modèle rapide (comme c'est le cas pour BERT). Si l'alignement permet de ne pas produire du texte raciste ou homophobe, par exemple, il peut aussi nuire aux usages scientifiques, en omettant les questions sensibles ou les sujets interdits (génocide, manifestations de Tian'anmen...). En somme, les LLMs sont en constante évolution, pour le modèle GPT comme pour les autres, afin de répondre à des demandes particulières.

En conclusion, les usages des LLMs en sciences sociales sont multiples, et il convient de comprendre comment ils fonctionnent afin de pouvoir s'en servir en toute connaissance de cause. En ce sens, la conférence de Julien Boelaert était particulièrement importante et éclairante, pour un public peu averti comme pour un public plus spécialiste. Elle a permis de rappeler les enjeux des LLMs en partant de leur création et de leur fonction première, dans un déroulement qu'il était aisé de suivre. Nous remercions vivement la SFR GANESHs ainsi que Julien Boelaert pour cette intervention très riche.