

Master II Sciences et Technologies du Logiciel (STL)

Rapport de stage de 5 mois

Application d'orientation au sein du Telegram



Encadrants universitaires : ...

Employeur : ...

Encadrant côté employeur : ...

Octobre 2023

Remerciements

Je tiens à remercier tout particulièrement Messieurs ... et ... pour avoir accepté d'encadrer le présent travail.

Mes remerciements s'adressent également à l'association ... pour avoir accepté ma candidature comme stagiaire et m'avoir ainsi permis de poursuivre mon cursus.

Je remercie M. ... pour sa patience et ses conseils avisés tout au long du stage réalisé, ainsi toute son équipe.

Enfin, je remercie ma famille sans le soutien de laquelle cette aventure professionnelle enrichissante n'aurait jamais vu le jour, et la correctrice, Mme ... pour ses qualités professionnelles sans précédent.

Sommaire

Table des matières

Application d'orientation au sein du Telegram.....	1
Remerciements.....	2
Sommaire.....	3
Introduction.....	4
Présentation du problème, l'importance de la question, enjeux.....	4
Le plan du rapport.....	4
Chapitre 1 : Présentation de l'association.....	5
Chapitre 2 : Présentation du Telegram.....	5
Chapitre 3 : Les marqueurs de propagande.....	6
Chapitre 4 : Revue des travaux menés dans la détection des messages de propagande au sein des médias.....	9
Chapitre 5 : Objectifs du stage.....	10
Chapitre 6 : Organisation des tâches.....	11
Chapitre 7 : la solution proposée.....	11
Limites des solutions proposées.....	12
Chapitre 8: Les développements principales.....	13
La vue d'ensemble.....	13
Sélection de chaînes.....	14
Intégration avec Telegram.....	14
Repérage les propriétés d'un message à partir des affirmations.....	14
Repérage les propriété d'un message par rapport aux caractéristiques de la propagande.....	15
Stockage des données.....	16
Conclusion sur le contenu d'une chaîne.....	17
Phase d'ajustement et d'apprentissage.....	17
Chapitre 9 : Perspectives touchant le traitement prochain de la détection des messages de propagande dans les média.....	21
Chapitre 10 : Bilan personnel et professionnel de l'expérience.....	22
Choses apprises durant le stage.....	22
l'Intérêt de mon approche, apport de ma contribution au domaine.....	22
Projets d'avenir.....	22

Introduction

Dans le cadre du Master II en Sciences et Technologies du Logiciel (STL), il est demandé aux étudiants de réaliser un stage en entreprise.

J'ai pu trouver une entreprise répondant à mes désirs de poursuite de ma vie professionnelle, qui a accepté ma candidature : l'association parisienne ... spécialisée dans les projets solidaires d'innovation alternative.

Au sein de cette dernière, j'ai mené un stage de 5 mois (entre le 03/04/2023 et le 12/09/2023).

Présentation du problème, l'importance de la question, enjeux

Au cours du stage, il m'a été donné de travailler sur des questions technologiques associées au thème de l'analyse de la propagande véhiculée dans les médias. Dans ce cadre, j'ai travaillé sur l'application de messagerie Telegram.

Plus précisément, mes principales missions ont consisté à réaliser des analyses numériques de l'information diffusée par ce média et de construire une application aidant les Internautes à s'orienter dans le monde de l'information, et ce, par le biais des diverses caractéristiques de la propagande et par la comparaison des affirmations diffusées.

Le plan du rapport

Afin de donner un aperçu des techniques apprises au cours de ce stage et d'expliquer en quoi ce dernier m'a permis d'acquérir des compétences professionnelles fort utiles et de m'intégrer à un environnement professionnel nouveau, je présenterai, dans un premier chapitre, l'entreprise où s'est déroulé le stage.

Puis, dans un second chapitre, sera abordé le média sur lequel il m'a été donné de travailler : Telegram.

Dans le cadre d'un troisième chapitre, je traiterai brièvement de la question de la propagande.

Le chapitre suivant proposera une revue des travaux menés dans la détection des messages de propagande au sein des médias.

Le chapitre 5 posera les objectifs numériques et professionnels du stage (en termes d'acquisition de compétences).

Tandis que le 6^e chapitre traitera de la question de l'organisation des tâches.

C'est à partir de là que seront étudiées en détails, dans le chapitre 7, les problématiques numériques mises au travail.

Le chapitre 8 abordera la création d'une application visant à analyser la fiabilité d'un média.

De là, le chapitre 9 indiquera les perspectives touchant le traitement attendu de la détection de ces marqueurs.

Enfin, le chapitre 10 dressera un bilan personnel et professionnel de l'expérience professionnelle vécue.

Chapitre 1 : Présentation de l'association ...

... est une association relevant de la loi 1901 basée à Paris, dans le 10^e arrondissement. Elle intervient auprès des populations vulnérables, pour apporter une aide en matière de santé, d'éducation ou encore de sécurité alimentaire. ... participe également à des campagnes de sensibilisation sur différentes problématiques sociales.

Mon stage s'est déroulé dans les locaux de deux partenaires d'..., à savoir l'association ... à Montreuil et l'association ... située dans le 13^e. En précisant que j'étais en télétravail un jour par semaine. Ces jours-là, je restais en communication par messagerie électronique avec l'encadrant et trois informaticiens de l'équipe d'Aeri.

Chapitre 2 : Présentation du Telegram

Au cours de ce stage, je me suis penchée sur la messagerie Telegram.

À noter que cette application est particulièrement importante du fait d'avoir été mise en place en 2013 par les frères Nikolai et Pavel Dourov, opposants au président Vladimir Poutine, mais aussi parce qu'en dépit de procès et de la censure, il a su conserver une certaine liberté de

ton ¹. Telegram fait figure de média d'opposition au gouvernement russe, ainsi que le relatent plusieurs médias français².

Notons aussi qu'initialement, les créateurs de Telegram avaient pour objectif de donner vie à un outil de communication capable de passer outre la surveillance du FSB (le service secret russe chargé des affaires de sécurité intérieure).

Ajoutons que l'application est particulièrement utilisée en Russie, en effet, 41% des Russes utilisent l'application au moins une fois par jour et le secteur Telegram russe a dépassé l'auditoire de WhatsApp en mars 2022 et ceux de Youtube en avril 2023 ^{3 4}.

Chapitre 3 : Les marqueurs de propagande

Selon le dictionnaire de l'académie française, la propagande désigne « l'ensemble des actions et des moyens mis en œuvre pour répandre et faire prévaloir une idée, une opinion, une doctrine ». Utilisé dans un sens péjoratif, ce terme désigne un « système de mobilisation, d'information et de contrôle de la société par un parti ou un État ⁵ ».

Selon J. Giry ⁶ : « Contrairement aux *fake news*, la notion de propagande ne requiert pas de faire nécessairement appel à la fabrication délibérée du faux. La propagande correspond bien davantage à une présentation tendancieuse, partielle et orientée d'éléments pouvant être réels mais qui cherchent à mettre en avant la position, les intérêts ou les préférences idéologiques et politiques de l'émetteur de l'énoncé (Breton, 2000).

Si la propagande (au sens premier) est la base des actions publicitaires menées dans le but de convaincre la population d'acheter leurs produits et services, sa logique n'est pas éloignée de celle utilisée dans le champ politique. En effet, la publicité commerciale et la politique ont en

1 En Russie, Telegram est l'une des rares sources d'information que le pouvoir n'a pas bloquée.

www.20minutes.fr/monde/ukraine/3292711-20220519-guerre-ukraine-telegram-cyber-soldats-amateurs-tendent-rallier-opinion-russe-garanties

2 Le Monde www.lemonde.fr/pixels/article/2022/03/10/telegram-tout-comprendre-a-ce-reseau-social-populaire-en-ukraine-et-en-russie_6116968_4408996.html

France Ouest www.ouest-france.fr/monde/guerre-en-ukraine/comment-fonctionne-telegram-au-coeur-de-la-guerre-de-l-information-entre-la-russie-et-l-ukraine-a91c98a9-04a8-434b-8e13-da4956e51a81

3 <https://www.forbes.ru/tekhnologii/487489-telegram-vpervye-obosel-youtube-po-ezednevnoj-auditorii-v-rossii>

4 <https://inclient.ru/how-many-users-in-telegram/>

5 Dictionnaire de l'académie française www.dictionnaire-academie.fr/article/A9P4598

6 Giry, Julien. « Les fake news comme concept de sciences sociales - Essai de cadrage à partir de notions connexes : rumeurs, théories du complot, propagande et désinformation » www.cairn.info/revue-questions-de-communication-2020-2-page-371.html

commun d'avoir pour but d'inciter les populations à réaliser une action. Toutefois, dans le cadre du présent rapport, c'est uniquement à la propagande politique que je vais m'intéresser.

Afin d'étudier les marqueurs de propagande présentes sur le site Telegram, j'ai listé quelques ses caractéristiques, partiellement repérables par OpenAI, de manière suivante :

- l'émission simultanée ou rapprochée de mêmes affirmations sur différents canaux ;
- les marqueurs lexicaux comme le nombre d' adverbess, des noms, des adjectifs, pronoms, citations, le mot techniques, des nombres, les verbes comme devoir, pouvoir, des mots d'introduction ⁷;
- une scientificité apparente mais sans référencement des sources ;
- la présence d'euphémismes (figure de style consistant à atténuer l'expression d'un fait ou d'une pensée supposé désagréable) venant remplacer l'information par une autre, afin de la décharger du contenu émotionnel ou d'en renforcer la force expressive. À titre d'exemple, l'expression « dommages collatéraux » est couramment utilisée à la place de l'énoncé « victimes civiles ». Je cite aussi un exemple pris dans un œuvre littéraire : Quand en 1971 trois cosmonautes soviétiques étaient morts, l'annonceur des nouvelles à la télévision, à la place de dire - Cher camarades, nous avons eu un malheur, - a annoncé d'un ton, comme toujours, solennel, que le vol du navire spatial avait été terminé avec succès, pendant le vol tels et tels progrès avaient été obtenus, les moteurs de frein s'étaient mis le contact à temps, le navire était entré dans les couches denses de l'atmosphère, avait atterri précisément dans la zone prévue, les cosmonautes avaient été trouvés sur ses places ... - à ce moment l'annonceur change de son ton solennel et triomphal pour ceux solennel et triste et termine la phrase ... - sans les indices de la vie. Tout était parfait, sauf qu'il n'y avait pas de quelques indices.⁸
- le recours à des généralités éblouissantes et à des termes chastes ;
- la justification de faits ou de pensées par des principes éthiques, comme l'amour de la patrie, le désir de paix, la liberté, la gloire, la justice, l'honneur, la pureté de la famille, la solidarité ;
- l'appel aux émotions et non pas à la raison ;

⁷<https://scientificrussia.ru/articles/kak-raspoznat-lozh-v-internete>

⁸ L'Union soviétique anti-soviétique, Vladimir Voïnovitch, 1985

- le recours à la peur. Ainsi, au cours du nazisme, Joseph Goebbels a-t-il exploité l'ouvrage de Theodore Kaufman (homme d'affaires et écrivain de confession juive) intitulé *L'Allemagne doit périr !*⁹ dans lequel il était écrit « peu importe que cet Allemand soit catholique, protestant, nazi, antinazi, communiste, socialiste, démocrate, etc., qu'il soit de Bohême, d'Autriche, de Prusse, d'Alsace-Lorraine, tous doivent y passer ». Et cela, afin de justifier l'idée selon laquelle les Juifs avaient tous pour objectif l'extermination du peuple allemand. Et par là même, de « justifier » la politique d'extermination mise en œuvre ;

- la citation de personnages importants pour soutenir une idée ;

- l'utilisation de moyens de raisonnement reposant sur l'insinuation (comme le fait d'insinuer qu'un mouvement de masse important et irréprouvable est déjà engagé dans la défense d'une idée), ou sur la suggestion (comme dans le fait de suggérer qu'une idée ou une action décrite a été adoptée par un groupe adverse) ;

- le fait de jeter l'anathème sur un individu ou un groupe d'individus ;

- la critique de l'adversaire en tant que personne venant remplacer la critique dialectique de ses arguments ;

- la mise en place de campagnes de diabolisation des teneurs de pensées contraires ;

- la présence de contradictions ;

- le fait que les informations soient partiales et/ou sélectionnées ;

- la diffusion dudit message ou de messages analogues sur une période longue ;

- la présence de répétitions ;

- grand nombre d'informations au profit de l'adversaire, mélangées parfois avec des informations fausses.

Ainsi, ce sont donc la présence de quelques de ces éléments que mon application vérifie de la présence au sein du média Telegram.

9 Kaufman, Theodore (1941). *L'Allemagne doit périr !*, New York.

Chapitre 4 : Revue des travaux menés dans la détection des messages de propagande au sein des médias

Outre la recherche théorique portant sur la propagande (son utilisation, ses effets etc.¹⁰), plusieurs initiatives ont vu le jour touchant la détection des fausses informations en général, et des messages de propagande en particulier, au sein des médias.

Ces initiatives ont notamment la forme de projets informatiques. Plus spécifiquement, concernant la détection automatique de fausses informations et de messages de discours, l'on peut citer :

- ceux liés à l'apprentissage d'une IA, comme le projet développé le Center for Mathematical Sciences Applied to Industry au Brésil (2022) ¹¹ ou le projet de Kung-Hsiang Huang, Kathleen McKeown, Preslav Nakov, Yejin Choi et Heng Ji incorpore des techniques de propagande dans des exemples de texte générés par des machines destinées à alimenter les détecteurs de fausses nouvelles (2022) ¹²;

- le projet développé par l'Atlantic Council's Digital Forensic Research Lab, visant à détecter des communautés de chaînes à une idéologie et à des intérêts communs ou semblables, en se fondant sur le nombre de liens et de messages transférés (2022) ;

- le projet développé par Kung-Hsiang Huang, Kathleen McKeown, Preslav Nakov, Yejin Choi et Heng Ji consistant à incorporer des techniques de propagande dans des exemples de texte générés par des machines destinées à alimenter les détecteurs de fausses nouvelles (2022) ;

- le projet développé par Huan Cao, de l'Institut de technologie de Pékin, capable de détecter 84% de faux *posts* adressés sur Twitter, en employant des techniques comme : l'estimation des messages des utilisateurs se trouvant géographiquement près de l'objet en tant que plus fiables ; l'historique de l'activité du compte ; l'analyse de la réaction des utilisateurs ;

¹⁰ Parmi ces travaux, l'on peut citer, en plus des ressources utilisées dans ce rapport : Colon, David (2021).

Propagande : la manipulation de masse dans le monde contemporain, Flammarion, 448 p.

Chomsky, Noam (2003). De la propagande, éd. 10X18, 443 p.

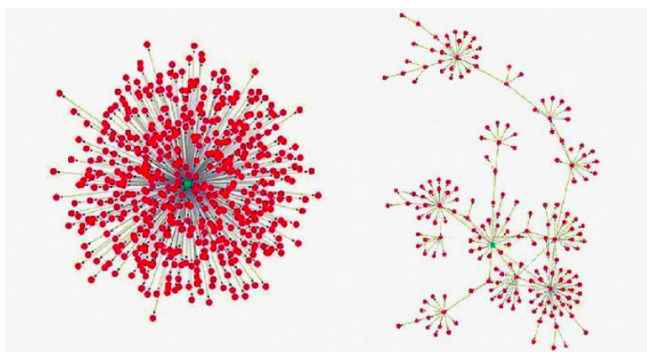
Bernays, Edouard ; Herson-Macarel, Eric et al. (2019). Propaganda. Comment manipuler l'opinion en démocratie, éd. Sixtrid.

Grinshpun, Yana (2023). La fabrique des discours de propagande contemporains : comment et pourquoi ça marche ?, L'Harmattan, 256 p. etc.

¹¹ <https://cemeai.icmc.usp.br/>

¹² <https://arxiv.org/abs/2203.05386>

l'analyse des traces de transmission des informations dans un réseau social, comme il le montre ce diagramme (2016) ¹³



- le projet Botometer ¹⁴ développé depuis 2014, par l'Observatory on Social Media et the Network Science Institute at Indiana University, qui analyse sur les comptes Twitter, les amis des Internauts, l'historique de leur activité, le langage utilisé, les sentiments etc., pour estimer l'authenticité des profils des utilisateurs de Twitter ;

- plusieurs plateformes de l'orientation commerciale, comme auxipress.be (depuis 2003), fever.ai (2018) propose fact-checking, buster.ai (2019) propose les vérifications des informations via une AI.

Chapitre 5 : Objectifs du stage

Les objectifs de cette période de formation étaient au nombre de cinq :

- aborder la vérification des informations contenues dans le Telegram ;
- valoriser mes compétences et en acquérir de nouvelles ;
- développer ma culture générale touchant différents domaines de l'informatique ;
- se lancer dans l'informatique à vocation sociale.

Je précise avoir grandement apprécié que l'association ... fasse montre d'une grande bienveillance à mon égard en intégrant, aux objectifs fixés, des buts concernant mon devenir professionnel.

¹³ https://hightech.fm/2018/08/28/fakenews?is_ajax=1&ysclid=ln2wvj9vsp325940854

¹⁴ <https://botometer.osome.iu.edu/faq#how-does-it-works>

Chapitre 6 : Organisation des tâches

Plusieurs tâches ou missions m'ont été confiées au cours du stage. Cependant, j'ignorais quel serait le temps nécessaire pour réaliser chacune d'elles ni quelles tâches pouvaient réussir. De ce fait, je devais explorer les outils qu'il était possible d'utiliser.

Toutefois, parmi les propositions touchant les missions que j'aurais à réaliser, certaines ont finalement été écartées car jugées inutiles ou par manque de temps. Ces missions ainsi annulées consistaient en :

- l'étude approfondie des réseaux de neurones artificiels ;
- l'extraction des idées contenues dans les messages ;
- la création de schémas efficaces pour le stockage et l'analyse ;
- l'utilisation de la file des tâches ;
- l'utilisation des flux d'événements ;
- le traitement du fil de l'actualité des médias ;
- l'utilisation d'Elasticsearch ;
- l'extraction de l'idée d'un message par deux ou trois AI.

J'ajouterai qu'aucune méthodologie spécifique, comme Agile or Scrum, n'a été prévue.

En revanche, en plus des missions confiées déjà mentionnées, j'ai accompli plusieurs tâches supplémentaires comme le fait de participer toute une semaine durant à l'atelier de codage Codingame pour enfants, et de terminer 6 projets en C pour l'École 42.

Chapitre 7 : la solution proposée

La solution en question repose sur deux piliers : l'analyse du contenu sémantique des messages (le sens et les connotations associées à chaque affirmation), mais aussi sur leurs aspects syntaxiques (la structure et la construction des phrases), les deux via OpenAI.

Grâce à cette double analyse, l'application est en mesure de détecter des incohérences, des exagérations ou d'autres indices qui pourraient trahir la présence de fausses informations.

Limites des solutions proposées

Sur le plan technologique, j'ai pu noter que GPT-3 traite environ 10 messages par minute, et GPT-4 traite 5 messages pas minute. Ce qui est un temps relativement long. L'on peut penser que, par la suite, ce chiffre augmentera en conséquence.

Par ailleurs, l'analyse réalisée par OpenAI comporte des erreurs du sens, d'autant plus nombreuses lorsque les textes analysés ont été rédigés dans une autre langue que l'anglais. Sur ce plan aussi, l'on peut estimer qu'avec le développement des technologies, la qualité d'analyse pourra s'améliorer prochainement.

J'ai noté aussi la présence quelques connexions échouées à API OpenAI.

En outre, au moment du choix des caractéristiques de la propagande, tout comme au moment de l'apprentissage, l'application repose sur l'opinion subjective de l'humaine.

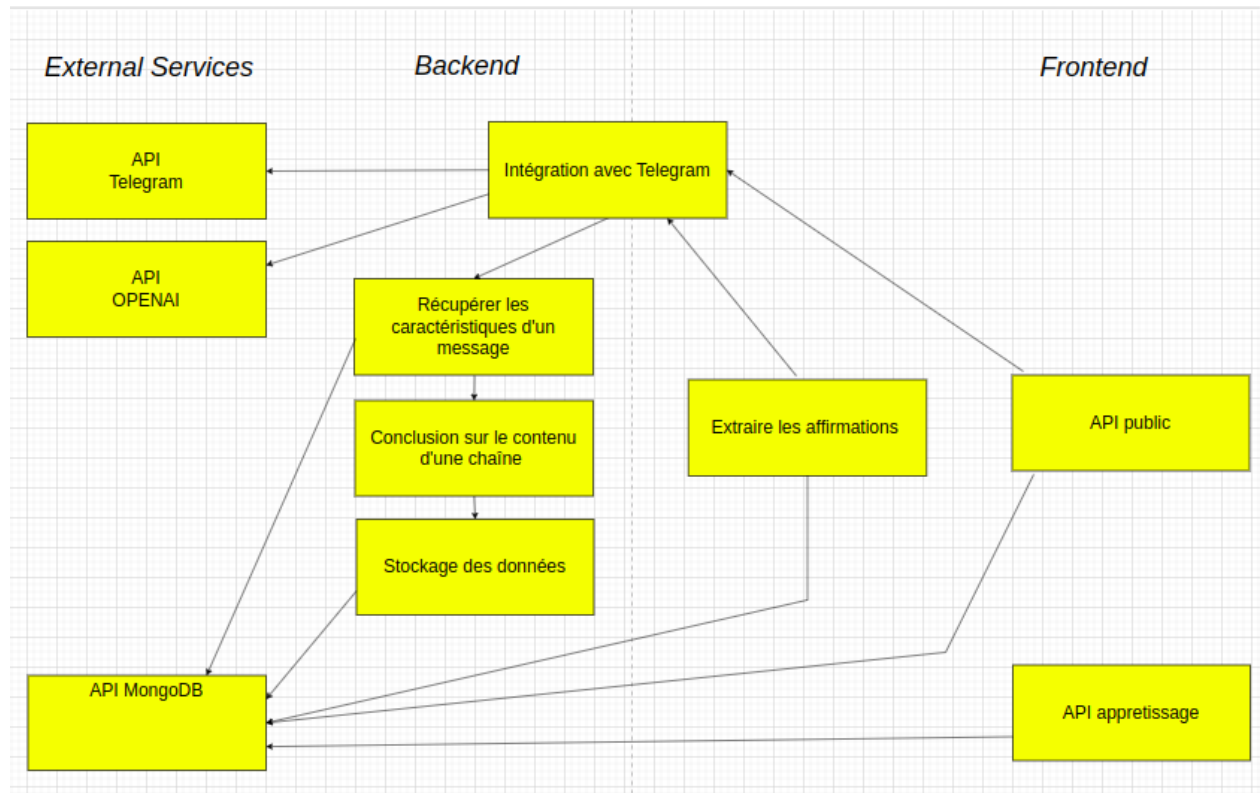
L'analyse est dépendant de la qualité et de l'exhaustivité des messages.

Sur un plan conceptuel, j'estime que l'analyse de marqueurs de propagande ne peut qu'aller de paire avec une étude des causes profondes de l'utilisation de cette dernière. Autrement dit, afin de mener une étude pleinement pertinente sur cette question, il me paraît logique de mener les recherches dans le but de comprendre pourquoi certains individus choisissent l'outil de la propagande et pourquoi d'autres, de nature manipulable, y sont sensibles.

En dernier lieu, en supposant qu'une application efficace existe, les personnes malveillantes ajusteraient leurs messages en les faisant passer par cette même application.

Chapitre 8: Les développements principales

La vue d'ensemble



Pour détecter certains marqueurs de propagande en temps réel, l'application repose sur quatre serveurs :

Le premier, codé en python, récupère un nombre défini des derniers messages du Telegram et fait analyser par OpenAI par rapport aux caractéristiques de la propagande.

Le deuxième, en Python, écoute quelques chaînes Telegram en temps réel également les fait analyser par OpenIA, mais d'une autre façon : cette fois-ci extraire les affirmations à partir du sens des messages, et les cartographier dans un graphe de relations selon la proximité du sens.

Le troisième serveur en node.js est dédié à l'apprentissage de la première approche, et le quatrième en node.js - à l'apprentissage du deuxième approche.

En précisant que Python possède beaucoup de bibliothèques liées à la data science comme NetworkX, Telethon, Pymongo, OpenAI, Matplotlib et celle permettant accéder aux API (OpenAI, Telegram et MongoDB) et ce, de manière asynchrone.

J'ai utilisé VSCode en tant qu'environnement de développement et des outils complémentaires comme Github, Draw.io, Google sheets et Google doc.

Le projet est divisé en plusieurs étapes clés, détaillées ci-dessous.

Sélection de chaînes

Sélection manuelle des chaînes qui seront utilisées dans l'analyse

Intégration avec Telegram

Établissement d'une interface API avec Telegram et collecte de données – en notant qu'en dehors du message lui-même, l'API permet de récupérer les informations comme si le message avait été épinglé, transféré, édité, envoyé via bot, etc.

Repérage les propriétés d'un message à partir des affirmations

L'application identifie et analyse la propagation d'affirmations provenant de canaux Telegram différents dans le but de détecter des schémas de propagande. Chaque message reçu sur un canal est analysé pour en extraire des affirmations, chacune affirmation soit vraie soit fausse. Ces affirmations sont stockées dans la base de données avec un horodatage.

Lorsqu'un nouveau message est reçu, son affirmation est comparée à celles précédemment stockées. Si l'affirmation est identique, alors une connexion dans un graphe est renforcée, non pas simplement de +1, mais de $+1 + P$ où P est une pondération basée sur la proximité temporelle des messages. Si l'affirmation est identique mais présente une valeur de vérité contraire, alors la connexion dans le graphe est réduite de 1.

La pondération P en fonction de la proximité temporelle des messages se calcule par

Une proximité des affirmations, calculé par de l'approche d'affirmations, ne signifie pas nécessairement une intention propagandiste.

L'objectif est d'identifier les schémas et les clusters d'affirmations du sens proche. Une concentration élevée de messages similaires avec des horodatages rapprochés pourrait indiquer une campagne de propagande coordonnée.

Repérage les propriété d'un message par rapport aux caractéristiques de la propagande

Pour cette étape, j'ai utilisé OpenAI GPT-4, car il est réputé pour être l'une des meilleures AI. J'ai expérimenté avec le paramètre température, ce qui n'a pas changé le résultat..

Un exemple d'une requête :

Analyze the following message and provide its characteristics. Reply by numbers, one number for each criterion and nothing else.

<message>

1. positive author attitude: {if true reply -1, if false reply 0, if neutral reply 0}
2. the message is advantageous for the actual political power: {if true reply -1, if false reply 0, if neutral reply 0}
3. exaggerations: {if true reply -1, if false reply 0, if neutral reply 0}
4. lack of object concordance: {if true reply -2, if false reply 0, if neutral reply 0}
5. appeal to the intellect: {if true reply 1, if false reply 0, if neutral reply 0}
6. appeal to emotions: {if true reply -1, if false reply 0, if neutral reply 0}
7. appeal to virtues: {if true reply -1, if false reply 0, if neutral reply 0}
8. appeal to ethics: {if true reply -1, if false reply 0, if neutral reply 0}
9. appeal to blinding generalities as fatherland, peace, glory, justice, honor: {if true reply -1, if false reply 0, if neutral reply 0}
10. appeal to fear: {if true reply -1, if false reply 0, if neutral reply 0}
11. appeal to the threat from other countries: {if true reply -1, if false reply 0, if neutral reply 0}
12. appeal to hatred: {if true reply -1, if false reply 0, if neutral reply 0}
13. appeal to authority: {if true reply -1, if false reply 0, if neutral reply 0}
14. appeal to human rights: {if true reply 1, if false reply 0, if neutral reply 0}
15. criticism of the actual political power: {if true reply 1, if false reply 0, if neutral reply 0}
16. generalizations: {if true reply -1, if false reply 0, if neutral reply 0}
17. seemingly scientific appearance, but only seemingly: {if true reply -1, if false reply 0, if neutral reply 0}
18. real scientific proofs of the information: {if true reply 1, if false reply 0, if neutral reply 0}
19. real proofs of the information: {if true reply 1, if false reply 0, if neutral reply 0}

- 20. euphemismes replacing offensive expressions: {if true reply -1, if false reply 0, if neutral reply 0}
- 21. euphemismes replacing war or victims: {if true reply -1, if false reply 0, if neutral reply 0}
- 22. formal register of the texte: {if true reply -1, if false reply 0, if neutral reply 0}
- 23. beneficiaries of the information regarding to politics: {if true reply -1, if false reply 0, if neutral reply 0}
- 24. flattery to the reader: {if true reply -1, if false reply 0, if neutral reply 0}
- 25. dilemma between only two choice: {if true reply -1, if false reply 0, if neutral reply 0}
- 26. criticisme of a person instead of criticism of his arguments: {if true reply -1, if false reply 0, if neutral reply 0}
- 27. underestimation of the proofs of the opponents: {if true reply -1, if false reply 0, if neutral reply 0}

Je note qu'une requête déposée sur OpenAI du type « Ce message possède-il les caractéristiques d'une propagande ? » n'a pas donné de réponse utile ;

Stockage des données

À ce sujet, sachant qu'un canal d'informations poste en moyenne un message toutes les 5 minutes, la base de données doit disposer d'assez d'espace de stockage.

J'ai stocké les messages sur MongoDB, dont le format BSON (JSON binaire) permet des schémas flexibles et dynamiques.

D'autres bases de données, comme CouchDB, offrent également un stockage orienté document. Toutefois, le BSON de MongoDB est souvent considéré comme étant plus performant en termes d'accès et de recherche. Contrairement à des bases de données comme Redis ou certaines bases de données de valeurs-clés, MongoDB offre un langage de requête plus élaboré pour les opérations CRUD, la recherche de texte et l'agrégation.

Conclusion sur le contenu d'une chaîne

Concernant le calcul du score de fiabilité d'un message, il correspond à la somme des scores attribués à chacune des caractéristiques recherchées.

Le score d'une chaîne est la somme des scores de ses messages, normalisé pour être compris entre 0 et 100.

Les tests réalisés montrent un écart compris entre 3 et 8 points de score moyen entre les chaînes considérées, selon ma connaissance des médias russes, comme « honnêtes » et celles considérées comme « malhonnêtes ».

Phase d'ajustement et d'apprentissage

Les ajustements du système peuvent être divisés selon 4 niveaux :

- Niveau 1 : les paramètres du programme. Pour ce faire, j'analyse par exemple deux groupes de médias – ceux « honnêtes » et ceux « malhonnêtes ». Puis, en consultant les logs, je peux ajuster les caractéristiques examinées ainsi que leurs poids. En précisant que les messages des chaînes changent de test en test en temps réel.

Exemple d'un log :

```
SET 1 ***** 80
@teamnavalny          81/100
@khodorkovski          82/100
@stalin_gulag          79/100
@viktor_shenderovich_izbrannoe 79/100
@nevzorovtv            76/100
@cherniy_kvartal       81/100
@eschulmann            83/100

  1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27
-----
 0.0 -0.1 -0.0 -0.0  0.1 -0.1 -0.0 -0.0 -0.0 -0.0 -0.1 -0.1 -0.1  0.2  0.2 -0.0 -0.1  0.0  0.0 -0.0 -0.0 -0.1 -0.2 -0.0 -0.0 -0.1 -0.0

SET 2 ***** 75
@RKadyrov_95           68/100
@MariaVladimirovnaZakharova 74/100
@SolovievVRNews        80/100
@medvedev_telegram     70/100
@vv_volodin            81/100

  1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27
-----
 0.0 -0.3 -0.1 -0.1  0.0 -0.2 -0.1 -0.1 -0.2 -0.0 -0.1 -0.2 -0.1  0.2  0.1 -0.2 -0.1  0.0  0.1 -0.0 -0.0 -0.2 -0.4 -0.1 -0.0 -0.1 -0.0
!
```

Niveau 2-A : la requête.

Un serveur en node.js est dédié à l'apprentissage d'OpenAI. Je corrige les réponses d'OpenAI via le browser, et openAI ajoute ces informations dans la limite de 15 messages par requête et 2000 symboles par requête au total, dans la requête OopenAI.

Ainsi, j'ai pris deux messages aux contenus assez « vifs » par rapport à un même événement ayant eu lieu le 25 septembre 2022.

Premier message analysé

« Internet se dispute. Est-il vrai qu'Adam Kadyrov a battu le brûleur du Saint Coran Nikita Zhuravel, complice des services spéciaux ukrainiens, ou pas ? Il l'a battu et a fait ce qu'il fallait.

En outre, je crois que quiconque ayant porté atteinte à une Écriture Sainte, y compris en la brûlant ostensiblement, en offensant par cela ainsi des dizaines de millions de citoyens de notre grand pays, doit être sévèrement puni.

Notre pays dispose d'une loi unique protégeant les Saintes Écritures, signée par le président russe Vladimir Vladimirovitch Poutine. Les valeurs religieuses des citoyens du pays sont protégées par l'État au plus haut niveau. Et cela ne peut pas être enlevé par des provocateurs !

Les actes d'empiétement sur le sacré ne sont caractéristiques que des personnes cyniques, immorales, molles et corrompues qui veulent semer la discorde dans la société, que notre État élève dans les idéaux de tolérance, de tolérance et d'amour mutuel. Ces ordures ne comprennent pas le langage de la persuasion, ils ne veulent pas construire un État ni le développer. Ils ne veulent pas que tout le monde en Russie vive côte à côte dans la paix et la compréhension. De tels provocateurs et traîtres sont une tumeur malade sur le corps qui doit être cautérisée.

Sans exagération, oui, je suis fier de l'action d'Adam. Il s'est toujours distingué par le désir de grandir non pas parmi ses pairs, mais parmi les plus âgés, grâce auquel il a formé des idéaux adultes d'honneur, de dignité et de défense de sa religion. Je respecte son choix. » ¹⁵

15 On retrouve un article similaire sur le lien <https://iqna.ir/fr/news/3486073/lauteur-dun-incendie-du-coran-battu-par-le-fils-de-ramzan-kadyrov>

Parameters	Old	Corrected	Explanation
positive author attitude (-1)	0	-1	here the attitude of the author is positive
the message is advantageous for the actual political power (-1)	-1		
exaggerations (-1)	0	-1	
lack of object concordance (-2)	0		
appeal to the intellect (1)	0		
appeal to emotions (-1)	0	-1	
appeal to virtues (-1)	-1		
appeal to ethics (-1)	-1		
appeal to blinding generalities as fatherland, peace, glory, justice, honor (-1)	-1		
appeal to fear (-1)	0		
appeal to the threat from other countries (-1)	0		
appeal to hatred (-1)	-1		
appeal to authority (-1)	-1		
appeal to human rights (1)	0		
criticism of the actual political power (1)	0		
generalizations (-1)	-1		
seemingly scientific appearance, but only seemingly (-1)	0		
real scientific proofs of the information (1)	0		
real proofs of the information (1)	0		
euphemismes replacing offensive expressions (-1)	0		
euphemismes replacing war or victims (-1)	0		
formal register of the texte (-1)	-1		
beneficiaries of the information regarding to politics (-1)	-1		
beneficiaries of the information regarding to money (-1)	0		
flattery to the reader (-1)	0		
dilemma between only two choice (-1)	0	-1	
criticism of a person instead of criticism of his arguments (-1)	-1		

Deuxième message analysé

« Les passages à tabac dans le centre de détention provisoire de Grozny sont sans aucun doute la meilleure chose qui soit arrivée en Russie. Une vidéo comme celle-ci pourrait coûter des milliards. Mais le stupide Adam a tout fait gratuitement. Cette vidéo manquait cruellement pour compléter le tableau.

Cette petite scène historique a réduit en cendres 20 ans de propagande.

Tous les points ont été placés. Les dernières illusions se sont évaporées. Le fiston de Kadyrov a donné le coup de grâce aux arguments des rares partisans de Poutine dans le monde civilisé. Le régime a arraché son dernier masque et a montré son visage. Si quelqu'un avait des doutes sur le fait qu'il était la quintessence du mal et de l'abomination, ils ne sont plus là.

Les grincements isolés et les « plaintes » de moutons n’y changeront rien. L’image est trop lumineuse et sans ambiguïté. La vidéo de Grozny s’est fermement ancrée dans les cerveaux de millions de personnes. Et c’est merveilleux.

Les orques 20 du front ont enfin découvert pourquoi ils mouraient : pour l’hitlérisme de Poutine et sa « verticale ». Les foules comprendront enfin qu’elles ne se battent pas du tout « pour leur patrie », mais pour avoir l’opportunité d’être battues, abaissées et salies au gré des caprices de n’importe quel roi fou.

Après tout, n’importe quel « citoyen » peut jouer le rôle de Zhuravel. Et vous n’êtes pas obligé de brûler le Coran pour cela. Il suffit simplement de ne pas plaire.

La scène dans la salle d’isolement a également informé l’intelligentsia de son avenir proche. Elle doit enfin comprendre ce qui l’attend en cas de dissidence, même la plus légère.

Cependant, l’obéissance sera désormais également récompensée par les coups de pied.

Il ne fait aucun doute que Poutine a conçu une construction sociale plus abrupte que celle de l’URSS et de la Corée du Nord. Kadyrov a simplement réalisé et publié une « présentation » visuelle de l’avenir russe. Un spoiler puissant et convaincant. Les « sentiments offensés » du croyant n’ont absolument rien à voir avec cela. »

- Niveau 4 : le tchat.

Le paramètre « rôle » permet de travailler sur les messages échangés sur un même tchat, ce qui permet d’améliorer les réponses grâce à la mémoire dans un tchat.

- Niveau 4 : au niveau d’OpenAI.

L’utilisation de GPT-4 présente une limitation majeure : son incapacité à apprendre de manière incrémentale. Cependant, une solution pourrait consister à intégrer une interface d’interaction, comme un serveur Node.js, pour annoter et corriger ses réponses en temps réel. Bien que cela ne change pas le modèle lui-même, en accumulant un ensemble de données de feedback, on peut orienter GPT-4 à fournir des réponses mieux adaptées. En ajoutant les corrections récentes à chaque nouvelle requête, on donne au modèle un contexte pour éviter de reproduire les mêmes erreurs.

Message: **Эрдоган считает, что Швеция делает недостаточно

GPT Extracted Affirmations:

Эрдоган считает, что Швеция выполнила свои	false
Террористические организации прекратили	false
Швеция изменила законодательство для	true
Эрдоган считает, что изменения в	false

Corrected Affirmations:

Эрдоган считает, что Швеция выполнила свои	false
Террористические организации прекратили	false
Швеция изменила законодательство для	true
Эрдоган считает, что изменения в	false

Chapitre 9 : Perspectives touchant le traitement prochain de la détection des messages de propagande dans les média

Il serait judicieux d’affiner et d’étendre ces méthodologies pour s'adapter à un paysage médiatique en constante évolution. Entre autre, j’envisage les perspectives suivantes comme la mise en place d’un système d’alerte pour des situations anormales, la détection systématique de la première chaîne à diffuser une information, le traitement croisé par plusieurs AI, la comparaison des résultats avec d'autres projets.

En outre, comme susmentionné, il serait pertinent de mener des études pluridisciplinaires associant Sciences et Technologies du Logiciel (STL), sciences politiques et sciences sociales – ainsi que d’autres pouvant, selon le sujet traité, entrer en ligne de compte (Histoire et Géographie, Économie, Démographie...)

Chapitre 10 : Bilan personnel et professionnel de l’expérience

Choses apprises durant le stage

Au cours des 5 mois passés au sein de l’..., je me suis familiarisée avec les langages de programmation Python et node.js. Pour la première fois, j’ai abordé un framework et appris beaucoup de détails (comme, par exemple, basculer entre les différentes versions de github, ou mettre les données sensibles de connexion dans le fichier .env).

l'Intérêt de mon approche, apport de ma contribution au domaine

Les missions qui m'ont été confiées m'ont donné l'opportunité de contribuer à un domaine d'étude passionnant car de nature pluridisciplinaire. En effet, ces tâches ont eu pour avantage de cumuler approche informatique, connaissances linguistiques et théories relevant des Sciences politiques et des Sciences sociales.

Concernant le caractère professionnalisant des missions réalisées, celles-ci ont, pour l'essentiel, été de nature expérimentale. En effet, l'essentiel des tâches a consisté à réaliser des tests sur OpenAI.

En dépit de ses limitations, OpenAI s'est révélé constituer un outil puissant pour traiter les informations en nombre.

Mais cette expérience a également été intéressante du point de vue de l'éthique.

Projets d'avenir

Si j'ai très rapidement trouvé ma place dans l'équipe d'... (en précisant que j'ai acquis une certaine expérience de la vie et de l'entrepreneuriat qui a rendu cette étape des plus aisées), le stage ne m'a déçu qu'en un point : sa brièveté. De fait, à ce jour, mon désir est de poursuivre dans cette voie et de travailler au sein d'une entreprise à vocation sociale.

En cela, ce stage m'a offert une base solide me permettant de mener des réflexions.

En termes de projets d'avenir, particulièrement intéressée par le domaine du numérique associé à la linguistique, et ayant étudié 5 années durant la linguistique à Moscou et Paris, je souhaite trouver ou créer une activité professionnelle me permettant de cumuler ces deux domaines de recherche. Notamment, j'ai pour ambition de créer une entreprise d'informatique afin de travailler, entre autres choses, sur les projets sociaux (comme la détection des situations facilitant le recours à la corruption, les manifestations citoyennes au sein des applications, un cryptomonnaie miné par la bonté, etc.) Effectivement, les technologies ont déterminé en grande partie les relations dans la société tout au long de l'histoire. En espèce, l'informatique aujourd'hui est un outil extraordinaire d'organisation de la vie sociale.