

BEDDAR Meyssa, BLACHON Tom, SIMOES Matthieu

# LE CONFLIT RUSSO-UKRAINIEN VU PAR TWITTER

# TABLE DES MATIERES

## Table des matières

Introduction	1
Méthodologie de construction du corpus	4
Choix du sujet et recherches préliminaires	4
Choix du réseau, captation et collecte	5
Périmètre des corpus d'analyse	8
Analyse du premier corpus – Ensemble des utilisateurs francophones	11
Utilisateurs les plus actifs	11
Utilisateurs les plus populaires	13
Champ lexical des 1000 tweets les plus populaires	15
Analyse du deuxième corpus – Candidats à l'élection présidentielle	18
Implication des candidats	18
Popularité des tweets	20
Champ lexical du corpus	22
Relations entre candidats	25
Conclusion	31

## Introduction

En février 2022, la situation politique se tend entre l'Ukraine et la Russie lorsque des mouvements de troupes militaires russes sont détectés à la frontière ukrainienne, menaçant d'une potentielle invasion. Confirmant les craintes, le président russe Vladimir Poutine reconnaît l'indépendance de deux régions ukrainiennes et envahi son voisin le 24 février. La guerre est officiellement déclarée entre les deux pays, suscitant de vives réactions dans toute l'Europe. Celles-ci proviennent tant de la part de personnalités politiques que de personnes désireuses de partager leur opinion de la situation.

En effet, au XXIème siècle, apogée des nouvelles technologies de l'information et de la communication, les réseaux sociaux constituent des places de choix permettant une libre circulation des informations. Les utilisateurs peuvent y suivre, presque en temps réel, l'évolution du conflit et s'exprimer à son sujet.

A l'échelle européenne, il s'agit d'une situation inédite : une guerre de proximité, dont les victimes s'expriment sur les réseaux et auxquels les européens s'identifient. De fait, le conflit génère un engouement massif, de l'intérêt, de l'empathie, mais également de la crainte et de la stupéfaction. En 2022, un tel déchainement des passions est partagé par le plus grand nombre à travers les canaux de communication en ligne. Ainsi, ces plateformes présentent de nombreux avantages. En premier lieu, elles favorisent la libre circulation des informations et restreignent la censure, volontaire où involontaire. Les nouvelles vont vite, sans nécessairement passer par la presse. Tout le monde est ainsi en mesure de s'informer de façon autonome, sans être contraint par les canaux traditionnels. D'autre part, les réseaux sociaux, comme Twitter, permettent de centraliser l'information et favorisent la multiplication des sources. Tout utilisateur est libre de consulter les médias qu'il désire pour s'informer et il peut croiser les sources pour vérifier la véracité d'un fait et se forger une opinion.

# INTRODUCTION

En outre, ces plateformes peuvent également générer des biais et avoir des effets nocifs quant à l'accès à l'information de la part de leurs utilisateurs. En effet, la libre expression et l'espace public offert par ces réseaux favorisent également la propagation de fausses informations dites « *Fakes news* ». S'ils ne sont pas vigilants, les membres peuvent alors les relayer involontairement et à leur tour, prenant ce qui est partagé par d'autres pour argent comptant. D'autant que le phénomène de groupe est exacerbé dans cet espace d'échange : plus un contenu est relayé, plus les utilisateurs sont susceptibles de le juger fiable et d'avoir une vision biaisée de la réalité. Nous pouvons citer à titre d'exemple une vidéo partagée de nombreuses fois au début de la guerre, exposant prétendument « le bombardement d'une centrale électrique à Kiev ». Il s'agissait en réalité d'une explosion ayant eu lieu dans une zone industrielle chinoise en 2015. Le caractère sensationnel de la vidéo d'origine a suffi à la rendre virale, au détriment de la réalité. Comme le disait déjà l'orateur grec Démosthène durant l'antiquité : « On croit ce que l'on veut croire. »...

Une autre limite de ces plateformes est l'enfermement communautaire. En effet, les algorithmes de recommandation de contenu proposé aux utilisateurs se base sur les centres d'intérêts et les relations de ces derniers avec d'autres personnes du réseau. De ce fait, un membre se retrouve rapidement enfermé dans un cercle de contenu en accord avec ses idées et sa façon de penser. Cette surexposition à un certain type de contenu peut alors conférer une vision biaisée de la réalité, passée au filtre d'un algorithme.

Ainsi nous avons entrepris de nous intéresser, lors de notre étude, à ce conflit européen à travers le prisme du réseau social Twitter et de ses utilisateurs. Nous venons de pointer du doigt les avantages et les mises en garde relatives au relai de l'information sur cette plateforme. Aussi, nous souhaitons observer la façon dont les actualités sont partagées sur ce réseau, par qui et de quelle manière. Il serait ainsi intéressant de déterminer quels sont les relais d'information privilégiés par les utilisateurs et les mots-clés les plus plébiscités pour évoquer le conflit. De la même manière, à l'aube de l'élection présidentielle française d'avril 2022, il nous semble

# INTRODUCTION

intéressant d'observer l'activité des candidats sur le réseau, leurs prises de positions sur la guerre ukraino-russe et les réactions suscitées auprès des autres utilisateurs.

De notre étude découlent plusieurs enjeux sociaux et politiques. D'une part, observer la compréhension d'un phénomène à travers ces nouveaux canaux d'informations. Il est en effet légitime de se demander si une ouverture globale et une discussion multiculturelle et internationale va mener à une compréhension commune du conflit, ou si, à l'inverse, la plateforme va renforcer les clivages entre les individus. D'autre part, un autre enjeu est celui de la puissance politique véhiculée par Twitter. En somme, si le conflit militaire est bien réel, nous pouvons nous questionner sur l'existence, en parallèle, d'une guerre numérique. Les personnalités politiques sont de plus en plus présentes sur les réseaux et il est intéressant de constater le poids des idéologies politiques sur ce conflit.

Ainsi, en tenant compte des différents enjeux relatifs à notre étude, nous nous demanderons **comment est perçue la guerre en Ukraine par la population française, à travers les utilisateurs francophones de Twitter, tant sur le plan social que politique**. Pour cela, nous bornerons notre étude entre le 24 février 2022, date de l'invasion de l'Ukraine par la Russie, et le 24 avril 2022, date de fin de nos observations et des élections présidentielles. Sur le plan géographique, nous avons choisi de nous limiter aux utilisateurs francophones pour analyser nos résultats dans un contexte culturel que nous connaissons. En effet, une analyse de la réception et de la sémantique n'a de sens que si l'on comprend la culture et la langue des utilisateurs observés.

Aussi, afin de présenter notre étude, nous décrirons dans une première partie notre méthodologie d'enquête et de collecte des données. Nous y aborderons les choix effectués en fonction des informations recherchées au préalable, jusqu'à parvenir à la constitution de notre corpus de données final. Puis, dans une seconde partie, nous analyserons les résultats obtenus et les graphiques choisis pour illustrer au mieux notre propos.

## Méthodologie de construction du corpus

En premier lieu, notons que notre travail de collecte et d'analyse a fait l'objet d'une méthodologie rigoureuse. En effet, une étude approfondie portant sur une thématique spécifique dont nous ne sommes pas familiers nécessite d'être préparée en amont afin de rendre cette dernière pertinente et intelligible, tant par notre équipe que nos lecteurs potentiels.

### CHOIX DU SUJET ET RECHERCHES PRELIMINAIRES

La première partie de notre travail a consisté à rechercher un sujet d'étude répondant aux contraintes de l'exercice. Pour rappel, il s'agissait de choisir une thématique libre, mais celle-ci devait être traitée à travers l'analyse d'un ou plusieurs réseaux sociaux. L'objectif serait alors d'extraire les données de ces réseaux afin de les interpréter et de les visualiser à l'aide d'autres outils. Par conséquent, il existe plusieurs restrictions de notre champ d'étude intrinsèques à cette consigne. D'une part, le sujet doit être actuel : sur les réseaux, l'information circule particulièrement vite, mais elle disparaît également rapidement et certaines thématiques, telles des « effets de modes », ne font l'objet que de discussions éphémères et disparaissent de la toile en quelques heures. Or, il est particulièrement ardu de retrouver des traces de sujets anciens et dépassés et ces plateformes, d'autant que la plupart des outils de collecte ne permettent qu'une extraction des publications datant des derniers jours.

D'autre part se pose une limite quantitative. La thématique choisie doit susciter un nombre important de réactions de la part des utilisateurs en ligne. Les réseaux sociaux sont des phénomènes de masse où les membres exposent leurs opinions, s'influencent et débattent

ensemble. Un faible nombre de réactions sur le sujet ne nous permettrait pas de constituer un réseau de publications suffisamment riche et ainsi de délivrer une analyse approfondie.

Partant de ces contraintes, nous avons donc entrepris la recherche d'un sujet d'actualité susceptible d'évoluer et de susciter des réactions et des avis potentiellement divergents. Nous nous sommes donc rendus sur l'onglet « Tendances en France » de Twitter afin d'y observer les sujets récurrents. Ceux-ci traitaient principalement des élections présidentielles 2022, alors en préparation, de protestations envers le gouvernement actuel, ainsi que de la tension naissante entre la Russie et l'Ukraine. Nous avons ainsi opté pour ce dernier sujet. Notons que nous ne savions pas, à ce moment-là, qu'il prendrait une telle ampleur, car nous avons réalisé notre sélection le 20 février 2022, soit avant le début officiel de la guerre.

Cependant, si nous avons choisi de nous concentrer sur ce sujet, nous avons également entrepris de nous renseigner sur le contexte politique qui l'entourait. En effet, nous souhaitons nous informer sur les raisons historiques menant à ce conflit afin de mettre en perspective ces événements, leurs tenants et leurs aboutissants. Une guerre si complexe ne peut pas être correctement analysée si nous n'en comprenons pas les enjeux. D'autant que, rappelons-le, notre analyse se concentre sur les opinions des utilisateurs de réseaux sociaux. De fait, les avis divergents et des *fakes news* circulent au rythme de l'actualité. Afin d'éviter l'écueil d'une perception biaisée et pour la pertinence de nos résultats et interprétations, il nous a donc semblé bon de réaliser un certain nombre de recherches préalables sur le sujet. Pour cela, nous avons privilégié la consultation de médias plus traditionnels, en ligne, télévisés et radiophoniques.

## CHOIX DU RESEAU, CAPTATION ET COLLECTE

La seconde étape de notre méthodologie de collecte des données a été de sélectionner le ou les réseaux sociaux nous semblant les plus appropriés, mais également les corpus d'étude et les outils d'extraction et d'analyse des données.

Concernant le choix du réseau tout d'abord, nous avons opté, après réflexion, pour ne retenir que Twitter dans le cadre de nos recherches. En effet, nous avons sélectionné notre thématique en nous basant sur les « tendances » de Twitter. Par conséquent, nous savions que ce sujet était particulièrement abordé et débattu sur cette plateforme, tandis que nous n'avions pas cette garantie pour les autres réseaux sociaux. Par ailleurs, d'autres plateformes telles que Facebook ou LinkedIn nous ont semblé trop orientées autour du réseau de proximité : lorsqu'un utilisateur s'y exprime, il s'adresse généralement à ses connaissances et non à un public global. L'interaction y est ainsi plus limitée et le débat nous y a donc semblé trop peu ouvert pour réaliser une collecte d'opinion pertinente. L'autre réseau ayant instauré une logique d'ouverture similaire à Twitter est Instagram, mais ce dernier est principalement destiné à l'expression à travers la publication d'images et de photographies. Il est alors bien plus difficile d'en extraire et d'en étudier les données, d'où son exclusion de notre étude. Enfin, notons que nous avons rejeté d'office les réseaux sociaux privilégiant le format vidéo, tels que TikTok, Youtube, ou encore Snapchat. Seul Twitter a donc été conservé.

Concernant la sélection de nos données d'étude, nous reviendrons plus tard sur nos critères d'extraction et le périmètre de notre étude, mais nous avons choisi de nous intéresser à deux corpus de données différents :

- Un premier corpus de **tweets global durant une semaine** : Ce premier corpus, global, vise à collecter 16 000 tweets (la capacité d'extraction maximum de notre outil) en français à un instant donné. Notre outil nous permettant de récolter les données datant au maximum de 7 jours avant la date d'extraction, ce corpus sera alors constitué de tweets portant sur le conflit durant 1 semaine, au cœur de la guerre. Les tweets sont alors filtrés, lors de la collecte, en se basant sur une série de mots-clés et de hashtags portant sur la guerre que nous détaillerons plus bas. L'objectif de ce corpus est ainsi de pointer du doigt les grandes tendances se dégageant du conflit : quels sont les utilisateurs les plus actifs, les plus suivis, ou encore quels sont les modes de publications et les formats privilégiés ?



- Notre second corpus se concentre sur **les candidats à l'élection présidentielle française**. Les deux événements ayant eu lieu à la même période, il nous a semblé intéressant d'étudier les tweets des différents candidats sur la guerre d'Ukraine. Ce conflit ayant un enjeu politique fort à l'échelle européenne, voire internationale, il s'est également rapidement placé au cœur du débat lors de la campagne présidentielle. Ainsi, l'observation de ce corpus, portant sur toute notre période d'étude, vise à déterminer les candidats les plus impliqués sur ce sujet, leurs opinions et leur popularité sur une thématique devenue clé à l'aube des élections.

Enfin, nous avons sélectionné nos outils de collecte et d'analyse de nos Tweets. Après nous être penché sur différentes options, nous avons privilégié l'utilisation du logiciel R pour la majorité de nos traitements. En effet, à l'aide de sa bibliothèque « *RTweet* » et d'un compte Twitter développeur, nous avons pu aisément réaliser l'extraction de nos corpus grâce à de multiples fonctionnalités de filtrage. Le sujet étant international, nous avons notamment pu appliquer une restriction sur la langue de collecte, nous permettant d'obtenir que les tweets francophones. De plus, notre équipe dispose de bonnes compétences en matière de manipulation de données sur R, nous permettant, sur le même outil, de procéder à un nettoyage sur mesure de nos jeux de données. Nous avons alors jugé que cette solution était plus adaptée à notre étude que son homologue *no code* : Netlytic. Toutefois, comme mentionné plus tôt, nous demeurons limités à une collecte de tweets datant de 7 jours maximum avant la date d'export.

Nous n'avons pas non plus entrepris d'utiliser le logiciel Gephi dans un but de collecte des données en temps réel. En effet, nous devions pour cela disposer d'un accès avancé à l'API de Twitter. Or, malgré nos demandes dès début mars, nous n'en avons obtenu l'accès que mi-avril. L'effervescence des réactions était alors passée et la quantité de données à disposition était donc insuffisante. Toutefois, nous avons utilisé Gephi lors de nos traitements, en modifiant et important directement les données collectées grâce à R, afin de réaliser une carte de relation des tweets des candidats aux élections présidentielles, comme nous le verrons plus tard. Du reste, la majorité de nos traitements ont été effectués directement sur R.

## PERIMETRE DES CORPUS D'ANALYSE

Une fois nos corpus d'analyse déterminés dans les grandes lignes, ainsi que le réseau social que nous souhaitons utiliser, la dernière étape avant de procéder à l'export des données consiste à délimiter avec précision le périmètre de nos corpus d'analyse. Pour cela, nous avons listé un certain nombre de critères de collectes à respecter impérativement lors de l'extraction :

- Données issues de **Twitter uniquement** : nous l'avons signalé plus haut, il s'agit du seul réseau sur lequel nous souhaitons baser notre étude.
- Données en **langue française** : pour le corpus des candidats à l'élection présidentielle, cela était indispensable. Pour le corpus général, nous ne souhaitons étudier qu'une seule population afin d'éviter de nous disperser face à la masse de données disponible. Comme nous ne pouvions restreindre notre collecte à une région géographique, nous avons alors pris la décision de nous limiter à une seule langue.
- Tweets **bornés chronologiquement** entre le 24 février 2022, date de déclaration de la guerre par la Russie, et le 24 avril 2022, date de fin du second tour des élections présidentielles et de la soumission de notre étude. Cette période de 2 mois nous a semblé être un cadre naturel à nos analyses.
- Des **utilisateurs ciblés** : nous l'avons également mentionné plus haut, mais les deux corpus que nous souhaitons constituer nous imposent une restriction naturelle des utilisateurs collectés. D'une part, les utilisateurs francophones de Twitter, de l'autre, les 12 candidats en lice pour l'élection présidentielle française.
- Données **filtrées par mots-clés** : Afin de centrer nos recherches sur le sujet précis de la guerre en Ukraine. Ainsi, nous avons établi une liste de mots-clés et de hashtags devant être présents au sein des tweets pour être sélectionnés. Ceux-ci devaient être à la fois

spécifiques à notre sujet, sans quoi nous risquerions d'exporter des tweets hors contexte susceptibles de biaiser notre étude, mais également suffisamment large pour omettre le moins de publications possibles. La sélection de ces mots-clés s'est également faite à l'aide de *Google Trend*, que nous avons utilisé afin de détecter les termes ayant connu un pic d'utilisation depuis le début de la guerre :

- **Corpus général** : « Ukraine », « Russie », « Ukrainewar », « Zelensky », « Poutine », « Donbass », « #kiev », « #ukrainerussie », « #laguerre », « #ukrainewar », « #invasion », « #worldwar3 », « #ukraine » et « #russie ».
- **Corpus des candidats** : « Ukraine », « Russie », « Ukrainewar », « Zelensky », « Poutine », « Donbass », « Kiev », « Ukrainien », « #Kiev », « #UkraineRussiaWar », « #UkraineRussie », « #UkraineInvasion », « #LaGuerre », « #UkraineWar », « #Invasion », « #Worldwar3 », « #Ukraine », « #Russie », « #Poutine ».

Pour ce dernier corpus, ne disposant pas d'une API Twitter premium, nous avons été contraints d'extraire l'intégralité des tweets de chaque candidat individuellement. Nous avons ensuite réuni ces 12 bases ensemble, puis nous avons appliqué notre filtrage des données par date et par mots-clés afin de ne conserver que les tweets s'inscrivant dans le périmètre de notre étude.

Nous avons ainsi pu réaliser nos extractions et constituer nos deux corpus afin de procéder à leurs analyses. Cependant, soulignons pour terminer cette section, une limite de notre méthodologie. En somme, si le corpus des candidats regroupe l'ensemble des tweets publiés par ces derniers sur l'intégralité de notre période d'analyse, il n'en va pas de même pour le second corpus, plus général. Ainsi, ce dernier a été extrait à un instant précis et ne regroupe que les tweets des 7 derniers jours précédant la collecte. De fait, il est le reflet des tendances sur une période d'une semaine parmi les 2 mois de notre analyse. Notre traitement de ce corpus n'incorpore donc pas d'éléments comparatifs avec les tendances à une autre date de la période. Il s'agit d'un parti pris de l'étude, supposant que la semaine collectée était en partie

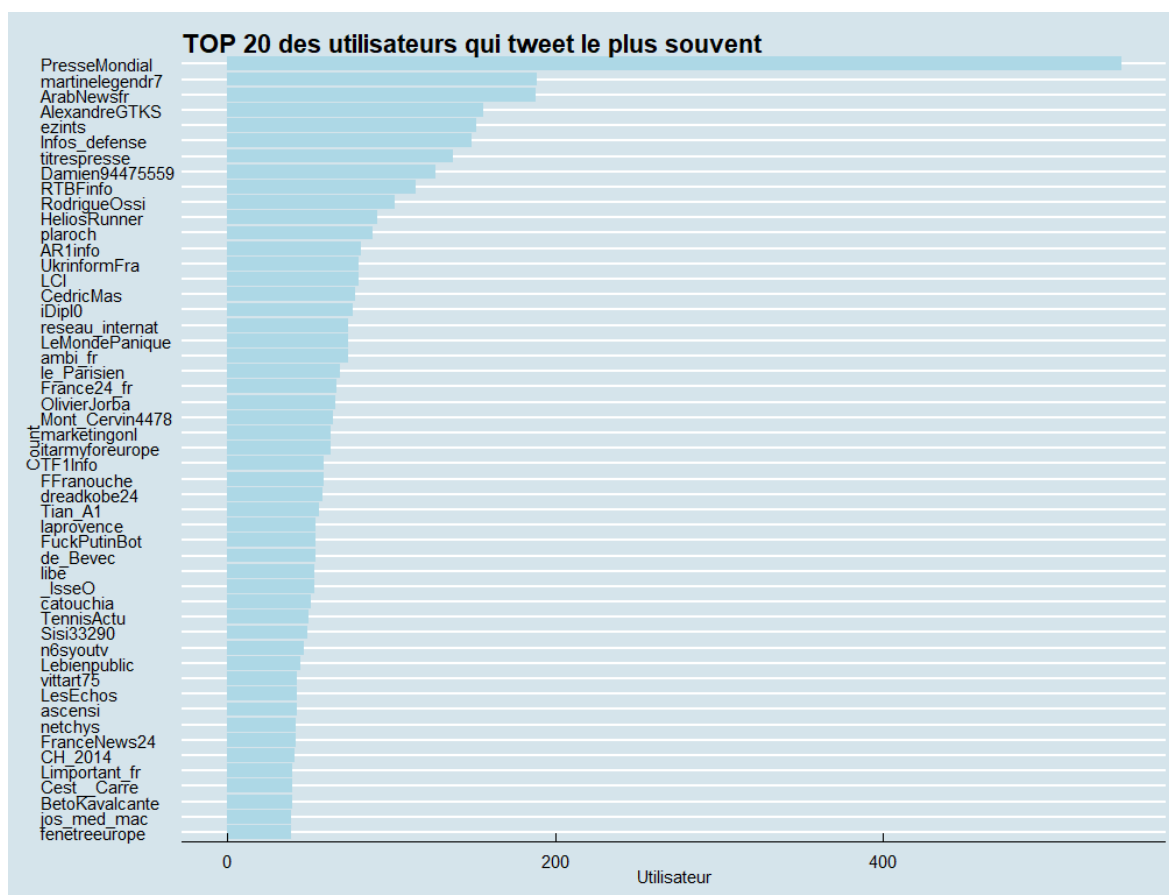
représentative de l'ensemble de la période. Toutefois, nous ne pouvons en avoir la confirmation et il est certain que certains phénomènes ont évolué en deux mois.

# ANALYSE DU PREMIER CORPUS – ENSEMBLE DES UTILISATEURS FRANCOPHONES

## Analyse du premier corpus – Ensemble des utilisateurs francophones

### UTILISATEURS LES PLUS ACTIFS

La librairie Rtweet propose, à partir de l'extraction des tweets, pas moins de 90 variables permettant de décrire chaque tweet. Nous avons décidé de nous concentrer sur un nombre restreint de variables, qui à notre sens apportait les informations les plus intéressantes au vu de notre analyse.



# ANALYSE DU PREMIER CORPUS – ENSEMBLE DES UTILISATEURS FRANCOPHONES

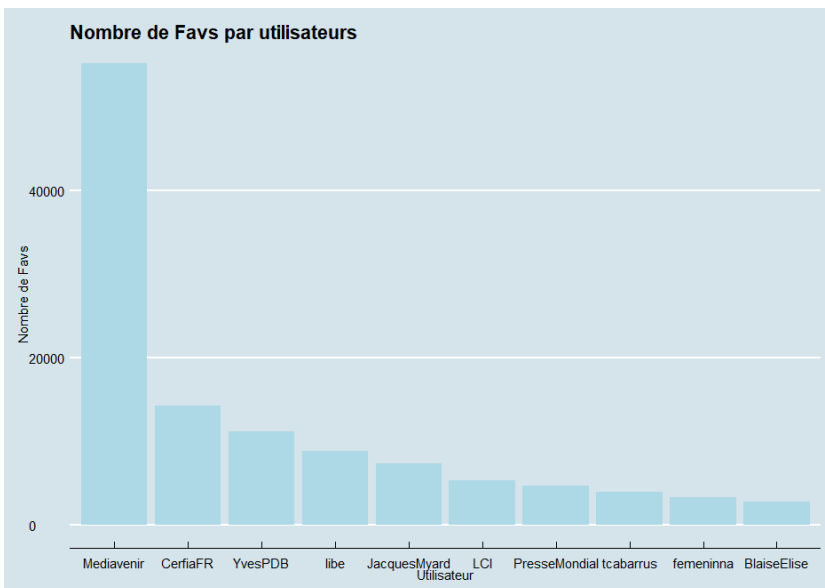
Premièrement, nous avons décidé de visualiser à travers le graphique ci-dessous quels sont les utilisateurs ayant le plus publié sur la plateforme au sujet du conflit russo-ukrainien. En effet, nous avons voulu observer quels sont les utilisateurs qui tweetent le plus souvent, afin de déterminer la part d'utilisateurs certifiés, et de voir s'il existe une grande disparité en terme de nombre de tweets par utilisateur.

On remarque ainsi que l'utilisateur ayant réalisé le plus de tweets est *PresseMondial*, avec plus de 500 tweets comptabilisés. On note également une nette différence entre le nombre de tweets de cet utilisateur et les autres. Effectivement, en comparaison avec le 2<sup>ème</sup> utilisateur ayant le plus contribué au commentaire du conflit, *PresseMondial* a publié plus de deux fois plus de tweets. Il apparaît important de préciser que *PresseMondial* est spécialisé sur l'information Franco-Russe. Il est donc peu étonnant d'observer une implication aussi forte de cet acteur quant à cette guerre.

Autrement, les types d'utilisateurs sont partagés entre des particuliers ou de la presse spécialisée dans un thème. De plus, si l'on regarde ces 20 utilisateurs, on voit que le volume de tweets par utilisateurs reste assez homogène hormis pour les 3 premiers utilisateurs.

# ANALYSE DU PREMIER CORPUS – ENSEMBLE DES UTILISATEURS FRANCOPHONES

## UTILISATEURS LES PLUS POPULAIRES

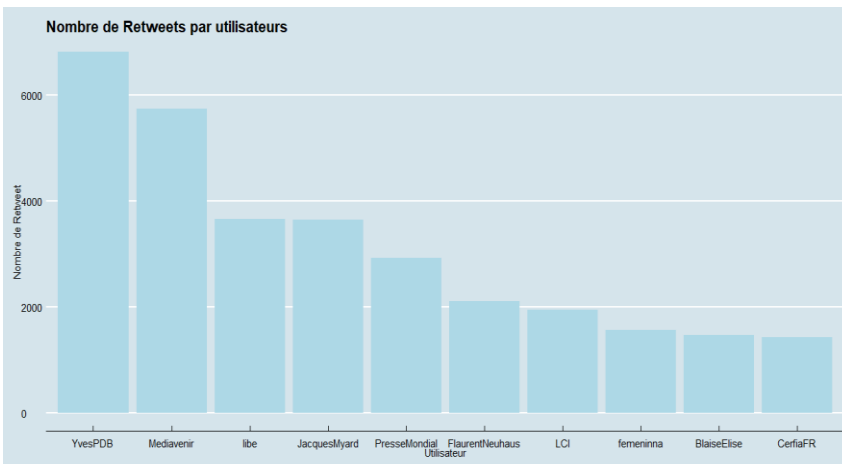


Le graphique ci-contre représente le top 10 des utilisateurs qui comptabilisent le plus de tweets en favoris.

On observe tout d’abord une grande disparité selon les utilisateurs. En effet, l’utilisateur comptant le plus de tweets en favoris est *Mediavenir*, et il est intéressant de préciser que ce dernier n’apparaissait pas parmi les utilisateurs les plus actifs. Malgré tout, on retrouve certains utilisateurs qui étaient dans le top 20 comme *PresseMondial*, et d’autres presses comme *Cerfia*, *Libération* ou encore le compte twitter de la chaîne de télévision *LCI*.

On remarque également la présence du sénateur et membre du conseil de Paris Yves Pozzo di Borgo et de la journaliste Blaise Elise.

# ANALYSE DU PREMIER CORPUS – ENSEMBLE DES UTILISATEURS FRANCOPHONES



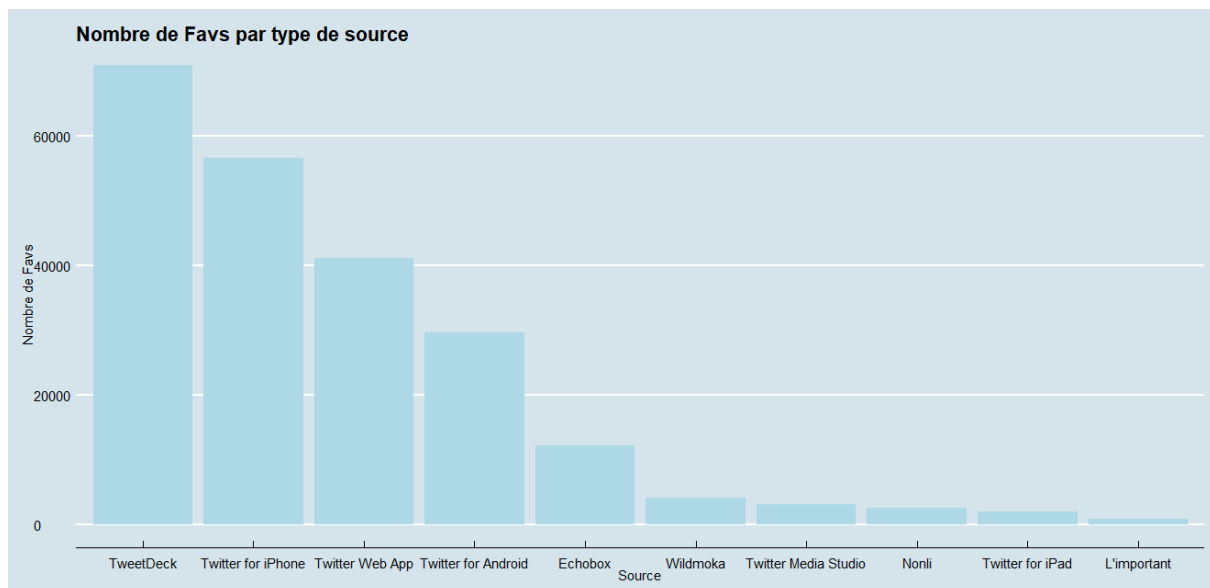
Ce graphique représente quant à lui le top 10 des utilisateurs les plus retweetés.

On remarque alors que ce top 10 comprend les mêmes utilisateurs que celui concernant le nombre de likes, bien que le classement s'en retrouve changé.

Ici, c'est l'utilisateur « YvesPDB » qui est le plus retweeté et donc le plus partagé. Mediavenir est malgré tout 2<sup>e</sup> en terme de nombre de retweets. On peut donc voir en combinant ces deux derniers graphiques que le top 10 des utilisateurs qui ont des tweets les plus likés ou retweetés ne sont pas forcément ceux qui tweetent le plus souvent.



# ANALYSE DU PREMIER CORPUS – ENSEMBLE DES UTILISATEURS FRANCOPHONES



Ce dernier graphique permet quant à lui d'appréhender que la source comptabilisant le plus de favoris est une source n'apparaissant pas dans le top 3 des sources les plus utilisées. En effet, *TweetDeck* semble majoritaire avec plus de 10 000 favoris par rapport à la 2<sup>e</sup> source (« Twitter for iPhone »). Ce classement est donc intéressant car cette source est notamment utilisée par *Mediavenir* et permet notamment de tweeter depuis plusieurs comptes facilement.

## CHAMP LEXICAL DES 1000 TWEETS LES PLUS POPULAIRES

Afin d'avoir une idée des positions dominantes des utilisateurs francophones quant au conflit russo-ukrainien, nous avons décidé de réaliser une rapide analyse textuelle d'un corpus de tweets constitué des 1000 tweets les plus populaires en termes de nombre d'ajouts en favoris.

## ANALYSE DU PREMIER CORPUS – ENSEMBLE DES UTILISATEURS FRANCOPHONES



	Terme	Total
1	Ukraine	868
2	russe	379
3	guerre	239
4	Poutine	166
5	russe	123
6	russes	109
7	La France	90
8	marioupol	80
9	macrone	73
dix	les forces	72
11	Ukrainiens	68
12	donbass	64
13	zelenski	64
14	paie	58
15	Président	58
16	soldats	51
17	armes	50
18	direct	47
19	mariupol	45
20	ville	44

Cependant, force est de constater que le classement des 20 termes les plus fréquents n'apporte pas de grande information quant à la prise de position des utilisateurs francophones de Twitter.

En effet, nous retrouvons naturellement les termes « Ukraine », « Russe », « Guerre », « Poutine », étant donné le filtrage ayant été réalisé au préalable sur les tweets.

On remarque toutefois que le terme le plus fréquemment utilisé, loin devant les autres, est « Ukraine », avec 868 occurrences. Le terme « Russie » n'apparaît étonnamment pas dans ce classement, et les mots « russe » ou « russes » lui sont préférés.

Aussi, en terme de personnage public, Poutine est le premier mentionné avec 166 occurrences, loin devant Zelenski (64 occurrences), Macron ou Biden.

En revanche, le nuage de mot fait apparaître certains mots évoquant des prises de position, tels que « Sanctions », « Paix », « Aide », « solidarité » ou encore « soutien ». Il n'est cependant pas possible de distinguer visuellement le poids de ces différents termes les uns comparativement

# ANALYSE DU PREMIER CORPUS – ENSEMBLE DES UTILISATEURS FRANCOPHONES

aux autres. Réaliser une analyse plus poussée pourrait être intéressant afin de dégager les prises de positions dominantes.

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE

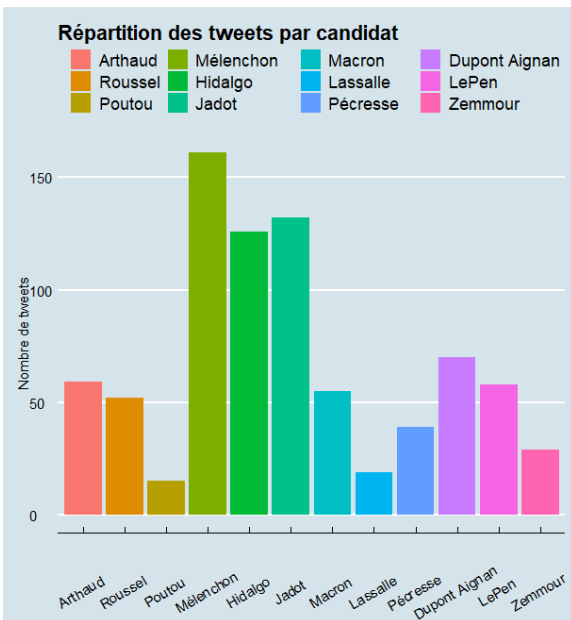
## Analyse du deuxième corpus – Candidats à l'élection présidentielle

Pour comprendre la façon dont le conflit a été appréhendé par la sphère politique française, nous avons décidé d'extraire les tweets des candidats à la présidentielle de 2022 durant une période précise : du 24/02/2022 – date d'invasion de l'Ukraine par la Russie – au 25/04/2022, lendemain du second tour de l'élection présidentielle. Le but est notamment d'interroger les positions sur ce conflit selon les candidats et les partis qu'ils représentent.

### IMPLICATION DES CANDIDATS

Sur la période du 24 février au 25 avril 2022, les 12 candidats à la présidentielle ont publié près de 9300 tweets au sujet du conflit russo-ukrainien selon le filtrage non exhaustif que nous avons réalisé, soit une moyenne de 775 tweets par candidat.

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE



Ce premier graphique permet de visualiser les candidats qui se sont le plus exprimés sur Twitter par rapport avec ce conflit.

On remarque que les candidats ayant réalisé le plus de tweets en rapport avec ce sujet sont les suivants :

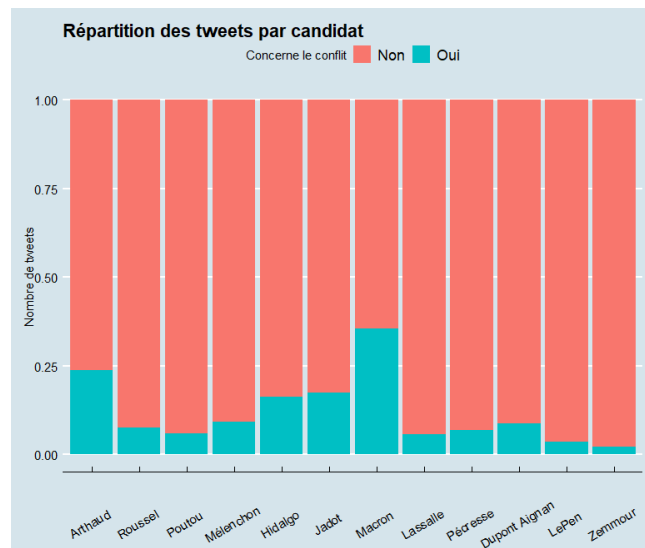
- Jean-Luc Mélenchon, avec plus de 150 tweets,
- Yannick Jadot, avec près de 140 tweets,
- Anne Hidalgo, avec 130 tweets.

Ainsi, les candidats s'étant le plus « impliqués » ont une orientation politique située à « gauche ».

A contrario, les candidats les moins présents dans cette base de tweets sont, par ordre décroissant, Philippe Poutou, Jean Lassalle et Eric Zemmour, avec moins de 50 tweets pour ces trois candidats.

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE

En revanche, sur l'ensemble des tweets durant cette période, on remarque qu'en termes de proportion, ce sont Macron et Arthaud qui ont le plus communiqué à ce sujet.



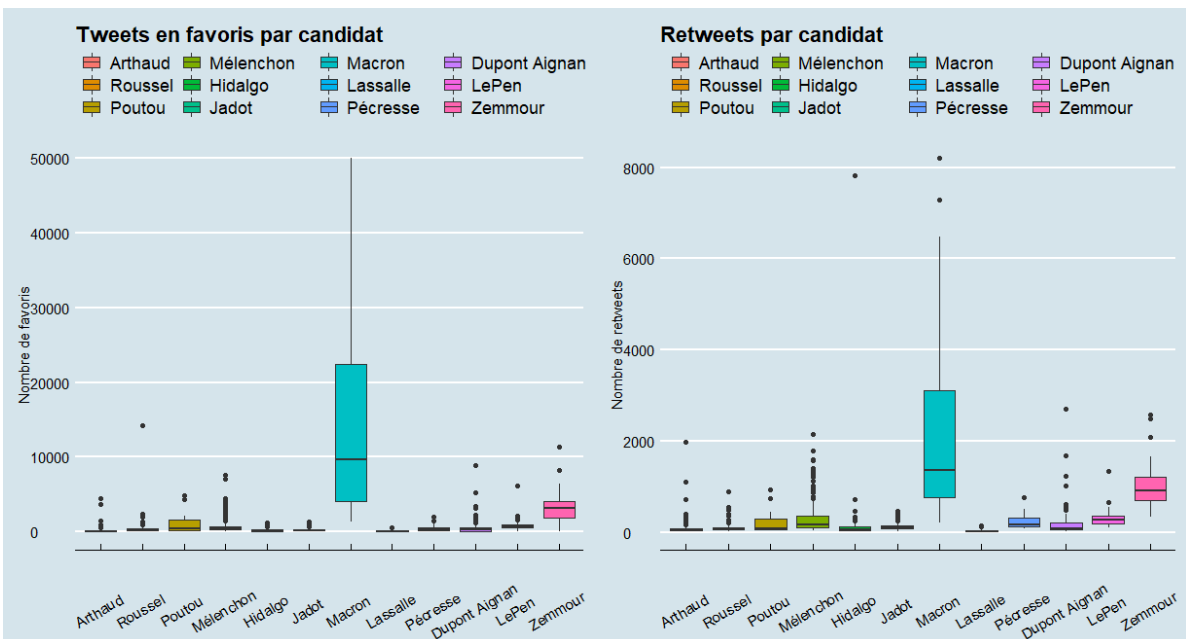
Effectivement, bien qu'Emmanuel Macron n'ait publié qu'une cinquantaine de tweets sur ce conflit, ils représentent près de 40% de l'ensemble de ses tweets sur cette période.

A contrario, si on pouvait penser que Jean-Luc Mélenchon était le candidat le plus impliqué, il apparaît que seuls 10% de ses tweets concernaient en réalité ce conflit.

On peut enfin noter que deux candidats se dégagent quant au nombre de tweets concernant ce conflit relativement à l'ensemble de leurs tweets. En effet, les deux candidats d'extrême droite Eric Zemmour et Marine Le Pen sont ceux ayant le moins communiqué sur ce sujet, lorsque l'on prend en compte le nombre de tweets qu'ils ont publié durant cette période.

## POPULARITE DES TWEETS

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE



Les deux graphiques ci-dessus permettent d'appréhender la popularité des tweets de chaque candidat, jugée à partir du nombre de favoris ainsi que du nombre de retweets par tweet.

Si Macron ne s'était pas distingué des autres candidats par le nombre brut de tweets publiés sur le conflit, il se démarque en revanche très clairement que ce soit par le nombre de favoris mais aussi par le nombre de retweets.

Egalement, si nous avons relevé que Zemmour faisait partie des candidats s'étant le moins exprimés, il se trouve en seconde position en termes de popularité de tweets.

Ainsi, le nombre de tweets sur ce sujet, témoignant de la volonté de s'exprimer quant à un sujet et de partager son opinion et son implication, n'implique pas que ces derniers seront populaires.

Par contre, cette répartition de la popularité ne semble pas spécifique au sujet. Après vérification sur l'ensemble des tweets de chaque candidat au cours de cette période, nous remarquons que dans l'ensemble, les candidats ayant le plus de favoris et retweets sont Emmanuel Macron et Eric Zemmour, ce qui coïncide avec les résultats obtenus sur les tweets concernant le conflit. Toutefois, on remarque que les tweets les plus populaires de notre base de données ont été

## ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE

publiés par Mélenchon et Poutou, mais force est de constater que ces derniers ne concernaient pas le conflit russo-ukrainien.

## CHAMP LEXICAL DU CORPUS

a. Ensemble des candidats



	Terme	Total
1	ukraine	372
2	guerre	274
3	poutine	231
4	russie	189
5	france	122
6	faut	108
7	russes	90
8	paix	86
9	l'ukraine	80
10	peuple	78
11	ukrainiens	77
12	sanctions	70
13	ukrainien	66
14	vladimir	66
15	doit	65
16	français	65
17	russe	65
18	président	60
19	faire	59
20	paris	51

Il est intéressant de noter que le conflit est considéré notamment par l'existence de deux grands protagonistes :

- L'Ukraine, en tant que nation,
- Vladimir Poutine, qui représenterait presque à lui seul le camp de la Russie.



# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE

Effectivement, le terme le plus fréquent du corpus et référencé pas moins de 370 fois est « Ukraine », contre 189 fois pour la Russie. Poutine est d'ailleurs plus cité que la Russie dans la globalité, puisque l'occurrence de son nom est de 231 fois. Aussi, on remarque que le nom « Zelensky », soit le président actuel de l'Ukraine, n'apparaît pas dans la liste des 20 termes les plus fréquents du corpus.

On peut alors supposer que l'intérêt se porte en particulier sur le sort de l'Ukraine, et Poutine serait la personnification de la Russie.

Aussi, plutôt que de parler de conflit, on utilise préférentiellement la notion de « guerre » pour décrire cet événement.

On peut également remarquer la prégnance du champ lexical du devoir, avec par exemple les termes « faut » ou encore « doit », ce qui témoigne d'une volonté d'agir ou à minima de prendre position.

Le nuage de mot met également en exergue certaines positions qui seraient dominantes mais relativement antagonistes :

- Les « sanctions », reflétant une prise de position contre la Russie et notamment Poutine (70 occurrences),
- Une volonté de « paix », de « cessez le feu », qui serait, si on se réfère au nombre de termes, légèrement plus populaire (86 occurrences pour la notion de paix par exemple),
- Bien moins populaire mais tout de même présent, un choix de neutralité.

Toutefois, on peut imaginer que les sanctions seraient un moyen de faire reculer Poutine et donc d'arriver à terme à un processus de paix.

Les champs lexicaux du soutien et de la solidarité sont également présents au sein des tweets des candidats à la présidentiel, comme en témoigne le nuage de mot faisant ressortir ces termes ou encore le hashtag « standwithukraine ».

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE

Les implications économiques sont également abordées avec notamment les questions du pétrole et du gaz russes.

## b. Par candidat

Après avoir ajouté en stop-words les mots relatifs aux programmes et donc à la candidature à la présidentielle des différents candidats, nous obtenons les mots distinctifs suivants pour chaque candidat :

**Mots distinctifs** (par rapport au reste du corpus) :

1. arhaud : **travailleurs** (7), **candidate** (8), **camp** (7), **écrasante** (5), **occidental** (4).
2. dptaignan : **neutralité** (16), **plan** (16), **donbass** (11), **élections** (6), **constitutionnelle** (5).
3. hidalgo : **standwithukraine** (35), **paris** (43), **ville** (19), **kyiv** (9), **ballet** (7).
4. jadot : **l'embargo** (8), **faireface** (8), **total** (18), **pétrole** (15), **crimes** (18).
5. lassalle : **qu'a** (3), **interviewé** (3), **l'agriculture** (2), **directrice** (2), **dalila** (2).
6. lepen : **bfmpolitique** (5), **agriculteurs** (5), **faceabfm** (4), **bfmtv** (8), **chine** (6).
7. macron : **président** (10), **50** (6), **union** (4), **putin** (4), **thoughts** (3).
8. melenchon : **monsieur** (13), **réacteurs** (12), **nucléaires** (15), **casques** (10), **bleus** (10).
9. pecresse : **disqualifiés** (4), **zemmour** (6), **vrai** (4), **présider** (3), **pleine** (3).
10. poutou : **corse** (5), **poutou** (3), **d'yvan** (3), **colonna** (3), **campagne** (8).
11. roussel : **droite** (5), **oligarques** (11), **tracts** (3), **réclamer** (3), **forfait** (3).
12. zemmour : **pologne** (4), **frontières** (3), **sarkozy** (2), **relation** (2), **définir** (2).

Nous remarquons alors plusieurs éléments intéressants :

- Jean-Luc Mélenchon, dont la campagne a beaucoup porté sur la question du nucléaire en France, a pu encore une fois, à travers ce conflit, évoquer les « réacteurs nucléaires » avec l'attaque de la Russie contre la plus grande centrale européenne,
- Nathalie Arthaud, dont le parti se nomme « Lutte ouvrière », a également abordé le conflit sous le prisme des « travailleurs »,

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE

- Anne Hidalgo, actuellement maire de Paris, évoque naturellement les mesures qu'elle y met en place, à travers le Conseil de Paris, afin d'aider les réfugiés ukrainiens et mettre en avant le soutien qu'elle apporte à Kyiv, capitale de l'Ukraine,
- Yannick Jadot souligne sa prise de position au niveau économique par l'évocation d'un embargo et la référence à l'entreprise Total, figurant parmi les 4 plus grands groupes pétroliers dans le monde, qui a continué à faire affaire avec la Russie malgré le conflit,
- Jean Lassalle - dont le programme appuyait le désir de mettre en avant les campagnes françaises - et Marine Le Pen ont quant à eux interrogé l'impact du conflit sur l'agriculture française,
- Eric Zemmour, pour qui la question de l'intégrité et du respect des frontières est primordiale, a pu reprendre cet élément de langage et dénoncer la violation des frontières par la Russie,
- Roussel, candidat du parti communiste français, a pu apporter le champ lexical des classes sociales en évoquant les oligarques russes.

Cette liste n'est pas exhaustive mais elle permet d'appréhender que chaque candidat a pu, par l'avènement de cette guerre, mettre en exergue les positions de son parti et ainsi l'utiliser pour sa campagne présidentielle.

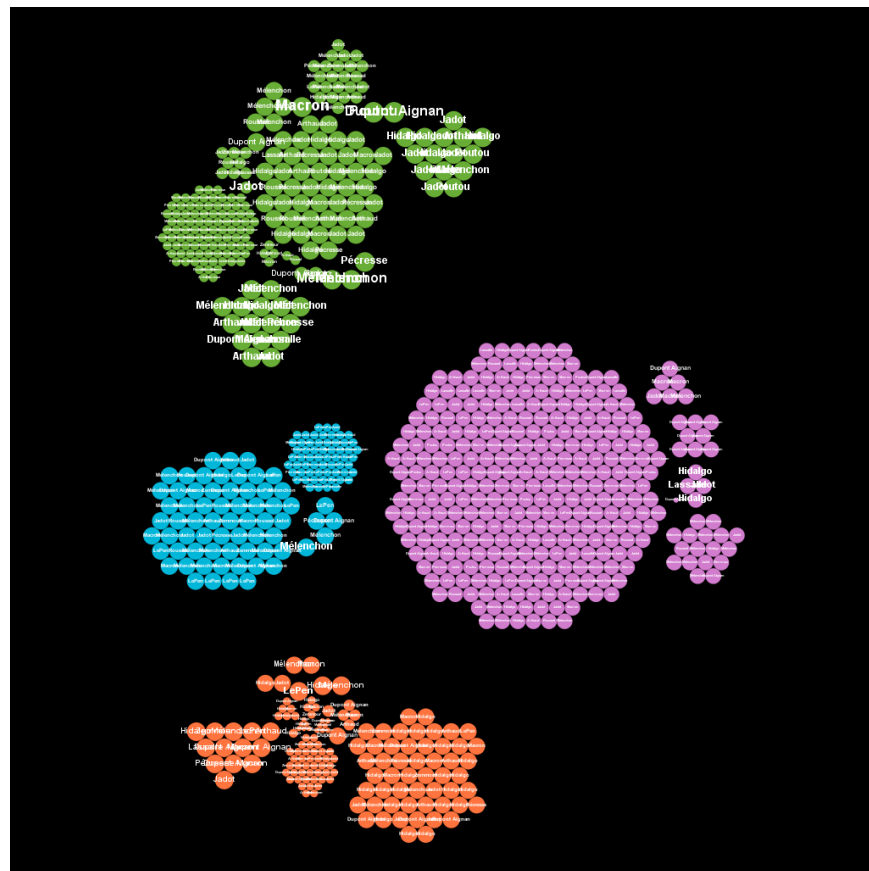
## RELATIONS ENTRE CANDIDATS

À l'aide du logiciel Gephi, nous avons enfin réalisé un graphique des relations entre les tweets des différents candidats à l'élection présidentielle. Pour cela, nous avons réalisé manuellement une table de nœuds et une table de liens entre les tweets. Chaque nœud correspond à un tweet et les liens correspondent à un ou plusieurs mots-clés communs entre deux tweets parmi une liste prédéfinies : « Ukraine », « Russie », « Zelensky », « Poutine », « Donbass », « Kiev », « Ukrainien », « UkraineRussiaWar », « UkraineRussie », « UkraineInvasion » et

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE

« Invasion ». Enfin, nous avons pondéré la relation entre deux tweets en fonction du nombre de mots-clés communs. C'est-à-dire que si deux tweets ont trois mots-clés en commun, alors leur relation sera plus forte que des tweets n'ayant qu'un seul mot-clé en commun.

Après avoir réalisé un graphique de relation clustérisé à l'aide de la représentation « Circle pack layout », nous obtenons le résultat suivant :

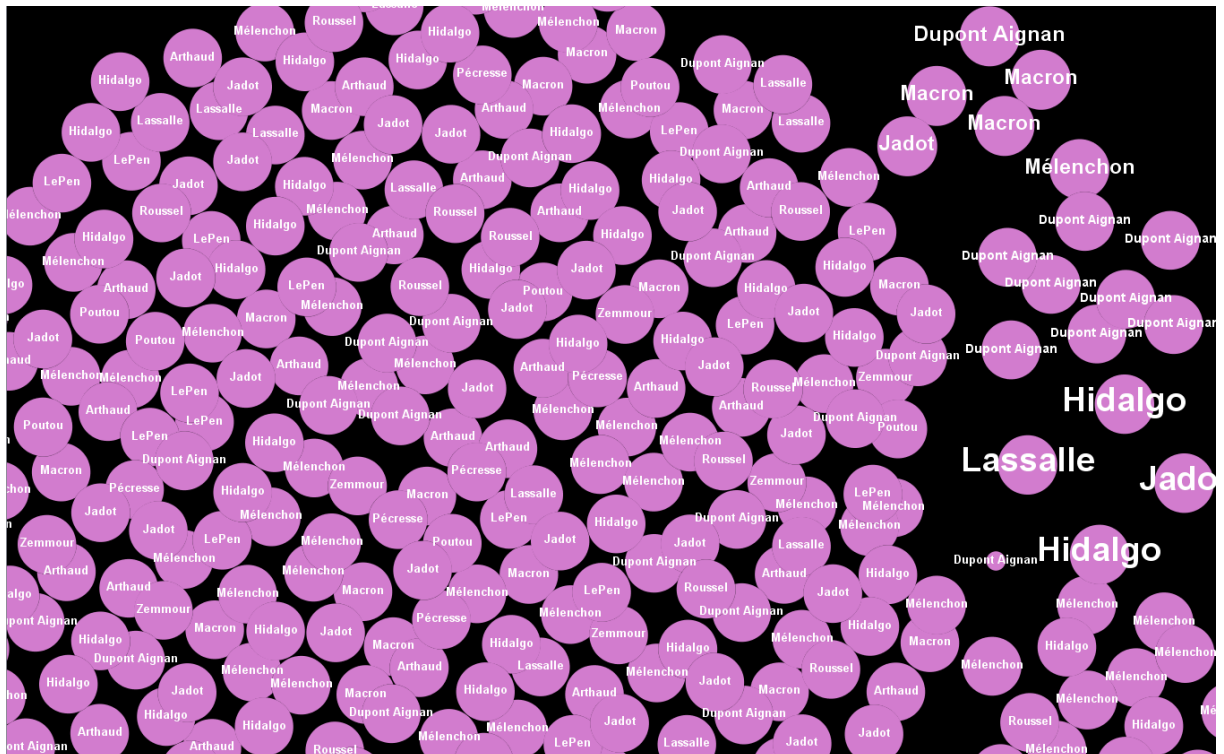


Par confort visuel, nous avons retiré l'affichage des liens entre les différents nœuds, car leur nombre conséquent rendait le graphique difficilement lisible. A première vue, nous distinguons 4 clusters dont un plus important et central dans le réseau (en violet). Ceux-ci ont été réalisés par modularité, c'est-à-dire en fonction de l'intensité des liens entre les individus au sein du cluster par rapport aux nœuds à l'extérieur du cluster. Cela signifie que les individus

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE

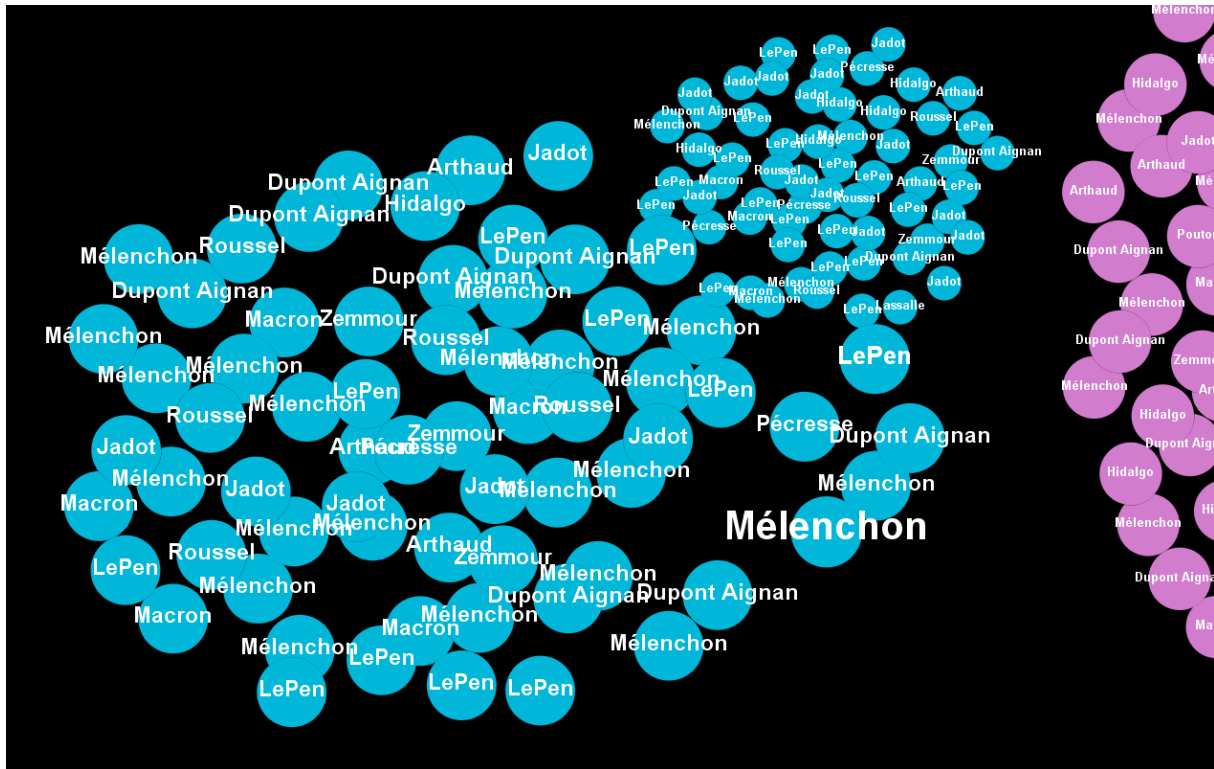
dans un même cluster ont plus de liens (donc plus de mots-clés en commun) entre eux qu'avec l'extérieur.

Observons maintenant plus en détail chaque cluster pour tenter d'en interpréter le sens :



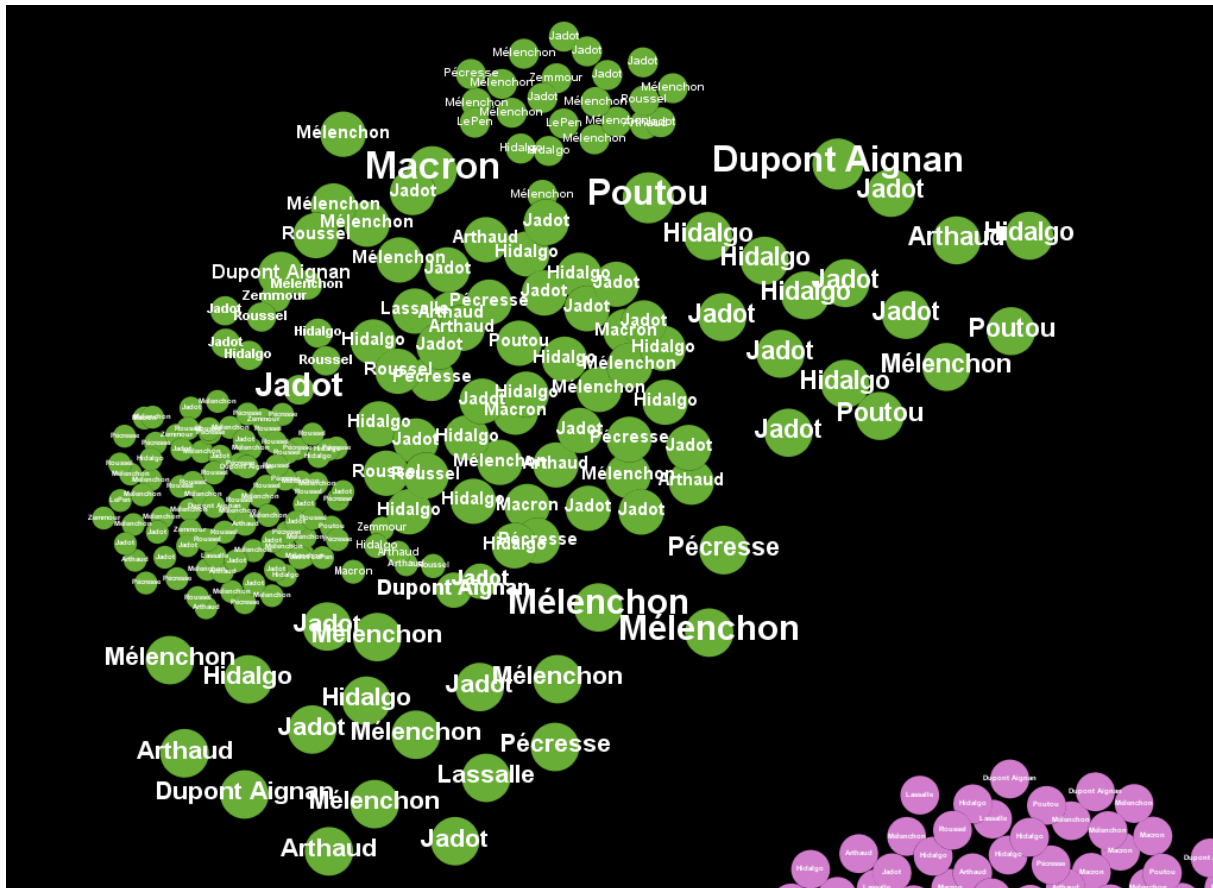
Les labels affichés sur chaque nœud indiquent le nom du candidat ayant publié le tweet en question et la taille du nom définit l'importance du tweet au sein du cluster. Au sein de ce grand regroupement, nous retrouvons le nom de tous les candidats et aucun d'entre eux ne se détache particulièrement. Ainsi, nous pouvons estimer que ce cluster principal regroupe les tweets de tous les partis, sans distinction sur la scène politique, du moment qu'ils partagent un mot-clé en commun. Parmi ce cluster, nous retrouvons sûrement des mots neutres tels que « Ukraine », mot qui revient 551 fois parmi les 816 tweets postés par les candidats durant notre période d'étude.

# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE



Dans ce cluster bleu, nous retrouvons notamment les noms de LePen, Dupont Aignan ou encore Zemmour, soit des partis jugés à l'« extrême droite » de la frise politique. Mais paradoxalement, un autre nom particulièrement présent est celui de Jean-Luc Mélenchon, pourtant situé à l'opposé de la scène. Ainsi, il est tout à fait possible que ce cluster présente les sujets faisant l'objet de clivage entre les partis de gauche et d'extrême droite. Nous savons notamment que ces sujets portent sur « Poutine » (222 occurrences) et la « Russie » (193 occurrences) de par sa proximité avec le rassemblement national dont il a soutenu le financement de campagne lors des élections présidentielles en 2017. Ce cluster atteste ainsi des dissensions politiques que l'on pouvait supposer sur un sujet de politique extérieure tel que celui-ci. N'excluons toutefois pas la possibilité que ces différents partis se rejoignent, de façon tacite, sur certaines idées communes.

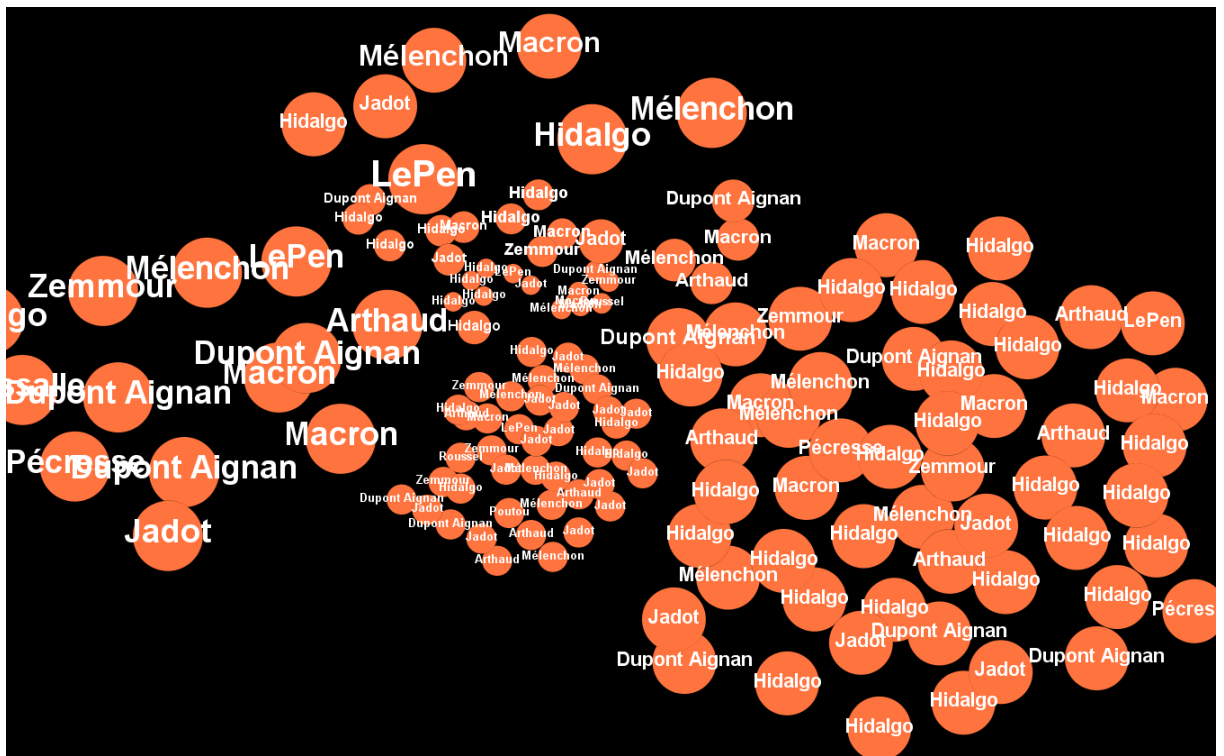
# ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE



Ce cluster vert et conséquent affiche presque exclusivement des noms de candidats jugés à « gauche » : Mélenchon, Poutou, Hidalgo ou encore Poutou. Ayant un avis plus tranché sur les actions russes et leur condamnation, nous pouvons supposer qu'ils se retrouvent à travers les mots-clés « Invasion » (34 apparitions), « UkraineInvasion » ou également « Poutine » à titre de critiques. Ainsi, il semblerait que l'ensemble de la gauche politique partage un certain nombre d'idées communes sur le conflit, que le reste des candidats ne partagent pas publiquement.



## ANALYSE DU DEUXIEME CORPUS – CANDIDATS A L'ELECTION PRESIDENTIELLE



Enfin, au sein de ce cluster orange, nous retrouvons une fois de plus le nom de nombreux candidats, sans que l'un d'eux ne se distingue particulièrement. Nous pouvons toutefois souligner la quasi-absence de l'extrême droite dans ce cluster, à l'exception de Nicolas Dupont-Aignan, dont le nom est récurrent et écrit en gros (à gauche sur le graphique). Ainsi, nous pouvons supposer qu'il s'agit d'un cluster regroupant des tweets peu clivants, mais probablement différents de ceux du cluster principal. Nous pouvons notamment penser à « Kiev » (15 occurrences), « Dombass » (14 occurrences), « Zelensky » (30 occurrences) qui sont des mots relativement neutres. Le mot « Ukrainien », présent dans 161 tweets, peut également se retrouver dans ce cluster, témoignant de l'empathie de la part de certains candidats envers le peuple victime de l'assaut. Cela pourrait expliquer la quasi-absence des candidats d'extrêmes droite, plus modérés quand aux agissements de la Russie.



## Conclusion

Pour conclure, nous avons véritablement pu observer, à travers notre étude, la perception et les comportements des utilisateurs de Twitter francophones vis-à-vis de la guerre en Ukraine. Sur le plan social, les comptes les plus actifs sont, en large majorité, des groupes de presse, indépendants ou non. Cependant, les utilisateurs les plus « likés » ne sont pas les plus prolifiques en termes de nombre de tweets. Ainsi, l'activité d'un compte est à décorréler de sa popularité. Notons également que, sur le plan social, le conflit est un sujet largement discuté et commenté avec ferveur par les utilisateurs du réseau, comme en témoigne la longueur moyenne des tweets à ce sujet. Il est en effet rare qu'un tweet soit rédigé en moins de 100 caractères sur les 300 autorisés, et la majorité des publications avoisinent même les 270 caractères.

Sur le plan politique, nous avons pu constater que cette guerre européenne en pleine campagne présidentielle a déclenché entre les candidats une véritable guerre idéologique. Certains candidats se sont ainsi particulièrement impliqués dans le conflit, à l'instar d'Emmanuel Macron, alors président français, ou Jean-Luc Mélenchon, particulièrement actif sur le réseau et ayant publié plus de 150 tweets sur la guerre en 2 mois. L'étude lexicale du corpus nous a également permis de constater, d'un point de vue politique, la condamnation des actes de Vladimir Poutine en personne. De la sorte, nous avons remarqué la construction d'un discours visant à opposer Poutine à l'Ukraine, et non un pays face à un autre. Puis, lorsque l'on se concentre sur le discours de chaque candidat, nous avons pu constater que la guerre a été une « opportunité » pour ces derniers de véhiculer leurs idées de campagne, parfois au détriment d'un réel soutien envers les victimes du conflit.

Cependant, nous l'avons dit, notre étude comporte des limites. L'observation du corpus général se concentre sur une semaine donnée et non sur les deux mois de notre étude. De ce fait, nous ne pouvons mettre en lumière les potentielles évolutions de tendances sur le sujet bien que le conflit ait évolué. Notre axe d'étude ne permet pas, notamment, de souligner certains moments phares ayant suscité de vives réactions : l'annonce du siège de Kiev, le bombardement de

## CONCLUSION

certaines grandes métropoles ukrainiennes, ou encore la publication de reportages télévisés sur le sujet à des heures de grande écoute début avril. Afin d'améliorer cet axe de notre étude, nous aurions pu, par exemple, effectuer des extractions quotidiennes durant ces 61 jours. Ainsi, nous aurions mis en lumière les sujets émergents ou en déclin concernant le conflit. Il serait toujours possible de conduire un tel projet aujourd'hui, en souscrivant à l'API premium de twitter – malheureusement particulièrement onéreuse – afin d'extraire des tweets sans restriction de durée.

Notons pour finir que cette guerre a tant fait, et continue de faire réagir les Français, de par sa proximité avec notre pays. La situation est inédite au XXIème siècle, où les réseaux sociaux et le web accentuent de façon exponentielle la vitesse de transmission des informations, quelles qu'en soient la fiabilité. Malheureusement, de nombreux conflits similaires subsistent et durent depuis bien longtemps aux quatre coins du globe. Il s'agit là d'un point de notre étude à souligner : l'implication d'une population, ainsi que l'empathie qu'elle développe vis-à-vis d'une lutte armée dépend de sa proximité avec celle-ci et de l'identification à sa population. Peut-être serait-il bon que les mentalités évoluent sur cette question si l'on aspire un jour à une cohabitation pacifique à l'échelle mondiale.