

Les drops de Q à l'épreuve des faits

Étude statistique des messages de « Q » (2017-2022) confrontés à l'actualité

Sylvain Geffroy

<https://s-geffroy.github.io/QAnon/> · doi:10.5281/zenodo.23473751 · Licence CC BY 4.0

Abstract. Between October 2017 and November 2022, an anonymous author signing "Q" posted 4,966 messages ("drops") on 4chan, 8chan and 8kun, the founding texts of the QAnon conspiracy theory. Most research studies QAnon's followers and platforms. We study the drops themselves. We check what they say against what actually happened.

Nothing suggests that Q anticipated the news. Trump's tweets do not follow the drops more often than chance once local activity is taken into account (4.2% vs 4.0%, $p = 0.23$). Of the checkable predictions, 28% came true (95% CI 21–36%), but only 16% once scheduled events are set aside. None of the announced arrests or indictments happened. In a blind "horoscope test", a language model must say which of two news lists goes with a drop. Three times out of four it finds no specific link. When it does, it favours the real period for the past and the future alike, and a second model sees a specific link in only 5 of 100 "future" comparisons. The timing of the drops is better described by a model of earthquake aftershocks than by a news feed. That model attributes most drops to earlier drops and almost none to Trump's tweets. Against the wider scene (six subreddits sampled every active day from 2016 to 2023, Wikipedia, Google Trends, 4chan's /pol/), Q draws on a small number of the scene's themes, and we detect no precedence either way at the scales we tested. Its own forums, though, react within hours. On /pol/, which we read in full (134.5 million posts), Q moved into threads already opened by Trump's "calm before the storm". It kept their conspiracies but not their explicit hate vocabulary, and its threads are calmer than the rest of the board.

The study combines descriptive statistics, text analysis, permutation tests, GDELT news data and a sourced check of 250 predictions. All data, code and verdicts are public.

Résumé

Entre octobre 2017 et novembre 2022, un auteur anonyme signant « Q » a publié 4 966 messages (« drops ») sur 4chan, 8chan puis 8kun, les textes fondateurs de la théorie du complot QAnon. La plupart des travaux étudient les adeptes de QAnon et ses plateformes. Nous étudions les drops eux-mêmes. Nous confrontons ce qu'ils disent à ce qui s'est réellement passé.

Rien n'indique que Q ait anticipé l'actualité. Les tweets de Trump ne suivent pas les drops plus souvent que le hasard une fois l'activité locale prise en compte (4,2 % contre 4,0 %, $p = 0,23$). Sur les prédictions vérifiables, 28 % se sont réalisées (IC 95 % 21–36 %), mais seulement 16 % une fois écartés les événements déjà programmés. Aucune des arrestations ou inculpations annoncées n'a eu lieu. Dans un « test de l'horoscope » à l'aveugle, un modèle de langage doit dire laquelle de deux listes d'actualité va avec un drop. Trois fois sur quatre, il ne trouve aucun lien précis. Quand il en trouve un, il préfère la vraie période pour le passé comme pour l'avenir, et un second modèle ne voit de lien précis que dans 5 comparaisons « avenir » sur 100. Le rythme des drops est mieux décrit par un modèle de répliques sismiques que par un fil d'actualité. Ce modèle attribue la plupart des drops à des drops antérieurs et presque aucun aux tweets de Trump. Face à la scène complotiste (six subreddits échantillonnés chaque jour actif de 2016 à 2023, Wikipédia, Google Trends, /pol/ de 4chan), Q ne reprend qu'un petit nombre des thèmes de la scène, et nous ne détectons aucune précédence, dans un sens ou dans l'autre, aux échelles testées. Ses propres forums, eux, réagissent dans les heures qui suivent. Sur /pol/, que nous

avons lu en entier (134,5 millions de posts), Q s'installe dans des fils déjà ouverts par le « calme avant la tempête » de Trump. Il en garde les complots, mais pas le vocabulaire de haine explicite, et ses fils sont plus calmes que le reste du forum.

L'étude combine statistiques descriptives, analyse textuelle, tests par permutation, données d'actualité de GDELT et vérification sourcée de 250 prédictions. Données, code et verdicts sont publiés.

Mots-clés : QAnon, théories du complot, prophéties, analyse textuelle, GDELT, modèles de langage, vérification des faits, processus de Hawkes, chaînes de Markov.

1. Introduction

QAnon est né le 28 octobre 2017 sur le forum /pol/ de 4chan, dans un milieu qui avait déjà produit collectivement la théorie du « Pizzagate » (Tuters et al., 2018). Un auteur anonyme, se présentant comme un initié du gouvernement américain détenteur d'une habilitation « Q », y annonçait qu'une cabale d'élites pédocriminelles serait bientôt démantelée par Donald Trump. Le mouvement s'est diffusé d'une plateforme à l'autre (de Zeeuw et al., 2020), a gagné une partie de l'électorat républicain (Sommer, 2023) et a été associé à l'assaut du Capitole du 6 janvier 2021 (Jensen & Kane, 2024).

La recherche a beaucoup étudié les adeptes – leur nombre (Enders et al., 2022 ; PRRI Staff, 2024), leur psychologie (Bloom & Moskalenko, 2021), leurs pratiques de « recherche » (Marwick & Partin, 2024) – et les plateformes. Les drops eux-mêmes ont été analysés surtout qualitativement (Linvill et al., 2021 ; McIntosh, 2022) ou sous l'angle de l'attribution d'auteur (Papasavva et al., 2022 ; Cafiero & Camps, 2023). À notre connaissance, il manque une confrontation systématique, sur tout le corpus, entre ce que Q a écrit et annoncé et ce qui s'est réellement passé.

Ce document de travail propose cette confrontation, avec trois exigences : mesurer plutôt qu'illustrer, comparer chaque observation à une référence explicite (un modèle nul, l'actualité mondiale, une période tirée au hasard), et publier toutes les données et tous les verdicts pour qu'ils puissent être discutés. Il accompagne un site bilingue où chaque chiffre est recalculé, sourcé et citable (<https://s-geffroy.github.io/QAnon/>).

2. Données

Le corpus. Il n'existe pas de collection officielle des drops, et les sites d'agrégation divergent à la marge (Papasavva et al., 2022). Nous utilisons le corpus JSON-QAnon (Kingsman, 2021) : 4 966 messages du 28 octobre 2017 au 27 novembre 2022 (heure de New York), avec leur forum, leurs images et les messages auxquels ils renvoient. Q a publié sur 533 jours distincts, avec un silence de 563 jours entre le 8 décembre 2020 et le 24 juin 2022.

Tweets de Donald Trump. Les tweets jusqu'au 8 janvier 2021 (CodeBreaker619, 2021), horodatés en UTC, soit environ 12,7 tweets par jour sur la période.

Actualité mondiale. Les fichiers quotidiens de GDELT 1.0 (The GDELT Project, s.d.) du 1er octobre 2017 au 31 janvier 2021 : 1 219 jours, 181 749 399 événements, agrégés par acteur et par pays. Pour le test de l'horoscope, nous utilisons aussi le portail « Current events » de Wikipédia, résumé éditorial quotidien de l'actualité avec sources : 17 701 items sur la même période.

Intérêt public. Les pages vues de 15 articles de Wikipédia (dont « QAnon » en onze langues) via l'API Wikimedia (Wikimedia Foundation, s.d.).

La scène complotiste américaine. Six subreddits (r/conspiracy, r/conspiracytheories, r/The_Donald, r/greatawakening, r/CBTS_Stream, r/QAnonCasualties), y compris ceux qui ont été bannis, via l'archive Arctic Shift : volumes quotidiens exacts de 2016 à 2023, activité horaire des deux forums de Q, et un échantillon quotidien systématique de messages, sans nom d'auteur (2 558 052 messages, tous les 8 946 jours actifs). Les pages vues de 25 articles de Wikipédia sur les grands thèmes de la scène (redirections comprises). L'intérêt de recherche Google (États-Unis) pour ces thèmes. Le forum /pol/ de 4chan de juin 2016 à novembre 2019 (Papasavva et al., 2020), lu une seule fois et réduit à des agrégats. Un lexique versionné de 459 motifs (formules de Q, grands récits complotistes, codes de haine, culture de /pol/, milices, élections, santé) sert aux mêmes comptages dans toutes ces sources. Sur /pol/, les 34 motifs ajoutés après sa lecture sont mesurés dans les fils Q, dans un échantillon de 1 % des posts et, pour les mots isolés, dans les comptes mensuels de mots de tout le forum.

Données éditoriales. Une chronologie de 31 événements datés et sourcés, un dictionnaire de 199 acteurs (motifs de reconnaissance, codes pays), un glossaire de 47 formules et une bibliographie vérifiée, tous versionnés avec le code.

3. Méthodes

Rythme. Intervalles entre drops consécutifs, coefficient de rafale B et coefficient de mémoire M (Goh & Barabási, 2008) ; sessions de drops séparés de moins de 30 minutes ; régimes d'activité par partition optimale des comptes hebdomadaires (Jackson et al., 2005 ; Killick et al., 2012), avec un coût de Poisson corrigé de la surdispersion et une pénalité BIC.

Discours. Thèmes par factorisation en matrices non négatives (Lee & Seung, 1999) des poids TF-IDF ; termes distinctifs de chaque année par log-odds avec a priori de Dirichlet informatif (Monroe et al., 2008) ; stylométrie exploratoire par le Delta de Burrows (Burrows, 2002) sur des blocs d'environ 1 500 mots ; lisibilité de Flesch (Flesch, 1948) et richesse lexicale MATTR (Covington & McFall, 2010).

Q et les tweets de Trump. Q affirmait que des tweets publiés quelques minutes après un drop en confirmaient l'authenticité (les « deltas »). Pour chaque fenêtre (1, 5, 10 et 60 minutes) et chaque sens, nous comparons la part de drops suivis (ou précédés) d'un tweet à sa distribution sous deux modèles nuls, par 1 000 permutations (Phipson & Smyth, 2010) : un modèle « jours », qui déplace chaque drop à la même heure d'un autre jour, et un modèle « local », qui le déplace de quelques heures et conserve donc l'activité du moment. Le premier surstime la synchronie, car Q et Trump publient tous deux davantage les jours de forte actualité. C'est le second qui juge la revendication.

Acteurs et actualité. Pour chaque acteur du dictionnaire, la sur-représentation est le rapport entre sa part dans les mentions de Q et sa part dans les événements GDELT. Nous calculons aussi des corrélations hebdomadaires décalées et des tests de causalité de Granger (Granger, 1969), sans correction pour comparaisons multiples (ils sont exploratoires).

Prédictions. Chaque drop non vide est soumis à un modèle de langage, avec une consigne et un schéma de réponse stricts, qui doit dire s'il contient une prédiction et laquelle (paraphrase, type d'événement, sujets, échéance, précision). Le modèle a été choisi parmi sept, sur un jeu de référence de 150 drops annotés à la main selon des consignes publiées (36 positifs), avec un intervalle bootstrap (Efron, 1979) sur le F1. Nous avons re-

tenu gpt-5.6-luna ($F1 = 0,86$; IC 0,78–0,94), pour un coût de 0,59 \$ sur le corpus, contre 0,90 et 4,34 \$ pour le meilleur modèle, l'écart n'étant pas significatif.

Vérification. Le modèle classe chaque prédiction extraite en précise ou vague. Les 290 annonces précises, ou vagues mais assorties d'une échéance datée, ont été vérifiées une à une, et celles qui se révèlent ne pas être des prédictions sont écartées. L'horizon est l'échéance énoncée, sinon douze mois après le drop. Chaque verdict (réalisée, partielle, non réalisée, invérifiable, pas une prédiction) est justifié et sourcé. Une prédiction est « triviale » si l'événement était déjà annoncé ou programmé quand Q l'a écrite. Les taux sont donnés avec des intervalles de Wilson (Wilson, 1927).

Test de l'horoscope. Si les drops anticipaient l'actualité, un lecteur devrait reconnaître, parmi deux périodes, celle qui suit réellement le drop. Pour 300 drops tirés au sort (au moins 15 mots, stratifiés par année), un modèle de langage reçoit, sans aucune date, le texte du drop et deux listes d'items d'actualité liés aux États-Unis, de même longueur et dans un ordre tiré au sort : la vraie période et une période tirée au hasard à plus de six mois du drop. Deux comparaisons sont faites : les 7 jours qui précèdent le drop (le drop commente-t-il l'actualité ?) et les 30 jours qui suivent (l'annonce-t-il ?). La consigne est sceptique : une formulation vague ne compte pas. Le modèle note chaque liste de 0 à 10 et choisit l'une d'elles, ou aucune. Nous testons la part de bons choix parmi les cas tranchés contre 50 %. Un second modèle repose les mêmes questions pour un tiers des drops.

Modèles stochastiques. Les comptes quotidiens de drops sont décrits par un modèle de Markov caché à émissions binomiales négatives (Rabiner, 1989), dont le nombre d'états est choisi par le BIC. Les thèmes NMF successifs d'une même session forment une chaîne de Markov, dont la persistance est comparée à 1 000 permutations des thèmes au sein des sessions. La dérive du style est l'entropie croisée d'un modèle de caractères d'ordre 3 entraîné sur un semestre et évalué sur le suivant, complétée par une distance de compression (Benedetto et al., 2002). Les horodatages des drops sont ensuite traités comme une séquence sismique : un processus de Hawkes (Hawkes, 1971) à noyau exponentiel ou d'Omori (Ogata, 1988), sur un fond propre à chaque régime avec cycle quotidien, estimé par l'algorithme EM (Veen & Schoenberg, 2008) et jugé par l'AIC et par l'adéquation des résidus après changement de temps. Un terme croisé mesure la part des drops que le modèle attribue aux tweets de Trump, comparée à 28 décalages témoins de -14 à $+14$ jours. Nous estimons aussi la mémoire longue par analyse des fluctuations redressées (Peng et al., 1994), la loi des tailles de sessions (Clauset et al., 2009 ; Vuong, 1989) et la phase du cycle quotidien de publication.

Q dans la scène. Pour chaque thème de la scène que Q emploie dans au moins dix drops, la part hebdomadaire de drops qui le mentionnent est comparée à l'attention de la scène (Reddit, Wikipédia, Google Trends) par corrélation croisée de -8 à $+8$ semaines après pré-blanchiment, avec 300 substituts IAAFT (Schreiber & Schmitz, 2000) et correction de Benjamini–Hochberg (Benjamini & Hochberg, 1995) : Q précède, suit, évolue en même temps, ou rien de significatif. L'entropie de transfert (Schreiber, 2000) cherche des flux d'information entre séries quotidiennes, une fois retirés les cycles hebdomadaires. La réponse des forums de Q à ses drops est mesurée heure par heure par analyse d'époque superposée, contre des instants tirés au hasard. Les plus grands pics d'attention de la scène sont classés en chocs extérieurs ou cascades internes selon la forme de leur montée et de leur retombée (Crane & Sornette, 2008). Sur /pol/, une formule « apparaît » la première semaine où son nombre de posts dépasse durablement le quantile 0,999 d'une loi de Poisson au taux d'avant Q. La toute première occurrence, souvent fortuite, n'est pas retenue. La toxicité est celle des scores Perspective fournis avec le jeu de données ; le vocabulaire propre aux fils Q est mesuré par log-rapport de cotes à prior de Dirichlet informatif (Monroe et al., 2008).

Prophéties recyclées et autocitations. Les prédictions sont regroupées par sujet (même type d'événement, même acteur) ; un sujet annoncé sur au moins trois mois distincts est « recyclé ». Nous mesurons enfin le délai des renvois de Q vers ses propres drops, qui sous-tendent la formule « Future proves past ».

4. Résultats

4.1 Un rythme de rafales, des régimes liés aux plateformes

Q ne publie pas au fil de l'eau mais par rafales : le coefficient B vaut 0,91 (0 pour un processus de Poisson). L'intervalle médian entre deux drops est de 14 minutes, et 26 % des drops suivent le précédent de moins de 5 minutes, 75 % de moins d'une heure. Les drops se regroupent en 1 743 sessions, et 85 % d'entre eux appartiennent à une session de plusieurs messages. Le record est de 64 drops le 28 mars 2019. Le volume hebdomadaire se découpe en six régimes (figure 1). Une phase intense dure jusqu'à l'automne 2018 (47 drops par semaine en moyenne). Après un creux vient un pic en février et mars 2019 (89 par semaine). Suit une phase faible d'avril 2019 au printemps 2020 (11 par semaine), qui couvre la fermeture de 8chan en août 2019 et le retour sur 8kun en novembre. L'activité reprend au printemps 2020 avec la pandémie (35 par semaine), puis s'éteint après l'élection de novembre 2020.

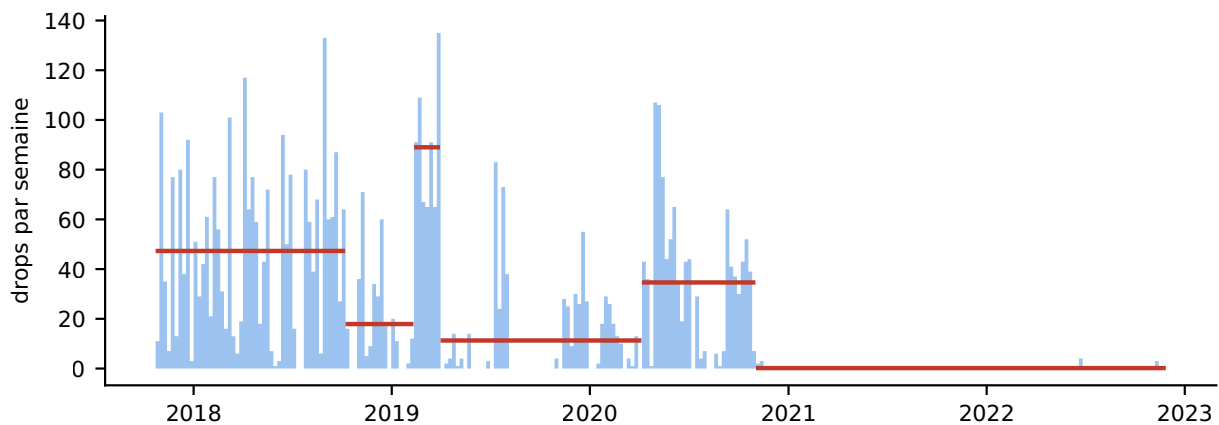


Figure 1. Drops par semaine (heure de New York) et moyenne de chacun des régimes d'activité détectés.

4.2 Une écriture qui délègue l'interprétation

Un drop contient en moyenne 1,39 point d'interrogation, et 38 % en contiennent au moins un. Les phrases sont courtes (7 mots en moyenne, indice de Flesch 65). Le thème dominant de la factorisation (43 % du poids) mêle le président (« POTUS »), le contrôle et le public. D'autres thèmes regroupent les formules rituelles (« fake news », « news unlocks », « future proves past », « God bless », « do you believe in coincidences? »). Q renvoie à un message d'anonyme dans 23 % des drops et à l'un de ses propres drops dans 8 %, signe d'une écriture en dialogue avec des lecteurs chargés de « faire leurs recherches », conformément à ce que décrivent les analyses qualitatives (McIntosh, 2022 ; Marwick & Partin, 2024). Le vocabulaire religieux ou moral reste minoritaire (6,3 % des drops) et ne progresse pas : il culmine dans les premières semaines.

La stylométrie exploratoire place la rupture de style la plus nette autour du 25 novembre 2017, c'est-à-dire au premier drop sur 8chan. Cela rejoint les études qui concluent à au moins deux plumes successives (OrphAnalytics, 2020 ; Cafiero & Camps, 2023), mais notre analyse ne permet pas de distinguer un changement d'auteur d'un changement de plateforme, et elle n'identifie personne.

4.3 Pas de synchronie avec les tweets de Trump

Avec le modèle nul local, un tweet de Trump suit un drop en moins de 5 minutes dans 4,2 % des cas, contre 4,0 % attendus ($p = 0,23$). Aucune fenêtre ne montre d'excès dans le sens « tweet après le drop ». Le modèle « jours » produit des écarts significatifs, mais dans les deux sens, y compris quand le tweet précède le drop, ce qui correspond à la co-activité des jours chargés plutôt qu'à une coordination (tableau 1).

Tableau 1. Tweets de Trump à proximité des drops, modèle nul local (1 000 permutations).

Fenêtre	Sens	Observé	Attendu	p
1 min	tweet après le drop	1,0 %	0,9 %	0,186
5 min	tweet après le drop	4,2 %	4,0 %	0,227
10 min	tweet après le drop	7,4 %	7,3 %	0,442
30 min	tweet après le drop	16,5 %	17,6 %	0,995
60 min	tweet après le drop	27,9 %	28,8 %	0,948
1 min	tweet avant le drop	0,7 %	0,9 %	0,887
5 min	tweet avant le drop	3,8 %	4,0 %	0,739
10 min	tweet avant le drop	7,4 %	7,4 %	0,463
30 min	tweet avant le drop	18,2 %	17,8 %	0,197
60 min	tweet avant le drop	30,9 %	29,9 %	0,041

Dans l'autre sens, Q réagissait à Trump : 59 drops citent un tweet présidentiel, en médiane 1 h 54 après sa publication.

4.4 Un choix d'acteurs très sélectif

Les acteurs les plus nommés sont Donald Trump (386 drops), les « médias » (241) et Barack Obama (166) ; 33 % des drops nomment au moins un acteur du dictionnaire. Rapportée à leur place dans l'actualité mondiale, la sur-représentation est massive pour George Soros ($\times 15,8$), Hillary Clinton ($\times 8,0$) et le Parti démocrate ($\times 3,1$), tandis que le Congrès ($\times 0,2$) ou l'Union européenne ($\times 0,45$) sont sous-représentés. Les pays cités par Q suivent pourtant l'actualité (corrélation de rang 0,61) : Q sélectionne donc ses acteurs bien plus que les pays dont il parle (figure 3).

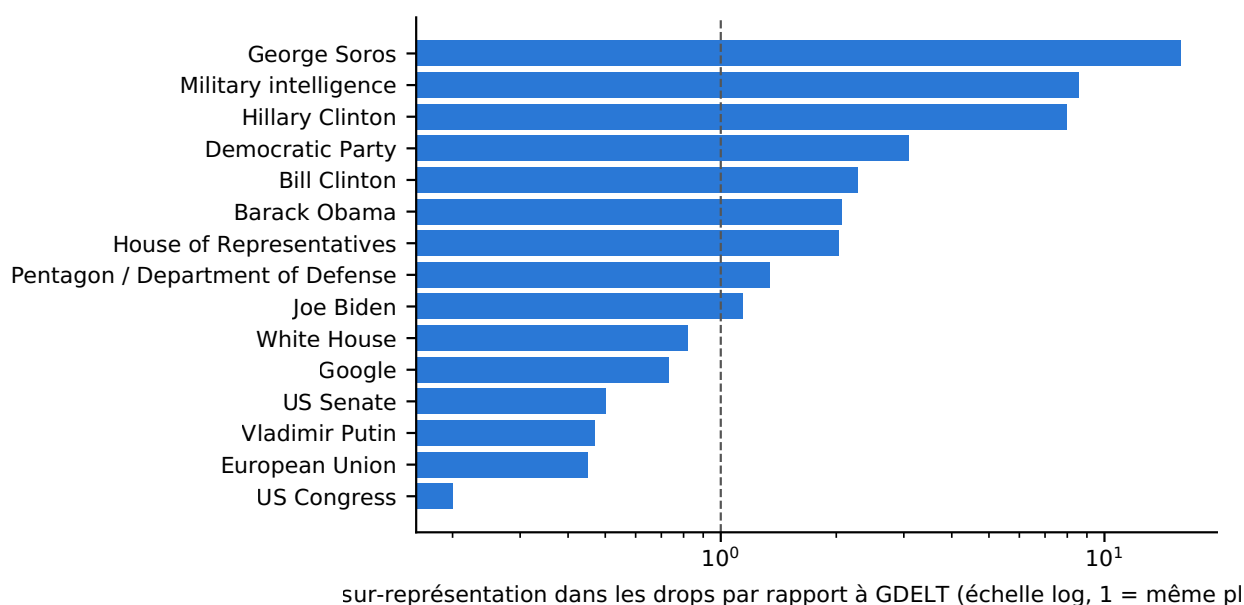


Figure 3. Place des principaux acteurs dans les drops rapportée à leur place dans l'actualité mondiale.

Les tests de Granger ne montrent pas que l'actualité d'un acteur précède les mentions de Q (une seule p-valeur sur quinze sous 0,05, ce qu'on attend du hasard).

4.5 Des prédictions rares, proches et rarement réalisées

Le modèle trouve au moins une prédiction dans 13 % des drops (598), soit 846 prédictions, dont 203 précises et 643 vagues. Les échéances datées sont proches : médiane de 7 jours, la moitié à moins d'une semaine. Sur les 290 annonces vérifiées, 40 se révèlent ne pas être des prédictions, et 110 des 250 restantes sont invérifiables. Des 140 prédictions vérifiables, 39 se sont réalisées, soit 28 % (IC 95 % : 21–36 %), et 22 en partie. Hors événements déjà annoncés ou programmés, le taux tombe à 16 % (IC 95 % : 10–24 %). Aucune des 8 arrestations ou inculpations vérifiables ne s'est produite, ni aucune des 10 déclassifications annoncées. Les réussites portent surtout sur des démissions (7 sur 14) et des décisions de justice déjà attendues (3 sur 3) (tableau 2).

Tableau 2. Prédictions vérifiables par type d'événement annoncé.

Type	Vérifiables	Réalisées	Partielles	Non réalisées	Taux
Autre	55	20	13	22	36 %
Révélation	25	3	3	19	12 %
Démission	14	7	1	6	50 %
Déclassification	10	0	1	9	0 %
Action militaire	9	2	3	4	22 %
Élection	7	1	0	6	14 %
Économie	5	3	1	1	60 %
Inculpation	5	0	0	5	0 %
Arrestation	3	0	0	3	0 %
Attaque	3	0	0	3	0 %
Décision de justice	3	3	0	0	100 %
Décès	1	0	0	1	0 %
Ensemble	140	39	22	79	28 %

Dix-sept sujets sont annoncés sur au moins trois mois différents, et dans sept d'entre eux l'annonce revient après un échec déjà vérifié – un comportement attendu des groupes prophétiques (Festinger et al., 1956 ; Westmark & McMahon, 2022).

4.6 Le test de l'horoscope

Dans environ trois quarts des comparaisons, le modèle ne trouve aucun lien précis entre le drop et l'une ou l'autre période. Quand il tranche, il reconnaît la vraie période dans 75 % des cas pour les 7 jours qui précèdent le drop (IC 95 % : 63–84 %, $p < 0,001$) et dans 68 % pour les 30 jours qui suivent (IC 95 % : 56–77 %, $p = 0,004$). Les deux taux ne diffèrent pas significativement (test exact de Fisher, $p = 0,44$). Le signal « avenir » ne suffit pas à conclure à une anticipation : il ne concerne qu'une minorité de comparaisons, des périodes tirées au hasard « correspondent » elles aussi, et il est compatible avec la simple continuité des sujets d'une semaine à l'autre, hypothèse que ce test ne mesure pas. Une période tirée au hasard obtient une note d'au moins 5 sur 10 dans 6 % des cas (figure 2). Des drops assez vagues « collent » à des événements survenus des années plus tard, comme une allusion à la Cour suprême rattachée à la mort de Ruth Bader Ginsburg en 2020. Pour tester la robustesse, un second modèle (gpt-6.1-sol, meilleur F1 de l'évaluation sur notre jeu de référence), a repris les mêmes questions pour 100 drops. Il est bien plus sceptique. Il ne voit aucun lien précis dans 95 des 100 comparaisons « avenir », contre 70 pour le modèle principal sur les mêmes drops, et ne tranche que 5 fois (4 en faveur de la vraie période). Les deux modèles font le même choix dans 75 % des cas, mais leur accord au-delà du ha-

sard reste modeste (kappa de Cohen 0,23). Les rares « correspondances » dépendent donc du lecteur. C'est une limite de la mesure plus qu'une preuve supplémentaire, et elle interdit d'y voir une anticipation reconnaissable.

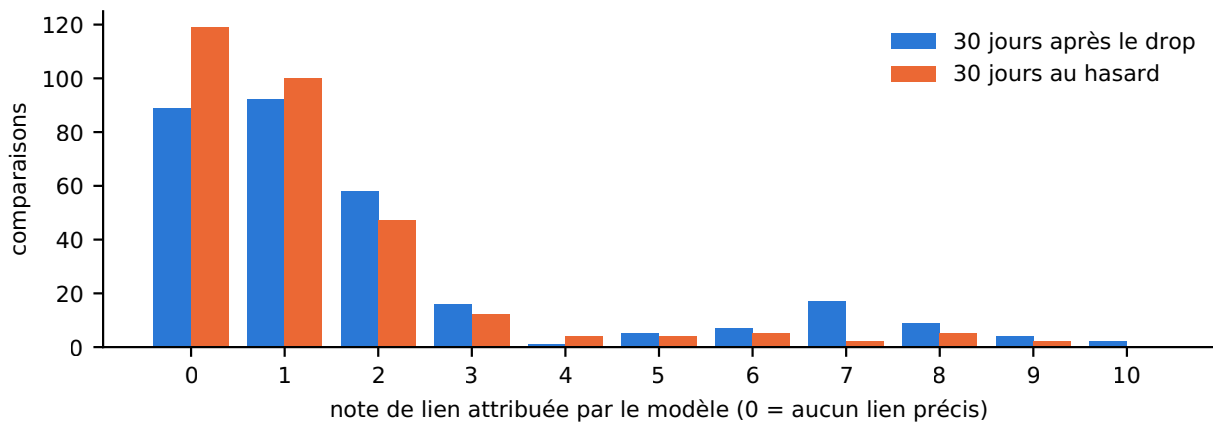


Figure 2. Test de l'horoscope : notes de lien pour la vraie période « avenir » et pour une période tirée au hasard.

4.7 L'intérêt public suit les événements, pas les drops

Les pages vues de l'article « QAnon » suivent à peine l'activité de Q : les corrélations hebdomadaires, avec des décalages de -4 à $+4$ semaines, restent faibles (0,17 au plus). Elles bondissent avec la visibilité du mouvement hors d'internet : $\times 13,3$ la semaine du meeting de Trump à Tampa (août 2018), où apparaissent les premières pancartes « Q » très médiatisées, et $\times 11,2$ après l'assaut du Capitole.

4.8 Un processus auto-excité, à mémoire longue, sans criticité

Le modèle de Markov caché retient 3 régimes, un silence, une veille et un régime actif (figure 4a). La veille compte 0,42 drop par jour, le régime actif 9,9. Ce dernier ne couvre que 26 % des jours, mais réunit 97 % des drops. Un séjour dure 7,0 jours en moyenne, pour 107 bascules en tout. Sur les 5 ruptures de la section 4.1, 3 tombent à moins d'une semaine d'une bascule du modèle, que rien ne contraint pourtant à les retrouver. Au sein des sessions, Q reste sur le même thème dans 38,6 % des enchaînements, contre 36,4 % si les thèmes étaient permutés dans chaque session ($p = 0,001$). L'ordre des drops n'est donc pas aléatoire, mais l'effet est faible (taux d'entropie de 2,38 bits contre 2,55 sans mémoire). Les deux plus fortes dérives du style tombent aux passages 2017 S2 $\rightarrow$ 2018 S1 (+0,56 bit par caractère) et 2019 S2 $\rightarrow$ 2020 S1 (+0,47), ce qui rejoint, avec une autre méthode, la rupture stylistique de la section 4.2.

Le processus de Hawkes décrit les horodatages bien mieux qu'un processus de Poisson, même doté du même fond par régime et du même cycle quotidien (écart d'AIC : 11 752). Le meilleur modèle attribue 80 % des drops à l'excitation par des drops antérieurs (75 % avec un noyau exponentiel). Le délai médian de déclenchement est de 17 minutes, et la retombée suit une loi d'Omori d'exposant 2,6, plus raide que celle des répliques sismiques. L'adéquation reste imparfaite (distance de Kolmogorov–Smirnov 0,044, contre 0,49 pour Poisson ; figure 4b). Le modèle n'attribue aux tweets de Trump que 0,5 % des drops, une part faible mais supérieure à celle de chacun des 28 décalages témoins ($p = 0,035$). Le lien avec Trump reste donc marginal dans ce modèle, ce qui rejoint les 59 citations de tweets de la section 4.3.

L'activité quotidienne a une mémoire longue (exposant de Hurst 0,81, contre 0,49 une fois les jours mélangés), du même ordre que celle des tweets de Trump (0,84). En revanche, rien n'indique une criticité auto-organisée : la queue des tailles de sessions (exposant 4,0 au-delà de 8 drops, 117 sessions sur 1 743) n'est pas distinguable

d'une loi lognormale ($p = 0,66$) ni d'une exponentielle ($p = 0,44$). Enfin, l'heure de publication bascule le 21 avril 2018, et le pic de publication passe de 20 h 48 à 17 h 12 (heure de New York ; $p = 0,005$ par permutation de la date de rupture). Mois après mois, la phase du cycle quotidien de Q reste accrochée à celle de Trump, environ 5 heures plus tard (valeur de verrouillage de phase 0,75). Ce changement d'horaire est compatible avec un changement de main ou de rythme de vie, sans que ces données permettent de trancher.

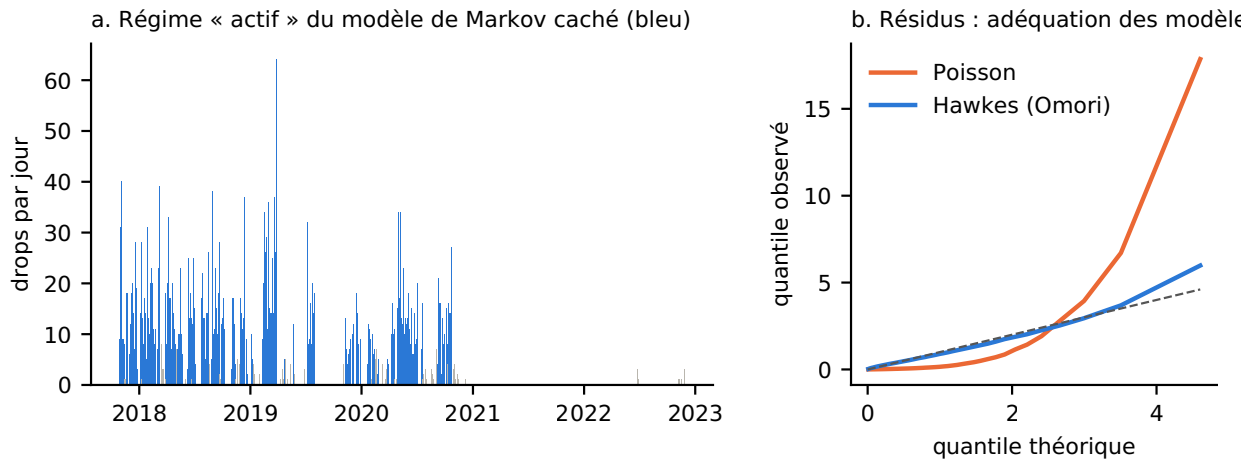


Figure 4. (a) Drops par jour ; les jours du régime « actif » du modèle de Markov caché sont en bleu. (b) Diagramme quantile-quantile des intervalles après changement de temps : un modèle bien ajusté suit la diagonale. Le processus de Poisson (même fond par régime et cycle quotidien) s'en écarte, le processus de Hawkes s'en approche.

4.9 Q dans la scène complotiste américaine

Sur 112 thèmes de la scène suivis, 66 n'apparaissent dans aucun drop et 35 dans moins de dix : Q puise dans un répertoire restreint. Pour les 15 comparaisons hebdomadaires possibles, aucune corrélation ne résiste à la correction pour comparaisons multiples. Aucune précedence n'est donc détectée, dans un sens ou dans l'autre, aux échelles testées. Le nombre de thèmes distincts abordés par trimestre culmine au quatrième trimestre 2017 (20) et au deuxième trimestre 2020 (21). Cette sélectivité prolonge celle des acteurs (section 4.4) : Q ne reprend pas l'ensemble du complotisme américain, il en retient quelques récits.

Entre 11 séries quotidiennes (106 liens orientés testés), aucun lien d'information ne reste significatif une fois retirés les cycles hebdomadaires et corrigées les comparaisons multiples. À l'échelle horaire, les forums de Q réagissent à ses drops : dans les six heures qui suivent une session, leur activité rapportée à la veille dépasse celle d'instant tirés au hasard de 9,6 % sur *r/CBTS_Stream* ($p = 0,008$, 118 sessions) et de 6,6 % sur *r/greatawakening* ($p = 0,012$, 395 sessions). Parmi les plus grands pics d'attention de chaque série de la scène (22 en tout), 4 ont la forme d'un choc extérieur et 2 celle d'une cascade interne ; les chocs extérieurs datés coïncident avec des événements de la chronologie (« 8chan mis hors ligne après la tuerie d'El Paso » ; « Assaut du Capitole » ; « Twitter suspend définitivement @realDonaldTrump »).

Le tableau 3 sépare deux cas. Sur les 15 formules de Q retrouvées sur Reddit, 5 y circulaient déjà avant lui, souvent comme expressions ordinaires de l'anglais (« The Great Awakening », « Enjoy the show », « Calm before the storm »). Les 10 autres n'y apparaissent qu'après le premier drop qui les emploie, de 0 à 28 jours plus tard (médiane 22 jours). Ces délais restent des bornes. L'échantillon couvre tous les jours actifs, mais pas tous les messages, et une formule absente de Reddit pendant quelques semaines pouvait circuler ailleurs, notamment sur */pol/* et *8chan*.

Tableau 3. Formules de Q : premier drop et première occurrence dans l'échantillon quotidien de Reddit (100 % des jours actifs échantillonnés ; un écart négatif signifie que la formule circulait avant Q).

Formule	Drops	Premier drop	Première occurrence sur Red-dit	Écart (jours)
Enjoy the show	67	29 oct. 2017	24 févr. 2016	-613
The Great Awakening	49	30 oct. 2017	22 févr. 2016	-616
Calm before the storm	8	2 nov. 2017	12 mars 2016	-600
Disinformation is necessary	8	4 nov. 2017	27 nov. 2017	23
Dark to light	28	12 nov. 2017	7 sept. 2016	-431
Sealed indictments	11	14 nov. 2017	28 mars 2016	-596
Expand your thinking	41	20 nov. 2017	18 déc. 2017	28
Future proves past	39	24 nov. 2017	22 déc. 2017	28
News unlocks map	12	24 nov. 2017	22 déc. 2017	28
Panic in DC	37	19 déc. 2017	13 janv. 2018	25
WWG1WGA	194	9 janv. 2018	9 janv. 2018	0
Trust the plan	28	5 févr. 2018	10 févr. 2018	5
Q+	36	19 févr. 2018	19 févr. 2018	0
Sheep no more	16	27 juin 2018	18 juil. 2018	21
Nothing can stop what is coming	16	21 nov. 2018	3 déc. 2018	12

4.10 /pol/, le berceau de Q

Le forum /pol/ de 4chan, où Q est apparu, a été lu en entier (Papasavva et al., 2020) : 134 529 233 posts dans 3 397 911 fils, dont 3 293 fils Q (405 183 posts). En novembre 2017, 6,9 % des posts de /pol/ sont écrits dans ces fils ; dès 2018, la part retombe autour de 0,12 % par mois. Sur 55 formules liées à Q assez fréquentes pour être datées, 16 étaient déjà en hausse avant le 28 octobre 2017, 4 naissent avec Q sans qu'il les emploie, 8 sont employées par /pol/ avant Q et 16 d'abord par Q (tableau 4). « CBTS » décolle dès la semaine du 2 octobre 2017, celle du « calme avant la tempête » lancé par Donald Trump le 5 octobre 2017 : Q s'installe dans des fils déjà ouverts. Les mentions de Q sur /pol/ augmentent dans les six heures qui suivent ses sessions, pendant sa période /pol/, jusqu'à son dernier drop sur /pol/ le 4 janvier 2018 (+0,54 par rapport au hasard, $p = 0,002$, 111 sessions) et après son départ pour 8chan, en 2018-2019 (+0,16 par rapport au hasard, $p = 0,002$, 1 059 sessions).

Tableau 4. Formules liées à Q sur /pol/ (au moins 500 posts) : taux par million de posts avant et après le 28 octobre 2017, semaine d'apparition (premier dépassement durable du taux d'avant Q) et premier drop qui emploie la formule.

Formule	Posts	Avant	Après	Apparition	Premier drop	Lecture
FBlanon (2016)	10 698	141,5	32,0	dès le début des données	—	déjà là avant Q
Autres « anons initiés »	2 216	26,6	8,7	1 août 2016	—	déjà là avant Q
Red pill	882 563	7 122,0	6 130,1	22 août 2016	19 janv. 2018	déjà là avant Q
HlIanon (High Level Insider)	805	9,5	3,3	29 août 2016	—	déjà là avant Q
Arrestation / inculpation d'Hillary Clinton annoncée	13 343	105,2	94,6	10 oct. 2016	28 oct. 2017	déjà là avant Q
Guerre des mêmes	31 816	335,5	160,6	7 nov. 2016	—	déjà là avant Q
Pizzagate	80 259	734,9	490,6	21 nov. 2016	—	déjà là avant Q

Formule	Posts	Avant	Après	Apparition	Premier drop	Lecture
					14 mai 2018	déjà là avant Q
WH insider anon	3 697	50,3	10,0	19 juin 2017	—	déjà là avant Q
Adrenochrome	5 364	16,9	57,5	3 juil. 2017	—	déjà là avant Q
Cabal	27 622	155,3	243,6	10 juil. 2017	—	déjà là avant Q
CBTS / « calme avant la tempête »	11 092	23,4	127,7	2 oct. 2017	2 nov. 2017	déjà là avant Q
MegaAnon	1 858	12,2	15,1	2 oct. 2017	—	déjà là avant Q
« The Plan to Save the World » / « Fall of the Cabal »	1 666	2,5	19,9	2 oct. 2017	2 nov. 2017	déjà là avant Q
White hats / Black hats	4 035	19,8	37,8	9 oct. 2017	11 avr. 2019	déjà là avant Q
Tribunaux militaires	3 495	7,0	40,5	16 oct. 2017	—	déjà là avant Q
JFK Jr.	1 119	3,7	11,9	16 oct. 2017	8 avr. 2018	déjà là avant Q
Sealed indictments	6 159	1,2	80,0	23 oct. 2017	14 nov. 2017	/pol/ avant Q
« Q a dit... » (attributions à Q)	30 066	0,3	394,5	30 oct. 2017	—	naît avec Q
QAnon	17 289	0,0	227,0	30 oct. 2017	6 avr. 2018	/pol/ avant Q
Q drop / Q post / qmap	9 783	0,0	128,4	30 oct. 2017	20 nov. 2017	/pol/ avant Q
Armée Q / adeptes de Q	7 441	3,2	95,3	30 oct. 2017	—	naît avec Q
« Q est une psyop / un LARP » (critiques internes)	5 452	0,0	71,6	30 oct. 2017	—	naît avec Q
Arrestations de masse	2 860	14,2	26,7	30 oct. 2017	11 mars 2019	/pol/ avant Q
The Great Awakening	2 854	5,4	33,4	30 oct. 2017	30 oct. 2017	Q d'abord
Q clearance	894	0,2	11,6	30 oct. 2017	2 nov. 2017	Q d'abord
« Godfather III »	878	0,4	11,2	30 oct. 2017	5 nov. 2017	Q d'abord
EBS (système d'alerte d'urgence)	711	4,1	6,2	30 oct. 2017	2 nov. 2017	Q d'abord
« Trust Sessions / Huber / Wray... »	5 305	29,9	46,7	6 nov. 2017	19 déc. 2017	/pol/ avant Q
Tableaux de Q sur 8chan/8kun	2 445	0,0	32,1	6 nov. 2017	15 déc. 2017	/pol/ avant Q
Codes entre crochets ([RR], [HUSSEIN], [DS], [-35]...)	1 888	11,2	16,2	6 nov. 2017	10 nov. 2017	Q d'abord
Tripcodes de Q	1 132	0,0	14,9	6 nov. 2017	9 déc. 2017	/pol/ avant Q
Trust the plan	8 595	0,0	112,8	19 mars 2018	5 févr. 2018	Q d'abord
« Boom week » / BOOM x N	548	2,3	5,5	19 mars 2018	29 avr. 2018	/pol/ avant Q
WWG1WGA (et variantes)	2 370	0,0	31,1	2 avr. 2018	9 janv. 2018	Q d'abord
« These people are sick/stupid »	1 919	8,5	18,7	2 avr. 2018	30 oct. 2017	Q d'abord

Formule	Posts	Avant	Après	Apparition	Premier drop	Lecture
Frazzledrip	1 019	0,0	13,4	9 avr. 2018	—	naît avec Q
Enjoy the show	3 508	17,6	32,5	14 mai 2018	29 oct. 2017	Q d'abord

Les posts liés à Q sont en moyenne moins toxiques que /pol/ (score Perspective 0,366 contre 0,406), mais jour par jour ils le sont davantage 58 % des jours. La ventilation par groupe (tableau 5) lève ce paradoxe apparent. Les fils Q sont plus calmes que /pol/ (0,317) et le sont 73 % des jours comparés : surtout pendant la naissance de Q, consacrée au déchiffrement des drops (0,291 de fin octobre à décembre 2017), mais encore en 2018-2019 (0,384 contre 0,407). Les posts isolés qui citent un promoteur, pour 76 % d'entre eux Alex Jones ou InfoWars, sont au contraire plus toxiques que /pol/ 67 % des jours (écart médian +0,018, $p < 0,001$), et les posts isolés au vocabulaire de Q à peine (écart médian +0,009). La moyenne d'ensemble est tirée vers le bas par le poids des fils Q, surtout en novembre 2017 ; les jours ordinaires, par les posts isolés.

Tableau 5. Toxicité (score TOXICITY de Perspective) des posts liés à Q sur /pol/, par groupe, juillet 2016 - novembre 2019. Écart médian : moyenne du groupe moins moyenne de tout /pol/, jour par jour, sur les jours où le groupe compte au moins 100 posts notés.

Groupe	Posts notés	Moyenne	Part > 0,8	Écart médian	Jours où le groupe dépasse /pol/	Jours comparés
tout /pol/	131 243 567	0,406	19,3 %	—	—	—
fils Q (CBTS, fils des drops)	391 258	0,317	13,7 %	-0,035	27 %	313
posts isolés au vocabulaire de Q	150 513	0,412	15,7 %	+0,009	58 %	683
posts isolés citant un promoteur	196 091	0,423	16,2 %	+0,018	67 %	866
posts isolés citant un pseudo-initié (FBIanon, HLIanon...)	31 258	0,398	13,2 %	-0,035	29 %	41

Les codes de haine, surtout antisémites, touchent 5,0 % des posts de /pol/ avant Q et 6,5 % après, 3,6 % des posts des fils Q et 0,1 % des drops. La culture du forum (argot, mêmes, accusations de « shills ») est en revanche plus présente dans les fils Q que sur /pol/ lui-même (3,9 % de leurs posts, contre 2,5 % avant Q et 2,1 % après), et « shills » y revient 3,2 fois plus souvent. Les mots les plus caractéristiques des fils Q (Monroe et al., 2008) sont « bread », « cbts », « larp », « qanon », « storm », « indictments », « podesta », « posts », « cia », « keystone ». Ce vocabulaire est celui du déchiffrement des drops, de ses cibles et d'un soupçon d'imposture qui revient sans cesse.

5. Discussion et limites

Ces résultats dessinent un auteur qui réagit à l'actualité, par rafales et sous forme de questions, choisit ses personnages et recycle ses annonces. Aucun de ces résultats ne désigne un initié qui l'anticiperait. Ils confirment, sur tout le corpus et avec des références explicites, ce que les études qualitatives et journalistiques avaient décrit (Rothschild, 2021 ; Tian, 2021), et précisent les ordres de grandeur.

Plusieurs limites s'imposent. Le corpus n'est pas canonique. L'annotation des prédictions est automatique, et sa qualité, mesurée ($F1 = 0,86$), n'est pas parfaite. La frontière entre question rhétorique, formule et annonce reste discutable. La vérification est éditoriale : chaque verdict est sourcé et publié pour être contesté. Le test de l'horoscope dépend du jugement d'un modèle de langage. La comparaison avec un second modèle montre que

les rares liens trouvés varient d'un modèle à l'autre (accord modeste). C'est une limite de la mesure, qui interdit d'interpréter finement les cas où un lien est vu. Ce test mesure en outre des liens précis, pas l'impression subjective de « confirmation » que recherchaient les lecteurs de Q. Les modèles de la section 4.8 sont des analogies : un processus de Hawkes décrit la structure temporelle des drops, pas leur cause, et une part « endogène » élevée est compatible avec un auteur qui écrit en série comme avec des réponses à des lecteurs. Les sources de la scène ont leurs propres biais : l'échantillon Reddit couvre chaque jour actif mais non chaque message, les bannissements de subreddits sont des censures et non des silences, Google Trends est un indice échantillonné et arrondi, et /pol/ s'arrête en novembre 2019. Les scores de toxicité de /pol/, calculés par Perspective, lisent mal l'ironie et l'argot de 4chan, et la date d'apparition d'une formule dépend d'un seuil (quantile 0,999, confirmation sur quatre semaines). Les 34 motifs ajoutés au lexique après la lecture de /pol/ n'y sont mesurés que dans les fils Q, un échantillon de 1 % et les comptes mensuels de mots, et la répartition des posts entre groupes de toxicité suit le premier terme reconnu. Enfin, les corrélations avec GDELT, Wikipédia, Reddit ou Google ne disent rien de la causalité.

6. Conclusion

Confrontés systématiquement aux faits, les drops de Q apparaissent comme un commentaire de l'actualité habillé en prophétie. La synchronie revendiquée avec Trump disparaît dès qu'on la compare au hasard. Les prédictions vérifiables se réalisent rarement (28 %), et nettement moins quand l'événement n'était pas déjà annoncé ou programmé (16 %, IC 95 % 10–24 %). Les « correspondances » avec l'avenir, rares, varient d'un modèle à l'autre et n'excèdent pas celles qu'on observe avec le passé. Replacé dans son milieu, Q n'apparaît ni comme l'inventeur de la scène complotiste américaine ni comme son écho : aucune précédence n'est détectée entre ses thèmes et les siens, dans les sources et aux échelles testées. Il s'installe sur /pol/ dans des fils et un vocabulaire qui existaient avant lui, n'en retient que quelques récits, et forge à son tour des formules que Reddit reprend en quelques jours ou semaines. Ses forums, eux, réagissent à ses drops dans les heures qui suivent. Le principal enseignement est méthodologique : chaque « preuve » avancée par le mouvement peut être transformée en hypothèse testable, à condition de lui donner une référence.

Données et code

Toutes les données dérivées, les verdicts de vérification, le jeu de référence et ses consignes, le code (licence MIT) et le site (licence CC BY 4.0) sont disponibles : <https://s-geffroy.github.io/QAnon/> et doi:10.5281/zenodo.23338131. Les chiffres de ce document sont ceux du site à la date de cette version (16 octobre 2026). Le site est recalculé à chaque évolution, et chaque version est archivée.

Références

- Benedetto, D., Caglioti, E. & Loreto, V. (2002). Language Trees and Zipping. *Physical Review Letters*, 88(4), 048702. doi:10.1103/PhysRevLett.88.048702
- Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289-300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Bloom, M. & Moskalenko, S. (2021). *Pastels and Pedophiles: Inside the Mind of QAnon*. Stanford University Press. doi:10.1515/9781503630611
- Burrows, J. (2002). 'Delta': a Measure of Stylistic Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing*, 17(3), 267-287. doi:10.1093/lc/17.3.267

- Cafiero, F. & Camps, J. B. (2023). Who could be behind QAnon? Authorship attribution with supervised machine-learning. *Digital Scholarship in the Humanities*, 38(4), 1418-1430. doi:10.1093/llc/fqad061
- Clauset, A., Shalizi, C. R. & Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*, 51(4), 661-703. doi:10.1137/070710111
- CodeBreaker619 (2021). *Donald Trump Tweets Dataset*. Kaggle. <https://www.kaggle.com/datasets/codebreaker619/donald-trump-tweets-dataset>
- Covington, M. A. & McFall, J. D. (2010). Cutting the Gordian Knot: The Moving-Average Type-Token Ratio (MATTR). *Journal of Quantitative Linguistics*, 17(2), 94-100. doi:10.1080/09296171003643098
- Crane, R. & Sornette, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*, 105(41), 15649-15653. doi:10.1073/pnas.0803685105
- de Zeeuw, D., Hagen, S., Peeters, S. & Jokubauskaitė, E. (2020). Tracing normification. *First Monday*, 25(11). doi:10.5210/fm.v25i11.10643
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1-26. doi:10.1214/aos/1176344552
- Enders, A. M., Uscinski, J. E., Klostad, C. A., Wuchty, S., Seelig, M. I., Funchion, J. R. et al. (2022). Who Supports QAnon? A Case Study in Political Extremism. *The Journal of Politics*, 84(3), 1844-1849. doi:10.1086/717850
- Festinger, L., Riecken, H. W. & Schachter, S. (1956). *When Prophecy Fails: A Social and Psychological Study of a Modern Group That Predicted the Destruction of the World*. University of Minnesota Press. <https://openlibrary.org/works/OL4788695W>
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221-233. doi:10.1037/h0057532
- Goh, K. I. & Barabási, A. L. (2008). Burstiness and memory in complex systems. *EPL (Europhysics Letters)*, 81(4), 48002. doi:10.1209/0295-5075/81/48002
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3), 424-438. doi:10.2307/1912791
- Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1), 83-90. doi:10.1093/biomet/58.1.83
- Jackson, B., Scargle, J. D., Barnes, D., Arabhi, S., Alt, A., Gioumoussis, P. et al. (2005). An algorithm for optimal partitioning of data on an interval. *IEEE Signal Processing Letters*, 12(2), 105-108. doi:10.1109/LSP.2001.838216
- Jensen, M. A. & Kane, S. (2024). QAnon-inspired violence in the United States: an empirical assessment of a misunderstood threat. *Behavioral Sciences of Terrorism and Political Aggression*, 16(1), 65-83. doi:10.1080/19434472.2021.2013292
- Killick, R., Fearnhead, P. & Eckley, I. A. (2012). Optimal Detection of Changepoints With a Linear Computational Cost. *Journal of the American Statistical Association*, 107(500), 1590-1598. doi:10.1080/01621459.2012.737745
- Kingsman, J. (2021). *QAnon Posts Dataset (JSON-QAnon)*. Kaggle / GitHub. <https://www.kaggle.com/datasets/jkingsman/qanon-drops>
- Lee, D. D. & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755), 788-791. doi:10.1038/44565
- Linville, D., Chambers, M., Duck, J. & Sheffield, S. (2021). Mapping the messenger: Exploring the disinformation of QAnon. *First Monday*, 26(12). doi:10.5210/fm.v26i12.11745
- Marwick, A. E. & Partin, W. C. (2024). Constructing alternative facts: Populist expertise and the QAnon conspiracy. *New Media & Society*, 26(5), 2535-2555. doi:10.1177/14614448221090201
- McIntosh, J. (2022). The sinister signs of QAnon: Interpretive agency and paranoid truths in alt-right oracles. *Anthropology Today*, 38(1), 8-12. doi:10.1111/1467-8322.12697
- Monroe, B. L., Colaresi, M. P. & Quinn, K. M. (2008). Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. *Political Analysis*, 16(4), 372-403. doi:10.1093/pan/mpn018
- Ogata, Y. (1988). Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *Journal of the American Statistical Association*, 83(401), 9-27. doi:10.1080/01621459.1988.10478560

- OrphAnalytics (2020-12-15). *Style analysis by machine learning reveals that two authors likely shared the writing of QAnon's messages at two different periods in time*. OrphAnalytics SA. <https://www.orphanalytics.com/en/news/whitepaper202012>
- Papasavva, A., Zannettou, S., De Cristofaro, E., Stringhini, G. & Blackburn, J. (2020-05-26). *Raiders of the Lost Kek: 3.5 Years of Augmented 4chan Posts from the Politically Incorrect Board*. *Proceedings of the International AAAI Conference on Web and Social Media*, 14. Association for the Advancement of Artificial Intelligence (AAAI). doi:10.1609/icwsm.v14i1.7354
- Papasavva, A., Aliapoulos, M., Ballard, C., De Cristofaro, E., Stringhini, G., Zannettou, S. et al. (2022). The Gospel according to Q: Understanding the QAnon Conspiracy from the Perspective of Canonical Information. *Proceedings of the International AAAI Conference on Web and Social Media*, 16, 735-746. doi:10.1609/icwsm.v16i1.19330
- Peng, C. K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E. & Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E*, 49(2), 1685-1689. doi:10.1103/PhysRevE.49.1685
- Phipson, B. & Smyth, G. K. (2010). Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn. *Statistical Applications in Genetics and Molecular Biology*, 9(1). doi:10.2202/1544-6115.1585
- PRRI Staff (2024-11-04). *The Rise and Impact of Q: The 2024 Election from the View of QAnon Believers*. PRRI Spotlight Analysis. PRRI. <https://prri.org/spotlight/the-rise-and-impact-of-q-the-2024-election-from-the-view-of-qanon-believers/>
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257-286. doi:10.1109/5.18626
- Rothschild, M. (2021). *The Storm Is Upon Us: How QAnon Became a Movement, Cult, and Conspiracy Theory of Everything*. Melville House. <https://openlibrary.org/works/OL24209566W>
- Schreiber, T. & Schmitz, A. (2000). Surrogate time series. *Physica D: Nonlinear Phenomena*, 142(3-4), 346-382. doi:10.1016/S0167-2789(00)00043-9
- Schreiber, T. (2000). Measuring Information Transfer. *Physical Review Letters*, 85(2), 461-464. doi:10.1103/PhysRevLett.85.461
- Sommer, W. (2023). *Trust the Plan: The Rise of QAnon and the Conspiracy That Unhinged America*. HarperCollins. <https://openlibrary.org/works/OL26341549W>
- The GDELT Project (s.d.). *GDELT Event Database*. <https://www.gdeltproject.org/>
- Tian, E. (2021-01-29). The QAnon Timeline: Four Years, 5,000 Drops and Countless Failed Prophecies. *Bellingcat*. <https://www.bellingcat.com/news/2021/01/29/the-qanon-timeline/>
- Tuters, M., Jokubauskaitė, E. & Bach, D. (2018). Post-Truth Protest: How 4chan Cooked Up the Pizzagate Bullshit. *M/C Journal*, 21(3). doi:10.5204/mcj.1422
- Veen, A. & Schoenberg, F. P. (2008). Estimation of Space–Time Branching Process Models in Seismology Using an EM–Type Algorithm. *Journal of the American Statistical Association*, 103(482), 614-624. doi:10.1198/016214508000000148
- Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, 57(2), 307-333. doi:10.2307/1912557
- Westmark, C. R. & McMahon, A. M. (2022). Identifying QAnon Conspiracy Theory Adherent Types. *New Political Science*, 44(4), 607-627. doi:10.1080/07393148.2022.2129927
- Wikimedia Foundation (s.d.). *Wikimedia Analytics API: Pageviews*. <https://doc.wikimedia.org/generated-data-platform/aqs/analytics-api/>
- Wilson, E. B. (1927). Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*, 22(158), 209-212. doi:10.1080/01621459.1927.10502953