



King's Research Portal

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Bounegru, L., Gray, J., Venturini, T., & Mauri, M. (2019). . Public Data Lab. <http://jcej.info/>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

調査のための
フィールドガイド

フェイス ニュース

日本語版

A COLLECTION OF RECIPES
FOR THOSE WHO LOVE TO COOK
WITH DIGITAL METHODS

COMPILED BY

Liliana Bounegru
Jonathan Gray
Tommaso Venturini
Michele Mauri

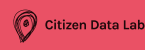


FIRST
DRAFT

DEN-
SITY
GN+

digital
methods
initiative

inria



A FIELD GUIDE TO "FAKE NEWS" AND OTHER INFORMATION DISORDERS

Compiled by **Liliana Bounegru**, **Jonathan Gray**, **Tommaso Venturini** and **Michele Mauri**.

This guide explores the use of digital methods to study false viral news, political memes, trolling practices and their social life online. It is a project of the **Public Data Lab** with support from **First Draft**.

The Public Data Lab (publicdatalab.org) is an interdisciplinary network seeking to facilitate research, democratic engagement and public debate around the future of the data society.

First Draft (firstdraftnews.com) is dedicated to improving

skills and standards in the reporting and sharing of information that emerges online.

© 2017 Public Data Lab. Amsterdam.

The guide is released under the Creative Commons Attribution License (creativecommons.org/licenses/by/4.0/). This means you can freely share, adapt and draw on this work as long as you give credit, as per the terms of the license.

If you reproduce or draw on material from this guide, we'd be grateful if you could credit and link back as per the following statement:

"This article draws on A Field Guide to "Fake News" and Other Information Disorders, a collaboration of the Public Data Lab and First Draft. For further details see: <http://fakenews.publicdatalab.org>"

The copyright for the research that this guide draws upon remains with its respective contributors.

This project would not have been possible without the contributions of researchers from the following institutions:

Centre for Journalism Studies, University of Ghent (NL)
Citizen Data Lab, Amsterdam University of Applied Sciences (NL)
DensityDesign Lab, Politecnico di Milano (IT)
Digital Methods Initiative, Media Studies, University of Amsterdam (NL)
Govcom.org Foundation, Amsterdam (NL)
Institut National de Recherche en Informatique et en Automatique (INRIA) (FR)
King's College London (UK)
Laboratoire d'Étude des Sciences et des Techniques (STS-Lab),
Université de Lausanne (CH)
Laboratoire Interdisciplinaire Sciences Innovations Sociétés (LISIS), Université Paris-Est Marne-la-Vallée (FR)
Médialab, Sciences Po, Paris (FR)
Techno-Anthropology Lab, Aalborg University Copenhagen (DK)
University of Siegen (DE)

Published by Public Data Lab.
Version 1.0.0
Released in January 2018.

A FIELD
GUIDE TO

"FAKE NEWS"

and other
INFORMATION DISORDERS



FIRST
DRAFT

DEN-
SITY
GN+

digital
methods
initiative

inria



IPR
Institute for
Policy Research



CJS

UNIVERSITÄT
SIEGEN

Unil
Université de Lausanne
Institut des sciences sociales
STSLab & LaDHUL

LISIS SciencesPo
MICALIS

Citizen Data Lab



目次

6 はじめに

14 ガイドで使用しているアイコンについて

第1章

17 フェイスブック上でフェイクニュースのホットスポットをマッピングする

20 1. フェイクニュースで盛り上がっているのはどんな人たちか？

40 2. フェイクニュースの伝達経路をいかに追跡するか？

52 3. ファクトチェックの取り組みはフェイクニュースの読み手に届くか？

第2章

61 ウェブ上のフェイクニュース拡散を追跡する

64 1. フェイクニュースはどこから生まれ、どのサイトで最初に再発信されるのか？

78 2. フェイクニュースに関わる最も明確なソースは何か？いつ、誰がソースに言及しているか？

第3章

95 ウェブページの背後にある広告用の「トラッカー」に注目し、フェイクニュースのサイトを特定する

97 1. フェイクニュースサイトは、どのようなトラッカーを使って、主要メディアからコンテンツを取り込んでいるのか？

103 2. トラッカーの使い方でフェイクニュースサイトと主要メディアのサイトを区別する方法

109 3. フェイクニュースサイトのトラッカーは、時間とともにどう変わっていくのか？

115 4. フェイクニュースサイトで使われているトラッカーIDは、他のどのようなサイトで使われているか

121 5. 政治的に極端に偏った内容や、誤解を招く情報を掲載するサイトが使うトラッカーは、言語圏によって異なるのか

第4章

- 127 フェイスブック上の政治ミームを調査する
- 130 1. フェイスブック上のミーム空間を明らかにする方法
- 139 2. 政治・メディア的イベントが、どんなミームになるのか
- 148 3. ミームのコンテンツを調査する方法

第5章

- 161 ツイッター上の荒らし行為をマッピングする
- 165 1. 政治家を攻撃対象にするツイッターアカウントを検出する方法
- 172 2. 荒らしを行うユーザーの特性を明らかにする
- 184 3. 荒らし行為はどのように特徴付けられるか

199 おわりに

205 用語集

214 Special Thanks

フェイクニュース調査の フィールドガイド

はじめに

- [1] See, for example, Robert Darnton, "The True History of Fake News", The New York Review of Books, February 2017, available at: <http://www.nybooks.com/daily/2017/02/13/the-true-history-of-fake-news/>
- [2] See, for example, "Sky Views: Facebook's fake news threatens democracy", Sky News: <http://news.sky.com/story/sky-views-democracy-burns-as-face-book-lets-fake-news-thrive-10652711>
- [3] See Edson Tandoc, Zheng Wei Lim & Richard Ling "Defining 'Fake News'", Digital Journalism, 811, August 2017. pp. 1-17. 2017. Available at: <https://www.tandfonline.com/doi/full/10.1080/21670811.2017.1360143>
- [4] See Claire Wardle & Derakhshan Hossein, Information Disorder: Toward an interdisciplinary

「フェイクニュース」とは何か。どう対処すればいいのか。誰に聞くかによって、さまざまな答えが返ってくる。「情報戦争における大変化の縮図」「不当に利益を得るための新興手段」「“オルタナ右翼”や、デジタルによって生み出された草の根的な政治動員を世界中で活気づける原動力」「新リベラルの“真理省”(ジョージ・オーウェル『1984年』参照)を求める勢力争いのスローガン」「デジタル社会を形成するオンライン・プラットフォームが生んだ、不必要な副産物」。もしくは「知的生産をめぐる従来の制度やプロセスに関する合意の崩壊を示唆し、政治や公的な生活における『ポスト・トゥルース』時代の到来を告げるカナリアの鳴き声」だとも言われる。

フェイクニュースは、単に「新しいボトルに入れた古いワイン」という批評家もいる。ミスリーディングな情報が蔓延する同様の現象は、プリント・メディアやその他のコミュニケーション・テクノロジーの歴史と同じくらい昔から存在しているというのだ^[1]。新しいオンライン・プラットフォームがフェイクニュースの拡散を加速させ、「増加させた」ことで、これまでになかった課題とダイナミクスを生み出したという意見もある^[2]。「フェイクニュース」という言葉の使用は徹底的に避けるべきだという主張さえある。その理由には、定義として曖昧すぎる^[3]、政治的な意味において脅威であること^[4]、過去に見られた偽情報の形態と区別できないこと^[5]、真実とは、現実をありのままに映し出すものだという、あまりにも単純な考えに立っていること^[6]、フェイクニュ

framework for research and policymaking (Report to the Council of Europe), 2017.

[5] See Caroline Jack, What's Propaganda Got To Do With It? Available at: <https://points.datasociety.net/whats-propaganda-got-to-do-with-it-5b88d78c3282>

[6] See Michael Lynch, Post-truth, alt-facts, and asymmetric controversies, 2017, First 100 Days, Available at: <http://first-100days.stsprogram.org/2017/02/06/post-truth-alt-facts-and-asymmetric-controversies-part-i/>

[7] See Henry Jenkins, Sam Ford, & Joshua Benjamin Green, Spreadable media. New York: New York University Press, 2017, Available at: <http://doi.org/10.1017/>

[8] See, John Naughton, "Facebook and Twitter could pay the price for hate speech", The Guardian, March 2017: <https://www.theguardian.com/commentisfree/2017/mar/19/john-naughton-germany-fine-social-media-sites-facebook-twitter-hate-speech>

[9] See, for example, Full Fact, "The State of Automated Factchecking", <https://fullfact.org/automated> and Lori Hawkins, "Austin startup wants to end fake news - and fake everything else - on the internet", 512Tech, February 2017: <http://www.512tech.com/technology/austin-start->

ース現象の最大の脅威は、虚偽性ではなく「拡散性」にあるという最も重要な視点が抜け落ちていること、などが挙げられている[7]。

対策として打ち出されたのが、メディアリテラシー、教育、ファクトチェックなどの新しい取り組みだ。新たな法律や政策、不適切なコンテンツを掲載するテクノロジー企業に対する罰金[8]、多くのスタートアップや新技術(コンテンツ認証や、自動でファクトチェックを行うプロジェクトなど)もある[9]。

さまざまな対応策が講じられる中、「フェイクニュース」という言葉はミスリーディングであり、悪質な捏造コンテンツは一言では定義できないというのが専門家の一致した意見だ。「フェイクニュース」という言葉は、この問題を初めて社会全体の懸念事項として認識させ、今回の実験的調査を行うきっかけとなった。そのため、このガイドでは偽ニュースを作り出す当事者を追跡し[10]、「フェイクニュース」という言葉の使用を排除していない。一方で、「フェイクニュース」という定義のもとで収集された捏造コンテンツには、さまざまな形態があることも認識している。それは、この問題を調査することが無意味だということではない。むしろ逆に、そのさまざまな形態こそが、本調査における重要な点であり、「フェイク」と「フェイクでないもの」の間に線引きをする簡単な方法はないという事実を受け入れることで、興味深い研究機会が生まれる。これはまさに、フェイクニュースの形態やコンテンツは主要メディアを模倣するものであるため——そして実に似たような経路で拡散されるため——「フェイク」の戦略やフォーマットだけでなく、デジタル時代におけるメディアや情報環境の政治学と構図を、より全般的に研究することができる。

このガイドの目的は、急速に展開し、論争の的になっているフェイクニュース問題について、一般の議論と協働リサーチを促すことにある。そのために、ネット上で実験的に研究を進め、マッピングし、調査できるさまざまな方法を提案する。最終的には、フェイクニュース問題と、それに付随する現象に対するより明確な説明を提供するだけでなく、人々がもっとこの問題に現実的な関わりを持つことを願っている。このガイドが広範な議論と参加を後押しし、フェイクニュースやその他の捏造コンテンツ拡散に関わるプラットフォームや政策、法律、インフラストラクチャー、テクノロジー、スタンダードの再形成に資することを願う。同

[up-wants-end-fake-news-and-fake-every-thing-else-the-internet/EcchWFgrl4PQmjPvm-kzycJ/](https://doi.org/10.1086/667828)

[10] Bruno Latour, *Reassembling the Social: An Introduction to Actor-Network-Theory*, Oxford: Oxford University Press, 2009
[CBO9781107415324.004](https://doi.org/10.1086/667828)

時に、こうしたさまざまな対応がもたらす予期せぬ結果や、その他の利害・懸念にも注意しなくてはならない。

このガイドでは、フェイクニュースとはオンライン上で拡散する一種のコンテンツというだけでなく、まさしくその「オンライン上で拡散、受容される」という性質こそがフェイクニュースを生み出しているとの観点に立つ。この意味において、フェイクニュースやその他の捏造コンテンツは、発せられるメッセージや内容形態にとどまらず、拡散を促す仲介インフラ、プラットフォーム、参加型カルチャーにおいても考察される。オンライン上の拡散という観点から離れては、フェイクニュースを真に理解することはできない。ニッチな分野の風刺として生み出されたものが、どのように、極端に意見が偏ったグループのクリックベイト記事として再形成され、広告マネーを生み出すのか。そして、悪質な政治デマ情報としてどうネット上で生き残っていくのか。拡散を追うことで、それらを理解することが可能になる。

よって、フェイクニュースの内容に個別に着目するのではなく、こういった文脈で拡散されるかに注目すべきだというのが、このガイドの主張だ。このシフトは、「不足モデル」アプローチ(フェイクニュースは、正しい情報が十分でないためにはびこっていると示唆する考え方)の限界を示している。このガイドは、疑わしい記事の特定・検証にとどまらず、フェイクニュースをマッピングし、対処する新たな方法を提案する。どうフェイクニュースが動き、人々を巻き込むかを深く理解するため、拡散に関するより「厚みのある」説明を提供する。フェイクニュース現象を的確に捉えるため、「フェイク」とその受容についての繊細な説明も行う。

[11] See Theodore Porter, "Thin Description: Surface and Depth in Science and Science Studies." *Osiris*, 2012: www.jstor.org/stable/10.1086/667828

オンラインやプラットフォームの指標は、セオドア・ポーターが「薄い記述」^[11]と呼ぶ方法でエンゲージメントを測るのにしばしば役立つが(例えば「いいね」やシェア、投稿の総数などの集計)、このガイドでは、異なる層がどのようにフェイクニュースに関与し意味を見出すのか、そしてその過程で、いかにプレーヤーが動員されるのかを調べるさまざまな方法を提案する。フェイクニュース調査からスタートするが、最終的にはオンライン上での草の根的な政治、メディア、参加型カルチャーを多岐にわたり明らかにし、中心にある社会的、政治的課題を解明することになる。これはネット上における草の根的政治アクションの在り方に疑問を投げかけ、再考を促すことになるかもしれない。

近年発行されているほかのガイドに倣い、土地や動植物について書き留める言葉や図像で、クラウドやデジタルインフラ、オンライン上の営みの世界を表す「フィールドガイド」のメタファーを採用した [12]。しかし、このメタファーには但し書きが必要だ。古典的な博物学の「フィールドガイド」では、自然現象を概念化して特徴を比較するための体系的な分類が目的であり、現象は文脈から切り離されることになる。このガイドはまさにその反対のことは行おうとしている。つまり、フェイクニュース現象を文脈から切り離すのではなく、文脈の中に位置づけなおすことだ。そのために、フェイクニュースを「野生の状態」の中で — つまり検索エンジン、デジタル・プラットフォーム、ファクトチェックの試み、そしてニュースサイト上で — 追跡する方法を提案する。

[12] For a recent overview see Shannon Mattern's "Cloud and Field", *Places Journal*, August 2016: <https://placesjournal.org/article/cloud-and-field/>

私たちは、フェイクニュースに関する唯一の説明を試みているのではない。解剖図的な説明や地図情報、サテライト画像や構造的図表を示そうとしているわけでも、フェイクニュースの特定に使えるかもしれない特徴のリストを示そうとしているわけでもない。その代わりに、読者自身が、フェイクニュース現象を詳しく調査するのに役立つだろうさまざまな方法と手段について説明する。その一環として、フェイクニュースを理解するために、公に利用可能な追跡方法のレパートリーを拡大したいと考えている。そのため、ここで提供している図式は、さらなる調査を後押しするための一時的なプレースホルダーだと理解してほしい。

[13] See Philip E. Agre, "Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI", in Geof Bowker, Les Gasser, Leigh Star and Bill Turner (eds), "Bridging the Great Divide: Social Science, Technical Systems and Cooperative Work", NJ: Erlbaum, 1997. Available at: <http://polaris.gseis.ucla.edu/pagre/critical.html>

また、フェイクニュースとして分類され、ラベル付けされるさまざまな方法、そして分類を行うにあたっての駆け引きにも注目する。この意味において、私たちは「クリティカル・テクニカル・プラクティス」[13] というものを構築したいと考えている。この場合、デジタル手法やデジタル・データ使用の省察と、これらがいかにして、直接的かつ独立的に抽出されうる現象を決定づけるのか、またこれらの手法そのものがまさに、フェイクニュースとは何かを定義づけるプロセスに関わる可能性の認識を意味する。シャノン・マッターンが言うように、オンライン上のフェイクニュースを調査する際、私たちは「フィールドを調査する自分たちの存在によって落とされる影」[14]を認識すべきである。現状を把握する目的のためにそうするのではなく、どんな役割を果たすのか、何のため、誰のためなのか。そして社会的、文化的、政治的装置としてどんな影響をもたらすのかを考慮するべきである。

[14] Shannon Mattern, "Cloud and Field", *Places Journal*, August 2016: <https://placesjournal.org/article/cloud-and-field/>

[15] See, for example, the work of Pamela Smith and her colleagues on the “Making and Knowing” project at Columbia University: <http://recipes.hypotheses.org/7430>, February 2016, and <http://www.makingand-knowing.org/>

フェイクニュース現象の説明より、調査手順の提供に注力するという点で、このガイドは「レシピ本」のようなものとも言える。最近の調査は、近世のレシピの作成と手続き的知識の間には、興味深い関連性があることを示唆している。料理や工芸のプロセスを記録する行為は、科学的メソッドの出現と関連があるというのだ[15]。過去数十年の間に、「レシピ」や「料理本」というメタファーは、ソフトウェア・プログラミングに関連しても使われるようになった。このガイドは、ネット上におけるフェイクニュース追跡調査のさまざまなアプローチを、一連の方法論的な「レシピ」の形で紹介する。多くの料理本と同様、我々の目的は、読者が特定のレシピを真似するだけでなく、むしろこのレシピを使い、別の料理法を見出すようサポートすることだ。このレシピを適用し、修正し、さらに先へ進んでほしい。ガイドでは、実践における「おすすめ手法」もいくつか挙げている。

このガイドのレシピが、ミステリー小説のナラティブ・アプローチのような形で、フェイクニュースや偽コンテンツの調査を補強するものになればと願う。つまり、一見シンプルな事件を解決しようとする中で、陰謀は膨らみ、登場人物は増え、謎は深まり、思わぬ展開や大局の変化に直面するのだ。ネット上でフェイクニュースが生まれ、拡散され、人々が反応する過程を追うことで、当初は調査しようとは思ってもいなかった事象に引き込まれるかもしれない。フェイクニュース発信者、プロパガンダを広めたい人々、trolle(ネット上の荒らし行為)やボットのメディア戦略。オンラインコンテンツの商業的・技術的な構造。バイラルコンテンツの策略やダイナミクス。そして、巨大なオンライン・プラットフォームやウェブサイトに対し、社会的営みがどのように適応し、進化し、革新を起こすのか。この意味において、フェイクニュースに関わるのは、一部の偽情報発信者や国家の陰謀にとどまらないということが、明確になるだろう。そして、社会におけるデジタル技術の役割や、我々とテクノロジーがどのように互いに影響し合っているかについて、重要で難しい問題が提起される。

[16] Edgar Allan Poe and Jacob Schwartz, “The Purloined Letter”, London: Ulysses bookshop, 1931.

エドガー・アラン・ポーの名作ミステリー「盗まれた手紙」では[16]、G警視総監と同僚たちが、スキャンダラスな情報が書かれた手紙を見つけようと、壁紙の裏やカーペットの下、家具の足やクッションの中を探し回る。だがその手紙は「ありふれた風景の中に隠されていた」のだ。同様に、アルゴリズムによって仲介される、デジタル・プラットフォーム上でのフェイクニュース拡散

[17] See, Noortje Marres, "The Redistribution of Methods: On Intervention in Digital Social Research, Broadly Conceived", *The Sociological Review*, June 2012; and Mike Ananny and Kate Crawford, "Seeing Without Knowing: Limitations of the Transparency Ideal and its Application to Algorithmic Accountability", *New Media and Society*, December 2016.

[18] Liliana Bounegru, Mette Simonsen Abildgaard, Andreas Birkebæk, Jonathan Gray, Mathieu Jacomy, Torben Elgaard Jensen, Anders Koed Madsen and Anders Kristian Munk, "Five Provocations about Fake News" (under review).

[19] See, for example, Richard Rogers, "Digital Methods", 2013, Cambridge, MA: MIT Press.

[20] See, for example, Noortje Marres, "Material Participation: Technology, the Environment and Everyday Publics", London: Palgrave Macmillan, 2012.

[21] See Tommaso Venturini, Anders Munk and Axel Meunier, "Data-Sprint: A Public Approach to Digital Research" in C. Lury, P. Clough, M. Michael, R. Fensham, S. Lammes, A. Last, & E. Uprichard (Eds.) "Interdisciplinary Research Methods", London: Routledge, 2017.

は、ヌーティ・マレスやキャロリン・ゲーリッツが言うところの「分散型アチーブメント」や、マイク・アナニーやケイト・クロフォードが言う「相関的アチーブメント」[17]と考えることができるだろう。これは、システムを「丹念に見ていく」ことから、オンラインの集合コンテンツのパターンに関わるさまざまな要素を「横断的に見る」ことへのシフトを伴う。

このガイドに関わった研究者の多くが、科学技術社会論 (Science and Technology Studies : STS) と呼ばれる分野のバックグラウンドを持つ。このガイドが調査する事象のいくつかは、STSがフェイクニュースの調査にどう役立つかに関する、今後発表予定の論文に基づいている[18]。また、レシピは「デジタル手法」の調査アプローチに基づく。STSの分野と交わることで発展し、ガイド著者の多くもティーチングやリサーチを通じて貢献してきた[19]。デジタル技術とデータ・インフラストラクチャーに対する関心も持ってもらいたい[20]。つまり私たちは、法的または技術的な解決策よりも、パブリック・エンゲージメントや民主的な議論の過程を促進することに、より重きを置いている。フェイクニュース問題に対してさまざまな見方があることへの好奇心と、多様なリアクションに対する想像力を刺激したいと考えている。

このガイドの内容は、アムステルダム、コペンハーゲン、ミラノでパブリック・データ・ラボが主催した、一連の「データ・スプリント」と調査ワークショップを通じて作成された。「データ・スプリント」は、STSとニューメディア研究の間で生まれたショートフォーム・ワーキング・フォーマットで、オープンソース・ソフトウェアの開発やオープンデータ、そしてシビック・ハッキングに関連したアプローチを用いている。異なる展望を持つ分野を含むさまざまな当事者が連携し、データや調査プロジェクトの共同生産を目的とするものだ[21]。

ガイド著者のうち2人は、データジャーナリズムのバックグラウンドを持つ。「データジャーナリズム・ハンドブック」を共同編集し、この分野でさまざまな取り組みを行ってきた[22]。このガイドは、データジャーナリストとデジタル研究者が生産的に互いにつながることを後押ししたいという、長年にわたる関心に基づいている。フェイクニュースは、この二つの分野における試みという点で、まさにぴったりの領域に見える。だが一方で、作家ホルヘ・

[22] Jonathan Gray, Liliana Bounegru and Lucy Chambers (Eds.) "The Data Journalism Handbook", Sebastopol, CA: O'Reilly Media, 2012, available at: <http://data-journalismhandbook.org/>

[23] Upon being invited to become Director of the Argentine National Library at a moment which coincided with the deterioration of his eyesight, Borges famously wrote: "No one should read self-pity or reproach / into this statement of the majesty / of God; who with such splendid irony / Granted me books and night at one touch". See J. L. Borges, "Poem of the Gifts" in "Selected Poems: Volume 2", London: Penguin Books, 2000, p. 95.

ルイス・ボルヘスが「書物と闇を同時に」与えられたと嘆き悲しんだように[23]、フェイクニュース問題に対する注目が増すにつれ、フェイクニュースを発信するウェブサイトが消えつつあるという皮肉な状況でもある。幸い、このガイドのアプローチと分析テクニックは、他の問題研究や議論においても、データジャーナリストとデジタル研究者の連携に活用されうる。また、オンライン活動を形作るプラットフォームやアルゴリズム、そしてインフラストラクチャーの媒介役としても活用できるだろう。

データ・スプリント・フォーマットは、新たな実験や実証研究を相対的に短期間で促進させるのにも役立ってきた。これは、フェイクニュース問題が展開しているペースを考えれば、注目すべき利点だ。この点において、以下の研究者、卒業生、学生に多大な感謝を表明したい。彼らのエネルギー、創造性、そして献身なしには、このプロジェクトは不可能だっただろう。

DensityDesign Lab (ミラノ工科大学)、*the Digital Methods Initiative* (アムステルダム大学)、*the European Journalism Centre*, *the Laboratoire Interdisciplinaire Sciences Innovations Sociétés* (パリ東大学)、*the médialab* (パリ政治学院)、*the Media of Cooperation research group* (ジューゲン大学)、*the STS-Lab* (ローザンヌ大学)、*the Techno-Anthropology Lab* (オールボー大学)

Jonathan Gray (@jwyg), *Liliana Bounegru (@bb_liliana)*, *Tommaso Venturini (@TommasoVenturini)*
ロンドン、2017年3月

ガイドで使用している アイコンについて

このガイドでは、👁️ (目)は可視化した図、🔧 (レンチ)はツール、→ (矢印)はコンセプトを表すシンボルとして使用している。読者の邪魔にならないよう、各レシピで初めて説明する際のみ、これらのアイコンを使用する。

さらに、各章のレシピは図表またはメソッド・マップを用いて説明され、結果に辿りつくための主要な分析ステップを表している。それぞれのメソッド・マップでは、矢印はアクションを、アイコンはその結果を表している。読者自身のレシピに使えるかもしれない材料として、これらのステップを参照することができる。

レシピのいくつかは、複数の結果に辿りつくようになっている。その場合、レシピ全体の完全なメソッド・マップが初めに示してあり(青色の背景)、個別のステップに関連する部分は白い背景で、説明とともに示されている。

以下は、メソッド・マップで使用している全アイコンのリストだ。



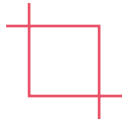
テーブル形式のデータセット。



バブルグラフ、ネットワーク図など、可視化した図。可視化モデルの全リストは、コンセプト用語集を参照。



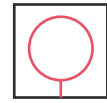
リスト。例えば、ウェブサイトやフェイスブックページのリストなど。



スクリーンショット。ウェブページのスナップショットを保管するためにブラウザから取る。



画像コーパス。同じ手法で取得された一連の画像。



ユーザーのプロフィール。ソーシャルネットワーク上のユーザーに関連するすべての情報を表す。例えば、ツイッター上では、ユーザー名、説明、プロフィール写真などを含む。



一時的な情報。例えば、フェイスブック投稿の作成日時など。



ハッシュタグ。ツイッターやフェイスブックなど多くのソーシャルネットワーク上で使われる。

アクション

関連するアクションはそれぞれのアイコンで表示することがある。下記リストを参照。



自動オペレーション。機械によって行われるアクション(例えば、アイテムをカテゴリー別に分ける時など)をハイライトするのに使われる。



手動オペレーション。手動によって行われるアクション(例えば、アイテムをカテゴリー別に分ける時など)をハイライトするのに使われる。



画像の比較。画像を視覚的に比較する必要がある際に使われる。



リストの融合。二つ以上のリストを一つにまとめる際に使われる。

第1章

フェイスブック上で フェイクニュースの ホットスポットをマッピングする

フェイクニュースで盛り上がって
いるのはどんな人たちか？

フェイクニュースの伝達経路を
いかに追跡するか？

ファクトチェックの取り組みは
フェイクニュースの読み手に
届くか？

<第1章のポイント>

フェイスブックで、どのような読者が、どんな
テーマのフェイクニュースを好んでシェアして
いるかを分析する。偽だと指摘する「暴露記
事」の広がりも調べる。

はじめに - 本章では、フェイスブック上でフェイクニュースの拡散状況を調べるための三つのレシピを示す。拡散に重点を置くのは、誤報やミスリード情報はそのまま「フェイクニュース」になる訳ではなく、多くの人を巻き込んで初めてフェイクニュースになるとの観点からだ。つまり目撃者や賛同者、「いいね!」や「シェア」をする人たち、さらには異議を唱え、誤りを正そうとする人たちもだ。フェイスブックはアクセスや認証システムの性質上、コンテンツそのものの拡散を調べるのが難しい。従ってここでは、フェイクニュースの読者の側に着目する。フェイスブックの公開ページ／グループであればアクセスが可能で、ひとまとまりの読者と見なすこともできるだろう。

2016年の米大統領選では、フェイスブック発で極めて党派性の高い「政治メディアマシン」が現れ、効果的に多数のフォロワーを集め、多くのエンゲージメントを生み出したと言われる[1]。本章では、こうしたフェイクニュースの拡散マシンとそれに

[1] See, John Herrman, "Inside Facebook's (Totally Insane, Unintentionally Gigantic, Hyperpartisan) Political-Media Machine", August 2016, The New York Times: <https://www.nytimes.com/2016/08/28/magazine/inside-face-books-totally-insane-unintentionally-gigantic-hy->

対する反応を調査する。最初の二つのレシピでは、フェイクニュースで盛り上がる読者と、フェイクニュースが伝播する経路をマッピングする。三つ目のレシピでは、ファクトチェックの

取り組みはフェイクニュースの読み手に届いているのか、その有効性を検討するアプローチを示す。一連のレシピにより、ソーシャルメディアにおけるフェイクニュースの流通やエンゲージメントに関し、数値のみに頼る「薄い」説明ではない、より「厚み」のある説明を目指したい。



フェイクニュースで 盛り上がっているのは どんな人たちか？

始める前に

このレシピの出発点は、フェイクニュースのリストだ。リストは既存のものを使っても、独自に作成してもよい。ここではレシピを分かりやすく説明するため、2016年の米大統領選に関連する22のフェイクニュースを集めた既存のリストを使用する。いずれもフェイスブック上で大きなエンゲージメントを生んだフェイクニュースで、バズフィードニュースがリスト化した。

レシピは四つの手順からなる。まずは記事が取り上げているテーマや主な政治上のイベントを特定することから始めよう(👁a)。次に、記事を最も多くシェアしているフェイスブックのページ／グループを特定する(👁b)。そしてユーザーが好む記事テーマを探り(👁c)、米大統領選のフェイクニュースで活気づいた人たちの姿を描き出す(👁d)。

START

list of 22 URLs of political fake news stories
Source: BuzzFeed News

input URLs in

CrowdTangle

output data

- > Fake news story URLs
- > Facebook pages and groups that share the URLs
- > Number of interactions per each page or group
- > Date of sharing of the story

identify time intervals with highest frequency of publication of fake news stories

identify key related events with

Google News



a
WHICH MEDIA AND POLITICAL EVENTS ARE SUCCESSFUL IN SETTING THE FAKE NEWS AGENDA?

import data in

RAWGraphs



b
WHICH FACEBOOK PAGES AND GROUPS PROMOTED THE HIGHEST NUMBER OF FAKE NEWS STORIES?

import data in

Gephi



c
DO FACEBOOK PUBLICS HAVE PREFERRED STORY THEMES?

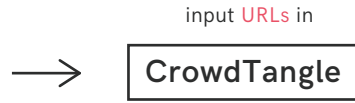
manually categorise Facebook pages



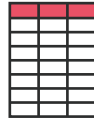
d
WHAT KINDS OF PUBLICS ARE ENERGISED BY FAKE NEWS?

START

list of 22 URLs of political fake news stories
Source: BuzzFeed News



output data



- > Fake news story URLs
- > Facebook pages and groups that share the URLs
- > Number of interactions per each page or group
- > Date of sharing of the story

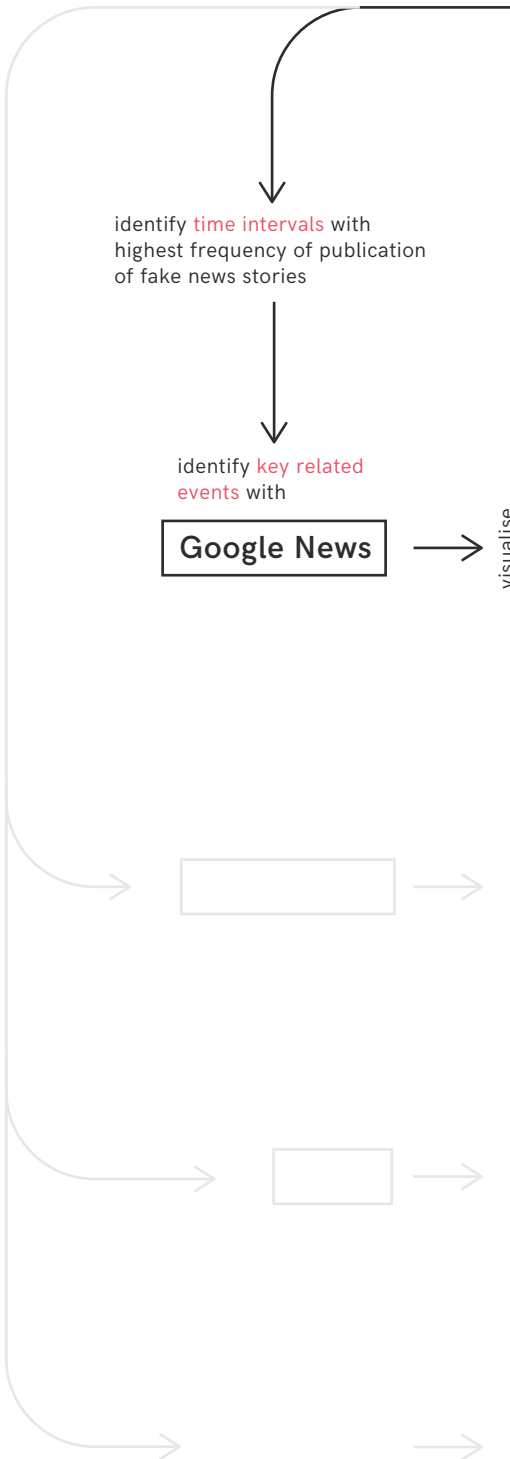
identify time intervals with highest frequency of publication of fake news stories

identify key related events with





a

WHICH MEDIA AND POLITICAL EVENTS ARE SUCCESSFUL IN SETTING THE FAKE NEWS AGENDA?



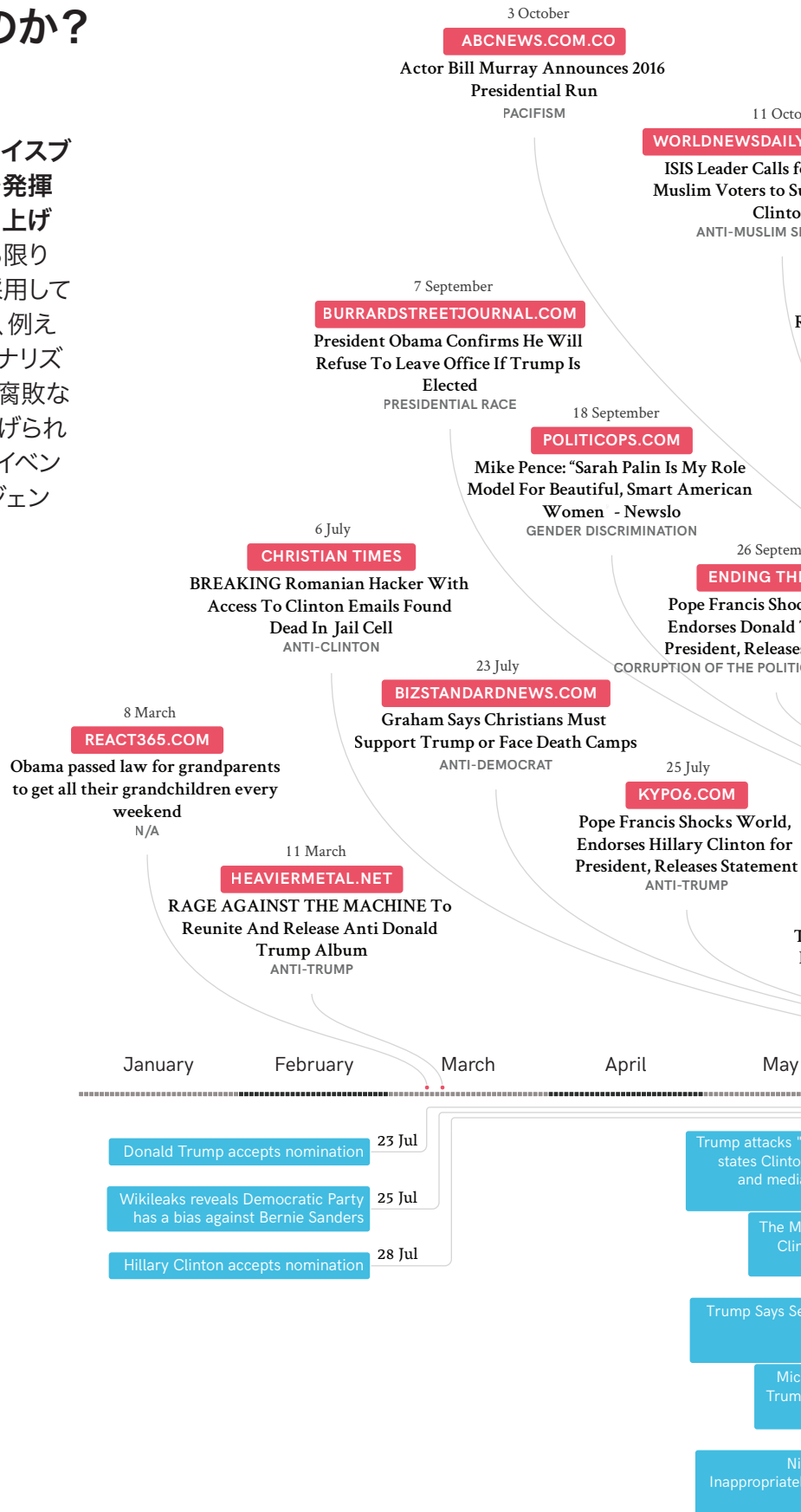
a. フェイクニュース記事が取り上げているテーマやイベントを特定する

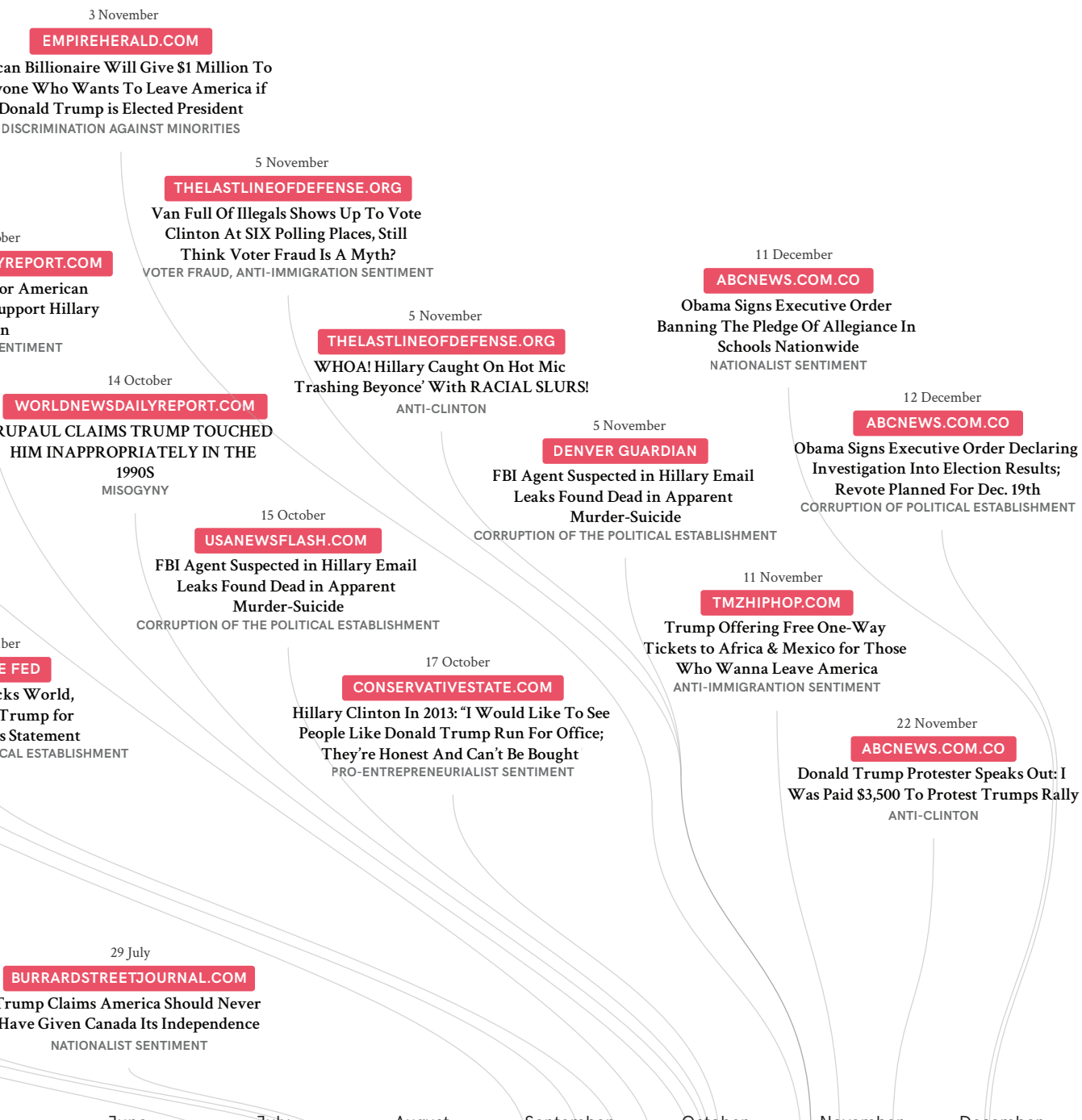
この分析では、各記事の内容を質的に検討するとともに、記事の論点に関わる、記事の発表前後に起きた政治／メディアイベントを特定する。この作業により、フェイクニュースの流通を担う読者がどんなトピックであれば盛り上がるのか、理解を深めることができる。

- ◇ フェイクニュース記事が元のURLに残っていない場合、非営利団体インターネット・アーカイブのサービス  **WayBack Machine** で、アーカイブされたURLが利用できないか確認してみるといいだろう。
- ◇ 記事の発表前後に起きた主な関連イベントを特定するには、 **グーグルニュース検索** などのニュースアグリゲーターや記事のアーカイブを活用できる。
- ◇ フェイクニュース記事と、同時期に起きた政治／メディアイベントを注釈付きの時系列にまとめることで、その関連性を考える出発点になるだろう。

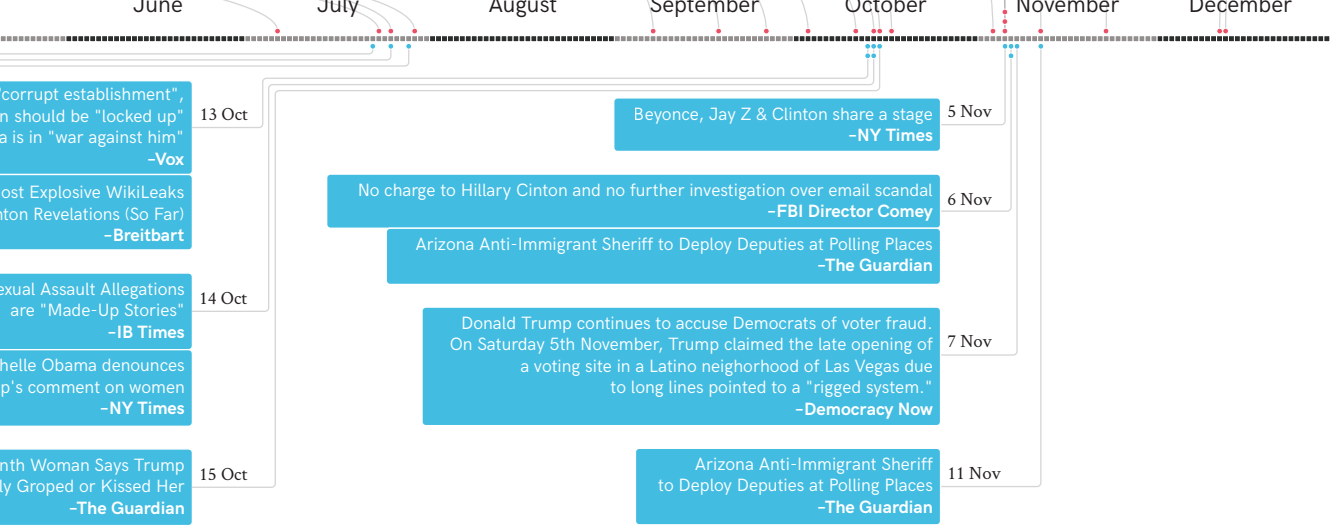
どのメディアと政治イベントが フェイクニュースのアジェンダ 設定に成功しているのか？

2016年の米大統領選に関し、フェイスブック上で最も高いパフォーマンスを発揮したフェイクニュースと、そこで取り上げられたイベントのタイムライン。見る限りでは、成功したフェイクニュースが採用しているのはポピュリズム的なテーマで、例えば反エスタブリッシュメント、ナショナリズム、反移民感情のほか、女性嫌悪や腐敗など、候補者の弱みと見られる点が挙げられる。7月末、10月半ばと11月初旬のイベントの多くが、フェイクニュースの「アジェンダ」設定に成功している。





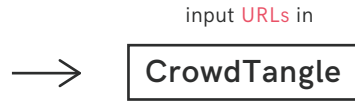
Fake news stories



Events


START

list of 22 URLs of political fake news stories
Source: BuzzFeed News

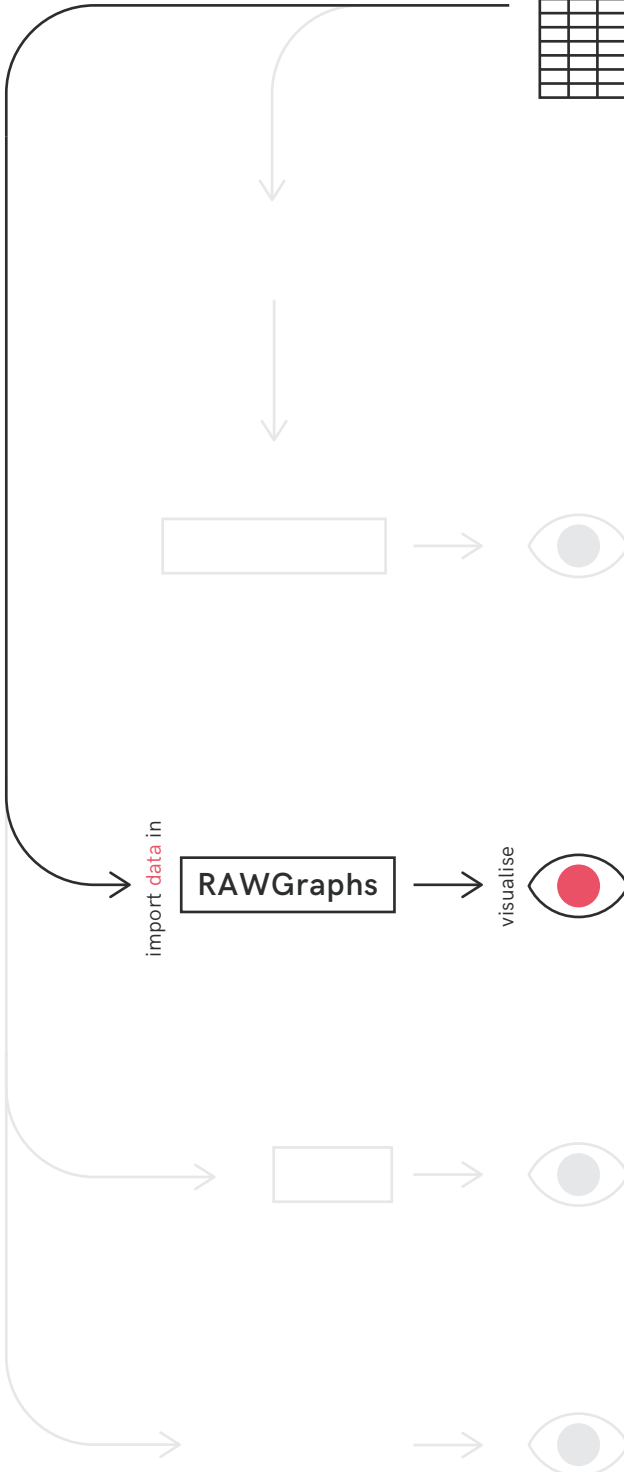


↓

output data



- > Fake news story URLs
- > Facebook pages and groups that share the URLs
- > Number of interactions per each page or group
- > Date of sharing of the story



b
WHICH FACEBOOK PAGES AND GROUPS PROMOTED THE HIGHEST NUMBER OF FAKE NEWS STORIES?

b. 記事をシェアしているフェイスブックのページやグループを特定する

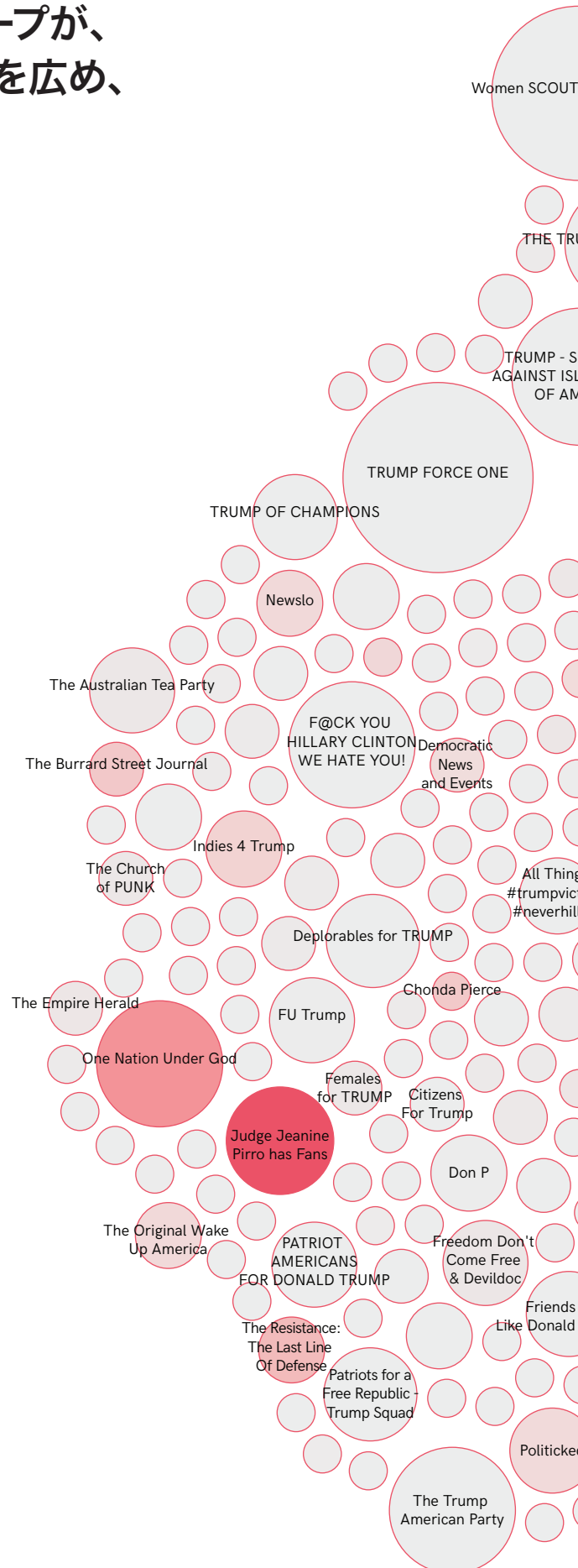
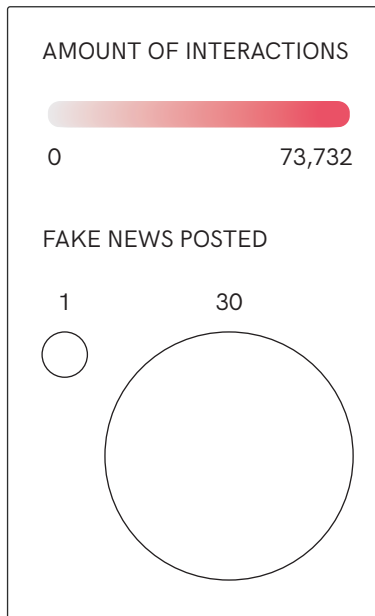
ここでは🔧 **CrowdTangle** のブラウザ拡張機能など、ソーシャルメディアのモニタリングツールを使う。フォロワー数、投稿ごとのインタラクション数（「いいね!」やシェアなどの数）が、フェイクニュースをシェアしているページ／グループの名前とともにスプレッドシートに記録される。

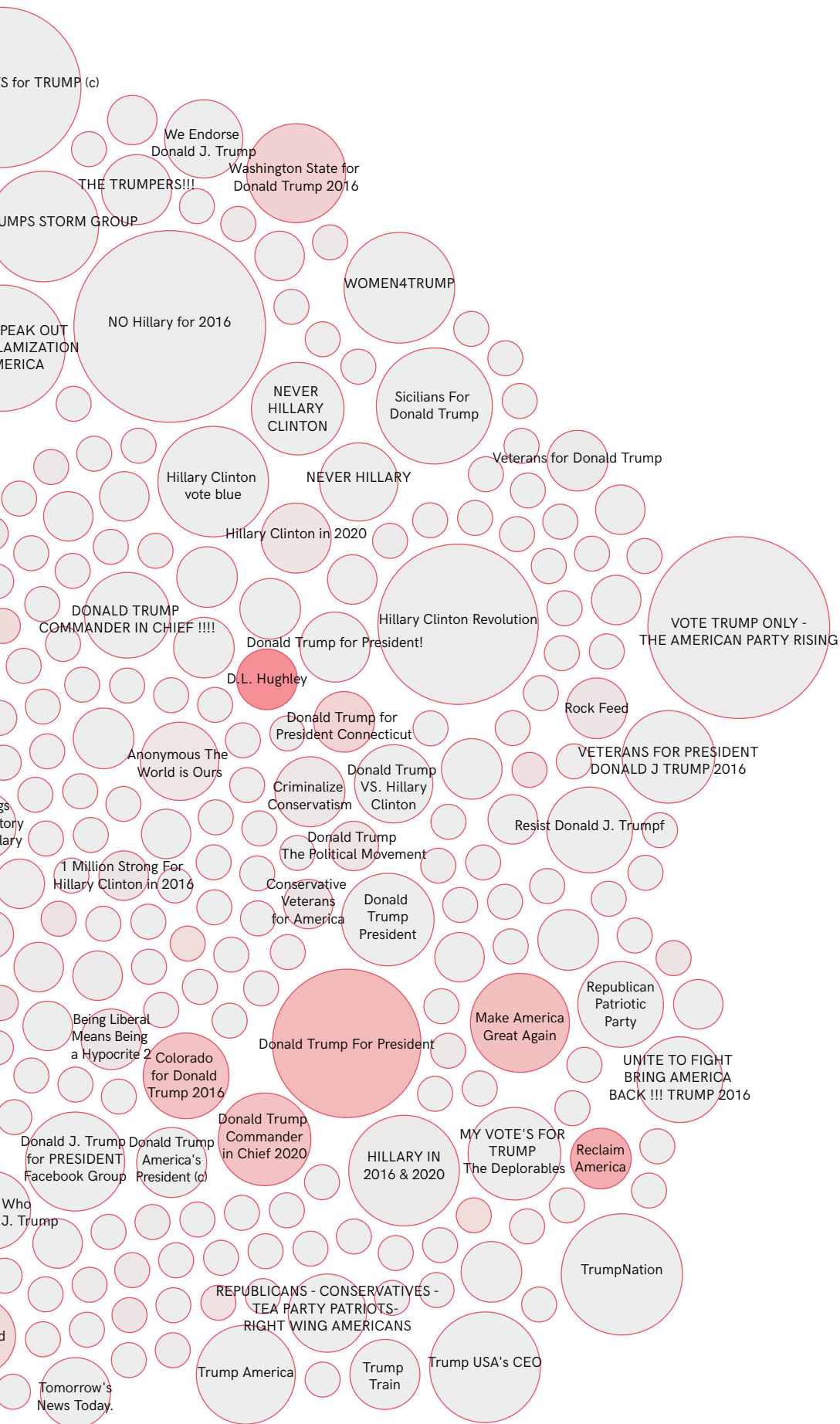
- ◇ フェイクニュースは多くの別サイトに再投稿されうるといふ点に注意が必要だ。したがって最初に、オリジナルのURLをシェアしているページ／グループだけを記録するのか、それとも、コピーされたサイトをシェアしているものも含むのか、方法論を定めておく必要がある。
- ◇ 最も多くのフェイクニュースをシェアするページ／グループに加え、各ページ／グループでのインタラクションの総数も記録しておく。
- ◇ 🔧 **CrowdTangle** はフェイスブックの場合、認証済みページとフォロワー数12万5千人以上のページから、最も人気のある公開設定の投稿500件を返す[1]。
- ◇ **サークルパッキング**📊 で、フェイクニュースをシェアしたページ／グループ、シェアされた記事数、インタラクション数を表現できる。データ可視化サービス 🔧 **RAWGraphs** を使うといいだろう。

[1]
See, Crowd-
Tangle's
"Frequently
Asked Questions", available
at: [https://
apps.crowd-
tangle.com/](https://apps.crowdtangle.com/)

どのフェイスブックページ/グループが、最も多くのフェイクニュース記事を広め、エンゲージメントを生み出したか

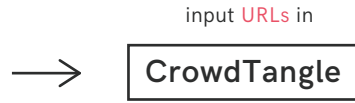
フェイクニュースをシェアした公開ページ/グループについて、シェアした記事数が円のサイズに、フォロワー数が色に反映されている。各ページは同じ記事を複数回シェアしている場合がある。最も多くの記事をシェアしているのは、主にトランプ支持者や反ヒラリーのグループだ。フェイクニュースをめぐるインタラクションが最も多かったのは、共和党のテレビコメンテーター、ジェニー・ピロのファンページだ。





START

list of 22 URLs of political fake news stories
Source: BuzzFeed News

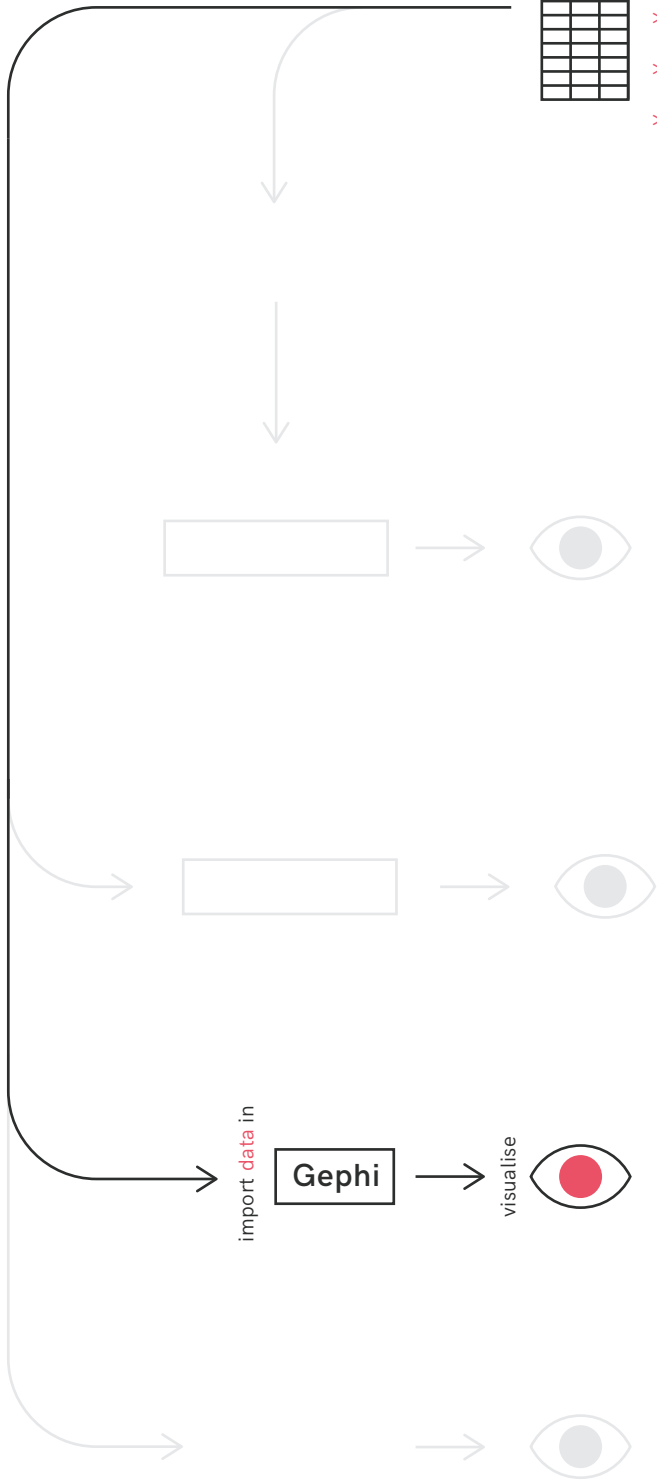


↓

output data






- > Fake news story URLs
- > Facebook pages and groups that share the URLs
- > Number of interactions per each page or group
- > Date of sharing of the story



c
DO FACEBOOK PUBLICS HAVE PREFERRED STORY THEMES?

c. ユーザーが好む記事のテーマを明らかにする

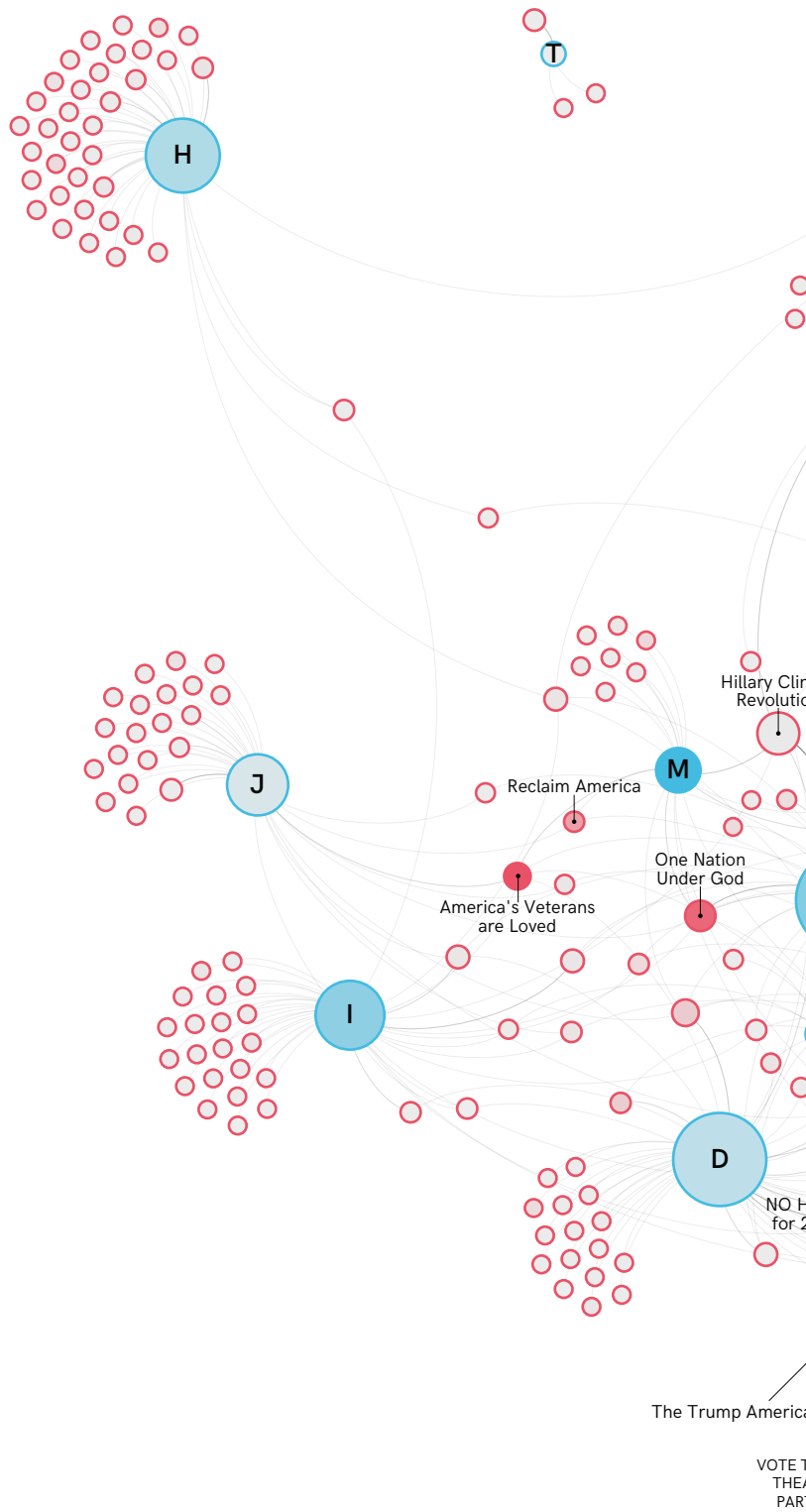
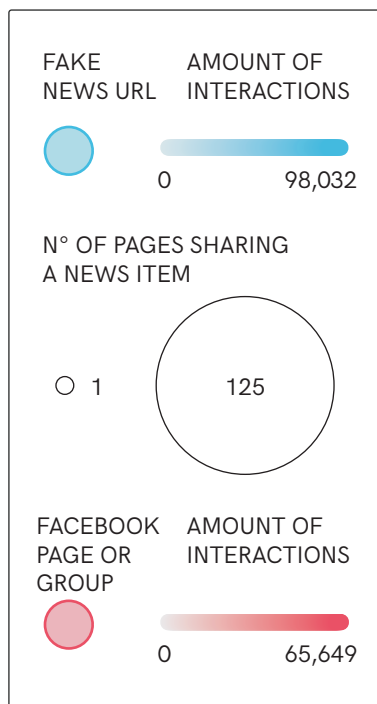
どのようなテーマの記事が読まれているのかを探り、そうした記事で盛り上がる読者を質的にプロフィールするには、フェイスブックの公開ページ／グループが共通してシェアしている記事を手掛かりにネットワーク分析をすればよいだろう。

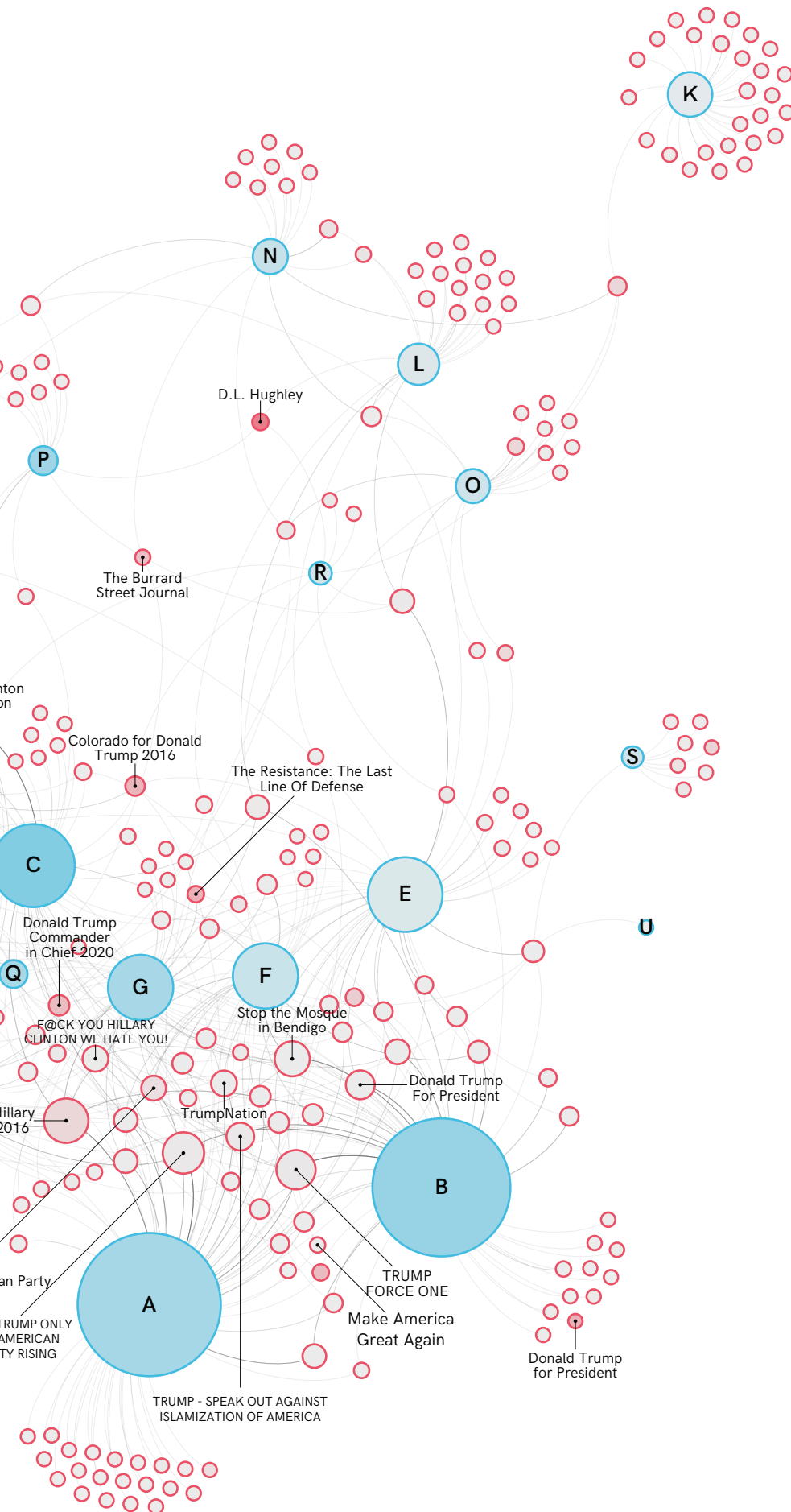
- ◇  **CrowdTangle**のブラウザ拡張機能で抽出したデータセットから、ネットワークファイルを作ることができる。あるフェイスブックページ／グループがフェイクニュースを投稿するたび、その記事との間にリンクが一つ形成される。
- ◇ パリ政治学院メディアラボが提供するツール、 **Table2Net**でCSVファイルをネットワークファイルに変換すれば、分析にはオープンソースのグラフ可視化ツール **Gephi**が活用できる。Gephiの「ForceAtlas2」[2]といった力学モデルの描画アルゴリズムが、抽出データの可視化に役立つだろう。
- ◇ ユーザーがどの記事で最も活気づいたかに加え、好まれる記事テーマはあるのかを明らかにする。

[2]
See, Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann and Mathieu Bastian, “ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software”, June 2014, PLoS ONE: <http://journals.plos.org/plosone/arti->

フェイスブックユーザーが好む記事テーマはあるか？

共通してシェアしたフェイクニュースでつながる、フェイスブック公開ページ/グループのネットワーク。注目すべきはネットワークの核だ。反ヒラリーの記事に盛り上がるトランプ支持者関係の一連のページ/グループが核を形成している。



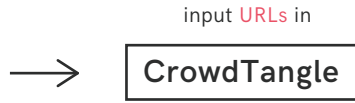


Fake News Headlines

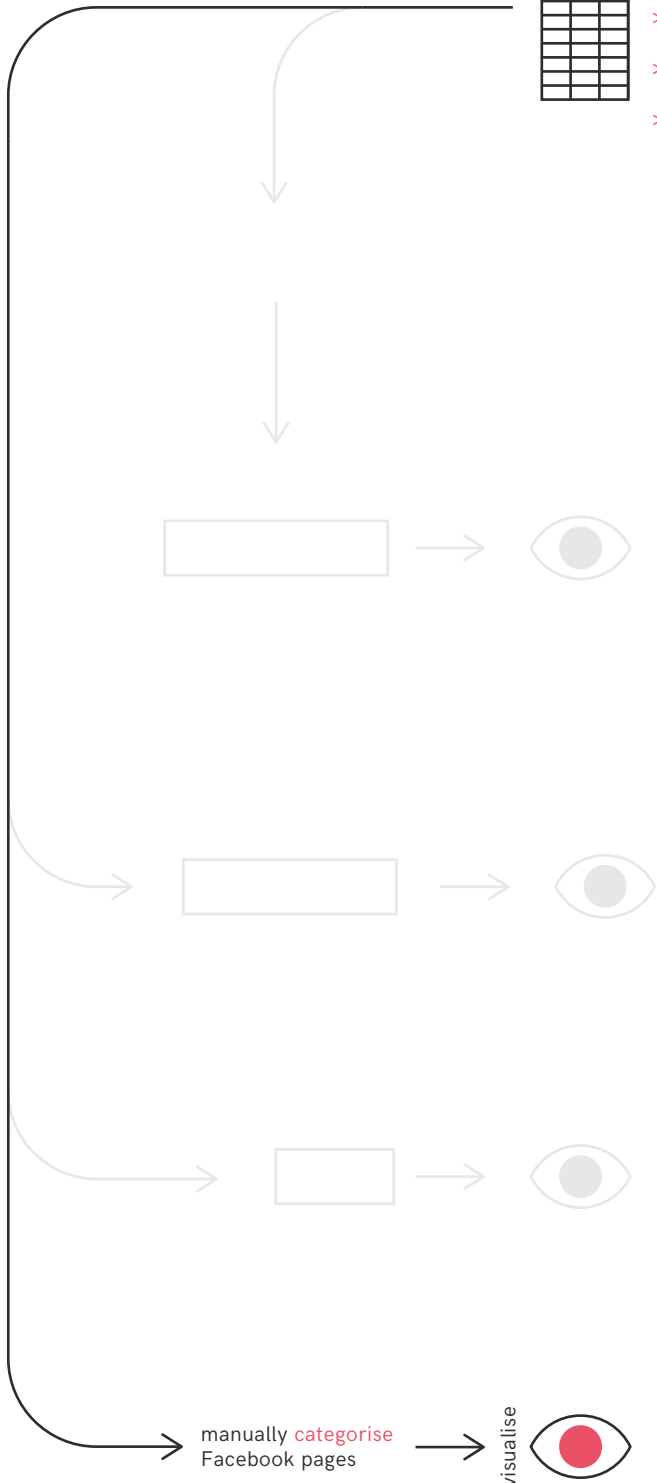
- A FBI Agent Suspected in Hillary Email Leaks Found Dead in Apparent Murder-Suicide
- B Hillary Clinton In 2013: "I Would Like To See People Like Donald Trump Run For Office; They're Honest And Can't Be Bought"
- C ISIS Leader Calls for American Muslim Voters to Support Hillary Clinton
- D Donald Trump Protester Speaks Out: "I Was Paid \$3,500 To Protest Trump's Rally"
- E Obama Signs Executive Order Declaring Investigation Into Election Results; Revote Planned For Dec. 19th
- F WHOA! Hillary Caught On Hot Mic Trashing Beyonce' With RACIAL SLURS!
- G Van Full Of Illegals Shows Up To Vote Clinton At SIX Polling Places, Still Think Voter Fraud Is A Myth?
- H RAGE AGAINST THE MACHINE To Reunite And Release Anti Donald Trump Album
- I Obama Signs Executive Order Banning The Pledge Of Allegiance In Schools Nationwide
- J BREAKING Romanian Hacker With Access To Clinton Emails Found Dead In Jail Cell
- K Actor Bill Murray Announces 2016 Presidential Run
- L Pope Francis Shocks World, Endorses Hillary Clinton for President, Releases Statement
- M Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement
- N Trump Claims America Should Never Have Given Canada Its Independence
- O Mike Pence: "Sarah Palin Is My Role Model For Beautiful, Smart American Women" - Newslo
- P RUPAUL CLAIMS TRUMP TOUCHED HIM INAPPROPRIATELY IN THE 1990S
- Q President Obama Confirms He Will Refuse To Leave Office If Trump Is Elected
- R Graham Says Christians Must Support Trump or Face Death Camps
- S African Billionaire Will Give \$1 Million To Anyone Who Wants To Leave America if Donald Trump is Elected President
- T Trump Offering Free One-Way Tickets to Africa & Mexico for Those Who Wanna Leave America
- U Obama passed law for grandparents to get all their grandchildren every weekend

START

list of 22 URLs of political fake news stories
Source: BuzzFeed News



- > Fake news story URLs
- > Facebook pages and groups that share the URLs
- > Number of interactions per each page or group
- > Date of sharing of the story



manually categorise Facebook pages



d
WHAT KINDS OF PUBLICS ARE ENERGISED BY FAKE NEWS?

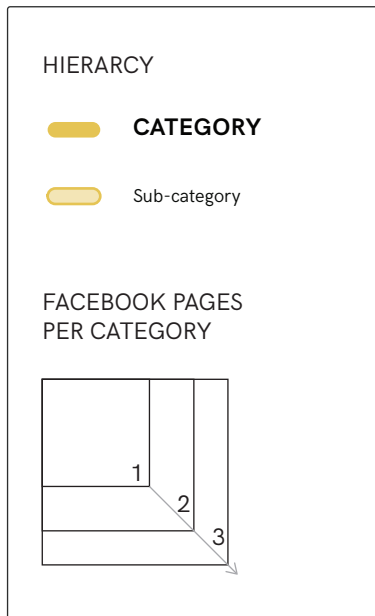
d. フェイクニュースで盛り上がっている読み手をプロフィールする

フェイクニュースをシェアした全てのフェイスブックページについて、「ページ情報」の紹介文を質的に分析する。

- ◇ ページ紹介文に現れるテーマを特定するには、文中に現れる語彙を分類するアプローチが有用だろう。カテゴリーは、包括的なもの(例えば「草の根活動」)から細かなもの(「反エスタブリッシュメント」)まで記録しておこう。
- ◇ 同じカテゴリーに属する全てのページのフォロワー数を合計する。
- ◇ **ツリーマップ図** にすれば、各カテゴリーの重みや階層を表現できる。🔧 **RAWGraphs** を使うとよい。

フェイクニュースで活気づいているのはどんなユーザーか？

フェイクニュースで盛り上がったフェイスブック上の読者の類型。記事をシェアしたページのカテゴリ分けは人の手で行った。各方面の草の根活動家、選挙立候補者の熱心な支持者、芸能人といったカテゴリーが目立つ。



WORKERS

Truckers	Small Business	Steelworkers
----------	----------------	--------------

OTHER

Italian Left	Climate Skeptics	Page Dedicated to the Philippines
--------------	------------------	-----------------------------------

SOCIAL CHANGE

Pro Latino	Human rights	Progressives
Against Gun Violence		

EARLY CAMPAIGNERS

Early 2020 Bernie Campaigners	Early 2020 Trump Campaigners	Early 2020 Hillary Campaigners
-------------------------------	------------------------------	--------------------------------

DISCUSSION SPACE

Political Commentary	Anti establishment
Political Commentary	
Anti-Democrat	

BUSINESS

Punk Tattoo	Fashion
Business	Motivational Speaker
Dating app	

OCCUPY MOVEMENT

WOMEN-FOCUSED

NEWS

News	Trending News	Liberal News	Con News
Critical News	Liberal news	News / News Aggregators	Rep News
News	News Aggregator	Alternative News	
Non-PC News			

Hillary Loyalists

Bernie Loyalists

Spiritual Community	Democr
Progressive Community	
Survivalist	
Pro-Black	Republi
	Religiou

GRASSROOTS ACTIVISTS/ CITIZEN VIGILANTES

For Free Society	Against Ted Cruz	Pro Bernie	For Disaster Preparedness	Against Islam	Against Financial Terrorism	Pro Latino
For Liberty and Property	White Lives Matter	For Information Freedom	Against Trump	Against Republicans		
Against Gun Owners						
Against Mainstream Media	Anti-conservatism			US Patriots		
Against Gun Control Advocates	Anonymous					
Against Corruption						
Pro-Trump						

PUBLIC FIGURE

Artist	Porn Actress
Actor	Republican Politician
Journalist	

ENTERTAINERS

LOYALISTS

Trump Loyalists

MUSIC/ENTERTAINMENT

COMMUNITY

Lat Community	Pro Military
ican Community	
is	

CLICKBAIT

おすすめ手法

このレシピから見えてくるのは、フェイクニュースを自分に都合よく解釈し、盛り上がっているユーザーの姿だ。

フェイクニュースの持つインパクトをより厚みを持って説明するには、そのバイラルな性質(単純なエンゲージメントの数値)だけではなく、読み手の行動に注目することも有益だろう。



フェイクニュースの 伝達経路を いかに追跡するか？

はじめに

このレシピでは、フェイクニュース記事を起点とし、そのURLを明らかにする。例として2016年の米大統領選に関する二つの記事を利用する。一つは反移民感情を利用しようとした「Trump Offering Free One-Way Tickets to Africa & Mexico for Those Who Wanna Leave America (トランプ候補: 米国から移住したい人にアフリカとメキシコ行きの片道切符を無償提供)」。二つめは反トランプ派の意識をあおった「Rage Against the Machine to Reunite and Release Anti Donald Trump Album (バンドRage Against the Machineが再結成、反トランプを歌うアルバムを発表)」。

このレシピは二つのステップがある。まずフェイクニュース元記事が掲載されたフェイスブックページ/グループを発見し、その動きを把握すること (👁️ a)。次に同じ分析を、その記事が再掲載された全てのウェブサイトに対して行う (👁️ b)。

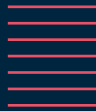
START

select a fake news story
you want to trace

design queries to identify
URLs where story is published

Google Search

compile list of URLs
where story is published



input each URL in

CrowdTangle

output data for each URL

- > Fake news story URLs
- > Facebook pages and groups that share the URLs
- > Number of followers per page or group
- > Date of sharing of the story

input data to

RAWGraphs

visualise



b

input URL in

CrowdTangle

output data

- > Fake news story URLs
- > Facebook pages and groups that share the URLs
- > Number of followers per page or group
- > Date of sharing of the story

input data to

RAWGraphs

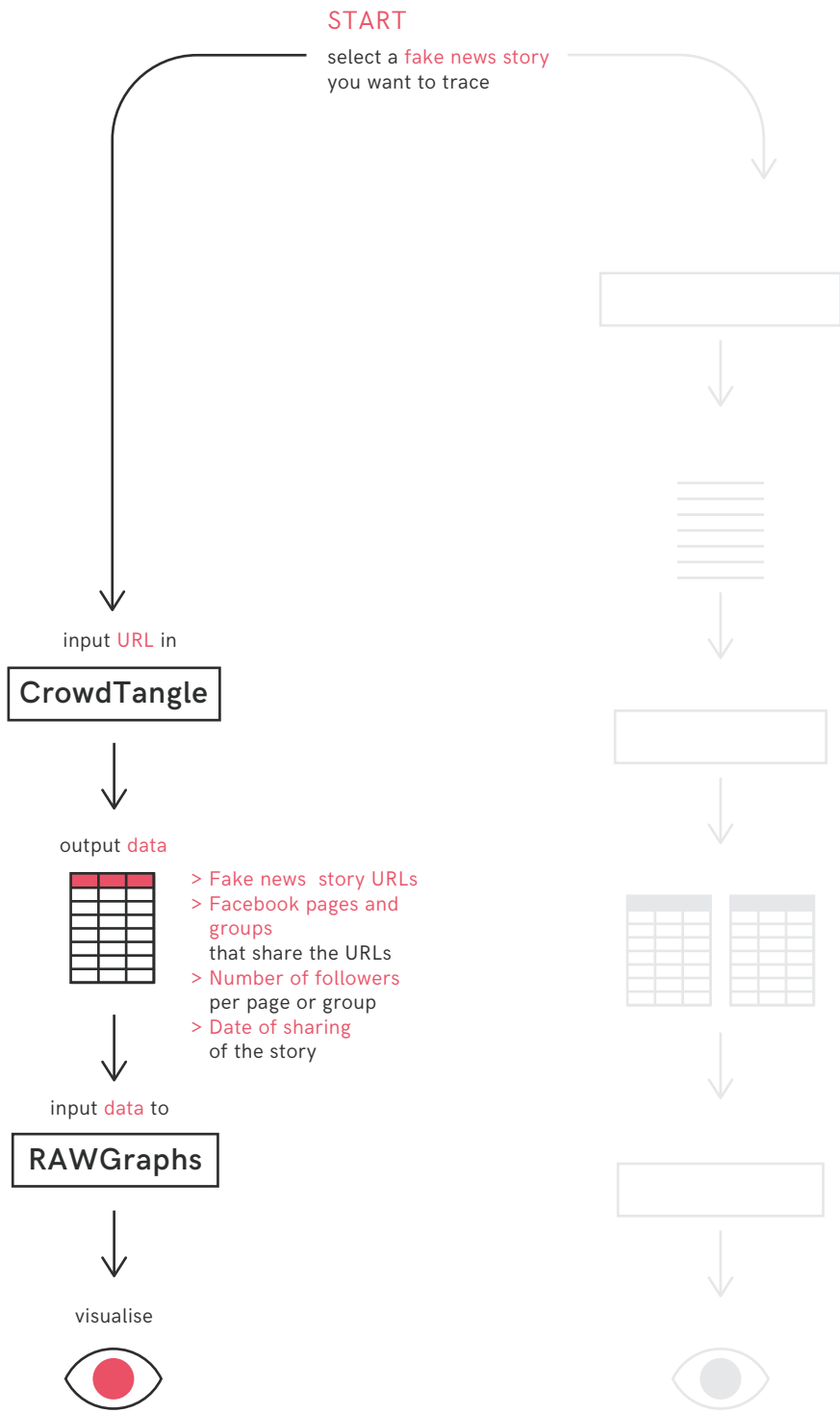
visualise



a

**HOW DOES THE STORY
"RAGE AGAINST THE
MACHINE TO REUNITE
AND RELEASE ANTI
DONALD TRUMP ALBUM"
TRAVEL ON FACEBOOK?**

**HOW DOES THE STORY
"TRUMP OFFERING FREE
ONE-WAY TICKETS TO AFRICA
& MEXICO FOR THOSE WHO
WANNA LEAVE AMERICA"
AND ITS DEBUNKED VERSIONS
TRAVEL ON FACEBOOK?**



a

HOW DOES THE STORY "RAGE AGAINST THE MACHINE TO REUNITE AND RELEASE ANTI DONALD TRUMP ALBUM" TRAVEL ON FACEBOOK?

a. 元記事のURLから、この記事が共有されたフェイスブックページ/グループを明らかにする

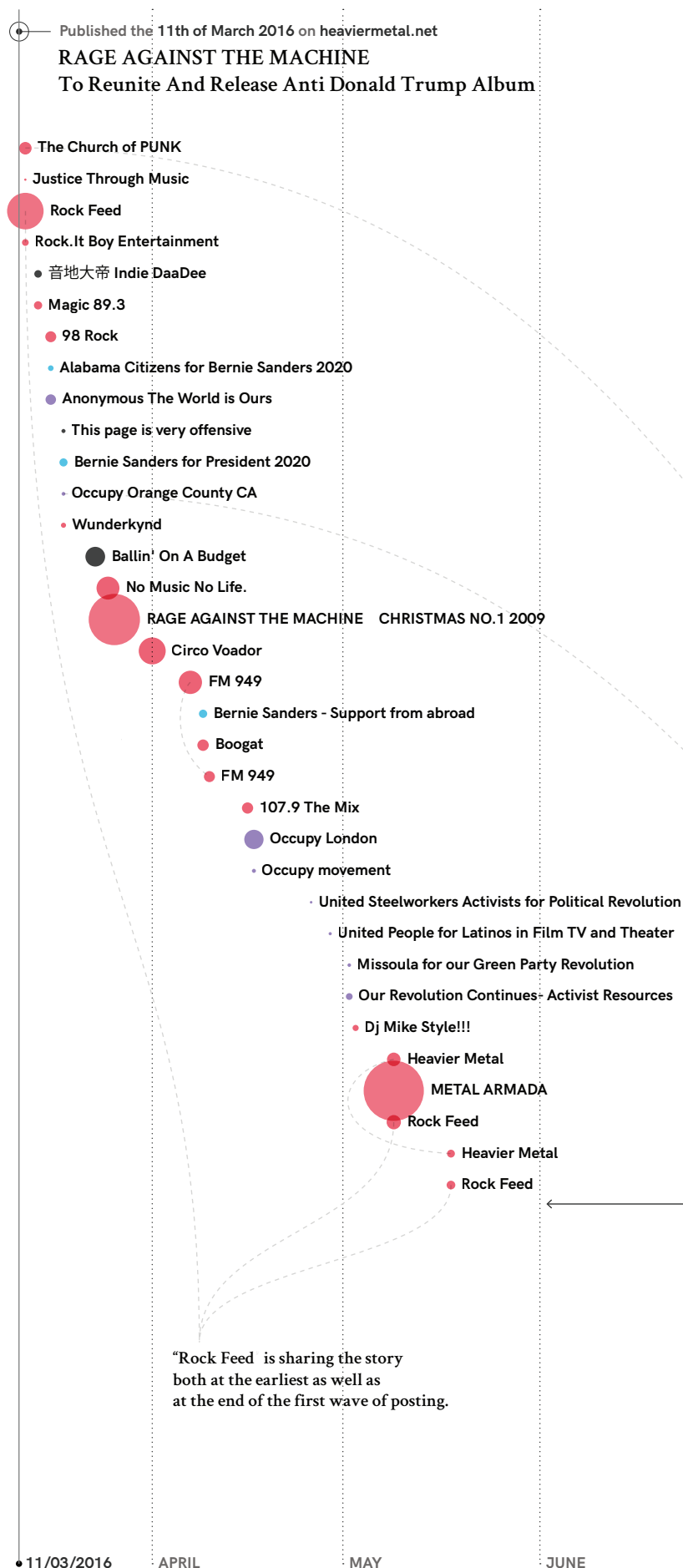
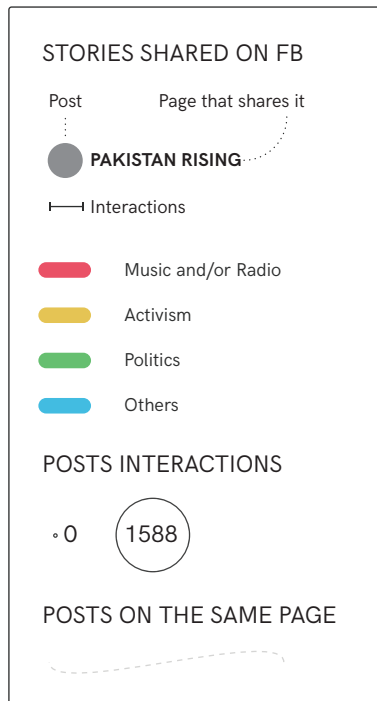
フェイクニュースを共有したページ/グループは、一般公開されていれば、🔗 **CrowdTangle**などソーシャルメディアモニタリングツールのブラウザ拡張機能で突き止められる。

- ◇ フェイスブックページ/グループ、フォロワー数、インタラクション、フェイクニュース記事の共有日を調べ、記事のURLごとにスプレッドシートにまとめる。
- ◇ フェイクニュースの拡散の動きを調べるには、タイムラインからページやグループをまたいだ伝達経路を時系列で辿る。🔗 **RAWGraphs** を使って可視化の基礎を作るのも良い。どんな人が記事を読んでいるのか、暴露記事(フェイクだと指摘する記事)は拡散に影響があるのか、観察する[1]。
- ◇ もう一步進めて、自分たちの主張を強化するために、どのようにフェイクニュースが使われているか、記事がシェアされる文脈において質的な分析をすることもできる。肯定的なのか、もしくは「この記事はフェイクだ」と書かれているのか。ただし、拡散したフェイクニュースを共有したフェイスブックページは、インターフェースやAPIから消されており、こういった分析は難しい場合もある。

[1]
この分析は、CrowdTangleが対応しているものは有効だが、フェイスブック上のフェイクニュースがすべてこの手法でカバーできるわけではない。詳しくはCrowdTangle dataのP.27へ。

「バンドRAGE AGAINST THE MACHINEが再結成、反トランプを歌うアルバムを発表」の記事はフェイスブック上でどのように伝播したか

記事がフェイスブックページ/グループにまたがって伝達した経路をCrowdTangleを使って抽出する。記事は音楽やエンターテインメント関連の英文のフェイスブックページで、2016年3~6月に風刺記事としてもっとも活発にシェアされたが、米大統領選後の11月にイタリア語の音楽や政治ページに再掲載され、再拡散したことが分かった。





Replace with "No prominent sharing of the story URL on public Facebook pages and groups from late May to early November according to CrowdTangle data"

The story reappears on Facebook on the 11th of November, shared by the page "Apartment Khunpa"

- Apartment Khunpa
- The Guitar Mag
- Banda Bassotti
- Marcello Belotti - Sinistra Italiana EU
- ROCK Mundial
- The Church of PUNK
- Radioactivo 98.5
- Los Angeles Punk Museum
- Veterans Against Rep. Ignorance
- FU Trump
- Occupy Orange County CA

JULY

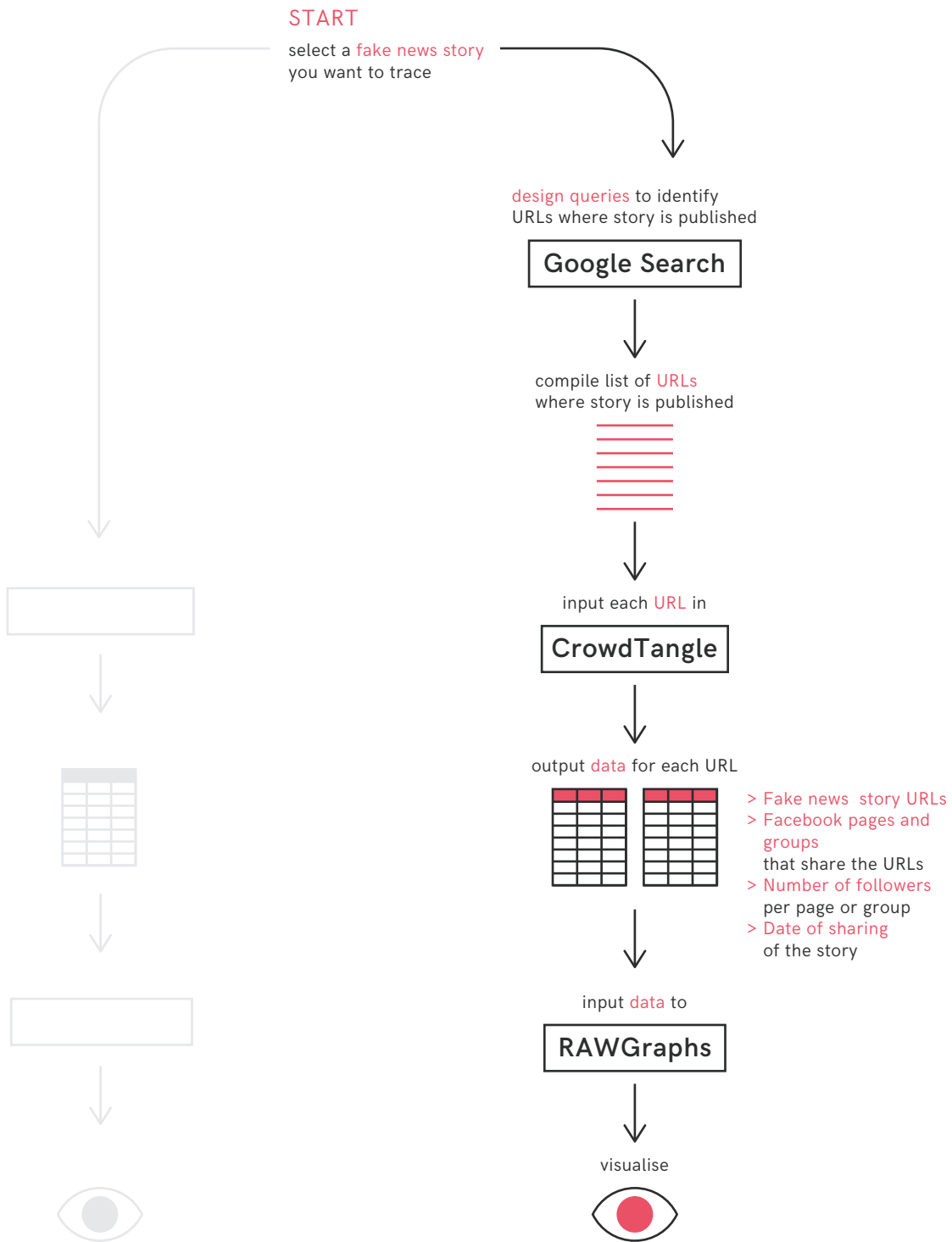
AUGUST

SEPTEMBER

OCTOBER

NOVEMBER

DECEMBER



b

HOW DOES THE STORY "TRUMP OFFERING FREE ONE-WAY TICKETS TO AFRICA & MEXICO FOR THOSE WHO WANNA LEAVE AMERICA" AND ITS DEBUNKED VERSIONS TRAVEL ON FACEBOOK?

b. コピーされたフェイクニュースを探す方法

フェイクニュース記事は多くのウェブページに再掲載される。従って、記事のオリジナルURLだけでなく、さまざまなサイトへの再掲載もすべて辿ると、1.2aの分析がより豊かになる。

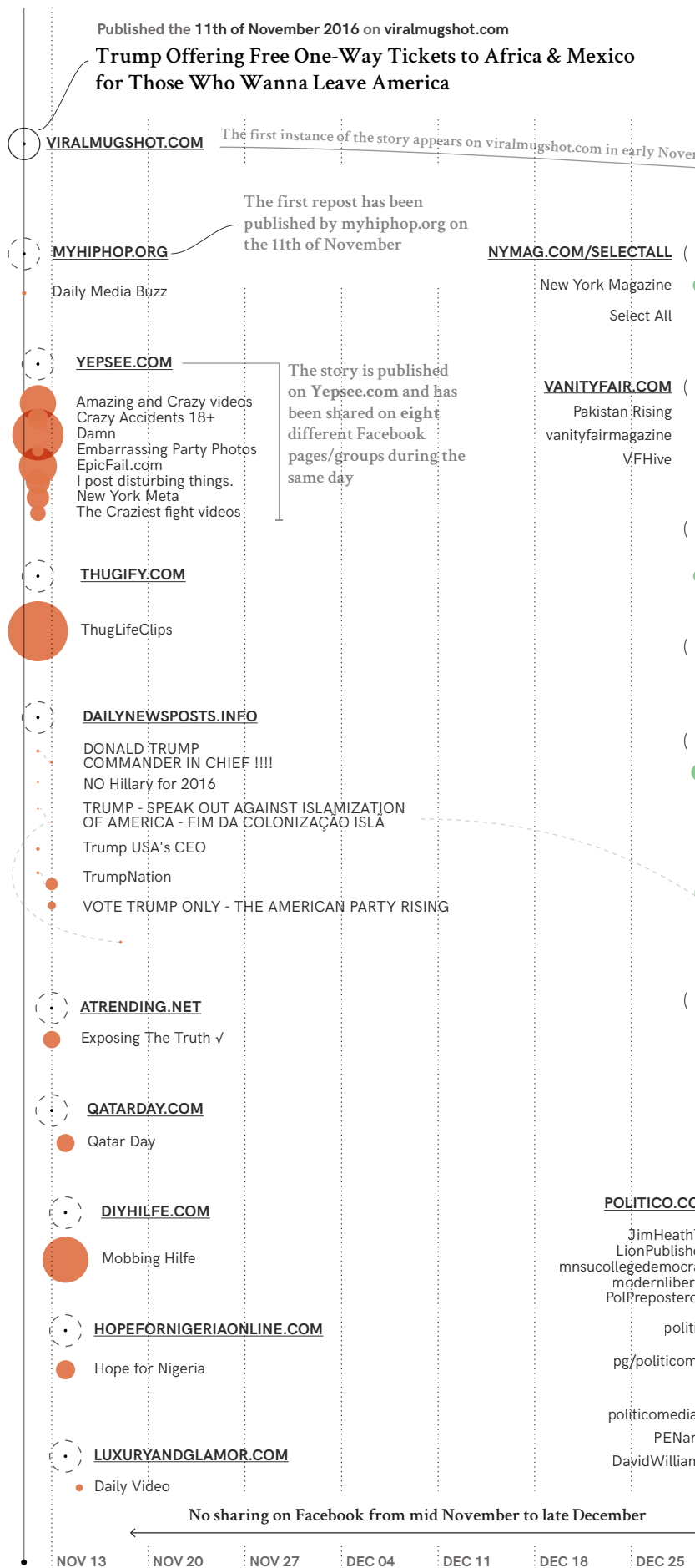
- ◇ フェイクニュース記事が再掲載されているウェブサイトや、暴露記事を特定するには、
✂ **グーグルウェブ検索**などの検索エンジンのリサーチブラウザ^[2]を使って、フェイクニュース記事のタイトルを検索し、該当するURLを抽出する。
- ◇ URLを✂ **CrowdTangle**などのモニタリングツールに入力し、フェイクニュース記事の再掲載や、暴露記事を共有したフェイスブックページやグループのリストが得られる^[3]。
- ◇ これらのページの動きを時間軸に描くことで、フェイスブック上の伝達経路はフェイクニュースの元記事によって異なるのか、暴露記事は認知されているのか、確かめられる。

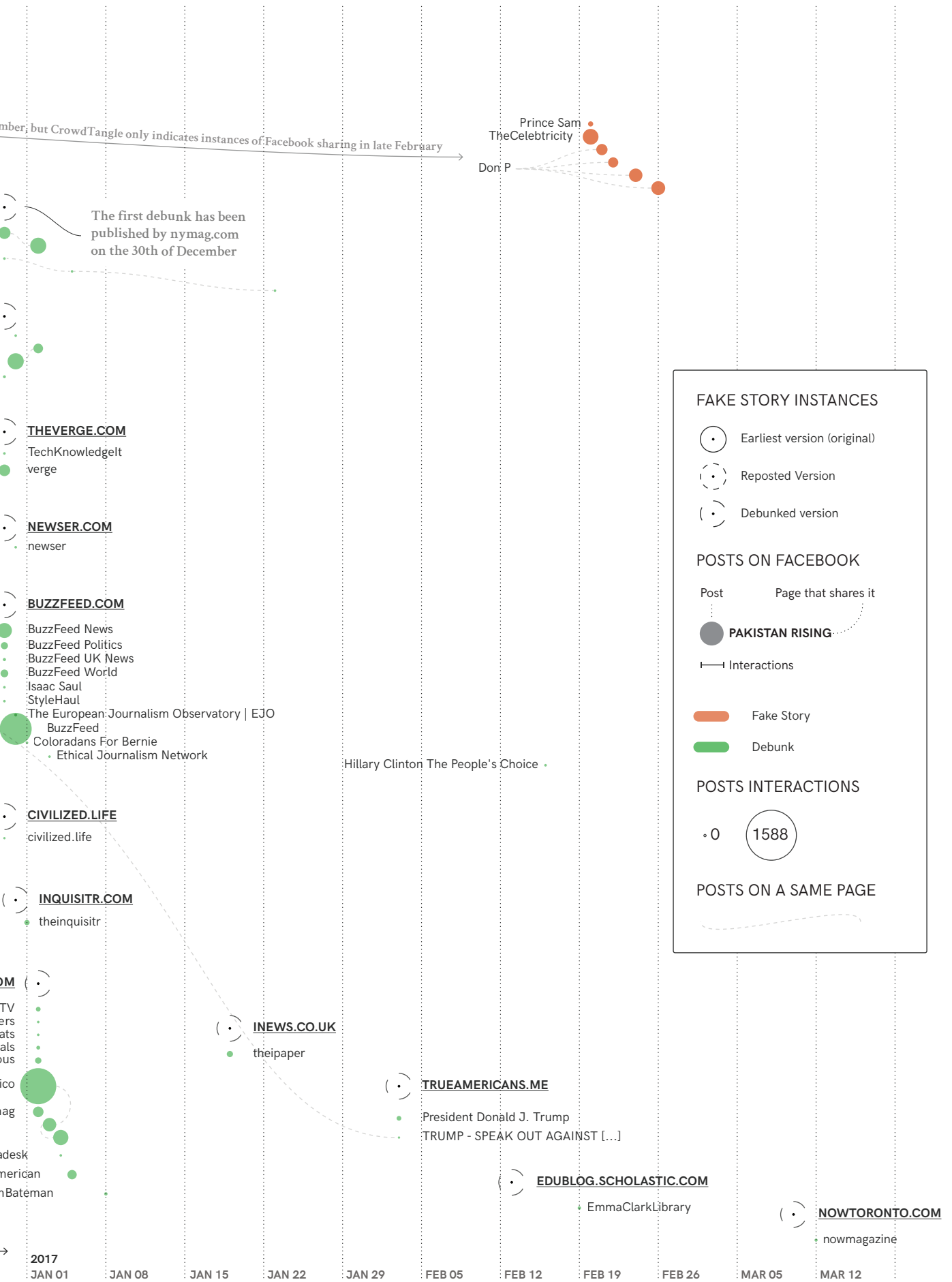
[2]
See instructions on how to set up a research browser in this video tutorial: <https://www.youtube.com/watch?v=bj65X-r9GkJM>

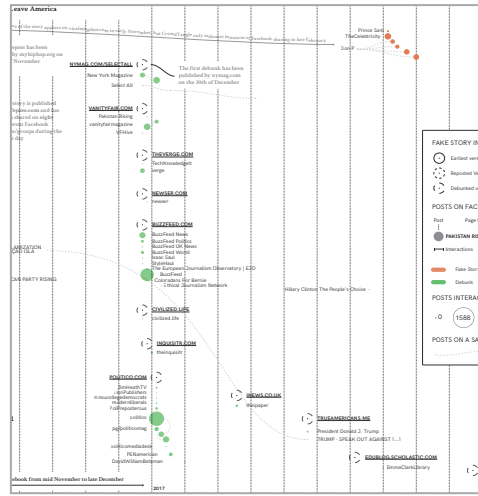
[3]
Please see the note on p.43 about the limitations of using CrowdTangle data.

記事「トランプ候補：米国から移住したい人にアフリカとメキシコ行きの片道切符を無償提供」と暴露記事はフェイスブック上でどのように伝播したのか？

記事「トランプ候補：米国から移住したい人にアフリカとメキシコ行きの片道切符を無償提供」と、ネット及びフェイスブック掲載の時系列。オリジナル記事が出た翌週、複数のクリックベイトサイトにコピーされたことで、フェイスブック上でさらに拡散した。フェイクニュースサイトで共有されると政治的関心のある読み手に影響がある一方、クリックベイトサイトで共有されると、バイラルメディアでさらにクリックベイトとして共有される現象が起きていた。従ってフェイクニュース自体に影響を受ける層と、暴露記事に影響を受ける層は異なることが分かった。







おすすめ手法

このレシピでは、フェイスブック上のフェイクニュースの伝達経路、拡散中の異なる段階、さらに拡散のきっかけ、及び媒体が理解できる。



ファクトチェックの取り組みは フェイクニュースの読み手に 届くか？

はじめに

このレシピは、フェイクニュース記事リストを起点とする。このリストは既存のものでも良いし、自分で作っても良い。ここでは、Buzzfeed Newsが2016年の米大統領選に関連して作った、22本の記事リストを利用する。

このレシピには二つのステップがある。まずは個々のフェイクニュースに関して、訂正情報や暴露記事を載せているウェブページのURLを特定すること(1.3a)。二つ目は、それらに対する、フェイスブックの公開ページの反応を調べる(👁b)。

START

List of 22 URLs of political fake news stories
Source: BuzzFeed News

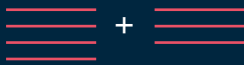
query
"Fake News Title 1" + Fake
"Fake News Title 2" + Fake
"... " in

Google Web Search

Retain the top ranked URL
of a debunk per fake news story



merge all URLs in a single list



input URLs in

CrowdTangle

output data



- > URLs of fake news story or debunked version thereof
- > Facebook pages and groups that share either the fake news story or its correction
- > number of followers per page or group

input data to

RAWGraphs

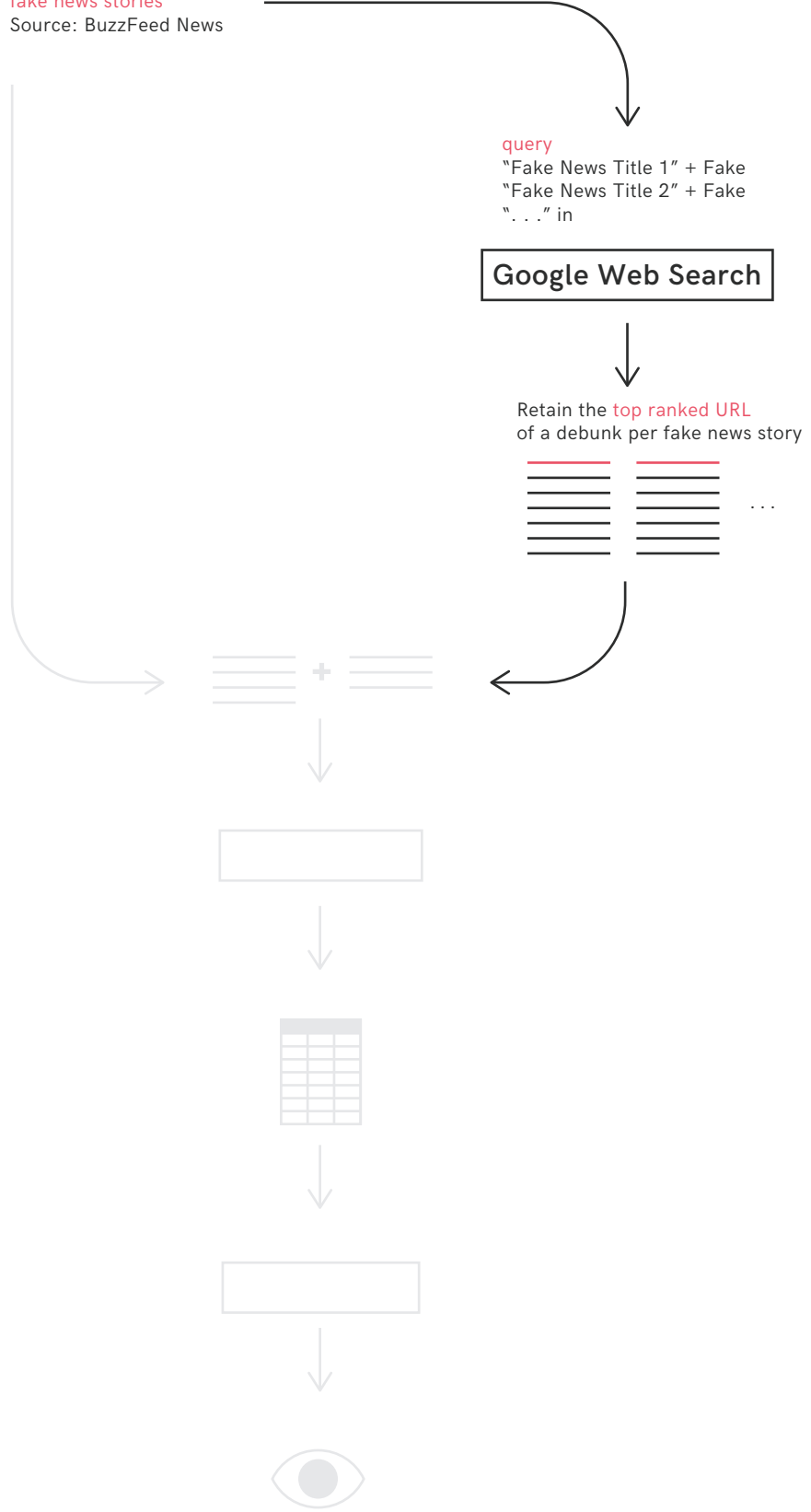
visualise




ARE DEBUNKING WEB PAGES ACKNOWLEDGED BY THE PUBLICS OF FAKE NEWS?

START

List of 22 URLs of political fake news stories
Source: BuzzFeed News



a. 暴露記事サイトを特定する

暴露記事を掲載しているウェブページを特定するには、 **グーグルウェブ検索** を活用するのがよいだろう。さらにそのフェイクニュースのキーワードを使って、ファクトチェックサイトを検索するのも良い。

- ◇ 暴露記事を見つけるには、リストの個々の記事について検索を工夫しなければならない。例えばフェイクニュースのタイトルを“ ”に入れ、続いてfakeと入力するなど
(例：“Trump Offering Free One-Way Tickets to Africa & Mexico for Those Who Wanna Leave America” fake)。
- ◇ 検索ランキングを訂正記事の信ぴょう性の指標とし、検索トップのURLを選ぶ。
- ◇ この作業を22本のフェイクニュース記事について行い、暴露記事のリストを作る。

START

List of 22 URLs of political fake news stories
Source: BuzzFeed News

query
"Fake News Title 1" + Fake
"Fake News Title 2" + Fake
"... " in

Google Web Search

Retain the top ranked URL
of a debunk per fake news story

merge all URLs in a single list

input URLs in

CrowdTangle

output data

- > URLs of fake news story or debunked version thereof
- > Facebook pages and groups that share either the fake news story or its correction
- > number of followers per page or group

input data to

RAWGraphs

visualise



ARE DEBUNKING WEB PAGES ACKNOWLEDGED BY THE PUBLICS OF FAKE NEWS?

b. フェイクニュース読者と、暴露記事の読者の重なりを調べる

フェイスブックの公開ページ/グループで、フェイクニュースと暴露記事を共有しているものは、

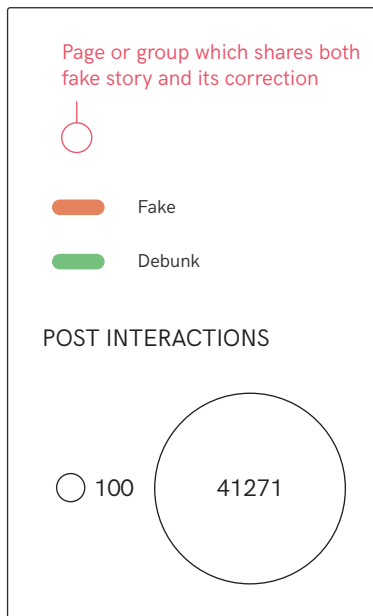
🔧 **CrowdTangle**の拡張機能で検索できる。

◇ 暴露記事サイトが、フェイクニュース記事をシェアしている読者に認識されているか調べるには、一般公開されているフェイスブックページ/グループで、フェイクニュースと暴露記事の両方を共有しているものを特定する。

◇ 🔧 **RawGraphs** を使ってサークルパッキングで可視化できる。

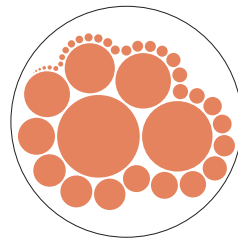
フェイクニュース読者に 暴露情報は認識されて いるか？

フェイクニュース記事と暴露情報の読者はフェイスブック上では異なる。CrowdTangleで得たデータによると、フェイスブック公開ページのうち、フェイクニュース記事をシェアし、その暴露情報を認識しているものは六つしかなかった。グーグルは暴露記事を優先しているものの、フェイスブック上で拡散しやすいのはフェイクニュースだ。保守派も革新派もフェイクニュースをシェアしているが、暴露記事を共有しているのは、革新派のフェイスブックページやジャーナリスト、ファクトチェック関連のページである。



Rage Against the Machine To Reunite And Release Anti Donald Trump Album

Published the 11th of March 2016 on heaviermetal.net



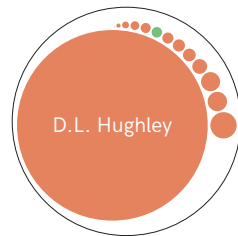
Actor Bill Murray Announces 2016 Presidential Run

Published the 3th of October 2016 on www.abcnews.com.co



Rupaul claims Trump touched him inappropriately in the 1990s

Published the 14th of October 2016 on worldnewsdailyreport.com



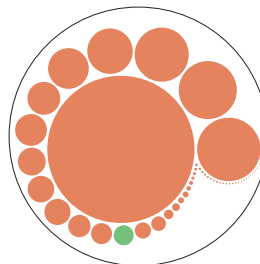
African Billionaire Will Give \$1 Million To Anyone Who Wants To Leave America if Donald Trump is Elected President

Published the 3th of November 2016 on www.empireherald.com



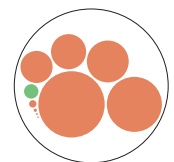
Hillary Clinton In 2013: "I Would Like To See People Like Donald Trump Run For Office; They're Honest And Can't Be Bought"

Published the 17th of October 2016 on conservativestate.com



Graham Says Christians Must Support Trump or Face Death Camps

Published the 23th of July 2016 on www.bizstandardnews.com



Obama Signs Executive Order Declaring Investigation Into Election Results; Revote Planned For Dec. 19th

Published the 12th of December 2016 on www.abcnews.com.co



Pope Francis Shocks World, Endorses Hillary Clinton for President, Releases Statement

Published the 25th of July 2016 on www.kypo6.com



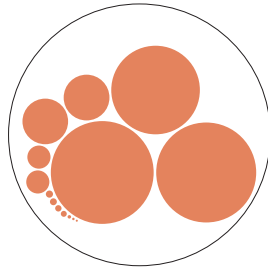
Mike Pence: "Sarah Palin Is My Role Model For Beautiful, Smart American Women"

Published the 18th of September 2016 on www.politicops.com



President Obama Confirms He Will Refuse To Leave Office If Trump Is Elected

Published the 7th of September 2016 on www.burrardstreetjournal.com



Trump Offering Free One-Way Tickets to Africa; Mexico for Those Who Wanna Leave America

Published the 11th of November 2016 on www.tnzhiphop.com



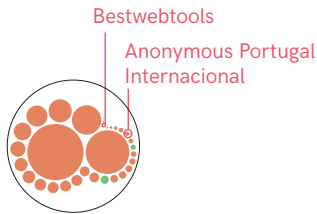
Obama passed law for grandparents to get all their grandchildren every weekend

Published the 8th of March 2016 on www.react365.com



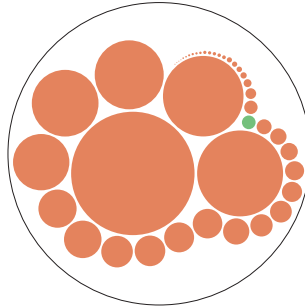
BREAKING Romanian Hacker With Access To Clinton Emails Found Dead In Jail Cell

Published the 6th of July 2016 on Christian Times



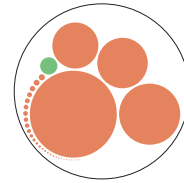
ISIS Leader Calls for American Muslim Voters to Support Hillary Clinton

Published the 11th of October 2016 on worldnewsdailyreport.com



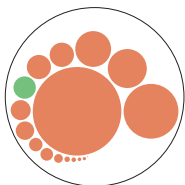
WHOA! Hillary Caught On Hot Mic Trashing Beyonce' With RACIAL SLURS!

Published the 5th of November 2016 on www.thelastlineofdefense.org



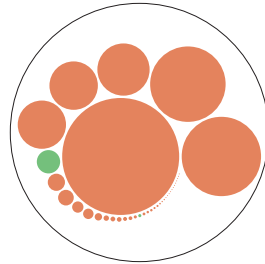
Trump Claims America Should Never Have Given Canada Its Independence

Published the 29th of July 2016 on www.burrardstreetjournal.com



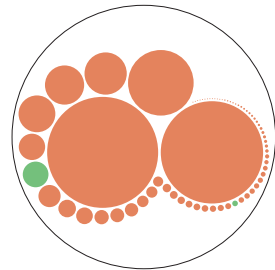
Van Full Of Illegals Shows Up To Vote Clinton At SIX Polling Places, Still Think Voter Fraud Is A Myth?

Published the 5th of November 2016 on www.thelastlineofdefense.org



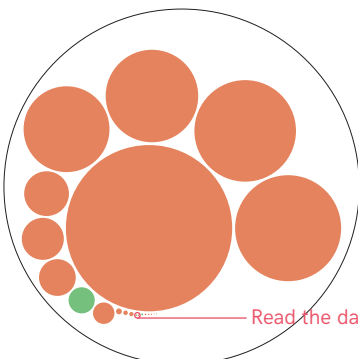
FBI Agent Suspected in Hillary Email Leaks Found Dead in Apparent Murder-Suicide

Published the 15th of October 2016 on www.usanewsflash.com



Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement

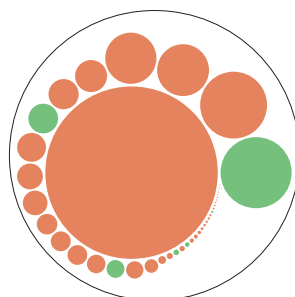
Published the 26th of September 2016 on Ending The Fed



Read the damn news

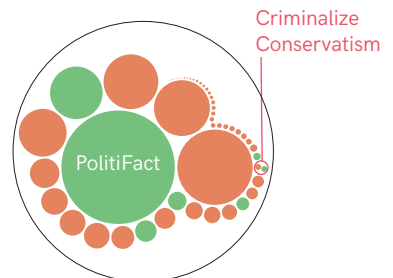
Obama Signs Executive Order Banning The Pledge Of Allegiance In Schools Nationwide

Published the 11th of December 2016 on www.abcnews.com.co

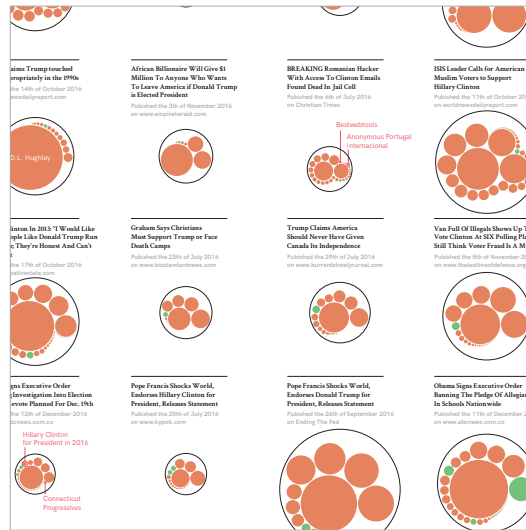


Donald Trump Protester Speaks Out: I Was Paid \$3,500 To Protest Trump's Rally

Published the 22th of November 2016 on www.abcnews.com.co



CHAPTER 1 → RECIPE 3



おすすめ手法

フェイクニュースに大きな関心を示しているフェイスブックページやシェアしている読者に暴露記事が認識されているか調べることによって、その効果が測定できる。

第2章

ウェブ上の フェイクニュース拡散を 追跡する

フェイクニュースはどこから生まれ、どのサイトで最初に再発信されるのか？

フェイクニュースに関わる最も明確なソースは何か？いつ、誰がソースに言及しているか？

<第2章のポイント>

フェイクニュースがどう拡散し、互いにどうつながっているかを調べる。

はじめに - フェイクニュースは単なる「虚偽のニュース」ではない。これらが興味深いのは、「本物のニュース」と内容やスタイルが異なるという点ではなく、主流のニュースと同様に(時にはそれ以上に)拡散されるという点だ。もし、あるブログが「ローマ法王がドナルド・トランプを支持している」と主張したとしよう。その時点では、それはただの嘘に過ぎない。しかし、それがほかの何十ものブログで言及され、何百ものウェブサイトに転載され、何千ものソーシャルメディアアカウントで紹介され、何十万という人が読むことで、フェイクニュースとなる。

ここでフェイクニュースの拡散について調べるのは、次のふたつの理由による。まず、政治的な視点からの必要性がある。多くの人々が、フェイクニュース対策としてよく用いられるテクニックや戦術に対し、期待に応えるものでないと不満を表明している。特に、ファクトチェックや、虚偽であることを指摘する試みは、作り話や噂の拡散を防げないでいる。反対にウェブでフェイクニュースをより目立たせてしまいかねない。オンライン上でフェイクニュースがどのように広がるかについて理解を深めることで、より効果的な対応ができる。次に、方法論的視点からの理由がある。フェイクニュースと「本物のニュー

ス」の間に「存在論的」な違いはない。よって、フェイクニュースの拡散について研究することで、他の種類のニュースがどのように広がるかについても知見を得られる。

このレシピには2種類ある。まずは、特にツールや技術的知識を必要としない、完全手動型のレシピを提案する。実行はしやすいが、時間を要する。フェイクニュースに言及したウェブページの検索、および公開日時と引用ソースの特定を手動で行う。次に、半自動型のレシピだ。完全手動型よりも大規模の調査ができるが、技術的スキルが要されるほか、手動での精査が必要となる場面もある。



フェイクニュースはどこから生まれ、 どのサイトで最初に 再発信されるのか？

始める前に

拡散を追跡したいフェイクニュースを選ぶ。ニュースが特徴的であるほど、拡散を追跡しやすくなる。例では、米大統領選時に広く拡散された「ローマ法王がトランプを支持」というフェイクニュースを取り上げる。

START

open Chrome browser
in *incognito* mode



query

[Pope AND endors* AND (Trump OR Clinton)
AND NOT (hoax OR "fake news" OR lie OR debunk)]

Google Web Search

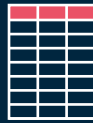


compile a list with
the top **resulting** URLs



a

input **URLs** in a spreadsheet



- > The page **advocating** or **debunking** the fake news item
- > **Rank** in the search engine results
- > the **date of publication** of each page

record **all sources** cited in the occurrences of your story, noting down:

- +
- +
- +

- > if they are cited as **evidence** or **counter-evidence**
- > if they are cited through **hyperlink, textual reference** or **copying/pasting** of its content



Make sure you visit both the **results** of your initial search and **all the sources** cited by those results



extract **network** of instances and references with

Table2Net



import **data** in

Gephi

visualise

b

HOW DO THE OCCURRENCES OF THE "POPE ENDORSES TRUMP" STORY CITE EACH OTHER?



Plot **URLs** on a **timeline** with

Graph Recipes

visualise

c

WHAT IS THE LIFE OF THE "POPE ENDORSES TRUMP" STORY ACCORDING TO PAGES IN SEARCH ENGINE RESULTS?

START

open Chrome browser
in **incognito** mode



query

[Pope AND endors* AND (Trump OR Clinton)
AND NOT (hoax OR "fake news" OR lie OR debunk)]

Google Web Search

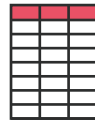


compile a list with
the top **resulting** URLs

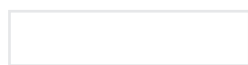
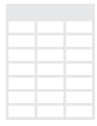


a

input URLs in a spreadsheet



- > The page **advocating** or **debunking** the fake news item
- > **Rank** in the search engine results
- > the **date of publication** of each page



a. 検索エンジンでフェイクニュースの掲載を特定、分析する

拡散しているフェイクニュースは常に、複数のウェブサイトに掲載される。このレシピの最初のステップは、その情報を特定し、収集することだ。

- ◇ 一つあるいは複数の検索エンジンを使い(例では **🔍 グーグルウェブ検索** を使用)、追跡したいフェイクニュースの掲載を特定する。拡散の過程でフェイクニュースは変化するので、さまざまなバリエーションに対応できるキーワードを考える必要がある。
- ◇ 関連性のランク付けについては検索エンジンを信頼し、最初に表示される検索結果に集中する(それらが最も拡散されているニュースであるという仮説の下で)。
- ◇ 「フィルターバブル」や個人の履歴を基にした検索結果を避けるため、調査専用のブラウザ使用を検討してほしい。^[1]
- ◇ 検索エンジンは、フェイクニュースよりも暴露記事を優先して表示する可能性がある。そのため、暴露記事の拡散を過大評価しないように気を付けたい。さらに、選択した検索エンジンを通して一連の現象を見ることとなり、それが追跡しているフェイクニュースあるいは調査に影響を及ぼすことを留意しておく必要がある。
- ◇ 関連するそれぞれの検索結果について、全てのメタデータの記録をとる。どれだけ特徴を収集してもいいが、そのページで該当のフェイクニュースがどのように言及されているかを必ず記録しておく(例えば、「信頼できるニュースソースとして」あるいは「嘘であると暴かれるべき、問題ある主張として」など)。「認知度」についても何らかの指標を記す(例えば、検索結果におけるランキング)。

[1] See instructions on how to set up a research browser in this video tutorial: <https://www.youtube.com/watch?v=bj65X-r9GkJM>

START

open Chrome browser
in **incognito** mode



query

[Pope AND endors* AND (Trump OR Clinton)
AND NOT (hoax OR "fake news" OR lie OR debunk)]

Google Web Search

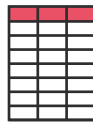


compile a list with
the top **resulting** URLs



a

input **URLs** in a spreadsheet



- > The page **advocating** or **debunking** the fake news item
- > **Rank** in the search engine results
- > the **date of publication** of each page

record **all sources** cited in the occurrences of your story, noting down:



- > if they are cited as **evidence** or **counter-evidence**
- > if they are cited through **hyperlink, textual reference** or **copying/pasting** of its content



Make sure you visit both the **results** of your initial search and **all the sources** cited by those results



extract **network** of instances and references with

Table2Net



import **data** in

Gephi

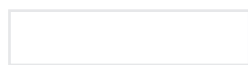


visualise



b

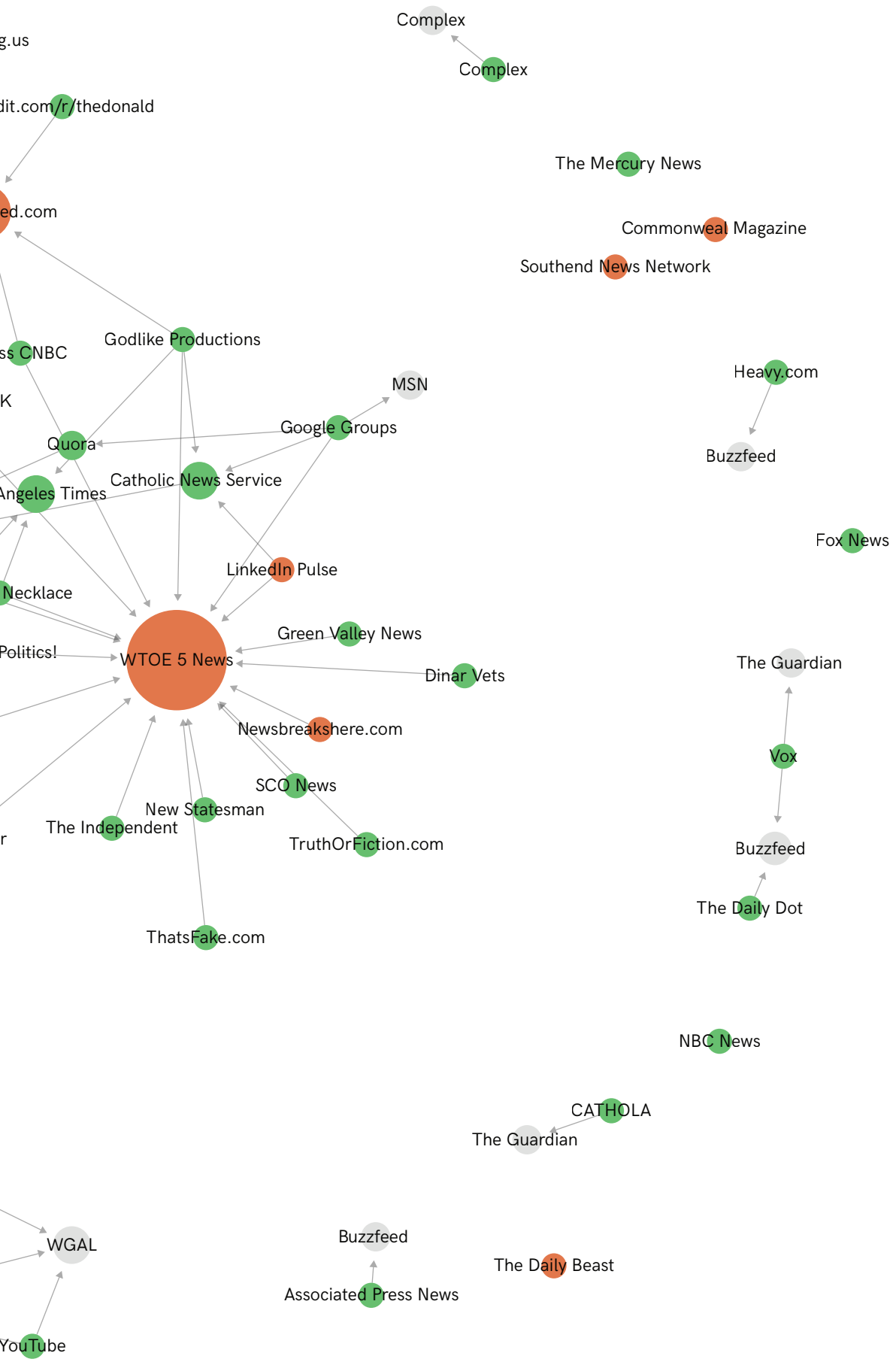
HOW DO THE OCCURRENCES OF THE "POPE ENDORSES TRUMP" STORY CITE EACH OTHER?



b. フェイクニュースに言及するネットワークを抽出する

フェイクニュースはそれに言及するネットワークに支持、あるいは反論されている。例えば、噂をまとめたサイトは他ページを引用して自分たちの主張を裏付ける。ファクトチェックサイトでは、有害なウェブサイトには警告マークを付けたり、有害サイトを否定しているソースに言及したりする。このステップでは、フェイクニュースに言及するネットワークを追跡する。

- ◇ 掲載されているフェイクニュースで、引用されている全てのソースを記録する。それぞれについて、ソースが証拠(エビデンス)あるいは反証のどちらかで引用されているかを記す。また、引用がハイパーリンク(「<http://snopes.com>」等、アドレス表記によるリンク)によるものか、テキストによる言及か(「WTOE 5 Newsのウェブサイト」等の記述)、またはコンテンツのコピーペーストによるものかを記録しておく。
- ◇ 最初の検索で出てきた全サイトだけでなく、これらのサイトで引用されている全てのソースを必ず訪問する。
- ◇ フェイクニュースの掲載と言及のネットワークを抽出する(🔧 **Table2Net**を利用できる)。
- ◇ 力学モデルグラフで、ネットワークをビジュアル化する(例えば 🔧 **Gephi** を使用する)。引用される数が多いほどノードを大きくし、フェイクニュースをどう報じているか(支持しているか、嘘だと暴れているか)によってノードの色を変える。



START

open Chrome browser
in **incognito** mode



query

[Pope AND endors* AND (Trump OR Clinton)
AND NOT (hoax OR "fake news" OR lie OR debunk)]

Google Web Search



compile a list with
the top **resulting** URLs



a

input **URLs** in a spreadsheet



- > The page **advocating** or **debunking** the fake news item
- > **Rank** in the search engine results
- > the **date of publication** of each page

record **all sources** cited in the occurrences of your story, noting down:

- + 
- + 
- + 

- > if they are cited as **evidence** or **counter-evidence**
- > if they are cited through **hyperlink**, **textual reference** or **copying/pasting** of its content



Make sure you visit both the **results** of your initial search and **all the sources** cited by those results



extract **network** of instances and references with

Table2Net

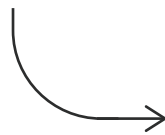
import **data** in

Gephi



Plot **URLs** on a **timeline** with

Graph Recipes



visualise



c

WHAT IS THE LIFE OF THE "POPE ENDORSES TRUMP" STORY ACCORDING TO PAGES IN SEARCH ENGINE RESULTS?

c. フェイクニュースの拡散を時系列でビジュアル化する

どのサイトが何を引用したかだけでなく、フェイクニュースがどの方向でどのように拡散したかを把握するため、前述のステップで抽出したネットワークは役立つ。拡散の様子を明らかにするには、このレシピの最初のステップで収集した日付を用いる。

- ◇ 前述のステップで抽出したネットワークを時系列で並べる。引用方法の違いを可視化するため、引用を示す線のデザインを変えるとよい（例示では、🔧 **Graph Recipes** のカスタムスクリプトを使用した）。

検索エンジンでの「ローマ法王がトランプを支持」という記事の寿命は？

「ローマ法王がトランプを支持」というフェイクニュースに言及したページで、相互引用のネットワークを時系列に表示した。ノードの色はページがニュースを肯定するか、あるいは嘘だと暴いているかによって変わり、ラインの種類はページがそれぞれをどう引用しているかを示している。緑からオレンジ及びグレーのノードに引かれる点線が多い。これは、ファクトチェックサイトはオリジナルソースに言及するが、リンクは張らない傾向にあることを示している。リンクを明示すると、オリジナルソースのオンライン上の認知度が上がってしまう。これを避けつつフェイクサイトに警告を出すため、こうした手法が採用されている。

- source in original list
- source identified through the analysis

CITED SOURCE SOURCE

TYPE OF PAGE

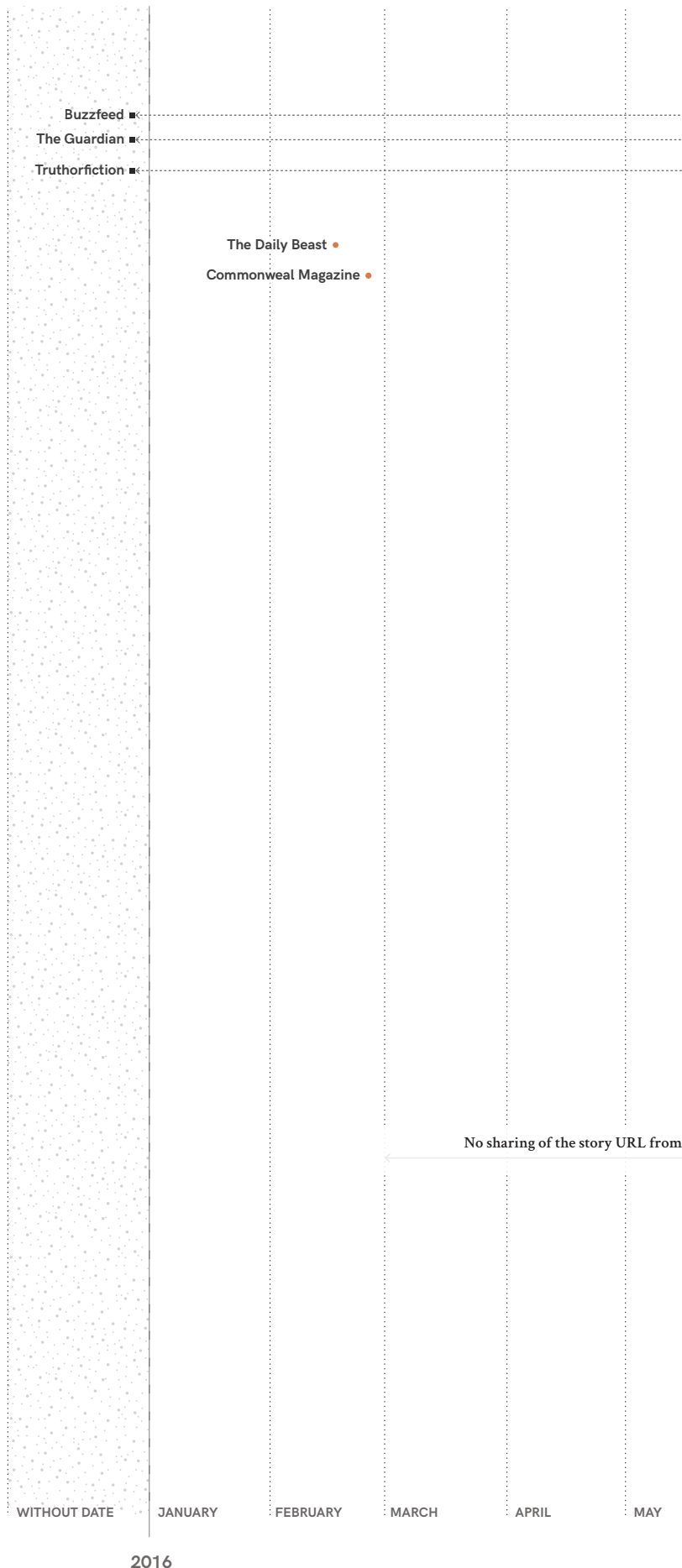
- Fake News
- Debunk

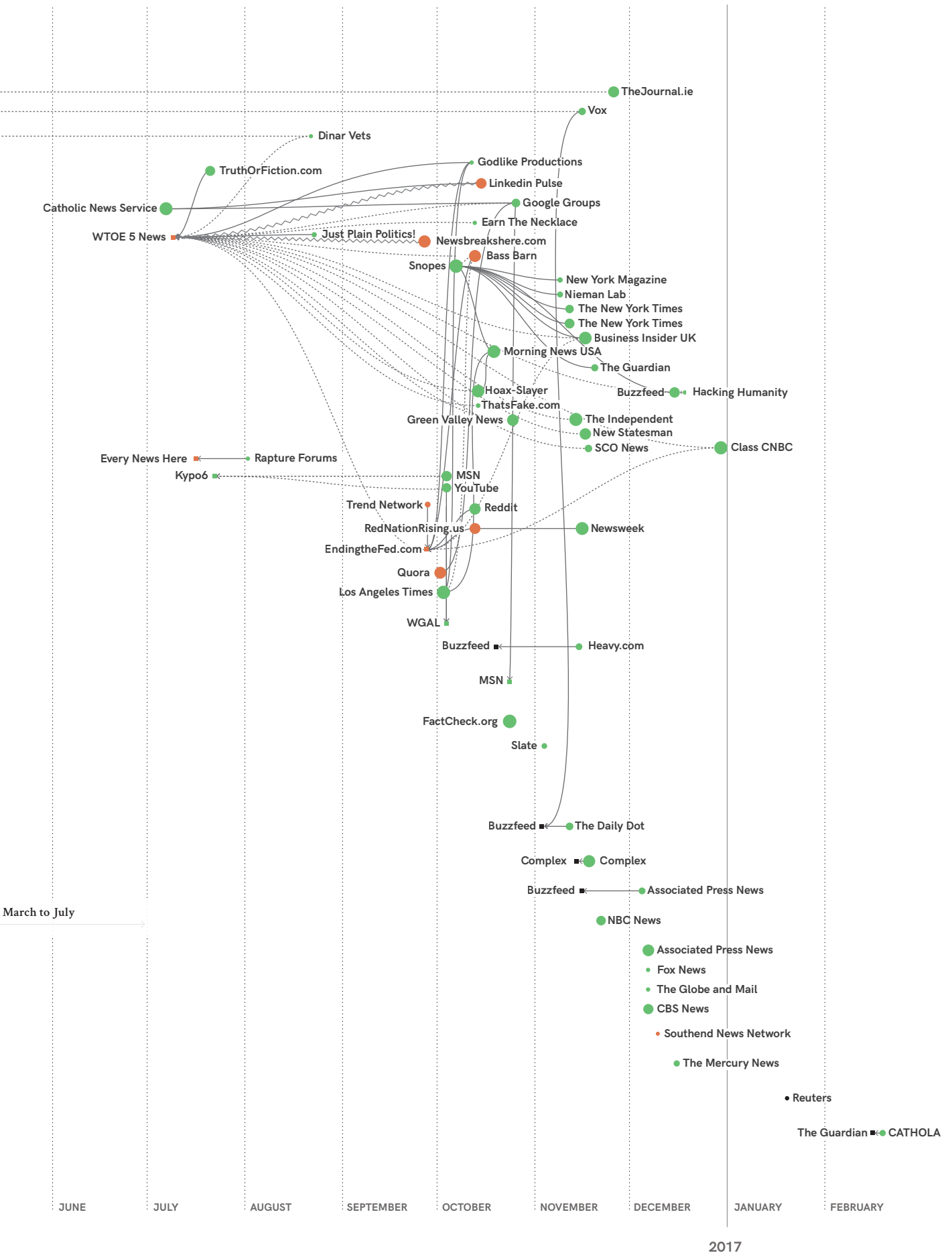
TYPE OF CITATION

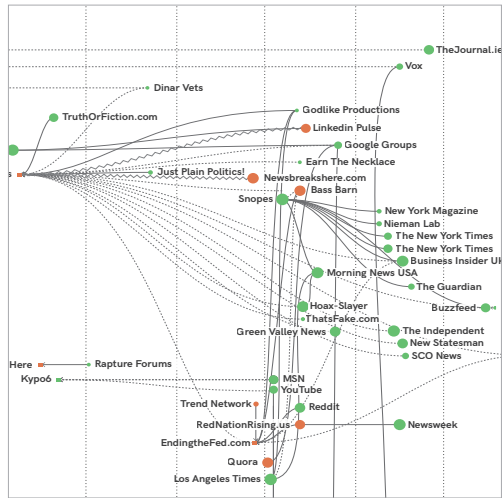
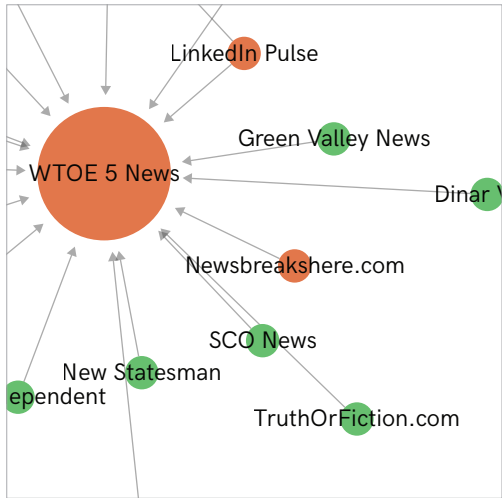
- source affirming the story cited as hyperlink
- source affirming the story cited as text
- source affirming the story cited as copy-paste

GOOGLE RANK

○ n. 1 ● n. 100







おすすめ手法



フェイクニュースとはさまざまなオンラインソースによって引用され、言及されているものである。このレシピを用いると、検索エンジンによるウェブ検索で得たデータから、あるフェイクニュースがどのように掲載されているか特定、追跡できる。時系列での拡散を遡って調べることも可能だ。嘘だと暴く活動がいつ起きて、誰がそれを促進し、結果としてフェイクニュースやウェブページの拡散にどのような影響を与えたかについても把握できるだろう。



フェイクニュースに関わる 最も明確なソースは何か？ いつ、誰がソースに言及しているか？

始める前に

このレシピは、レシピ1によるアプローチのスケールアップを可能にするものだが、技術的な知識や、より大きなデータセットも必須となる。特に次へのアクセスが必要となる。

- ◇ ウェブアーカイブ(ここでは  Radarly を使用)
- ◇ 追跡したいフェイクニュースが掲載されている可能性がある全てのウェブソースのリスト(ここでは  Le Monde Décodexによって選定されたリストを使用)

このレシピの解説のため、2017年のフランス大統領選で拡散されたエマニュエル・マクロン氏が同性愛者であるとするフェイクニュースを例として取り上げる。

START

define a **query** identifying your fake news story

identify the **occurrences** of your story in

Radarly

export all data

- > Full text of each occurrence
- > Date of publication
- > Metadata

compile a list of potential sources

websites collected and curated by the project **Décodex** by Le Monde

crawl the websites with

Hyphe

export a **graph file**

it represents the **general connectivity** among potential sources

search mentions of **potential sources** in the instances of your story

Csv Rinse Repeat

export a **graph file**

create a **basemap** network with

Gephi

visualise



a

WHAT ARE THE MAIN SPHERES IN THE FRENCH MEDIA SYSTEM?

Project the **instances** of your story on the basemap

visualise



b

WHICH ARE THE SOURCES CITED IN THE OCCURRENCES OF THE FAKE NEWS STORY?

produce **time-sliced networks** using

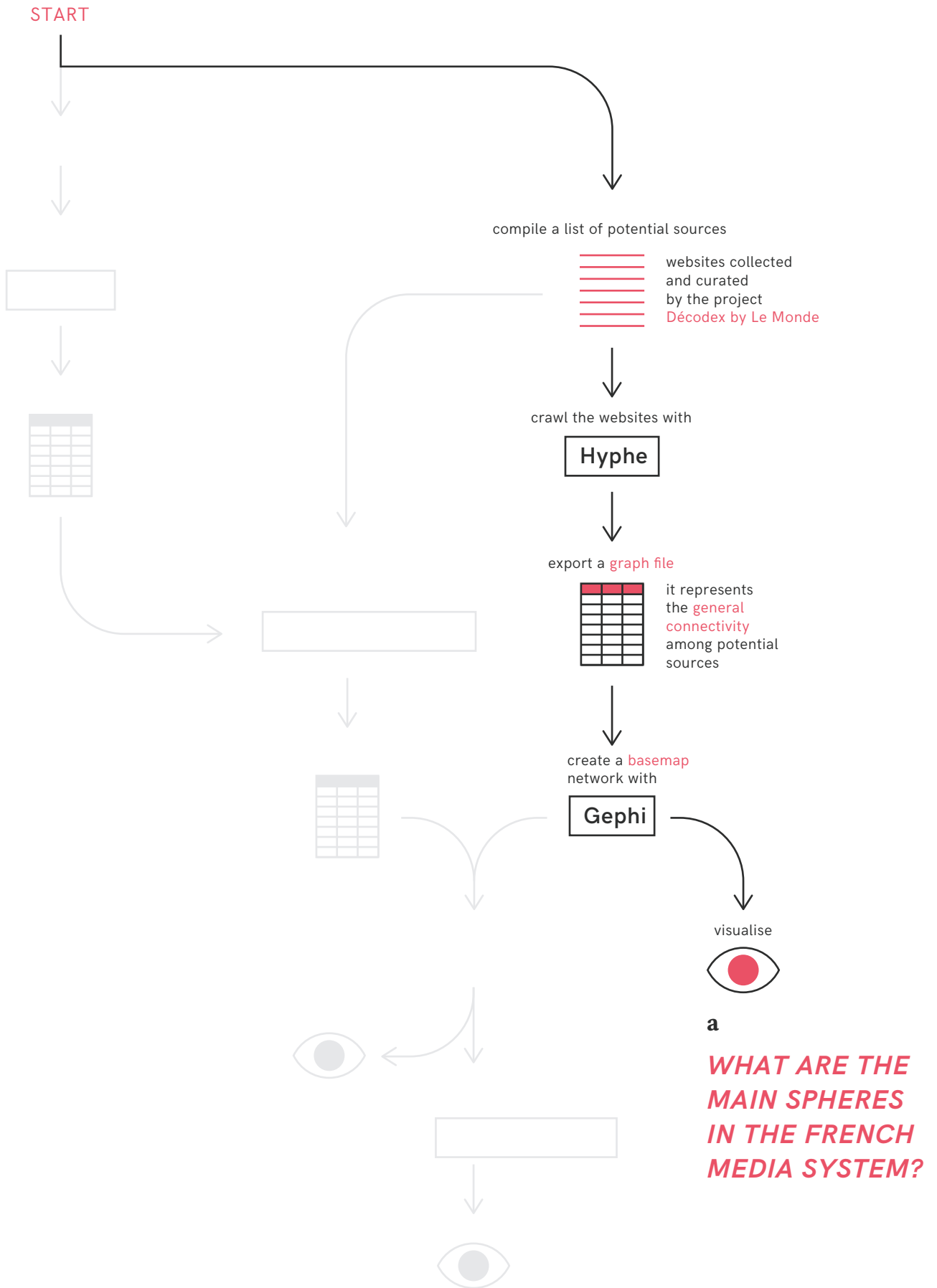
Graph Recipes

visualise



c

HOW MANY OCCURRENCES OF THE FAKE NEWS ARE PUBLISHED IN EACH PERIOD OF TIME AND WHAT SOURCES DO THEY CITE?



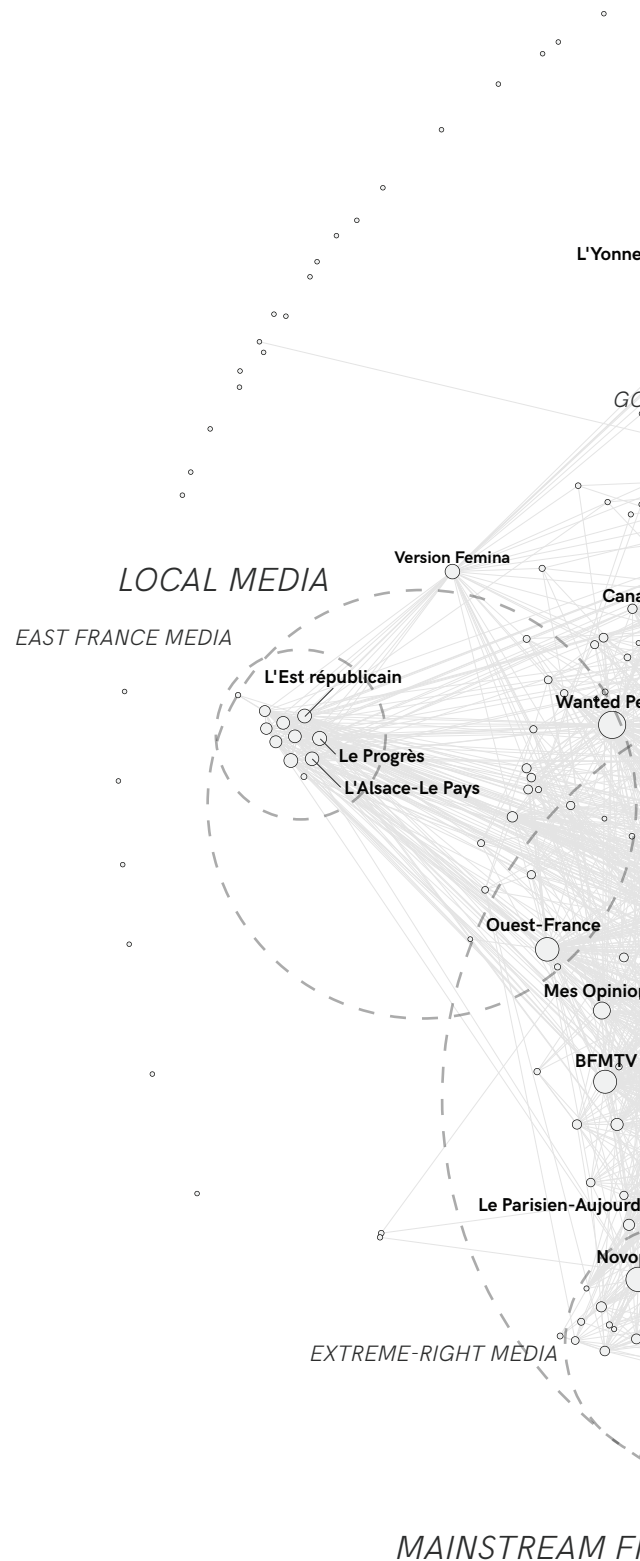
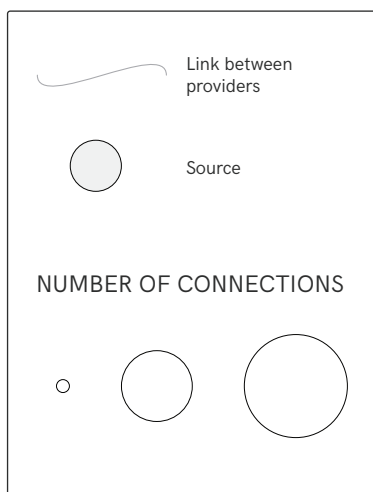
a. メディアのベースマップを作成する

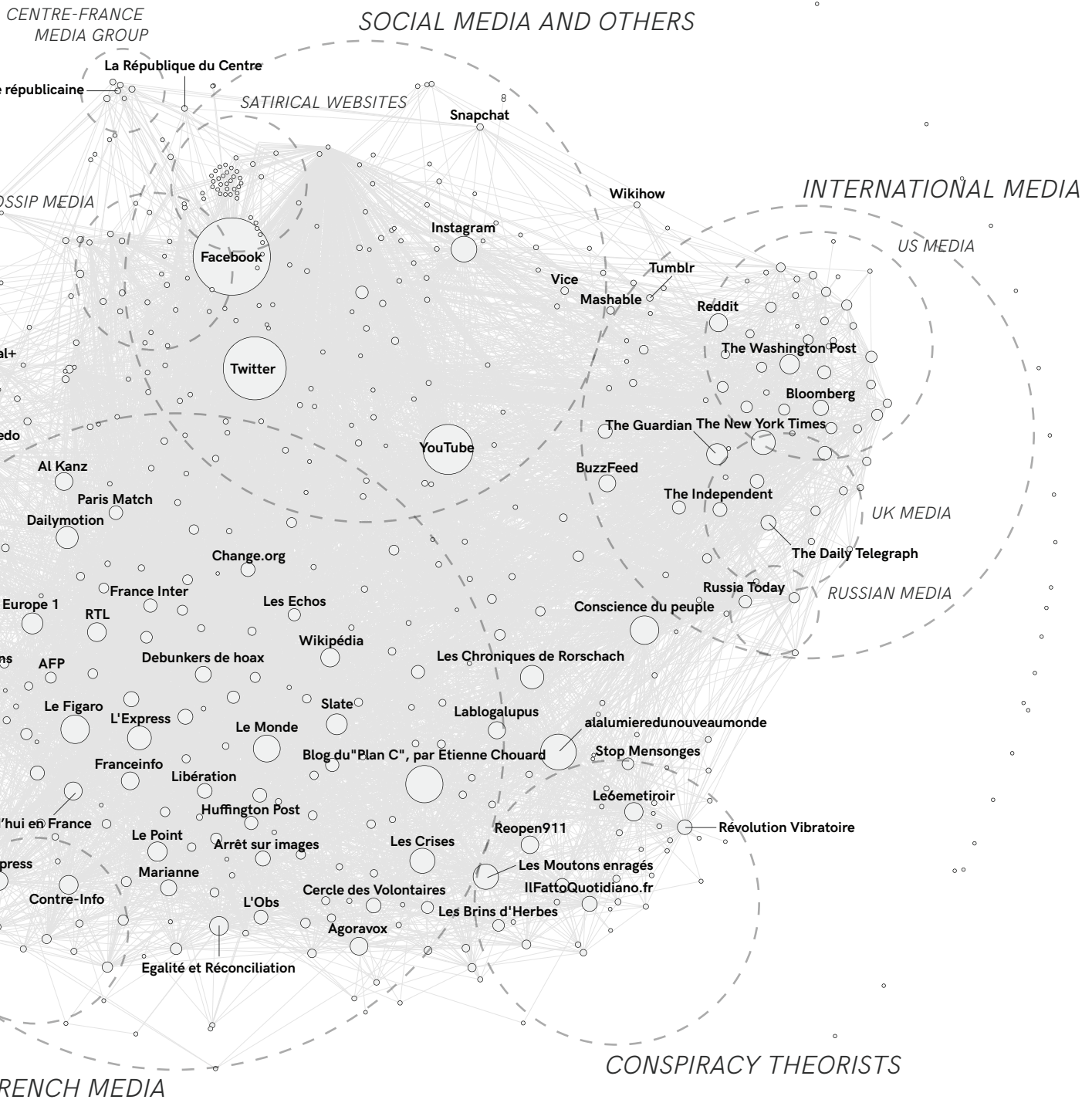
追跡したいフェイクニュースが掲載されていると思われる全てのウェブソースのリストを特定、あるいは作成する(可能な限り網羅する)。ファクトチェック団体による、フェイクニュースサイトのリストを使用してもよいし、それを主要マスメディアのリストと組み合わせてもよい。

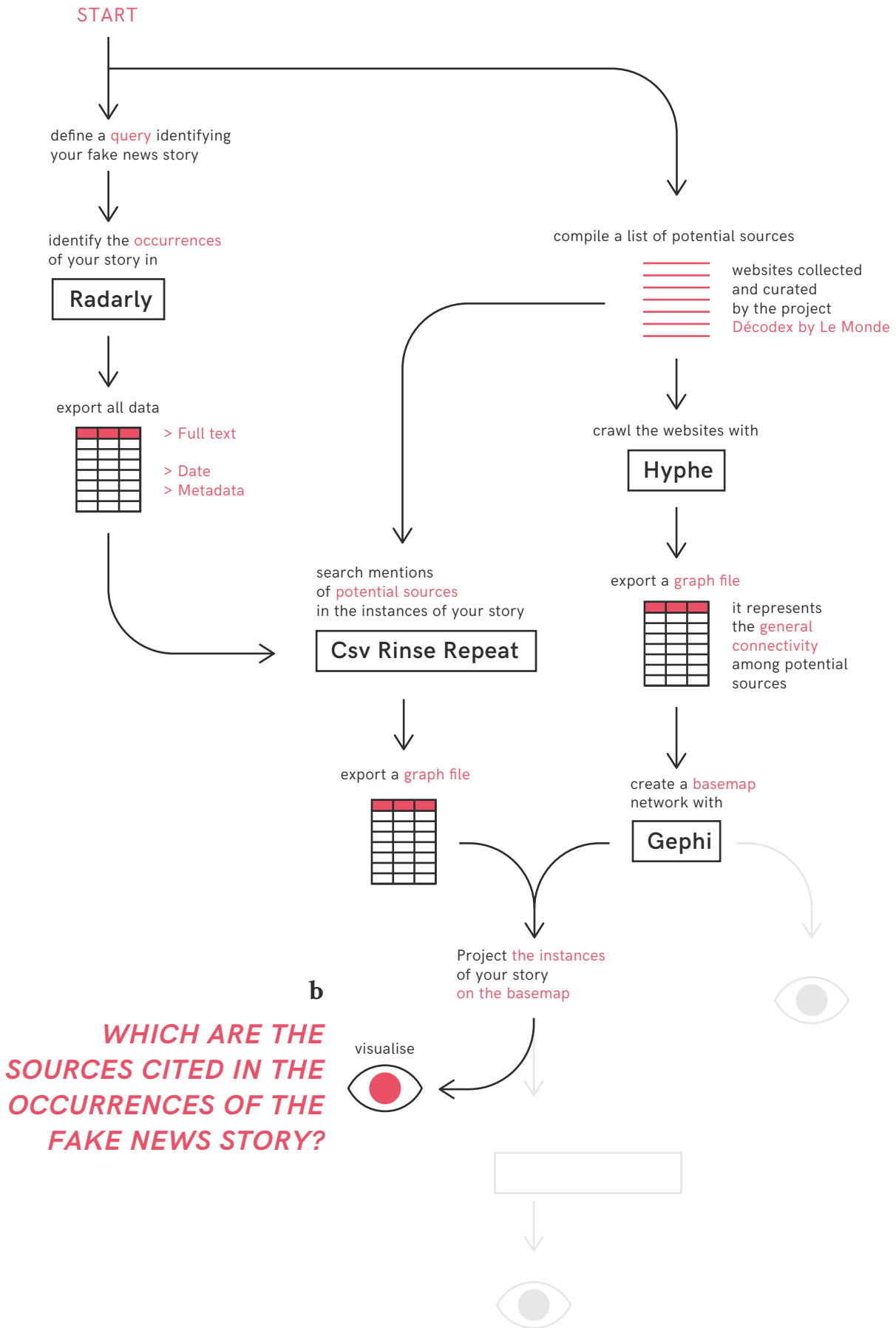
- ◇ リストのソースがいかに相互に関連しているかをウェブ巡回、リンク分析により特定する。ここでは、🔧 **Hyphe** を使用した。
- ◇ 出現したネットワークをビジュアル化し、力学モデルグラフのアルゴリズムを適用してソースをクラスター化する。この手順には🔧 **Gephi** を使える。
- ◇ 手動でクラスターをハイライトし、名前を付ける。

フランスの 言論空間における 主要集団とは？

フランスの言論空間で発信しているメディアのネットワーク分析を行う。図はル・モンドのDécodexプロジェクトによってリスト化されたニュースソース及びそれらを相互に関連付けるリンクを示す。ウェブサイトの主要クラスターとそれぞれに関連と位置づけを明らかにするために力学モデルグラフを適用している。









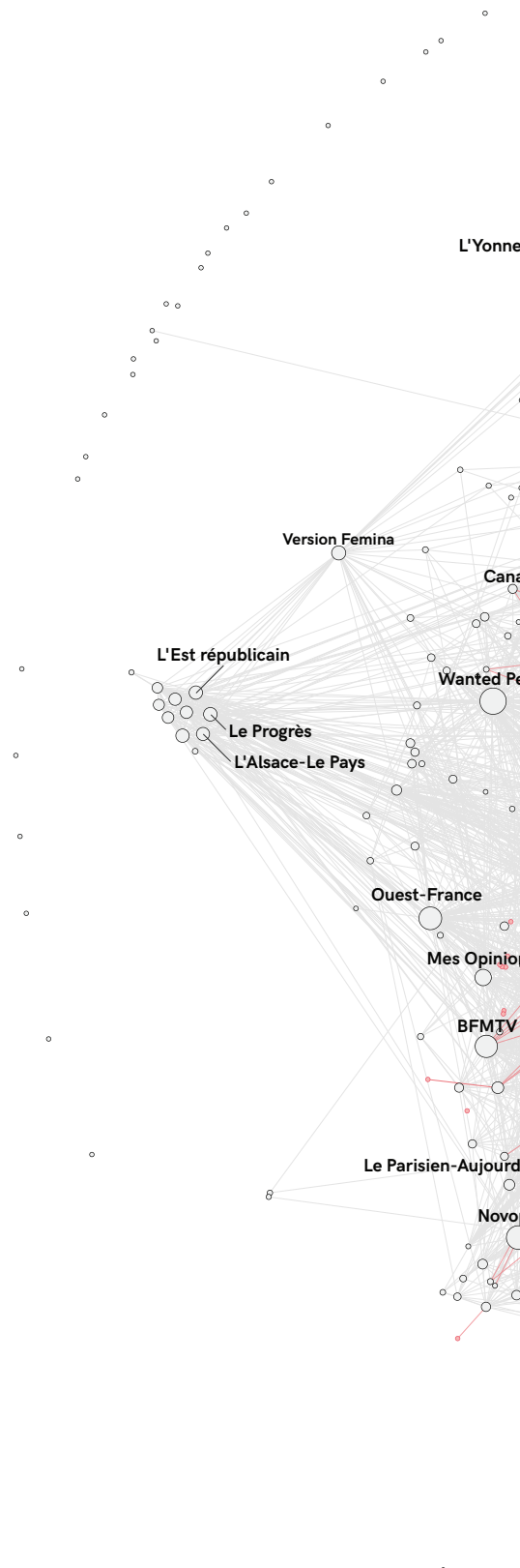
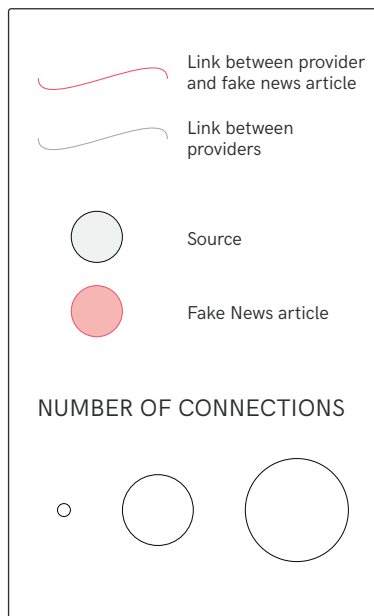
b. ベースマップ上でフェイクニュース掲載をハイライトする

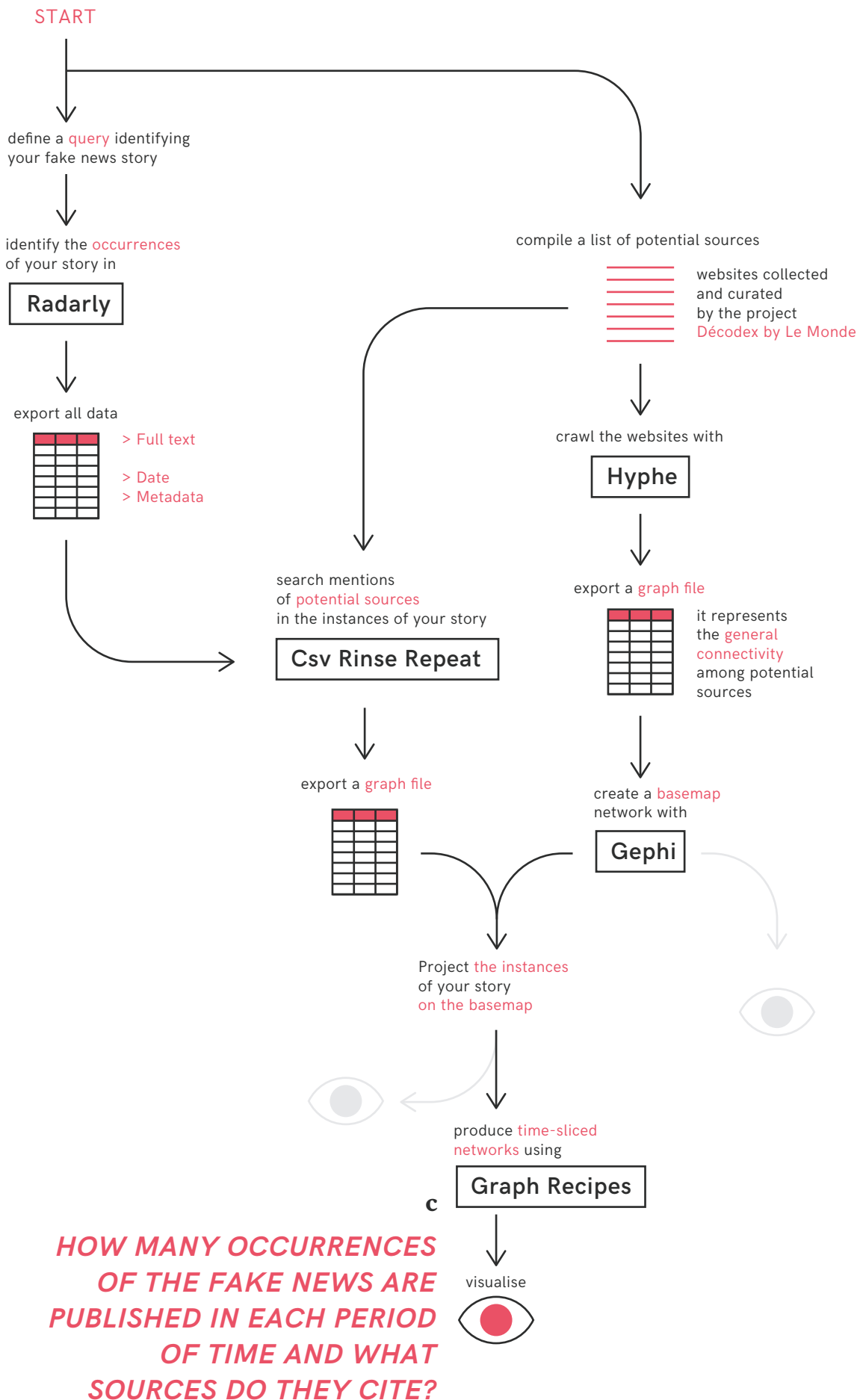
このステップでは、フェイクニュースがさまざまなソースとどのように関連付けられているかを調べる。この調査は興味深い。というのは、例えば最初は単なる風刺として広がり始めたフェイクニュースが、拡散されるにつれて風刺ではないソースとの関連が深くなっていくということがあるからだ。ここでは、フランスメディアのベースマップにあるソースのうちどれが、フェイクニュース掲載時に言及されたかを特定する。

- ◇ 追跡したいフェイクニュースを特定するクエリを作成する。ニュースと明確に関連付けられるキーワード、及び「誤判定」を排除するためのストップワードを使用する。
- ◇ 選択したアーカイブでクエリを走らせ、該当のフェイクニュースの掲載を特定する。それぞれの検索結果について、全文と公開日を記録する。
- ◇ 該当のフェイクニュースが掲載されていれば、そこでベースマップにあるソースへの言及を調べる。URLやソース名を検索する(例えば、sputniknews.comあるいはSputnik)。例では  **CSV Rinse Repeat.** のカスタムスクリプトを用いた。
- ◇ フェイクニュースの掲載状況をベースマップに重ね、それぞれを言及したソースと関連付ける。ソースのノードは固定し、力学モデルの立体アルゴリズムを適用して( **Gephi**を用いる)、フェイクニュース掲載を表すノードを、それらが最も引用したベースマップのクラスターに近づける。

フェイクニュース掲載時に 最も引用された ソースは何か？


フェイクニュース掲載状況をメディアソースのネットワークに重ね合わせる。この図では、フェイクニュース記事は、引用したソースに基づき、ベースマップ上に位置付けられている。このニュースについては、ソーシャルメディアやRussia Today、Sputnik International への明らかな言及傾向が見られる。





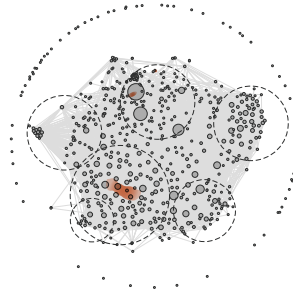
c. ベースマップ上でフェイクニュースの拡散をビジュアル化する

このステップでは、これまでに特定した言及パターンが時間を経るにつれてどう変化するかを明らかにする。

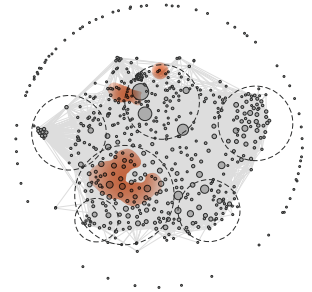
- ◇ 該当するフェイクニュースの拡散スピードに合わせ、月または週、一日単位で、記事と引用のネットワークを表示する。この例では、月ごとに表示した後、拡散が最も顕著な期間については4日ごとに示した。
- ◇ ソースのベースマップを固定し、フェイクニュースの掲載状況が時期によって異なることをビジュアル化する。
- ◇ 変化とパターンをより読みやすくするため、ノードではなく密度を表すヒートマップを用いてフェイクニュースの掲載状況を表現してもよい(例示の図は  **Graph Recipes**)を用いて作成)。

フェイクニュースの 掲載量と、引用された ソースを時期別に見る

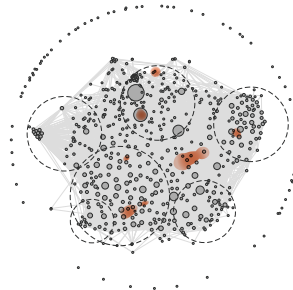
調査対象期間において、フェイクニュースの掲載状況がどのように変化していったかを調べる。この図では、フェイクニュースの出現を4週ごと（2週間は重複）に表示している。ニュースの出現はノードではなくヒートマップで表した。ニュースへの言及は1年以上続いていたが、急速に広がったのは、ロシアのウェブサイトSputnik International が新たなフェイクニュース記事を公開した2017年2月と見られる。



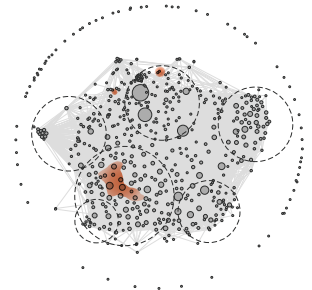
14 MAR 2016 - 11 APR 2016



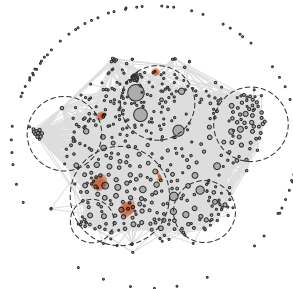
28 MAR 2016 - 25 APR 2016



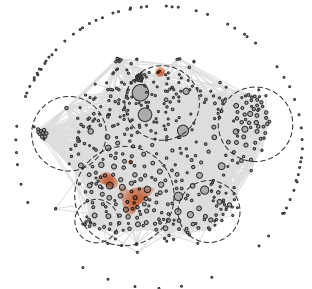
6 JUN 2016 - 4 JUL 2016



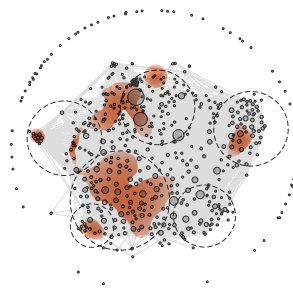
20 JUN 2016 - 18 JUL 2016



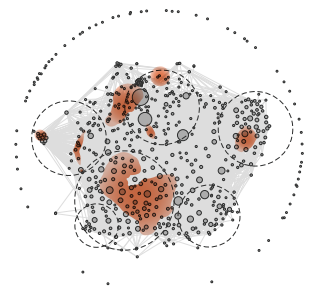
29 AUG 2016 - 26 SEP 2016



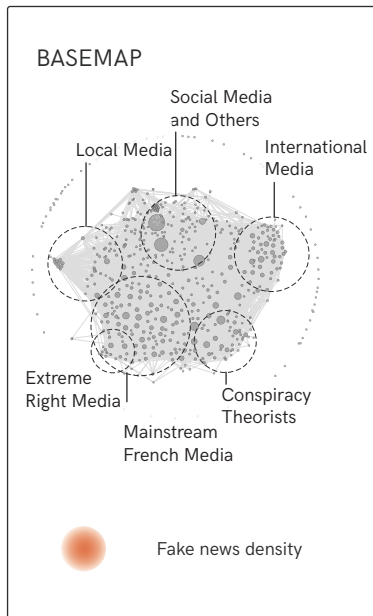
12 SEP 2016 - 10 OCT 2016

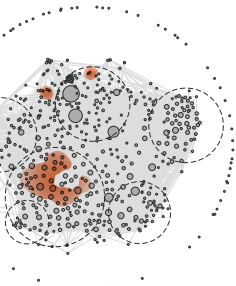


21 NOV 2016 - 19 DEC 2016

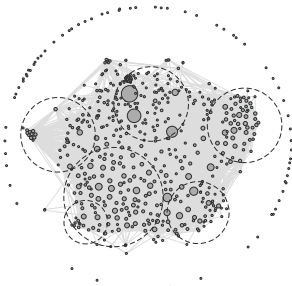


5 DEC 2016 - 2 JAN 2017

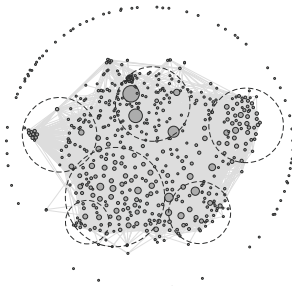




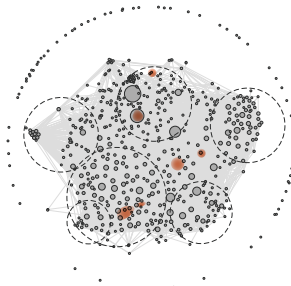
APR 2016 - 9 MAY 2016



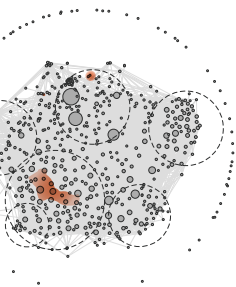
25 APR 2016 - 23 MAY 2016



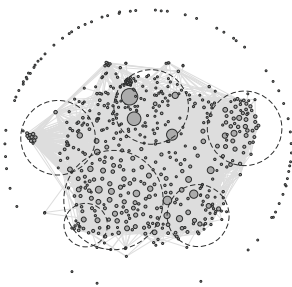
9 MAY 2016 - 6 JUN 2016



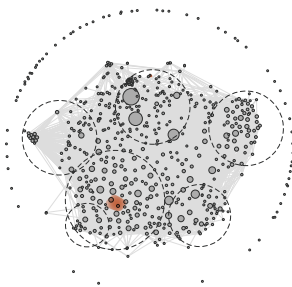
23 MAY 2016 - 20 JUN 2016



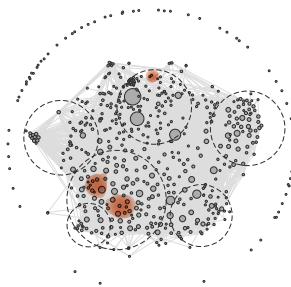
JUL 2016 - 1 AUG 2016



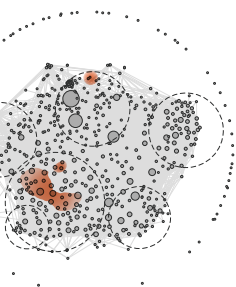
18 JUL 2016 - 15 AUG 2016



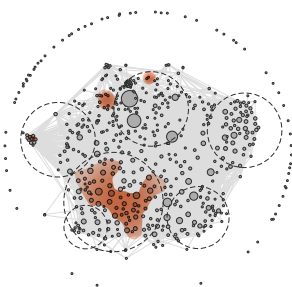
1 AUG 2016 - 29 JUL 2016



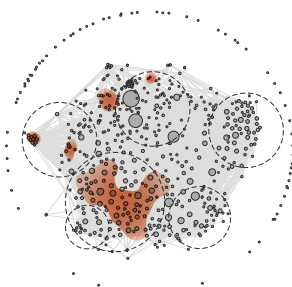
15 AUG 2017 - 12 SEP 2016



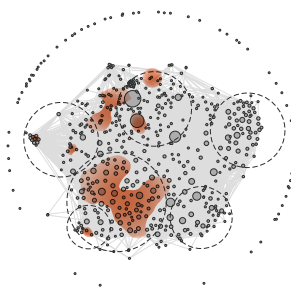
SEP 2016 - 24 OCT 2016



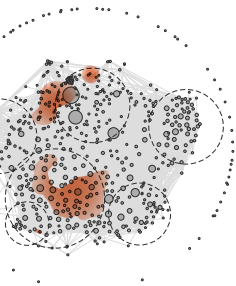
10 OCT 2016 - 7 NOV 2017



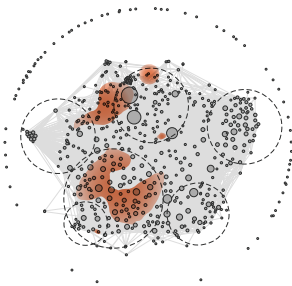
24 OCT 2017 - 21 NOV 2017



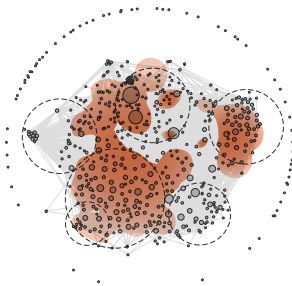
7 NOV 2016 - 5 MAY 2016



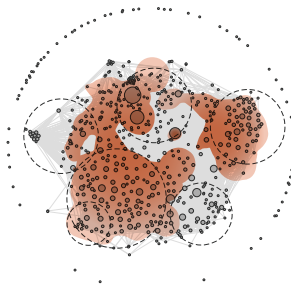
DEC 2016 - 16 JAN 2017



2 JAN 2017 - 30 JAN 2017



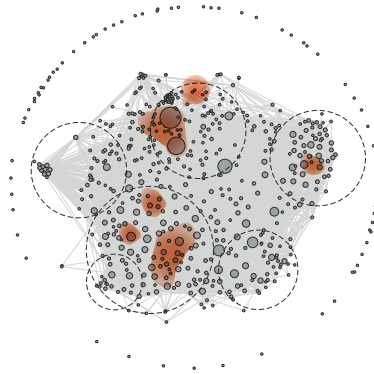
16 JAN 2017 - 13 FEB 2017



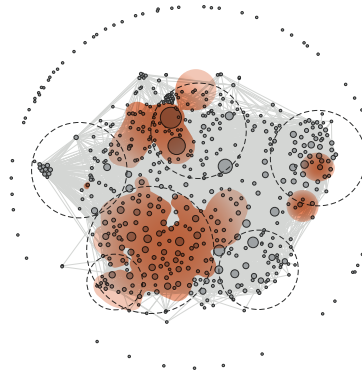
30 JAN 2017 - 27 FEB 2017

2017年2月に掲載された フェイクニュースの数と、 引用元を見る

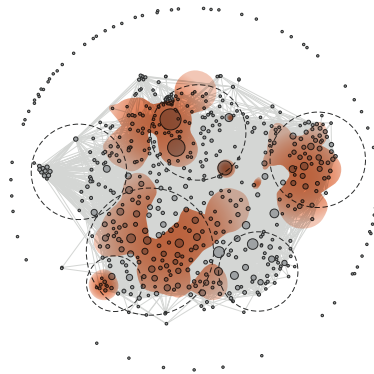
2017年2月のフェイクニュースの拡散の様子である。この図は、前述した図のうち、時期を絞ってズームアップしたものだ。フェイクニュース記事の出現を4日ごとに示している(2日間は重複)。



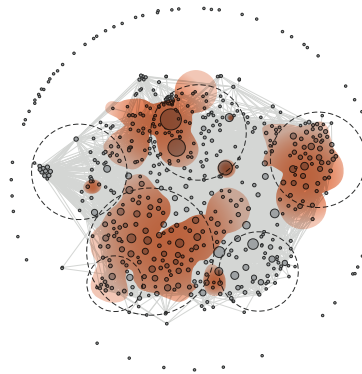
1 FEB 2017 - 5 FEB 2017



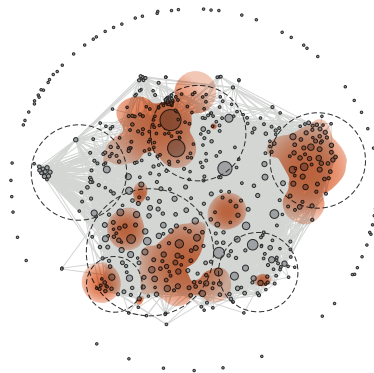
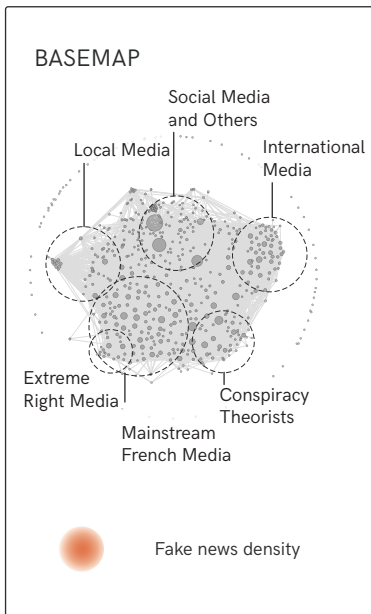
3 FEB 2017 - 7 FEB 2017



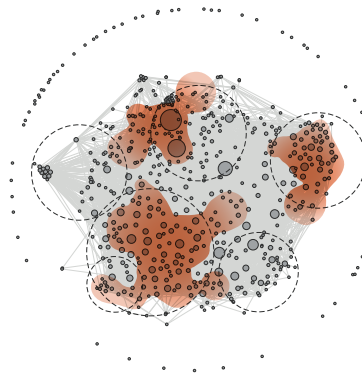
11 FEB 2017 - 15 FEB 2017



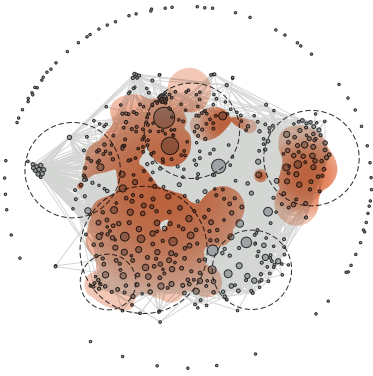
13 FEB 2017 - 17 FEB 2017



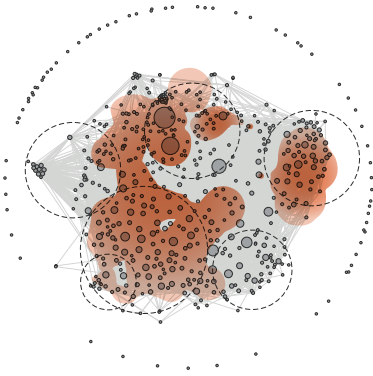
21 FEB 2017 - 25 FEB 2017



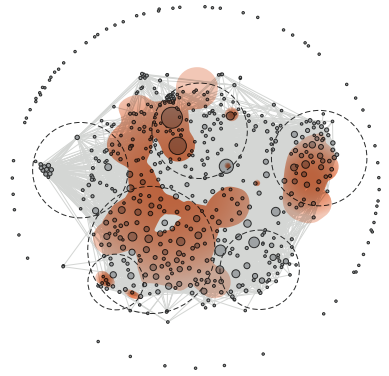
23 FEB 2017 - 27 FEB 2017



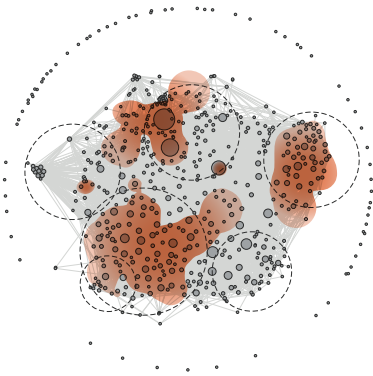
5 FEB 2017 - 9 FEB 2017



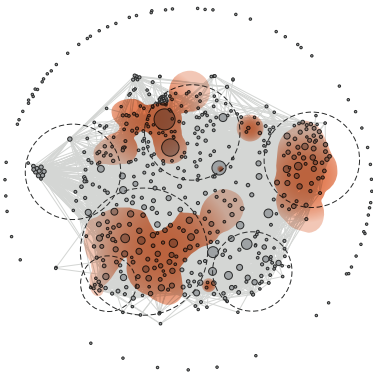
7 FEB 2017 - 11 FEB 2017



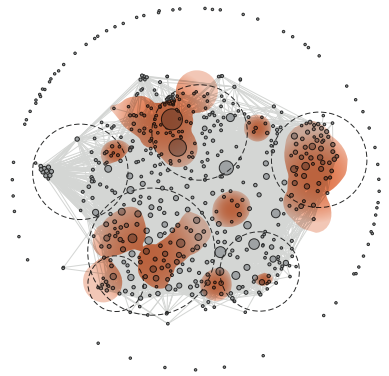
9 FEB 2017 - 13 FEB 2017



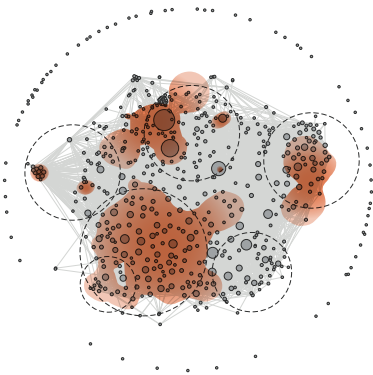
15 FEB 2017 - 19 FEB 2017



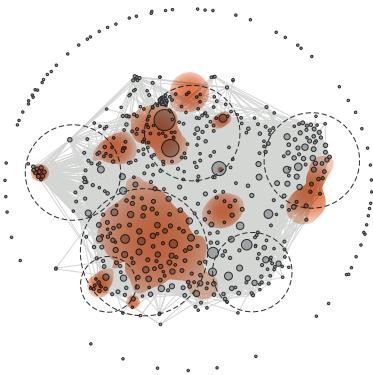
17 FEB 2017 - 21 FEB 2017



19 FEB 2017 - 23 FEB 2017

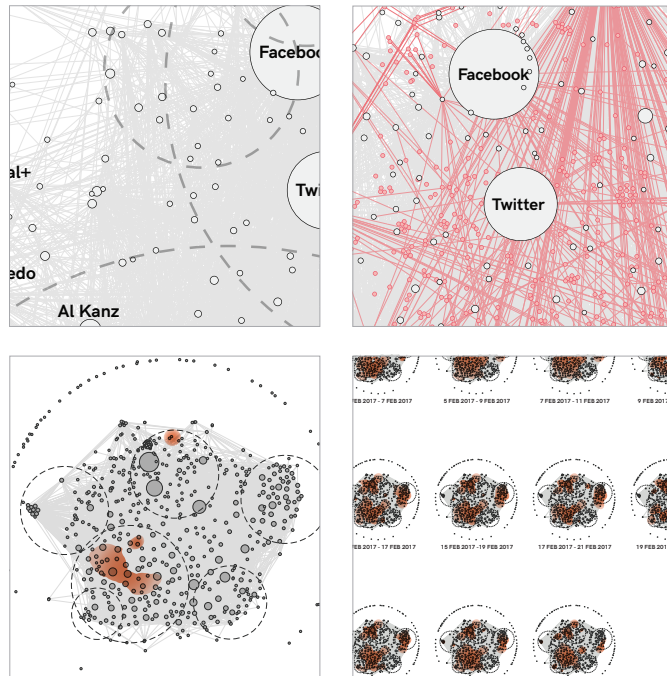


25 FEB 2017 - 1 MAR 2017



27 FEB 2017 - 3 MAR 2017

CHAPTER 2 → RECIPE 2



おすすめ手法

このレシピを使うと、さまざまなメディアグループにおいて、フェイクニュースに頻繁に言及するウェブサイトを特定できる。こうしたサイトはフェイクニュースの発信元とは限らないが、拡散に寄与する影響力のあるメディアであることが多い（噂としてであれ、暴露記事としてであれ）。

第3章

ウェブページの背後にある 広告用の「トラッカー」に注目し、 フェイクニュースの サイトを特定する

フェイクニュースサイトは、どのようなトラッカーを使って、主要メディアからコンテンツを取り込んでいるのか？

トラッカーの使い方でフェイクニュースサイトと主要メディアのサイトを区別する方法

フェイクニュースサイトのトラッカーは、時間とともにどう変わっていくのか？

フェイクニュースサイトで使われているトラッカーIDは、他のどのようなサイトで使われているか

政治的に極端に偏った内容や、誤解を招く情報を掲載するサイトが使うトラッカーは、言語圏によって異なるのか

<第3章のポイント>

「トラッカー」は、ウェブサイトに埋め込まれているアクセス解析を行うための仕掛けで、広告ページに移動するユーザー数の計測や、記事の拡散状況を記録するものなどがある。トラッカーを分析し、記事の信頼性を判断する。

はじめに - 誤った情報や虚偽の情報はこれまで数十年にわたって、政治家やジャーナリストなどの公的な立場にいる人たちの指摘や告発によって注目を集めてきた。どのような指摘がなされ、それがどのように広がっているのかを調べるだけでなく、指摘を受けているウェブサイトの構造に注目することで、ウェブ上のフェイクニュースを見極める方法はないのだろうか。

多くのウェブサイトでは訪れた人の数や行動、それに広告の効果など、サイトのエンゲージメントを計るために、「トラッカー」と呼ばれるコードをいくつも埋め込んでいる。本章ではフェイクニュースや、誤解を与える情報を見極める道具として、トラッカーのデータを使いこなす方法について説明する。



フェイクニュースサイトは、 どのようなトラッカーを使って、 主要メディアからコンテンツを 取り込んでいるのか？

始める前に

この項目で作業を始めるにあたって、フェイクニュースサイトと、主要メディアのサイトのURLをそれぞれ一覧にした二つのリストを準備しよう。既存のリストを使っても良いし、さまざまなメディアで公開されているリストを編集し直して自分で新たなリストを作っても良いだろう。しかし、この二つのリスト作りがこの後の作業の核になることを肝に銘じてほしい。出発点となるリストの選び方によって結論の見方にも大きく影響する。この項目ではBuzzFeed Newsが作ったフェイクニュースサイトのリスト（🔪 **BuzzSumo**の分析でもっともコンテンツが閲覧されているリスト）、そして、BuzzFeed NewsとAlexaから得られた主要メディアのリストを使うことにする。

START

obtain seed lists
of fake news URLs



input URLs in

DMI Tracker Tracker



output data to



- > Extract list of trackers associated with web pages on each list
- > Calculate tracker usage per site type



import data in

RAWGraphs



visualise



**DO MAINSTREAM
MEDIA AND FAKE NEWS
WEBSITES SHARE THE SAME
TRACKER ECOLOGY?**

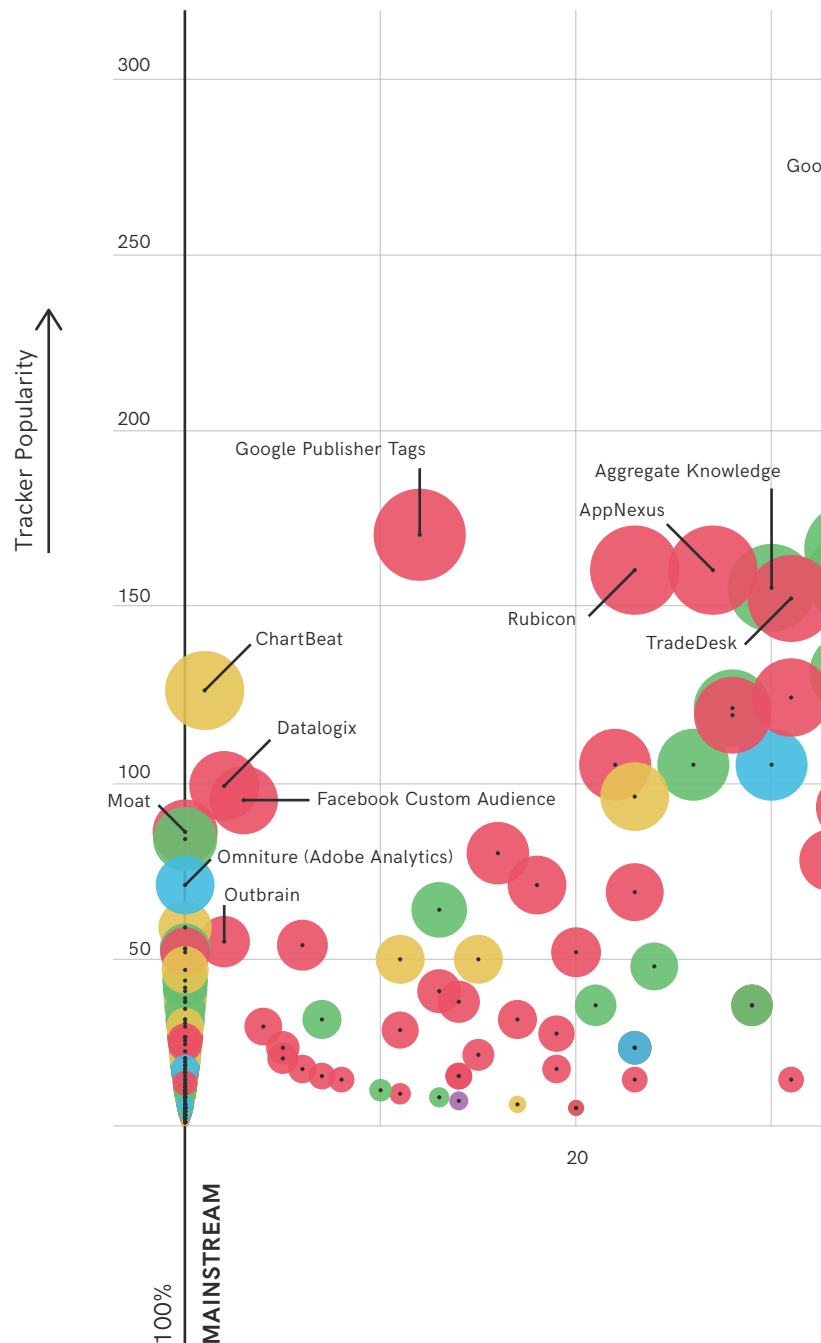
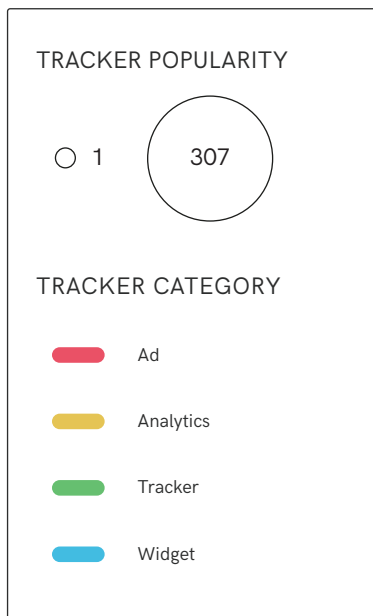
サイトのタイプごとに、使われているトラッカーを数える

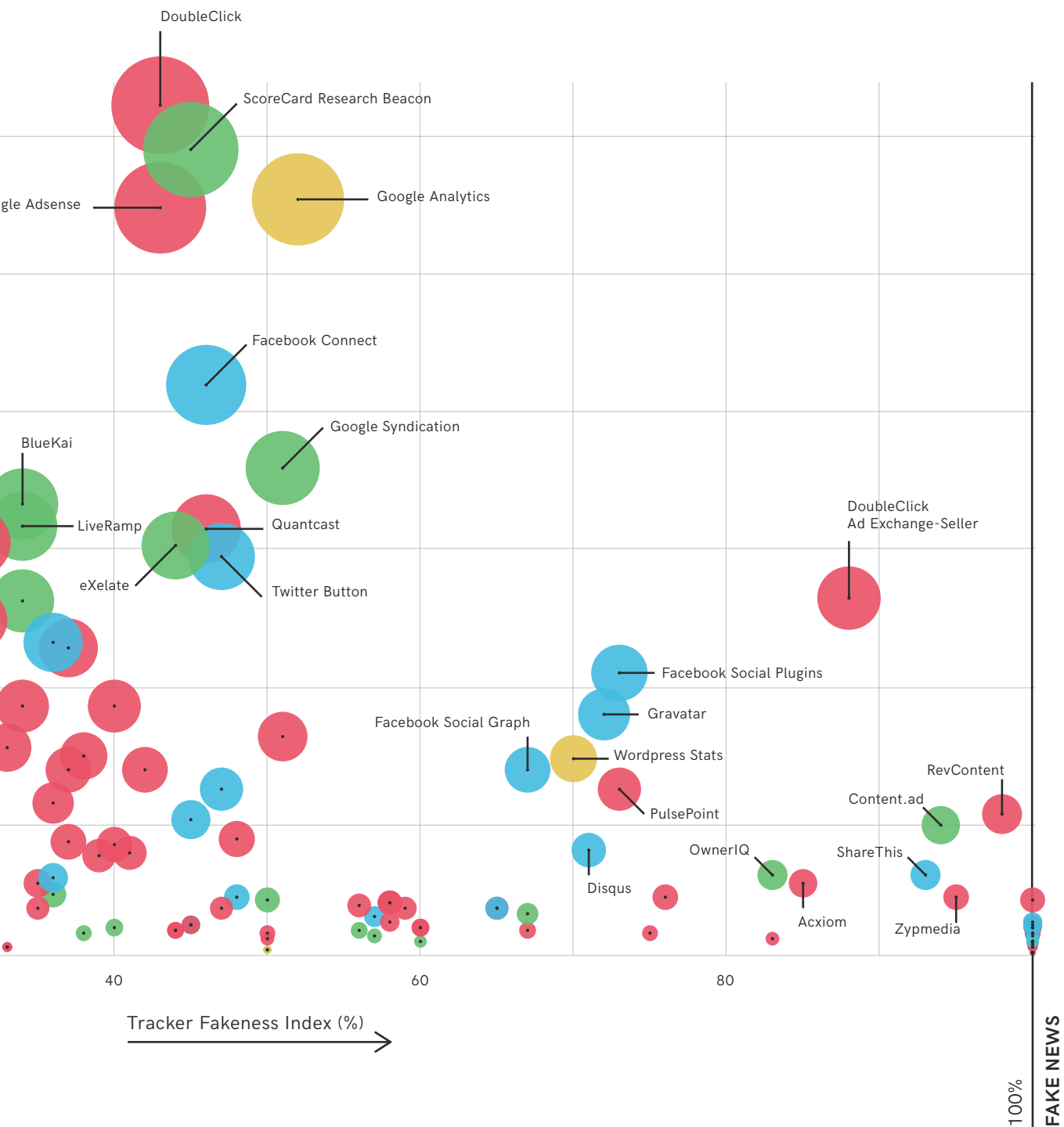
ウェブページのソースコードを見ると、多くの場合、どういうトラッカーのサービスが使われているかということがわかる。

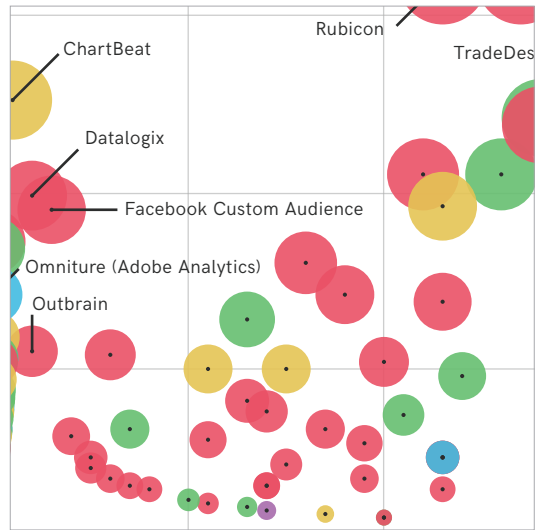
- ◇ 二つのリストにあるウェブページのトラッカーに関するデータを集めてみよう。こうした情報を集めるのに、🔧 **DMI Tracker Tracker** のようなツールを使うと便利だ。
- ◇ フェイクニュースサイトと主要メディアのサイトとで、それぞれいくつのトラッカーが使われているのか、数えてみよう。
- ◇ 調べたデータは散布図を使って可視化するとわかりやすい。ひとつひとつの円は、それぞれひとつのトラッカーを表していて、カテゴリー別に色分けされている。横軸は、例えばトラッカーがフェイクニュースサイトで使われている割合の分布を、縦軸ではトラッカーが使われた総数について表示すると良いだろう。ここでは、🔧 **RAWGraphs** というツールを使って可視化を試みた。

主要メディアと フェイクニュースサイトで 使われるトラッカーの傾向

フェイクニュースサイトと、主要メディアのサイトでのトラッカーの使い方を示す散布図。グーグルアドセンスやDoubleClick、グーグルアナリティクスなどがフェイクニュースサイトと主要メディアサイトの双方で人気がある一方で、主要メディアが使うトラッカーは厳選され、種類が限られていることがわかる。








おすすめ手法

この方法を使えば、さまざまなウェブサイトにはめ込まれているトラッカーの使われ方や、広く普及しているトラッカーの種類について調べ上げることができる。フェイクニュースサイトでしか使われていないようなトラッカーを特定したり、人気のあるトラッカーがどう使われているかについて調べることができる。利用頻度が少なく使い方が特化された、いわゆる「ロングテール」の部分に属するトラッカーについて調べるときにも有効な方法だ。



トラックターの使い方 フェイクニュースサイトと 主要メディアのサイトを 区別する方法

始める前に

この項目で作業を始めるにあたって、フェイクニュースサイトと主要メディアのサイトのURLを、それぞれ一覧にした二つのリストを準備してほしい。既存のリストを使っても良いし、さまざまなメディアで公開されているリストを編集し直して自分で新たなリストを作っても良いだろう。しかし、この二つのリスト作りがこの後の作業の核になることを肝に銘じてもらいたい。出発点となるリストの選び方によって、結論の見方にも大きく影響する。この項目では、BuzzFeed Newsが作ったフェイクニュースサイトのリスト( **BuzzSumo** の分析でもっともコンテンツが閲覧されているリスト)、そして、BuzzFeed NewsとAlexaから得られた主要メディアのリストを使うことにする。

START

obtain seed lists
of fake news URLs



input URLs in

DMI Tracker Tracker



output data to



> Extract list of trackers
associated with web pages
on each list



import data in

Gephi



visualise



***HOW DO FAKE NEWS SITES
AND MAINSTREAM MEDIA
CLUSTER ACCORDING THEIR
TRACKER USAGE?***

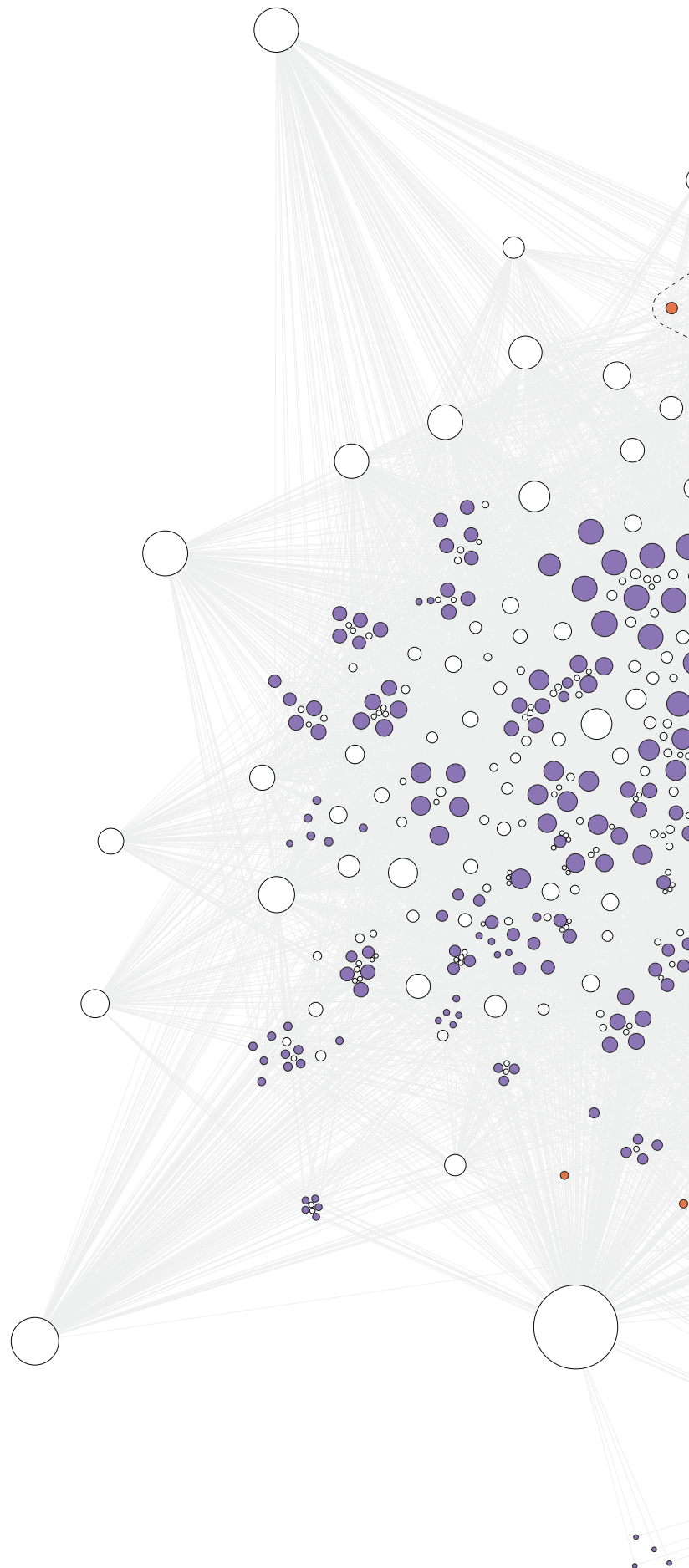
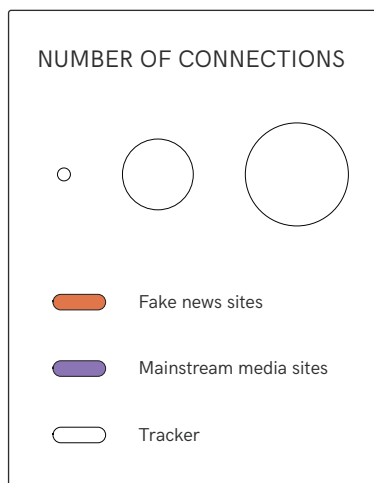
ウェブページとトラッカーとの関係を可視化する

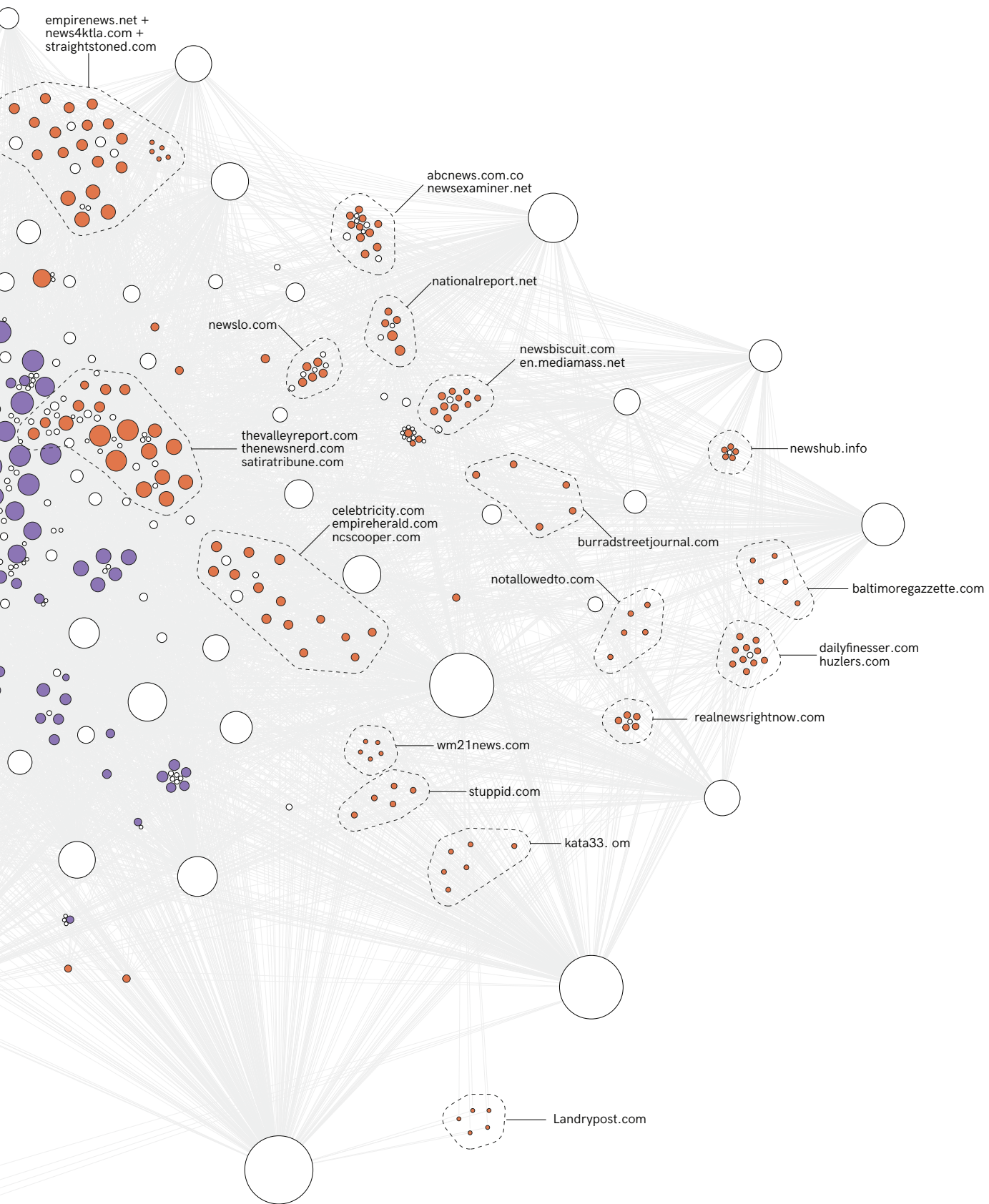
別々のURLのウェブページであっても、似たようなトラッカーの使い方をしていることがある。こうした関係をネットワーク図で表すことで、それぞれのウェブページと、トラッカーとの関連を強調して示すことができる。

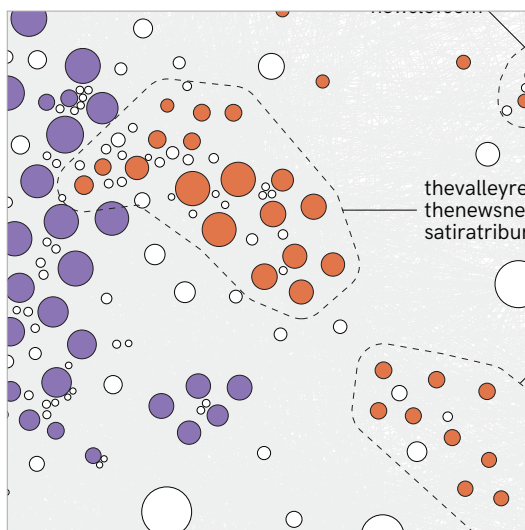
- ◇ **DMI Tracker Tracker** のようなツールを使って、最初に作ったフェイクニュースと、主要メディアのリストにそれぞれ関連するトラッカーを抽出し、リストアップしてみよう。
- ◇ それぞれのウェブページと、トラッカーの使われ方との関係を示すためのネットワーク図を作ってみよう。ここではGephiを使い、力学モデル(**force directed network layout**)のレイアウトで関係性を可視化した。
- ◇ 似通ったウェブサイトの集団を強調するために色を変えたり注釈を付けたりするなどして、グループ化するのも良いだろう。(フェイクニュースサイトのグループ／主要メディアのグループ、など)

フェイクニュースサイトと 主要メディアのサイトの トラッカーの使い方による グループ

フェイクニュースサイトと主要メディアのサイトのトラッカーの使い方によるグループ。トラッカーと、それを使うウェブサイトの対応を示すネットワーク図。同じトラッカーを使っている複数のウェブサイトをグループ化することで、それらのサイトでどのような目的でトラッカーが使われているかを明確にすることができる。また、グループ化によって「フェイクニュースのグループ」を浮き上がらせ、フェイクニュースサイトの検出にも役立つ。







おすすめ手法

この項目では、トラックの使われ方に応じて複数のウェブページをグループ化する方法を学んだ。こうしてトラックとウェブサイトとの関係を可視化することで、リストアップされていたウェブサイトや、その定量的な分析(もっとも使われている/使われていないトラックなど)についてわかりやすく見渡すことができる。例えばフェイクニュースの潜在的な「メディアグループ」を浮き上がらせることで、本格的な調査を始めるにあたっての出発点にもなる。またフェイクニュースのウェブページでのトラックの使われ方や、そのパターン、そして実際のフットプリントを調べてそれぞれの違いや類似点を挙げることで、フェイクニュースの特徴をつかむことができる。さらに取り上げている地域やテーマ、それにソースがまったく異なるサイトを比較することも有効だ。



フェイクニュースサイトの トラッカーは、時間とともに どう変わっていくのか？

始める前に

この項目では、同じウェブページ(または複数のウェブページ)の二つの異なる時点でのソースコードを使って作業を進める。同じページのソースコードを保存し、蓄積していくのは手作業でも自動的にもできるし、🔧 **The Wayback Machine**のようなインターネット上のアーカイブプロジェクトを活用することもできる。

START

obtain seed lists
of fake news URLs



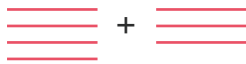
input URLs in

The Wayback Machine

obtain past versions
of the pages



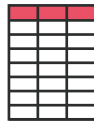
merge all URLs in a single list



input URLs in

DMI Tracker Tracker

output data to



> Extract list of trackers
associated with web pages
on each list

import data in

Gephi

visualise



***HOW DO FAKE NEWS SITES
ADAPT THEIR TRACKER USAGE
IN RESPONSE TO BLACKLISTING
FROM MAJOR AD NETWORKS?***

ウェブページとトラッカーとの関係を可視化する

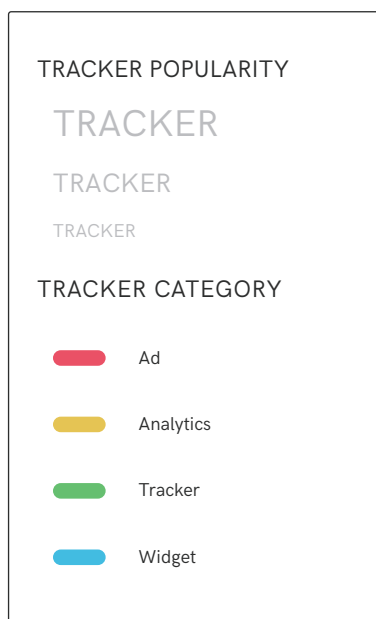
この項目では、同じウェブページでも二つの異なる時点で、トラッカーの使われ方がどのように変わったのかについて調べていく。グーグルとフェイスブックは2016年11月に、フェイクニュースを掲載するサイトには広告サービスを使えなくするプログラムを導入した。例えばこうした出来事の影響や、その前後の対応について調べるうえで、トラッカーの使われ方の変化を明らかにしていくのは有効な調べ方のひとつだ。

- ◇ ウェブページのアーカイブをいくつか用意しよう。🔧 **The Wayback Machine** を使えば、あるページが時間とともにどのように変わったのかがわかる。
- ◇ 過去と現在のウェブページで、それぞれどのようなトラッカーが使われているのか、調べてみよう。🔧 **DMI Tracker Tracker** を使えば、比較的簡単に調べることができる。
- ◇ 過去のページにしか使われていないトラッカー、現在のページにしかないもの、そして双方で使われているトラッカーに分類してみよう。
- ◇ これらの三つのグループをそれぞれリストアップし、さらにわかりやすいように、トラッカーのカテゴリーに応じて別々の色を付けてみよう。

主要な広告ネットワークから 拒絶されたフェイクニュースサイトは、 その後どのようにトラックターの 使い方を変えるのか

主要広告ネットワークから拒絶されたフェイクニュースサイトのトラックターの動向。

フェイクニュースサイトのソースコードには、広告ネットワークから拒絶された後でも、使えなくなったはずのこうしたトラックターの形跡がいくつか残っている。現在は使われていないトラックターでも、サイトのソースコードには、コードが消えておらず、残っているケースがあるからだ。一方で、図を見ると、フェイクニュースサイトがその後、新たに開拓した広告ネットワークも浮かび上がってくる。ウェブサイトに広告を掲載するための仕組みを丹念に見ていくことで、分析をさらに進め、フェイクニュースのさらなる発掘にも役立つ可能性がある。



Admarvel
Clicksor
Drawbridge
Facebook Custom Audience
Google Publisher Tags
Gumgum
Media.net
Sekindo
Doubleverify
Visible Measures
Omniture (Adobe Analytics)

TRACKERS PRESENT
ONLY BEFORE DECEMBER

DoubleClick

Google Analytics

Google AdSense

DoubleClick Ad Exchange-seller

Google Syndication

Scorecard Research Beacon

Gravatar

Wordpress Stats

Facebook Connect

Exelate

Disqus

Brightroll

Adobe Audience Manager

Bluekai

Amazon Associates

Appnexus

Bidswitch

Criteo

Mediamath

Openx

Pubmatic

Tradedesk

Facebook Social Plugins

Taboola

Twitter Button

Adtech

Advertising.com

Index Exchange (Formerly Casale Media)

Pulsepoint

Quantcast

Rubicon

Spoutable

Stickyads

Tapad

Teads

Turn Inc.

Yahoo Ad Exchange

Tubemogul

Krux Digital

Liveramp

Acloudimages

Acuity Ads

Adscale

Eyeview

Revcontent

Smart Adserver

Sovrn (Formerly Lijit Networks)

Twitter Advertising

Zyppmedia

At Internet

Twitter Analytics

Aggregate Knowledge

Lotame

Owneriq

Rocket Fuel

Videology

Addthis

Facebook Social Graph

Lockerz Share

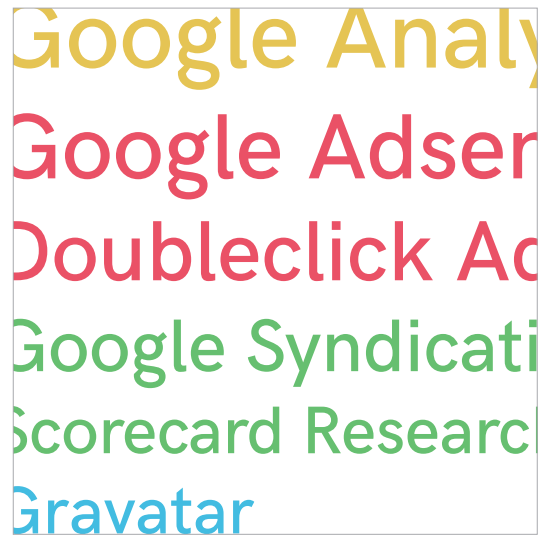
Pinterest

Sharethis

Twitter Badge

Typekit By Adobe

TRACKERS
ALWAYS PRESENT



Google Analy
Google Adser
DoubleClick Ad
Google Syndicati
Scorecard Research
Gravatar

おすすめ手法

フェイクニュースの広告収入を断つための方法について考えるとき、この項目で紹介したノウハウは、トラッカーサービスや、グーグルやフェイスブックなどのプラットフォーム、そして広告主がこれまでとってきた対策を、フェイクニュースサイトがどう乗り越えているかという実態を理解するのに役立つだろう。同時にこうした対処が、フェイクニュースへの対策としていかに有効かということも示せるはずだ。異なる時点でどのトラッカーが使われなくなり、どういったものが残り、そして新たに採用されたトラッカーは何かということを明らかにできるからだ。

START

obtain seed lists
of fake news URLs



obtain the Google Analytics ID
associated to each page



input IDs in

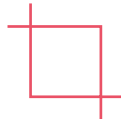
Spyonweb



get the list of associated pages



obtain screenshots
of associated pages



categorise websites groups




visualise



**WHAT MEDIA GROUP STRATEGIES
CAN BE DETECTED THROUGH
SHARED GOOGLE ANALYTICS IDS?**

同じトラッカーIDを使っているページを調べる

この項目では、リストアップされたウェブページで使われているトラッカーIDと同じものを含むウェブサイトを調べていく。

- ◇ 用意したリストのすべてのURLについて、グーグルアナリティクスIDを抽出しよう。この作業は、ソースコードの中で、“UAxxxxxxx”という文字列を手作業で探し出すこともできるし、ウェブスクレイピングや、ほかのツールを使って自動的にリストアップすることもできる。今回はウェブサイトのメタデータから情報を抽出するための専用のスクリプトを書いた。
- ◇ 同じグーグルアナリティクスIDを含むウェブページのリストを作ろう。ここでは、リストを作るのに  **Spyonweb.com** のAPIを使った
- ◇ リストアップされたそれぞれのページのスクリーンショットを撮ろう。ここでは、スクリーンショットの作業を自動化できるプログラムを使った。それぞれのウェブサイトをいくつかのグループに視覚的に分類することが最終的な目標だ
- ◇ 同じIDを使っているスクリーンショットを並べ、グループ内／グループ間で、ページの類似点や違いを調べてみよう。

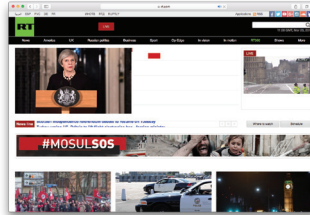
同じグーグルアナリティクスIDを共有するサイトを調べ、メディア戦略を探る

EU Disinformation Reviewから作ったリストのうち、同じグーグルアナリティクスIDを共有するサイトの一覧。この一覧からは、ロシア関連の虚偽情報が掲載されたとされているウェブサイトのルーツの多様性が見て取れる。Russia Todayといった大規模メディアグループから、軍事や神秘主義といったテーマ別のグループ、それにカナダなどの地理的なグループまで、幅広く分布している。こうしたページを見た目で比べると、類似したスタイルや、同様のCMS(=ウェブページのコンテンツ管理システム)で制作されたのではないかと思わせる特徴の見られるページもある。誤解を招く情報を掲載しているウェブページの配信やコミュニケーションの方策を含め、メディアとしてどのような戦略を描いているのかということを探るのにも役立つ。

rt.com

19 disinformation stories

UA-5773642

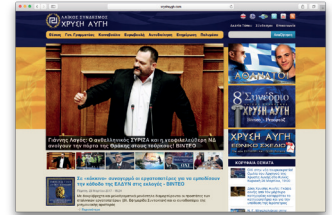


Media group
(Russia Today)

xryshaygh.com

1 disinformation story

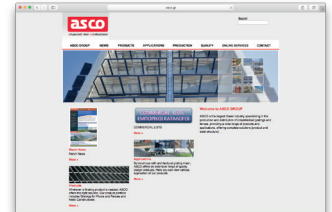
UA-4839940



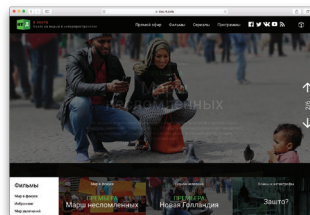
Lone
Webmaster



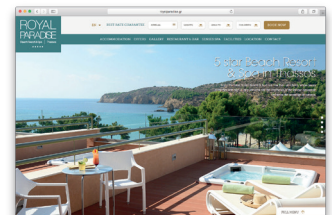
assange.rt.com



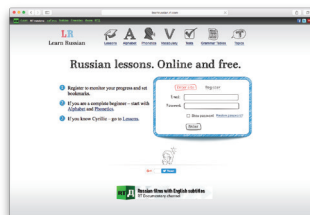
asco.gr



doc.rt.com



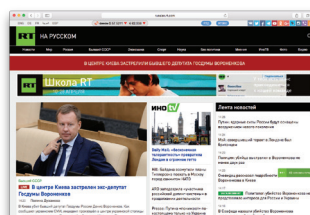
royalparadise.gr



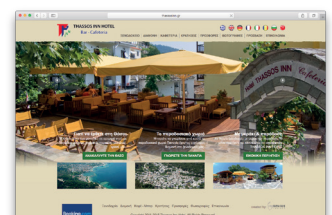
learnrussian.rt.com



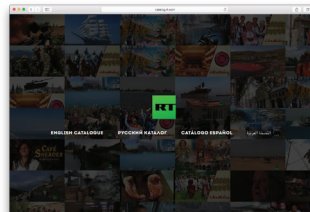
discoverthassos.com



russian.rt.com



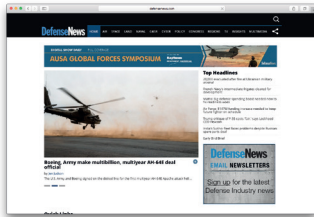
thassosinn.gr



catalog.rt.com

defensenews.com

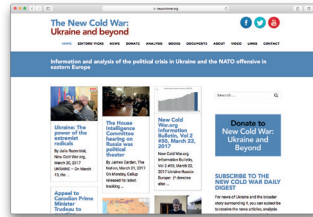
1 disinformation story
UA-841082



Themed network
(Navy, Airforce)

newcoldwar.org

7 disinformation stories
UA-15942468



Canadian socialism
and unions

almanach.cz

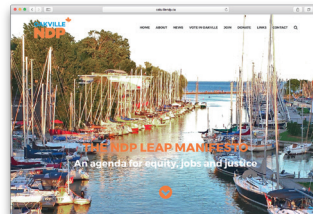
3 disinformation stories
UA-3004323



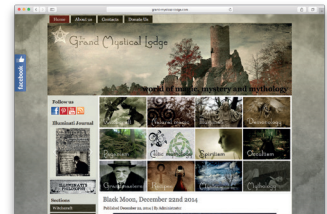
Themed network
(Mysticism, Liberland)



armedforcesjournal.com



oakvillendp.ca



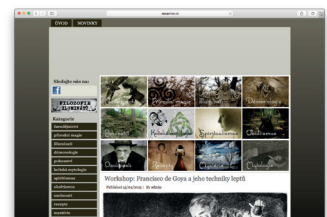
grand-mystical-lodge.com



sightlinemediagroup.com



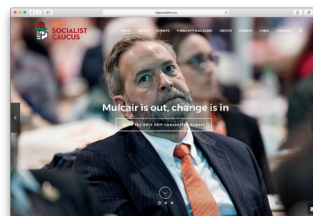
socialiststudies.com



malachim.cz



militarytimes.com



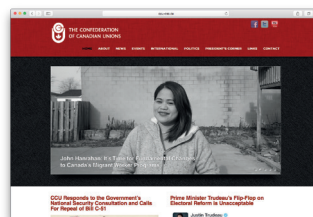
ndpsocialists.ca



liberlandpress.com



airforcetimes.com



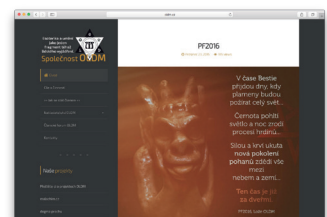
ccu-csc.ca



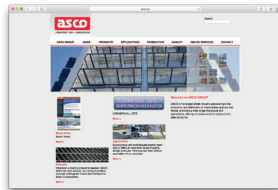
illuminati-journal.com



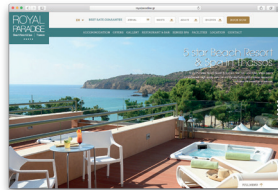
marinecorpstimes.com



oldm.cz



asco.gr



royalparadise.gr

おすすめ手法

この項目で学んだ方法を利用すれば、同じトラッカーIDを含む他のウェブサイトのリストを抽出することで、フェイクニュースに関連する可能性があるサイトの対象を広げていくことができる。また、フェイクニュース供給者をグループ化し、そのデジタル戦略を可視化するためにも有効なノウハウでもある。



政治的に極端に偏った内容や、 誤解を招く情報を掲載する サイトが使うトラッカーは、 言語圏によって異なるのか

始める前に

作業を始めるにあたって、異なる言語によるフェイクニュースや極端に偏った政治的な主張、それに誤解を招く情報が掲載されたウェブサイトのリストを用意しておく必要がある。このリストをもとに、トラッカーの種類や使われ方について調べていく。ここではオランダ語、英語、ドイツ語のサイトを参照しながら説明を進めていくことにする。

START

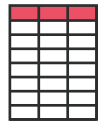
obtain seed lists
of fake news URLs



input URLs in

DMI Tracker Tracker

output data to



> Extract list of trackers
associated with web pages
on each list

compare trackers lists with

DMI Triangulation Tool

visualise

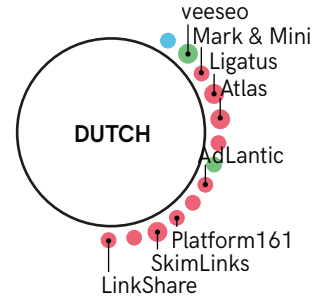


**DO MISINFORMATION AND
HYPER-PARTISAN WEBSITES
IN DIFFERENT LANGUAGE
SPHERES HAVE DISTINCT
TRACKER ECOLOGIES?**

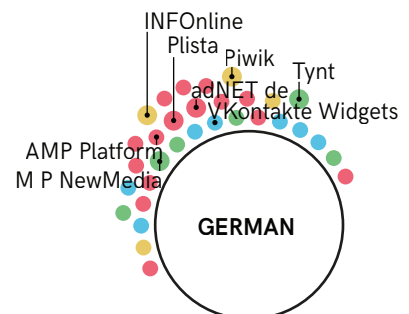
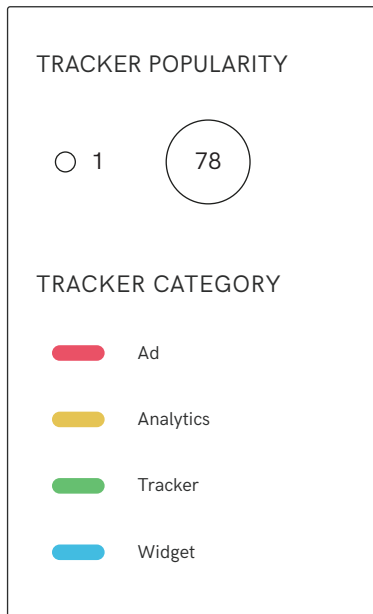
言語ごとにトラッカーを調べる

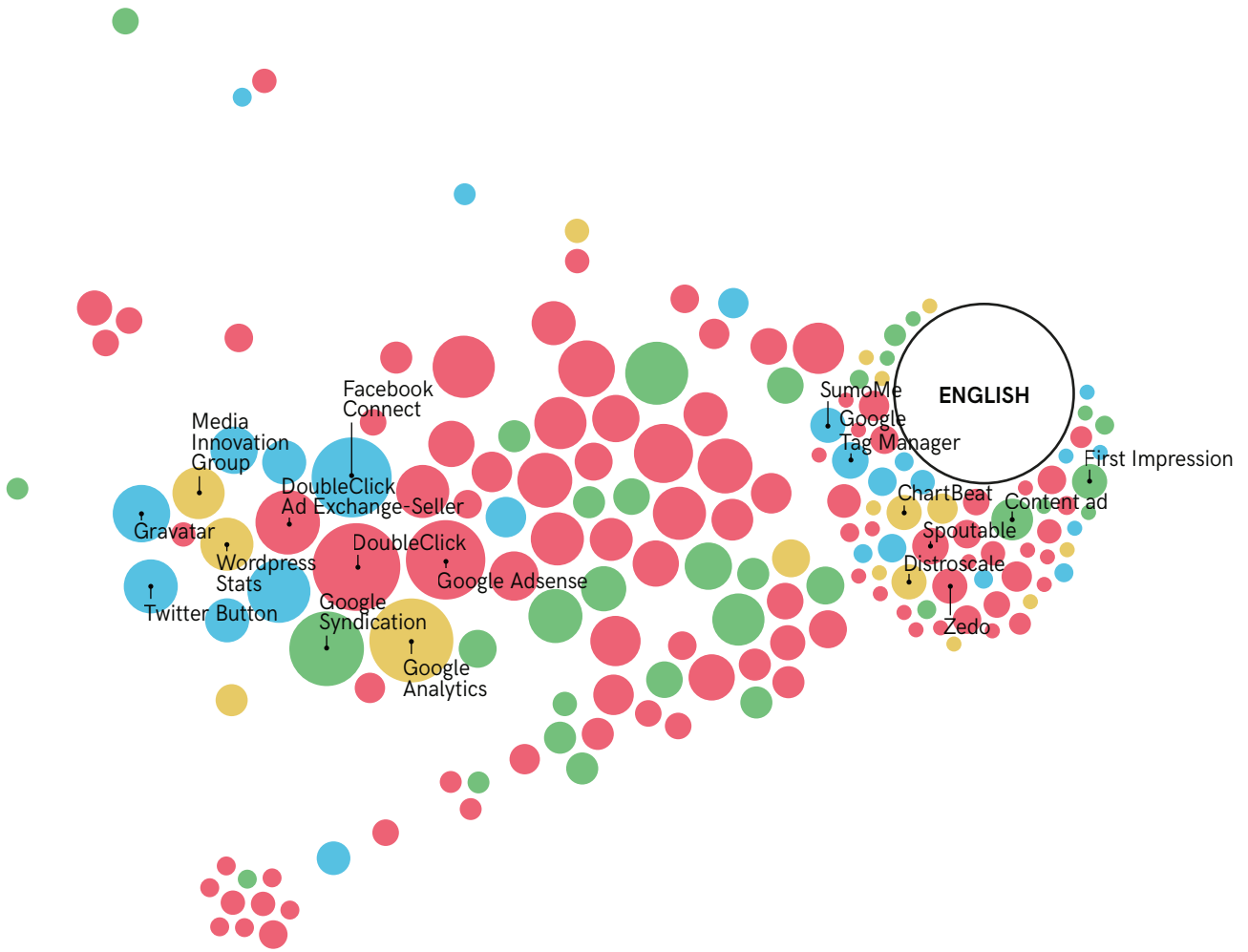
- ◇ 言語ごとにリストアップしたウェブページで使われているトラッカーをすべて抽出しよう。ここでは、🔧 **DMI Tracker Tracker** を使って抽出した。
- ◇ 異なる言語のサイトの間で共通して使われているトラッカーと、特定の言語のサイトでしか使われていないトラッカーを調べてみよう。ここでは、🔧 **DMI Triangulation** を使って調べた。
- ◇ ここで調べた結果については、磁石をイメージして考えるとわかりやすい。三つの言語がそれぞれ頂点となって三角形を作っている。トラッカーは三角形の中で、使われている言語の種類に応じて分布している。つまり三つの言語のすべてで使われていれば、三角形の内部に、二つの言語で使われていれば一辺の線上に配置されるということだ。

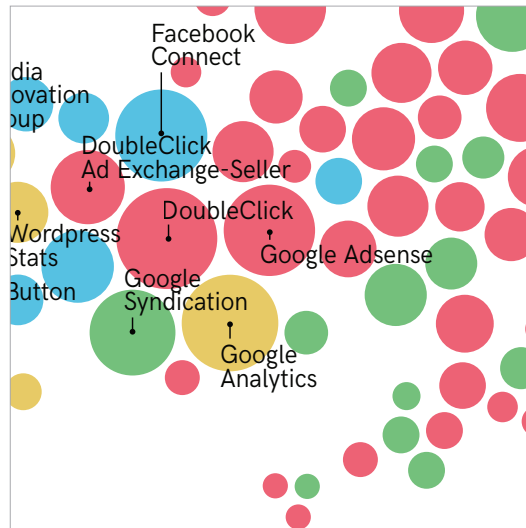
誤解を招く情報や 極端に内容が偏ったサイトの、 言語によるトラッカーの 使い方の傾向



政治的に極端に偏った内容や、誤解を招く情報を掲載したサイトで使われるトラッカーの傾向の可視化。DoubleClick、グーグルアドセンス、フェイスブックコネクトといった、広く普及しているサービスはすべての言語で使われているが、その一方で、言語圏に特化したサービスも見られる。例えば、主にロシア語圏で普及しているMail.ruグループに関連するトラッカーは、ドイツ語圏のウェブサイトでのみ使われている。







おすすめ手法

この項目で学んだ方法を活用すると、言語圏特有のトラッカーについても、言語圏をまたぐものについても、より詳しく調べることができる。フェイクニュースの課金モデルや技術的な背景について異なる言語で特有のものを調べるといふ、新たな調査の方向性を発展させていくのにも役立つ可能性もある。

第4章

フェイスブック上の 政治ミームを調査する

フェイスブック上のミーム空間を
明らかにする方法

政治・メディア的イベントが、
どんなミームになるのか

ミームのコンテンツを調査する方法

<第4章のポイント>

ミームとは、インターネット上で広まるさまざまな形態のうわさを意味する。この章では画像ミームを分析する。

はじめに - ミームとは、ネット上で広がる噂を意味する。このガイドのレシピはニュースを模倣するようなメディアコンテンツに焦点を当ててきた。だが、広く影響を及ぼすハイパー・パルチザンのコンテンツ、誤情報、意図的に人をだます目的で作られられた情報やプロパガンダは、必ずしもBBCやCNN、無数のその他のサイトによく見られるような、「見出し・写真・文章」のコンビネーションで現れるとは限らない。実際のところ画像、特に(合成画像などの)画像ミームは、ソーシャルメディアの生態系において、先に述べたようなコンテンツと同等か、それ以上に拡散されている。

「コロンビア・ジャーナリズム・レビュー」の記事によると、ブライトバートのフェイスブックページで最も多いエンゲージメントを生み出したメディア形式は、画像ミームだ[1]。よって、フェイクニュースサイトだけでなく、ミームのようなバイラルコンテンツのジャンルが、どのように政治アジェンダの設定や政治プロセス、政治文化に関わっているのかを考察することが必要不可欠だ。

本章では、政治的ミームを調査するいくつかのアプローチを提供する。政治的なトピック、人物、出来事を題材と

するミームに焦点を当てる。2016年米大統領選の際にフェイスブックで見られた、オルタナ右翼とトランプ支持派のミームをケーススタディとして扱い、これらのアプローチを説明する。本章では、ミームに関連してユーザーが取る多様なアクション(ミームの拡散、模倣、変換など)を指して、「ミーム的活動」という言葉を用いる。レシピ1では、フェイスブック上でミームを特定し、マッピングする方法を紹介する。レシピ2は、ミーム的活動を通じて、フェイスブックユーザーがどう政治的出来事に関わっているかを調査する方法に焦点を当てる。最後のレシピ3は、ミームのコンテンツを分析する一連のアプローチを提供する。

[1] See, Nausicaa Renner, "Memes Trump articles on Breitbart's Facebook page", Columbia Journalism Review: 2017. Available at: http://www.cjr.org/tow_center/memes-trump-articles-on-breitbart-faces



フェイスブック上の ミーム空間を 明らかにする方法

始める前に

フェイスブック上で話題になっている政治的ミームに、誰が関わっているかを調査するには、フォロワー数が多く、ミーム的活動が見られるページを初めに特定しよう。例としてここでは、トランプ支持派で、オルタナ右翼的活動を活発に行っている「Disdainus Maximus」というページを選んだ。このページに関連するネットワークを追跡し、トポロジー（接続形態）を調べた。

START

Choose a Facebook page
with **memetic activity**

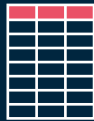


trace the **network of connections** with

Netvizz



extract page **like network** (depth 2)



import **data** in

Gephi



Quantitative Analysis

- InDegree
- OutDegree
- Betweenness Centrality
- Netviz Fan Count Metrics



Qualitative Analysis

- Prominent clusters identification
- Cluster thematic annotation
- Exploration of content and self-description of pages



visualise



a

**WHAT ISSUES ANIMATE
THE INTER-LIKED FACEBOOK
PAGE NETWORK SEEDED
BY A PRO-TRUMP POLITICAL
MEME REPOSITORY?**

START

Choose a Facebook page with **memetic activity**

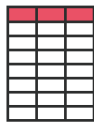


trace the **network of connections** with

Netvizz



extract page **like network** (depth 2)



import **data** in

Gephi



- Quantitative Analysis
- InDegree
 - OutDegree
 - Betweenness Centrality
 - Netviz Fan Count Metrics



- Qualitative Analysis
- Prominent clusters identification
 - Cluster thematic annotation
 - Exploration of content and self-description of pages



visualise



a

WHAT ISSUES ANIMATE THE INTER-LIKED FACEBOOK PAGE NETWORK SEEDED BY A PRO-TRUMP POLITICAL MEME REPOSITORY?

ミーム的活動をしているページのネットワークを調べる

ミーム的活動が見られるページのネットワークを調べるためには、該当ページが「いいね!」をしているほかのページを調べる必要がある(ユーザーからの「いいね!」とは区別する)。

- ◇ ページに関連する「いいね!」のネットワーク抽出には、フェイスブック・クローラーを使うといい。ガイドでは🔧Netvizzの「page like network」モジュールを使い、「互いに“いいね!”でつながっているページのネットワークを構築」した。
- ◇ クローラーの深さを2にセットし、調査対象のページが「いいね!」をしたページ、さらにそれらが「いいね!」をしたページを抽出した。そして、ノードがページで、エッジが「いいね!」を表す有向ネットワークファイルを取得した。
- ◇ 🛠️ Gephi などのネットワーク分析ツールを使い、互いにつながっているフェイスブックページのグラフを調べることもできる。力学モデルのレイアウト・アルゴリズム(ForceAtlas2など)は、ネットワークの形状を可視化し、ミーム的活動が行われている相互につながった空間を詳しく見るのに役立つ。

[1]
See Netvizz, Facebook application, version: 1.3, 2017, at: <https://apps.facebook.com/netvizz/>

START

Choose a Facebook page with **memetic activity**

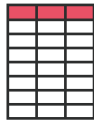


trace the **network of connections** with

Netvizz

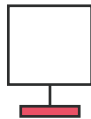


extract page **like network** (depth 2)



import **data** in

Gephi



Quantitative Analysis

- InDegree
- OutDegree
- Betweenness Centrality
- Netviz Fan Count Metrics



Qualitative Analysis

- Prominent clusters identification
- Cluster thematic annotation
- Exploration of content and self-description of pages



visualise



a

WHAT ISSUES ANIMATE THE INTER-LIKED FACEBOOK PAGE NETWORK SEEDED BY A PRO-TRUMP POLITICAL MEME REPOSITORY?

ミーム拡散を盛り上げる話題やテーマをプロフィールする

一つ前のステップで得られた、フェイスブックページ・ネットワークの構成と配列に関する質的・量的分析を行う。→ ネットワークグラフの構造を質的・量的に分析することで、際立つクラスターとノードを特定する。このステップでは次を行う。

- 量的分析:

- ◇ indegreeなどのグラフメトリックを使い、ネットワーク内で一番人気があるページを特定、つまりネットワーク内の他のページからの「いいね」数を見る。
- ◇ 「いいね」しているページを特定、つまりネットワーク内の他のページへの「いいね」数を見る。
- ◇ Betweenness centralityなどのグラフメトリックを使い、ネットワーク内の異なるクラスターをつなぐページを特定する。
- ◇ フェイスブックの「fan count」メトリックを使い、フェイスブックユーザー間で最も人気のあるページを特定する。

- 質的分析:

- ◇ グラフ内にあるノード集団の形態や密集度を視覚的に分析し、際立ったクラスターを特定する。
- ◇ ネットワーク内のページがシェアしているコンテンツの内容やタイトル、ページ説明を調べ、それぞれのクラスター内で共通して取り上げられている話題を特定する。

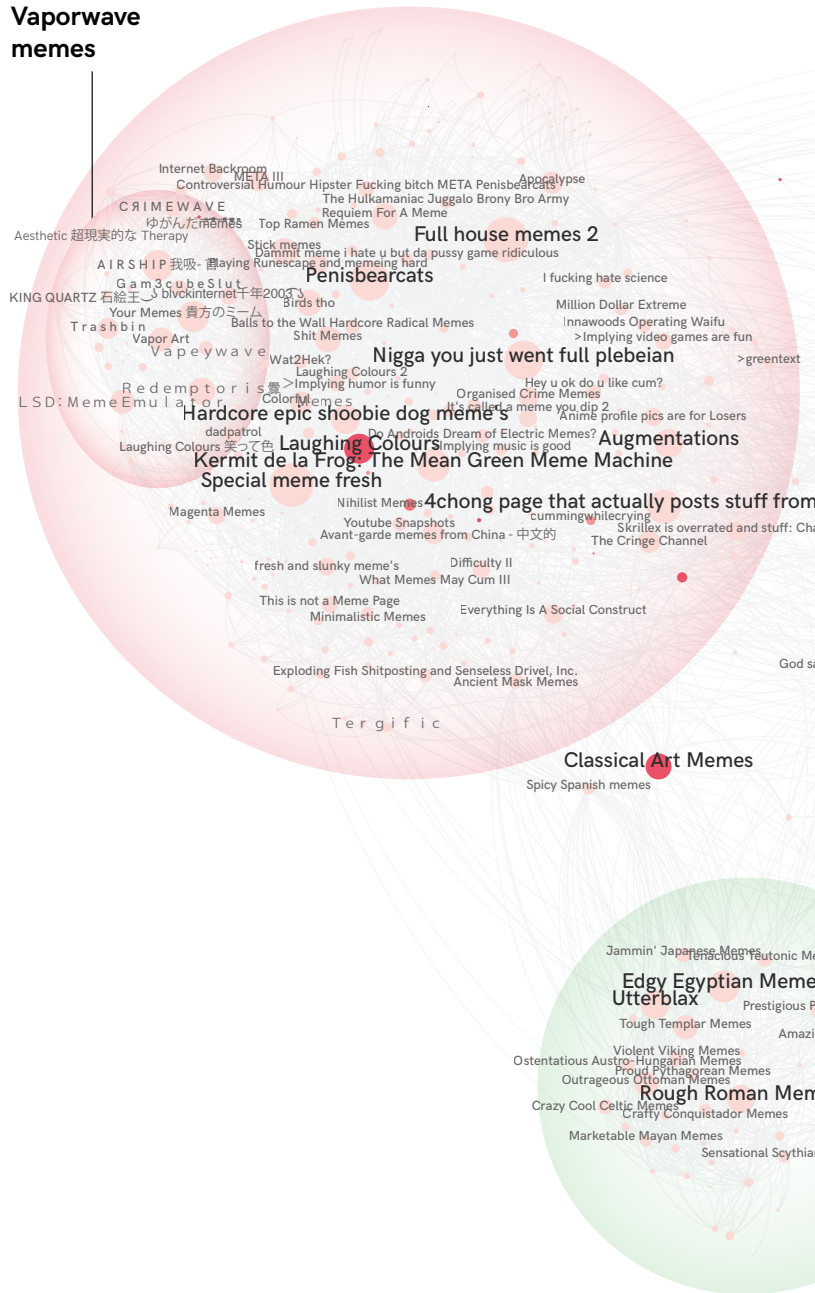
ネットワーク図の可読性を上げるために、クラスターのテーマ別分類の注釈を付ける。

トランプ支持派の 政治的ミームによってつながる フェイスブックページの ネットワークは、どんな話題で 盛り上がっているのか？

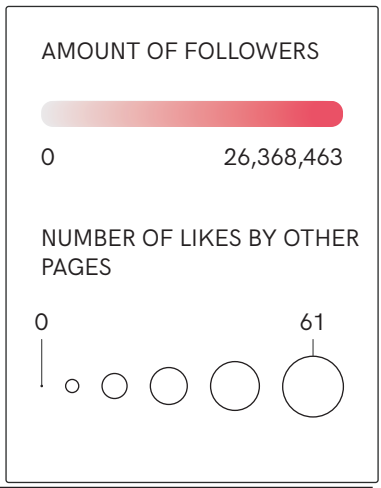
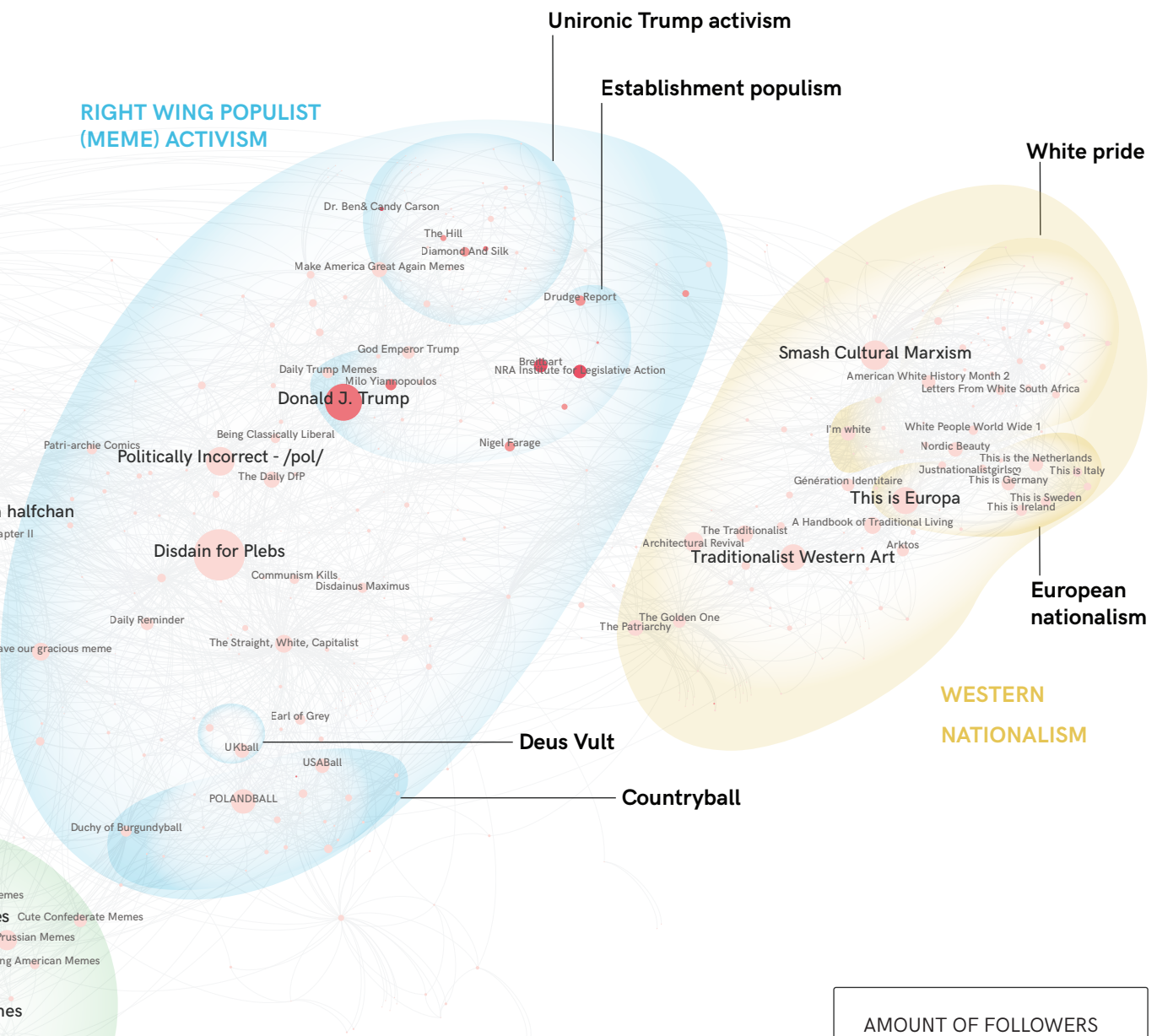
トランプ支持のミームを大量に拡散するページ「Disdainus Maximus」から派生する、フェイスブックページのネットワーク図。このネットワークは751のフェイスブックページで構成され、ミーム的活動が見られる。ページは、互いに「いいね！」し合うことでつながっている。力学モデルのレイアウト・アルゴリズムによって空間化されている。クラスターの特定には、モジュール方式のアルゴリズムが使われている。ノードのサイズは「いいね！」の数に応じて示され、フォロワー数によって色付けされている。トランプを支持する内容のミームは、フェイスブック上のミーム文化において目を引くものだ。これは、もともと政治的ではないページも、トランプのミームを拡散するページとつながっていることから分かる（さまざまな形態のナショナリズムやポピュリズムに特化したページと同様に）。ドナルド・トランプの公式ページがこのネットワークの中心に位置していることから、トランプが、オルタナ右翼のミーム空間を活気づけていることが示唆される。

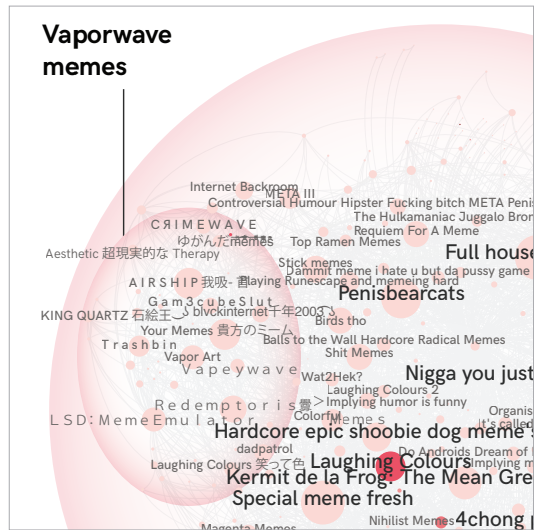
PLURALIST APOLITICAL MEMES

Vaporwave memes



IMPERIAL MEMES





おすすめ手法

このレシピは、フェイスブック上で政治的ミームを大量に拡散するページの特定に使うことができる。そうしたページの関係性や、ミーム化されるテーマの調査にも役立つだろう。



政治・メディア的イベントが、 どんなミームになるのか

始める前に

政治的イベント、またはメディアイベントがどんなミームになるかを明らかにするために、まずは調査対象となるイベントを特定する必要がある。レシピを説明するために、ここでは米大統領選に関連する以下の主要な出来事を例として使用した。

日時	説明
2016年3月3日	共和党第11回討論
2016年6月3日	メキシコ系判事に対するトランプの発言
2016年9月9日	クリントン氏の(トランプ支持者は)「哀れ」発言
2016年10月8日	トランプ氏の女性侮辱的発言(録音音声)
2016年10月29日	クリントン陣営幹部のEメール流出
2017年1月27日	トランプ氏による最初の渡航禁止令発表

START

Choose a set of Facebook pages with memetic activity



Make a list of relevant political events

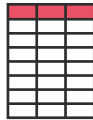


Netvizz

extract photos timeline metadata (including image URLs) for the analysed timeframe



get images URLs



input images URLs in

DownThemAll!
or
Tab Save

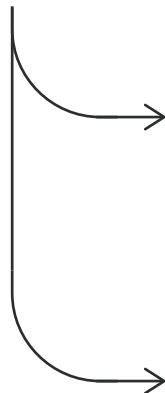


create a corpus of images



Create an Image Montage with

Image J



visualise



a

HOW DO MEMETIC REACTIONS TO DIFFERENT POLITICAL EVENTS COMPARE?

visualise



b

HOW DO MEMETIC REACTIONS AROUND A SINGLE POLITICAL EVENT COMPARE ACROSS FACEBOOK PAGES?

イベントに関連するミームを見つける

説明のため、ここではレシピ4.1で集めたトランプに関わる751のフェイスブックページから、46ページを選んで使用する。量的・質的な基準(エンゲージメント数や、テーマ別クラスターなど)に基づき選定を行った。

- ◇ ミームを選ぶ期間を設定する。ここでは、調査対象となるイベントが起きた日から3日間とした。
- ◇ 設定した期間中に、対象ページに投稿されたミームのリストを抽出する。画像リストや関連するメタデータの収集には、
🔧 **Netvizz**などのツールが使える。
- ◇ 各URLの画像をダウンロードする。Chromeには 🔧 **Tab Save**、Firefoxには 🔧 **DownThemAll!** などのブラウザ拡張機能が使える。
- ◇ 画像イメージを並べて置き、イベントをグループ化することで、それぞれのイベントへのミームの反応を比べることができる。
🔧 **ImageJ** などのツールが使える (4.2.a)。
- ◇ さまざまなページが一つのイベントに対してどう反応しているかを調査することで、ページにまたがる比較ができる (4.2.b)。

さまざまな 政治的イベントに対する ミームの反応を 比較するには

2016年米大統領選に関連した六つの主要イベントにおける、フェイスブックページ上のミーム的活動を画像処理したもの。さまざまなイベントを取り巻くミーム的活動を、そのスケールと構成において比較できる。

Eleventh Republican
debate; blood coming out
of her... wherever
3rd of March 2016

Trumps' Mexican Judge
remark
3rd of June 2016

Basket of Deplorables comment

11th of September 2016



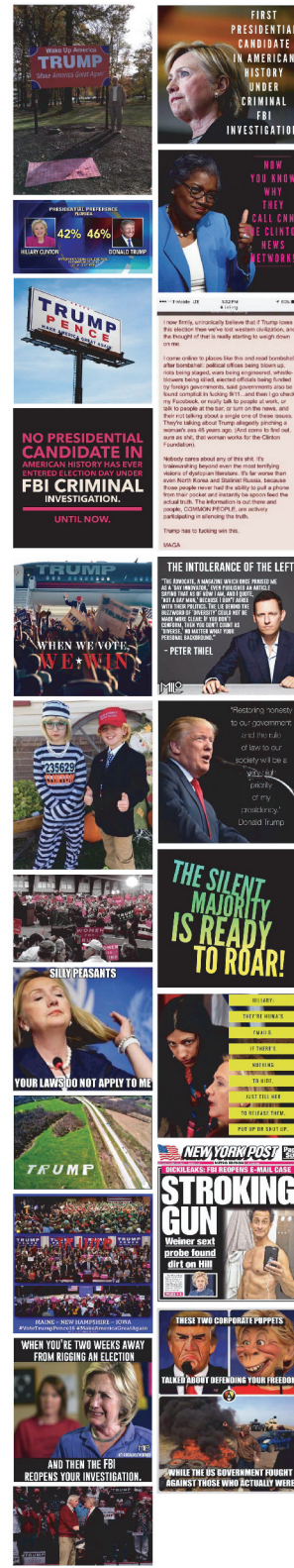
Trump's Taped Comments About Women

8th of October 2016



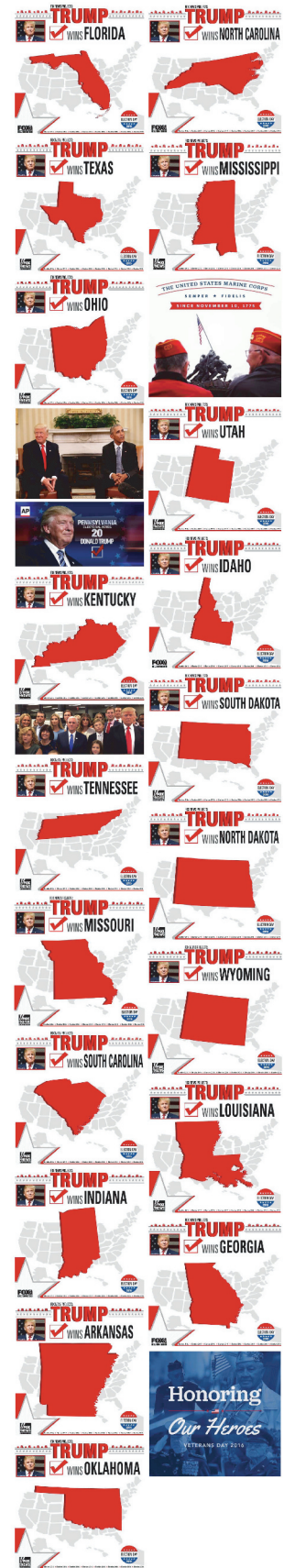
Podesta emails

29th of October 2016



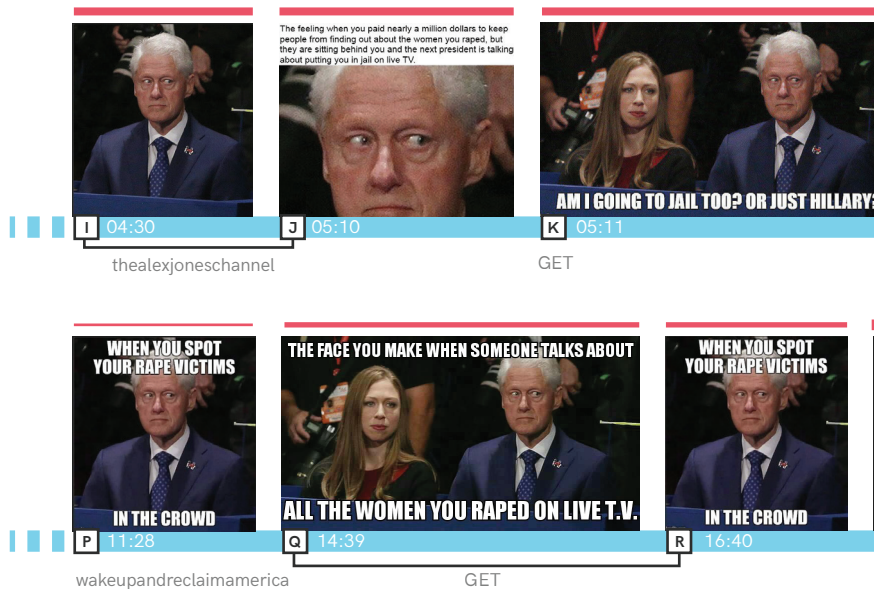
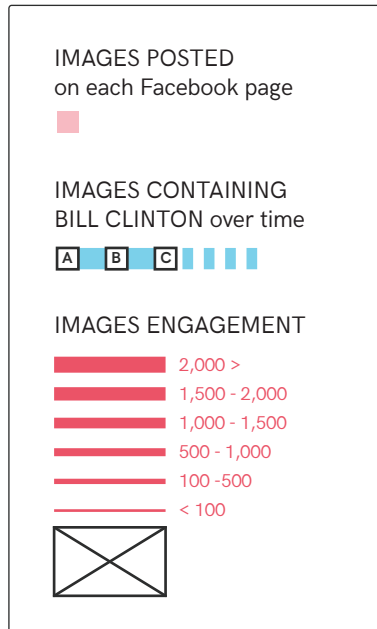
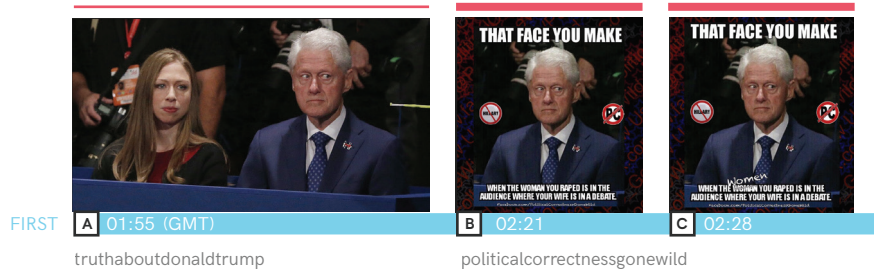
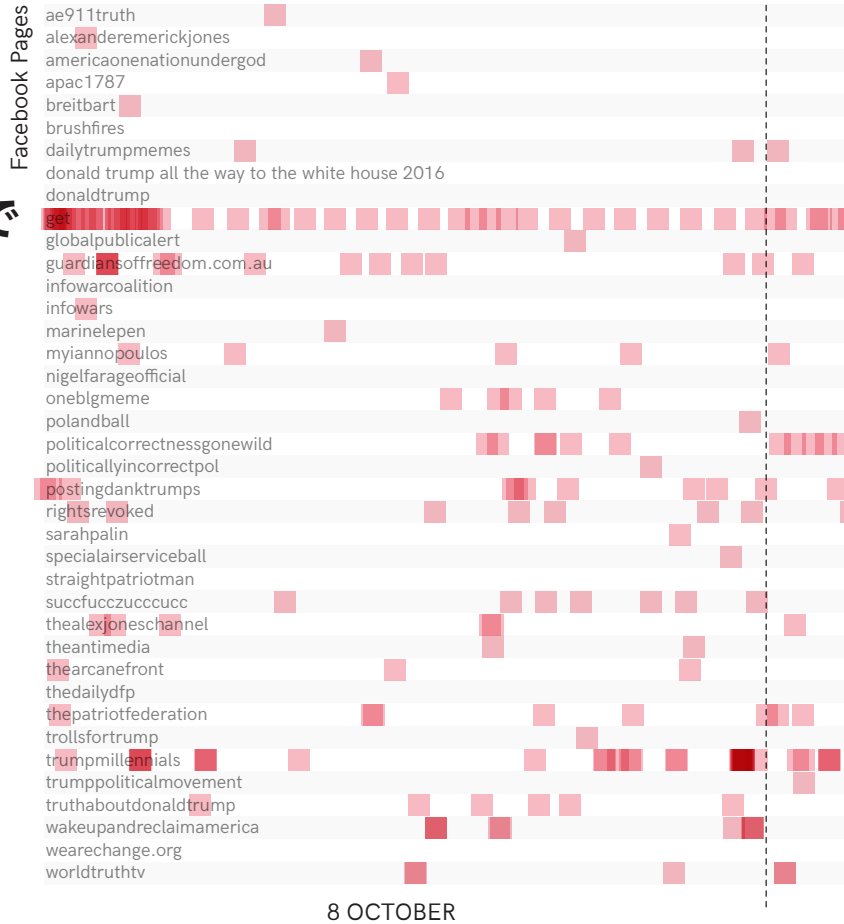
Election

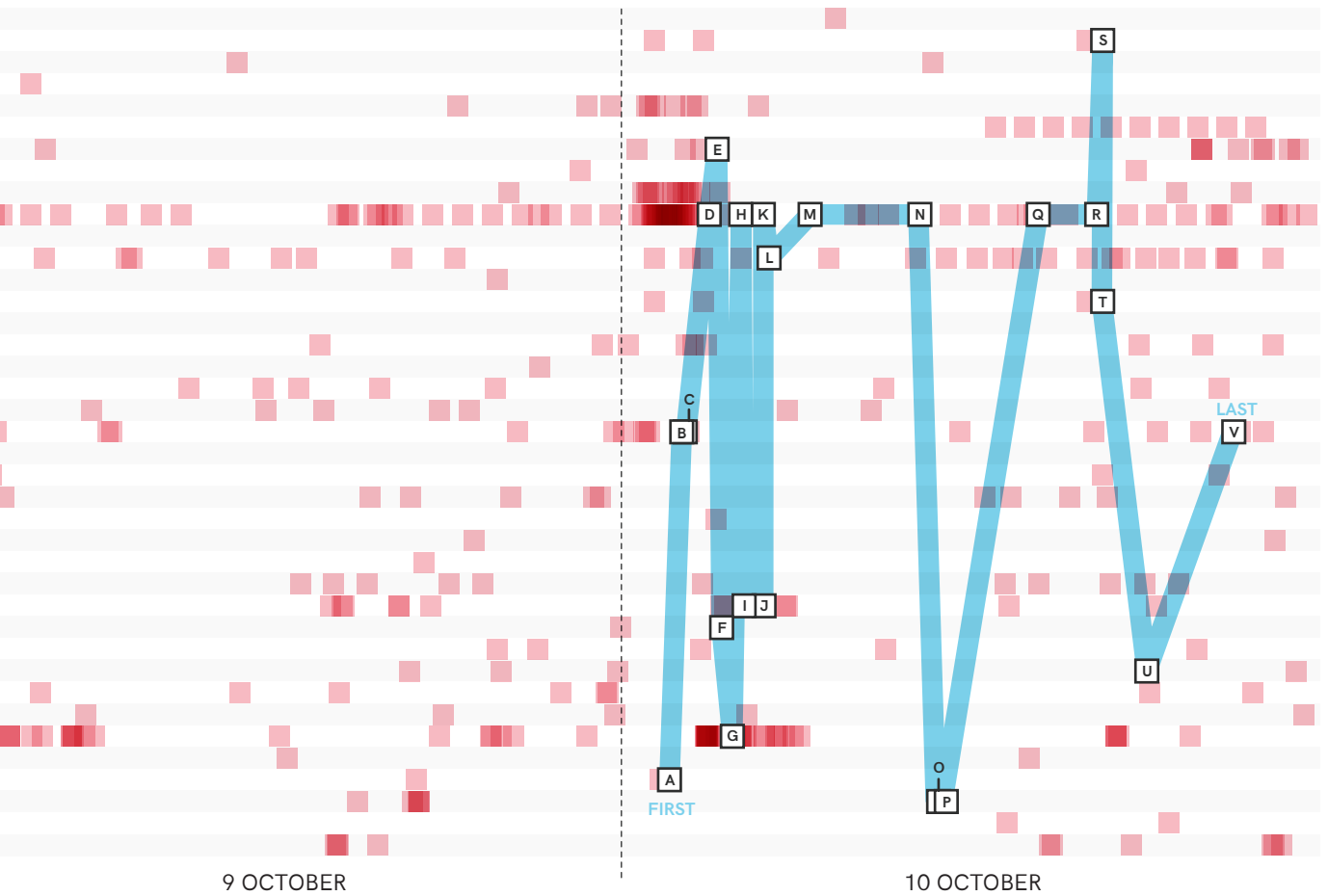
9th of November 2016



さまざまな 政治的イベントに対する ミームの反応を、異なる フェイスブックページ上で 比較するには

特定の説明が加えられたミームを画像処理したもの。トランプ支持派のフェイスブックページ上で、同氏の女性侮辱的発言(録音音声)に関するミーム的活動で顕著に見られたものだ。ビジュアル化することで、トランプが女性への「わいせつ行為」を語る音声が公になった際、ビル・クリントンの過去のセクハラ疑惑に関する画像を投稿したページがあることが分かる(クリントンが、自身によるセクハラ被害者とされる人物と同じ場所に居合わせたという内容の画像)。特にこうした説明的ミームは、ブライットバートが使う手法と同類のようだ。





D 03:18 GET



E 03:34 dailytrumpmemes



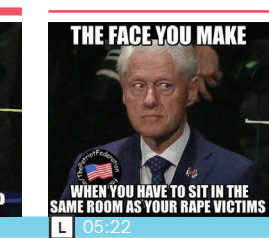
F 03:43 theantimedia



G 04:05 trumpmillennials



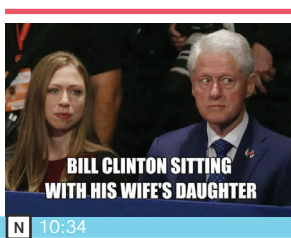
H 04:23 GET



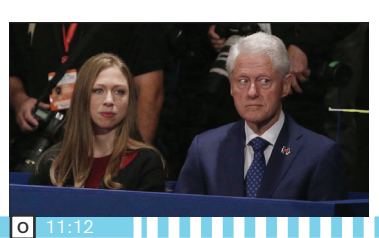
L 05:22 guardiansoffreedom.com.au



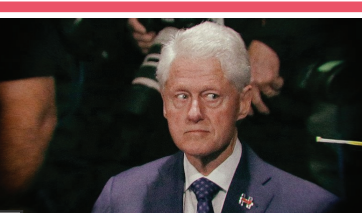
M 06:46 GET



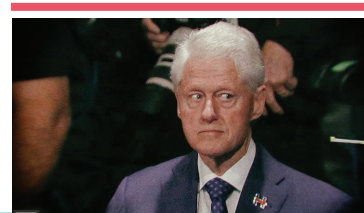
N 10:34



O 11:12 wakeupandreclaimamerica



S 16:52 alexanderemerickjones



T 16:52 infowars



U 18:23 thedailydfp



V 21:23 politicalcorrectnessgonewild

LAST

おすすめ手法

このレシピは、政治的イベントを取り巻くビジュアル文化に見られる、一般参加型コンテンツ作成の調査にも使える。また、それがアジェンダ設定の動きにどんな影響を与えているかも見ることができる。



ミームのコンテンツを 調査する方法

始める前に

このレシピは、ミーム的活動のジャンルやスタイル特定だけでなく、ミームに含まれているテキストや画像の分析方法について解説する。例として、トランプ支持のミームを扱うブライtbartと、God Emperor Trumpのフェイスブックページを使用する。これらのページに投稿される画像は、古典的なミームのフォーマット(画像にテキストを加えた「イメージ・マクロ」など)と合致しないとしても、拡散性、ユーザーが手を加えていく手法、模倣、間テキスト性など、ミームの特徴を備えている。ブライtbartは、オルタナ右翼文化を盛り上げるのに中心的役割を果たしていることから選んだ^[1]。God Emperor Trumpは、ガイド執筆時点で24万5千を超える「いいね!」とフォロワーを持ち、トランプ支持派、オルタナ右翼の最も人気あるミームページの一つだ。ブライtbartのページで画像を投稿できるのは管理者だけだが、God Emperor Trumpでは、ユーザーが管理者にコンテンツを提供し、投稿することができる。

[1] Y., Faris, R., Roberts, H., & Zuckerman, E. (2017, March 3). Study: Breitbart-led right-wing media ecosystem altered broader media agenda. Retrieved March 8, 2017, from <http://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php>

START

Choose one or more Facebook pages with memetic activity



input pages in
Netvizz



get images URLs



input images URLs in

DownThemAll!
or
Tab Save

create a corpus of images



analyse corpus with

Google Vision API



Use optical character recognition (OCR) to extract the text contained within each image



computational linguistic analysis

Cortext

filter images according to computational linguistic analysis



visualise



a

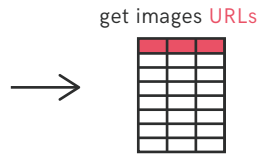
WHAT THEMES LEND THEMSELVES TO MEMETIC ACTIVITY ON BREITBART'S FACEBOOK PAGE?

b

CAN WE DETECT DISTINCT VISUAL STYLES WITHIN A MEME REPOSITORY?

START

Choose one or more Facebook pages with memetic activity



create a corpus of images



analyse corpus with



Use optical character recognition (OCR) to extract the text contained within each image

computational linguistic analysis



filter images according to computational linguistic analysis



a
WHAT THEMES LEAD THEMSELVES TO MEMETIC ACTIVITY ON BREITBART'S FACEBOOK PAGE?



b
CAN WE DETECT DISTINCT VISUAL STYLES WITHIN A MEME REPOSITORY?

選定したフェイスブックページに投稿された画像と、関連するメタデータを集める

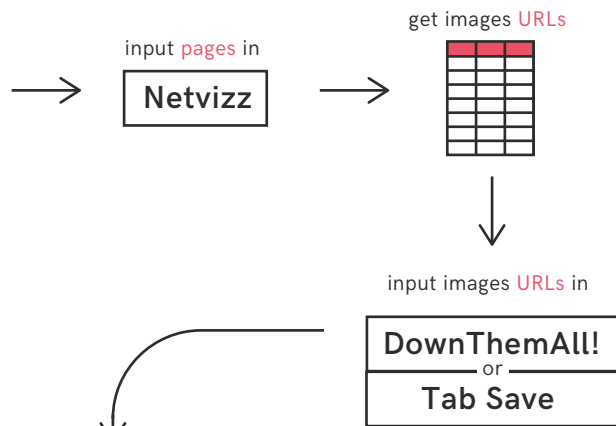
コーパスを作成するには、関心のある期間(例えば、特定の政治・メディア的イベントに近い日時。レシピ 4.2を参照)を選ぶか、フェイスブックページに投稿されたすべての画像を使用する。

- ◇ フェイスブックのAPIを使うと、ページに投稿された画像に関連するメタデータの抽出が可能になる。「写真」タブからアクセスできる。
- ◇  **Netvizz** などのデータ抽出ツールで「page timeline images」モジュールを使っても、メタデータの収集ができる。
- ◇  **Netvizz** の「page timeline images」モジュールを使用することで、それぞれの画像に関連したメタデータ(作成日時、URL、いいね!、リアクション、コメント数)などを含むタブ区切りファイルが取得できる。
- ◇ Chromeの  **Tab Save**や、Firefoxの  **DownThemAll!** などのブラウザ拡張機能を使えば、画像をダウンロードできる。
- ◇ 画像に含まれるテキストを抽出するには、光学文字認識(OCR)ソフトウェアが必要になる。このレシピでは、 **グーグルの Vision API** と、画像URLのリストをVision APIに送るスクリプトを使用した^[2]。
- ◇ 画像分析ソフトウェアを使えば、エンティティ、オブジェクト、属性の検出とラベリングを通じて、画像データに追加のメタデータを生成することもできる。

[2] See memespector script written by Bernhard Rieder, University of Amsterdam, at <https://github.com/bernorieder/memespector>

START

Choose one or more Facebook pages with memetic activity

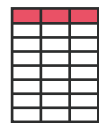


create a corpus of images



analyse corpus with

Google Vision API



Use optical character recognition (OCR) to extract the text contained within each image

computational linguistic analysis

Cortext

filter images according to computational linguistic analysis

visualise



a

WHAT THEMES LEND THEMSELVES TO MEMETIC ACTIVITY ON BREITBART'S FACEBOOK PAGE?

visualise



b

CAN WE DETECT DISTINCT VISUAL STYLES WITHIN A MEME REPOSITORY?

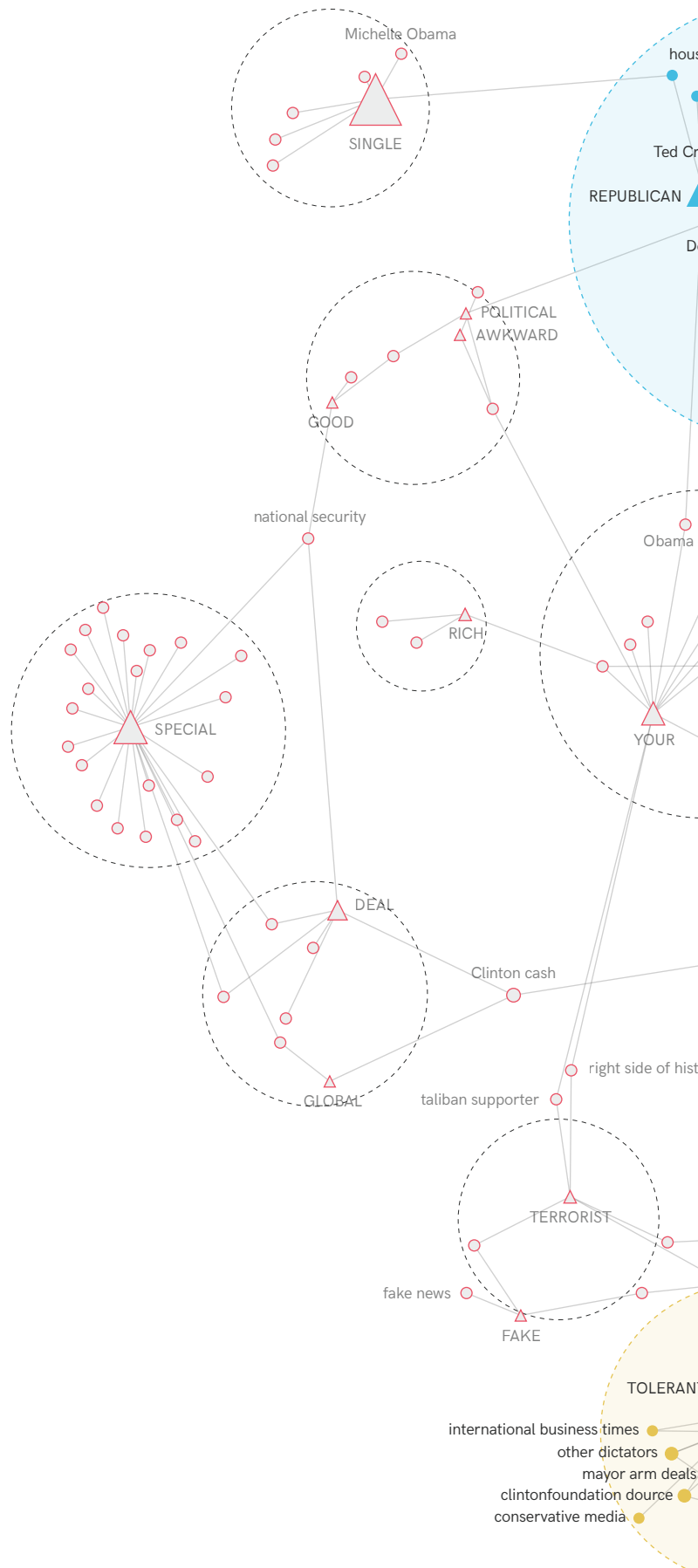
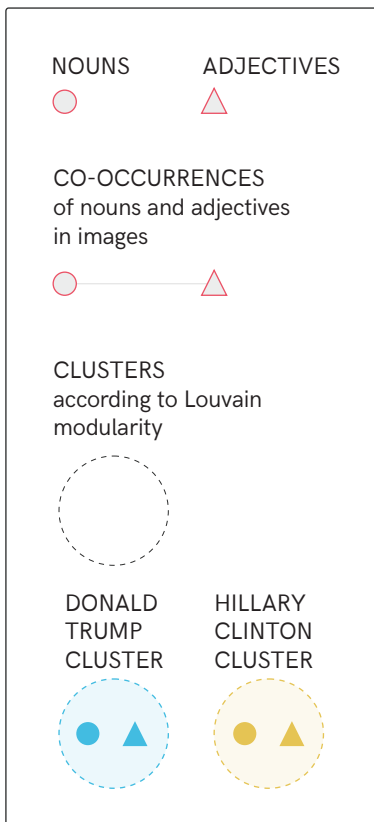
テキスト分析で、ミームが題材にしているテーマを調査する(実験的試み)

ミーム的活動のトリガーとなる話題を調査するには、手動の質的分析、もしくは半自動のセマンティック解析で画像からテキストを抽出し、分析する。この分析の結果は、OCRの質に左右される。数理言語学ツール(🔧 **CorText**など)が使えるが、抽出したテキストの関連性評価や、ツールのパラメータ設定には、分析者の判断が非常に重要だ(どのアルゴリズムを使用するか、どの単語を残すか、頻度はどうするか、など)。

- ◇ 字句解析を行えば、ミームに使われている最も関連性のある言葉の特定に役立つ。
- ◇ OCRアウトプットでクエリを実行することで、特定の話題がどう共鳴し合っているかを調査できる。
- ◇ ミームデータの用語共起を分析することで、主な話題と、話題がどう関連しているかを調査することができる。主なテーマクラスターや、それらの関係性を特定するのに、用語共起のネットワークを視覚的に検証するのもよい。また、主要な政治的話題が、ミームに使われている言葉の関連においてどう表現されているかも特定することができる。

ブライトバートの フェイスブックページで、 ミーム的活動を盛り上げる テーマは何か？

2016年にブライトバートのフェイスブックページに投稿された画像の中で、セットで使われた名詞と動詞のネットワークを示した図(二つの単語が同じ画像に使われていた場合、つながっている)。クラスターは、Louvain法に基づいて色分けしてある。最も目を引く二つのクラスターは、ドナルド・トランプ(上部、青色)とヒラリー・クリントン(下部、黄色)を中心に構成されている。ヒラリー・クリントンのクラスターと、その周辺クラスターに含まれる言葉を、構成とアジェンダ設定の観点から調査するのもよいだろう。



START

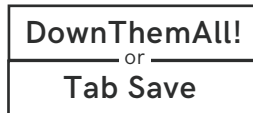
Choose one or more Facebook pages with memetic activity



get images URLs



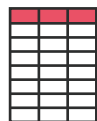
input images URLs in



create a corpus of images



analyse corpus with



Use optical character recognition (OCR) to extract the text contained within each image

computational linguistic analysis



filter images according to computational linguistic analysis



visualise



a
WHAT THEMES LEAD THEMSELVES TO MEMETIC ACTIVITY ON BREITBART'S FACEBOOK PAGE?


visualise



b
CAN WE DETECT DISTINCT VISUAL STYLES WITHIN A MEME REPOSITORY?

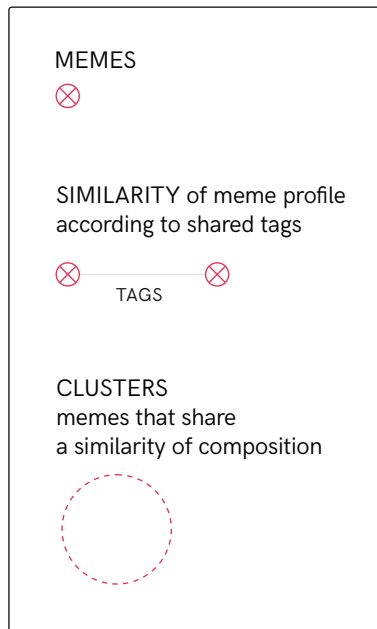
画像分析でミームのスタイルを調査する

あるミーム群に用いられているスタイルを調査するには、ステップ4.2で説明した画像分析ソフトウェアのアウトプット(画像データで検出されたエンティティおよび属性を記述するために生成した、ラベルまたはタグ)が使用できる。この分析を、God Emperor Trumpのページで説明する。

- ◇ グーグルVisionで画像を分析し、画像やテキストのコンテンツを記述するタグを抽出する。
- ◇  **CorText** を使い、共有ラベルを持つ画像同士の関連を調べる。
- ◇ ネットワークグラフの構成を視覚的に調べ、クラスターごとのスタイルを特定する。
- ◇ さまざまなページの形態を比べるには、この手順を繰り返す。

ミーム群内の、 特徴的なスタイルを 特定するには

共有タグに基づいた、God Emperor Trumpコーパスの画像ネットワーク。ミーム・プロフィールの類似性によりリンクされている。はっきりしない境界線は、ミームが似たような構成を持っていることを示しており、ミームの特徴と関連づけられる。ミーム活動のジャンルや形態のいくつかが特定できる。スクリーンショットのほか、漫画キャラクター「カエルのペペ」など、コミックや漫画を基にしたクラスターがあることが分かる。



Cluster A

- COMMUNITY ECOSYSTEM
- PEOPLE WATERCRAFT
- AUDIENCE ESTATE
- TEAM PALACE
- CROWD SELFIE
- YOUTH SOCIAL GROUP
- LAWN PROTEST
- LANDMARK

Cluster B

- SPEECH
- PERSON
- PROFESSION
- NEWS
- SOLDIER
- CAR
- SPEAKER
- VEHICLE
- AUTOMOTIVE DESIGN
- TROOP
- MODEL
- NEWSCASTER
- CONVERSATION
- AUDIENCE

Cluster C

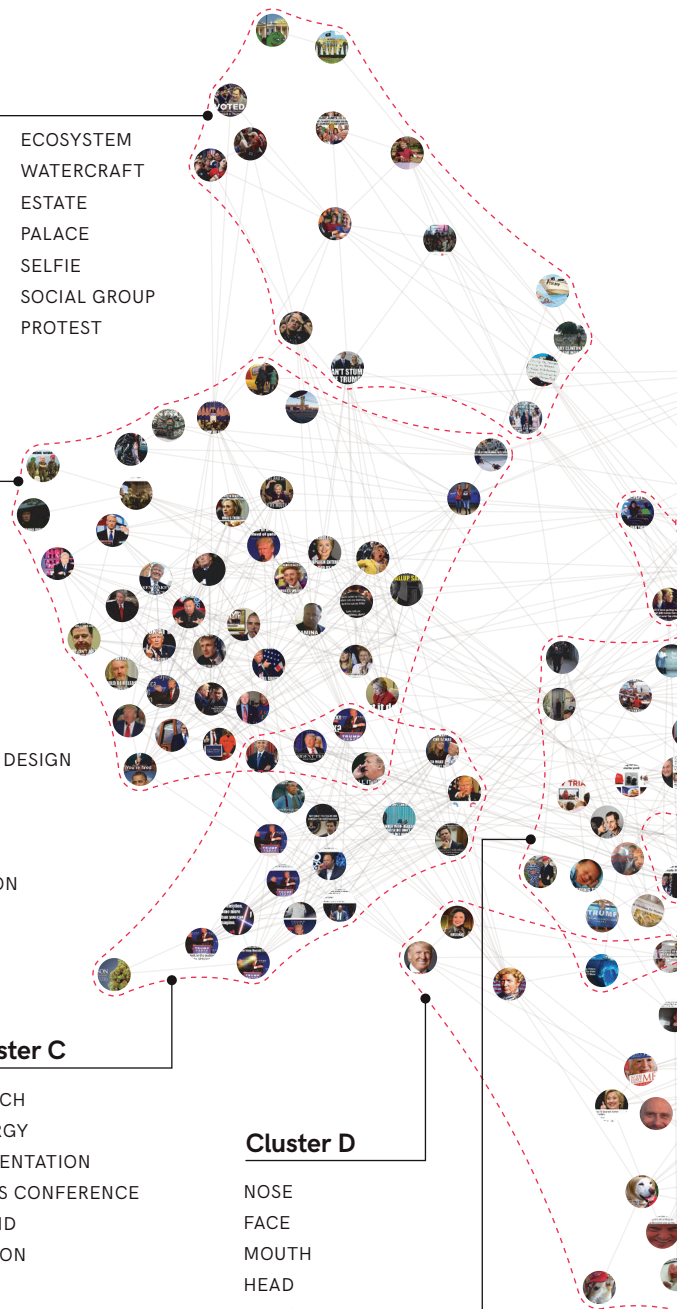
- SPEECH
- ENERGY
- PRESENTATION
- NEWS CONFERENCE
- BRAND
- PERSON

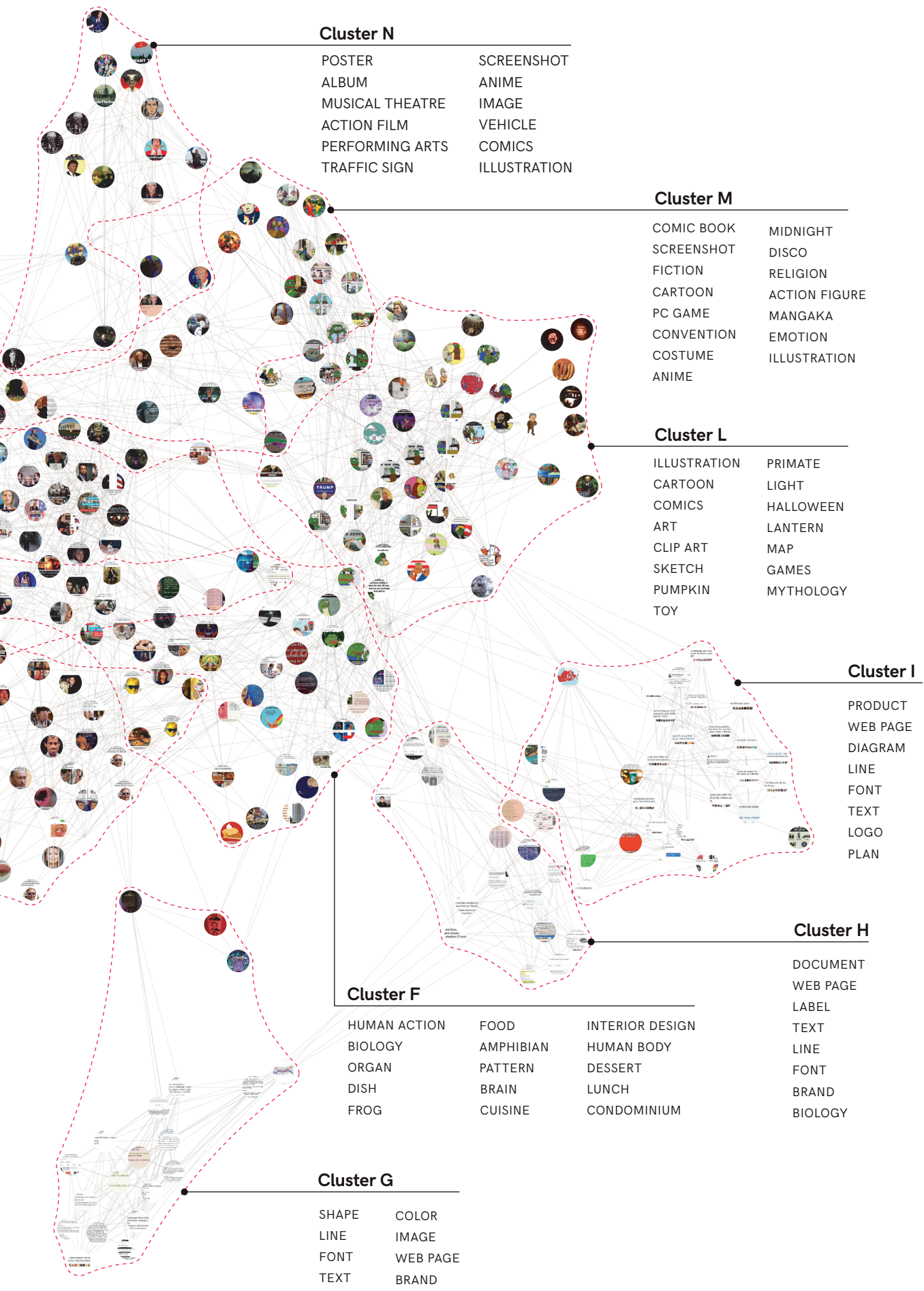
Cluster D

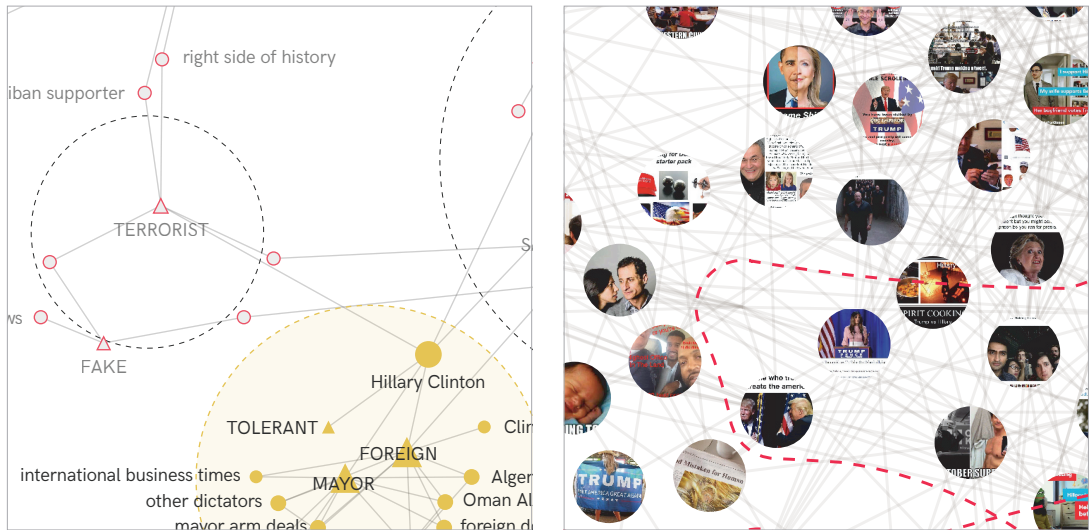
- NOSE
- FACE
- MOUTH
- HEAD
- HAIRSTYLE
- HAIR
- VISION CARE
- GLASSES
- FACIAL HAIR
- SKIN
- JAW
- FINGER
- FACIAL EXPRESSION
- DOG
- ORGAN

Cluster E

- PRESENTATION
- SENSE
- MAN
- BANNER
- CAP
- EDUCATION
- BRAND
- CONVERSATION
- SEA
- CHILD
- IMAGE
- PLAN
- INTERACTION
- EMOTION
- MUSCLE







おすすめ手法

このレシピは、政治的・社会的話題やイベント、人物が、「一般参加型プロパガンダ」^[3]とも言える形で、どんなミームになっているかの調査に使える。

[3] See, Alicia Wanless, "Have You Fallen Down a Participatory Propaganda Rabbit Hole?", Politics Means Politics, at: <https://politicsmeanspolitics.com/have-you-fallen-down-a-participatory-propaganda-rabbit-hole-6f71c83f-04fa>

第5章

ツイッター上の 荒らし行為を マッピングする

政治家を攻撃対象にする
ツイッターアカウントを
検出する方法

荒らしを行うユーザーの特性を
明らかにする

荒らし行為は
どのように特徴付けられるか

<第5章のポイント>

ツイッターのプロフィールや画像、フォロワーなどの情報を分析し、荒らし行為をするユーザーや偽アカウントの特徴を調べる。

はじめに - ネット上の荒らし行為(トローリング)やボット、「自作自演(ソックスパペット)」アカウントの使用といった戦術は、政治的な偽情報、プロパガンダ拡散と関連している[1]。

2017年のオランダ議会の総選挙に至るまでの間に、ジャーナリストたちは、いくつかの政党が自分たちのメッセージをオンライン上で増幅し、ソーシャルメディアで政敵を攻撃するために自作自演のアカウント(すなわち意見に影響を与えたり、だましたりする目的と思われるオンライン上の偽の身分)を使用したと指摘した[2]。

このレシピでは、2017年のオランダにおける選挙戦をケーススタディとし、荒らしやミスリーディングな情報を拡散する行為を分析するための手法を提供する。政治家に対する攻撃という狭い意味で、荒らし行為に焦点を当てる。

荒らし行為の3つの側面、つまり荒らしを行うユーザー、それら行為の特徴、そして攻撃対象を見る。

[1] See, for example, Alice Marwick and Rebecca Lewis, "Media Manipulation And Disinformation Online", Data Society, May 2017: <https://datasociety.net/output/media-manipulation-and-disinfo-online/>

[2] See, Andreas Kouwenhoven and Hugo Logtenberg, "Hoe Denk met 'trollen' politieke tegenstanders monddood probeert te maken", NRC, 2017. Available at: <https://www.nrc.nl/nieuws/2017/02/10/de-trollen-van-denk-6641045-a1545547>

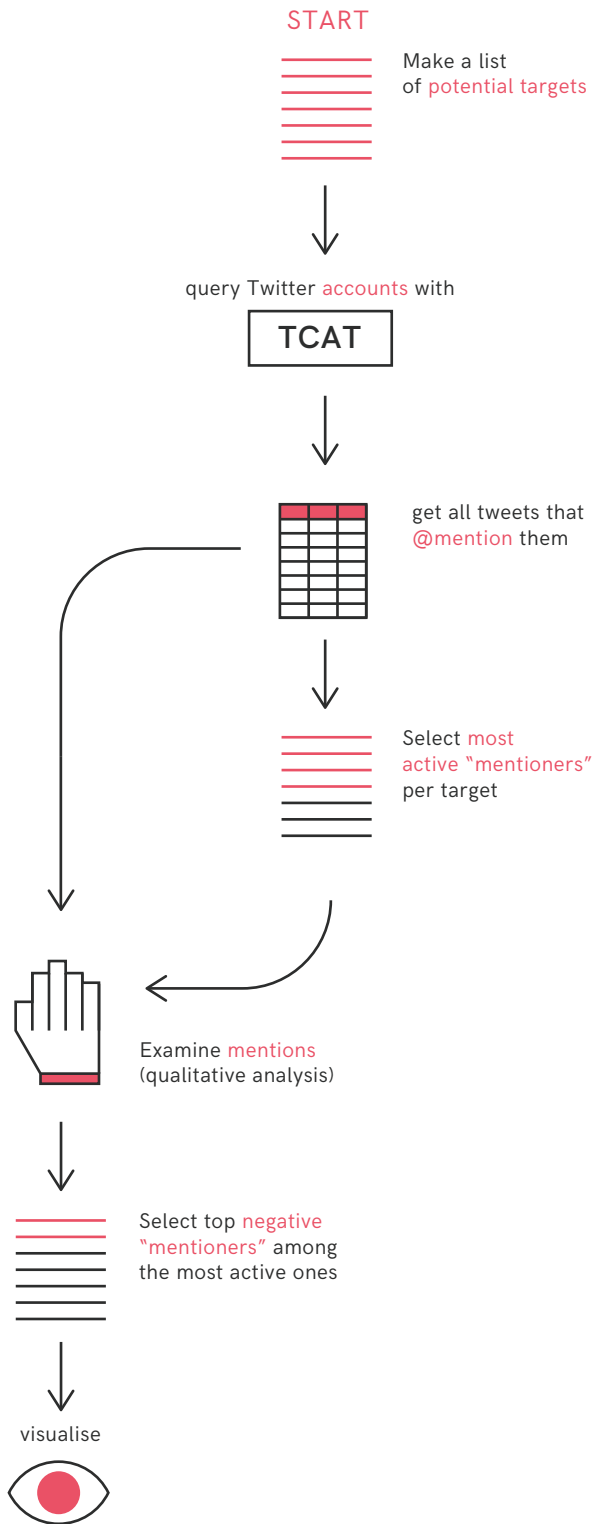


政治家を攻撃対象にする ツイッターアカウントを 検出する方法

始める前に

政治的な荒らし行為と、そうした行為を行うユーザーを発見するには、潜在的なターゲットのリストを作成することから始める必要がある。例えば、政治家に関連する一連のツイッターアカウントなどだ。

私たちは、2017年のオランダ総選挙に参加した政党のリーダーに関連する28のツイッターアカウントを使用した。



a

HOW ARE ATTACKS DISTRIBUTED ACROSS THE POLITICAL SPECTRUM IN THE NETHERLANDS?

ツイッターで政治家を攻撃対象にしているユーザーを特定する

ツイッターで政治的リーダーについて否定的な発言をしている人物を特定するには、政治家のアカウントのすべてのreferences(言及・引用)を特定することから始める。ツイッターの慣習に従って、“@”マークで定義された“言及”(@マークを含むツイート)で参照・引用を使用することができる。

- ◇ 一連のアカウントの@ユーザー名の発言をすべて集める。ここでは、🔪 **TCAT**というツイッターデータの抽出と分析ツールセットを使用した。
- ◇ 2017年オランダ総選挙が開かれた前月の2月8日から3月8日の間、28の政治的リーダーのツイッターアカウントの少なくとも一つに言及しているすべてのつぶやきを集めた。519,245件のツイートが収集され、この章のすべてのレシピで使用している。
- ◇ 最も活発な「発言者」を特定するため、対象となるアカウントの一つについて、一定基準以上に言及しているすべてのアカウントを探す。ここでは、私たちのデータセットの中の政治家1人について100回以上の発言をしたアカウントを対象に残した。
- ◇ 前のステップで特定された最も活発な“発言者”による言及内容を検証して、それぞれの発言の性質を判断する(すなわち、その政治家を支持しているのか、それとも攻撃しているのか)。
- ◇ 一人または複数の政治家を常に否定的に攻撃するユーザーだけを残す(ここでは、ユーザーリストを25人に絞り込んだ)。

オランダの 政治家に対する攻撃は どう行われているか？

ツイッターでの肯定的または否定的な発言の標的となったオランダの政治リーダーの視覚化(一カ月間に100回以上言及したユーザーからのもの)。赤い丸は攻撃的な発言をしたユーザーを表し、緑色の丸は好意的な発言をしたユーザーを表す。円のサイズは、政党のリーダーに言及しているユーザーの総数を表す。政治家に対する荒らし行為の分布を見ると、その非対称さが目立つ。否定的な発言の対象となっているのは、ほとんどの場合、左派の政治家だ。

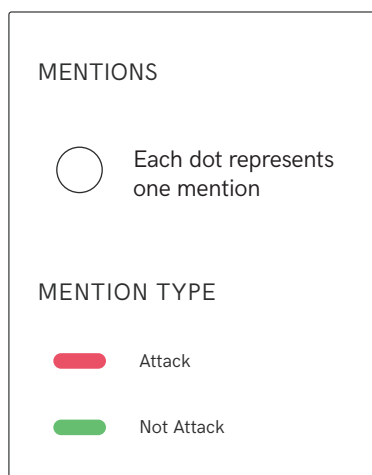
@sybrandbuma (CDA)



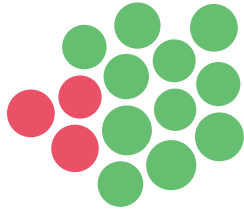
@thierrybaudet (Forum voor Democratie)



@JacquesMonasch (Nieuwe Wegen)



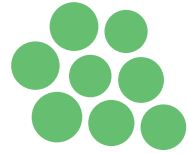
@MinPres (VVD)



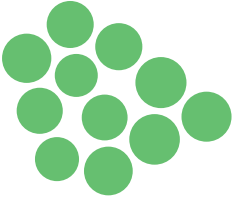
@markrutte (VVD)



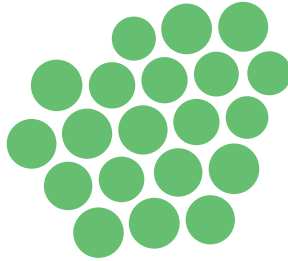
@mariannethieme (PvdD)



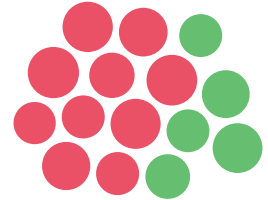
@ncilla (Pirate Party)



@geertwilderspvv (PVV)



@LodewijkA(PvdA)



Gertjansegers(CU)



@HenkKrol (50plus)



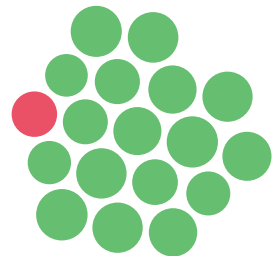
@emileroemer (SP)



@APechtold (D66)



@LavieJanRoos (VNL)



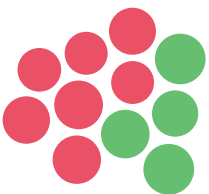
@Tunahankuzu (DENK)



@keesvdstaaij (SGP)



@jesseklaver (GroenLinks)

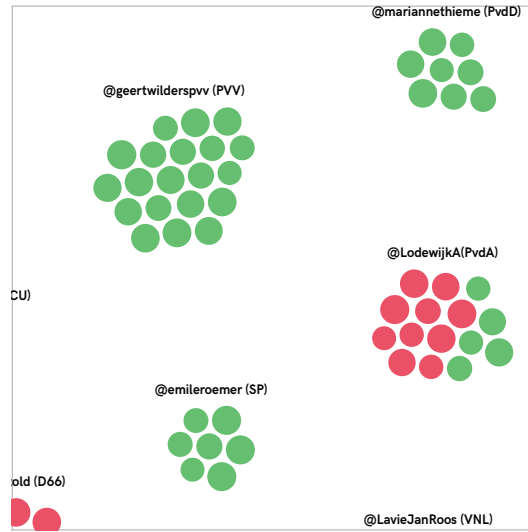


@jndkgrf (GeenPeil)



@KleinNorbert (VZP)





おすすめ手法

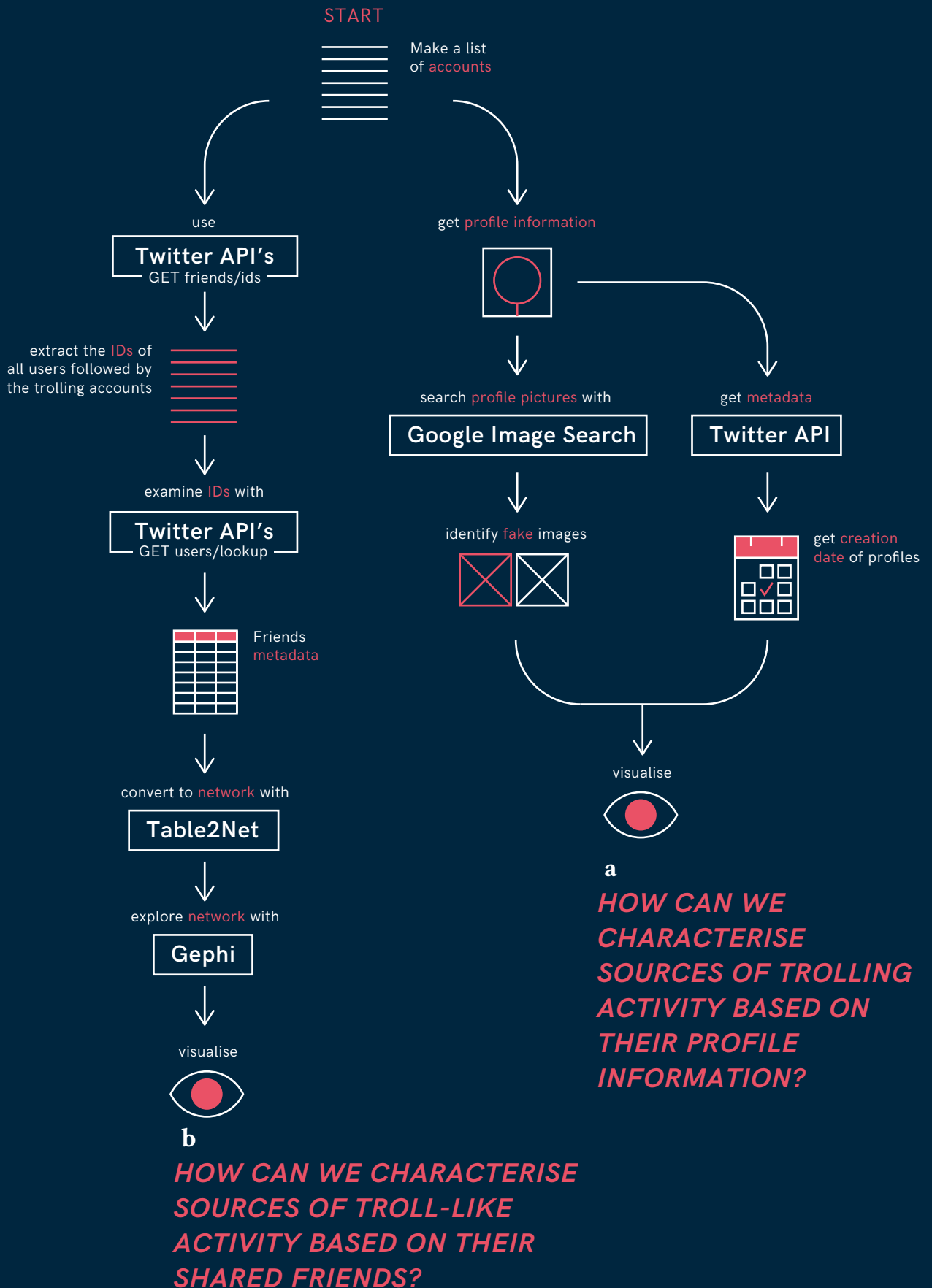
このレシピは、ツイッターで個人攻撃をするアカウントの特定に使用することができ、政治的な荒らし行為の文脈を超えて活用できる。



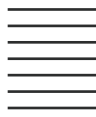
荒らしを行うユーザーの 特性を明らかにする

始める前に

このレシピでは、前のレシピで特定された、少なくとも1人の政治家に100回以上言及した25のアカウントを出発点とする(アクティブではない一つのアカウントは除いた)。



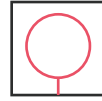
START



Make a list of accounts



get profile information



search profile pictures with

Google Image Search

get metadata with

Twitter API

identify fake images



get creation date of profiles



visualise



a

HOW CAN WE CHARACTERISE SOURCES OF TROLLING ACTIVITY BASED ON THEIR PROFILE INFORMATION?



use

Twitter API's
GET friends/ids

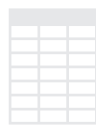
extract the IDs of all users followed by the trolling accounts



examine IDs with

Twitter API's
GET users/lookup

retrieve friends metadata



convert to network with

Table2Net

explore network with

Gephi

visualise




b

HOW CAN WE CHARACTERISE SOURCES OF TROLL-LIKE ACTIVITY BASED ON THEIR SHARED FRIENDS?

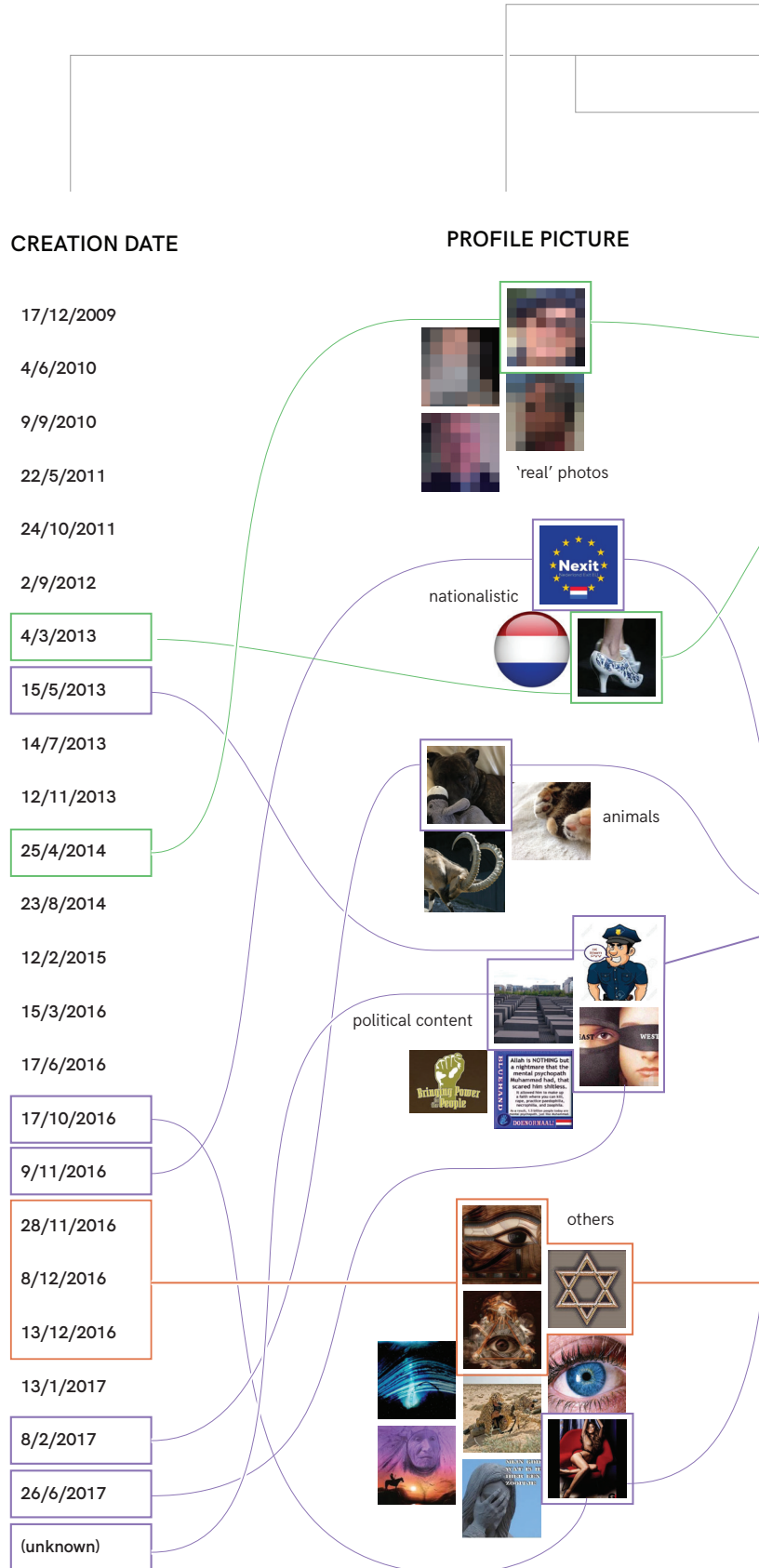
a. プロフィール情報から、荒らし行為を行うユーザーの特徴をつかむ

荒らし行為を行うユーザーの特徴をつかむ一つの方法は、プロフィール情報を調べることだ。

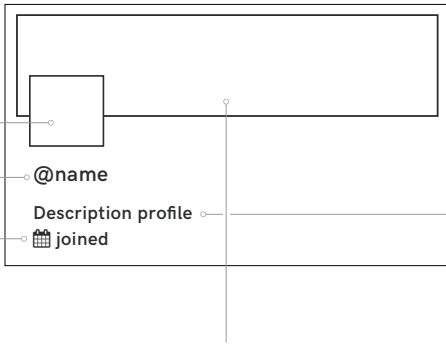
- ◇ 各アカウントにアクセスして、ツイッターのユーザー情報を収集する(プロフィール説明、プロフィール画像、バナー画像)。
- ◇ この情報を分析して、政治的関心、言及されたハッシュタグ、およびつながりを特定する。
- ◇ プロフィール画像の有無を記録し、識別された画像を  **グーグル画像検索** にアップロードして、いずれかのアカウントが偽のプロフィール画像を使用しているかどうかを検出する。
- ◇ ツイッター APIを使用して、アカウントの作成日を抽出し、複数のアカウントが同じ日付に作成されているかどうかを調べる。

プロフィール情報に基づいた、 荒らし行為を行う ユーザーの特徴

オランダ総選挙で荒らし行為を行った24のアカウントのクラスター分析。プロフィール情報は、類似性に従ってクラスター分析された。3人のユーザーが非常によく似たプロフィールを持ち、短い時間で作成された(これは、荒らし行為のために作成された「ソックスパペット」アカウントの識別に役立つ)。他の6人は同じ反イスラムアジェンダを宣伝しているが、偽のアカウントではない。



ELEMENTS FOR CHARACTERIZATION ANALYSIS



BANNER

BIO DESCRIPTION

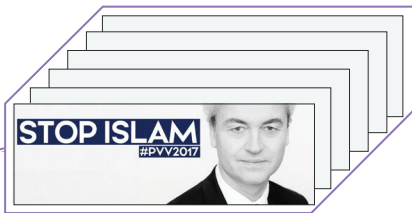
TROLL OR NOT TROLL?



- @D***er
BOER, Economist, Father of six, Doctor/MBA Patriotically correct. Romantically conservative, islamophobe, aspirant lid PVV, lid FvD
- @M***te
(no bio description)

?

✗



- @H***an
(no bio description)
- @l***dy
#de-islamize #stoppolice #stopterror#voteMarine #voteWilders#minderminder #lessless#maroccons
- @M***00
Ben fascist noch racist,maar realist.Links is het nieuwe rechts en blank het nieuwe zwart.Stop Brussels EU omvolking,Stop massa migratie,Stop Rutte & Timmermans
- @A***rt
Een kern van (mijn) waarheid overgoten met sarcasme, sadisme en galgenhumor. Neem niets persoonlijk op, maar neem het wel op!
- @g***91
Twitter Censors Everything, their Board members are Izlamic. MOVE to Gab.ai for true FREE SPEECH FOR ALL <http://gab.ai/ger2519> #PVV #MEGA #MHGA #BanIslam
- @Y***NL
(account suspended)

?

?

?

?

?

?

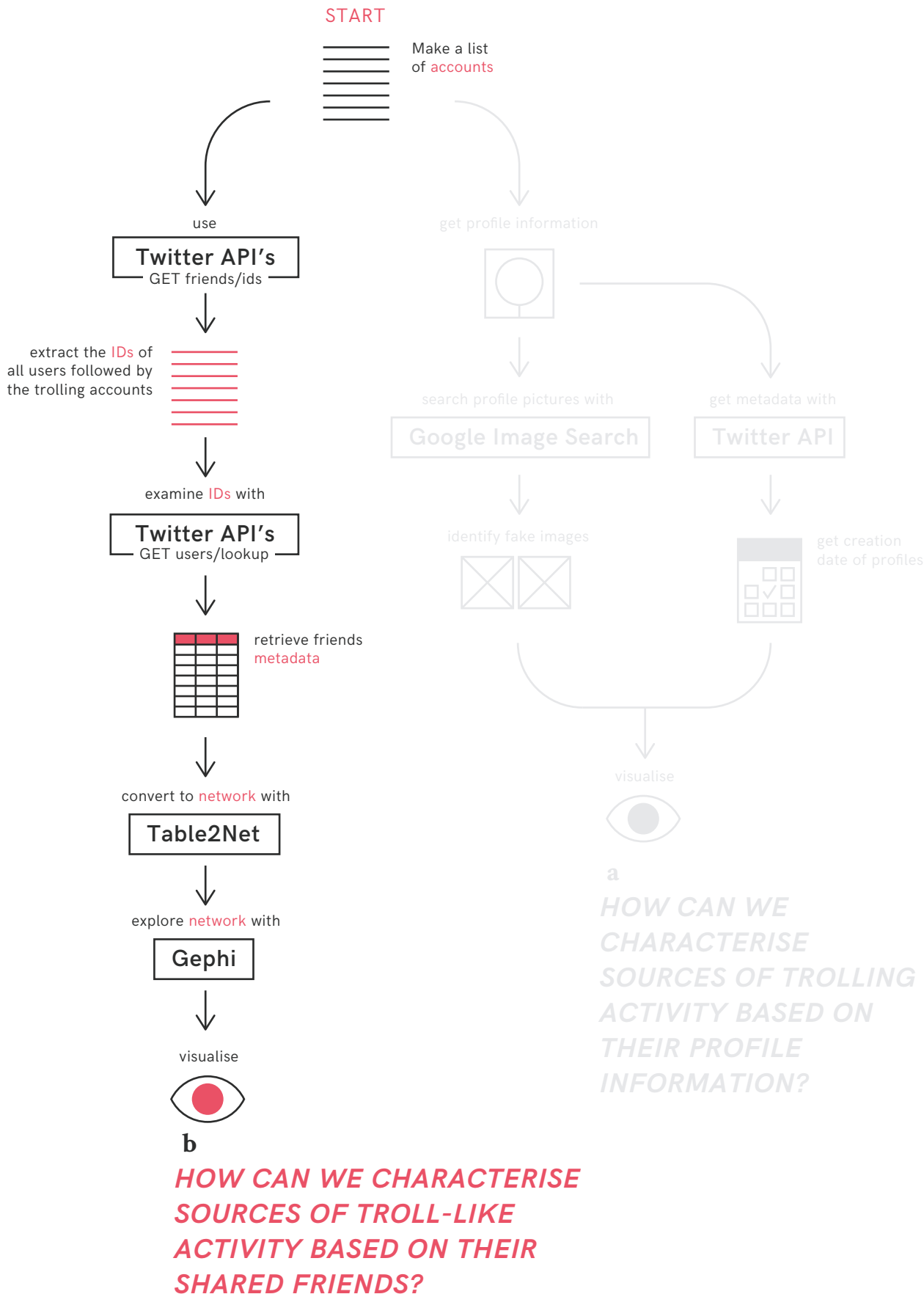


- @j***33
PVV Geert Wilders Next, Trump Trooper MAGA, Love Le Pen Vive la France MEGA, Anti Islam, De-islamization Forever
- @j***33
Geert Wilders PVV MEGA Next De-Islamization worldwide, Love Le Pen Superwoman, Trump MAGA, Close borders, Anti islam, Anti EU, Jail Obama, Jail Hillary...
- @k***33
Geert Wilders Next, Vive Le Pen, Trump MAGA, AfD Germany, Brexit, Pro Israël, Anti EU, De-islamization...

✓

✓

✓



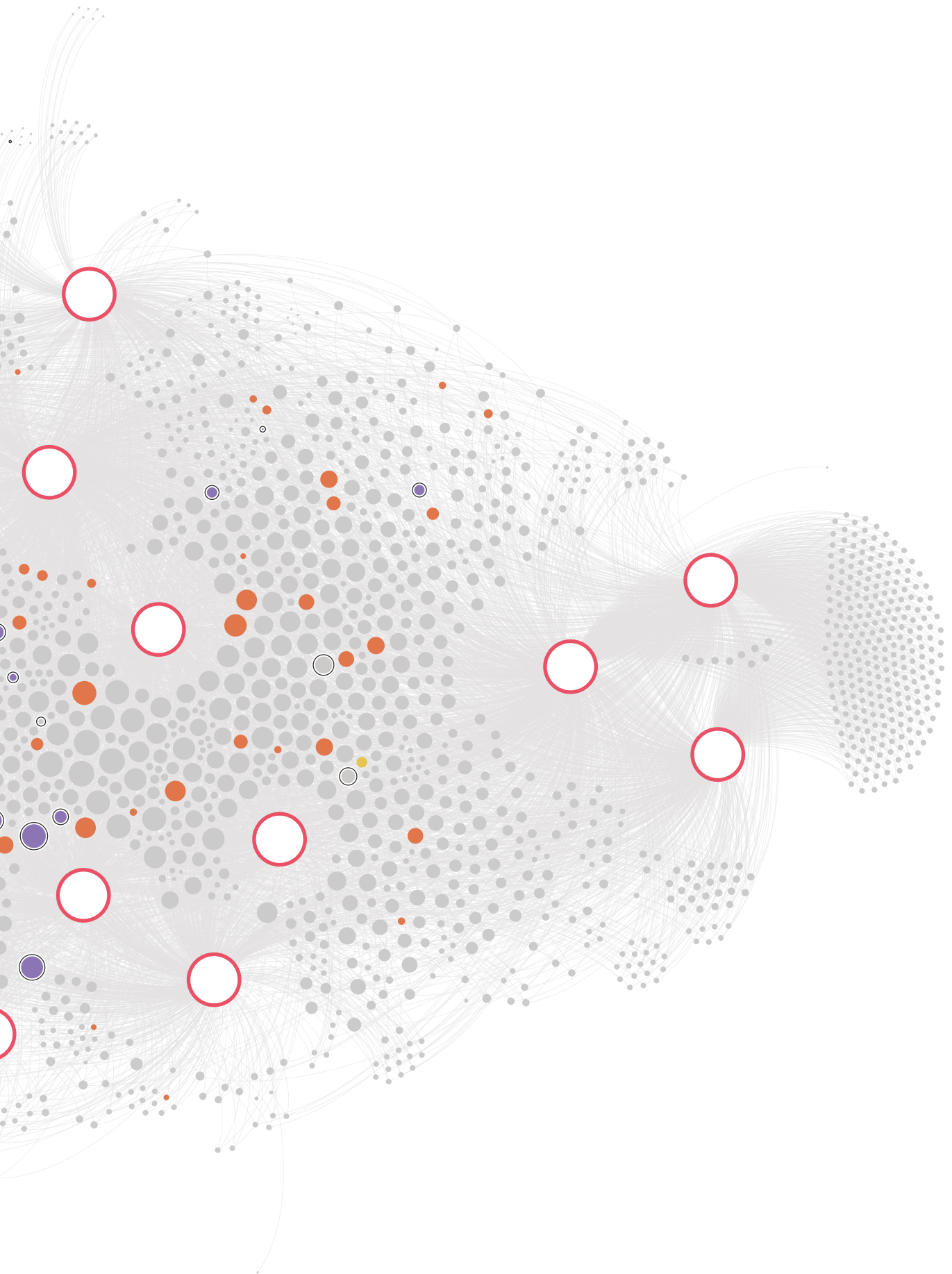
b. ツイッター上の友人を通して見られる、荒らし行為を行うユーザーの特徴

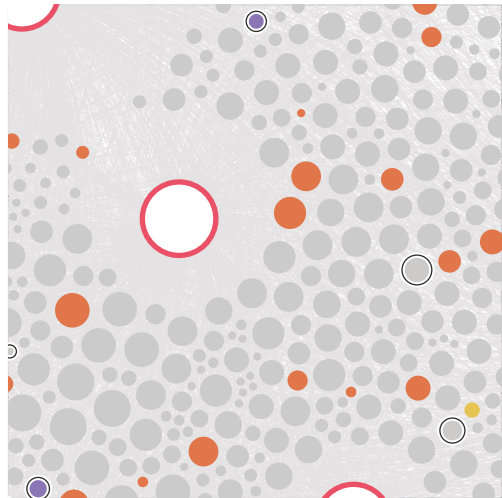
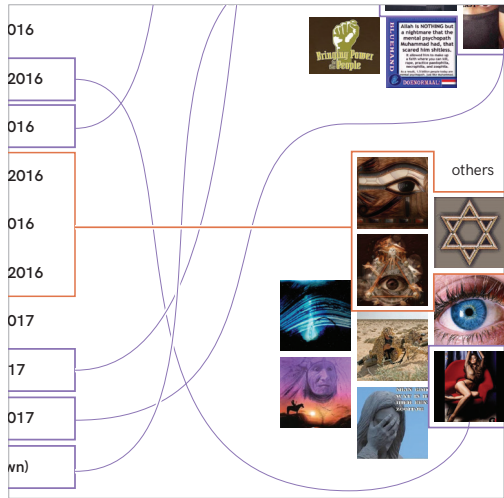
荒らし行為を行うユーザーの特徴を分析する別の方法は、ツイッターで友人として誰をフォローしているかを調べることだ。

- ◇ ツイッターAPIの "GET friends / ids"機能を使用して、荒らし行為を行うアカウントにフォローされているすべてのユーザーのIDを抽出する。^[1]
- ◇ ツイッター APIの "GET users/lookup"機能を使用して、各友達のプロフィール情報を取得する。^[2]
- ◇ 🔧 **Table2Net** を使用して、調査したアカウントとその友人を含む一覧表(テーブル)を、彼らの「友情」関係によって結ばれたツイッターアカウントのネットワークへと変換する。
- ◇ 🔧 **Gephi** などのネットワーク分析と視覚化ツールを使用して、荒らしを行うアカウント間で共有されている友人や、フォローしているアカウントを探ることができる。

[1] See Twitter Developer Platform, API reference: GET friends/ids, 2017, at: <https://developer.twitter.com/en/docs/accounts-and-users/follow-search-get-users/api-reference/get-friends-ids>

[2] See Twitter Developer Platform, API reference: GET users/lookup, 2017, at: <https://developer.twitter.com/en/docs/accounts-and-users/follow-search-get-users/api-reference/get-users-lookup>





おすすめ手法

このレシピは、他の政治キャンペーンに関する荒らしアカウントの分析にも使用できる。



荒らし行為は どのように特徴付けられるか

始める前に

このレシピは、一つ前のレシピで特定されたアカウントと、調査期間中にそれらのアカウントから投稿されたツイートを出発点にしている(レシピ5.1参照)。

START



make a list of **accounts** engaging in negative targeting



capture tweets posted by these **accounts** with

TCAT



extract **hashtags** by their frequency



examine **hashtags** (qualitative analysis)



a

WHAT ISSUES ARE PRESENT IN TWEETS @MENTIONING A POLITICAL CANDIDATE?



extract **hosts** by their frequency



examine **hosts** (qualitative analysis)



visualise



b

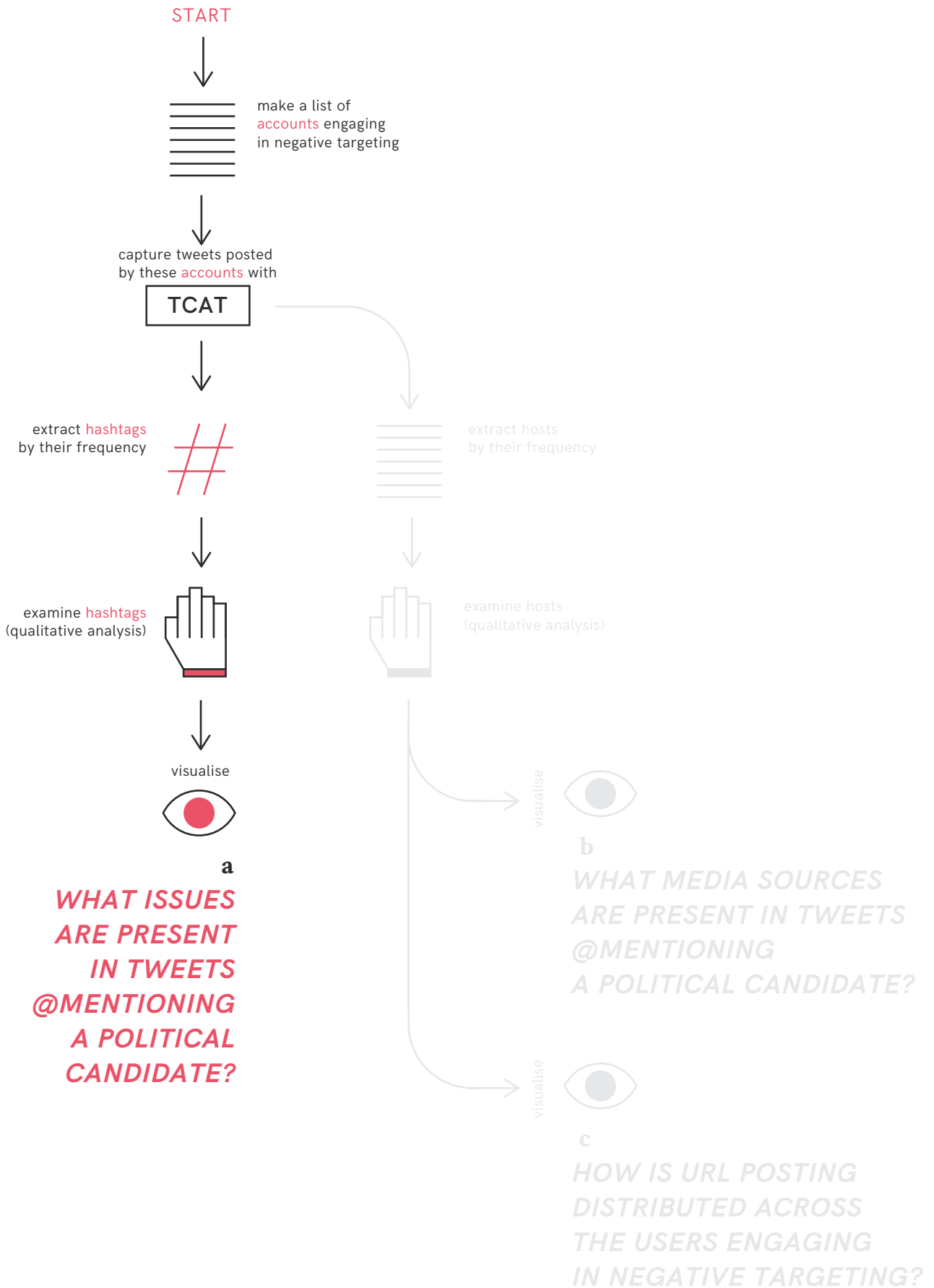
WHAT MEDIA SOURCES ARE PRESENT IN TWEETS @MENTIONING A POLITICAL CANDIDATE?

visualise




c

HOW IS URL POSTING DISTRIBUTED ACROSS THE USERS ENGAGING IN NEGATIVE TARGETING?



荒らし行為を行うユーザーが使用するハッシュタグを調べる

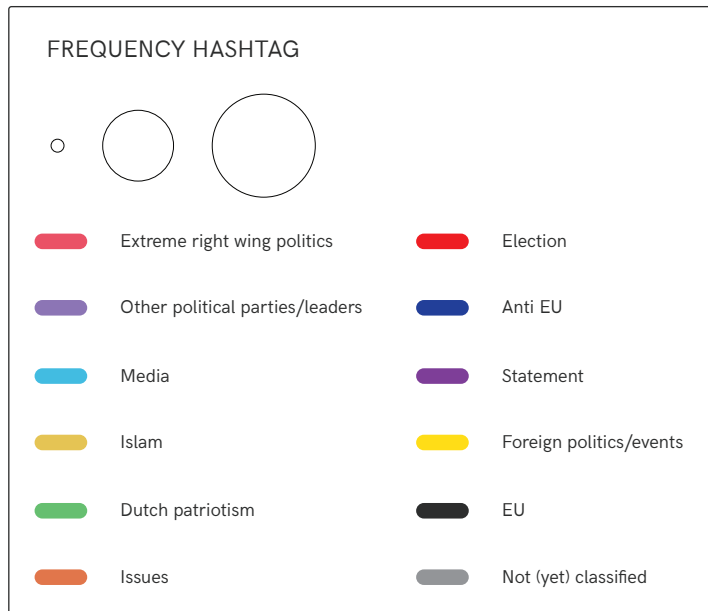
荒らし行為に関連する話題を特定するには、頻繁に攻撃を行うユーザーが投稿した、政治家に言及するつぶやきで使用されるハッシュタグを調べる。

- ◇  TCATの「hashtag frequency」機能を使用し、ハッシュタグの頻度をランク付けする。
- ◇ 最も頻繁に使用されるハッシュタグを手動で分析して、荒らし行為に関わるユーザーの活動が盛り上がっている話題を特定する。

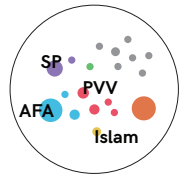
選挙候補者に @で言及するツイートは 何を話題にしているのか？

2017年のオランダ総選挙で、荒らし行為に関わった24のアカウントが、@候補者で言及したツイート投稿の中で、ハッシュタグ付きで触れられた話題のバブルグラフ。論点は種類別に色分けされ、発生頻度によってサイズが決められ、ツイートに含まれる候補者別に依じてグループ化されている。ハッシュタグ付きのほとんどのつぶやきは、右翼のポピュリストのGeert Wilders候補に言及している。最も顕著なのは、PVVの政治的メッセージ(「Nexit」訳注: 欧州連合からの英国脱退「Brexit」のオランダ版、「Stoplislam」、「Banlislam」)ならびにオランダの愛国心の表現に属するものだった。

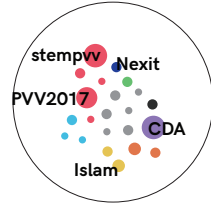
概して言えば、右派の政治家は主に「荒らし行為をするユーザー」からの支持を受け、他の政治家は攻撃の標的であると結論づけることができる(レシピ5.1でも説明)。



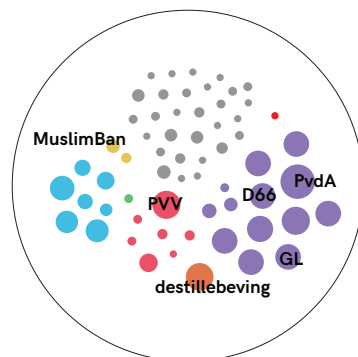
@emileroemer



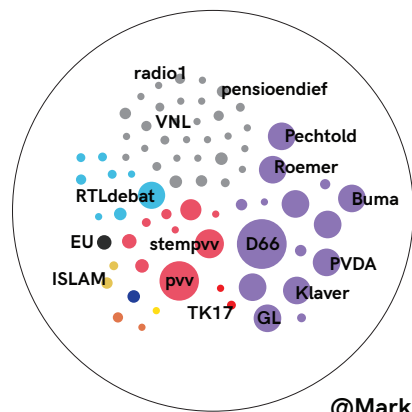
@sybrandbuma



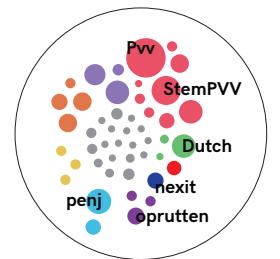
@LodewijkA

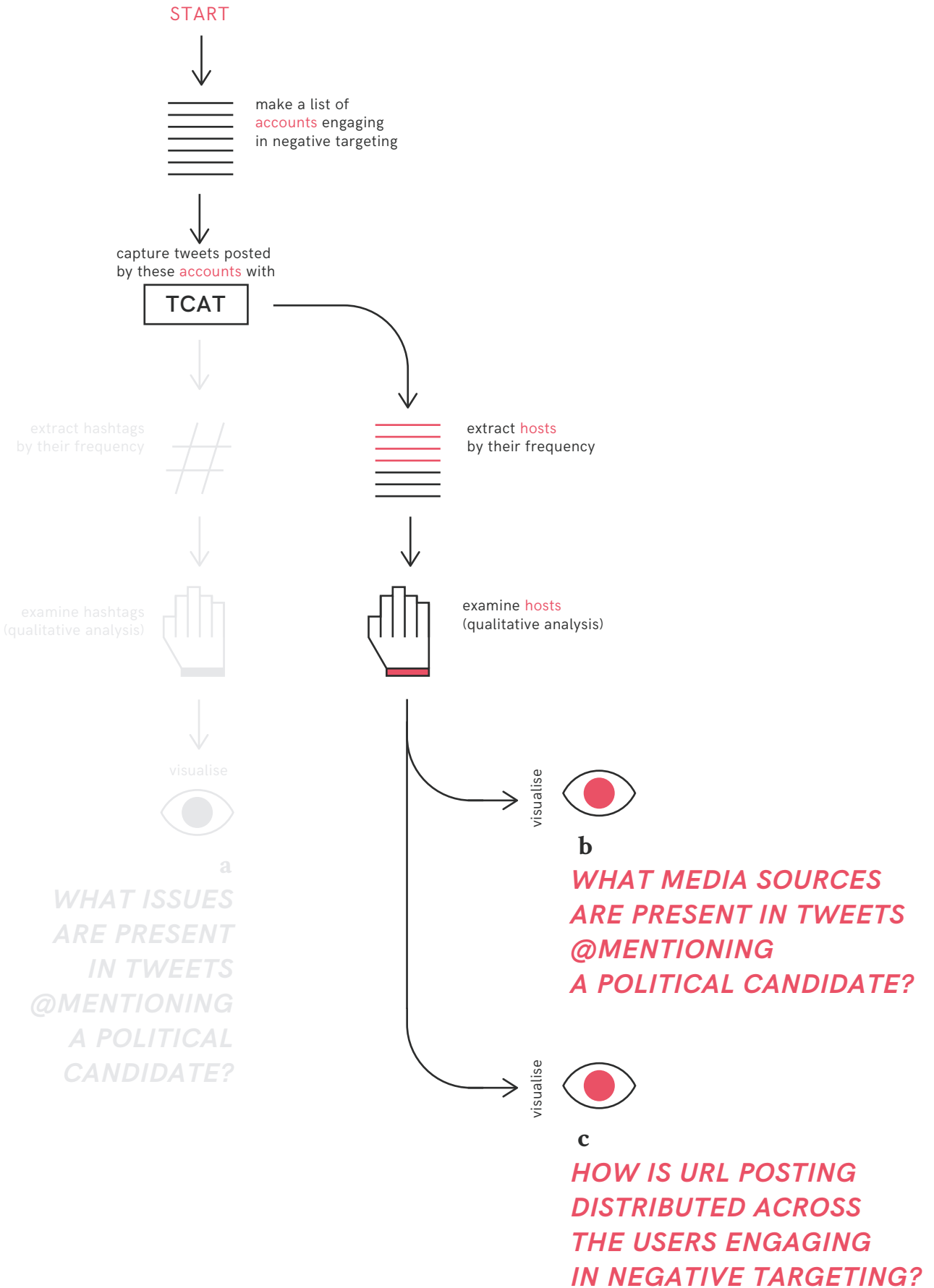


@APechtold




@MarkRutte





攻撃を行うアカウントで共有されているメディアソースを調べる

荒らし行為をするユーザーが投稿したツイートで共有されたコンテンツを特定するため、そのツイートの挿入されたURLを調べる。

- ◇  **TCAT** の "hosts frequency" 機能を使用して、出現頻度でランク付けされたメディアソースを抽出する。
- ◇ 最も頻繁に使用されるメディアソースを手動で分析して、プロフィールを決定する。
- ◇ ユーザー全体のURLをシェアする行動を分析することは、荒らし行為を発見する手段として使用され得る。

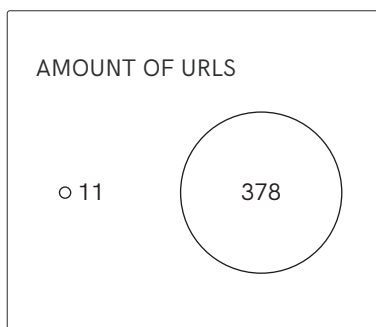
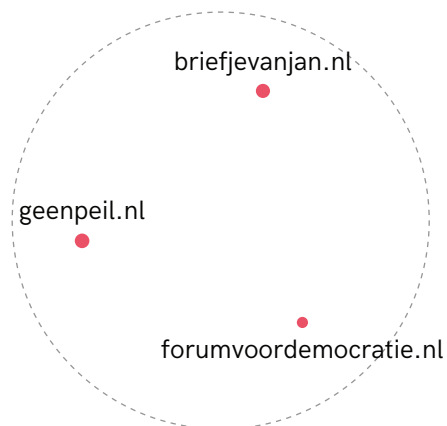
政治の候補者に @で言及しているツイートには、 どのようなメディアソースが あるか？

2017年のオランダ総選挙で、荒らし行為に参加した24のアカウントによって投稿された候補者に言及したツイートの中で、最も反響があったメディアソースの図。もっともツイートされた情報源はオランダのオルタナ右派ブログfenixx.orgで、反イスラムサイトの「Jihad Watch」と右翼のシンクタンク「Gatestone Institute」にフォローされている。

OTHER MEDIA



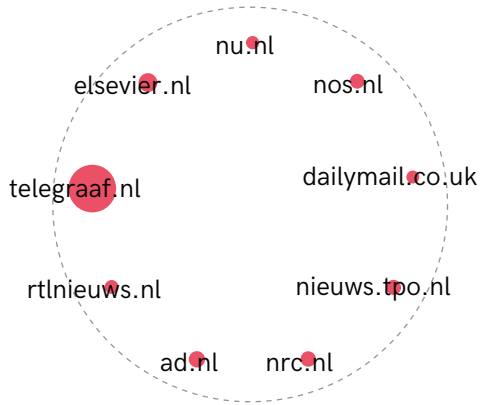
POLITICAL PARTY



PUBLIC BROADCASTER

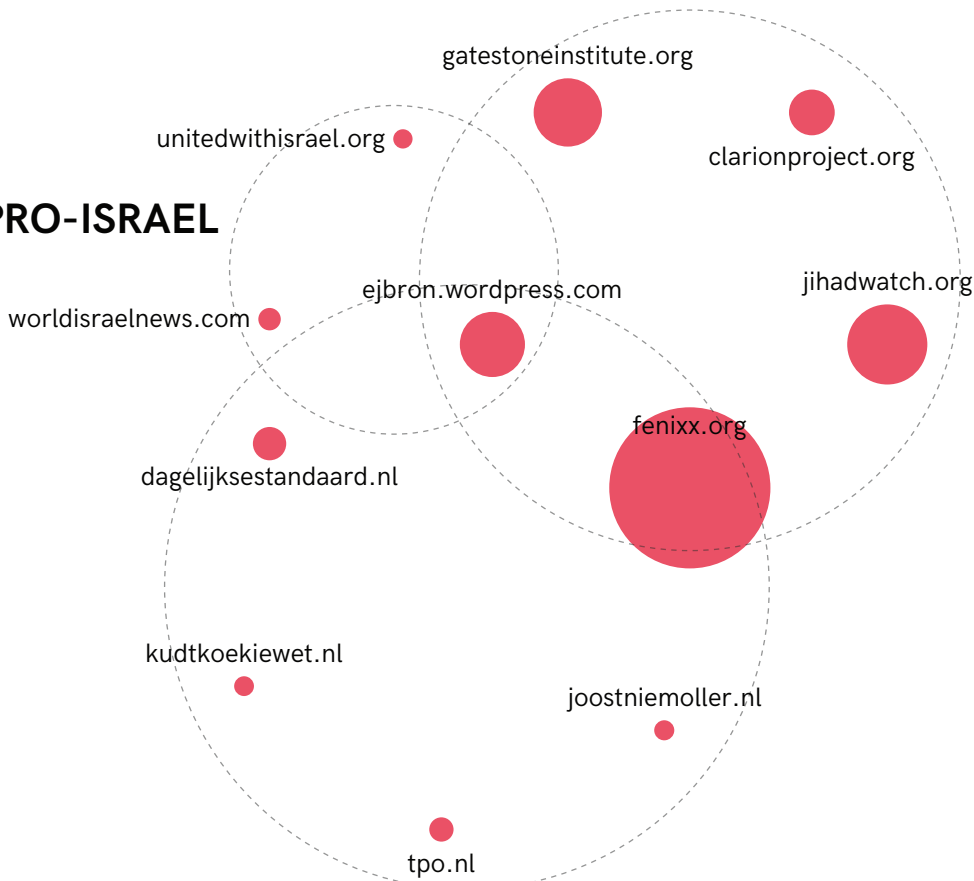


NEWS



ANTI-ISLAM

PRO-ISRAEL



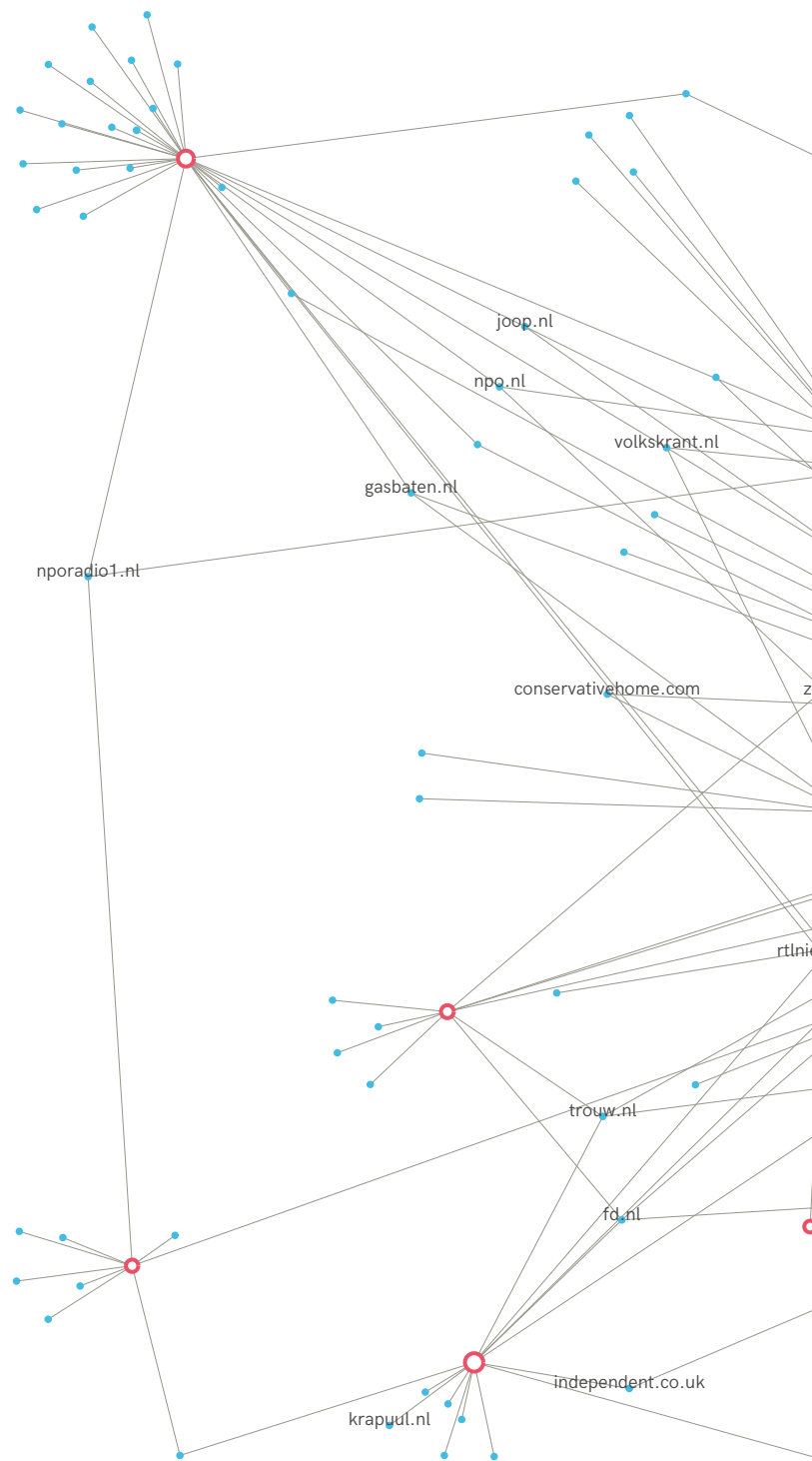
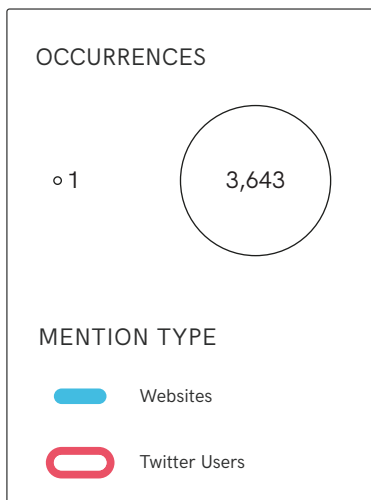
DUTCH-RIGHT

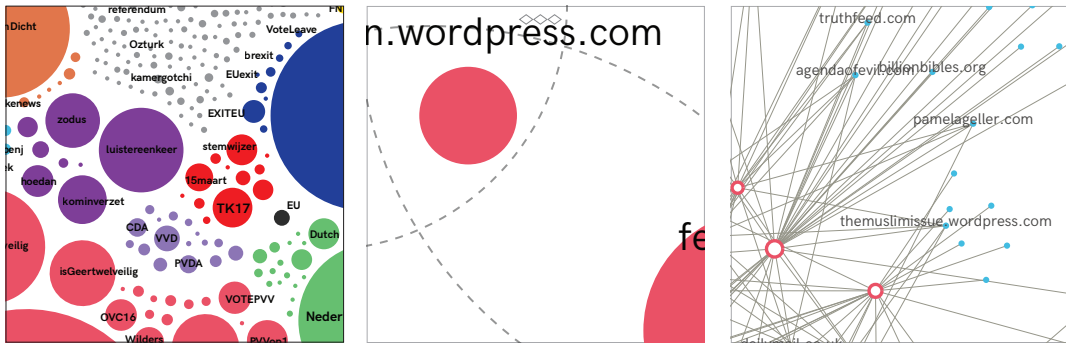
ALT-RIGHT



攻撃に関わるユーザー間で 同じURLが 投稿されているか

オランダ総選挙の前月に政治家を攻撃した24人のユーザー全体で共有されたURL分布のネットワークグラフ。調査期間中にURLの大半が2人のユーザーによって投稿されたことを示している。





おすすめ手法

このレシピは、荒らし行為に携わるソーシャルメディアアカウントに関連した話題やメディアソースを分析するために使用でき、荒らし行為やその状況についてより深く理解することができる。

もし、個人のアカウントを調べる場合は、潜在的な法的、倫理的影響を考慮し、公共の利益とプライバシーへの配慮をはかりに掛けるべきだ。複数のアカウント間のネットワークや関係に注目するかどうか、個人識別情報の削除や修正を行うかどうかも検討が必要だ。

おわりに

フェイクニュースは、避けては通れない問題だ。毎日のように新しい新聞記事、ブログ投稿、調査報告書やプロジェクトが発表され、公共の議論が巻き起こり、多くの研究が行われている。プロジェクトを通じて私たちは、この問題を理解し、調査や研究を進めようとする世界中のメディア組織やジャーナリスト、市民団体、公的機関、企業、研究者、学生と協力してきた。

激しい公共の討論とメディア化された状況の中で、「フェイクニュース」という言葉を使うべきではないという提案もされてきた。「はじめに」で述べたように、とりわけ、フェイクニュースという言葉はあいまいであり、政治的に危険であり（さまざまな政党の戦術的な言葉として用いられる）、従来のプロパガンダの形や偽情報、誤情報と区別がつかないと言われる。このガイドでは、「フェイクニュース」という言葉を使用しているが、別の手段でここに挙げた懸念への対処を試みている。

お気づきのように、このガイドは、コンテンツに基づいてフェイクニュースを定義したり、特定したりすることに焦点を当てていない。そうした関心は確かに重要だが、関連する現象の領域を分類したり区別したりする試みは、さまざまな事例の特徴だけでなく、そのソーシャルメディア上の営みについての実証的調査であるべきだ。そのような取り組

みが、フェイクニュースに関連する事象を説明するために、精度を高めた分析的で、概念的で、理論的な語彙の発展に役立つことを願っている。

フェイクニュースに対して何をすべきか? 「はじめに」で触れたように、このフィールドガイドで達成したいことが一つあるとすれば、ソーシャルメディア上のニュースの動きや、その周辺のデジタル環境にもっと本質的な焦点を当て、公共の議論やジャーナリズム、研究の重点の幅を広げることだ。この取り組みが、問題を研究して理解を深めるだけでなく、変化を起こすことができる人々を奮い立たせ、動かし、協力体制を作るという壮大な課題に、少しでも貢献できることを願っている。

2016年の米大統領選の際にこのプロジェクトを開始してから1年後、偽ニュース問題は「グローバル化」した。登場人物の配役は多様化し、初期の物語に登場した党派心の強い草の根の宣伝者や、ご都合主義のマケドニアの若者、米国をターゲットとするロシアの政治的な運営者たちから、より多様にグーグルやフェイスブック、欧州委員会、中国共産党、イタリアの「五つ星運動」、英国の諜報機関「MI6」、ウィキペディア、ウェブの発明家ティム・バーナーズ＝リー、選挙ボット、メッセージングアプリ、核の脅威、テック系のスタートアップ、セキュリティ会社や世界各国の「ダーク」マナーなどが登場するようになった。

そして「フェイクニュースについて何をすべきか?」という疑問は、ただ単純にオンライン上のミスリーディングな情報に関してではなく、オンライン・プラットフォームや広範なデジタル文化、関連するテクノロジーと習慣などに連なる問題として広がりを持ってきた。攻撃的な記事や常軌を逸したユーザーを特定し、淘汰するための問題として始まったことが、オンライン上の生活全般に関するより大きな疑問や議論へと広がっている。情報や知識、文化を創造し、点検し、共有し、活用し、そして意義づける土台としてのインフラや制度もまた問われているのだ。

企業、市場、国家、政治家、政策立案者、メディア組織、教育機関、市民団体などにその責任(と批判)をどう配分すべきかについて、さまざまな方策や概念があるのはもちろんだが、すべての論争と同じように、問題を診断して定義し、すくい上げるには多様な手段がある。そして、多くの危機と同じように、この問題に関わる人々は、異なる課題を持ちながら、「良い危機を無駄にはしない」という共通の考えも共有している。

フェイクニュースの対応は急がれるが、このガイドで探究するアプローチにより、読者が拙速に陥ることなく状況を評価し、問題を診断し、対策を提案していくことを願う。このガイドの一連のレシピを通して、この現象にどのように時間を費やすかについて、いくつかの指針を提供したいと考えている。特に、オンライン・プラットフォームとウェブの媒介能力と文化という広範な関心事の一つの側面について、このガイドが研究、調査、公の議論に情報提供し、サポートしていくことを願っている。提供したレシピ一式は、オンライン上の偽ニュースや他の偽造行為に重点を置いているが、レシピの多くは、知識をめぐる政治や論争、その舞台となるオンライン空間とデジタルインフラについて、さまざまな角度から調べるために簡単に応用できる。

知識政治を研究するためデジタル記録を再利用する際には、主張の中身を単独で考察、評価するのではなく、それを取り巻くネットワークにこそ目を向けるべきだ。したがって、フェイクニュースコンテンツを個別に研究するのではなく(その拡散という文脈は別だが)、ネットワーク化されて広がるというその性質、そして周りがそれを意味づけするという社会的、文化的な慣習、またフェイクニュース拡散を「下支え」するメディア・システムに注目すべきだ。言い換えれば、調査者はフェイクニュースと分類されるアイテムについて、その真偽や、世界で起きていることをどの程度正確に伝えているかということだけではなく、どのように、誰にシェアされているか、何に依拠しているか、そして受け止める人の立場でニュースの意味合いや価値が異なることについても考慮してほしい。

なぜ、そのような転換が必要なのか。第一に、デジタルコン

コンテンツが社会的、文化的に意味づけされる過程をより深く理解できれば、また違った問いの立て方ができる可能性があるからだ。グループ内で(ニュースを)シェアする理由は、文字通り真実であると考えているから?あるいは面白いから?適切だから?ばかばかしいから?我慢できないから?あるいは人の信仰やバックグラウンドに共感するから? コンテンツの流れを途切れさせることは、グループ内で共有する信仰や慣習、関心を無視する(あるいは真逆の働きをする)効果を持つ可能性がある。ファクトチェックをして訂正する作業は、ポイントやオチが分かっていると意味がない恐れがある。その主張が心に訴え、ののしられ、面白がられる背景について知識が必要になる。それは主観性と客観性の違いを埋めたほうがいいということではない。その両方が、社会制度の共通背景と、全員にとって当たり前とは言えない文化習慣に依拠しているということだ。

第二に、一つ一つのフェイクニュースに関する調査から、ネットワーク化された調査へとシフトすることで、デジタルなプラットフォームやインフラによって社会制度や文化的な習慣がうまくいったり、ぎこちなくなったり、組織化されたりする具体的な道筋についてもっと知ることができる。このガイドは博物学のメタファーを取り入れたが、調査するオンライン空間は自然の生態系ではなく、具体的なデバイスやアルゴリズムに連なって社会生活や文化生活が展開されている作られた空間であると理解する必要がある^[1]。同時に、ウェブとオンライン・プラットフォーム自身は、どう使用されるかを定めることはできず、技術政治的な「形」だけでなく、オンライン生活がどう広がるかを「見通す」ことが必要になる。フェイクニュースやその他のうそを研究する中で、オンライン・プラットフォームとインフラの能力、ユーザーの社会的習慣を探ることができる。

このフィールドガイドを通して、フェイクニュースを扱うことが、デジタル環境の力学を研究する良い機会となることを示してきた。その力学は、実証的に調査されるべきだ。力学のさまざまな側面を説明しようと、私たちはネットワーク化された偽ニュースの特性と、その技術性、つまり偽ニュースが形作られ、発注され、計量され、データ化されデジタル・プラットフォームとともに生み出される手順を見てきた^[2]。

[1] Here we are inspired by work on “technological landscapes” in the study of science and technology, see, for example, Richard Rogers, *Technological Landscapes*, London: Royal College of Art, 1999.

そして第1章と第2章では、これらのプラットフォーム上のユーザーと、拡散について取り上げた。第3章では、オンラインコンテンツが埋め込まれたトラッキングのネットワークやデータを調べた。第4章では、オンラインで拡散するメディアコンテンツである画像ミームを分析し、第5章ではプラットフォームの特徴が政治家に対する攻撃にどう利用されているかを説明した。

[2] Sabine Niederer, *Networked Content Analysis: The Case of Climate Change*, unpublished manuscript, 2016.

フェイクニュースの状況や、その他の偽造コンテンツを研究する実証的なアプローチは重要だが、一方で多くの課題もある。デジタル・プラットフォームの現在の構造の性質上、そうした調査ができないことも多い。結果として、本書に記載したすべてのレシピは、あくまでも公開された範囲でのフェイクニュースの拡散を研究するものだ。フェイスブックに関する研究は、そうした課題の典型的な例を示している。ソーシャルネットワークのAPIで、研究者は公開ページのコンテンツを取得することはできるが、個人アカウントで交流した情報にアクセスすることはできない(一部は公開ページを通じてアクセスすることが可能ではあるが)。オンライン・プラットフォームは、それぞれの判断で公開と非公開の境界を決める。フェイクニュースの場合、技術面での対策が結果的に調査活動を難しくすることもあるのだ。フェイスブックの場合、問題のある投稿をプラットフォームから削除するという対策により、研究者はユーザーがこれらの投稿にどう接触しているのかを調べることができない。プラットフォーム、規制当局、政策立案者、ユーザーおよび他の人々が、新しく生まれた公開とプライバシーの空間、説明責任のメカニズムといった構造について広がっている問題を、どのように協議していくのかはこれからだ。

デジタル空間がどのように研究され、組織され、再構築されるかという疑問を超えて、このガイドが現代のデータ政治を研究し、そして入り込むためにデジタルな手法をどう使うかというインスピレーションの一助になることを願っている。データの作成と使用方法を具体化することができるのは誰か?議論を終わらせるだけでなく、議論を豊かにするために、どのようにデータを活用できるか?異なる種類のデータは、一つの視点だけでなく、複数の異なる視点を通じて、客観性を追求するのに役立つか?デジタルインフラ

の構成は、私たちの行動を、どのように規定するのか？ こうしたデータ社会から恩恵を受けるのは誰か？「The Field Guide to Fake News」は、パブリック・データ・ラボが継続している活動と実験のスタートだ。私たちはこのテーマについて、今後も引き続き調査していく。

**Jonathan Gray (@jwyg),
Liliana Bounegru(@bb_liliana),
Tommaso Venturini (@TommasoVenturin)
Paris, November 2017**

ツール用語集

このガイドで使われているさまざまなツールについて、簡単な説明とリンクを紹介する。ここにある説明を読めば、ガイドを理解するのに十分な情報が得られるようにした。どんなプロジェクトや調査においても、これらのツールがどう役立ち、何に使えるか(そして使えないか)を把握することが大切だ。実際にプロジェクトでツールを使う前に、下記に挙げてあるウェブサイトの詳細を確認することをおすすめする。

- 🔑 **BuzzSumo:** 特定のトピックやドメインに関し、最も多くのエンゲージメントを生み出しているコンテンツを調査するのに使える、ソーシャル・アナリティクス・サービス。言語や国、文字数やコンテンツの種類(記事、インフォグラフィック、インタビュー、動画)で検索できる。(http://buzzsumo.com/)
- 🔑 **CorText:** テキストデータの分析に使うオンライン・アプリケーション。統計やネットワークのさまざまな可視化ができる。(http://www.cortext.net/)
- 🔑 **CrowdTangle:** ウェブ上におけるコンテンツの拡散を追跡できるデジタルツール。フェイスブック、ツイッター、ユーチューブ、インスタグラム、バイン上で、投稿やアカウントのパフォーマンスも分析できる。(http://www.crowdtangle.com)
- 🔑 **CSV Rinse Repeat:** CSVファイルを構築(フィルタリング、クラスター化、構文解析、マーキング、マッチングな

ど)する、JavaScriptのツール。

- 🔑 **DMI Tracker Tracker:** Ghosteryプロジェクトのデータを使い、与えられたURLにおいて、分析ツール、ウィジェット、ソーシャルプラグインやその他のトラッカーの900を超える「フィンガープリント(特徴)」を検知するウェブツール。
(<https://tools.digitalmethods.net/beta/trackerTracker/>)

- 🔑 **DMI Triangulation Tool:** DMI Triangulation Tool : 複数のリストに共通するアイテムを特定する。
(<https://tools.digitalmethods.net/beta/triangulate/>)

- 🔑 **DownThemAll!**: ウェブページに含まれるすべてのリンクや画像が収集できる、Mozilla Firefoxのブラウザ拡張機能。(<https://addons.mozilla.org/en-US/firefox/addon/downthemall/>)

- 🔑 **Gephi:** ネットワーク分析・可視化ソフトウェア。大きなデータセットにおけるパターンやトレンド、特徴を見つけるのに特に有効なツール。(<http://gephi.org>)

- 🔑 **グーグル画像検索:** グーグルが提供する検索サービス。キーワードやクエリに関連する画像を収集できる。(<http://images.google.com>)

- 🔑 **グーグルニュース検索:** ニュース記事の検索結果を、掲載日時順に提供するグーグルのツール。

- 🔑 **グーグルVision API:** 画像分析ツール。写真をカテゴリ化し、物体や個人の顔を検知。テキストのコンテンツも抽出できる。

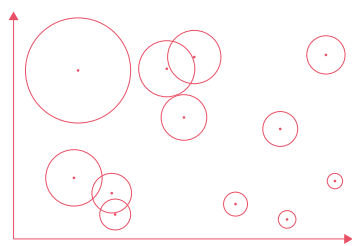
- 🔑 **グーグルウェブ検索:** 「ページランキング」に基づき結果を表示する検索エンジン(コンセプト用語集を参照)。

- 🔧 **Graph Recipes:** オンラインのJavaScriptツール。統計画像を作成したり、ネットワークの統計値を計算したりできる。デフォルトのスクリプトがいくつか提供されているが、ユーザーが追加もできる。(http://tools.medialab.sciences-po.fr/graph-recipes)
- 🔧 **Hyphe:** 半自動のウェブ・クローラー。ウェブページのリンクを特定・追跡し、ウェブサイトのコーパスを特定・カテゴリ化できる。ウェブ・エンティティのネットワークと、それらのつながりを作る。(http://hyphe.medialab.sciences-po.fr/)
- 🔧 **Image J:** JavaScriptの、オープンソース・プログラム。ビジュアルデータの編集、修正、加工、測定と分析ができる。(https://imagej.nih.gov/ij/)
- 🔧 **Le Monde Décodex:** オンライン上で拡散する情報の出どころを確認したり、うわさや偏った情報を特定したりするのに役立つ。(http://www.lemonde.fr/verification/)
- 🔧 **Netvizz app:** フェイスブックのアプリ。フェイスブック上で、グループやファンページ、検索機能などさまざまなセクションから、データを抽出できる。(https://apps.facebook.com/netvizz/)
- 🔧 **Radarly:** ソーシャルメディアのモニターツール。特定の話題や人物、イベントに関してどんなオンライン上で会話がされているかを追跡できる。(http://www.lemonde.fr/verification/)
- 🔧 **RAWGraphs:** データの可視化ツール。svgフォーマットに基づいており、カスタマイズがしやすい。可視化データは、ベクターグラフィック・アプリケーションで取り込み、編集が可能。または、ウェブページに直接埋め込むことができる。(http://rawgraphs.io)
- 🔧 **Spyonweb:** 登録されたユーザーのWHOISプロトコルまたは、インターネットリソース受託者のクエリで、

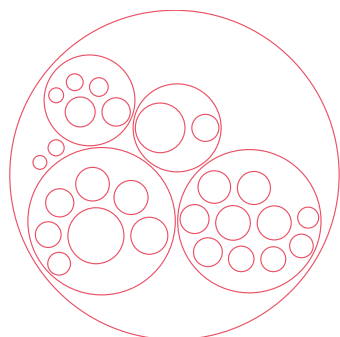
同じIDに関連するウェブサイトを特定できる。(http://spyonweb.com)

- 🔧 **Tab Save:** グーグルクロームのブラウザ拡張機能。PDFファイルや画像、URLのリストなどを収集し、保存できる。(https://chrome.google.com/webstore/detail/tab-save/lkngoeaeclaebmpkgapchgjdbaikacki?hl=en)
- 🔧 **Table2Net:** 表(.cvs)をネットワーク(.gexf)に変換できる。(http://tools.medialab.sciences-po.fr/table2net)
- 🔧 **TCAT:** ツイッターからデータを収集するツール。キーワードやユーザー、ハッシュタグ検索でデータを取り出せる。(https://wiki.digitalmethods.net/Dmi/ToolDmiTcat)
- 🔧 **The Wayback Machine:** 一定期間ごとに、更新されるウェブサイトのバージョンをアーカイブする。Internet Archiveによるプロジェクト。(https://archive.org/web/)

コンセプト用語集



バブルグラフ



円パッキング

- **API (Application Programming Interface):** 二つのソフトウェアが互いに通信することを可能にする、明確に定義された通信方式。ウェブリサーチでは、データベースに直接アクセスすることなく、公開または非公開のデータセット(通常はソーシャルプラットフォームで収集されたもの)からのデータ抽出によく使われる。
- **バブルグラフ:** 投影点のサイズが第3の変数に比例する散布図の一種(用語集の定義を参照)。
- **円パッキング:** 階層構造のデータを視覚化する、データビジュアライゼーションの一種。各クラスターやグループは円で表される。円の内側に、サブグループを表す小さな円が示される。円のサイズは、異なる定量的特性を表す。
- **クリックベイト:** ユーザーの注意を引き、特定のページへのリンクをクリックするよう促すオンライン上のコンテンツ。
- **CMS (Content Management System):** デジタルコンテンツや、特にウェブサイトの作成・管理に使うソフトウェア・アプリケーション
- **エマージェント・コーディング:** 観察前に前提とされていないが、探査プロセスで反復的に定義されているカテゴリーを通して項目を分類する技術。このタイプのコーディングは、データをあらかじめ設定されたカテゴリーに適合させるのではなく、調査対象が使用しているカテゴリーに可能な限り近づけることを目的としている。
- **フェイスブックページまたはフォロワー:** あるフェイスブックページに「いいね!」したか、グループに参加したユーザーの数。

→ **力学モデルのネットワーク・レイアウト:** ネットワーク内の項目を空間化し、データを理解するためのグラフ描画アルゴリズム。力学モデルのレイアウトは、ノード間に反発力を使用し、隣接ノード間に引力を加える。

→ **グーグルアナリティクスID:** ユーザーのアカウントを特定するため、グーグルアナリティクス(ウェブサイトのトラフィックを追跡、レポートするグーグルのサービス)によって割り当てられたID。

→ **ヒートマップ:** データ・ビジュアライゼーションの一種。図または値が、色の勾配によって表される。

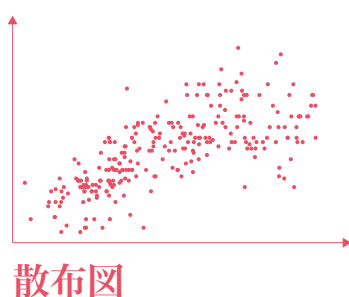
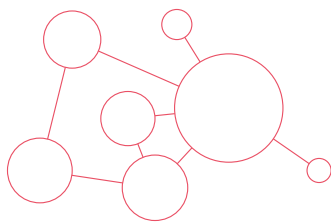
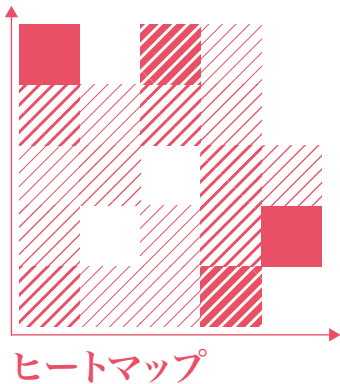
→ **インタラクション/ エンゲージメント:** フェイスブック投稿へのいいね!やシェア、コメントの合計数。(参照: <http://www.crowdtangle.com/resources/glossary>)

→ **ネットワーク分析と可視化:** 個人やウェブページ、アカウント、その他グループのつながりや関係性を調査するプロセス。Gephiなどの可視化ツールを使うと、ノード(個人、集団、項目)において関連する事象や、それらに関連付けるつながりやリンクの特徴が分かる。

→ **ネットワークグラフ:** データ・ビジュアライゼーションの一種。エンティティ間の関係を強調するために使用される。ノード(またはポイント)がエンティティを、ライン(または円弧、エッジ)は、それらの関係性を表す。

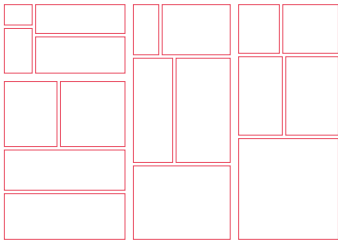
→ **ページランキング:** グーグル検索で用いられるアルゴリズム。検索結果をランク付けする。特定の検索ワードにおいて、ウェブサイトが表示される位置に影響を与える要因はいくつかあるが(公表されていない方法で組み合わせが行われている)、基本となっているのはそのサイトへの言及の再帰カウントだ。

→ **散布図:** データ・ビジュアライゼーションの一種。二つの異なる変数(図の軸に対応する)上にある値に従い、デカルト図に点が配置される。二つの変数間の相関を明



らかにする際に最もよく使用される。

- **ソースコード:** HTMLやJavaScriptなどのプログラミング言語で記述された一連の命令。ソフトウェアの機能やドキュメントの表示方法を定義する。
- **CrowdTangleスプレッドシートの購読者カウント:** 投稿が掲載された時点における、アカウントの購読者数。アカウント上で特定できる、今現在の購読者数とは異なる。



ツリーマップ

- **ツリーマップ:** データ・ビジュアライゼーションの一種。入れ子になった長方形で、階層的な分類を表すのに使用される。各カテゴリーは、サイズがグループの重要度、または重みに比例する長方形で表され、サブグループを表す小さな長方形がその内側にある。
- **ウェブ・クロール:** ウェブサイト、またはウェブページのアンサンブルをつなぐハイパーリンクのネットワークを抽出するプロセス。クロールは普通「スパイダー」と呼ばれる、自動または半自動ツールによって実行され、HTMLページに存在するすべてのハイパーリンクを識別、追跡する。
- **ウェブスクレイピング:** 構造化された情報やコンテンツをウェブサイトから抽出する方法(表形式で保存する)。

CONTRIBUTORS AND ACKNOWLEDGEMENTS

Contributors

Co-investigators

Liliana Bounegru, University of Groningen, University of Ghent

Jonathan Gray, King's College London

Tommaso Venturini, Univ Lyon, Inria, ENS de Lyon, CNRS, UCB

Michele Mauri, Density Design, Politecnico di Milano

Research and Editorial Assistant

Daniela Demarchi, University of Amsterdam

Editorial Design

Ángeles Briones, Politecnico di Milano

Carlo De Gaetano, Politecnico di Milano

DensityDesign Researchers

Agata Brilli, Politecnico di Milano

Ángeles Briones, Politecnico di Milano

Carlo De Gaetano, Politecnico di Milano

Gabriele Colombo, Politecnico di Milano

Mariasilvia Poltronieri, Politecnico di Milano

Michele Invernizzi, Politecnico di Milano

Michele Mauri, Politecnico di Milano

Paolo Ciuccarelli, Politecnico di Milano

Tommaso Elli, Politecnico di Milano

Facilitators

Alex Gekker, University of Amsterdam

Anders Munk, Aalborg University

Bilel Benbouzid, University of Paris-Est Marne-la-Vallée

Erik Borra, University of Amsterdam

Esther Weltevrede, University of Amsterdam

Jonathan Gray, King's College London

Jorn Preuss, University of Siegen

Liliana Bounegru, University of Ghent, University of Groningen

Marc Tuters, University of Amsterdam

Mathieu Jacomy, Sciences Po

Natalia Sánchez-Querubín, University of Amsterdam

Nicolas Baya-Laffite, University of Lausanne

Paolo Ciuccarelli, Politecnico di Milano

Richard Rogers, University of Amsterdam

Sabine Niederer, Hogeschool van Amsterdam

Tommaso Venturini, Univ Lyon, Inria, ENS de Lyon, CNRS, UCB

Analysts

Anders Grundtvig, Aalborg University

Anna Keuchenius, University of Amsterdam

Antonio Martella, University of Amsterdam

Asbjørn Fleinert Mathiasen, Aalborg University

Asger Gehrt Olesen, Aalborg University

Carlo Santagiustina, University of Amsterdam

Charlotte Leclercq, University of Amsterdam

Daniel Bach, Aalborg University

Daniela Demarchi, University of Amsterdam

Ecesu Erol, University of Amsterdam

Emil Jørgensen, Aalborg University

Joep Voorn, University of Amsterdam

Jörn Preuss, University of Siegen

Kaspar Beelen, University of Amsterdam

Katerina Gladkova, University of Amsterdam

Lieke Kersten, University of Amsterdam

Lisanne Blomberg, University of Amsterdam

Manon van Hoek, University of Amsterdam

Maria Hayat, University of Amsterdam

Marlene Scherf, University of Amsterdam

Michel Blonk, University of Amsterdam

Mintsje de Witte, University of Amsterdam

Mischa Benjamin Szpirt, Aalborg University

Pieter Vliegenthart, University of Amsterdam

Rina Tsubaki, European Journalism Centre

Ronja Ingeborg Lofstad, Aalborg University

Sal Hagen, University of Amsterdam

Stefani Mans, University of Amsterdam

Stefanie Voortman, University of Amsterdam

Talia Castellanos Usigli, University of Amsterdam

Zoë Versteegen, University of Amsterdam

Contributors by chapter

chapter 1

MAPPING FAKE NEWS HOTSPOTS ON FACEBOOK

Facilitators: Erik Borra, University of Amsterdam
Liliana Bounegru, University of Groningen, University of Ghent
Jonathan Gray, King's College London
Natalia Sánchez-Querubín, University of Amsterdam
Esther Weltevrede, University of

Members: Lisanne Blomberg, University of Talía Castellanos, University of Tommaso Elli, DensityDesign Research
Mintsje de Witte, University of Stefanie Voortman, University of Stefani Mans, University of Amsterdam
Antonio Martella, University of Rina Tsubaki, European Journalism
Manon van Hoek, University of Zoë Versteegen, University of Joep Voorn, University of Amsterdam

chapter 2

TRACING THE CIRCULATION OF FAKE NEWS ON THE WEB

Facilitators: Mathieu Jacomy, Sciences Po
Anders Grundtvig, Aalborg University
Tommaso Venturini, Univ Lyon, Inria, ENS de Lyon, CNRS, UCB

Members: Agata Brilli, DensityDesign Research
Daniela Demarchi, University of Ronja Lofstad, Aalborg University
Anders Kristian Munk, Aalborg

chapter 3

USING TRACKER SIGNATURES TO MAP THE TECHNO-COMMERCIAL UNDERPINNINGS OF FAKE NEWS SITES

Facilitator: Liliana Bounegru, University of Groningen, University of Ghent
Jonathan Gray, King's College London
Richard Rogers, University of Amsterdam

Members: Michele Invernizzi, DensityDesign
Mischa Szpirt, Aalborg University

chapter 4

HOW TO STUDY POLITICAL MEMES ON FACEBOOK

Facilitators: Nicolas Baya-Laffite, University of Lausanne
Bilel Benbouzid, University of Paris-Est Marne-la-Vallée
Marc Tuters, University of Amsterdam

Members: Daniel Bach, Aalborg University
Carlo De Gaetano, DensityDesign
Sal Hagen, University of Amsterdam
Emil Jørgensen, Aalborg University

chapter 5

MAPPING TROLL-LIKE PRACTICES ON TWITTER

Facilitators: Erik Borra, University of Amsterdam
Sabine Niederer, Hogeschool van Jörn Preuß, University of Siegen
Esther Weltevrede, University of

Members: Ángeles Briones, DensityDesign Research
Michel Blonk, University of Amsterdam
Lieke Kersten, University of Amsterdam
Carlo Santagiustina, University of Marlene Scherf, University of Amsterdam
Pieter Vliegthart, University of

SPECIAL THANKS

五十嵐茉莉子、伊藤亮介、稲増五生、内田友紀、鵜殿有正、永瀧和貴、小笠原盛浩、岡本真、
小木曾克典、越智慎一郎、小野秀子、開沼博、勝野正博、加納亜弥、亀松太郎、河合優一郎、
間間理、北野章子、儀武勝希、日下紗代子、櫛引素夫、小泉豪紀、小林かや、坂田友梨子、
桜井陽、佐々木千代子、芝拓耶、清水暁子、酒造唯、杉山まゆ、鈴木さくら、曾根田和久、
高木恒一、高橋梢、武末文男、田幸大輔、田子環、多田哲馬、徳田順子・慶太、富谷瑠美、
永里真弓、中根靖明、根本望、花田一郎、林恭一、藤村厚夫、Furuta Daisuke、三浦なつみ、
道下健弘、宮澤京子、耳塚志津子、耳塚紳一、村岡和彦、銘苅拓也、望月千恵、本橋健、
モモセタケシ、百瀬渚、森保之、森なを子、森廣陽子、安田菜津紀、矢田部匡、矢田部祐佳、
山口由美子、山田雅俊、山村哲史、横山晃嗣、吉見憲二、與那覇里子(敬称略、50音順)
クラウドファンディングなどを通じてご支援いただいた皆様、ありがとうございました!

邦訳チーム / TRANSLATION TEAM

明知隼二(中国新聞) / Junji Akechi, The Chugoku Shimbun
足立義則(NHK) / Yoshinori Adachi, Japan Broadcasting Corporation
井上直樹(西日本新聞) / Naoki Inoue, The Nishinippon Shimbun
木下真寿美(徳島新聞) / Masumi Kinoshita, The Tokushima Shimbun
栗原岳史(NHK) / Takeshi Kurihara, Japan Broadcasting Corporation
小石克(佐賀新聞) / Osamu Koishi, The Saga Shimbun
藤代裕之(JCEJ) / Hiroyuki Fujishiro, JCEJ
耳塚佳代(JCEJ) / Kayo Mimizuka, JCEJ
森廣陽子(ライター・翻訳家) / Yoko Morihiro, Writer and Translator

監修: 日本ジャーナリスト教育センター

/ Japan Center of Education for Journalists(JCEJ)



A FIELD
GUIDE TO

**"FAKE
NEWS"**

and other

INFORMATION DISORDERS

<http://fakenews.publicdatalab.org/>

JCEJ 日本ジャーナリスト教育センター
Japan Center of Education for Journalist