

FROM RESEARCH TO INDUSTRY

cea tech



DÉCOUVERTE DE POINTS DE VUE SUR LES MÉDIAS SOCIAUX

OPHÉLIE FRAISIER

Encadrée par Mohand Boughanem, Romaric Besançon, Guillaume Cabanac et Yoann Pitarch

Séminaire « Savoirs, Réseaux, Médiations »

13 avril 2018

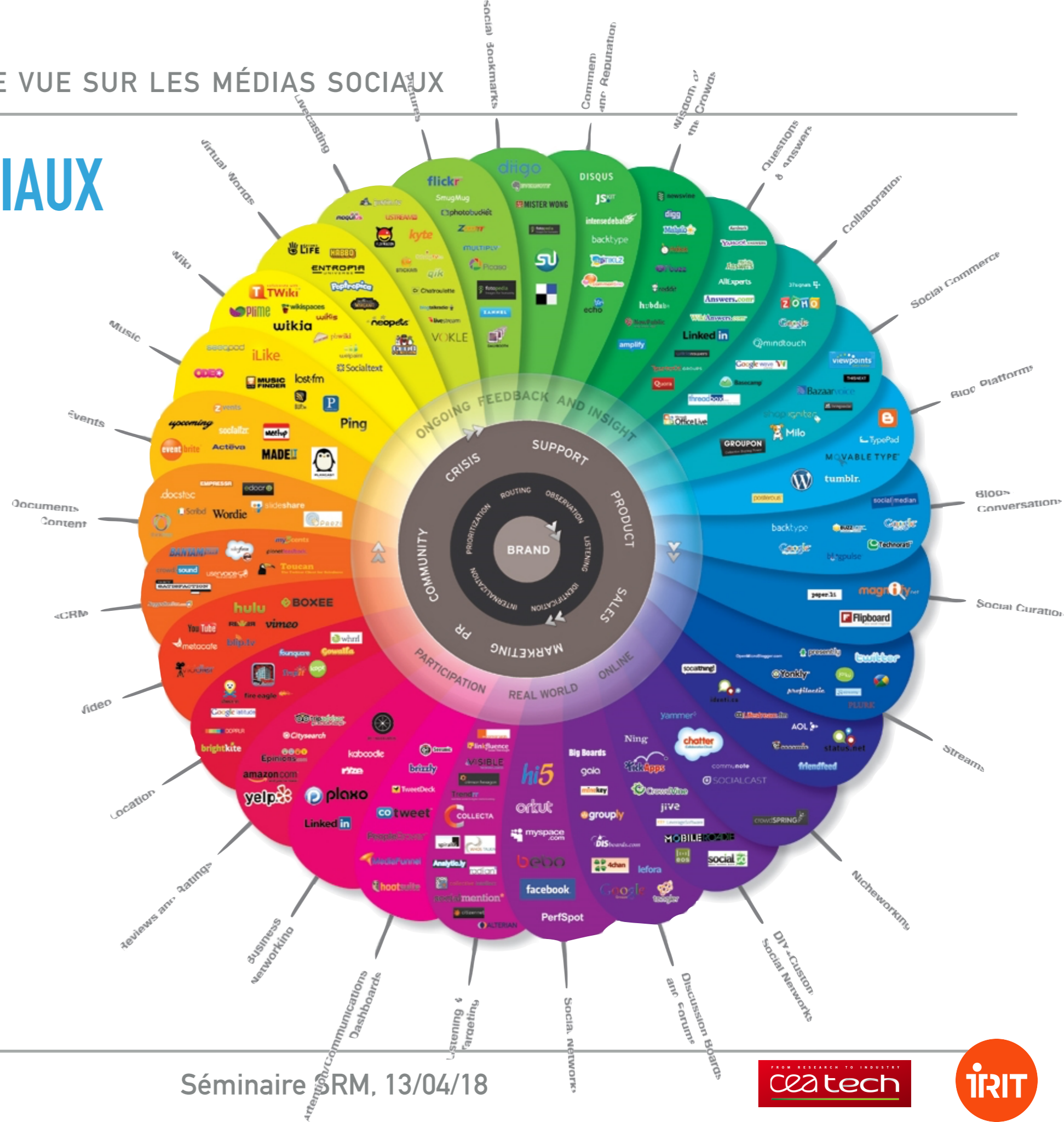
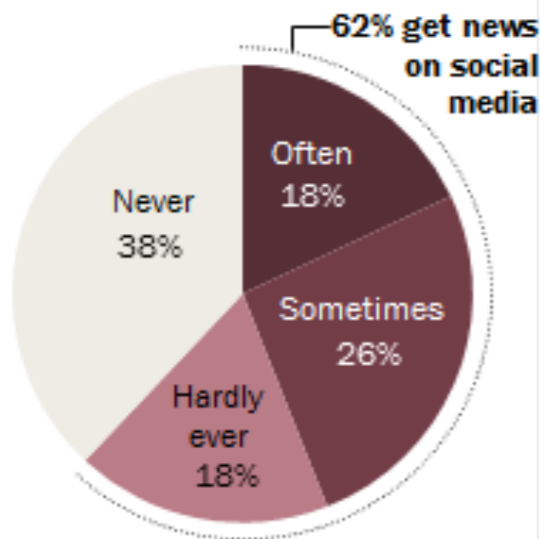


PROJET COFINANCÉ PAR LE FONDS EUROPÉEN DE DÉVELOPPEMENT RÉGIONAL

DES MÉDIAS SOCIAUX OMNIPRÉSENTS

About 6-in-10 Americans get news from social media


% of U.S. adults who get news on a social networking site ...

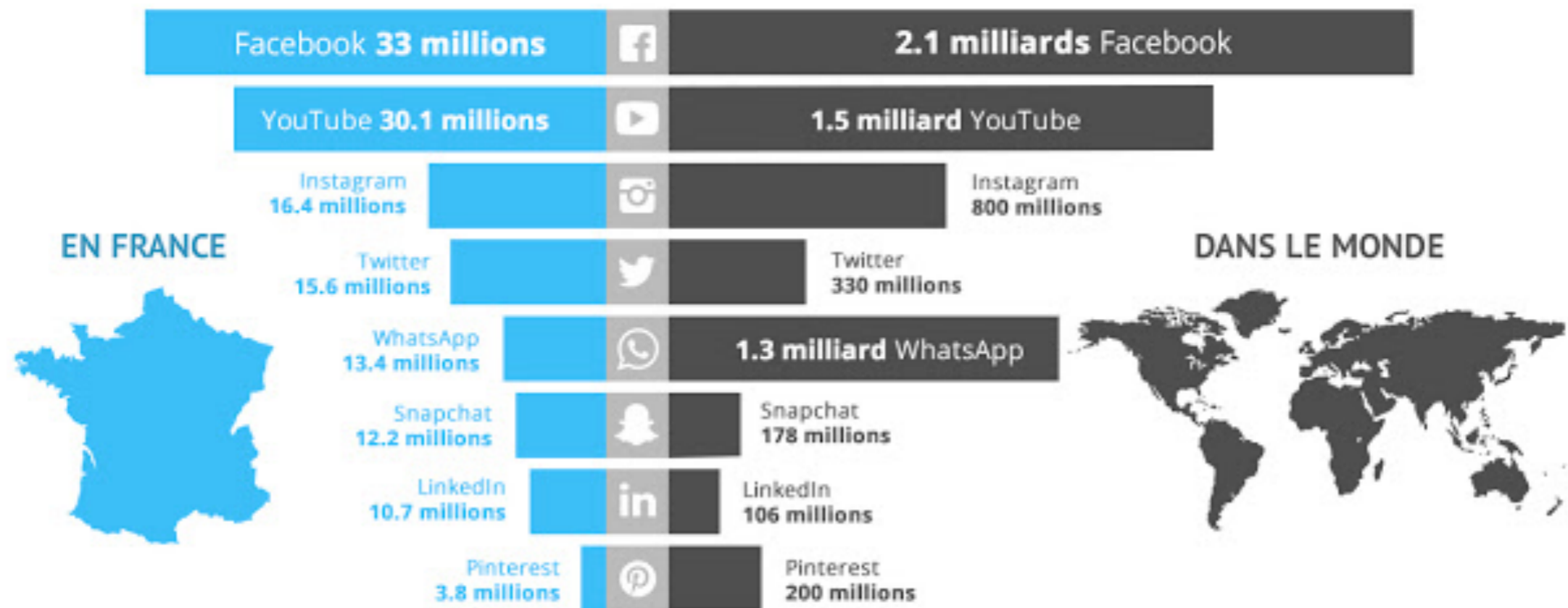


DES MÉDIAS SOCIAUX OMNIPRÉSENTS

LES RÉSEAUX SOCIAUX

NOMBRE D'UTILISATEURS ACTIFS

Janvier 2018
Infographie réalisée par 



FOUILLE D'OPINION, ANALYSE DE SENTIMENT, POINT DE VUE...

Opinion personnelle ciblée

- ▶ Habituellement détectée comme positive, négative ou neutre
- ▶ Ex : Critiques de produits ou de services en ligne



Positionnement

- ▶ Réponse élaborée, basée sur un ensemble de principes et de valeurs
- ▶ Ex : Positionnements politiques

➔ Point de vue

FOUILLE D'OPINION, ANALYSE DE SENTIMENT, POINT DE VUE...

Opinion personnelle ciblée

- ▶ Habituellement détectée comme positive, négative ou neutre
- ▶ Ex : Critiques de produits ou de services en ligne



Marvel's most unusual film of its Cinematic Universe is big on world-building, empowerment, strong (female) characters and jabs at the current state of the world. All that mixed with the ingredients of a blockbuster film, a great cast, good humorous timing and...

[More](#)



Jens S.

★ Super Reviewer

Positionnement

- ▶ Réponse élaborée, basée sur un ensemble de principes et de valeurs
- ▶ Ex : Positionnements politiques

➔ **Point de vue**

POINTS DE VUE POLITIQUES SUR TWITTER ?

- ▶ Allport (1937)
 - ▶ Verbalisations produites par des personnes,
 - ▶ Sur des thématiques importantes pour elles,
 - ▶ Conscientes que d'autres réagissent à ces thématiques,
 - ▶ Et que leur comportement peut leur permettre d'atteindre un but.

Limites

- ▶ Peu de profondeur
- ▶ Qualité difficilement mesurable

POINTS DE VUE POLITIQUES SUR TWITTER ?

▶ « Fake news »

- ▶ Truthy: mapping the spread of astroturf in microblog streams (Ratkiewicz et al., 2011)
- ▶ Fake tweet buster: a webtool to identify users promoting fake news on Twitter (Saez-Trumper, 2014)
- ▶ Rumor Gauge: Predicting the Veracity of Rumors on Twitter (Vosoughi et al. 2017)

▶ Prédiction d'élections

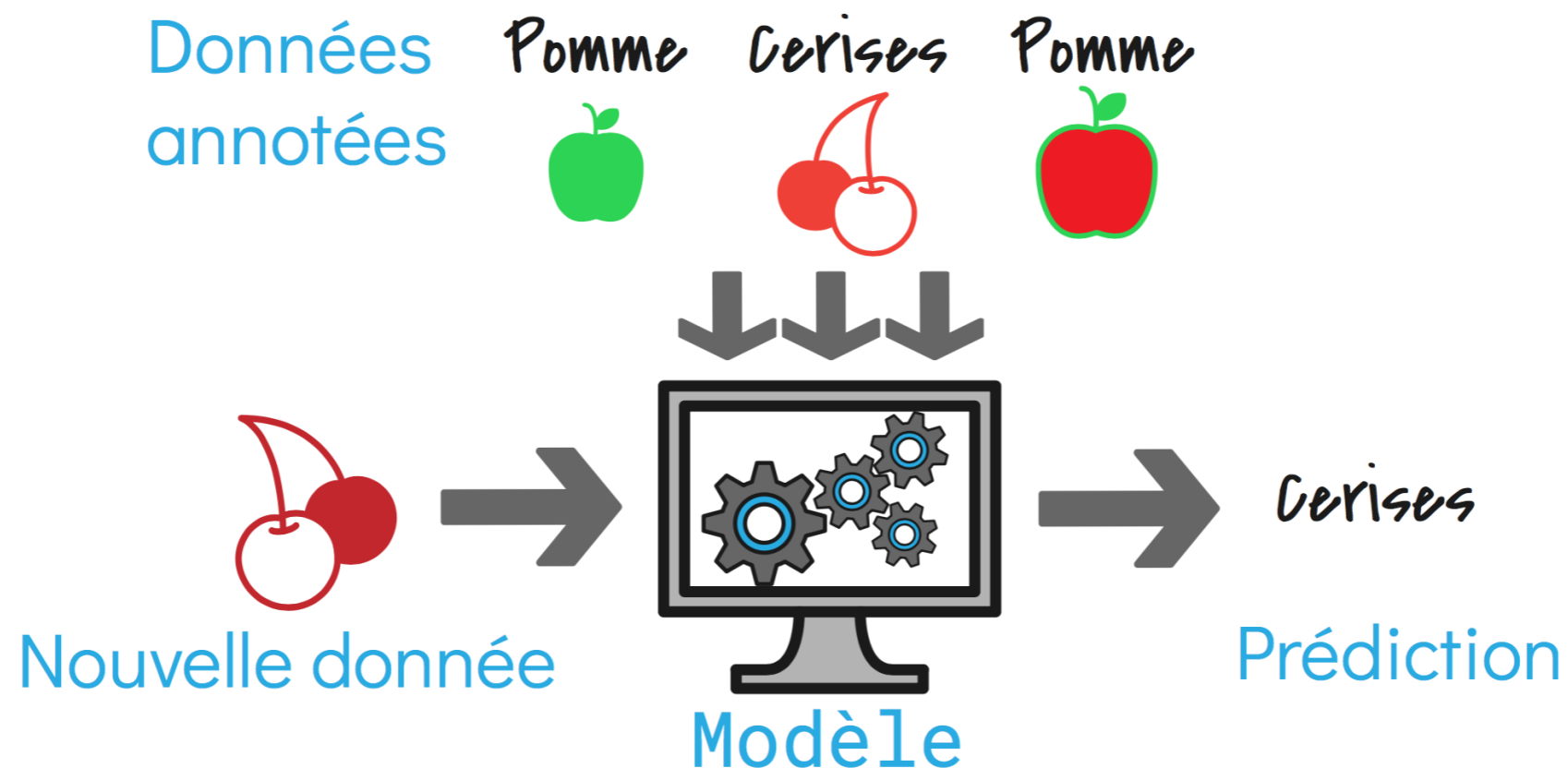
- ▶ Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment (Tumasjan et al., 2010)
- ▶ Predicting the 2011 Dutch Senate Election Results with Twitter (Sang et Bos, 2012)
- ▶ Why the Pirate Party Won the German Election of 2009 or the Trouble with Predictions (Jungherr et al., 2012)
- ▶ No, You Cannot Predict Elections with Twitter (Gayo-Avello, 2012)

POINTS DE VUE POLITIQUES SUR TWITTER ?

- ▶ Détection de points de vue
 - ▶ Democrats, Republicans and Starbucks Afficionados: User Classification in Twitter (Pennacchiotti et Popescu, 2011)
 - ▶ Predicting Political Preference of Twitter Users (Makazhanov et Rafiei, 2013)
 - ▶ Predicting the Political Alignment of Twitter Users (Conover et al., 2011)

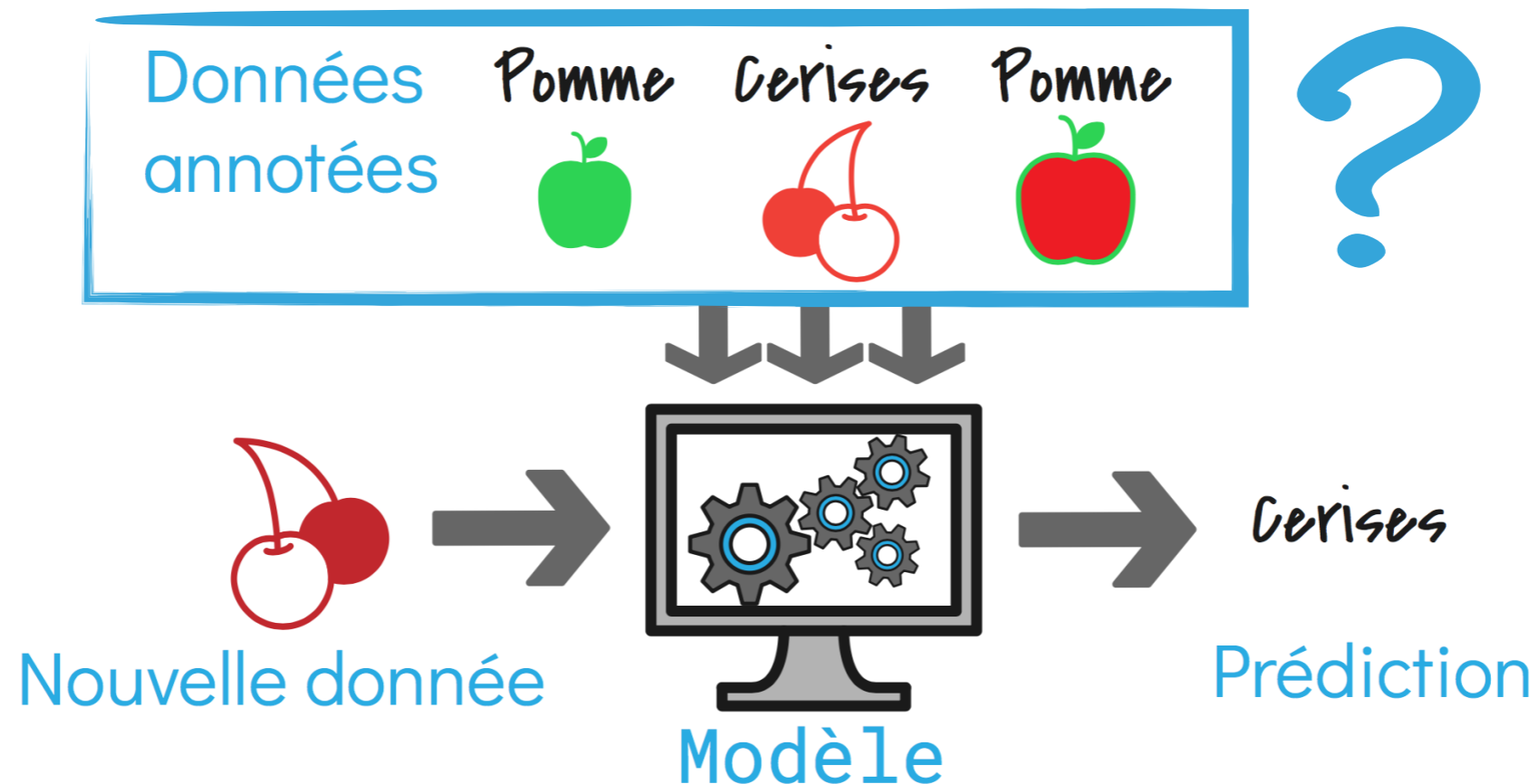
POINTS DE VUE POLITIQUES SUR TWITTER ?

- ▶ Détection de points de vue
 - ▶ Democrats, Republicans and Starbucks Afficionados: User Classification in Twitter (Pennacchiotti et Popescu, 2011)
 - ▶ Predicting Political Preference of Twitter Users (Makazhanov et Rafiei, 2013)
 - ▶ Predicting the Political Alignment of Twitter Users (Conover et al., 2011)



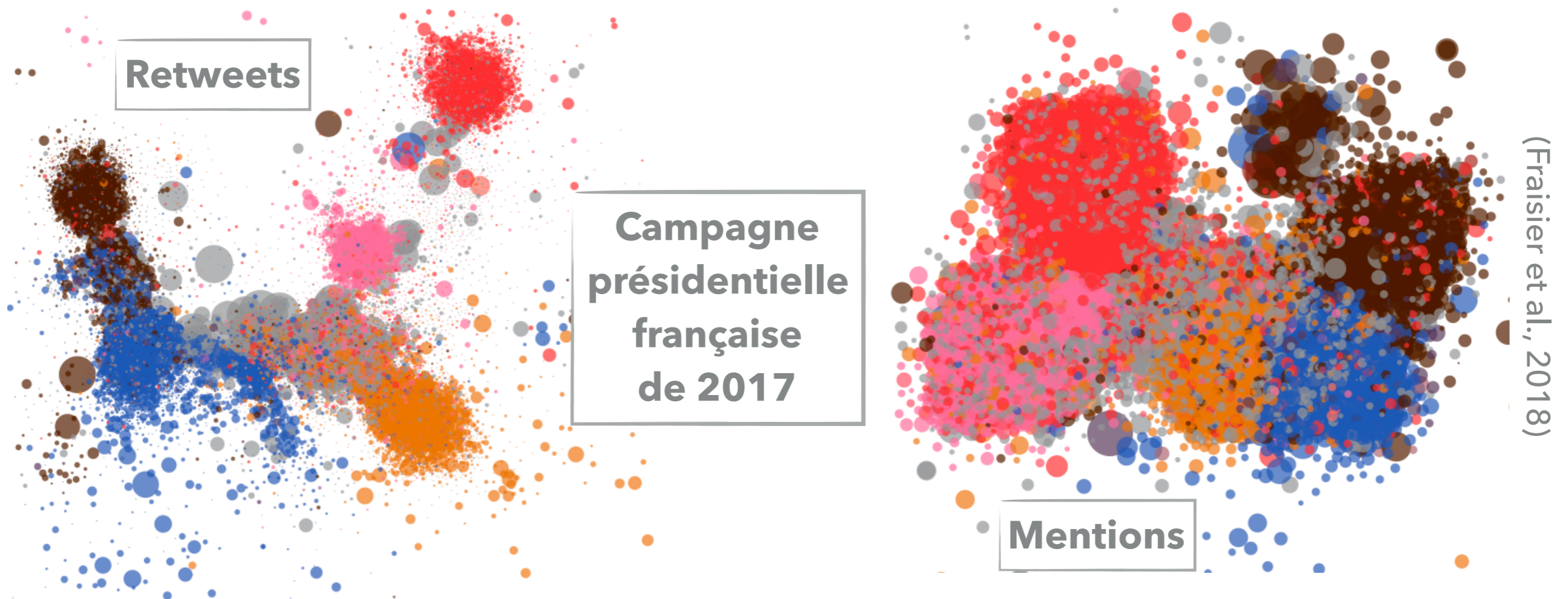
POINTS DE VUE POLITIQUES SUR TWITTER ?

- ▶ Détection de points de vue
 - ▶ Democrats, Republicans and Starbucks Afficionados: User Classification in Twitter (Pennacchiotti et Popescu, 2011)
 - ▶ Predicting Political Preference of Twitter Users (Makazhanov et Rafiei, 2013)
 - ▶ Predicting the Political Alignment of Twitter Users (Conover et al., 2011)



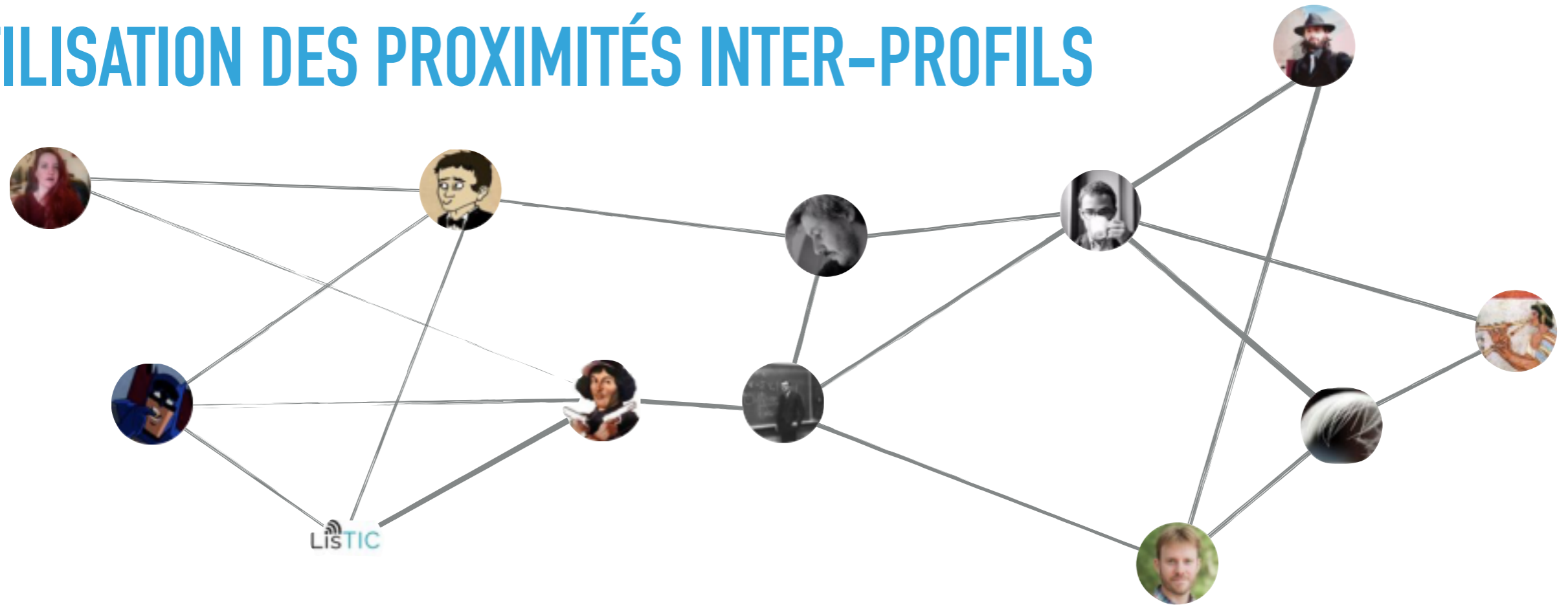
POLARISATION ET CHAMBRES D'ÉCHOS

- ▶ Homophilie existante sur les RSN (McPherson et al, 2001)
- ▶ Twitter très polarisé sur les sujets politiques (Sunstein, 2009; Conover, 2011)



« An echo chamber is a metaphorical description of a situation in which information, ideas, or beliefs are amplified or reinforced by communication and repetition inside a defined system. »

UTILISATION DES PROXIMITÉS INTER-PROFILS



▶ Mesures de proximités en fonction de différents aspects

- ▶ Contenu textuel
- ▶ Contexte social
- ▶ Contexte géographique

I'm #YesBecause it's a vote for the people.
People make Scotland.
vine.co/v/MLbvK2gYbmp #indyref #yes

♥ Thibault Le Hégarat 📺 a aimé

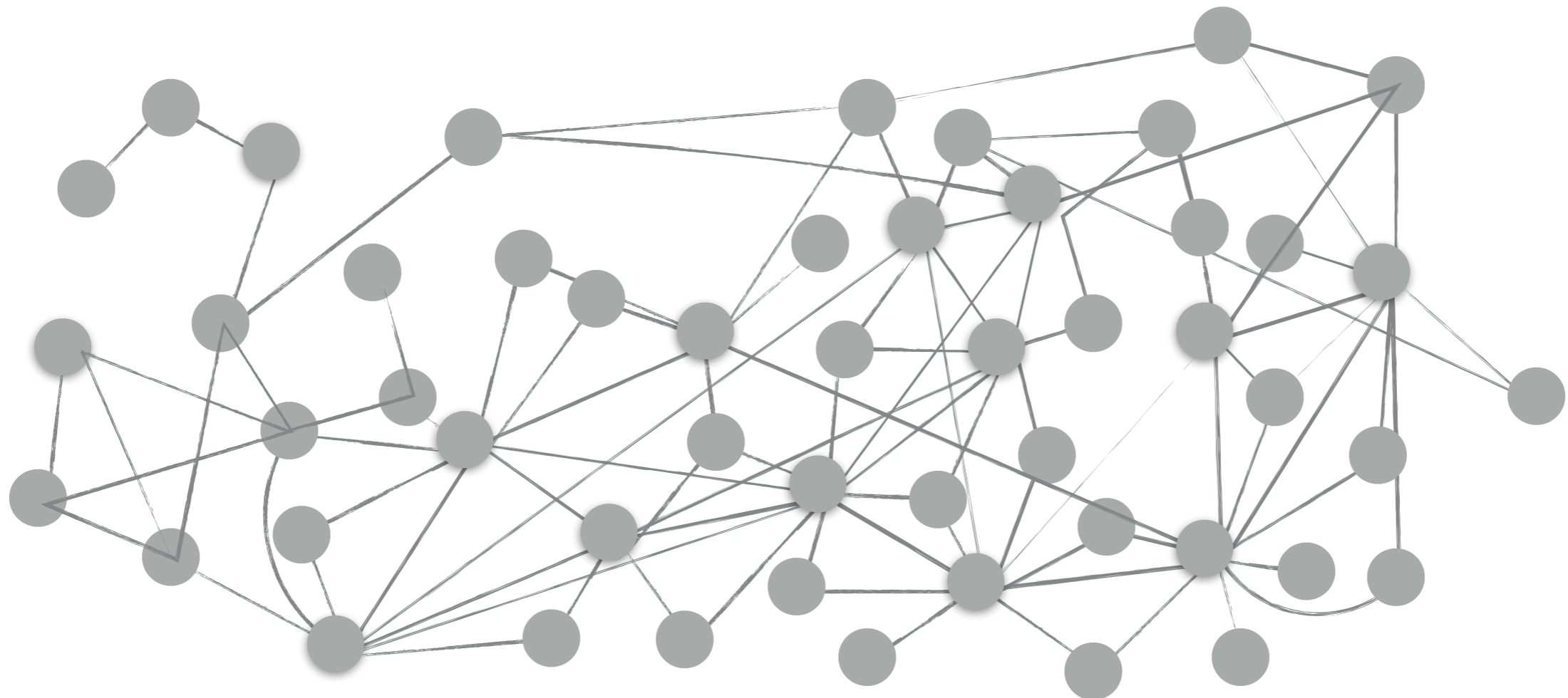
👤 -Boulet- et 5 autres suivent

📍 Glasgow

🔄 Xavier Molénat a retweeté

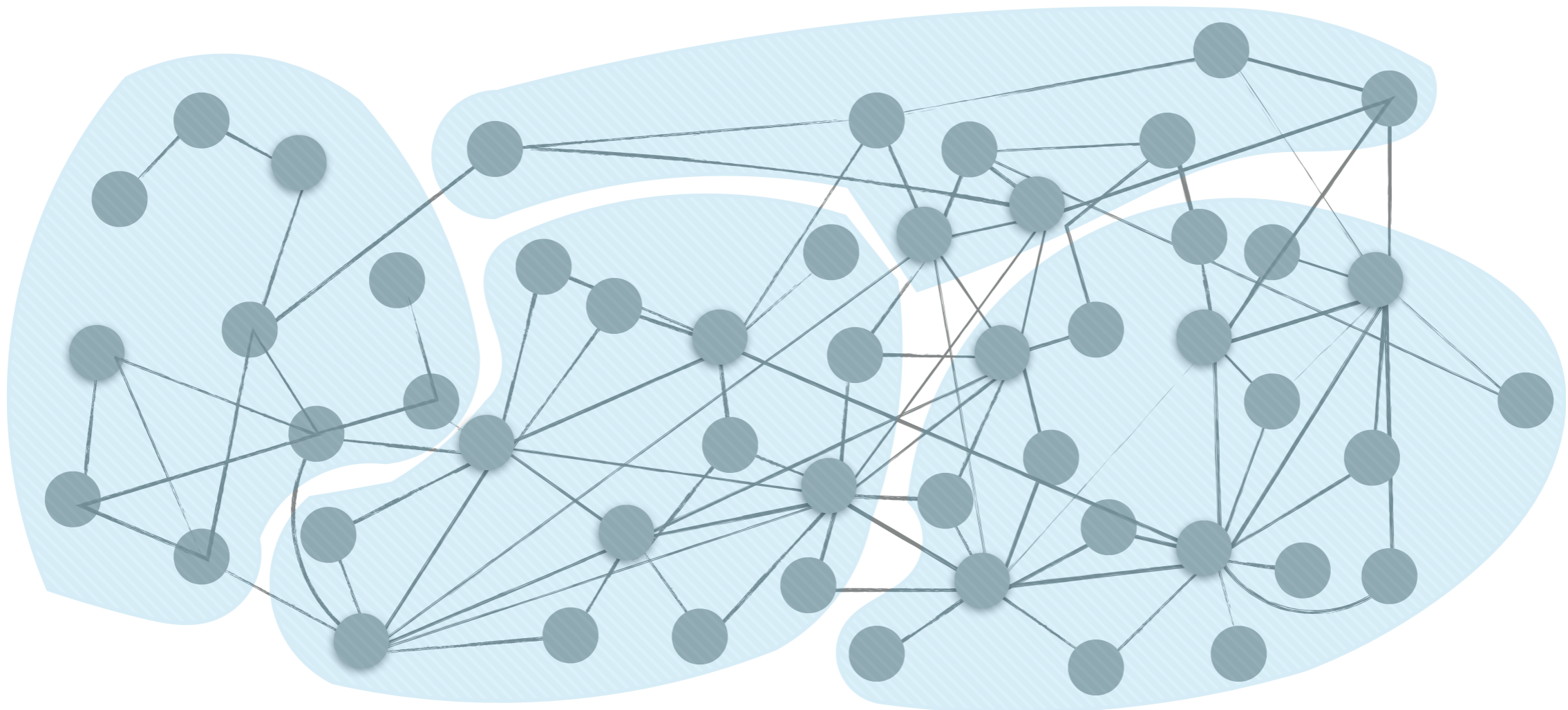
UTILISATION DES PROXIMITÉS INTER-PROFILS

- ▶ Utilisation des communautés détectées sur les différents graphes
- ▶ Au départ : seulement une poignée de profils dont on connaît les points de vue, répartis dans différentes communautés



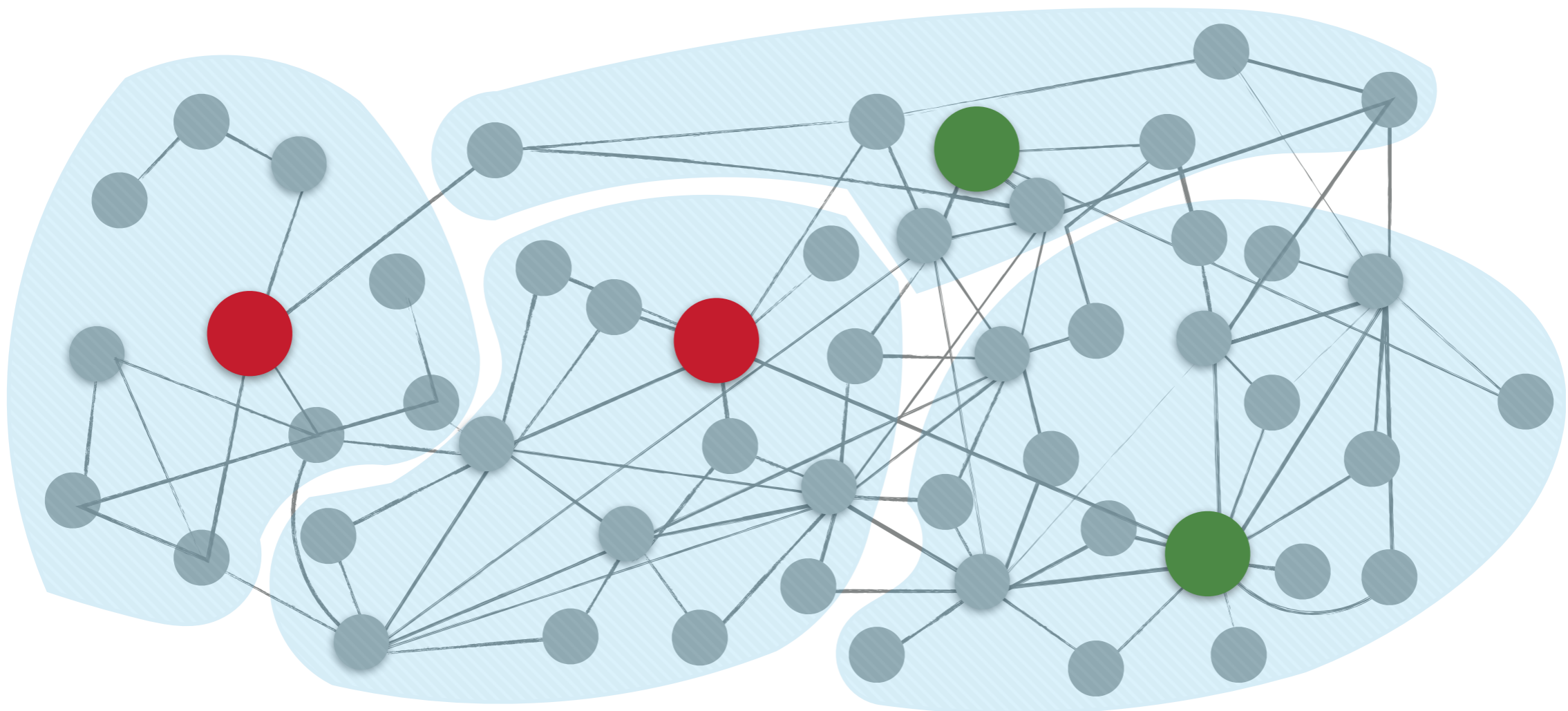
UTILISATION DES PROXIMITÉS INTER-PROFILS

- ▶ Utilisation des communautés détectées sur les différents graphes
- ▶ Au départ : seulement une poignée de profils dont on connaît les points de vue, répartis dans différentes communautés



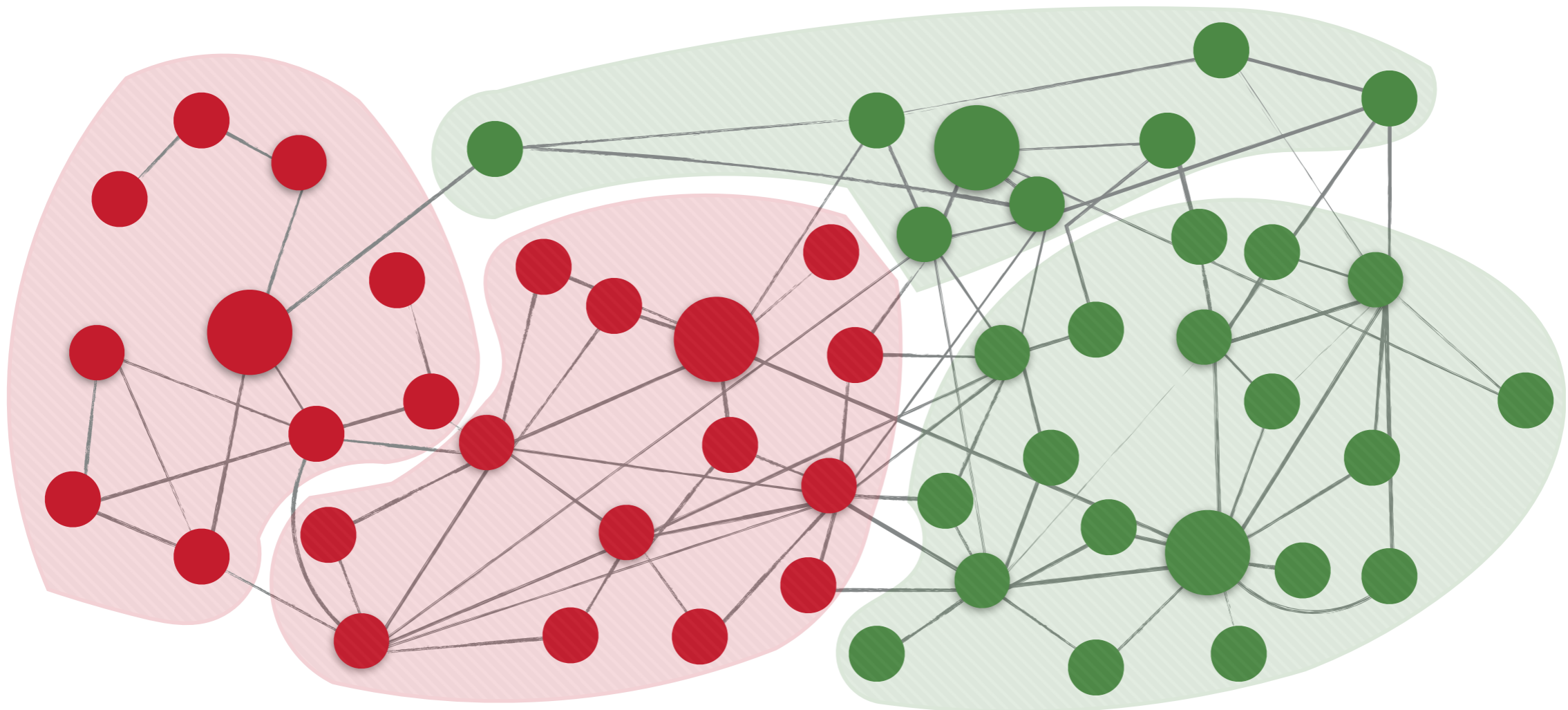
UTILISATION DES PROXIMITÉS INTER-PROFILS

- ▶ Utilisation des communautés détectées sur les différents graphes
- ▶ Au départ : seulement une poignée de profils dont on connaît les points de vue, répartis dans différentes communautés



UTILISATION DES PROXIMITÉS INTER-PROFILS

- ▶ Utilisation des communautés détectées sur les différents graphes
- ▶ Au départ : seulement une poignée de profils dont on connaît les points de vue, répartis dans différentes communautés



UTILISATION DES PROXIMITÉS INTER-PROFILS

Jeu de données

Points de vue

Effectif

Jeu de données	Points de vue	Effectif
SR Référendum sur l'indépendance écossaise [2014]	Oui, Non	1043 profils <i>(Brigadir et al, 2015)</i>
ME Élections législatives 2014 aux USA	Républicain, Démocrate	1457 profils <i>(Brigadir et al, 2015)</i>
PU Élections présidentielles 2016 aux USA	Républicain, Démocrate	908 profils <i>(Littman et al, 2016)</i>
PF Élections présidentielles 2017 en France	FI, PS, EM, LR, FN	18649 profils <i>(Fraisier et al, 2018)</i>

Précision des points de vue prédits

Social	SR	ME	PU	PF
Retweets	91 %	97 %	99 %	91 %
Retweets réciproques	100 %	100 %	100 %	96 %
Mentions	25 %	72 %	74 %	90 %
Mentions réciproques	78 %	94 %	94 %	84 %
Abonnements	97 %	97 %	98 %	89 %
Abo. réciproques	99 %	25 %	97 %	89 %
Icônes	55 %	25 %	26 %	67 %

Texte	SR	ME	PU	PF
Urls	51 %	87 %	76 %	48 %
Hashtags	25 %	57 %	26 %	5 %

Géographique	SR	ME	PU	PF
Villes	60 %	64 %	65 %	41 %
Régions	56 %	56 %	58 %	37 %

MODÈLE MULTI-PROXIMITÉS

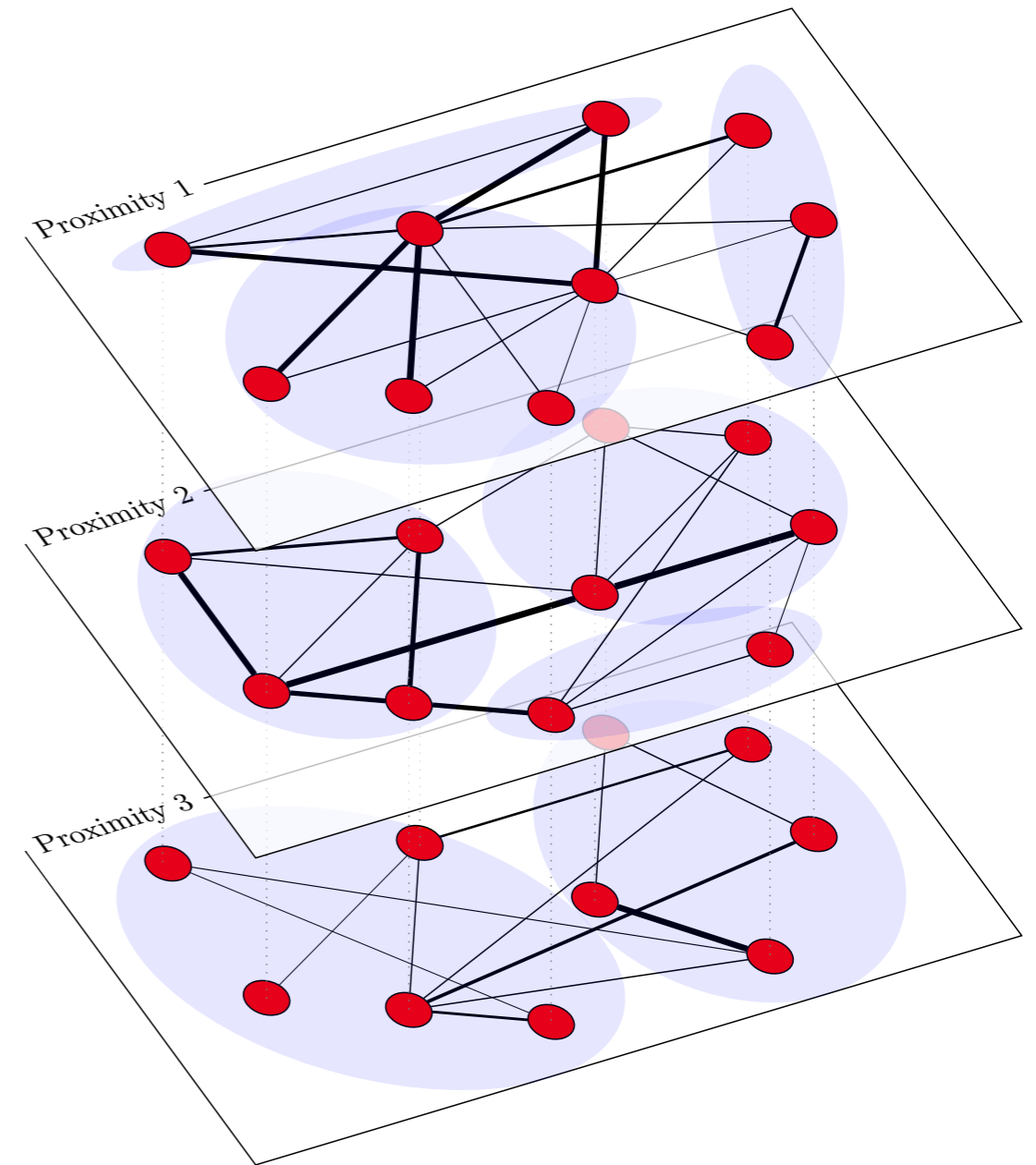
- ▶ Proximités les plus efficaces : communautés généralement petites

Rappel des points de vue prédits

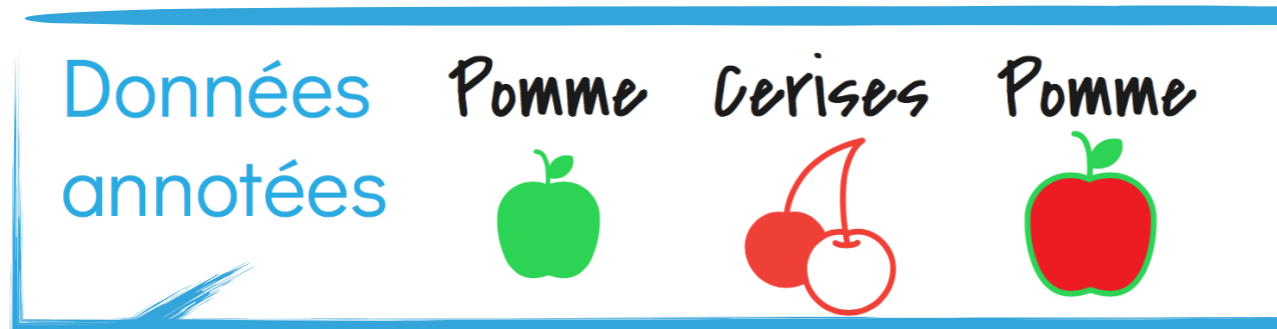
	Retweets réciproques
SR	34 %
ME	3 %
PU	29 %
PF	22 %

MODÈLE MULTI-PROXIMITÉS

- ▶ Proximités les plus efficaces : communautés généralement petites
- ➔ Itérations sur plusieurs proximités
- ➔ Proximités les plus homogènes utilisées en premier



SCORES



		SR			ME			PU			PF		
		p	r	f1	p	r	f1	p	r	f1	p	r	f1
80 %	SVM-Profil	92 %	92 %	92 %	88 %	87 %	87 %	89 %	89 %	89 %	75 %	76 %	75 %
	RF-Citation	94 %	91 %	90 %	93 %	92 %	92 %	94 %	93 %	92 %	91 %	89 %	89 %
	RF-Associé	95 %	94 %	94 %	96 %	93 %	93 %	98 %	97 %	97 %	87 %	82 %	83 %
3 %	LP-Citation	89 %	76 %	78 %	20 %	33 %	24 %	12 %	11 %	9 %	21 %	19 %	20 %
	LP-Associé	95 %	93 %	94 %	11 %	8 %	7 %	91 %	93 %	91 %	11 %	9 %	10 %
	SCSD	95 %	95 %	95 %	95 %	95 %	95 %	99 %	98 %	98 %	90 %	88 %	89 %

OBSERVATIONS

*Retweets
Urls*

- ▶ Construction du discours par partage d'arguments favorables

Mentions

- ▶ Interpellations entre adversaires existantes, mais réponse plus probable entre alliés

Hashtags

- ▶ Majorité des mots-clés liés à un sujet plus qu'à un point de vue

Abonnés

- ▶ 2 comportements observables concernant les profils suivis
 - ▶ Sélection partisane des profils suivis
 - ▶ Suivi des alliés et des adversaires : veille ?

DÉCOUVERTE DE POINTS DE VUE SUR LES MÉDIAS SOCIAUX

- ▶ Allport, F. H. (1937). Toward a Science of Public Opinion. *Public Opinion Quarterly*, 1(1), 7--23. [\[doi\]](#)
- ▶ Brigadir, I., Greene, D., & Cunningham, P. (2015). Analyzing Discourse Communities with Distributional Semantic Models. In *Proceedings of the ACM Web Science Conference* (p. 1–10). ACM Press. [\[doi\]](#)
- ▶ Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Flammini, A., & Menczer, F. (2011). Political Polarization on Twitter. In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media* (p. 89--96). [\[lien\]](#)
- ▶ Conover, Michael D., Goncalves, B., Ratkiewicz, J., Flammini, A., & Menczer, F. (2011). Predicting the Political Alignment of Twitter Users (p. 192–199). IEEE. [\[doi\]](#)
- ▶ Fraiser, O., Cabanac, G., Pitarch, Y., Besancon, R., & Boughanem, M. (2017a). Uncovering Like-minded Political Communities on Twitter. In J. Kamps, E. Kanoulas, & M. de Rijke (Éd.), *International Conference on the Theory of Information Retrieval (ICTIR)*, Amsterdam, 01/10/17-04/10/17. ACM. [\[doi\]](#)
- ▶ Fraiser, O., Cabanac, G., Pitarch, Y., Besancon, R., & Boughanem, M. (2017b). Uncovering Like-minded Political Communities on Twitter. In J. Kamps, E. Kanoulas, & M. de Rijke (Éd.) (p. 261–264). Présenté à *International Conference on the Theory of Information Retrieval (ICTIR)*, Amsterdam: ACM. [\[doi\]](#) [\[preprint\]](#)
- ▶ Fraiser, O., Cabanac, G., Pitarch, Y., Besancon, R., & Boughanem, M. (2018a). #Élysée2017fr: The 2017 French Presidential Campaign on Twitter. Présenté à *International Conference on Weblogs and Social Media (ICWSM)*, Stanford, California, États-Unis: AAAI Press. [\[preprint\]](#)
- ▶ Fraiser, O., Cabanac, G., Pitarch, Y., Besancon, R., & Boughanem, M. (2018b). Stance Classification through Proximity-based Community Detection. Présenté à *ACM Conference on Hypertext & Social Media (HT)*, Baltimore, Maryland, USA: ACM. [\[preprint\]](#)
- ▶ Gayo-Avello, D. (2012). No, You Cannot Predict Elections with Twitter. *IEEE Internet Computing*, 16(6), 91–94. [\[doi\]](#)
- ▶ Jungherr, A., Jürgens, P., & Schoen, H. (2012). Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welp, I. M. "Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment". *Social Science Computer Review*, 30(2), 229–234. [\[doi\]](#)
- ▶ Littman, J., Wrubel, L., & Kerchner, D. (2016). 2016 United States Presidential Election Tweet Ids. Harvard Dataverse. [\[doi\]](#)

DÉCOUVERTE DE POINTS DE VUE SUR LES MÉDIAS SOCIAUX

- ▶ Makazhanov, A., & Rafiei, D. (2013). Predicting political preference of Twitter users. In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (p. 298–305). ACM Press. [\[doi\]](#)
- ▶ McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1), 415–444. [\[doi\]](#)
- ▶ Pennacchiotti, M., & Popescu, A.-M. (2011). Democrats, republicans and starbucks aficionados: user classification in twitter. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (p. 430). ACM Press. [\[doi\]](#)
- ▶ Pew Research Center. (2016). News Use Across Social Media Platforms. [\[lien\]](#)
- ▶ Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Patil, S., Flammini, A., & Menczer, F. (2011). Truthy: mapping the spread of astroturf in microblog streams. In Proceedings of the 20th international conference companion on World wide web (p. 249--252). ACM Press. [\[doi\]](#)
- ▶ Saez-Trumper, D. (2014). Fake tweet buster: a webtool to identify users promoting fake news on twitter. In Proceedings of the 25th ACM conference on Hypertext and social media (p. 316–317). ACM Press. [\[doi\]](#)
- ▶ Sang, E. T. K., & Bos, J. (2012). Predicting the 2011 Dutch Senate Election Results with Twitter. In Proceedings of the Workshop on Semantic Analysis in Social Media (p. 53–60). Avignon, France: Association for Computational Linguistics. [\[lien\]](#)
- ▶ Sunstein, C. R. (2009). Republic. Com 2. 0. Princeton University Press. [\[lien\]](#)
- ▶ Tauzin, A. (2018). Combien d'utilisateurs des réseaux sociaux en 2018 en France. [\[lien\]](#)
- ▶ Thonet, T. (2017, novembre). Modèles thématiques pour la découverte non supervisée de points de vue sur le Web (Thèse de doctorat). Université Paul Sabatier, Toulouse, France. [\[lien\]](#)
- ▶ Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welp, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media.
- ▶ Vosoughi, S., Mohseni, M., & Roy, D. (2017). Rumor Gauge: Predicting the Veracity of Rumors on Twitter. *ACM Transactions on Knowledge Discovery from Data*, 11(4), 1–36. [\[doi\]](#)