LE PROJET POLITOSCOPE

David Chavalarias
Director of the Paris IdF Complex Systems Institute
CNRS Research director @ EHESS/CAMS
http://chavalarias.com

Journées Nationales de la Science Ouverte
6 Dec. 2018

Origine du projet

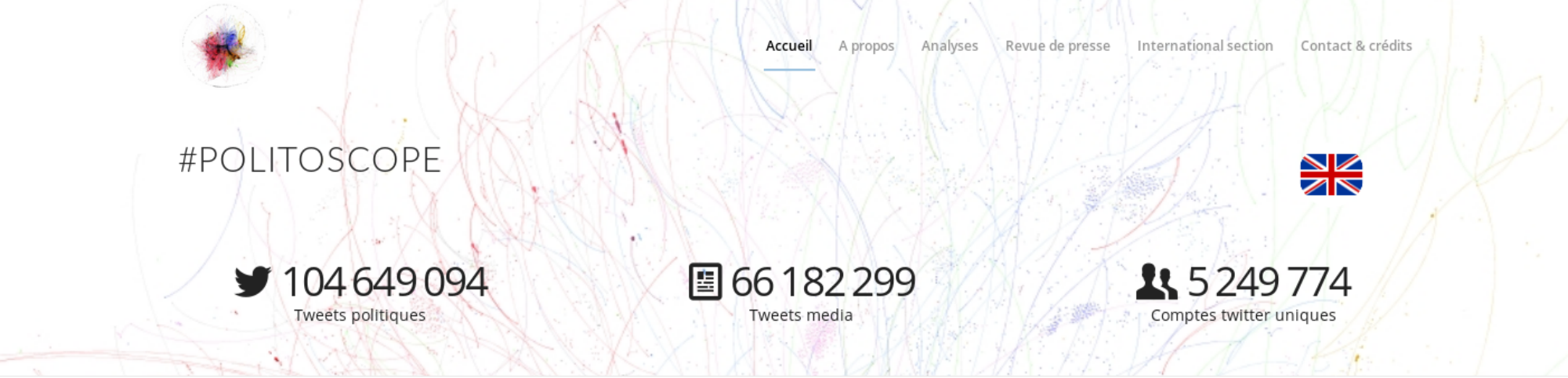
- Il y a des raisons théoriques qui lient numérisation des interactions sociales via le web et les réseaux sociaux aux changements radicaux des dynamiques du débat public et les possibilité de manipulation (Brexit, Trump, Bolsonaro, etc.),
- le Politoscope a pour but d'accompagner les recherches sur le sujet par la constitution de corpus de données et d'environnement de travail adaptés,
- Le premier objet d'étude a été la présidentielle de 2017.

Origine du projet

- Il y a des raisons théoriques qui lient numérisation des interactions sociales via le web et les réseaux sociaux aux changements radicaux des dynamiques du débat public et les possibilité de manipulation (Brexit, Trump, Bolsonaro, etc.),
- le Politoscope a pour but d'accompagner les recherches sur le sujet par la constitution de corpus de données et d'environnement de travail adaptés,
- Le premier objet d'étude a été la présidentielle de 2017.

Origine du projet

- Il y a des raisons théoriques qui lient numérisation des interactions sociales via le web et les réseaux sociaux aux changements radicaux des dynamiques du débat public et les possibilité de manipulation (Brexit, Trump, Bolsonaro, etc.),
- le Politoscope a pour but d'accompagner les recherches sur le sujet par la constitution de corpus de données et d'environnement de travail adaptés,
- Le premier objet d'étude a été la présidentielle de 2017.



Capture de données en flux

Étude "présidentielle":
60M de tweets ; 2.4M de comptes

(With N. Gaumont and M. Panahi)

http://politoscope.org

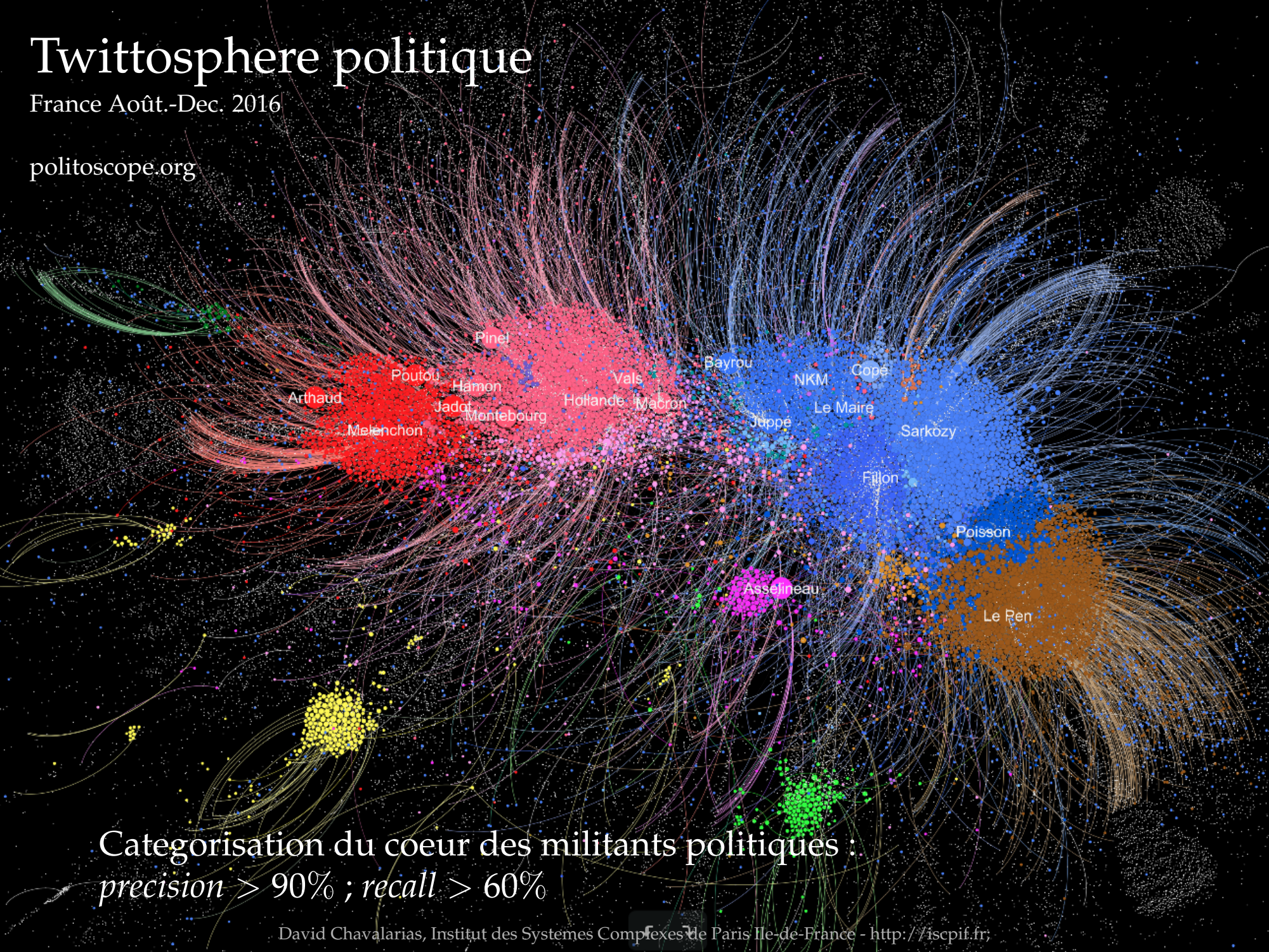
Infrastructure Multivac.iscpif.fr

Gaumont, N., Panahi, M., Chavalarias, D., forthcoming. Reconstruction of the socio-semantic dynamics of political activist Twitter networks - Method and application to the 2017 French presidential election. (2018) PLoS ONE.

Twittosphere politique

France Août.-Dec. 2016

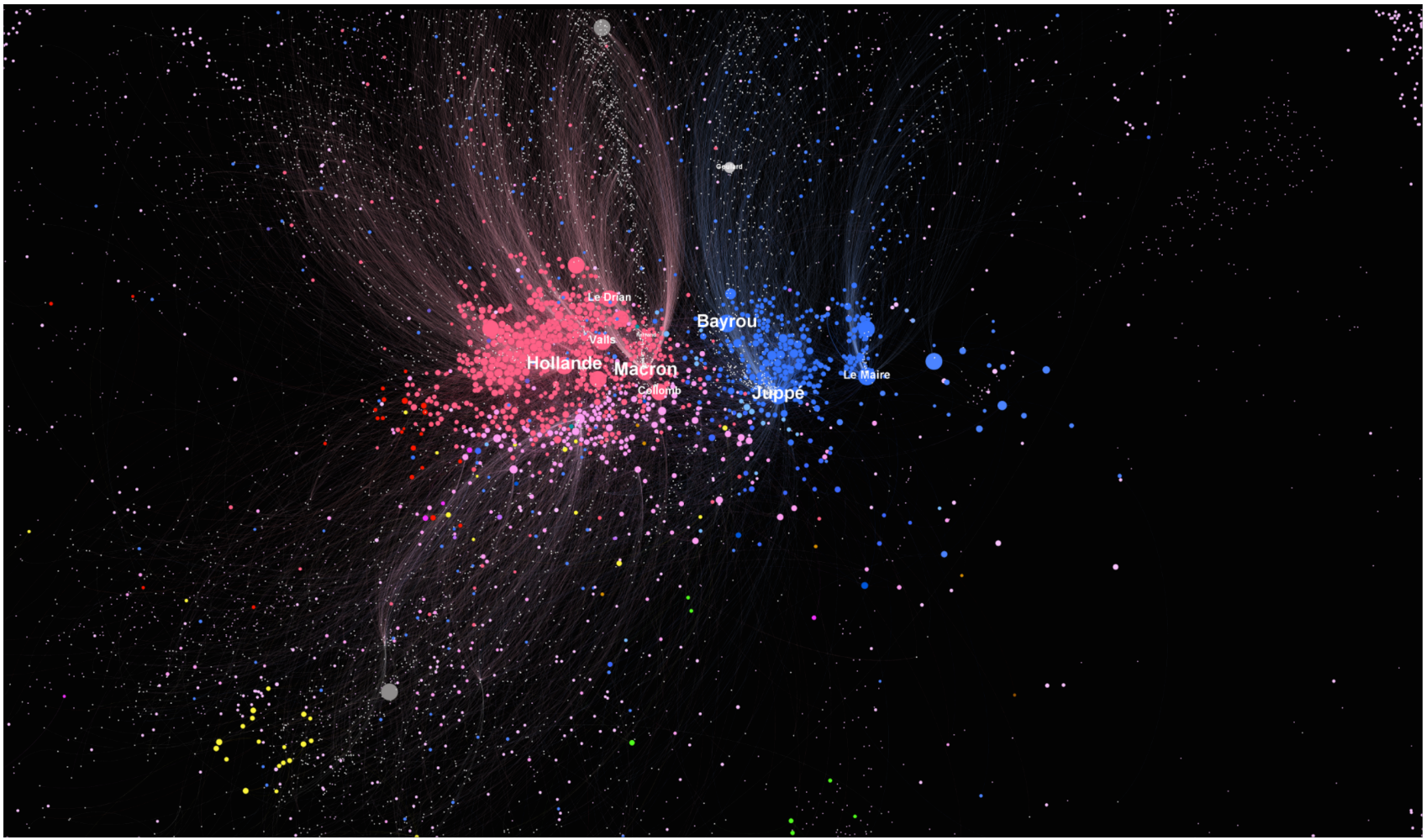
politoscope.org



Categorisation du coeur des militants politiques :
precision > 90% ; *recall* > 60%

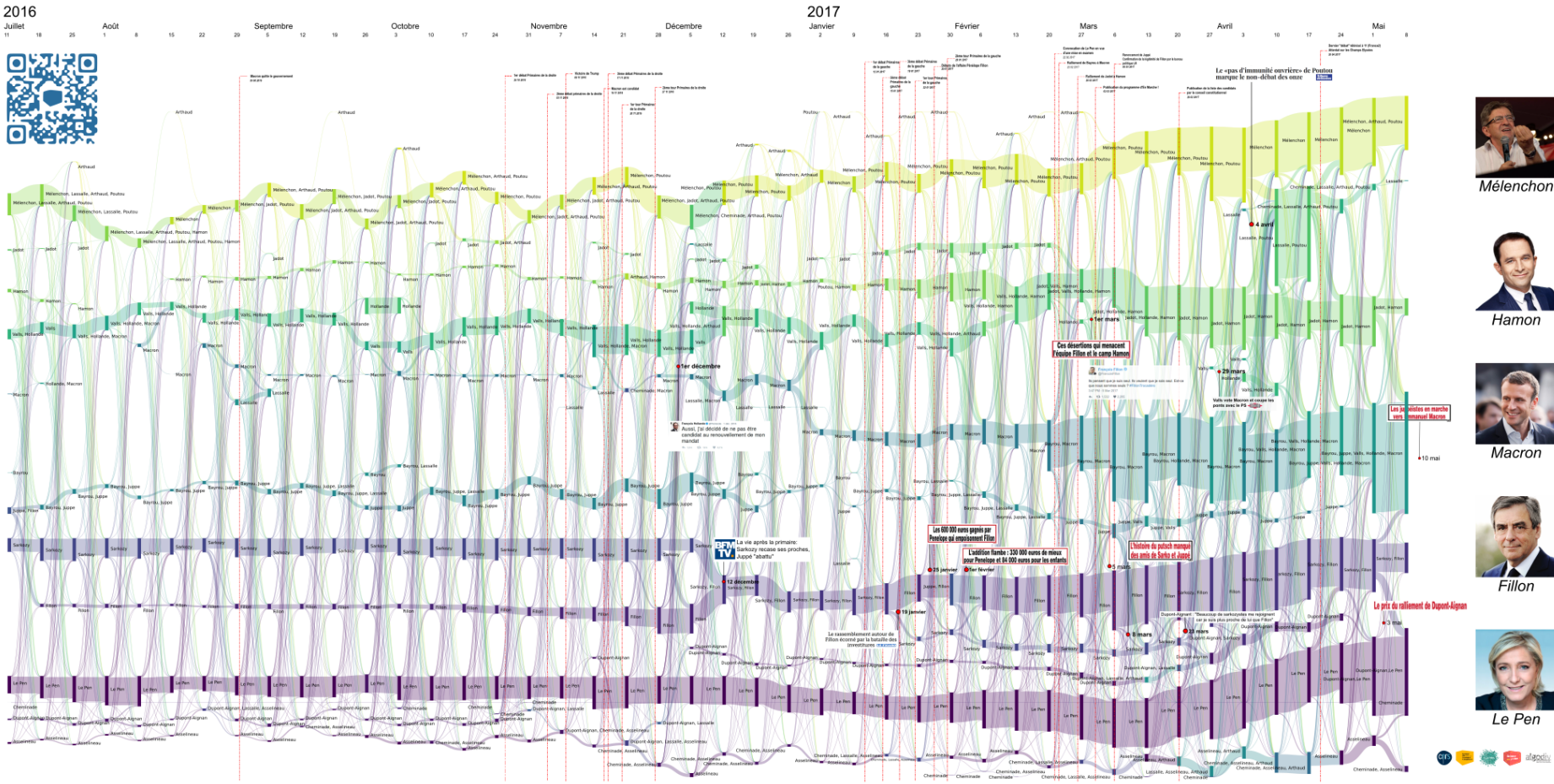
David Chavalarias, Institut des Systemes Complexes de Paris Ile-de-France - http://iscpif.fr

Evolution des communautés



Comptes qui ont été au moins une fois dans la communauté Macron au cours des trois derniers mois de la campagne.

Evolution of the French political landscape

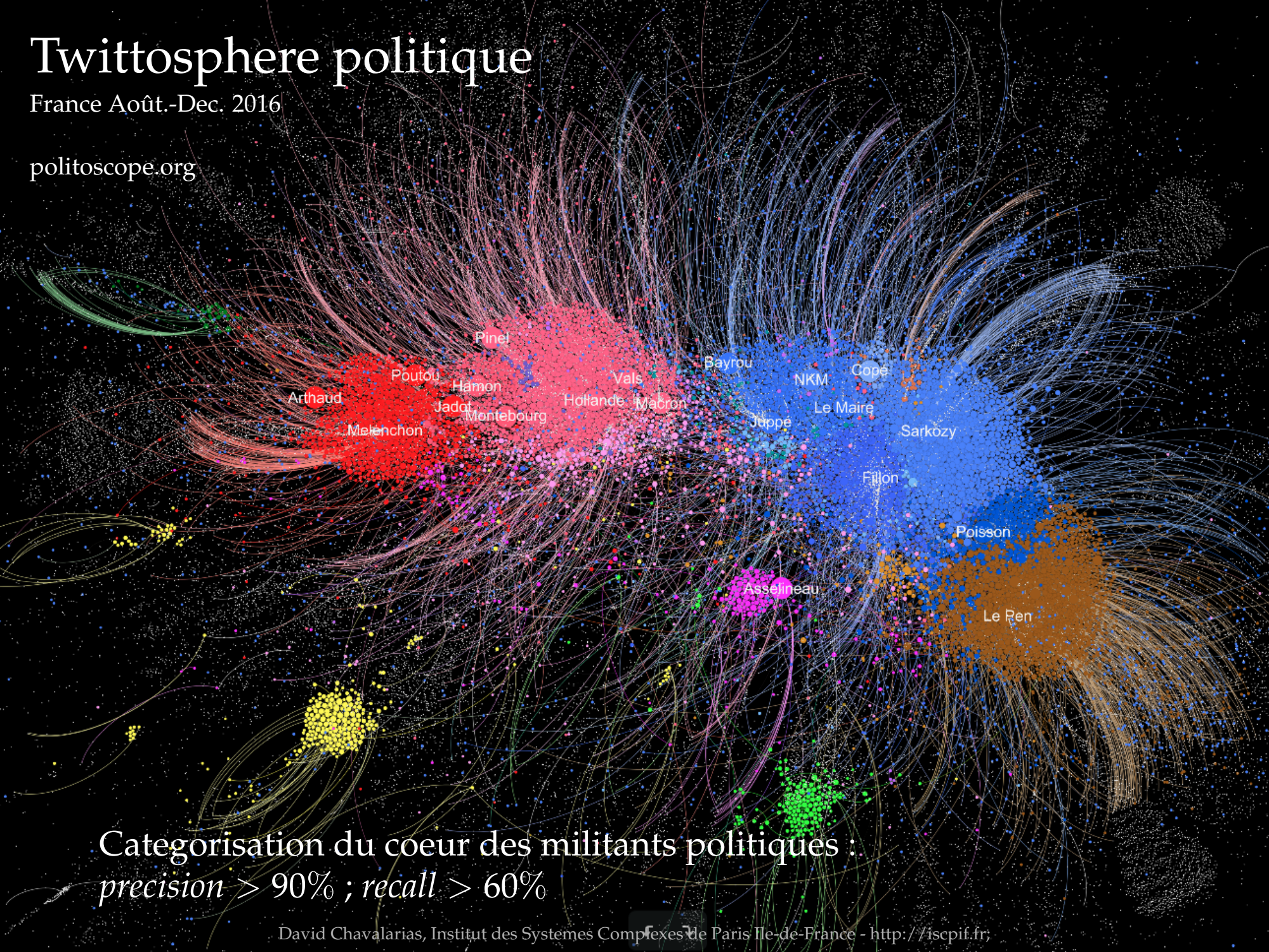


Noe Gaumont, Maziyar Panahi, David Chavalarias (2018) Reconstruction of the socio-semantic dynamics of political activist Twitter networks – Method and application to the 2017 French Presidential election. PLoS ONE.

Twittosphere politique

France Août.-Dec. 2016

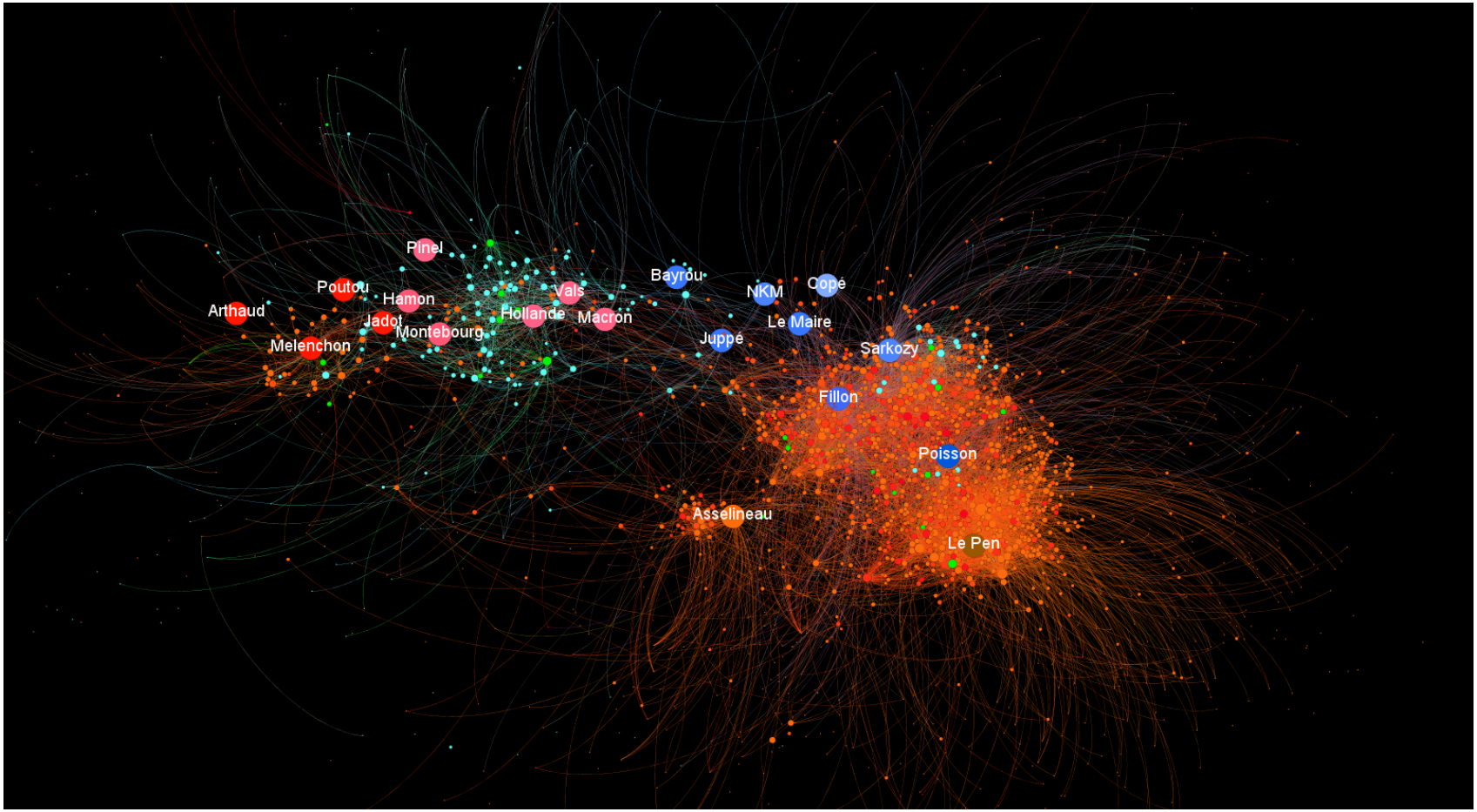
politoscope.org



Categorisation du coeur des militants politiques :
precision > 90% ; *recall* > 60%

David Chavalarias, Institut des Systemes Complexes de Paris Ile-de-France - http://iscpif.fr

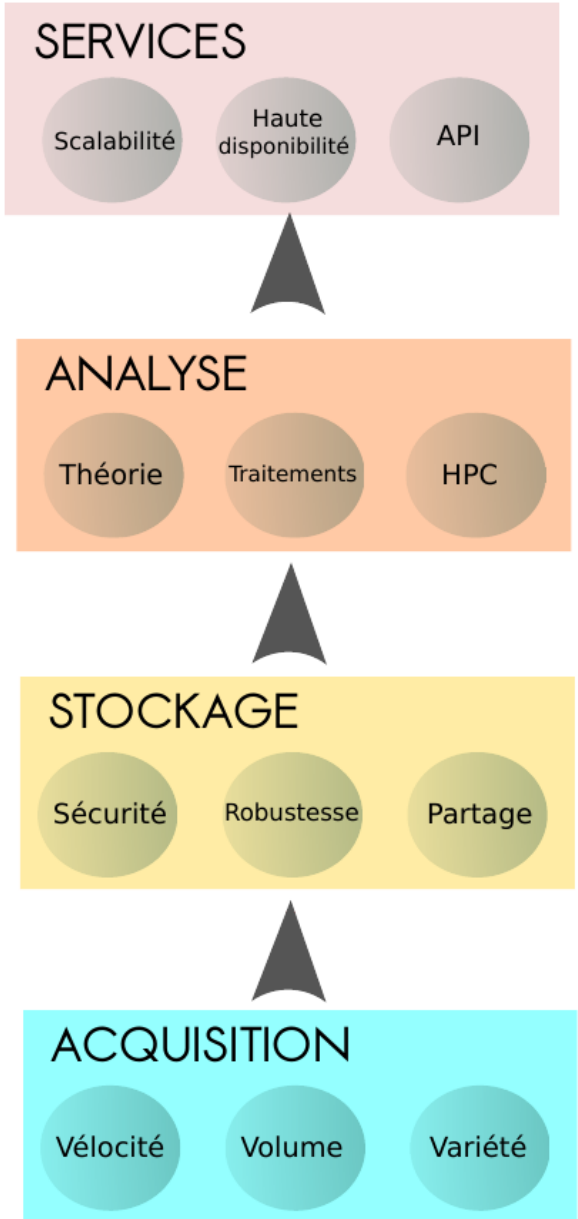
Circulation des fausses informations

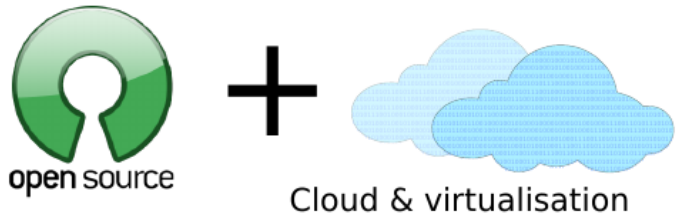


En orange : Comptes ayant relayé au moins une fausse information référencée par le Decodex ; *en bleu :* comptes ayant relayé au moins une fois un correctif du Decodex.

Gaumont, N., Panahi, M., Chavalarias, D., forthcoming. Reconstruction of the socio-semantic dynamics of political activist Twitter networks - Method and application to the 2017 French presidential election. (2018) PLoS ONE.

Statégie science ouverte de l'ISC-PIF





Les problèmes de la recherche sur les données des réseaux sociaux

- Ce type de données ne peut être écarté par la recherche en raison de leur très fort impact sociétal,
- Ce ne sont pas des données ouvertes, mais seulement des données publiques avec restrictions,
- Les chercheurs sont liés aux TofU des plateformes et aux évolutions de leurs APIs,

Les problèmes de la recherche sur les données des réseaux sociaux

- Ce type de données ne peut être écarté par la recherche en raison de leur très fort impact sociétal,
- Ce ne sont pas des données ouvertes, mais seulement des données publiques avec restrictions,
- Les chercheurs sont liés aux TofU des plateformes et aux évolutions de leurs APIs,

Les problèmes de la recherche sur les données des réseaux sociaux

- Ce type de données ne peut être écarté par la recherche en raison de leur très fort impact sociétal,
- Ce ne sont pas des données ouvertes, mais seulement des données publiques avec restrictions,
- Les chercheurs sont liés aux TofU des plateformes et aux évolutions de leurs APIs,

Les problèmes de la recherche sur les données des réseaux sociaux

- Ce type de données ne peut être écarté par la recherche en raison de leur très fort impact sociétal,
- Ce ne sont pas des données ouvertes, mais seulement des données publiques avec restrictions,
- Les chercheurs sont liés aux TofU des plateformes et aux évolutions de leurs APIs,

Comment fait-on de la science ouverte sur de tels sujets ? De par leur nature, ces données, ou une certaine forme d'agrégation ne devraient-elles pas être considérées comme un bien commun, accessibles à la recherche ?

Merci de votre attention

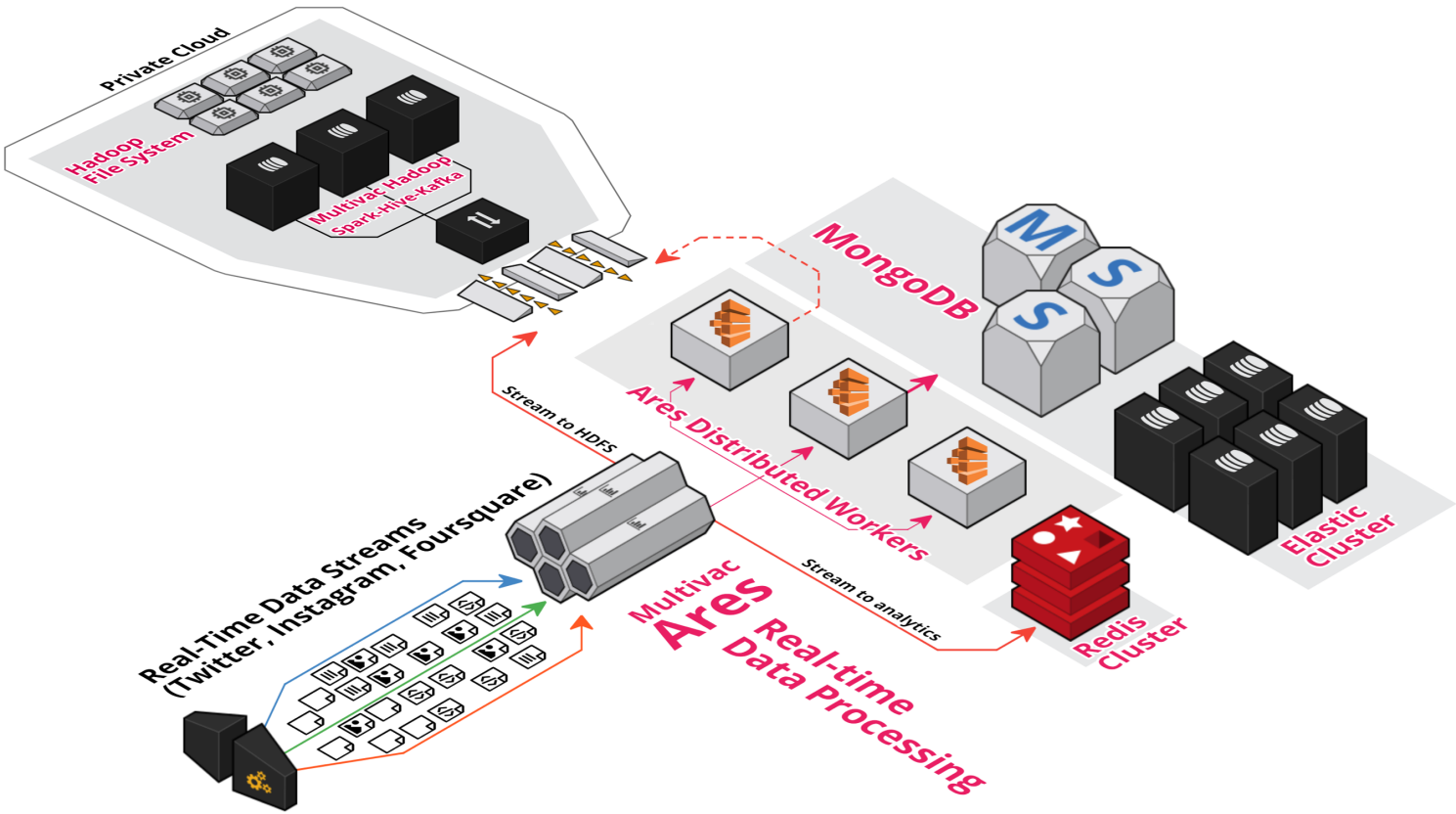
David CHAVALARIAS
Web : http://chavalarias.com
david.chavalarias@iscpif.fr

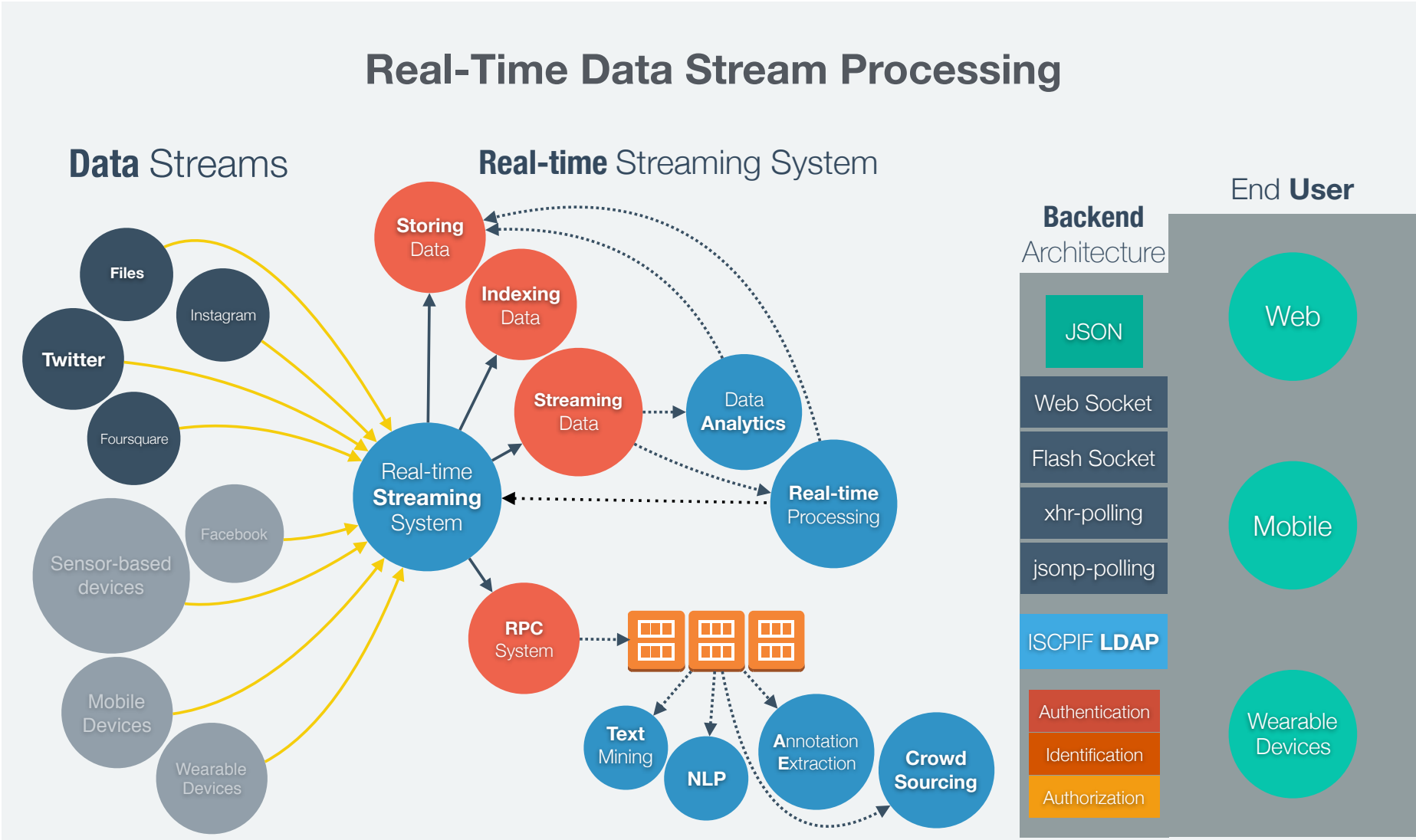
En collaboration avec :
Noé Gaumont & Maziyar Panahi

Visit us !
iscpif.fr

Les coulisses du Politoscope : Multivac.iscpif.fr

Maziyar Panahi : Hybride solution for real-time social networks analysis



Les coulisses du Politoscope : Multivac.iscpif.fr

Maziyar Panahi : Hybride solution for real-time social networks analysis