# Les mots/maux du Plan S

Journées nationales de la Science Ouverte
4-6 décembre 2018

**Frédérique Bordignon, DirDoc, École des Ponts ParisTech, France**
**Marianne Noel, LISIS, UPEM, CNRS, INRA, ESIEE Paris, France**
**Guillaume Cabanac, IRIT, Université de Toulouse, France**

**Objectifs :** caractériser « l'activisme »/le militantisme autour du thème de l'Open Access, décrypter la controverse liée au Plan S.

**Hypothèse :** le Plan S (ensemble de 10 principes) joue le rôle d'élément déclencheur de conversations dans des espaces numériques disjoints.

**Questions :** où retrouve-t-on les principes du Plan S ? Quels discours, préconceptions, points de vue, etc peuvent être révélés par ce type d'enquête ?

**Méthode :** analyse lexicale des échanges captés sur Twitter, extraction et spatialisation des termes.

## 3 CORPUS, 1 TEXTE FONDATEUR, 1 LETTRE OUVERTE

### 1 Corpus de tweets

Collecte des tweets du 4/09/18 au 4/11/18 en 3 passes (transcrites en 8 requêtes sur Twitter) pour délimiter le sujet et éviter le bruit :

- une *pure OA* (idée : pendant cette période, si ça parle OA, ça parle plus probablement Plan S)
- une "plan S" + une longue liste de termes liés à la publication scientifique (idée : avoir de façon certaine des tweets sur le Plan S, impossibles à récupérer sans croiser avec ces mots)
- une "#OA" + la même liste de termes (idée : profiter du hashtag OA, mais impossible à exploiter tout seul parce qu'il ramène trop de bruit).

**5 997 tweets** dédoublonnés

**Data Miner** pour la collecte
**Twint** pour caractériser les tweetos (nombre de followers)

### 2 Corpus de textes cités dans les tweets

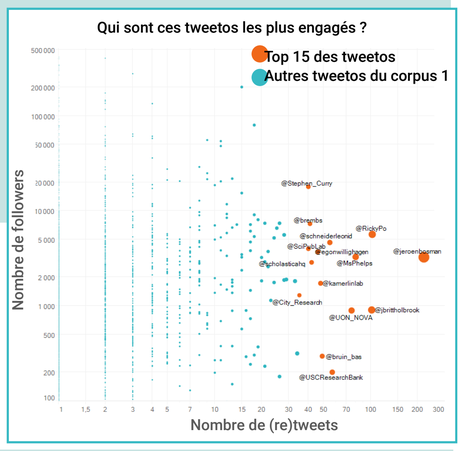
- Identification des URLs citées dans les tweets + conversion des URLs raccourcies (**longurl.eti.pw**)
- Élimination des liens vers des articles scientifiques (DOI, URL d'éditeurs etc…)
- Collecte manuelle des textes dans **Evernote** et conversion en tableur avec une règle **Zapier** (**Evernote** to **Google Sheets**)

**162 textes** dédoublonnés

### 3 Corpus de tweets du Top 15 des tweetos

- Identification des tweetos qui ont le plus tweeté ou retweeté dans le premier corpus
- Constitution du corpus de tweets du Top 15 des tweetos

**1 035 tweets** dédoublonnés




### 4

Le texte du Plan S :
https://www.scienceeurope.org/coalition-s

### 5

Le texte de l'Open Letter (pétition contre le plan S) :
https://sites.google.com/view/plansopenletter/open-letter

## CE QU'A DÉCLENCHÉ LE PLAN S DEPUIS SEPTEMBRE 2018

**Ce qui n'était pas (explicitement) dans le Plan S mais qui se retrouve au cœur des débats**

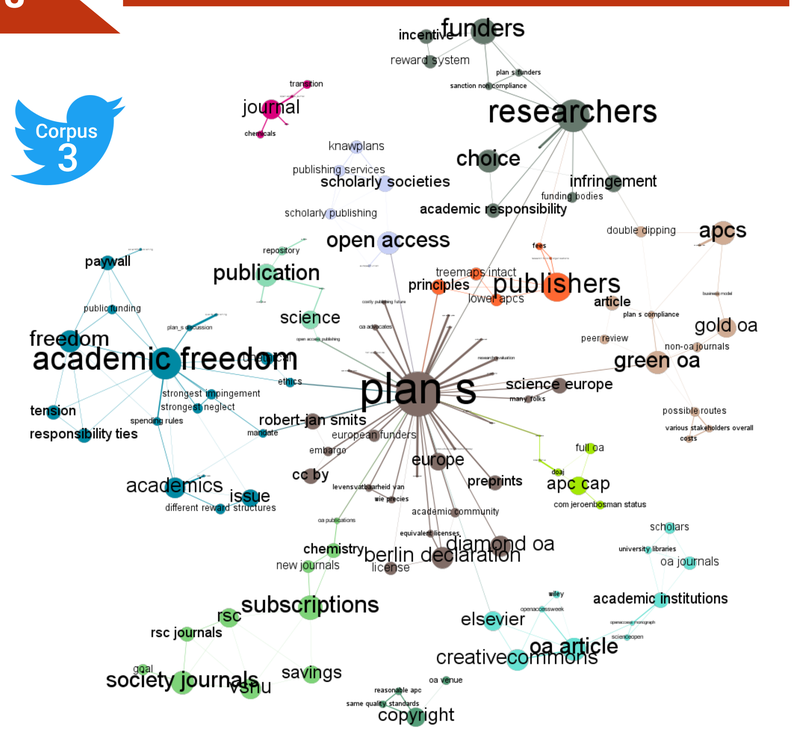
**Comparaison des concepts entre les corpus (lissés avec un label générique)**

Uniquement ceux qui représentent au moins 1% des occurrences dans au moins 1 des corpus.

| | Plan S | Tweets | Top15 tweetos | Articles | Open Letter |
|---|---|---|---|---|---|
| publishers | | 4% | 3% | 5% | |
| open science | | 2% | 1% | 1% | |
| subscriptions | | 1% | 2% | 1% | 1% |
| paywall | | 1% | 0% | 1% | |
| APCs | | 1% | 2% | 2% | 2% |
| scholarly societies | | 1% | 2% | 0% | 3% |
| freedom | | 1% | 5% | 1% | 1% |
| gold oa | | 1% | 1% | 1% | 0% |
| costs | | 0% | 0% | 1% | 1% |
| green oa | | 0% | 3% | 1% | 0% |
| preprints | | 0% | 1% | 0% | 0% |
| ban | | | | | 2% |
| risk | | | | | 1% |
| peer review | | 0% | 1% | 1% | 1% |
| diamond oa | | 0% | 2% | 0% | |
| impact | | 0% | | 0% | 3% |
| funders | 11% | 2% | 2% | 3% | 1% |
| plan s compliance | 9% | 1% | 2% | 1% | 0% |
| open access platforms | 5% | | 0% | 0% | |
| implementation | 5% | 0% | 0% | 1% | 0% |
| funds | 3% | 0% | | 1% | |
| open access publication fees | 2% | | | 0% | |
| public grants | 2% | | | 0% | |
| quality | 2% | | 0% | | 2% |
| books | 2% | 1% | 1% | 2% | |
| CC-BY | 2% | 1% | 3% | 1% | 0% |
| task force | 1% | | | | |
| incentives | 1% | | | 0% | 0% |
| sanction | 1% | | 0% | 0% | 0% |
| berlin declaration | 1% | 0% | 1% | 0% | |
| APCs cap | 1% | 0% | 1% | 0% | |
| copyright | 1% | 0% | 1% | 1% | |
| repositories | 1% | 1% | 0% | 2% | |
| hybrid model | 1% | 1% | 1% | 1% | 1% |
| libraries | 1% | 1% | 1% | 1% | |

**Termes (nœuds) dans les tweets des Top 15 tweetos, spatialisés par représentation conditionnelle (arcs), colorés par détection de communautés entre termes**

Corpus 3

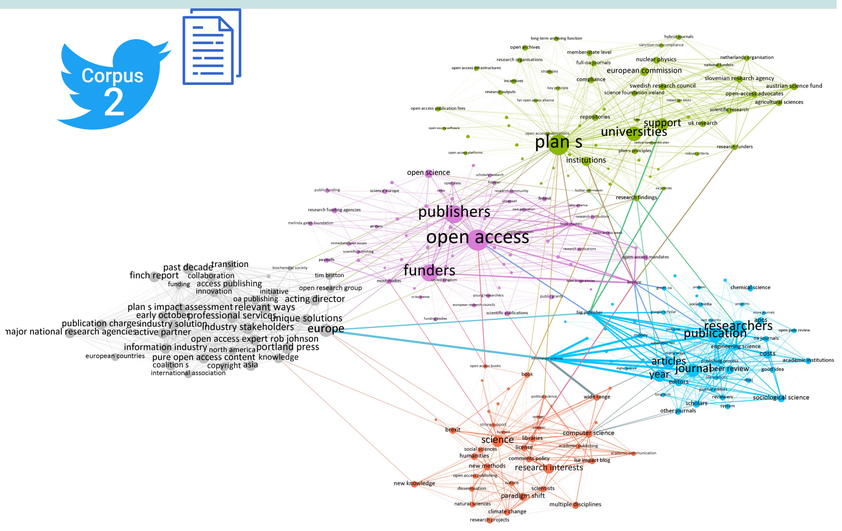



**Analyse**

- Les échanges d'experts actifs comme révélateurs de la controverse
- Les cercles concentriques du Plan S
- Les mots de la controverse : *academic freedom, choice, subscriptions, society journals*

**Termes (nœuds) dans les textes tweetés, spatialisés par représentation conditionnelle (arcs), colorés par détection de communautés entre termes**

Corpus 2




**Analyse**

- Centre : re-polarisation sur les acteurs (*publishers* et *funders*) dans le contexte (*open access*)
- Périphérie : contexte historique (*Finch report, past decade*) et économique (gris), impact sur l'activité scientifique (orange) et sur le process (bleu)

## CONCLUSIONS ET PERSPECTIVES

**Résultats :**

- Éprouver des méthodes numériques pour détricoter une forme d'activisme en ligne
- Nécessité de recourir à des matériaux complémentaires aux tweets (blogs/déclarations/policies mobilisés pour argumenter, compléter, convaincre)

**Limites :**

- Courte période d'observation (2 mois) alors que la controverse perdure (open letters, pétitions, tours d'horizon ou articles d'opinion, etc.)
- Analyse à chaud à approfondir (avec des enquêtes de terrain en cours)

**Perspectives :**

- Caractérisation des acteurs : qui suit qui ?, qui produit du contenu nouveau ?, qui va au contact des autres ?
- Rapprocher ce mouvement actuel des initiatives militantes passées (pétition en Suède 2010 lancée par la chimiste C. Moberg)