

Annonces de politiques monétaires: allier R et
recherche reproductible pour une analyse
linguistique longitudinale

Thierry Warin, HEC Montréal, WCFIA (Harvard U.), CIRANO
&
William Sanger, Polytechnique Montréal, CIRANO

R @ Québec - Université Laval - 25-26 mai 2017


Déroulement de la présentation


1. Sciences de données et plateforme technologique à HEC
Montréal & CIRANO
2. Recherche en sciences sociales
 - ▶ analyse textuelle des discours de la Banque Centrale Européenne

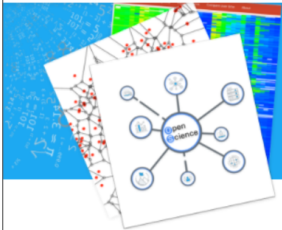
OpenScience: data science & affaires internationales

www.openscience.nuance-r.com

Open Science @ HEC Montreal

powered by 

[Team](#) [Open Government](#) [Open Finance](#) [Research](#) 







HEC Montreal's Data Science Initiative in International Business

A NEW "OS" FOR DATA SCIENCE. This website is dedicated to the International Business field of research with Data Science as the core methodology. We also value reproducible research, and open source developments.

Open Science Toolkit

Navigate through our set of tools and servers dedicated to the production of academic content using only open source code.

 <p>Mondo: International Data In partnership with CRNO, millions of observations useful for International Business.</p>	 <p>Open Research Available to HEC Montreal students, Dr Hector's Research platform.</p>	 <p>Open Teaching Dr Hector's course website, with some open MOOCs.</p>	 <p>Data Science for International Business HEC Montreal's Data Science committee: www.dshb.org</p>
---	--	---	--

OpenScience: data science & affaires internationales



OpenScience: data science & affaires internationales



Code collaboratif

- Git pour étudiants et chercheurs

OpenScience: data science & affaires internationales



Dr. HECtoR

- Plateforme de cours en ligne utilisant R (MOOC, moodle)

OpenScience: data science & affaires internationales



Cartographie

- Partenariat avec Nüance-R
- GeoNode

OpenScience: data science & affaires internationales



Mondo.international

- Partenariat avec CIRANO
- Données internationales

OpenScience: data science & affaires internationales



Console RStudio

- Console en ligne

OpenScience: data science & affaires internationales



RConnect

- Partenariat avec RStudio
- Gestion de documents dynamiques en ligne

OpenScience: data science & affaires internationales



Econométrie

- Logiciel d'économétrie en ligne

OpenScience: data science & affaires internationales



Recherche bibliographique

- Partenariat avec la Leibniz Foundation
- Conférences et articles scientifiques en économie

OpenScience: data science & affaires internationales



ds4ib.org

- Comité étudiant en science de données

Nüance-R: data science & sciences sociales



Nüance-**R**

Historique du projet

- ▶ 2014:
 - ▶ Développement d'une plateforme de données internationales au CIRANO (www.mondo.international)
- ▶ 2015: Harvard University et développement de Nüance-R
 - ▶ Weatherhead Center for International Affairs [Gary King]
 - ▶ TH. Chan School of Public Health
- ▶ 2016:
 - ▶ Lancement de Dr. HECtoR à HEC Montréal (www.drhector.org)
 - ▶ Sciences de données pour les affaires internationales (M.Sc.)
- ▶ 2017:
 - ▶ Lancement de Open Science à HEC Montréal (www.openscience.nuance-r.com)
 - ▶ Partenariats avec RStudio, la Leibniz Foundation

Analyse de la politique monétaire de la Banque
Centrale Européenne à travers ses conférences
de presse - une perspective de science de
données

Cartographie



Figure 1: Francfort, siège de la BCE; et pays (bleu) de la zone euro

Banque Centrale Européenne

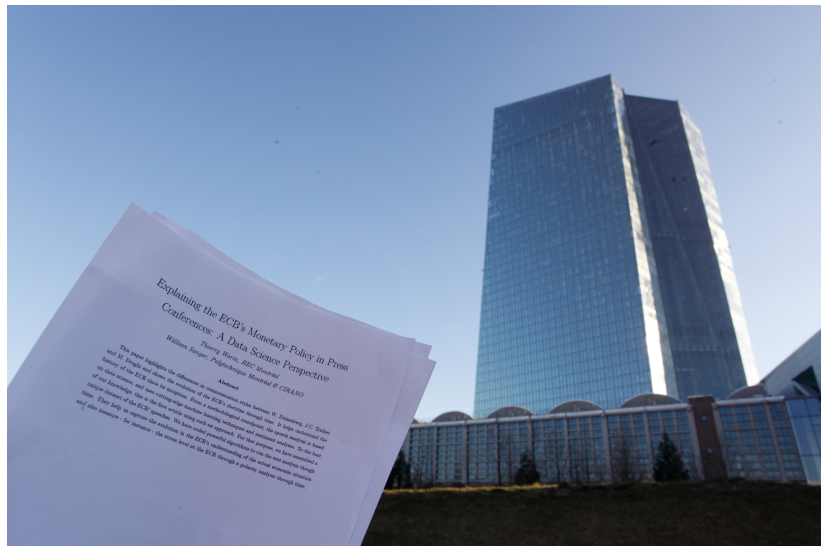


Figure 2: Siège de la BCE, Francfort

Résumé du papier

Contexte

BCE - jeune institution (20 ans); crise de l'euro en 2008 & Brexit

Revue de littérature

Répondre à la crise de la zone euro, une approche
{quali/quantitative}tative des banquiers centraux

Question de recherche

Comment la communication institutionnelle diffère de la réaction à l'actualité touchant la zone euro?

Données

203 discours et 3501 réponses aux journalistes analysés

Méthodologie

Science de données: analyse de sentiment & de polarité

Résultats

Perceptions différentes entre les deux types de discours, quantifiables à travers le temps

Contexte

La Banque Centrale européenne est une jeune institution (naissance en 1998), avec comme défi d'établir sa crédibilité (par rapport à des institutions plus anciennes: FED, Banque Centrale d'Angleterre).

Au cours de son histoire, elle a connu trois présidents successifs:

- ▶ Wim Duisenberg (NL)
- ▶ Jean-Claude Trichet (FR)
- ▶ Mario Draghi (IT)

La crise financière de 2008 toucha de plein fouet l'Europe puis se répercuta à travers les différents pays européens. Comment un tel choc se reflète dans les prises de position de la BCE?

Revue de littérature

Cartographie de la littérature

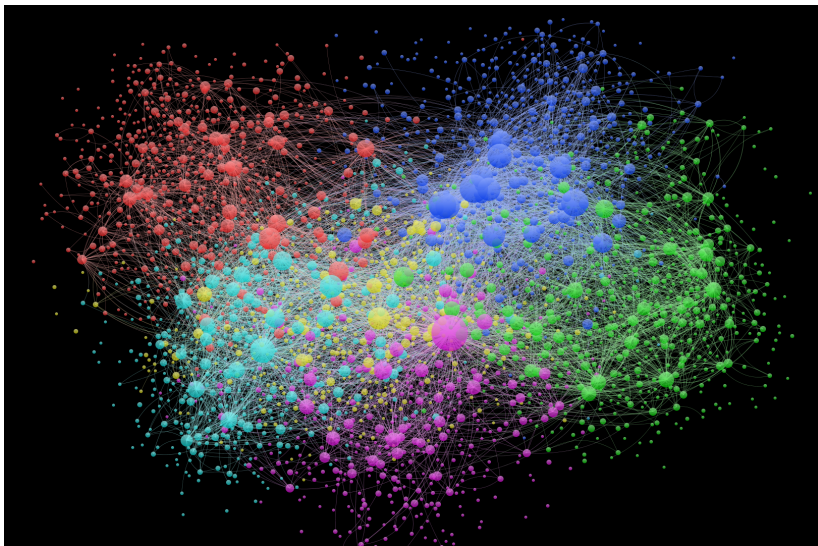


Figure 3: Panorama de la littérature - 36'000 articles

Revue de littérature - Avant la crise

Comment gérer les asymétries dans l'eurozone?

- ▶ Est-ce que la BCE peut gérer les larges asymétries dans l'eurozone? (Brissimis & al., 2008; Gambacorta, 2003; Aksoy & al., 2002; Arnorld, 2006)
- ▶ Anticipations à partir des perceptions (Berger & al., 2009)
- ▶ Mesures politiques dans le contexte du Pacte de Stabilité et de Croissance (Clausen & al. 2005; Eichengreen 2005)

Politique monétaire

- ▶ Politique optimale en Europe (Adolfson et al, 2008, Ascari, 2007); Gestion des excédents de réserves et un débalancement global (Bindseil & al., 2006; Bini Smaghi, 2007)
- ▶ Indicateurs monétaires :
 - ▶ Crédibilité (Breedon & Hume, 2007) et confiance (Wälti, 2012)
 - ▶ Indépendance de la banque centrale (Weymark, 2007)

Revue de littérature - Après la crise

- ▶ Gestion de crise et réponse de la BCE (Pronobis 2014; Breuss & others 2016)
- ▶ Politiques et outils monétaires non-standards (Darracq-Paries & De Santis 2015; Dreger & Wolters 2015; Eser & Schwaab 2016)
- ▶ Analyse institutionnel comparée (Cukierman, 2013)
- ▶ Lessons à tirer de la crise (Drakos & Kouretas, 2015)
- ▶ Réponses à une crise de la dette souveraine suite à une crise financière (Roman & Bilan, 2012)
- ▶ Réponse différente de la BCE et de la FED par rapport au développement des marchés financiers (Hoffmann, 2013)
- ▶ Changement à la règle de Taylor en présence de taux d'intérêt égaux à 0 (Belke & Klose, 2013)

Plusieurs auteurs appellent à une intégration financière plus soutenue en Europe (Dabrowski, 2010).

Revue de littérature - Approche qualitative

La Banque Centrale Européenne envoie des signaux de différentes façons, notamment à travers sa communication.

L'interprétation des décisions de politiques monétaires met en évidence l'importance d'une dimension qualitative (*soft dimension*):

- ▶ Anticipations des marchés (Andersson & al., 2006; Kholodilin & al. 2009)
- ▶ Niveau de transparence perçue de la BCE (Crujisen & Eijffinger 2010; Rosa & Verga 2005)

Revue de littérature - Approche qualitative

Analyse des communications et des discours de la BCE

- ▶ Conférences de presse et conséquences potentielles sur les marchés financiers (Amaya & Filbien, 2015)
- ▶ Utilisation de mots de références (Blinder & al., 2008)
- ▶ Codification des discours et des propositions de la BCE (Ehrmann & Fratzscher, 2007; Jansen & Jakob Haan, 2009)

Stratégies de communication des institutions

- ▶ FED vs. ECB vs. Banque d'Angleterre concernant les protocoles de communication (Ehrmann and Fratzscher, 2007)
- ▶ *Timing* des communication (Ehrmann and Fratzscher, 2007)
- ▶ Styles de communication des présidents de la BCE (Rosa & Verga, 2005, Fourçans & Vranceanu, 2007)
- ▶ Comparaison entre la FED et la BCE (Fourçans & Vranceanu, 2007): la FED possède un rôle plus "activiste" que la BCE

Revue de littérature - Approche quantitative

- ▶ Transformer l'information qualitative en une échelle ordonnée (Rosa & Verga, 2007)
- ▶ Codifier les attentions de la BCE et de la FED (Rosa, 2008)
 - ▶ Les annonces de la FED sont plus directescourtes (15)
- ▶ Niveau de transparence des institutions pour prédire les futures actions de la BCE (Rosa & Verga, 2005, Rosa & Verga, 2005, Rosa, 2009)
- ▶ Niveau de similarité entre les discours de la BCE augmente au fil du temps (Amaya & Filbian, 2015)

Revue de littérature

Sujet	Auteurs
Avant la crise Gérer les asymétries dans l'eurozone	Brissimis & Skitida, 2008; Gambacorta, 2003; Aksoy & al., 2002; Arnorld, 2006; Berger & al., 2009; Clausen & Wohltmann 2005; Eichengreen 2005
Politique monétaire	Adolfson et al, 2008, Ascari, 2007; Bindseil & al., 2006; Bini Smaghi, 2007; Breedon & Hume, 2007; Wälti, 2012; Weymark, 2007
Après la crise Réponses à la crise	Pronobis 2014; Breuss & others 2016; Darracq-Paries & De Santis 2015; Dreger & Wolters 2015; Eser & Schwaab 2016; Cukierman, 2013; Drakos & Kouretas, 2015; Roman & Bilan 2012; Hoffmann, 2013; Belke & Klose, 2013
Approche qualitative des banques centrales Importance d'une <i>soft dimension</i>	Andersson & al., 2006; Kholodilin & al. 2009; Cruijssen & Eijffinger 2010; Rosa & Verga 2005
Analyse des communications et des discours de la BCE	Amaya & Filbien, 2015; Blinder & al., 2008; Ehrmann & Fratzscher, 2007; Jansen & Jakob Haan, 2009
Stratégies de communication des institutions	Ehrmann & Fratzscher, 2007a; Ehrmann & Fratzscher, 2007b; Rosa & Verga, 2005; Fourçans & Vranceanu, 2007
Évaluations quantitative de la communication de la BCE	Rosa & Verga, 2007; Rosa, 2008; Rosa & Verga, 2005a, Rosa & Verga, 2005b, Rosa, 2009; Amaya & Filbian, 2015

Figure 4: Résumé de la littérature

Question de recherche et hypothèses

Question de recherche

Quelles sont les différences de perception entre le discours institutionnel de la BCE et les réponses des Présidents de la BCE?

Hypothèses

1. La première partie des discours est associée à la communication institutionnelle de la BCE
2. Lors des questions des journalistes, les Présidents de la BCE réagissent aux événements touchant la zone euro
3. La science des données permet de mettre en valeur ces différences structurelles de perception des discours

Données

Discours

Chaque mois, la BCE explique sa politique monétaire. Entre 1998 et 2016, la BCE a tenu 203 discours divisés en deux parties:

1. une première où la BCE dévoile ses positions en termes de politique monétaire: **203 discours**
2. une seconde où les Présidents de la BCE répondent aux questions des journalistes: **3501 réponses aux journalistes**

Discours

Au total, la base de données comporte 917'730 mots, soit:

- ▶ 38'703 tweets
- ▶ 10'880 génériques de Star Wars épisode IV
- ▶ Les Misérables (V. Hugo) et Guerre et Paix (L. Tolstoi)

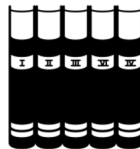


Figure 5: Équivalent de 917'730 mots

Discours: première partie

Media

- + Press releases
- + Governing Council decisions
- Press conferences
 - 2017
 - 2016**
 - 2015 - 1998
- + Monetary policy accounts
- + Speeches
- + Interviews
 - Weekly schedule
- + TV services & webcasts
- + Calendars
 - Media contacts

Introductory statement to the press conference (with Q&A)

Mario Draghi, President of the ECB,
Vitor Constâncio, Vice-President of the ECB,
Ewald Nowotny, Governor of Oesterreichische Nationalbank,
Vienna, 2 June 2016

[Jump to the transcript of the questions and answers](#)

Ladies and gentlemen, the Vice-President and I are very pleased to welcome you to our press conference. I would like to thank Governor Nowotny for his kind hospitality and express our special gratitude to his staff for the excellent organisation of today's meeting of the Governing Council. We will now report on the outcome of our meeting.

Based on our regular economic and monetary analyses, we decided to keep the **key ECB interest rates** unchanged. We continue to expect them to remain at present or lower levels for an extended period of time, and well past the horizon of our net asset purchases. Regarding **non-standard monetary policy measures**, we confirm that the monthly asset purchases of €80 billion are intended to run until the end of March 2017, or beyond, if necessary, and in any case until the Governing Council sees a sustained adjustment in the path of inflation consistent with its inflation aim. As a next step, on 8 June we will start making purchases under our **corporate sector purchase programme (CSPP)**. Moreover, starting on 22 June, we will conduct the first operation in our new series of **targeted longer-term refinancing operations**. Further information on implementation aspects of the CSPP will be released after the press conference on the ECB's website.

Find out more

- [Calendar of Governing Council meetings and press conferences](#)
- [Monetary policy decisions \(key interest rate decisions\)](#)
- [Webcasts of press conferences](#)
 - press conferences' photos



Figure 6: Première partie des discours

Discours: réponses aux questions des journalistes

We are now at your disposal for questions.

** **

Question: I might have missed it, but I did not hear just now whether you expressed a view on whether the risks to inflation are now broadly balanced or tilted to the downside, whether that has changed. Could you please let us know?

My second question is, at the accounts of your April meeting, you said that there was no conclusive evidence yet on whether there are second-round effects, and I'm wondering whether since then you have seen substantial evidence pointing in either direction.

Draghi: To the first question I'll answer, basically, reading again what is in the introductory statement: additional stimulus beyond the impetus already taken into account is expected from the monetary policy measures still to be implemented, namely the TLTRO II and the CSPP, and will contribute to further rebalancing the risks to the outlook for growth and inflation. In other words, what we are saying is that we have the June Eurosystem staff projections, and what's called the probability mass of risks over the years 2017 and 2018 has shifted slightly upward. So that is what it is. It's nothing dramatic.

On the second point, no, we haven't seen second-round effects. But at the same time, in the midst of what one would qualify as good news as far as growth and employment are concerned, we still see no substantial pressures on the wage and price-setting mechanisms, with the possible exception of Germany. It goes without saying that the Governing Council remains watchful and alert to the presence of any second-round effect, and stands ready to act using all the instruments within its mandate.

Figure 7: Seconde partie des discours



Discours

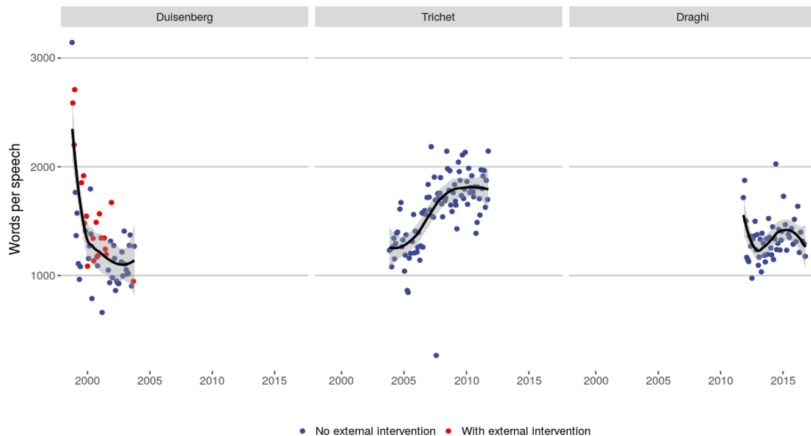


Figure 8: Longueur des discours entre 1998 et 2016

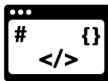
Méthodologie

Méthodologie - Plateforme technologique

Plateforme de science de données, **Nüance-R**: 3 serveurs en ligne possèdent un compilateur R, PostgreSQL, Ubuntu 16.04 LTS et d'autres logiciels.

Cette plateforme nous permet de **collecter**, créer des **visualisations** et **analyser** des données de manière efficace, rapide et robuste à l'aide d'algorithmes développés dans notre programme de recherche. Nous utilisons l'interface de navigation de RStudio.

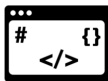
Méthodologie - Recherche reproductible



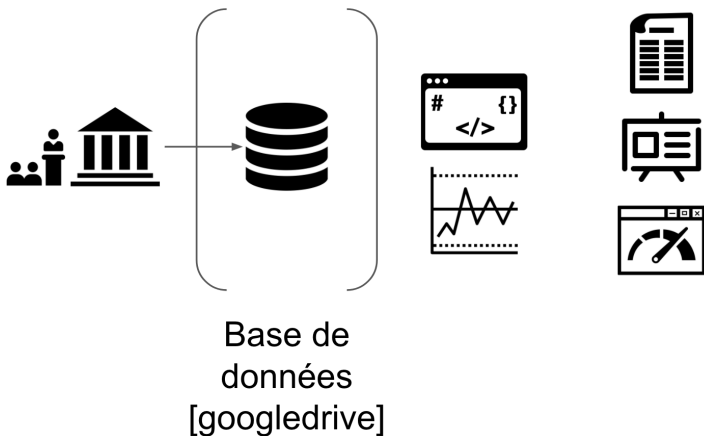
Méthodologie - Recherche reproductible



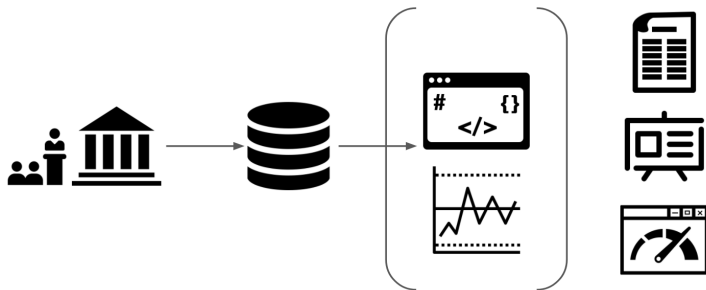
Discours
mensuels
de la BCE



Méthodologie - Recherche reproductible

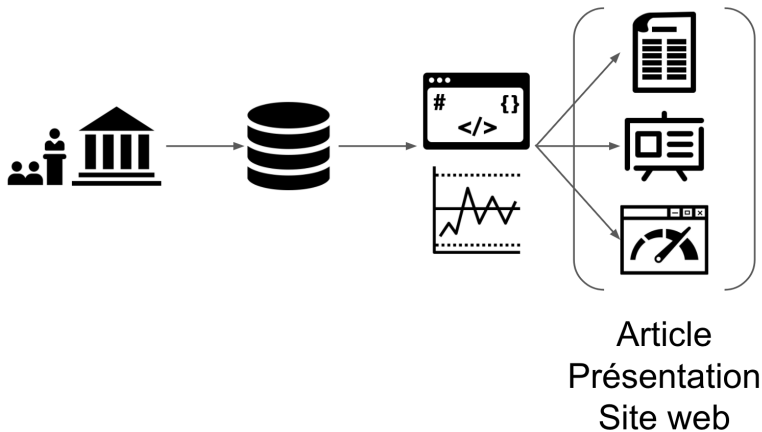


Méthodologie - Recherche reproductible

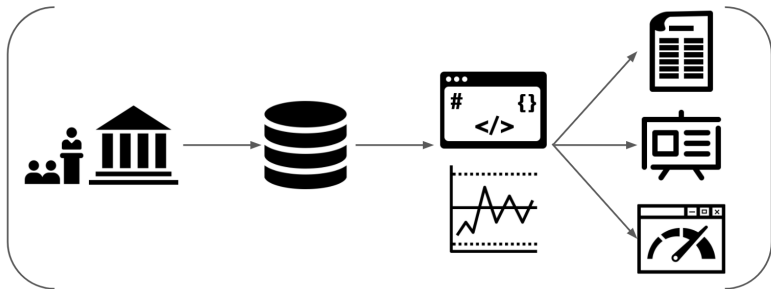


RMarkdown
Économétrie
Analyse textuelle

Méthodologie - Recherche reproductible



Méthodologie - Recherche reproductible



Méthodologie - Analyse textuelle [tm; RTextTools]

Stratégie de cadrage: pour chaque Président de la BCE, nous dissociions ses discours de ses réponses aux questions des journalistes. Par la suite, nous utilisons l'analyse lexicale (Lucas et al., 2015) et les différents packages en R (*tm*, Feinerer et al., 2008, *RTextTools*) afin de développer des algorithmes linguistiques.

Transformations textuelles

- ▶ Séparation des mots à partir des espaces
- ▶ Chaque mot est normalisé (sans majuscule)
- ▶ Élimination de "mots vides" (*of, a, the, ...*)

Modélisation des sujets

Utilisation d'expressions régulières (racine de mots) pour caractériser la fréquence d'occurrence de:

- ▶ confiance: [Cc]onfiden[tc]
- ▶ Grèce: [Gg]ree[ck]

Méthodologie - Analyse de sentiment [sentiment]

- ▶ Association de chaque mot à **six** catégories d'émotion: *colère*, *dégoût*, *peur*, *joie*, *tristesse*, *surprise*.
- ▶ Chaque catégorie est identifiée par un lexique de référence (total: 1'542 mots) afin d'obtenir des scores d'émotions tels que:

$$\left| \ln \left(\frac{nb_{emotion}}{nb_{emotion,lexique}} \right) \right| \quad (1)$$

Méthodologie - Analyse de polarité [sentiment]

- ▶ Association de chaque mot à **deux** catégories lexicales: *positive, négative*.
- ▶ Chaque catégories est identifiée par un lexique de référence (total: 6'518 mots - 2'324 positifs; 4'175 négatifs) pour obtenir un score de polarité tel que:

$$\left| \ln \left(\frac{n_{positive,negative}}{m_{polarity,lexical}} \right) \right| \quad (2)$$

Résultats

Wim Duisenberg (1998 - 2002)

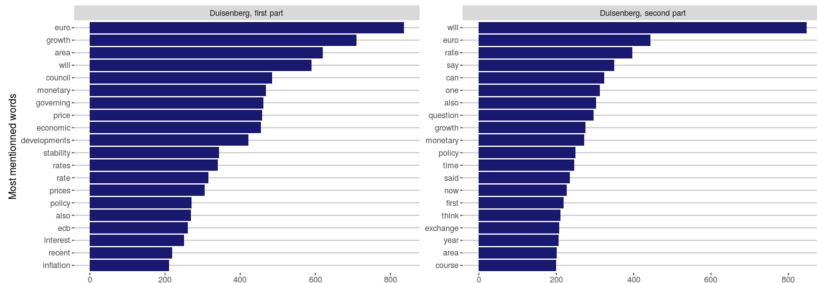


Figure 9: Mots les plus utilisés par Wim Duisenberg

Jean-Claude Trichet (2002 - 2012)

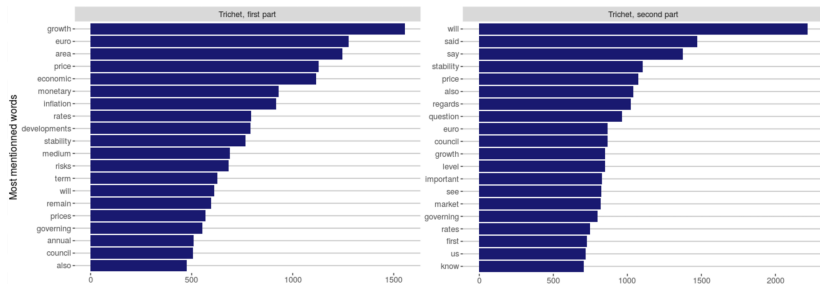


Figure 10: Mots les plus utilisés par Jean-Claude Trichet

Mario Draghi (2012 - actuel)

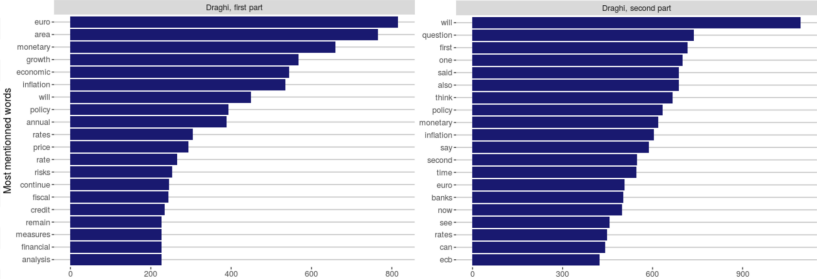


Figure 11: Mots les plus utilisés par Mario Draghi

Résultats - Notions économiques

	Wim Duisenberg	Jean-Claude Trichet	Mario Draghi
euro	14.5 (1)	13.14 (2)	15.2 (1)
croissance	12 (2)	15.6 (1)	10.5 (4)
inflation	3.7 (19)	9.4 (7)	10 (6)
risques	2.5 (30)	6.5 (12)	2.4 (54)
stabilité	5.7 (12)	7.5 (10)	4.4 (15)

Résultats - Mentions de la Grèce

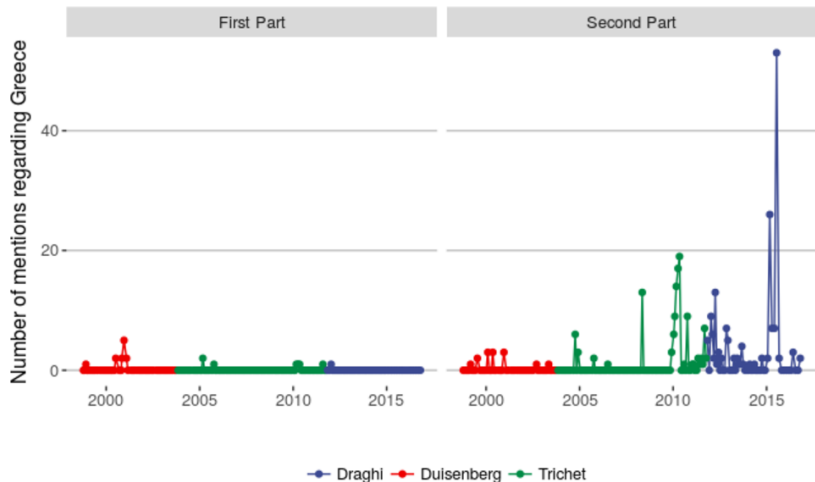


Figure 12: Mentions de la Grèce

Résultats - Confiance

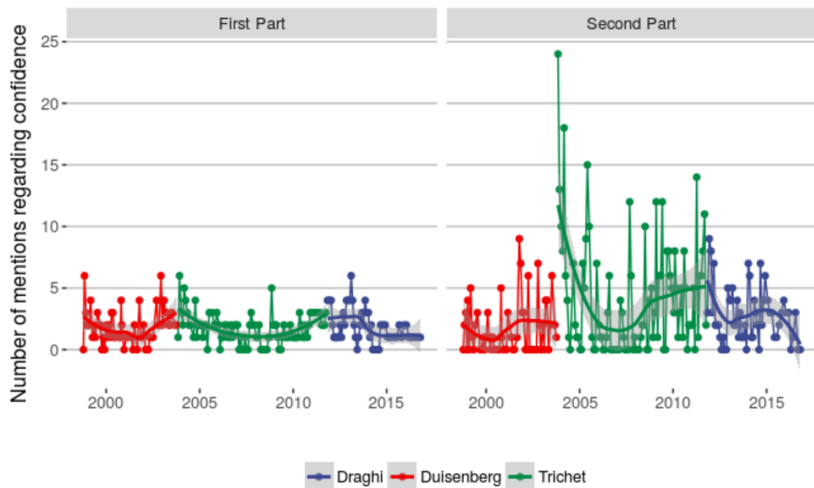


Figure 13: Confiance/confidence

Résultats - Analyse de sentiment

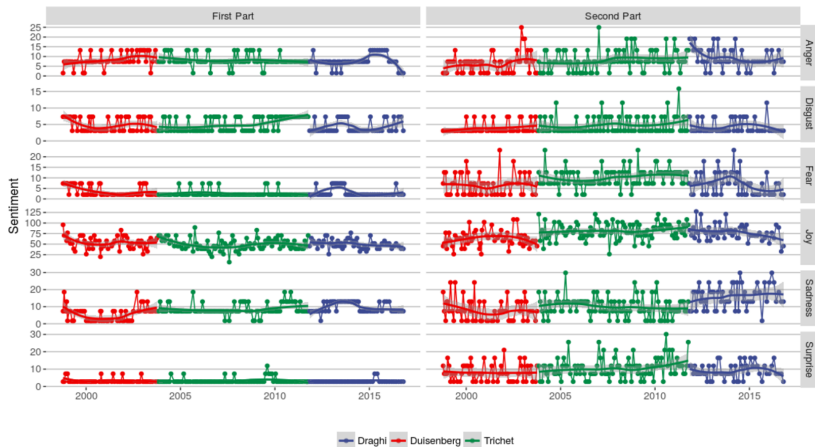


Figure 14: Évolution de l'importance des sentiments

Résultats - Analyse de polarité

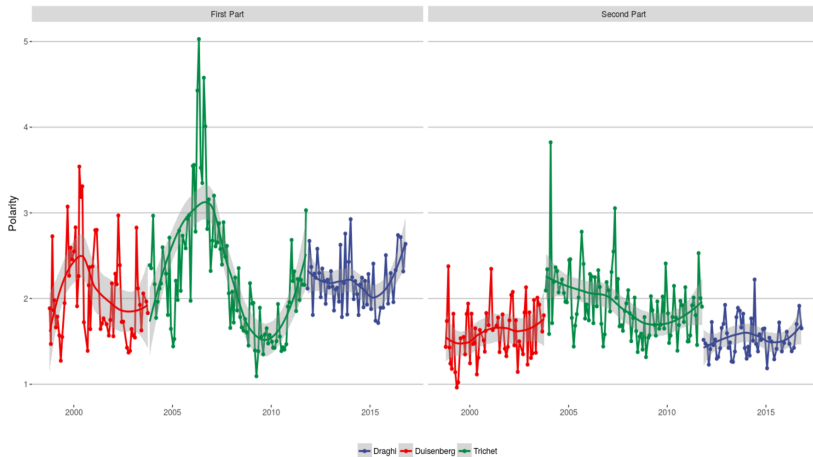


Figure 15: Évolution de la polarité des discours

Discussion et conclusion

Discussion

1. Utilisation de techniques d'analyse de texte afin de quantifier des concepts qualitatifs
2. Base de données unique, prenant en compte les questions des journalistes comme résonance des évènements présents dans l'eurozone

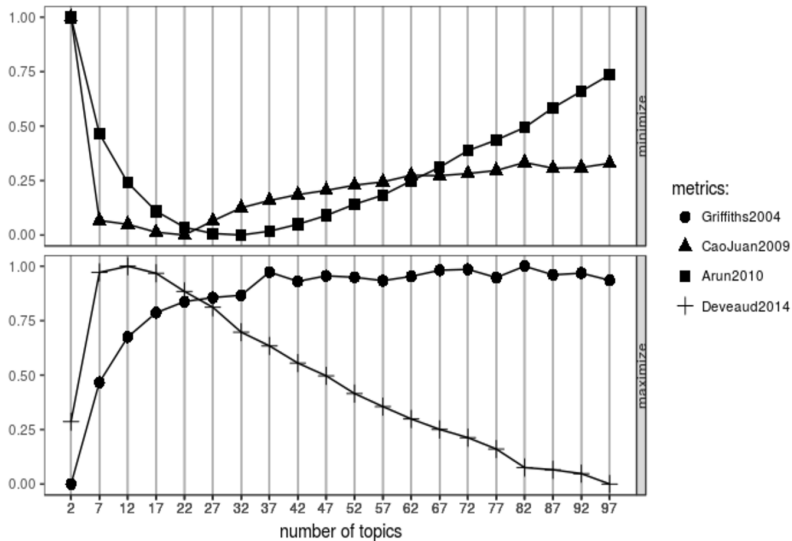
Limites

- ▶ Coupler ces indicateurs aux perceptions des marchés financiers

Prochains développements

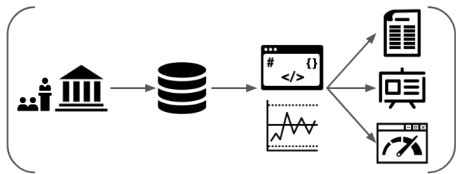
- ▶ LDA (Latent Dirichlet Allocation) afin de faire émerger les thématiques les plus importantes
- ▶ Prendre en compte les réponses aux questions des membres du Parlement Européen

LDA [ldatuning; stm]



Résumé de la présentation

Utilisation de R en sciences sociales Plateforme de sciences de données [Nüance-R & OpenScience]



Contexte

BCE - jeune institution (20 ans); crise de l'Euro en 2008 & Brexit

Résultats

Perceptions différentes entre les deux types de discours, quantifiables à travers le temps