

Événements dans le texte, événements dans le monde

Habilitation à Diriger les Recherches
Xavier Tannier

Travaux non abordés aujourd'hui

Systèmes de questions-réponses

avec V. Moriceau, B. Grau, etc.

Crawling orienté

avec C. de Groc, Syllabs

Cohésion thématique

avec C. de Groc, Syllabs

Petits outils divers

...

Co-organisation de campagnes d'évaluation

- Tâche « NLP » à INEX 2006 (avec le Queensland University of Technology)
- Tâche « QA » puis « Tweet Contextualization » à INEX 2010 à 2013 (avec LIA, LSIS, IRIT)

Activités d'encadrement

Encadrement		Année	Type
Doctorats	Béatrice Arnulphy	2009-2012	Quaero
	Clément de Groc	2010-2013	CIFRE
Post-doctorats	Rémy Kessler	2011-2012	ANR Chronolines (resp. scientifique LIMSI)
	Kiem-Hieu Nguyen	2013	ANR Chronolines
	Kiem-Hieu Nguyen	2014-2015	Digiteo ASTRE (porteur)

Au menu

- **Définition et Représentations**
 1. Un mot dans la phrase et dans le document
 2. Un schéma informationnel
 3. Une description textuelle complexe
 4. Un pic d'activité thématique
- **Enjeux**
- **L'événement dans le texte et dans le monde**
- **Bilan & Perspectives**

Définition

Définition

« On peut dire, grossièrement, que les événements sont des choses qui arrivent à un certain moment »

Davidson, 1980

Représentations

L'appel du général de Gaulle fut lancé le 18 juin 1940.

En 1944, les alliés débarquaient en Normandie.

Représentations

- 1. Un mot dans la phrase
et dans le document***
-

L'**appel** du général de Gaulle fut lancé le 18 juin 1940.

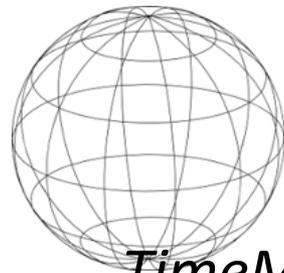
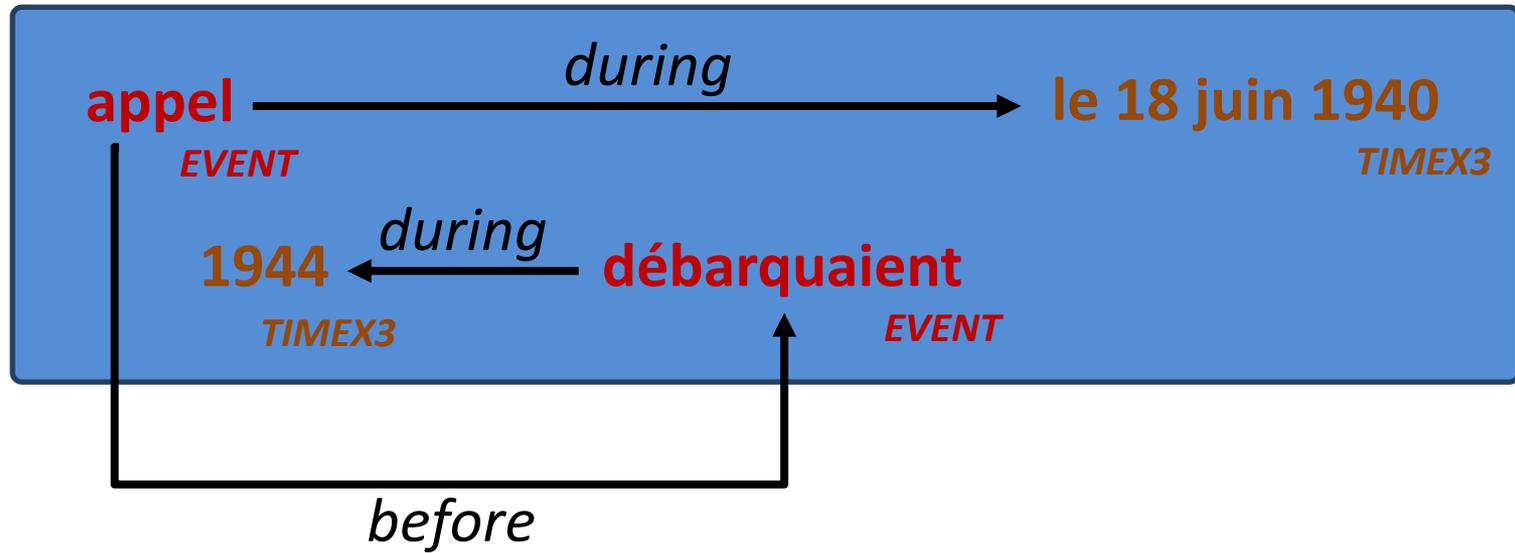
EVENT

TIMEX3

En 1944, les alliés **débarquaient** en Normandie.

TIMEX3

EVENT



TimeML



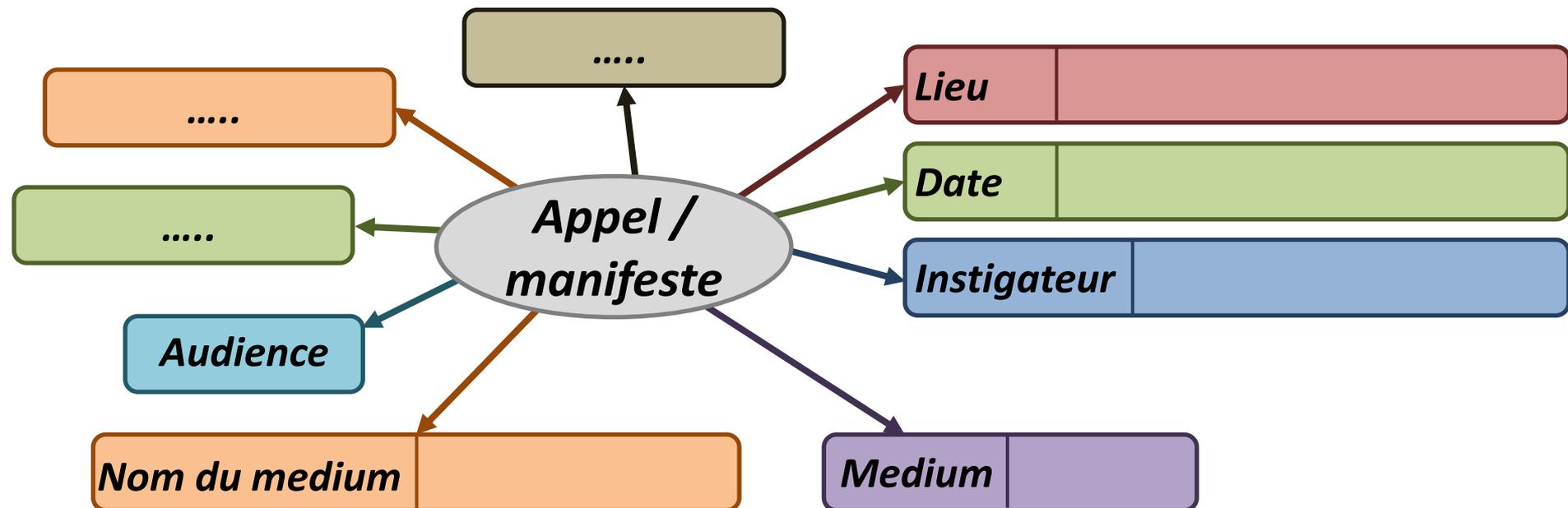
TempEval

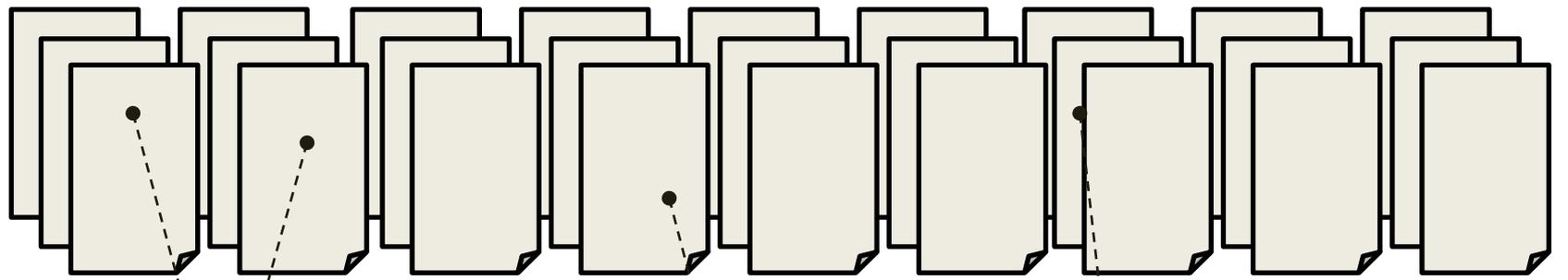
Représentations

2. Un schéma informationnel

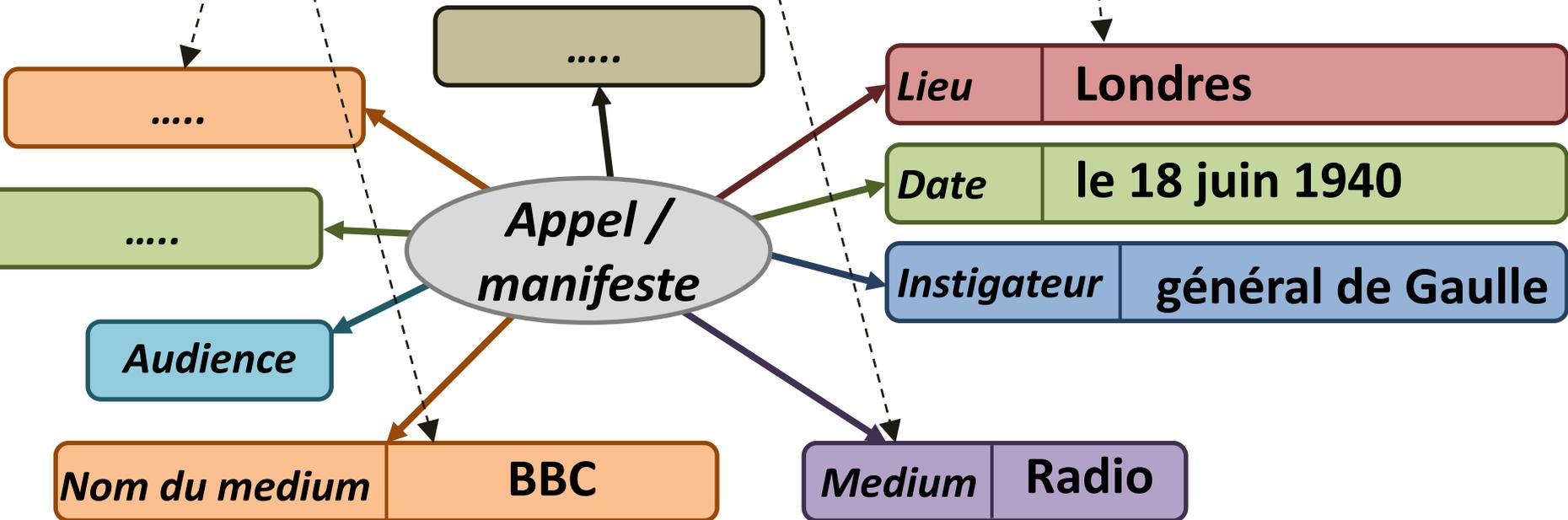
L'appel du général de Gaulle fut lancé le 18 juin 1940.

En 1944, les alliés débarquaient en Normandie.





MUC & C°



Représentations

3. Une description textuelle complexe

L'appel du général de Gaulle fut lancé le 18 juin 1940.

En 1944, les alliés débarquaient en Normandie.

(réalisé avec trucage)

Titres Du Jour

BBC : L'appel du général de Gaulle à la résistance en France



Ce 18 juin sur la radio britannique BBC, le général Charles de Gaulle, exilé depuis l'invasion de la France par l'Allemagne, a...

Le Point.com



Blocus total décrété par Hitler sur le Royaume-Uni

Le Führer allemand a mis en place un blocus total sur l'ensemble du territoire britannique, mettant à profit l'occupation récente de la...

Le Matin.com



Les colonies européennes placées sous tutelle américaine

Pour éviter qu'elles ne tombent sous contrôle allemand, les colonies françaises et néerlandaises ont été placées sous la tutelle...

Liberation.com

Monde

France

Résumé/Synthèse automatique

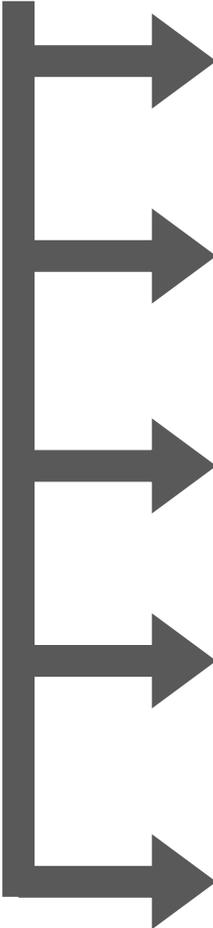
*Lorem ipsum dolor sit amet, consectetur adipiscing elit. In et justo ligula. Donec semper nisl a lorem euismod, sit amet condimentum eros luctus. Mauris gravida interdum hendrerit. Proin tincidunt lorem eros, at vulputate diam adipiscing et. Aenean porttitor turpis ut lectus sagittis elementum. **L'appel du général de Gaulle fut lancé le 18 juin 1940.** Nunc tempus mauris consequat luctus euismod.*

Nulla a eros ornare enim dictum placerat eget in sapien. Aliquam in justo pellentesque, cursus libero adipiscing, ultrices sem. Cras vel arcu egestas, posuere libero nec, bibendum libero.

Chronologies automatiques



2nde guerre mondiale



3 septembre 1939 : Invasion de la Pologne

18 juin 1940 : L'appel du général de Gaulle fut lancé

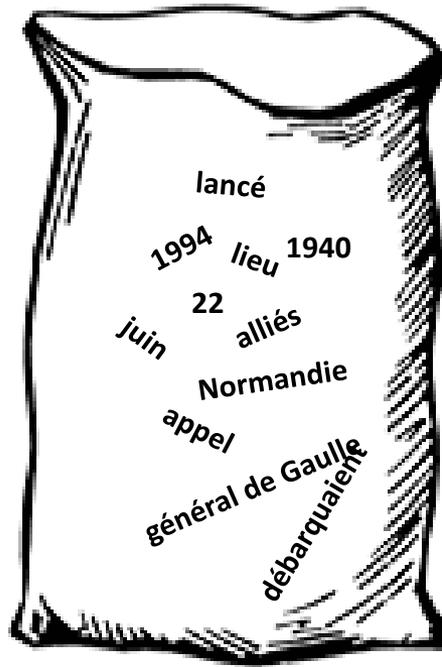
8 novembre 1942 : Les alliés débarquent en Algérie

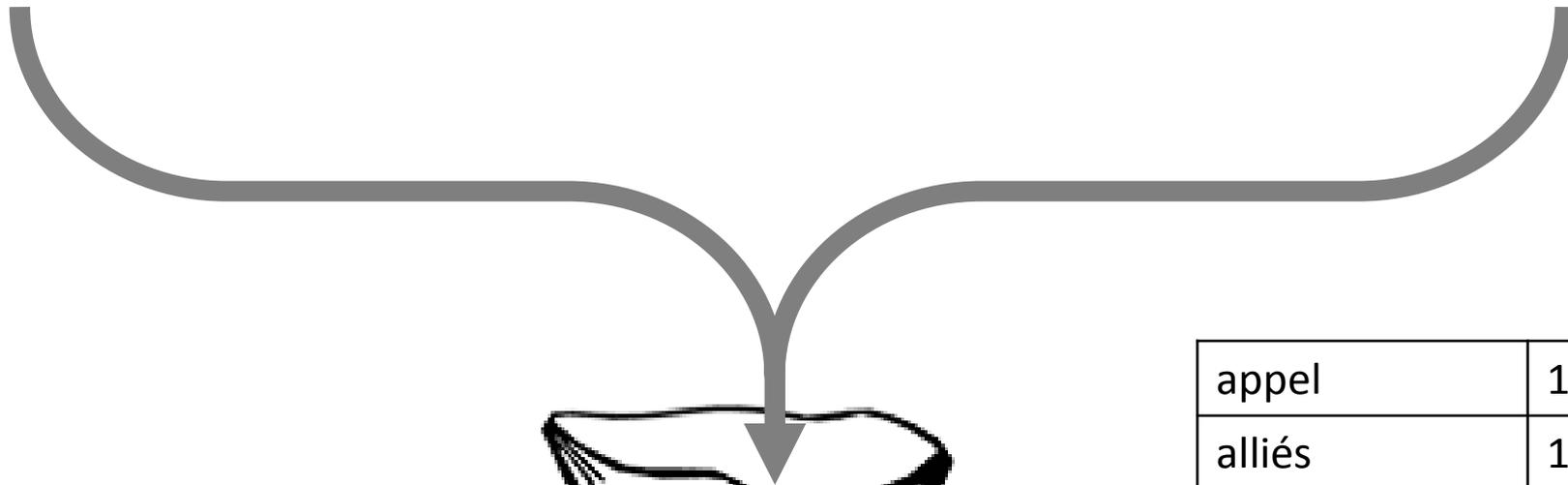
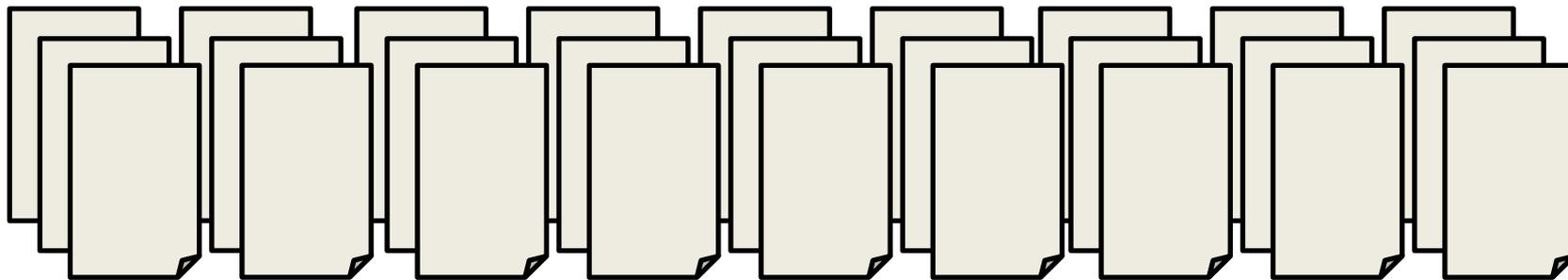
6 juin 1944 : Les alliés débarquent en Normandie

8 mai 1945 : L'Allemagne capitule

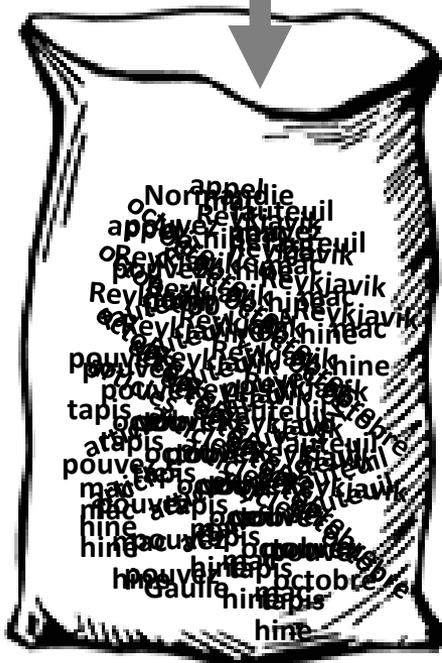
L'appel du général de Gaulle fut lancé le 18 juin 1940.

En 1944, les alliés débarquaient en Normandie.





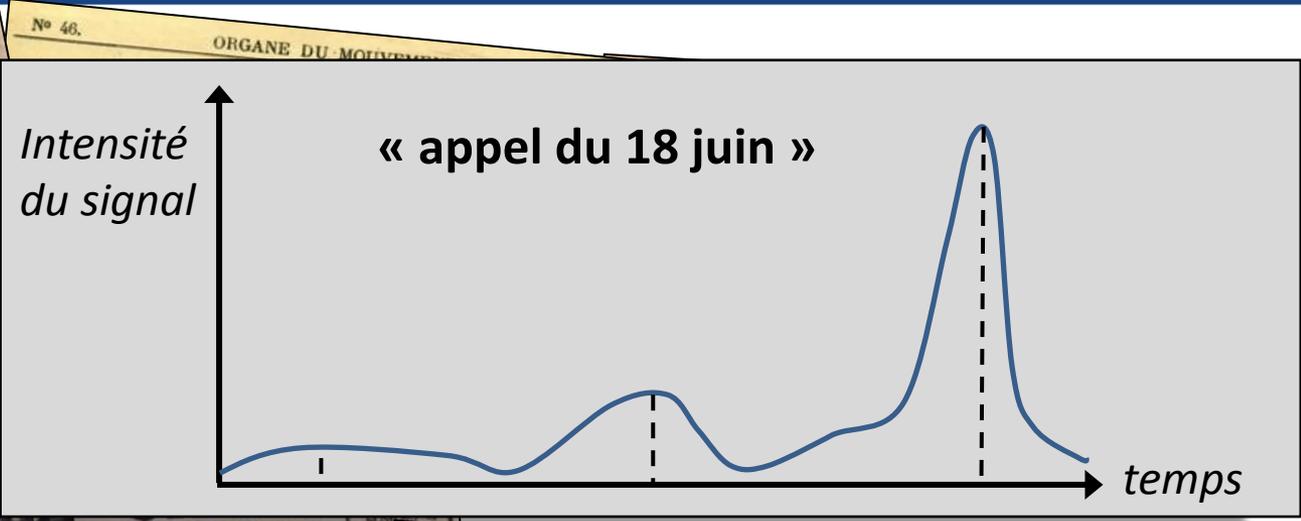
DUC
TAC
TDT
?



appel	1435
alliés	1642
débarquaient	345
Gaulle	530
général	654
lancé	38943
Normandie	134
1940	423
1944	577

Représentations

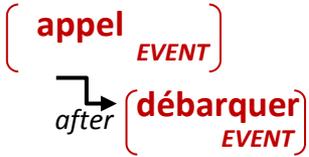
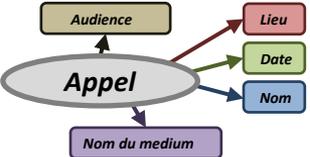
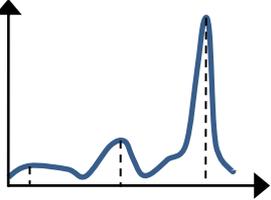
4. Un pic d'activité thématique



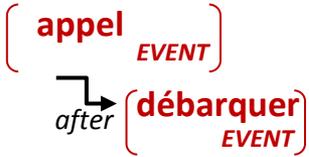
(réalisé avec trucage)

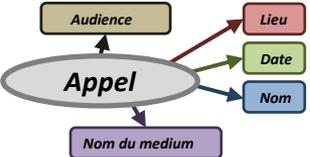
Bilan & Enjeux

Enjeux et domaines associés

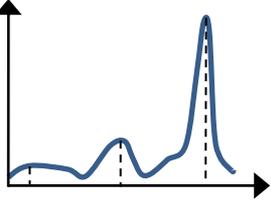
<p>1. </p>	<p>Cibler</p> <p><i>Extraction d'information</i></p>
<p>2. </p>	<p>Agréger</p> <p><i>Extraction d'information</i> <i>Bases de connaissance</i></p>
<p>3. </p>	<p>Hiérarchiser</p> <p><i>Résumé</i> <i>Recherche d'information</i></p> <p>Contextualiser</p>
<p>4. </p>	<p><i>Recherche d'information</i> <i>Détection d'anomalie</i></p>

Perspectives

1. 

2. 

3. 

4. 

Cibler

Agréger

Hierarchiser

Contextualiser

Observer l'évolution

Questions posées

Événement dans le texte et dans le monde

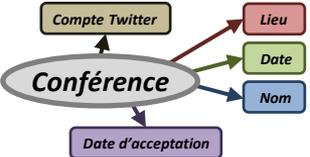
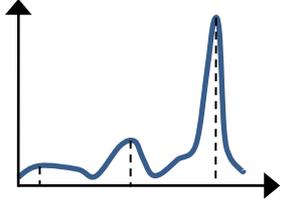
Précision et couverture dans les **gros volumes de données**
(**redondance** et **variation**)

Dimension temporelle

Évaluation

Liens avec d'autres acteurs

Travaux sur les événements

<p>1. </p>	<p>Relations temporelles dans les textes</p>	<p><i>avec P. Muller, C. Hagège, puis P. Zweigenbaum, C. Grouin, S. Rosset</i></p>
<p>2. </p>	<p>Événements nominaux</p>	<p>Métriques d'évaluation</p>
<p>3. </p>	<p>Relations multi-documents</p>	<p><i>avec V. Moriceau</i></p>
<p>4. </p>	<p>Génération de chronologies</p>	<p><i>avec R. Kessler, K.H. Nguyen, V. Moriceau</i> <i>Xerox, MoDyCo, Exalead, AFP</i></p>

Au menu

- Définition et Représentations
- Enjeux
- **L'événement dans le texte et dans le monde**
 1. Événements nominaux
 2. Relations temporelles dans les textes
 3. Relations multi-documents
 4. Génération de chronologies
 5. Journalisme de données
- **Bilan & Perspectives**

L'événement dans le texte

1. Événements nominaux

avec **B. Arnulphy**, A. Vilnat

(CicLing 2012, CORIA 2014,
LREC 2012, TALN 2011, TALN 2010)



Cibler

Événements nominaux

Le général de Gaulle a **appelé** à la résistance

→ L'**appel** du 18 juin

Les alliés **débarquent** en Normandie

→ Le **débarquement** du 6 juin 1944

- Noms d'événement et ambiguïté :

- Polysémie « classique »

salon triathlon

- Événement vs attribut de l'événement

construction (⇒ résultat)

shampooing (⇒ instrument)

exil (⇒ résultat)

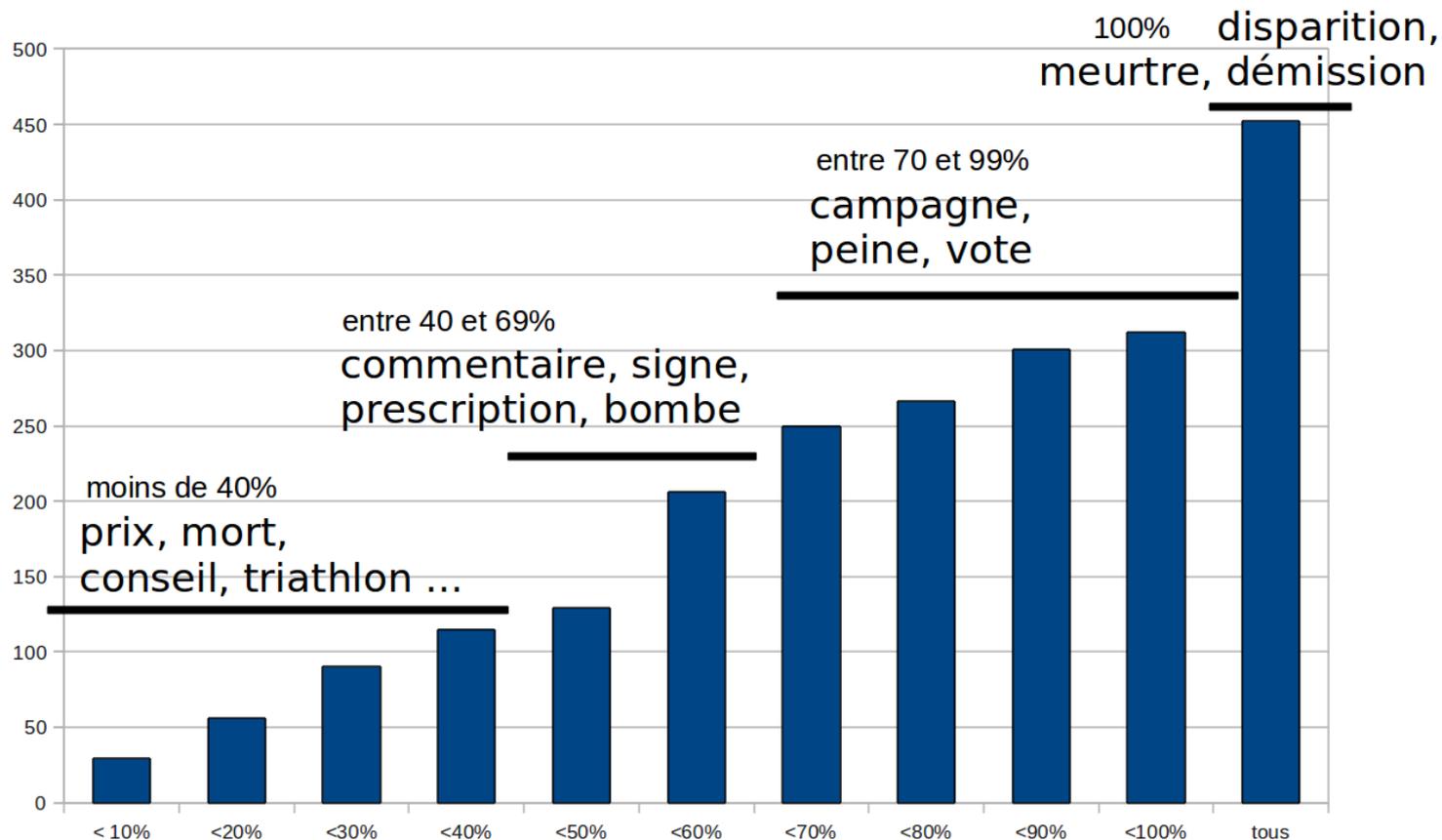
étalage (⇒ localisation)

- Métonymie

Tchernobyl **11 septembre**

21 avril « L'UMP se remettra-t-elle de **Bygmalion** ? »

Étude en corpus (ambiguïté)



Un lexique pondéré des noms d'événement

Mise en place de règles d'extraction **précises** et sur un **large corpus** (AFP)

Ils ont libéré des activistes { pendant
lors de
au début de } **l'Intifada**
à la suite de
à l'issue de

Le **sommet du G8** { se produit
a lieu
est organisé
provoque
entraîne

Un lexique pondéré des noms d'événement

« Poids relatif d'événementialité »

$$\text{ERW}(w) = \frac{\text{Nombre d'occurrences du mot extraites par les règles}}{\text{Nombre total d'occurrences du mot dans le corpus}}$$

Un lexique pondéré des noms d'événement

(CicLing 2012)

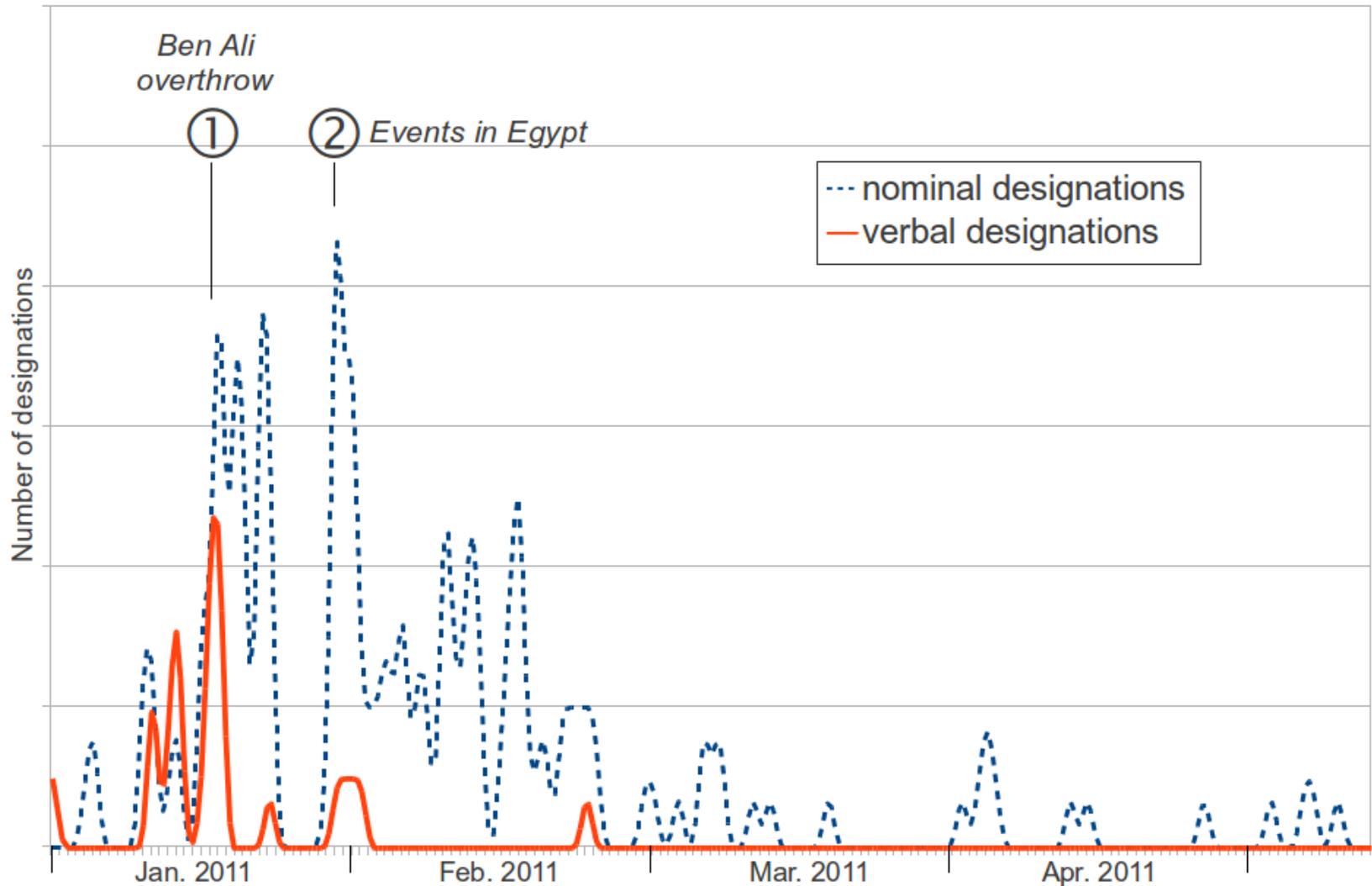
Mots communs avec le lexique VerbAction	ERW
chute	0,166
clôture	0,134
élection	0,128
guerre	0,098
coopération	0,003
subvention	0,002

Mots n'apparaissant pas dans le lexique VerbAction	ERW
Anschluss	0,750
méchoui	0,600
krach	0,118
RTT	0,084
demi-finale	0,063
11-Septembre	0,003

Étude en corpus (évolution)

(LREC 2012)

Révolution tunisienne



L'événement dans le texte

2. Relations temporelles dans les textes

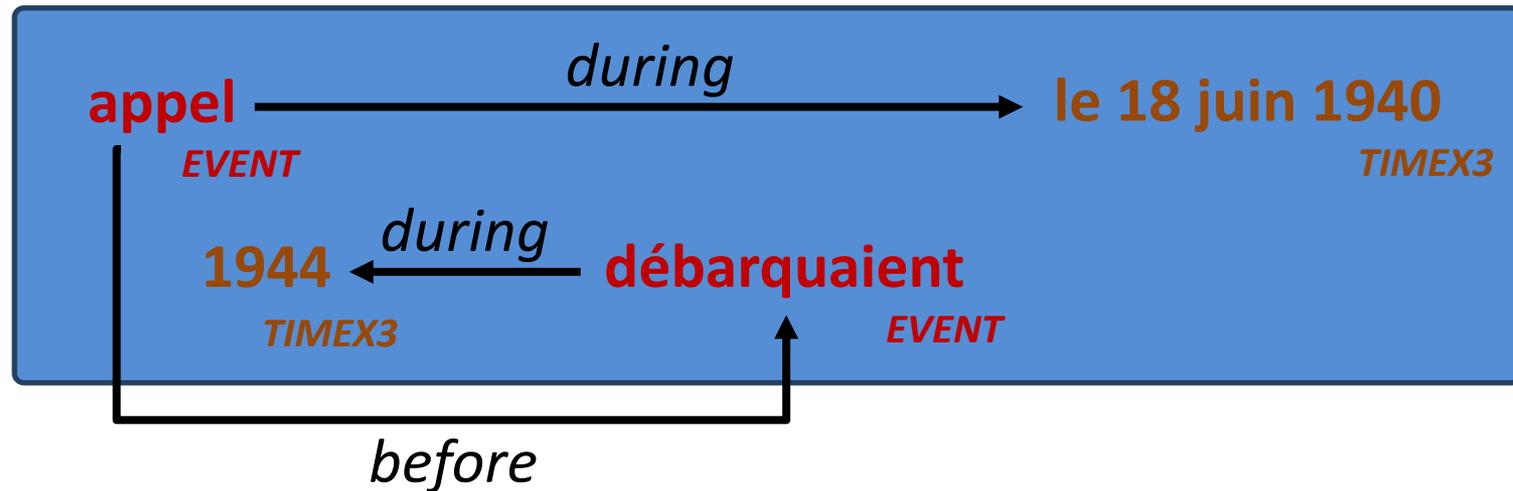
*avec P. Muller, C. Hagège, P. Zweigenbaum,
C. Grouin, S. Rosset*

*(JAMIA 2013, JAIR 2010, Coling 2004,
CicLing 2008, SemEval 2007, LREC 2008,
TALN 2013, TALN 2004)*



Cibler

Relations temporelles dans les textes



↪ Système à base de règles (avec P. Muller, C. Hagège)
(*Coling 2004, CicLing 2007*)

↪ Par apprentissage et combinaison de classifieurs (domaine biomédical,
avec P. Zweigenbaum, C. Grouin, S. Rosset) (*JAMIA 2013, TALN 2013*)

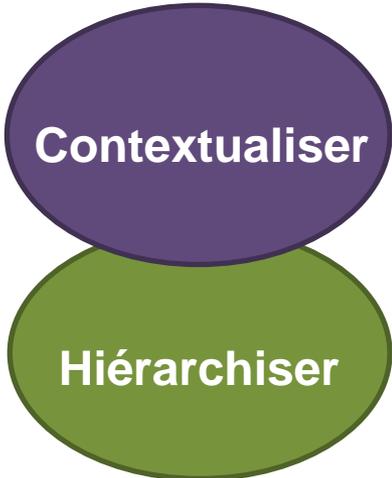
↪ Évaluation des graphes temporels (P. Muller) (*JAIR 2010*)

L'événement dans le monde

3. Génération de chronologies

avec **R. Kessler**, **K.H. Nguyen**, V. Moriceau,
Xerox, MoDyCo, Exalead, AFP

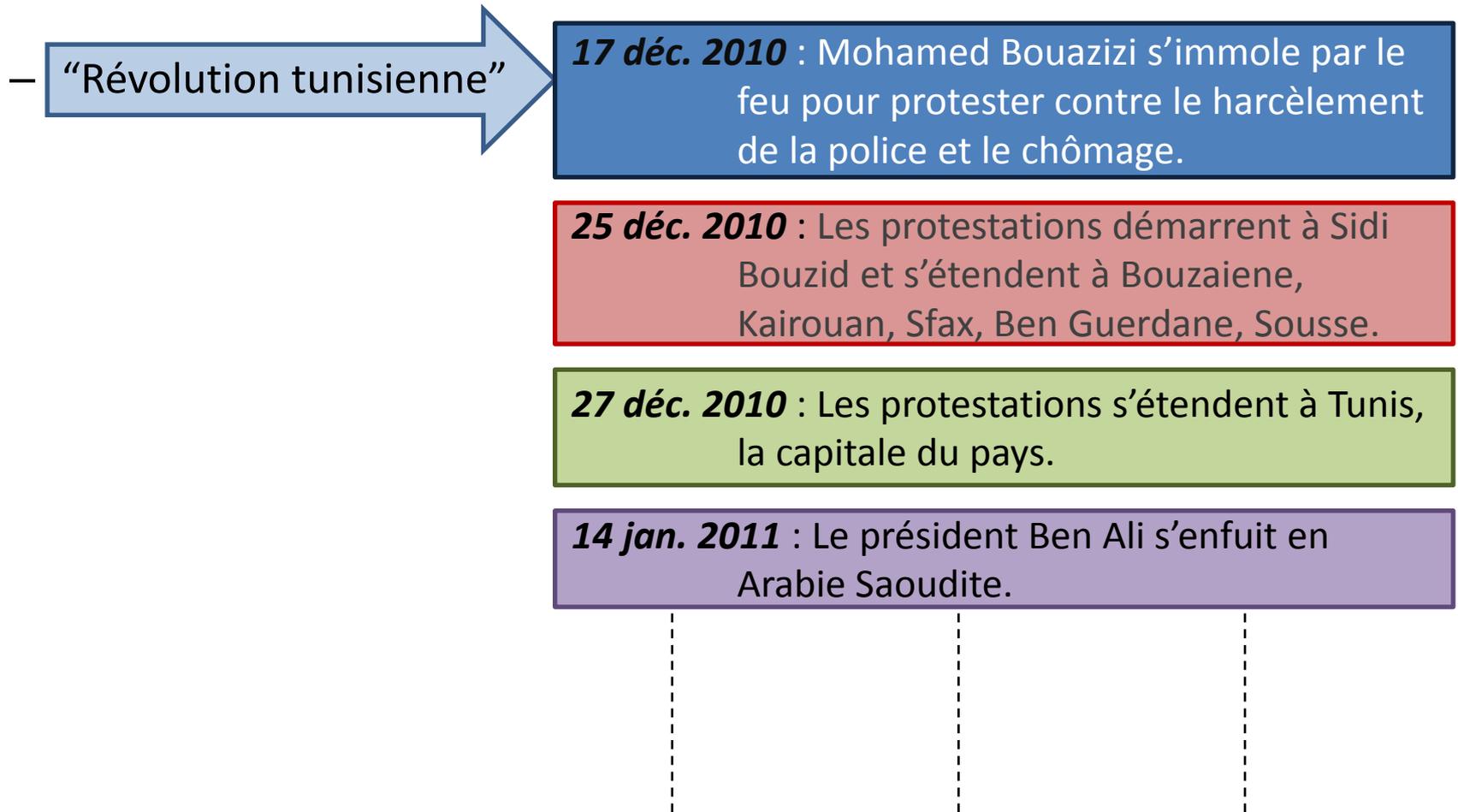
(ACL 2012,
COLING 2014, TAL 2013)



Contextualiser

Hiérarchiser

Objectif



Approche

INDEX

Requête

Plusieurs milliers de documents

Le président égyptien Hosni Mubarak, qui a démissionné vendredi, et le président Zine El Abidine Ben Ali, qui a quitté le pouvoir le **14 janvier**, ont fait face à des protestations populaires sans précédent.

Zine el-Abidine Ben Ali a annoncé **jeudi soir** qu'il ne se représentera pas en 2014 au poste qu'il occupe

Ben Ali a signé sa démission **vendredi** après une vague de protestations déclenchée par le suicide d'un étudiant de 26 ans que la police avait empêché de vendre des fruits et des légumes pour vivre.

Des manifestants ont également blessés **vendredi**, y compris Chawki Belhoussine El Hadri

Ben Ali a signé sa démission **vendredi** après une vague de protestations déclenchée par le suicide d'un étudiant de 26 ans que la police avait empêché de vendre des fruits et des légumes pour vivre.

La révolte s'est propagée **le 24 décembre** dans le centre-du pays, notamment à Menzel Bouzaiane, où Mohamed Ammari est tué par balle dans la poitrine par la police.

Approche

INDEX

Requête

Plusieurs milliers de documents

Clusters temporels

Regroupés par dates normalisées

Ordonnés par l'importance de la date

- Redondance
- Autres critères (apprentissage)

plus important

14 jan. 2011

Ben Ali a signé sa démission **vendredi** après une vague de protestations déclenchée par le suicide d'un étudiant de 26 ans que la police avait empêché de vendre des fruits et des légumes pour vivre.

Ces révélations ont lieu après la révolte tunisienne qui a mit fin à 23 ans de règne de Ben Ali, qui s'est enfui de Tunisie pour l'Arabie Saoudite **vendredi**.

24 déc. 2010

La révolte s'est propagée **le 24 décembre** dans le centre-du pays, notamment à Menzel Bouzaiane, où Mohamed Ammari est tué par balle dans la poitrine par la police.

Des manifestants ont également blessés **vendredi**, y compris Chawki Belhoussine El Hadri

13 jan. 2011

Zine el-Abidine Ben Ali a annoncé **jeudi soir** qu'il ne se représentera pas en 2014 au poste qu'il occupe

Ben Ali a donné **jeudi** l'ordre à la police de ne plus tirer sur les manifestants

INDEX

Approche

Requête

Plusieurs milliers de documents

Clusters temporels

- Regroupés** par dates normalisées
- Ordonnés** par l'importance de la date
- Redondance
 - Autres critères (apprentissage)

Choix des descriptions d'événements

- Maximiser la **pertinence**
- Minimiser la **redondance**

plus important

14 jan. 2011

Ben Ali a signé sa démission **vendredi** après une vague de protestations déclenchée par le suicide d'un étudiant de 26 ans que la police avait empêché de vendre des fruits et des légumes pour vivre.

Ces révélations ont lieu après la révolte tunisienne qui a mit fin à 23 ans de règne de Ben Ali, qui s'est enfui de Tunisie pour l'Arabie Saoudite **vendredi**.

24 déc. 2010

La révolte s'est propagée **le 24 décembre** dans le centre-du pays, notamment à Menzel Bouzaiane, où Mohamed Ammari est tué par balle dans la poitrine par la police.

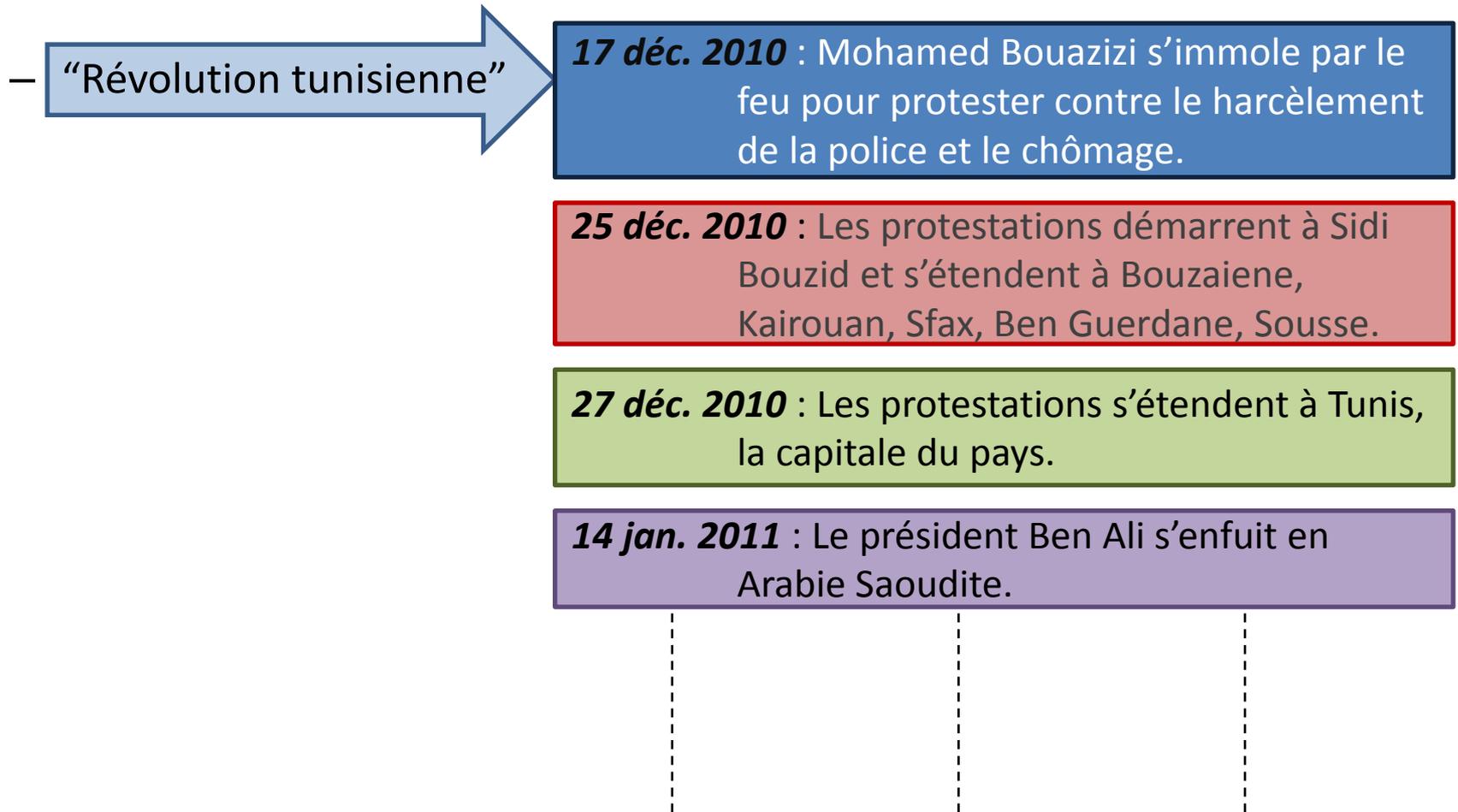
Des manifestants ont également blessés **vendredi**, y compris Chawki Belhoussine El Hadri

13 jan. 2011

Zine el-Abidine Ben Ali a annoncé **jeudi soir** qu'il ne se représentera pas en 2014 au poste qu'il occupe

Ben Ali a donné **jeudi** l'ordre à la police de ne plus tirer sur les manifestants

Chronologie

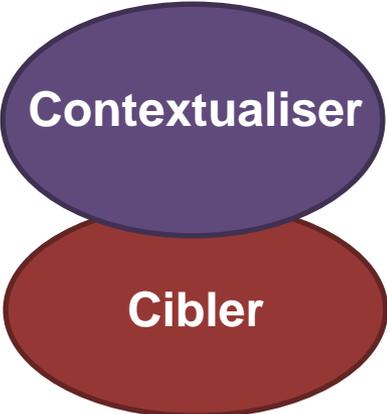


L'événement dans le monde

4. Relations multi-documents

avec V. Moriceau, AFP

(EMNLP 2013)

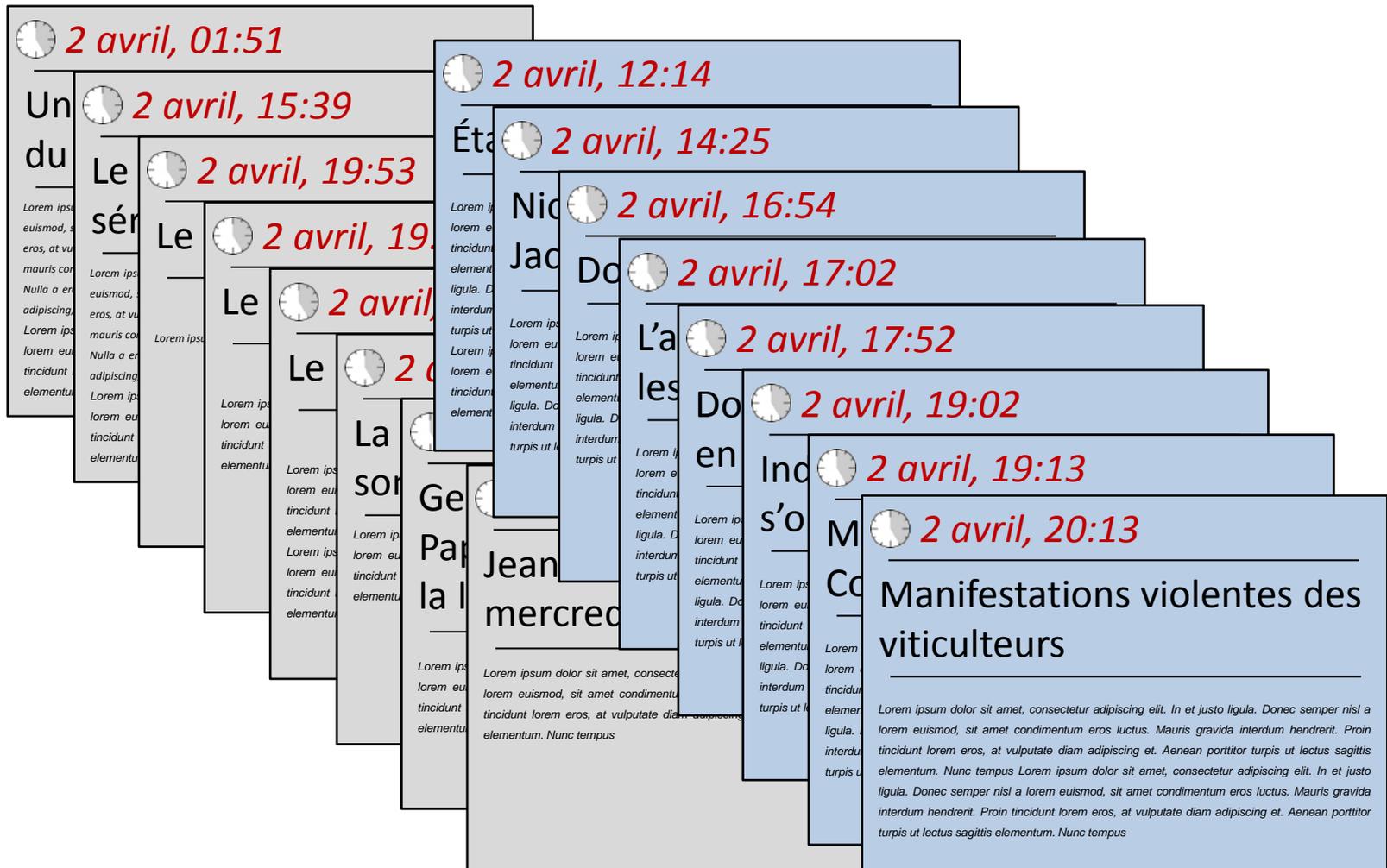


Contextualiser

Cibler

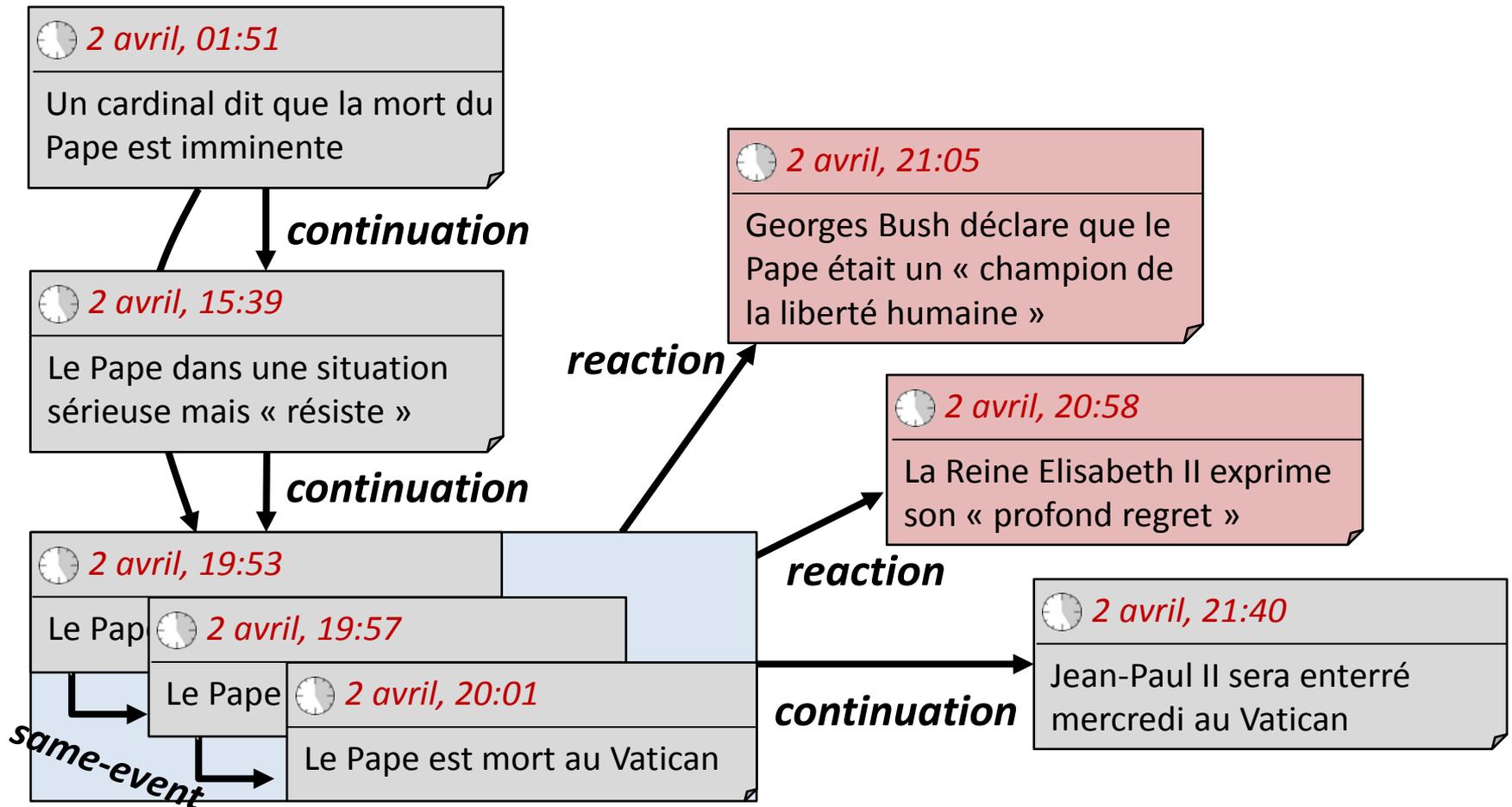
Objectif

- Une approche multi-document des relations temporelles ou causales entre événements

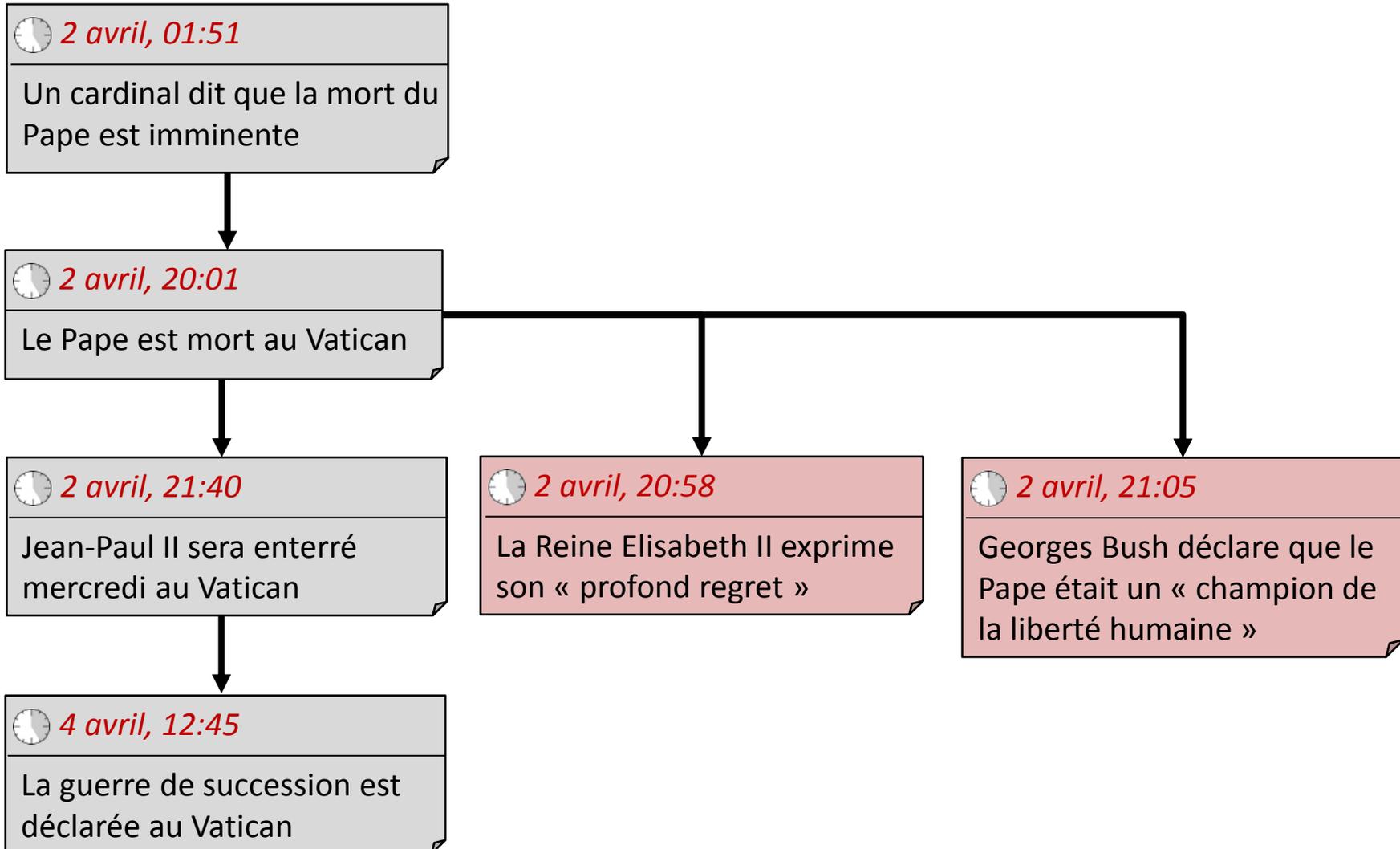


Objectif

- Une approche multi-document des relations temporelles ou causales entre événements

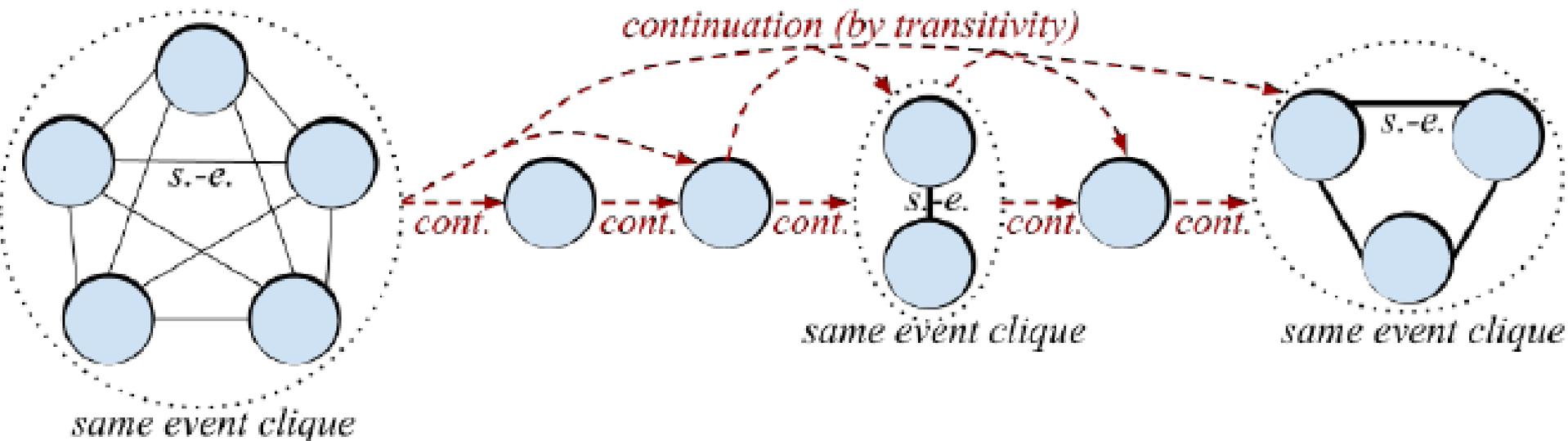


Fil temporel



Approche (à peaufiner)

- Allier traitements linguistiques fins, manipulation de sacs de mots et utilisation de la redondance des informations
 - Extraction de relations par des classifieurs SVM
 - Similarité lexicale
 - Indices temporels
 - Indices lexicaux
 - Utilisation de la redondance et vote



L'événement dans le monde

5. Extraction d'événement et journalisme de données

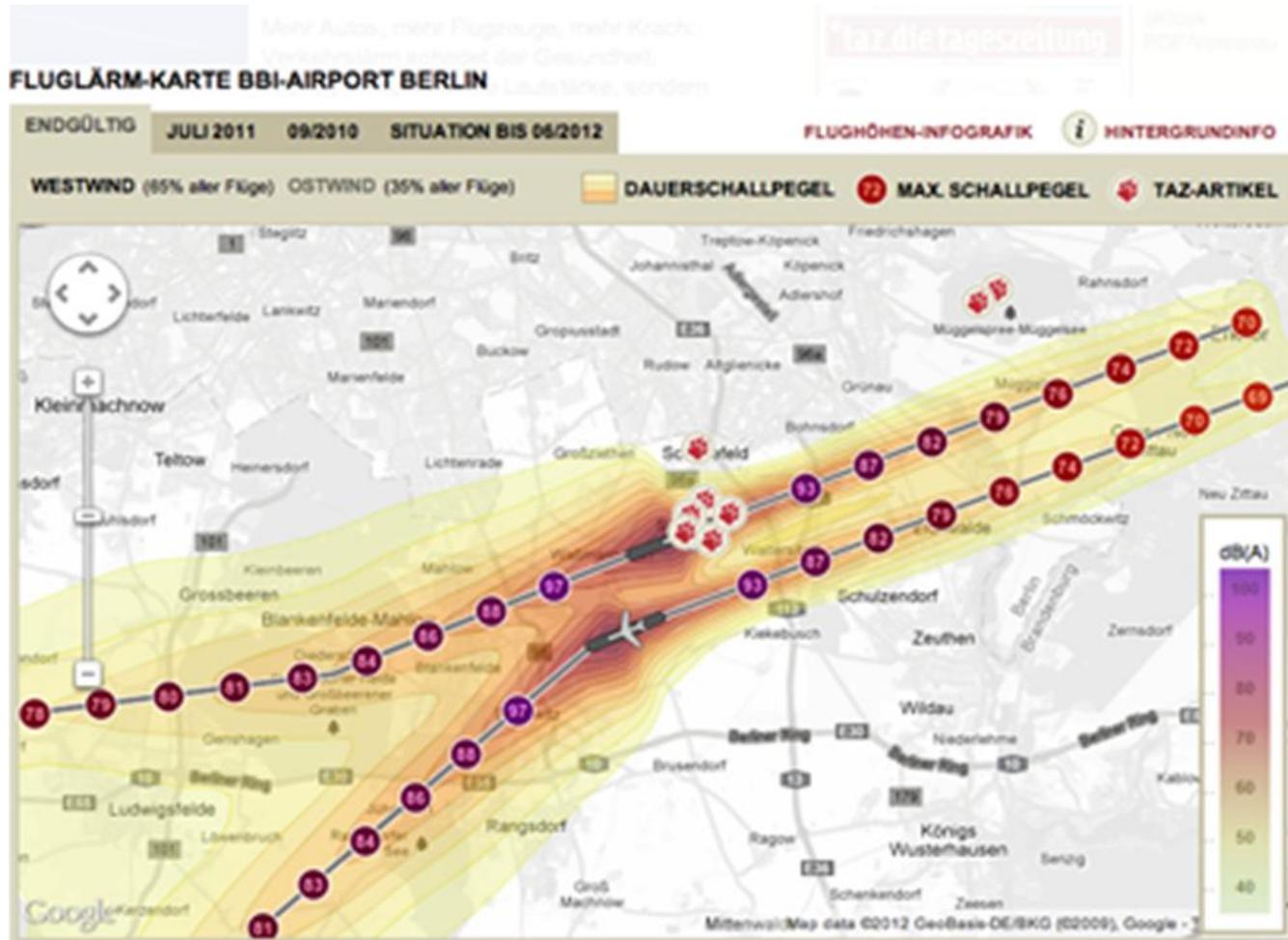
avec AFP

(en cours de soumission)

Contextualiser

Agréger

Journalisme de données : La carte des nuisances sonores autour de l'aéroport de Berlin



Journalisme de données :

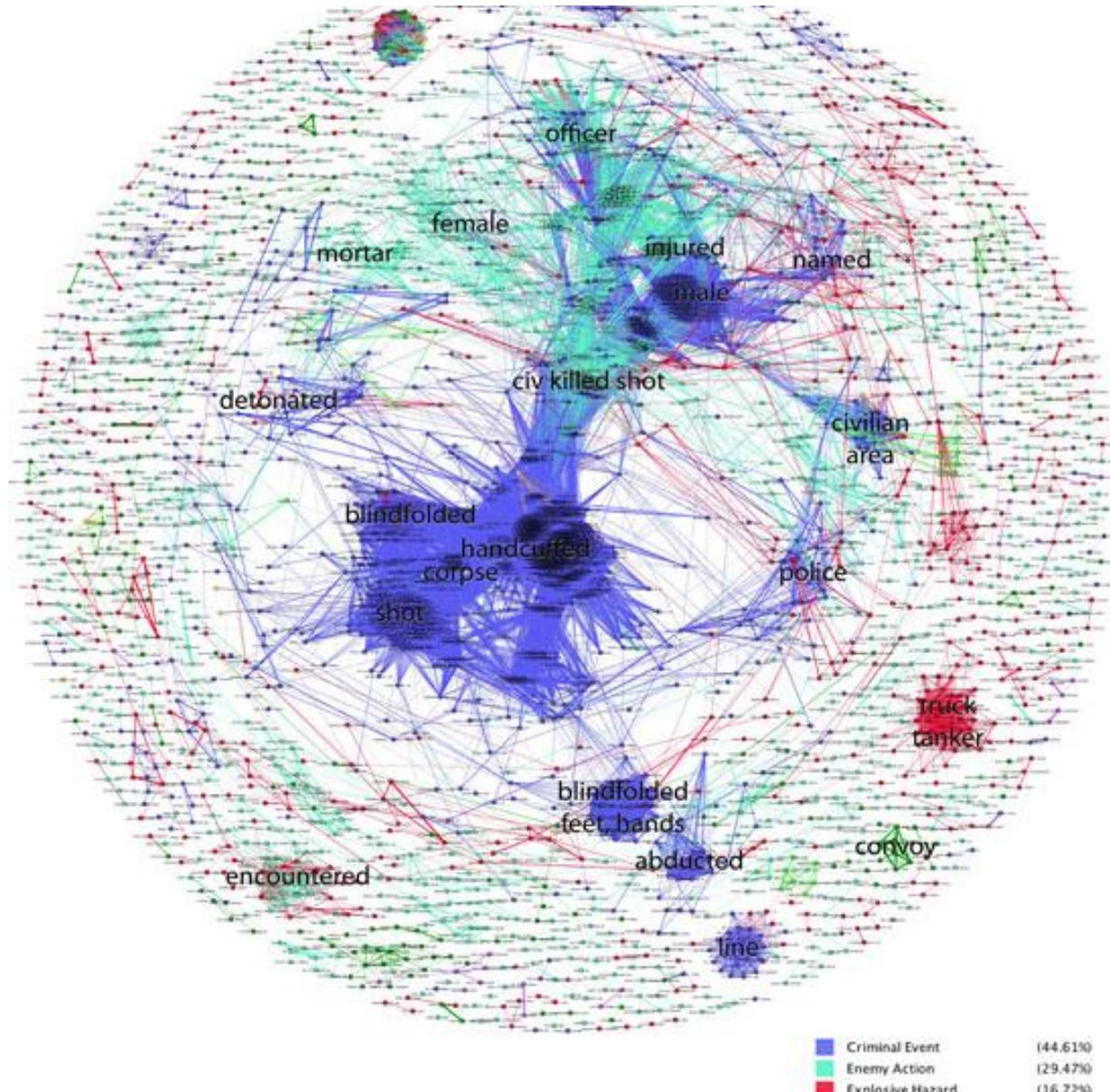
Le test « gagnez-vous autant qu'un homme ? »

LE PARITEUR
– Changez le sexe de votre fiche de paie –

Vous êtes une FEMME
Gagnez-vous autant
qu'un HOMME ?

FAITES
LE TEST
∨

Journalisme de données : Le graphe thématique des câbles Wikileaks sur l'Irak



Journalisme de données

- Utilisation de données numériques
- Visualisation
- Interactivité
- Dimension temporelle
- Assez peu de TAL

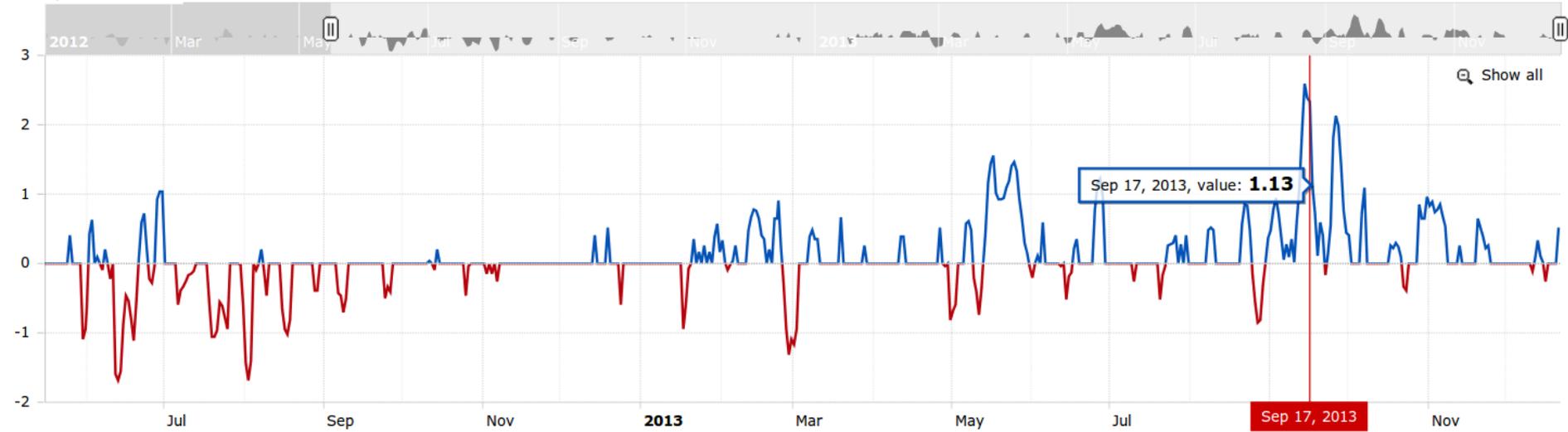
Proposition

- Extraire des données numériques fiables à partir d'un corpus
 1. **Extraire** des informations dans les documents
 2. **Agréger** ces informations
 3. **Atténuer le bruit**
 4. Conserver la **dimension temporelle**
 5. Proposer une **visualisation interactive** adéquate

Proposition

Exemple : relations entre les États-Unis et la Russie sur le dossier syrien

chart by amcharts.com



Tuesday, September 17th, 2013

- UN chief Ban Ki-moon said Tuesday he would meet the foreign ministers of the five key nations in New York on the sidelines of the UN General Assembly, and three days later, with the top diplomats of the US and Russia, to help resolve the Syria crisis.
- What matters now is that this rapprochement between Russia and the United States is confirmed, lasts, and that together these two leading countries mobilise other leading countries so that we finally move towards a political solution, he said in an interview with Swiss public broadcaster RTS.
- WASHINGTON: President Barack Obama says that a US deal reached with Russia at the weekend has the potential to end the threat of Syria's chemical arsenal.
- Oil markets continue to retreat in the wake of the agreement between the US and Russia to work to eliminate Syria's stockpile of chemical weapons, said Addison Armstrong of Tradition Energy.



Proposition

1. *Alliances et oppositions internationales*

Indonesia **voiced support** for **East Timor's** bid to join the ASEAN.

⇒ POS(Indonesia, East Timor)

China earlier **protested** **Obama's** meeting with the Dalai Lama, [...].

⇒ NEG(China, U.S.A.)

↳ Extraction par apprentissage

- Priorité à la précision (déclencheurs explicites)
- Classifieurs SVM (SMO)
- Traits lexicaux et syntaxiques

⇒ Précision \approx 80 %

Proposition

2,3,4. **Agrégation et lissage pour chaque unité de temps t**

Étant donné un thème (requête)

Pour chaque paire de pays dans chaque unité de temps :

$$poids(e_1, e_2, t) = \log \frac{1 + POS(e_1, e_2, t)}{1 + NEG(e_1, e_2, t)}$$

Lissage dans une fenêtre temporelle :

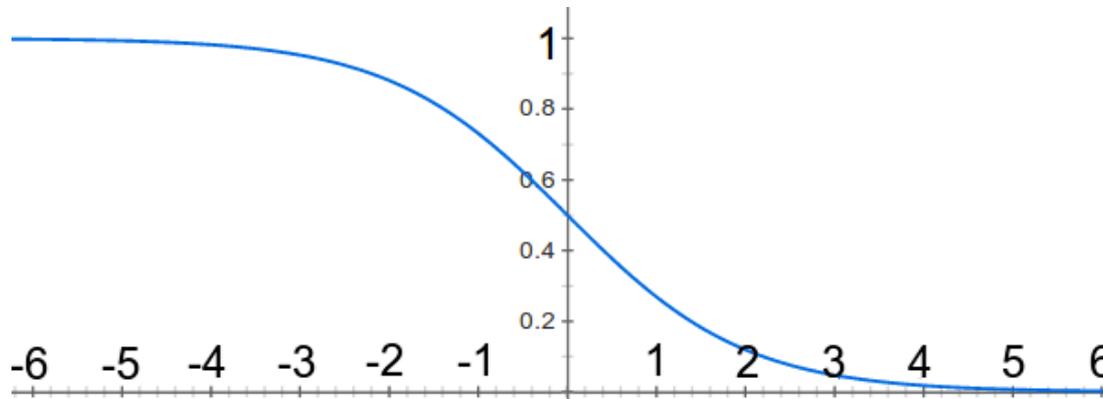
$$poids_lissé(e_1, e_2, t) = \text{moyenne pondérée de } poids(t)$$

Proposition

2,3,4. **Agrégation et lissage pour chaque unité de temps t**

Transformation des poids en distance

$$distance(e_1, e_2, t) = 1 - \frac{1}{1 + e^{-poids(e_1, e_2, t)}}$$

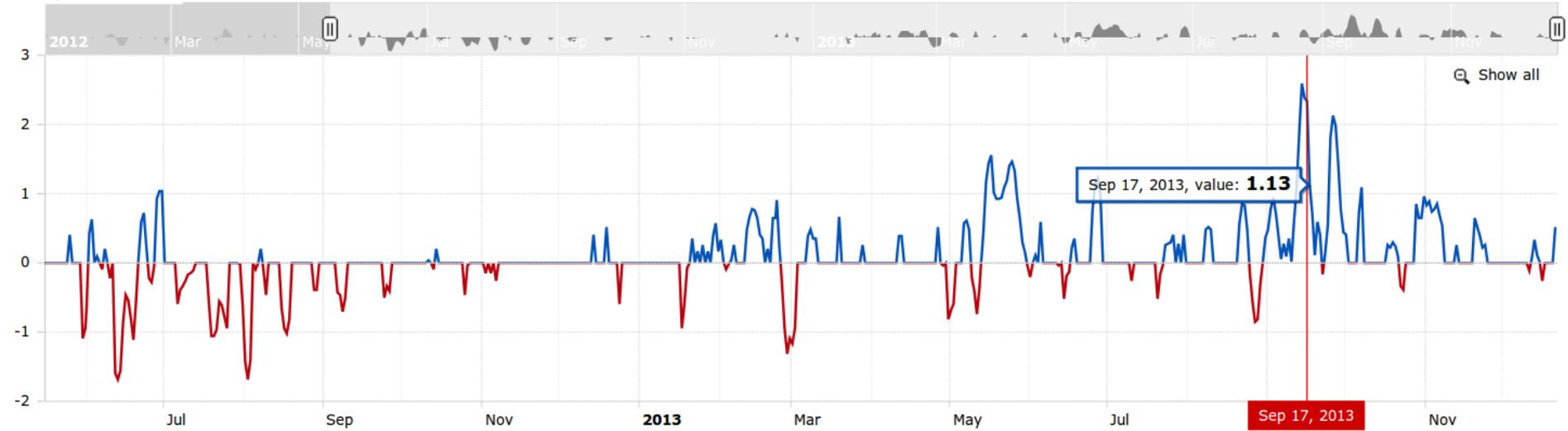


Proposition

5. *Visualisation*

Requête : « syria » entre USA et Russia

chart by amcharts.com



Tuesday, September 17th, 2013

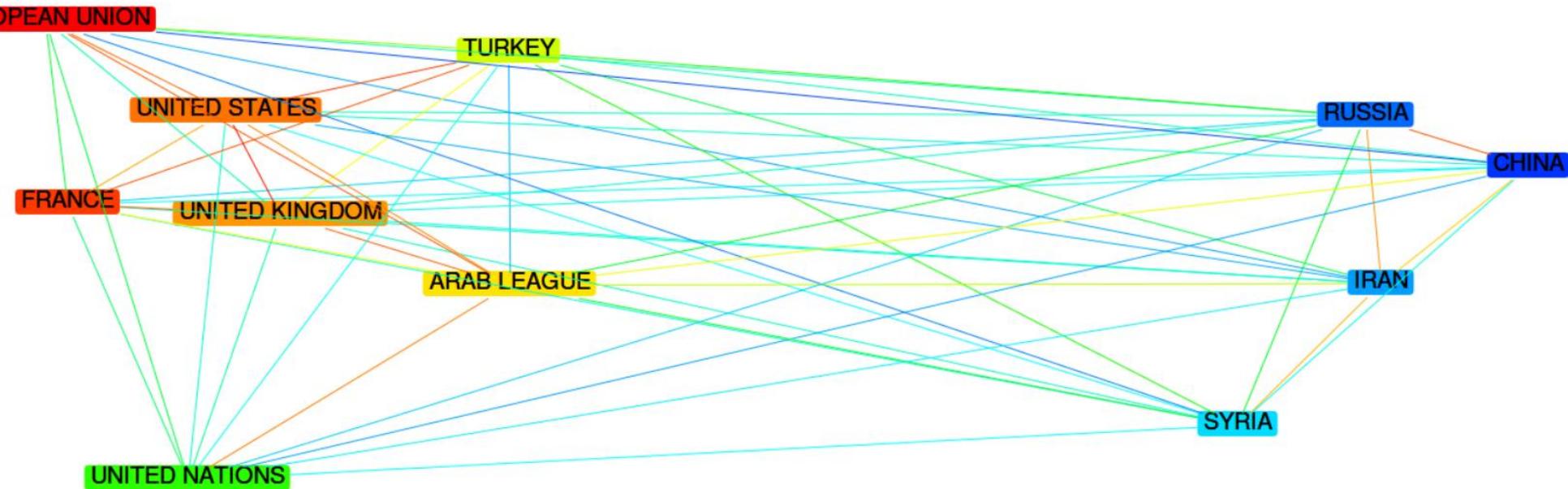
- UN chief Ban Ki-moon said Tuesday he would meet the foreign ministers of the five key nations in New York on the sidelines of the UN General Assembly, and three days later , with the top diplomats of the US and Russia, to help resolve the Syria crisis.
- What matters now is that this rapprochement between Russia and the United States is confirmed, lasts, and that together these two leading countries mobilise other leading countries so that we finally move towards a political solution, he said in an interview with Swiss public broadcaster RTS.
- WASHINGTON: President Barack Obama says that a US deal reached with Russia at the weekend has the potential to end the threat of Syria s chemical arsenal.
- Oil markets continue to retreat in the wake of the agreement between the US and Russia to work to eliminate Syria s stockpile of chemical weapons, said Addison Armstrong of Tradition Energy.



Proposition

5. *Visualisation*

Requête : « syria » en 2012

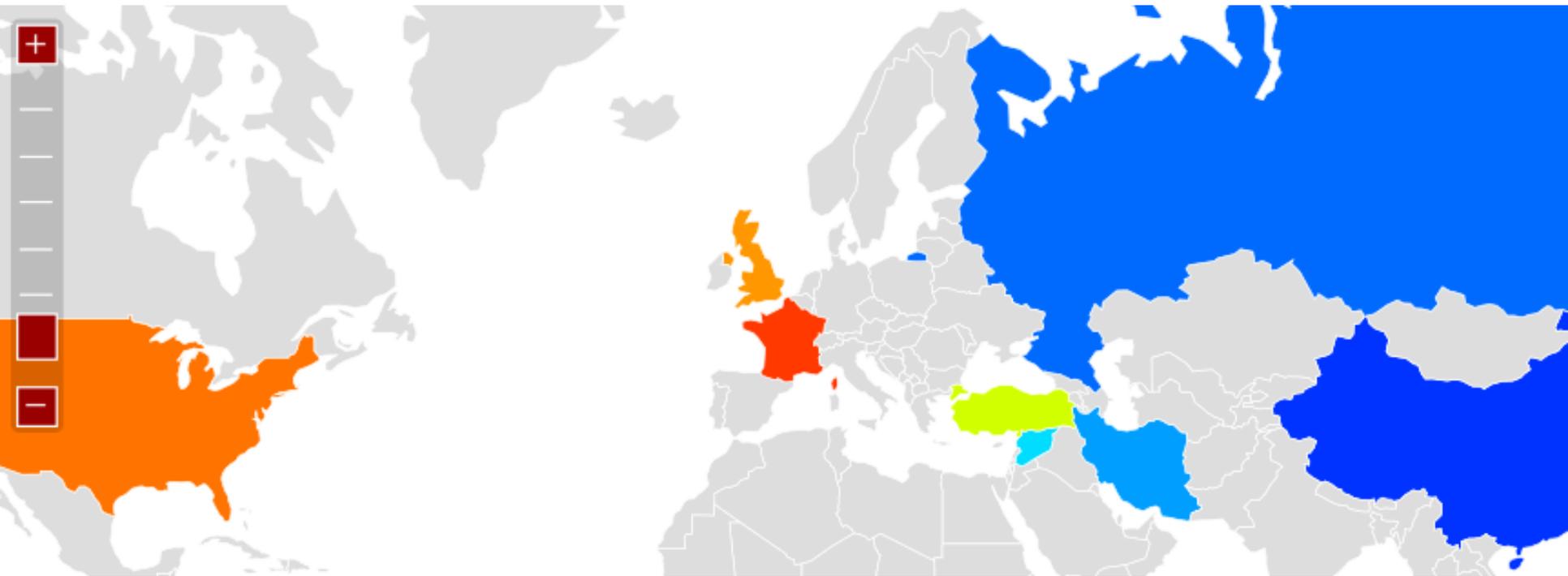


*Nœuds placés par l'algorithme de Barnes-Hut (force-directed layout)
Couleurs choisies en fonction du plus court chemin dans le graphe (plus proches voisins)*

Proposition

5. *Visualisation*

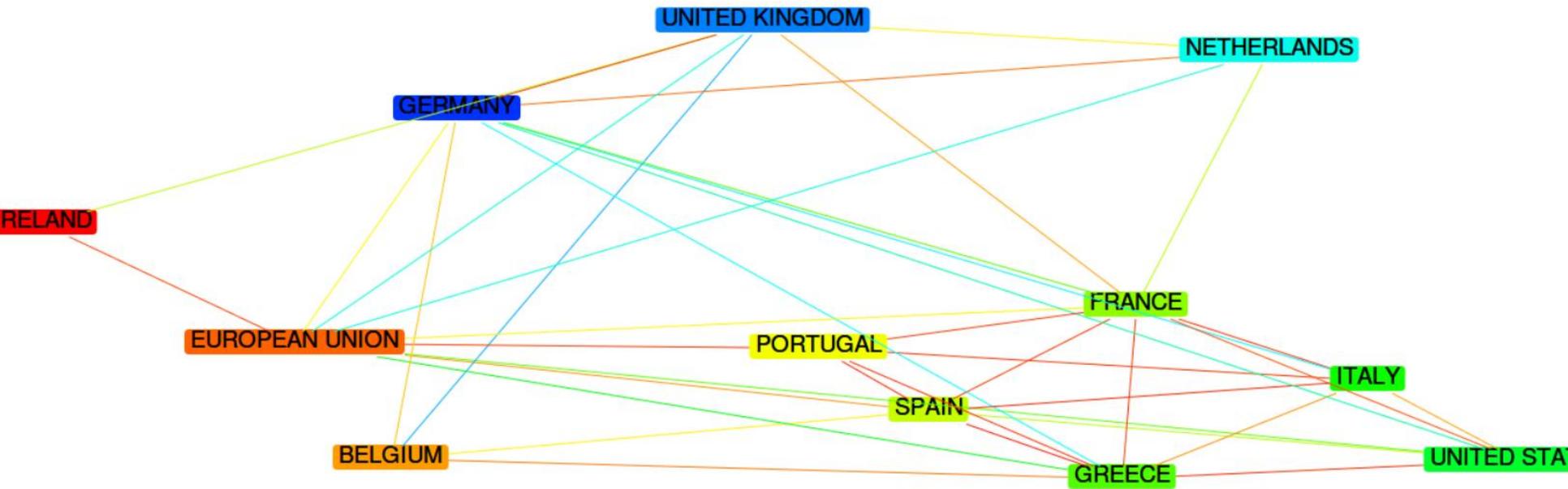
Requête : « syria » en 2012



Proposition

5. *Visualisation*

Requête : « *austerity* » en 2012-2013



Au menu

- Définition et Représentations
- Enjeux
- L'événement dans le texte et dans le monde
 1. Événements nominaux
 2. Relations temporelles dans les textes
 3. Relations multi-documents
 4. Génération de chronologies
 5. Journalisme de données
- **Bilan & Perspectives**

Questions posées

Événement dans le texte et dans le monde

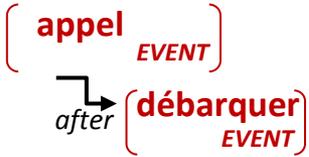
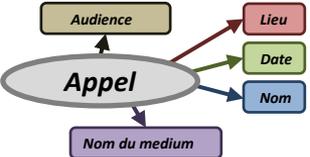
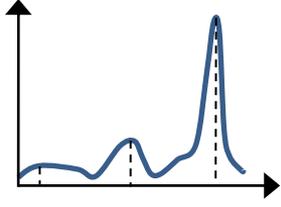
Précision et couverture dans les **gros volumes de données**
(**redondance** et **variation**)

Dimension temporelle

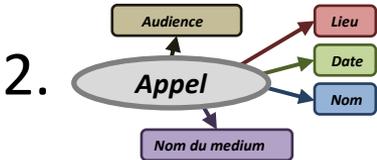
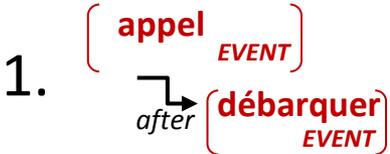
Évaluation

Liens avec d'autres acteurs

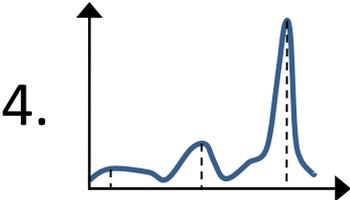
Travaux sur les événements

<p>1. </p>	<p>Relations temporelles dans les textes</p>	<p><i>avec P. Muller, C. Hagège, puis P. Zweigenbaum, C. Grouin, S. Rosset</i></p>
<p>2. </p>	<p>Événements nominaux</p>	<p>Métriques d'évaluation</p>
<p>3. </p>	<p>Relations multi-documents</p>	<p><i>avec V. Moriceau</i></p>
<p>4. </p>	<p>Génération de chronologies</p>	<p><i>avec R. Kessler, K.H. Nguyen, V. Moriceau</i> <i>Xerox, MoDyCo, Exalead, AFP</i></p>

Perspectives à court terme



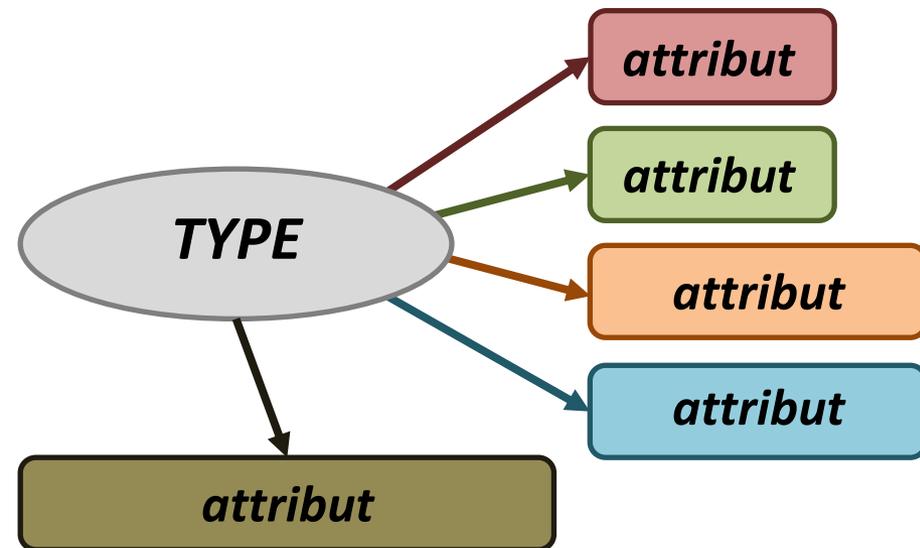
Induction de schémas



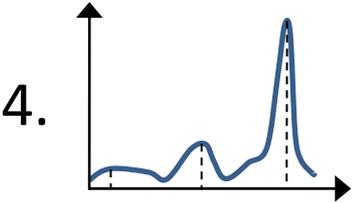
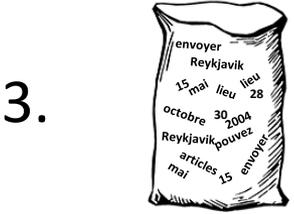
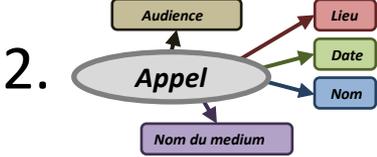
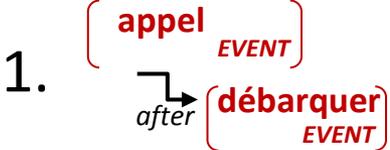
Induction de schémas

- **Découvrir automatiquement** les schémas à partir d'un large corpus
- **Construire, structurer** et **remplir** une base de connaissances
 - Identification
 - Nommage
 - Annotation
 - Recherche

Lorem ipsum dolor sit amet, **consectetur adipiscing** elit. In et justo ligula. Donec semper nisl a lorem euismod, **sit amet** condimentum **eros luctus**. Mauris gravida interdum hendrerit.



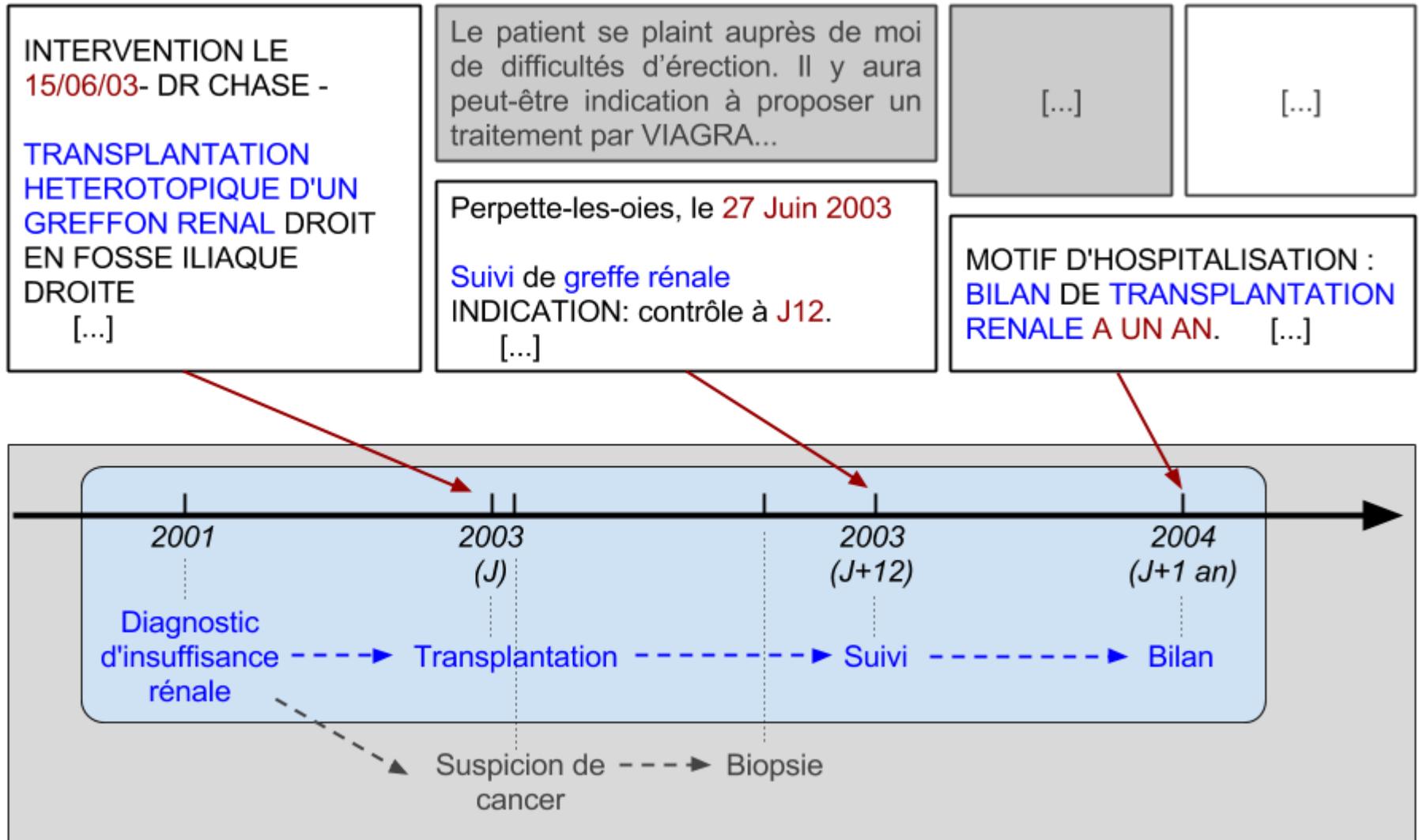
Perspectives à court terme



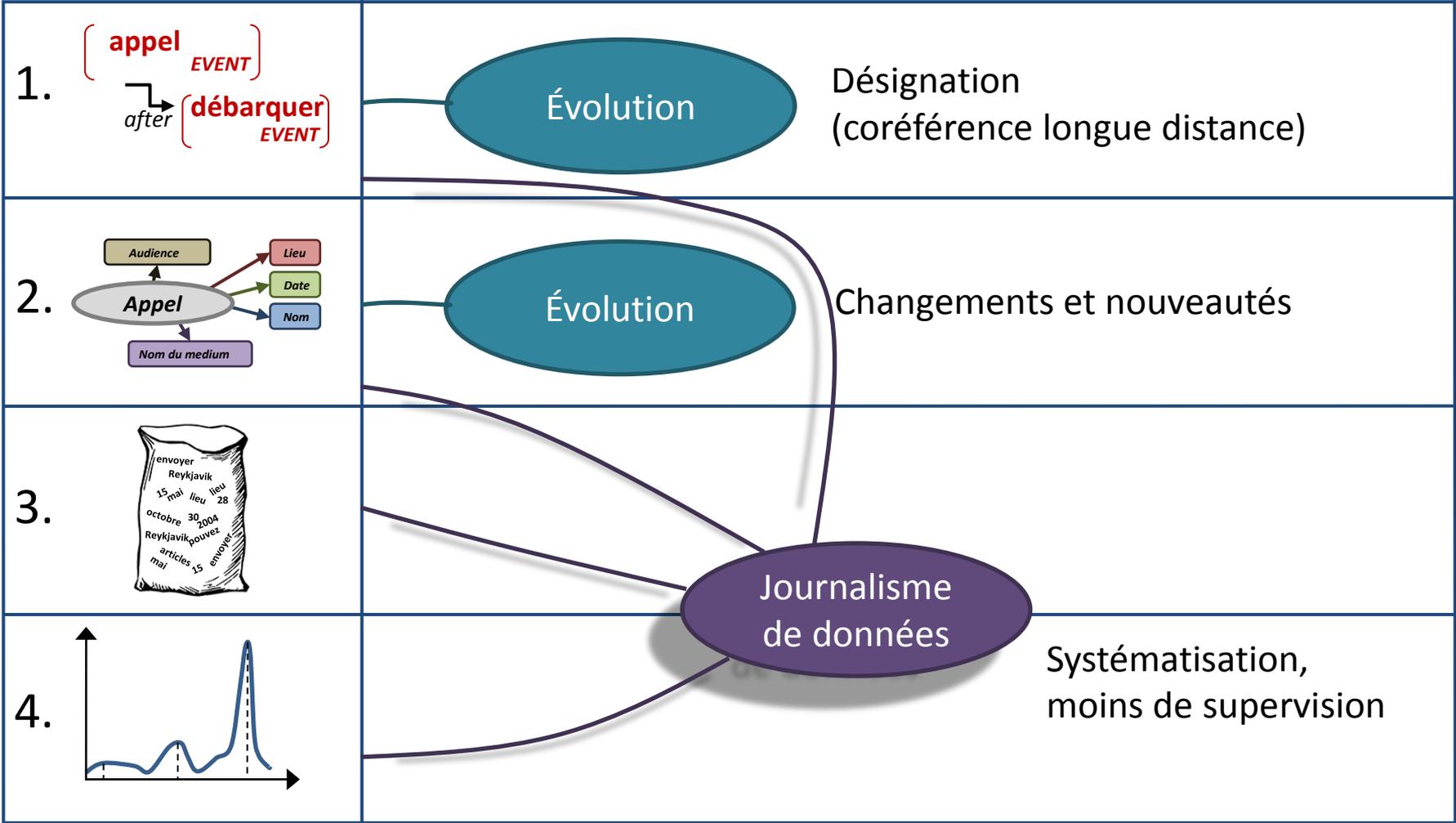
Induction de schémas

Chronologie du parcours de soin

Chronologie du parcours de soin

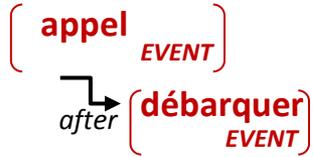


Perspectives à moyen terme

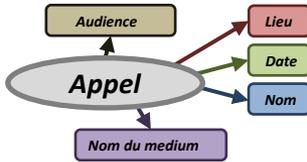


Perspectives à moyen terme

1.



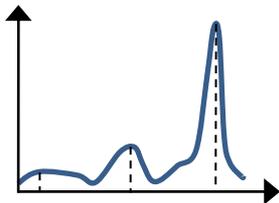
2.



3.



4.



Fact
Checking

Formalisation
des
représentations

Perspectives

- **À plus long terme**

Les robots qui écrivent les articles

La mort du journalisme

Les moteurs de recherche
rendent stupide

Déchéance de l'esprit humain

Libérer du temps de cerveau
disponible

Singularité technologique

La substantifique moelle

L'aiguille dans la botte de foin

Les moteurs de recherche
rendent libre

Révolution de la
plasticité cérébrale

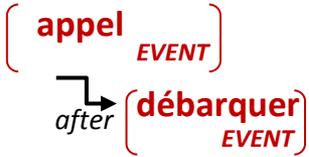
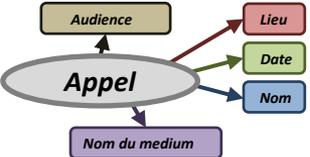
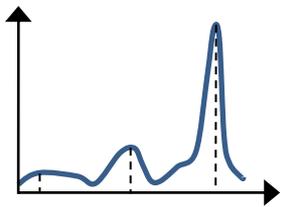
Libérer du temps de cerveau
disponible

Citoyenneté numérique



« Si l'information est importante, elle viendra à moi »

Merci !

	Campagnes	Angle de vue	Difficultés & Limites
<p>1.</p> 	<p>TempEval</p> 	<p><i>Relations</i></p>	<ul style="list-style-type: none"> - Mono-document - Application ?
<p>2.</p> 	<p>MUC & C°</p>	<p><i>Structure</i></p>	<ul style="list-style-type: none"> - Spécificité au domaine
<p>3.</p> 	<p>DUC TAC</p>	<p><i>Importance</i> <i>Redondance</i></p>	<ul style="list-style-type: none"> - Peu exploré - Évaluation complexe
<p>4.</p> 	<p>TDT</p>	<p><i>Importance</i> <i>Nouveauté</i></p>	<ul style="list-style-type: none"> - Peu d'évolution



Le patient a été ~~placé~~ sous Neupogen.

~~EVENT~~

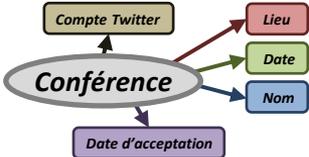
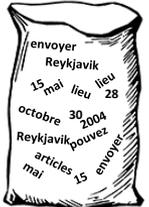
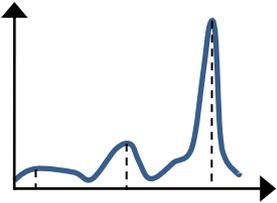
Le patient a été placé sous **Neupogen**.

EVENT

Thesaurus MeSH

Name of Substance	Filgrastim
Entry Term	Neupogen
Heading Mapped to	*Granulocyte Colony-Stimulating Factor
Heading Mapped to	Recombinant Proteins
Note	A nonglycosylated protein with an amino acid sequence that is identical to the human granulocyte colony-stimulating factor except for the addition of an N-terminal methionine.
Unique ID	C554038

Travaux sur les événements

<p>1. </p>	<p><i>avec P. Muller, C. Hagège</i></p> <p>Relations temporelles dans les textes</p> <p>Événements nominaux</p>	<p><i>avec P. Zweigenbaum, C. Grouin, S. Rosset</i></p> <p>Analyse de comptes-rendus d'hospitalisation</p> <p>Métriques d'évaluation</p>
<p>2. </p>	<p><i>avec B. Arnulphy, A. Vilnat</i></p> <p><i>avec P. Muller</i></p>	
<p>3. </p>	<p>Relations multi-documents</p> <p>Génération de chronologies</p>	<p><i>avec V. Moriceau</i></p> <p><i>avec R. Kessler, K.H. Nguyen, V. Moriceau</i></p>
<p>4. </p>	<p>Application au journalisme de données</p> <p><i>Xerox, MoDyCo, Exalead, AFP</i></p>	

- Trois types de construction et d'ambiguïté :

- Apparentée à un verbe d'action

construction	(⇒ résultat)	shampooing	(⇒ instrument)
exil	(⇒ résultat)	étalage	(⇒ localisation)

- Noms décrivant des événements de façon intrinsèque

festival	tsunami	salon	(⇒ polysémie classique)
match	apartheid	triathlon	

- Nominalisation par métonymie

Tchernobyl	11 septembre	« L'UMP se remettra-t-elle de Bygmalion ? »
21 avril		

(⇒ ambiguïté systématique)

L'événement dans le texte

2. Analyse temporelle de comptes-rendus d'hospitalisation

*avec P. Zweigenbaum,
C. Grouin, S. Rosset*

*(JAMIA 2013,
I2B2 2012, TALN 2013)*



Compte-rendu d'hospitalisation

ADMISSION DATE : 10/17/95
date

DISCHARGE DATE : 10/20/95
date

HISTORY OF PRESENT ILLNESS :

This is a 73-year-old man with squamous cell carcinoma of the lung, status
problem
post lobectomy and resection of (...)
treatment *treatment*

HOSPITAL COURSE :

He was started on Neupogen, 400 mcg. subq. q.d. (...)
treatment

He was discharged home
occurrence

(...).

Compte-rendu d'hospitalisation

ADMISSION DATE : 10/17/95

DIS
HIS
This
pos
HO
He

Identifier les **relations temporelles** décrivant la **chronologie** du séjour hospitalier d'un patient

- **Meilleure connaissance** de l'enchaînement des problèmes, des traitements, des rendez-vous, des opérations
- **Aide aux analyses** et aux traitements des médecins ou des systèmes de surveillance automatique

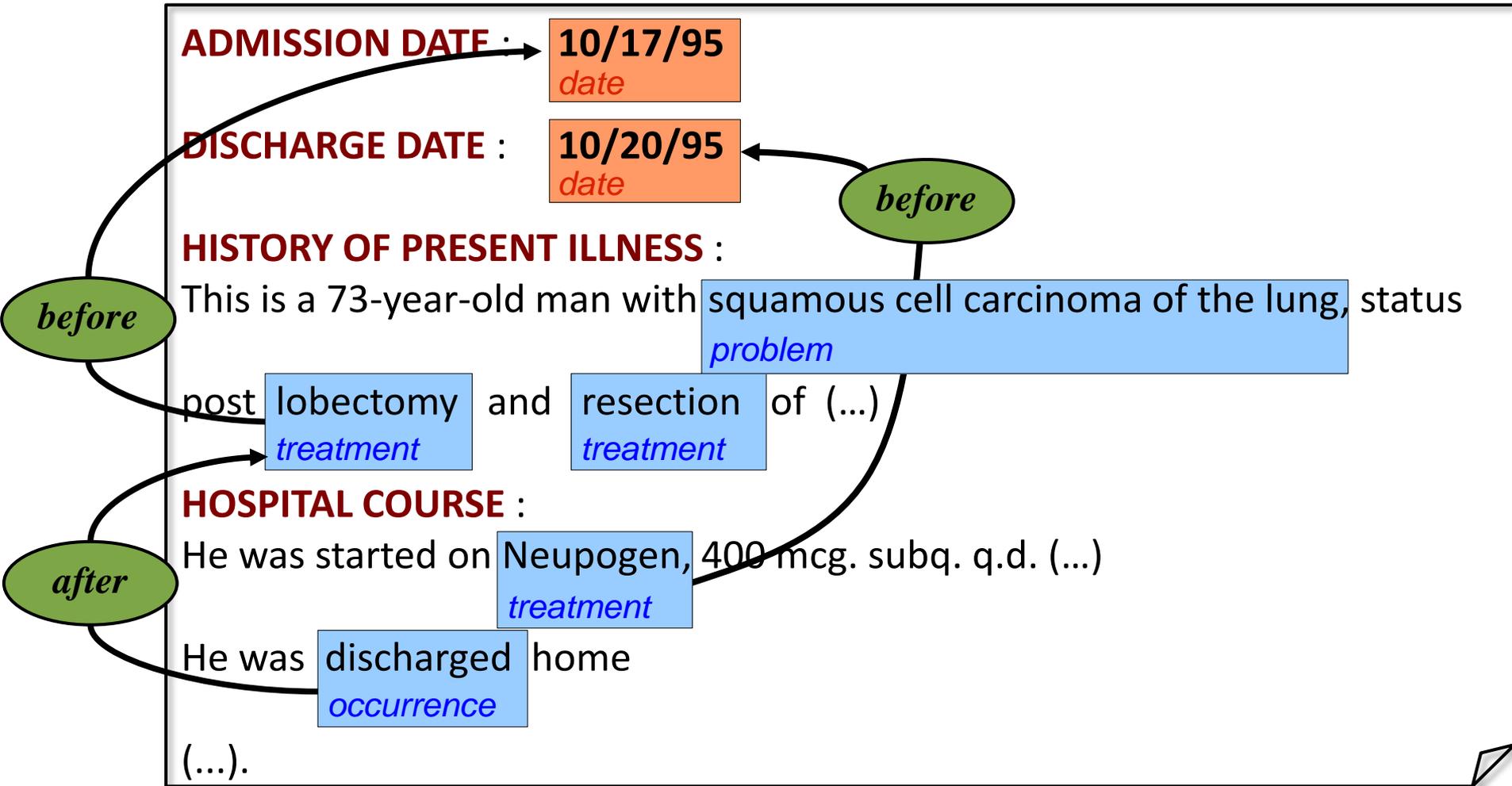
↪ Identifier uniquement les relations « utiles »

He was discharged home

occurrence

(...).

Objectif



Approche

- Pour chaque **paire d'éléments** (« événement » ou date) on cherche à étiqueter leur relation par **BEFORE**, **AFTER**, **OVERLAP** ou **NIL**.



- Différentes **dimensions** :
 - **Section** de la source et de la cible (history of present illness, hospital course...)
 - **Type** de la source et de la cible (événement ou date)
 - **Distance** entre les éléments (même phrase, phrases contiguës, autre)

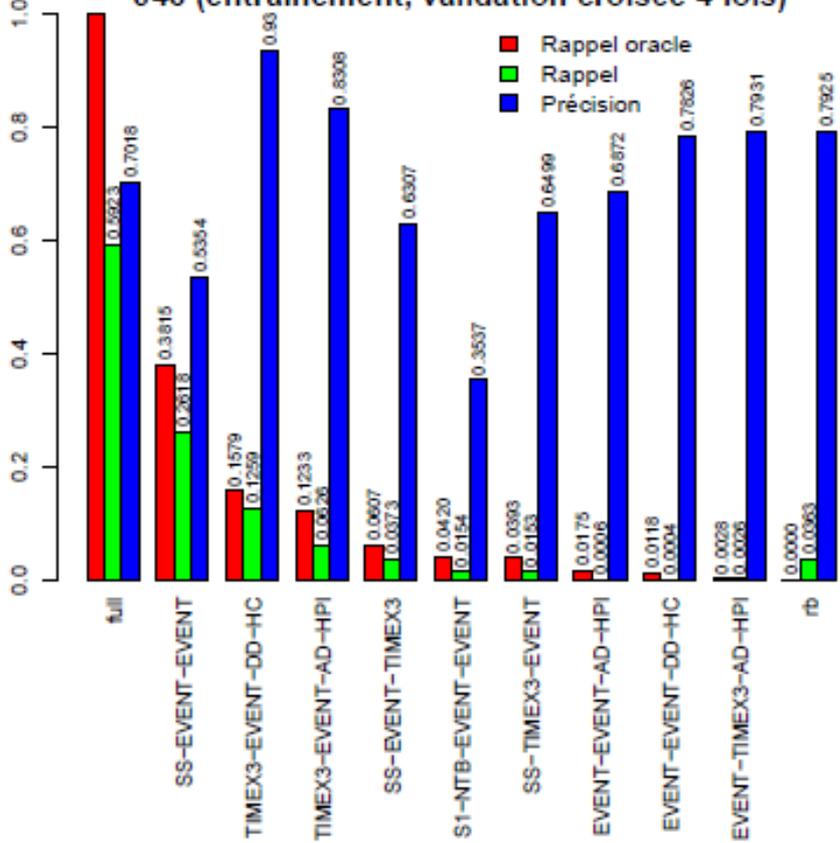
↪ **56 combinaisons**

Approche

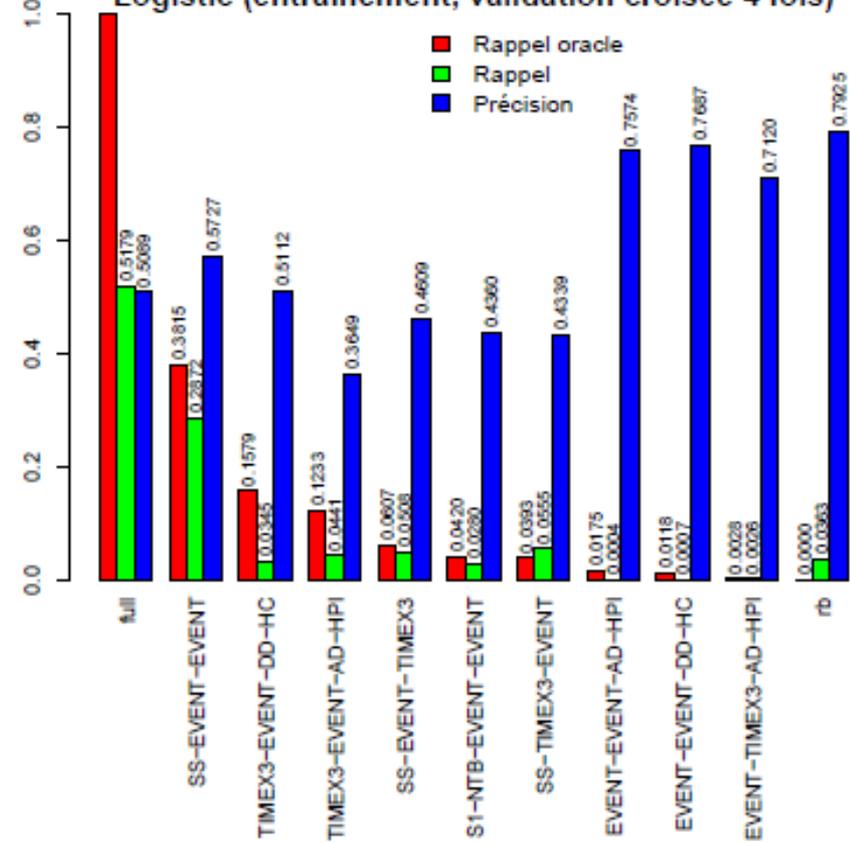
- Quelques **heuristiques préalables** (l'admission est avant la sortie, etc...)
- Étude systématique par validation croisée des **meilleurs classifieurs pour chaque combinaison**
 - Arbres de décision, JRip, K étoile, classifieur bayésien naïf, SVM, k plus proches voisins, MaxEnt
 - (meilleur classifieur en moyenne : J48)

Approche

J48 (entraînement, validation croisée 4 fois)



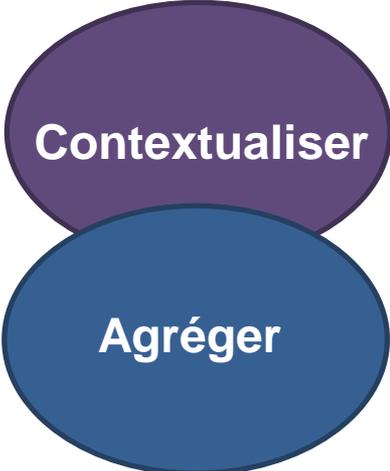
Logistic (entraînement, validation croisée 4 fois)



L'événement dans le monde

4. Acquisition de structures d'événements

*avec V. Moriceau, K.H. Nguyen,
le CEA et l'AFP*



Contextualiser

Agréger

L'événement dans le texte

- Quand et comment le **nom** représente-t-il un événement dans les textes ?
- Comment identifier les **relations temporelles** entre événements dans un document ?
- Comment analyser temporellement le **dossier médical** d'un patient ?
- Comment **évaluer** un « graphe temporel » d'événements ?

L'appel du général de Gaulle fut lancé le 18 juin 1940.

En 1944, les alliés débarquaient en Normandie.



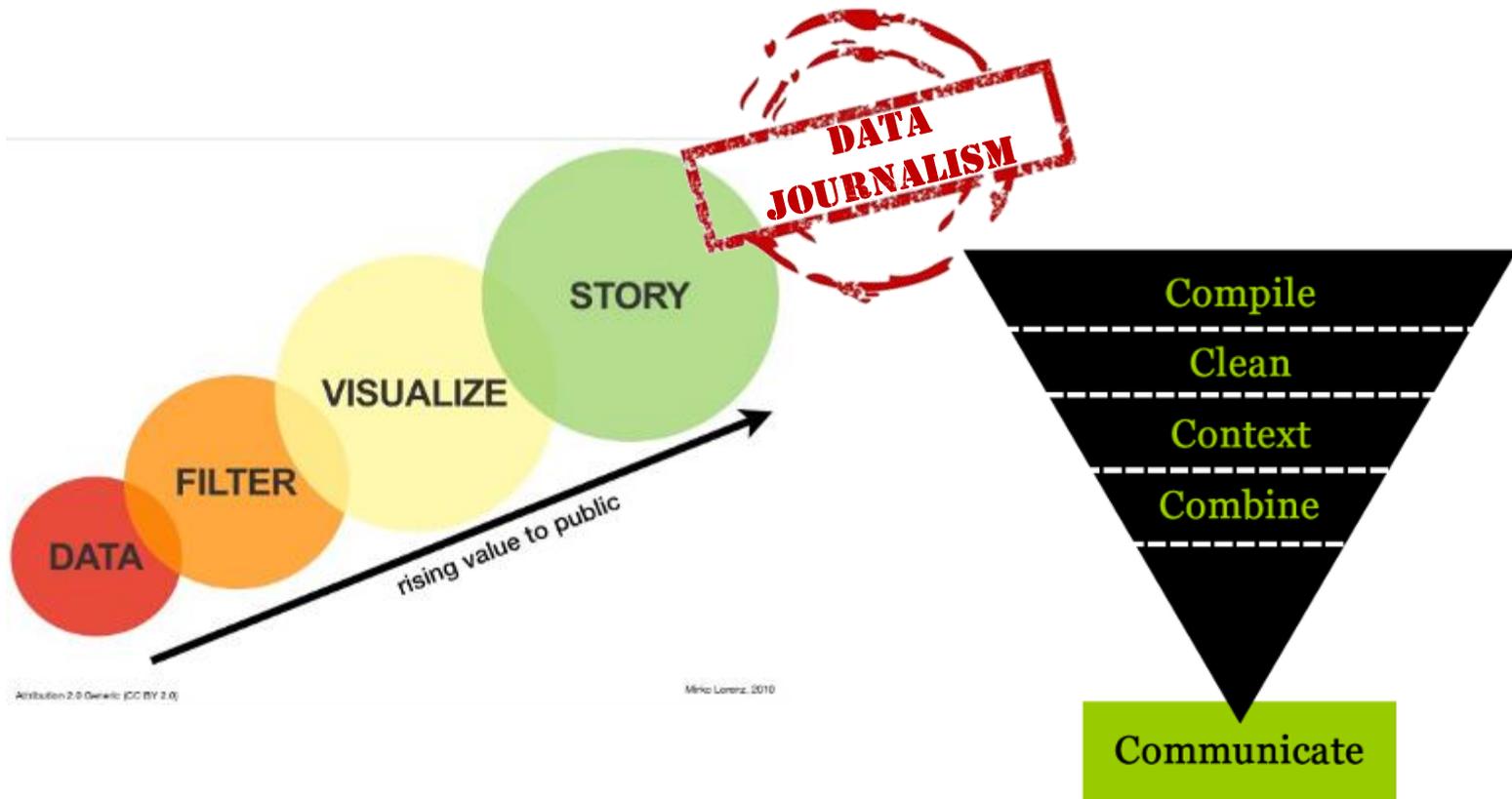
Domaines liés

- Faire le lien avec des **bases de connaissances**



Domaines liés

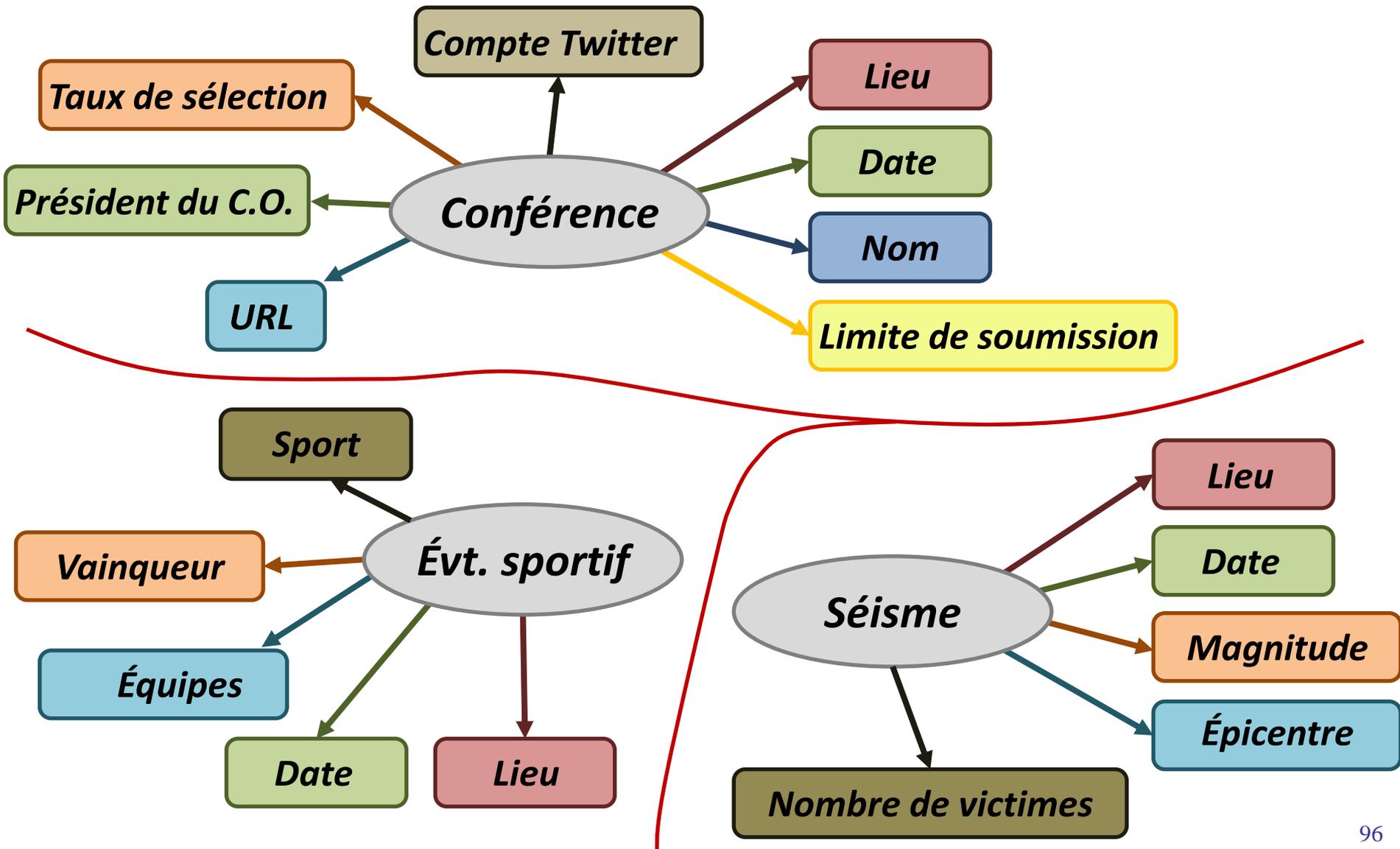
- Visualiser



Perspectives

- **À court terme**
 - Consolidation
 - Le tourbillon (maîtrisé) des projets...
 - Exemples...

Induction de schémas



Perspectives

- **À moyen terme**
 - **Évolution** des informations
 - Désignation des événements (coréférence longue distance)
 - Changements et nouveautés
 - **Journalisme de données**
 - Systématisation
 - Moins de supervision
 -
 - **Fact-checking**
 - **Formalisation** des représentations