



HAL
open science

Democracy Big Bang: What data management can(not) do for journalism

Ioana Manolescu

► **To cite this version:**

Ioana Manolescu. Democracy Big Bang: What data management can(not) do for journalism. 34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications, Oct 2018, Bucarest, Romania. , 2018. hal-01968347

HAL Id: hal-01968347

<https://inria.hal.science/hal-01968347>

Submitted on 2 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Democracy Big Bang

What data management can(not) do for journalism

Ioana Manolescu
Inria Saclay and Ecole polytechnique

MOTIVATION

Bad memories: Romania, 1989



Bad memories: Romania, 1989



Ceaușescu re-elected
at the 14th congress!

Bad memories: Romania, 1989



Ceaușescu re-elected at the 14th congress!

He was in power since 1965.

Bad memories: Romania, 1989



Bad memories: Romania, 1985



Things get better



... kind of



1000 dead (approx)
No one convicted.

Democratic societies crucially need the press

- To debate and express dissent
- To analyze, confirm or refute public statements
- To expose and explain society functioning



Fact-checking
(see next)



Democratic societies crucially need the press

- To debate and express dissent
- To analyze, confirm or refute public statements
- To expose and explain society functioning



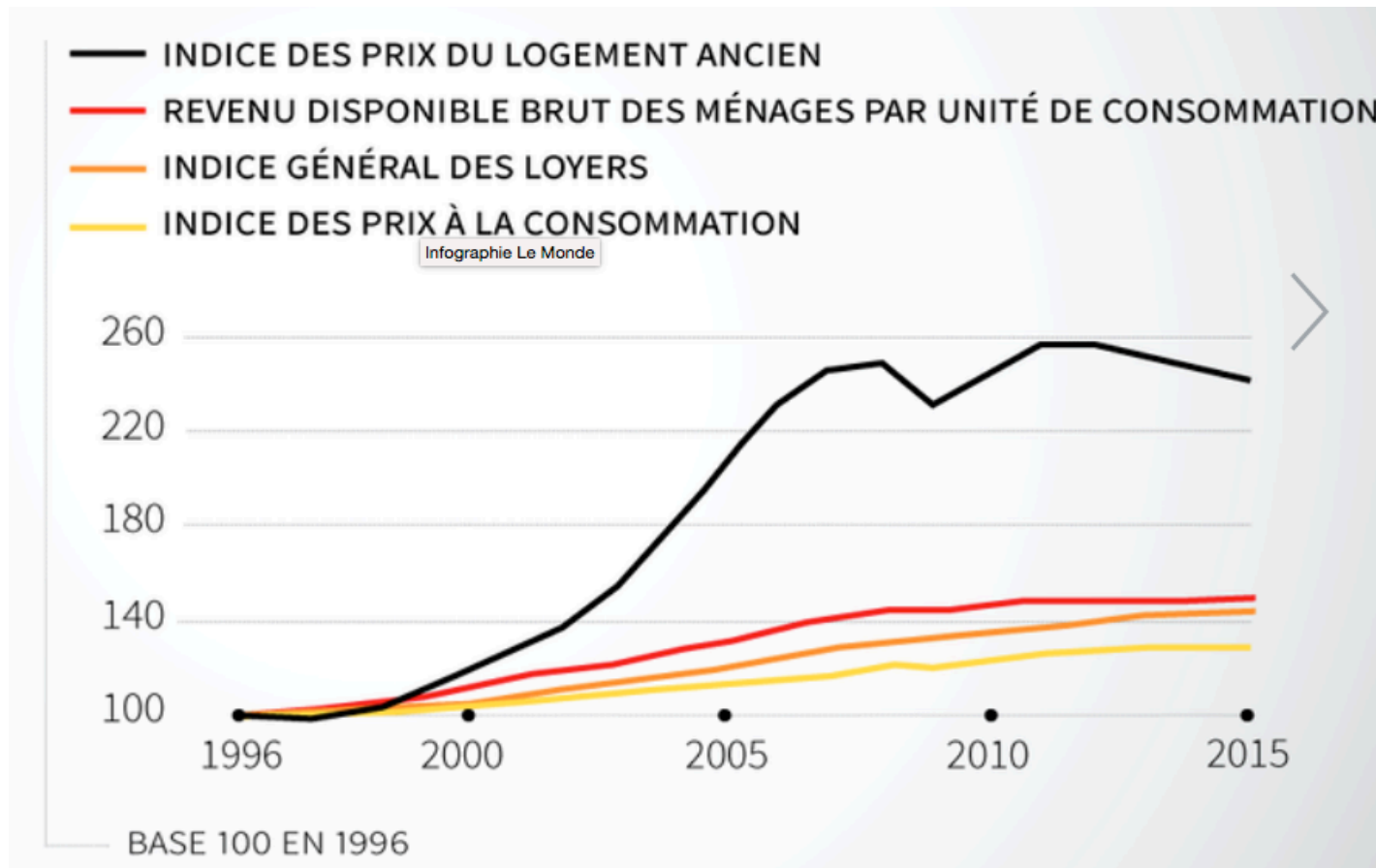
Fact-checking
(see next)



DATA JOURNALISM AND JOURNALISTIC FACT-CHECKING

Data journalism

Investigative journalism based on **complex and/or large data**



Data journalism

Panama Papers (International Consortium of Investigative Journalism, ICIJ)

The screenshot shows the ICIJ website page titled "Panama Papers The Power Players". The page features a navigation bar with the ICIJ logo and the title "The Panama Papers". Below the navigation bar, there is a search bar and social media icons. The main content area is divided into several sections. At the top, there is a filter bar with "All", "Europe", and "Oceania" selected. Below this, there is a "Country leaders" section with a sub-section for "Politicians/public officials". The main content area displays a grid of 10 portraits of individuals, each with a caption describing their role. The captions are: "Advised former Greek prime minister", "Former government minister in France", "Iceland's minister of finance", "Iceland's Interior Minister", "Malta's minister of energy and health", "Former member Hungarian Assembly", "Deputy chief justice of Kenya's Supreme Court", "Former Warsaw mayor, former member EU parliament", "Baroness and lifetime member of U.K. Parliament", and "Member of British Parliament".

ICIJ The Panama Papers

Panama Papers The Power Players

All Europe Oceania

Africa Latin America & the Caribbean

Asia Middle East

Country leaders Politicians/public officials Filters Europe Country

Advised former Greek prime minister

Former government minister in France

Iceland's minister of finance

Iceland's Interior Minister

Malta's minister of energy and health

Former member Hungarian Assembly

Deputy chief justice of Kenya's Supreme Court

Former Warsaw mayor, former member EU parliament

Baroness and lifetime member of U.K. Parliament

Member of British Parliament

Data journalism

Panama Papers (International Consortium of Investigative Journalism, ICIJ)

The screenshot shows a web browser displaying the ICIJ website. The URL is https://panamapapers.icij.org/the_power_players/. The page features a profile for Jérôme Cahuzac, a French politician, and a corporate structure diagram.

Jérôme Cahuzac
Budget minister at the Ministry of the Economy, Finance and External Trade (2012-2013); Deputy, National Assembly of France (1997-2002, 2007-2012)

Related countries
France

The lies told by Jérôme Cahuzac in 2013 triggered one of the most spectacular downfalls of a public official in the annals of French scandals. As a government minister waging a campaign against tax evasion, Cahuzac was forced to admit he lied to President François Hollande, former colleagues in Parliament and the French people when he repeatedly denied owning foreign bank accounts. He said he stashed over \$750,000 in a Swiss bank account for 20 years, moving the money to Singapore in 2009. His ex-wife disclosed an account opened in Great Britain in 1997. Cahuzac, who made a fortune as a cosmetic surgeon, resigned his ministry post and awaits trial for tax fraud.

Inside the Mossack Fonseca data » Offshore company held bank account for minister accused of tax fraud [Read more...](#)

[Offshore glossary](#)

Corporate Structure Diagram:

- MONFORT CAPITAL PARTNERS JLT (registered)
- CERMAN GROUP LIMITED (Beneficial owner)
- TALWAY INTERNATIONAL CORP (Shareholder)
- Jérôme Cahuzac (Beneficiary)
- Mr. Jerome Andre C. (Beneficiary)
- 85 avenue de Breteuil Paris
France (registered address)

Fact-checking (since 1930 approx.)

Fact-checking: verification of facts mentioned **in media content**

- ❑ To protect media reputation and avoid legal action

“The day I became a fact-checker at The New Yorker, I received **one set of red pencils** [...] for underlining **passages on page proofs of articles that might contain checkable facts.** [...] confirmed **with the help of reference books** from the magazine’s library”



<http://www.nytimes.com/2010/08/22/magazine/22FOB-medium-t.html>

Fact-checking (2012 – ongoing)


Not everyone agrees, however, that Democrats are not flip-flopping on the issue.

Mark Krikorian, executive director of the Center for Immigration Studies, a think tank that advocates for lower immigration, said that because the public doesn't know exactly what border barriers the Trump administration wants to build, Mulvaney's statement is not an "exact" comparison. But, he said, to dismiss it simply on that basis would be "tendentiously literal."

"The fact is that, other than the 'Mexico will pay for it' stuff, Trump is simply channeling the 2006 Secure Fence Act, and Schumer et al. who voted for it out of political calculation are indeed hypocrites for opposing the attempt to finally bring that law to fruition," Krikorian told us via email.

At the surface level, it is true in a broad sense that Democrats including Schumer, Obama and Clinton have in the past supported border fencing. All three voted for the Secure Fence Act of 2006, and all three supported the 2013 Senate immigration overhaul that passed the Senate, and which called for tougher border security including some additional fencing. But to claim that those measures are the same as what Trump is proposing is a stretch.

Share The Facts





Mick Mulvaney
Director, Office of Management and Budget

"We don't understand why the Democrats are so wholeheartedly against [President Trump's border wall]. They voted for it in 2006."

Fox News Sunday – Sunday, April 23, 2017

SHARE READ MORE







A Project of The Annenberg Public Policy Center

HOME ARTICLES ASK A QUESTION VIRAL SPIRAL ARCHIVES ABOUT US SEARCH MORE

THE WIRE

Did Democrats Once Support Border Wall?

By Robert Farley Posted on April 26, 2017

Like 835 Tweet Pin It Share 11

White House Office of Management and Budget Director Mick Mulvaney made an apples-to-oranges comparison when he said he couldn't understand why Democrats opposed supplemental funding for a border wall since many of them were for it back in 2006.

Mulvaney is referring to the Secure Fence Act of 2006, which called for construction of 700 miles of fencing and enhanced surveillance technology, such as unmanned drones, ground-based sensors, satellites, radar coverage and cameras. Sen. Chuck Schumer and then-Sens. Barack Obama and Hillary Clinton were among a bipartisan majority that voted in favor of the legislation, and it was signed into law by President George W. Bush.

In a very general sense, the Democrats named by Mulvaney supported a bill to build more

ASK FACTCHECK

Like 983 Tweet Pin It Share 98

Q: Did the Supreme Court rule that public schools cannot teach students about Islam?

A: No. That false claim was spread by a network of fake news websites.

It's not just **checking**

- Most aspects of modern reality are complex
- **Explaining** can be as important and useful as checking
 - Helps also analyze the future



The screenshot shows the 'LES DÉCODEURS' section of the Le Monde.fr website. The page features a dark header with the site's logo and navigation links. Below the header, there is a main navigation bar with categories like INTERNATIONAL, POLITIQUE, SOCIÉTÉ, ÉCO, CULTURE, IDÉES, PLANÈTE, SPORT, SCIENCES, PIXELS, and CAMPUS. The main content area is titled 'LES DÉCODEURS' in large, colorful letters, with the subtitle 'VENONS-EN AUX FAITS'. Below this, there is a sub-navigation bar with options like 'Datavisualisation', 'Vérification', 'Nanographix', 'Contexte', 'Evasion fiscale', and 'Le blog du Décodex'. The main article is titled 'Une sortie de l'euro ferait-elle exploser la dette française ?' and is written by Maxime Vaudano. A sidebar on the right contains a section titled 'Les décodeurs, mode d'emploi' which explains the site's mission: 'Les décodeurs du Monde.fr vérifient déclarations, assertions et rumeurs en tous genres ; ils mettent l'information en forme et la remettent dans son contexte; ils répondent à vos questions.' Below this sidebar is a link to 'LIRE CHARTER'.

Libération *Désintox*: cost of saving Cyprus

The image shows a screenshot of a web browser displaying a blog post on the 'Désintox' website. The browser's address bar shows 'desintox.blogs.liberation.fr/blog/'. The page header includes the 'DÉSINTOX le blog' logo and the 'Libération' logo. The main content is a blog post dated 'mardi 23 avril 2013' with the title 'Bobards en stock sur les plans de sauvetage européens'. The post text includes quotes from Nicolas Dupont-Aignan, Florian Philippot, and Marine Le Pen. On the left side of the page, there are social media sharing buttons for '0 Commentaires', '11 Tweet', and '34 Like'. On the right side, there are sections for 'À PROPOS DE CE BLOG', 'ALERTE EMAIL', and 'RECHERCHER'. Three blue arrows point from the left towards the text of the blog post, with the following annotations: '\$2-3 bn' pointing to the first quote, '\$10 bn' pointing to the second quote, and '\$15 bn' pointing to the third quote.

\$2-3 bn ←

\$10 bn ←

\$15 bn ←

Saving Cyprus: how much does it cost?

The article reads:

- ❑ The money is not **given** but **lent**
- ❑ The European Mechanism of Stability will lend **\$9bn**
- ❑ Out of which France contributes 20% or a bit less than **\$2bn**

However, things are complicated because

- ❑ The initial contribution of France to the EMU (\$16 bn) is counted toward French public debt
- ❑ The remainder contribution (\$124) is not



INTOX En période de rigueur budgétaire, les plans de sauvegarde successifs offrent depuis trois ans un boulevard aux détracteurs de l'euro et de l'Europe. Nicolas Dupont-Aignan et Florian Philippot ont ainsi, de concert, dénoncé récemment les milliards déversés à Chypre au moment où les Français se serrent la ceinture. «Les Français vont devoir donner 2 à 3 milliards d'euros pour des banques à Chypre. D'un côté on supprime les infirmières, on surtaxe les PME en France [...] mais quand il s'agit de donner de

l'argent de l'UE, c'est-à-dire des Français, à des banques pourries à Chypre, on le donne», dénonçait Nicolas Dupont-Aignan. En version Florian Philippot, cela donne : «C'est nous qui allons encore verser de l'argent puisque le MES s'est engagé à hauteur de 10 milliards d'euros, dont 2 milliards au titre de la France et des contribuables français». Un propos qui fait écho -notamment- à celui entendu quelque deux ans avant, dans la bouche de Marine Le Pen, au sujet cette fois du plan d'aide à la Grèce : «Comment peut-on imposer aux Français, aux classes populaires et moyennes, aux petites entreprises, de nouvelles taxes, des taxes sur les sodas, une hausse de la CSG de 550 millions d'euros, et d'un autre côté alimenter de 15 milliards d'euros supplémentaires en Grèce l'incendie de la zone euro ?»

DÉSINTOX Quel impact ont donc les plans de sauvetage successifs de l'Euro? Et le dernier en date, en direction de Chypre, va-t-il contraindre les Français à allonger 2 milliards d'euros?

Saving Cyprus: how much does it cost?

The article reads:

- ❑ The money is not **given** but **lent**
- ❑ The European Mechanism of Stability will lend **\$9bn**
- ❑ Out of which France contributes 20% or a bit less than **\$2bn**

However, things are complicated because

- ❑ The initial contribution of France to the EMS (\$16 bn) is counted toward French public debt
- ❑ The remainder contribution (\$124) is not



INTOX

En période de rigueur budgétaire, les plans de sauvegarde successifs offrent depuis trois ans un boulevard aux détracteurs de l'euro et de l'Europe. Nicolas Dupont-Aignan et Florian Philippot ont ainsi, de concert, dénoncé récemment les milliards déversés à Chypre au moment où les Français se serrent la ceinture. «Les Français vont devoir donner 2 à 3 milliards d'euros pour des banques à Chypre. D'un côté on supprime les infirmières, on surtaxe les PME en France [...] mais quand il s'agit de donner de

l'argent de l'UE, c'est-à-dire des Français, à des banques pourries à Chypre, on le donne», dénonçait Nicolas Dupont-Aignan. En version Florian Philippot, cela donne : «C'est nous qui allons encore verser de l'argent puisque le MES s'est engagé à hauteur de 10 milliards d'euros, dont 2 milliards au titre de la France et des contribuables français». Un propos qui fait écho -notamment- à celui entendu quelque deux ans avant, dans la bouche de Marine Le Pen, au sujet cette fois du plan d'aide à la Grèce : «Comment peut-on imposer aux Français, aux classes populaires et moyennes, aux petites entreprises, de nouvelles taxes, des taxes sur les sodas, une hausse de la CSG de 550 millions d'euros, et d'un autre côté alimenter de 15 milliards d'euros supplémentaires en Grèce l'incendie de la zone euro ?»

DÉSINTOX

Quel impact ont donc les plans de sauvetage successifs de l'Euro? Et le dernier en date, en direction de Chypre, va-t-il contraindre les Français à allonger 2 milliards d'euros?

Saving Cyprus: how much does it cost?

The article reads:

- ❑ The money is
- ❑ The European
lend **\$9bn**
- ❑ Out of which France contributes 20% or a bit less than **\$2bn**

The initial statements are mostly false

However, things are complicated because

- ❑ The initial contribution of France to the EMS (\$16 bn) is counted toward French public debt
- ❑ The remainder contribution (\$124) is not



INTOX En période de rigueur budgétaire, les plans de sauvegarde successifs offrent depuis trois ans un boulevard aux détracteurs de l'euro et de l'Europe. Nicolas Dupont-Aignan et Florian Philippot ont ainsi, de concert, dénoncé récemment les milliards déversés à Chypre au moment où les Français se serrent la ceinture. «Les Français vont devoir donner 2 à 3 milliards d'euros pour des banques à Chypre. D'un côté on supprime les infirmières, on surtaxe les PME en France [...] mais quand il s'agit de donner de

l'argent de l'UE, c'est-à-dire des Français, à des banques pourries à Chypre, on le donne», dénonçait Nicolas Dupont-Aignan. En version Florian Philippot, cela donne : «C'est nous qui allons encore verser de l'argent puisque le MES s'est engagé à hauteur de 10 milliards d'euros, dont 2 milliards au titre de la France et des contribuables français». Un propos qui fait écho -notamment- à celui entendu quelque deux ans avant, dans la bouche de Marine Le Pen, au sujet cette fois du plan d'aide à la Grèce : «Comment peut-on imposer aux Français, aux classes populaires et moyennes, aux petites entreprises, de nouvelles taxes, des taxes sur les sodas, une hausse de la CSG de 550 millions d'euros, et d'un autre côté alimenter de 15 milliards d'euros supplémentaires en Grèce l'incendie de la zone euro ?»

DÉSINTOX Quel impact ont donc les plans de sauvetage successifs de l'Euro? Et le dernier en date, en direction de Chypre, va-t-il contraindre les Français à allonger 2 milliards d'euros?

Saving Cyprus: how much does it cost?

The article reads:

- The money is
- The European
lend **\$9bn**
- Out of which France
less
than **\$2bn**

The initial statements are
mostly false

The complete picture is
complicated

However, things are complicated
because

- The initial contribution of France to the EMS
(\$16 bn) **is** counted toward French public
debt
- The remainder contribution (\$124) **is not**



INTOX

En période de rigueur budgétaire, les plans de sauvegarde successifs offrent depuis trois ans un boulevard aux détracteurs de l'euro et de l'Europe. Nicolas Dupont-Aignan et Florian Philippot ont ainsi, de concert, dénoncé récemment les milliards déversés à Chypre au moment où les Français se serrent la ceinture. «Les Français vont devoir donner 2 à 3 milliards d'euros pour des banques à Chypre. D'un côté on supprime les infirmières, on surtaxe les PME en France [...] mais quand il s'agit de donner de l'argent à l'UE, c'est-à-dire des Français, à des banques pourries à Chypre, ça se donne», dénonçait Nicolas Dupont-Aignan. En version

Florian Philippot, cela donne : «C'est nous qui allons encore verser de l'argent puisque le MES s'est engagé à hauteur de 10 milliards d'euros, dont 2 milliards au titre de la France et des contribuables français». Un propos qui fait écho -notamment- à celui entendu quelque deux ans avant, dans la bouche de Marine Le Pen, au sujet cette fois du plan d'aide à la Grèce : «Comment peut-on imposer aux Français, aux classes populaires et moyennes, aux petites entreprises, de nouvelles taxes, des taxes sur les sodas, une hausse de la CSG de 550 millions d'euros, et d'un autre côté alimenter de 15 milliards d'euros supplémentaires en Grèce l'incendie de la zone euro ?»

DÉSINTOX

Quel impact ont donc les plans de sauvetage successifs de l'Euro? Et le dernier en date, en direction de Chypre, va-t-il contraindre les Français à allonger 2 milliards d'euros?

Saving Cyprus: how much does it cost?

The article reads:

- ❑ The money is
- ❑ The European
lend **\$9bn**
- ❑ Out of which France
less
than **\$2bn**

However, things are complicated because

- ❑ The initial contribution of France to the EMS (\$16 bn) is counted toward French public debt
- ❑ The remainder contribution (\$124) is not

The initial statements are mostly false

The complete picture is complicated

I would not have been able to give the answer



INTOX En période de rigueur budgétaire, les plans de sauvegarde successifs offrent depuis trois ans un boulevard aux détracteurs de l'euro et de l'Europe. Nicolas Dupont-Aignan et Florian Philippot ont ainsi, de concert, dénoncé récemment les milliards déversés à Chypre au moment où les Français se serrent la ceinture. «Les Français vont devoir donner 2 à 3 milliards d'euros pour des banques à Chypre. D'un côté on supprime les infirmières, on surtaxe les PME en France [...] mais quand il s'agit de donner de l'argent à Chypre, c'est-à-dire des Français, à des banques pourries à Chypre, ça se donne», dénonçait Nicolas Dupont-Aignan. En version

«C'est nous qui allons encore verser de l'argent à hauteur de 10 milliards d'euros, à hauteur de 10 milliards d'euros, en France et des contribuables français». Un moment- à celui entendu quelque deux ans auparavant par Marine Le Pen, au sujet cette fois du plan d'aide à la Grèce : «Peut-on imposer aux Français, aux classes populaires et moyennes, aux petites entreprises, de nouvelles taxes, des taxes sur les sodas, une hausse de la CSG de 550 millions d'euros, et d'un autre côté alimenter de 15 milliards d'euros supplémentaires en Grèce l'incendie de la zone euro ?»

DÉSINTOX Quel impact ont donc les plans de sauvetage successifs de l'Euro? Et le dernier en date, en direction de Chypre, va-t-il contraindre les Français à allonger 2 milliards d'euros?

Saving Cyprus: how much does it cost?

The article reads:

- ❑ The money is
- ❑ The European
lend **\$9bn**
- ❑ Out of which France
less
than **\$2bn**

However, things are complicated because

- ❑ The initial contribution of France to the EMS (\$16 bn) **is** counted toward French public debt
- ❑ The remainder contribution (\$124) **is not**

The initial statements are mostly false

The complete picture is complicated

Forming a competent opinion is non trivial



INTOX En période de rigueur budgétaire, les plans de sauvegarde successifs offrent depuis trois ans un boulevard aux détracteurs de l'euro et de l'Europe. Nicolas Dupont-Aignan et Florian Philippot ont ainsi, de concert, dénoncé récemment les milliards déversés à Chypre au moment où les Français se serrent la ceinture. «Les Français vont devoir donner 2 à 3 milliards d'euros pour des banques à Chypre. D'un côté on supprime les infirmières, on surtaxe les PME en France [...] mais quand il s'agit de donner de l'argent à Chypre, c'est-à-dire des Français, à des banques pourries à Chypre, ça se donne», dénonçait Nicolas Dupont-Aignan. En version

«C'est nous qui allons encore verser de l'argent à hauteur de 10 milliards d'euros, pour sauver des banques à Chypre et des contribuables français». Un exemple à celui entendu quelque deux ans auparavant par Marine Le Pen, au sujet cette fois du plan d'aide à la Grèce : «Peut-on imposer aux Français, aux classes populaires et moyennes, aux petites entreprises, de nouvelles taxes, des taxes sur les sodas, une hausse de la CSG de 550 millions d'euros, et d'un autre côté alimenter de 15 milliards d'euros supplémentaires en Grèce l'incendie de la zone euro ?»

DÉSINTOX Quel impact ont donc les plans de sauvetage successifs de l'Euro? Et le dernier en date, en direction de Chypre, va-t-il contraindre les Français à allonger 2 milliards d'euros?

Do we need to understand?

"Populism is telling people that there are simple answers to complex problems"

<https://www.express.co.uk/news/world/1034797/matteo-salvini-italy-budget-crisis-european-union-eu-news-alto-adig>

Salvini's populism on the march as EU-backing north Italy turns on Brussels

<https://www.ft.com/content/53bf2caa-3d6c-11e8-b9f9-de94fa33a81e>

Populism in Europe

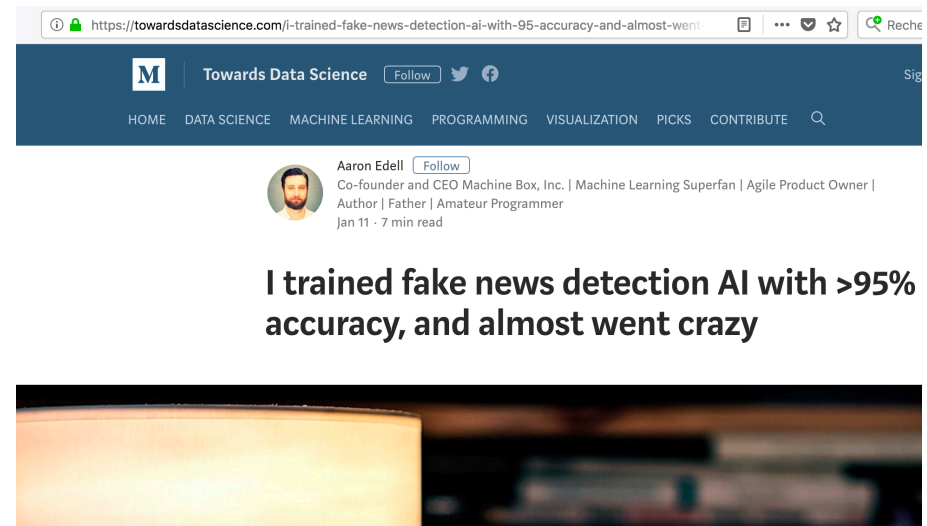
+ Add to myFT

Orban's populism prospers by challenging EU taboos

Hungary's prime minister leads the nationalist charge by challenging liberal taboos

Fact checking vs. fake news detection

- Fact checking is based on some background information source
- Fake news detection may or may not use a source
- E.g., text classifier (true, fake) trained with major news agency content, resp. known fake content (often virulent style)



A CONTENT MANAGEMENT PERSPECTIVE

Objectives

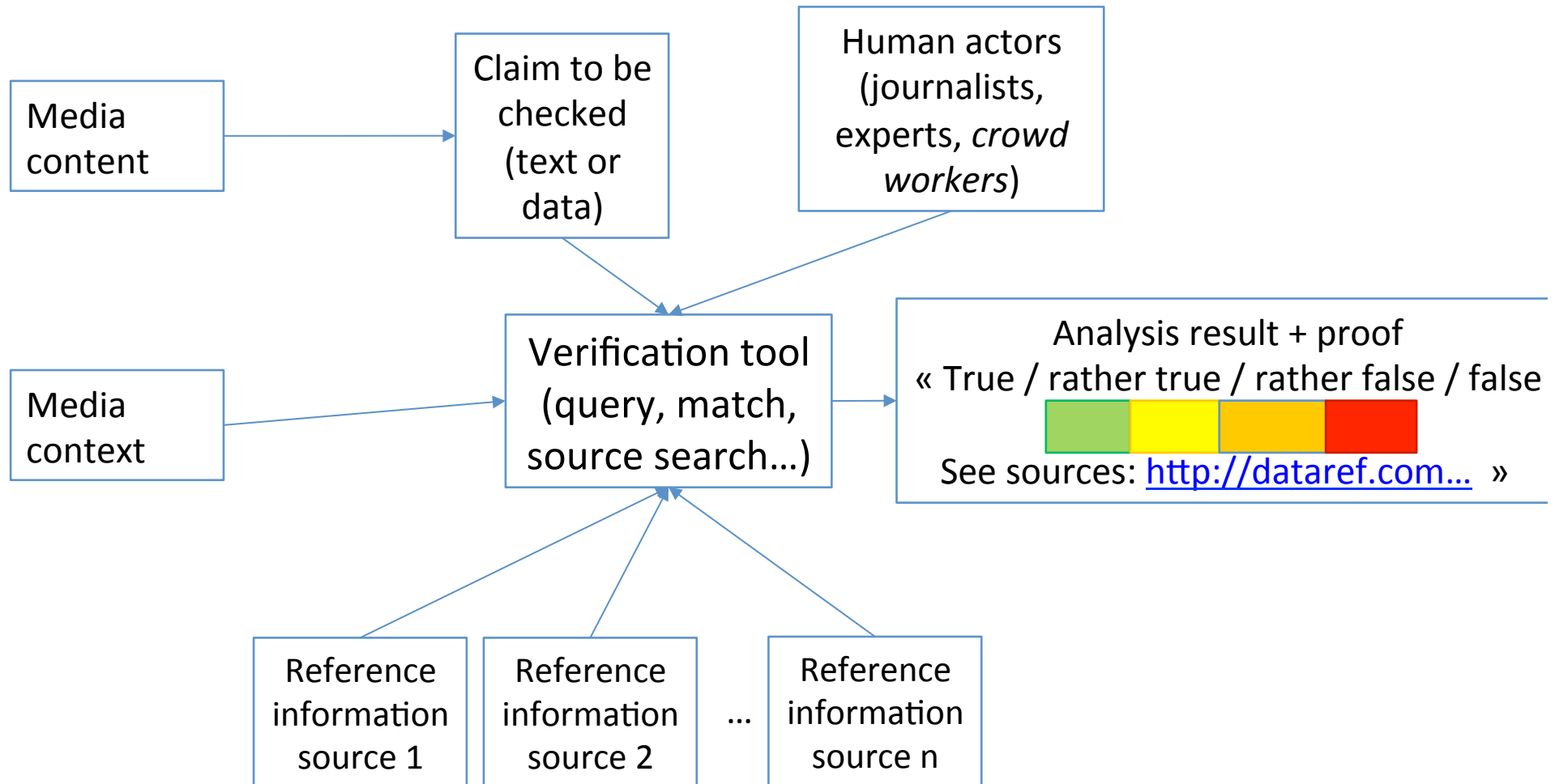
1. **Model fact-checking** through a data and information management perspective
2. Identify **existing tools and techniques** which could be directly applied
 - In a special journalistic context (see next)
3. Devise **new models, tools and techniques** for fact-checking and data journalism problems

ANR ContentCheck project

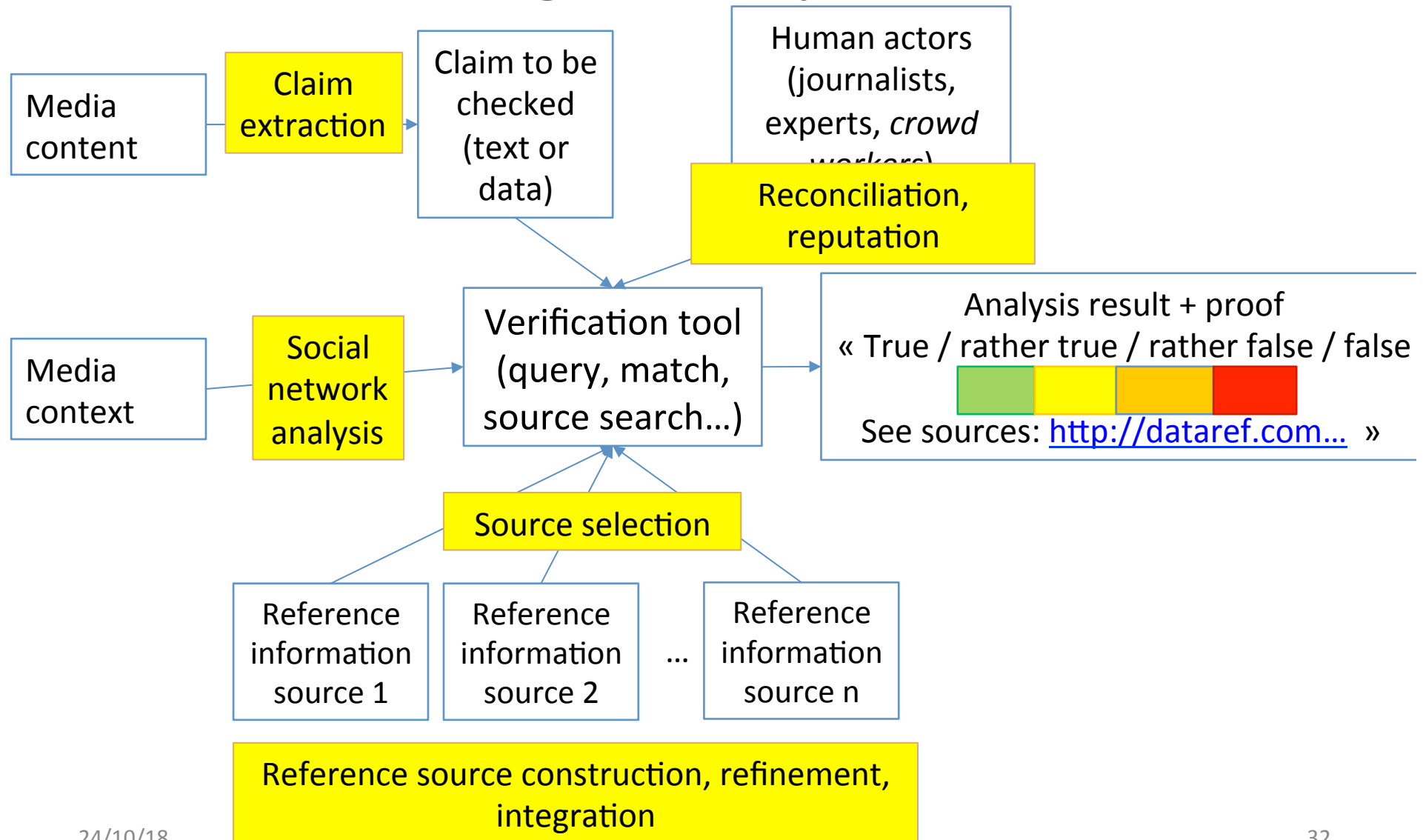


- 2016-2019
- Partners:
 - Inria Saclay: I. Manolescu
 - LIMSI (now Sorbonne Université): Xavier Tannier
 - U. Lyon: Sylvie Cazalens, Philippe Lamarre, Jean-Marc Petit, M. Plantevit
 - U. Rennes 1: François Goasdoué
 - Le Monde: Les Décodeurs
- Also: WebClaimExplain team with AIST Japan (Julien Leblay)

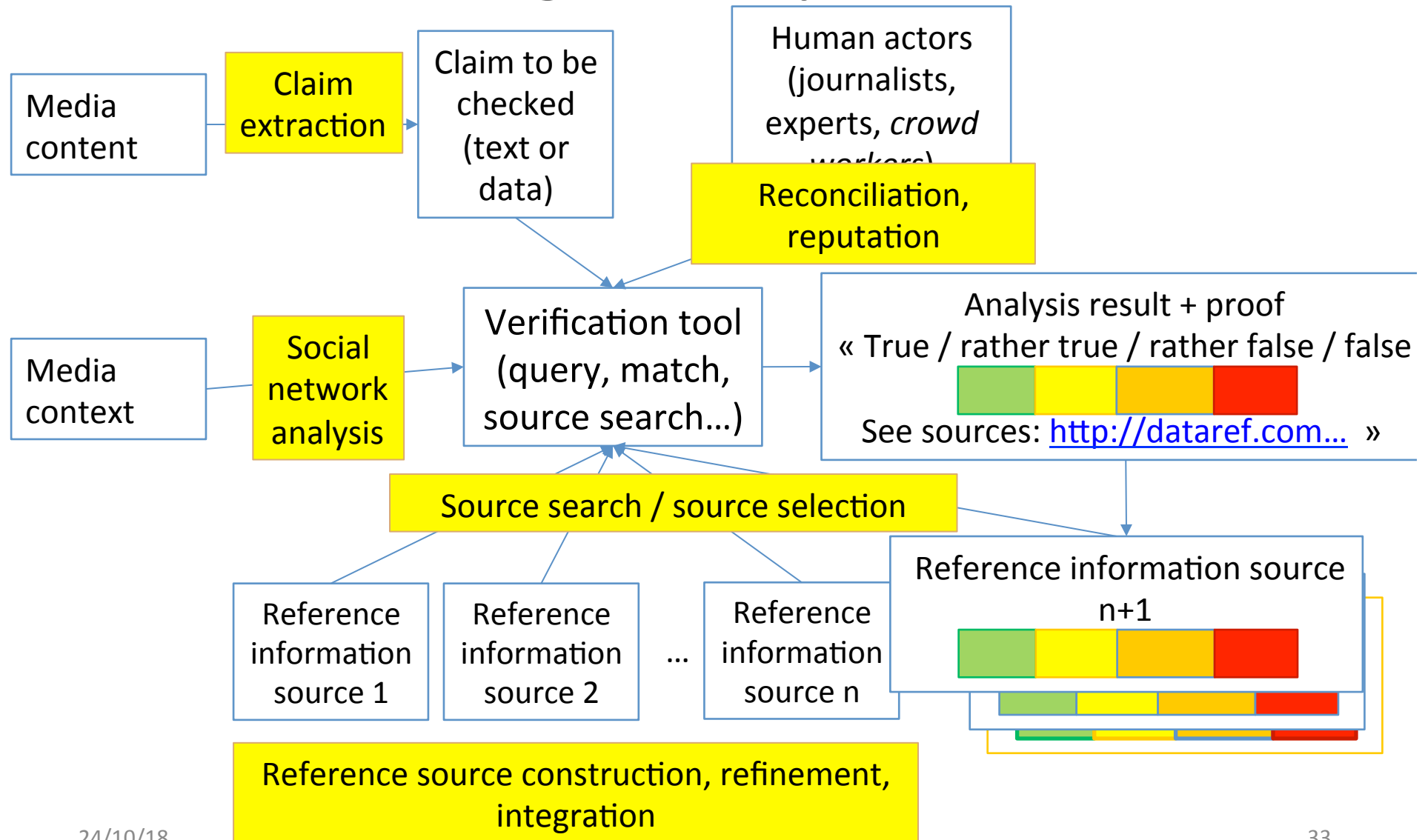
Fact-checking as a content management problem



Fact-checking as a content management problem



Fact-checking as a content management problem

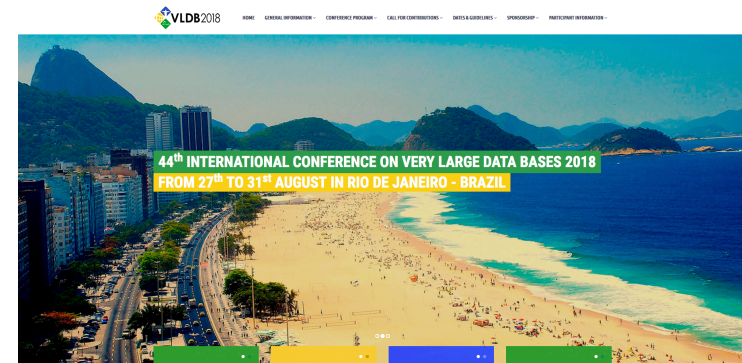


Fact-checking as a content management problem

[WWW2018] "A Content Management Perspective on Fact-Checking", S. Cazalens, J. Leblay, I. Manolescu, X. Tannier (fact-checking track)

[WWW2018 tutorial] "Computational fact-checking: problems, state of the art, and perspectives", J. Leblay, I. Manolescu, X. Tannier

[VLDB2018 tutorial] "Computational fact-checking: a content management perspective", S. Cazalens, J. Leblay, P. Lamarre, I. Manolescu, X. Tannier



Which existing technologies can be used?

- **Database queries?**
 - Journalists do not, generally, build databases.
 - It's seen as "not part of our job"
 - Persisting data is novel to some
 - SQL unfriendly
 - Not always helped by journal information systems
- Remarkable exception: Ouest France



Which existing technologies can be used?

- **Database queries?**
 - Journalists do not, generally, build databases.
 - It's seen as "not part of our job"
- **Curse of the coverage:** they need to cover (almost) any topic
- **Curse of noteworthiness:** write about hot topics of today (or tomorrow)
- They work under **strong time pressure**

Which existing technologies can be used?

- **IR/NLP?**
 - Better options
- The outcome of fact-check has to be **explainable**
 - Not (only) "an ML algorithm said so"
 - But they don't like to train ML systems, either.
- Extremely **picky on data sources**
 - "A majority of Web sources claim it" is not an option

Automatic source selection: FactMinder

[SIGMOD2013demo] F. Goasdoué, K. Karanasos, Y. Katsis, J. Leblay, I. Manolescu, S. Zampetakis: "Fact-checking and analyzing the Web"

- Plug-in
- Brings relevant data from document and knowledge bases

The screenshot displays the FactMinder interface. On the left, a news article snippet is shown with the headline "Bill Clinton apporte son soutien à Barack Obama". The main content area is partially obscured by a sidebar on the right. The sidebar contains several panels:

- Concepts:** A list of entities extracted from the document, including Bill Clinton (dbpedia:Person), Barack Obama (dbpedia:Person), Des Moines, Iowa (dbpedia:Place), Hillary Clinton (dbpedia:Person), Michelle Obama (dbpedia:Person), Bruce Springsteen (dbpedia:Person), and Pennsylvanie (dbpedia:Place).
- Related stories:** A list of related news items, such as "Ohio, Florida, ... être compliqué" and "Des indicateurs ... d'Obama".
- Facts & figures:** A panel titled "Curriculum" showing biographical information for Bill Clinton, including his birth date (August 19, 1946) and political party (Democratic Party).
- Quotes:** A panel containing quotes from the document, such as "Launching the Africa Regional Media Hub in Johannesburg" and "President Obama speaking LIVE for the last time before the election".
- Sources:** A panel listing the sources of the information, including "Daily Press Briefing: November 5, 2012".

- « Second screen »

Improving access to reference data sources

[BDA2018] T.-D. Cao, I. Manolescu, X. Tannier. "Extracting Linked Data from statistic spreadsheets" (WebDB 2018, SBD 2017)

“Créations d’entreprises en France en 2015” → we return:

Créations d'entreprises dans quelques pays de l'Union européenne en 2015

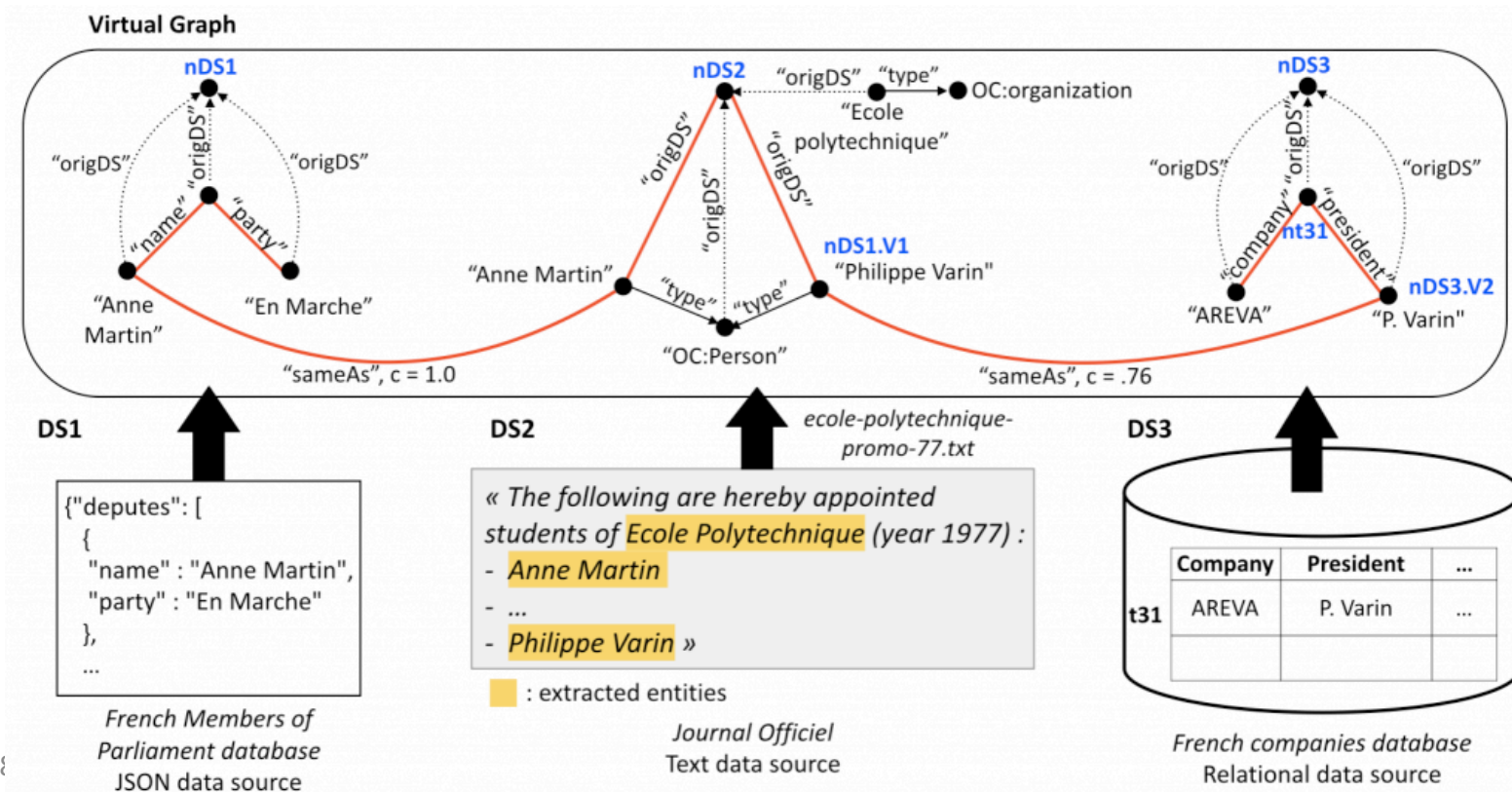
en %

Pays	Taux de création
Allemagne	7,1
Belgique	6,2
Espagne	9,5
France (1)	9,5
Italie	7,5
Pays-Bas	10,1
Portugal	15,7
République tchèque	8,2
Royaume-Uni	14,3

Keyword search across heterogeneous sources

[VLDB2018demo] C. Chaniel, R. Dziri, H. Galhardas, J. Leblay, M. Le Nguyen and I. Manolescu: "ConnectionLens: Finding Connections Across Heterogeneous Data Sources"

- Also today



Temporal beliefs:

A structured data model for fact-checking

With F. Goasdoué and L. Duroyon (U. Rennes 1)

Modeling who said what when

- "On Jan 22, 2017, Le Canard Enchaîné wrote that François Fillon had stated on Jan 21 that Penelope worked for the NA from 2002 to 2008."
- What has one actor heard of?
- Who knew this at that time?
- Who reversed their position on a topic?

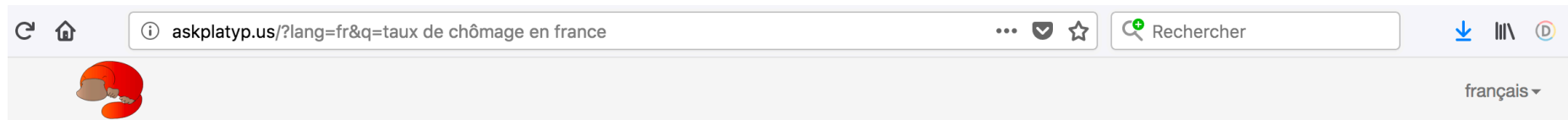
WRAP-UP

WHERE WE STAND, WHAT CAN WE DO?

Perspectives

- Any **question answering framework** which can be plugged on trusted data sources is useful
 - Usually no time nor skills to integrate the data
- No IT infrastructure can be counted upon
 - More "a jumble of tools"
- ML promises a lot but delivery is hard(er)
 - Also: journalists ambivalence
- Educate: the general audience, journalists, other social scientists
- Strong and increasingly IT-literate community
 - E.g. <https://www.poynter.org/channels/fact-checking>

NLP is hard



Demandez-moi n'importe quoi Version alpha



taux de chômage en france



Résultat

Après sept millions et demi d'années de calcul j'ai déterminé que la réponse est 42.

Cet outil de réponse aux questions est sponsorisé par [Lexistems SAS](#) et [ENS de Lyon](#).

Is this worth it?

“Some people will never be convinced”

- Correct.
- “Facts have a liberal bias” (Paul Krugman)
- **Source:** <https://www.nytimes.com/2017/12/08/opinion/facts-have-a-well-known-liberal-bias.html>
- Scientists and humanity scholars believe in a constructed, logical discourse, and believe **humans yield to reason**. Businesspeople **know this is not true**, in general. Businesspeople have thus an **advantage** in winning political competitions
George Lakoff, former Berkeley professor
<https://georgelakoff.com/2016/11/22/a-minority-president-why-the-polls-failed-and-what-the-majority-can-do/>
- Conspiracy theory adepts believe two obviously contradicting theories [Wood et al., 2012]

THANK YOU / QUESTIONS?

[HTTP://CONTENTCHECK.INRIA.FR/](http://contentcheck.inria.fr/)