OPSCIDIA

The value of knowledge

Can Open Access help fight fake news ?

NISO PLUS 2021

Presentation of Opscidia

Scientific articles – a huge source of knowledge but many locks

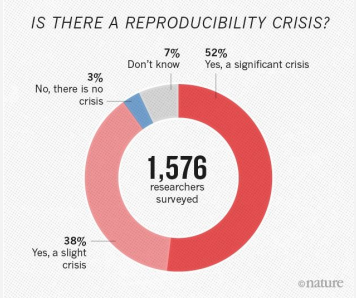
Access :

Can I access information ?

OPEN ACCESS

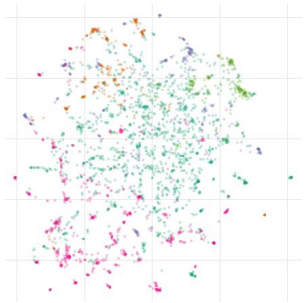
Reproducibility :

Is information trustworthy ?



Discoverability :

Can I find the information I need ?



Our purpose *: to ease the reusability of scientific results in society as a whole !*

What we do

Open Access Publishing

Opscidia - SciTA Scientific Text Analysis

Diamond Open Access

- CC-BY Licence
- APC Free
- Open Source (OJS)

SaaS platform

- Concept detection
- Smart navigation
- Standard dashboards
- Links to research articles

Projects

- Data fusion
- Weak signal detection
- Personalized dashboards
- Automated flux

Scientific “fact-checking”

Our “Science-checking” prototype

Does an *Agent X*

Cure
Cause
Prevent

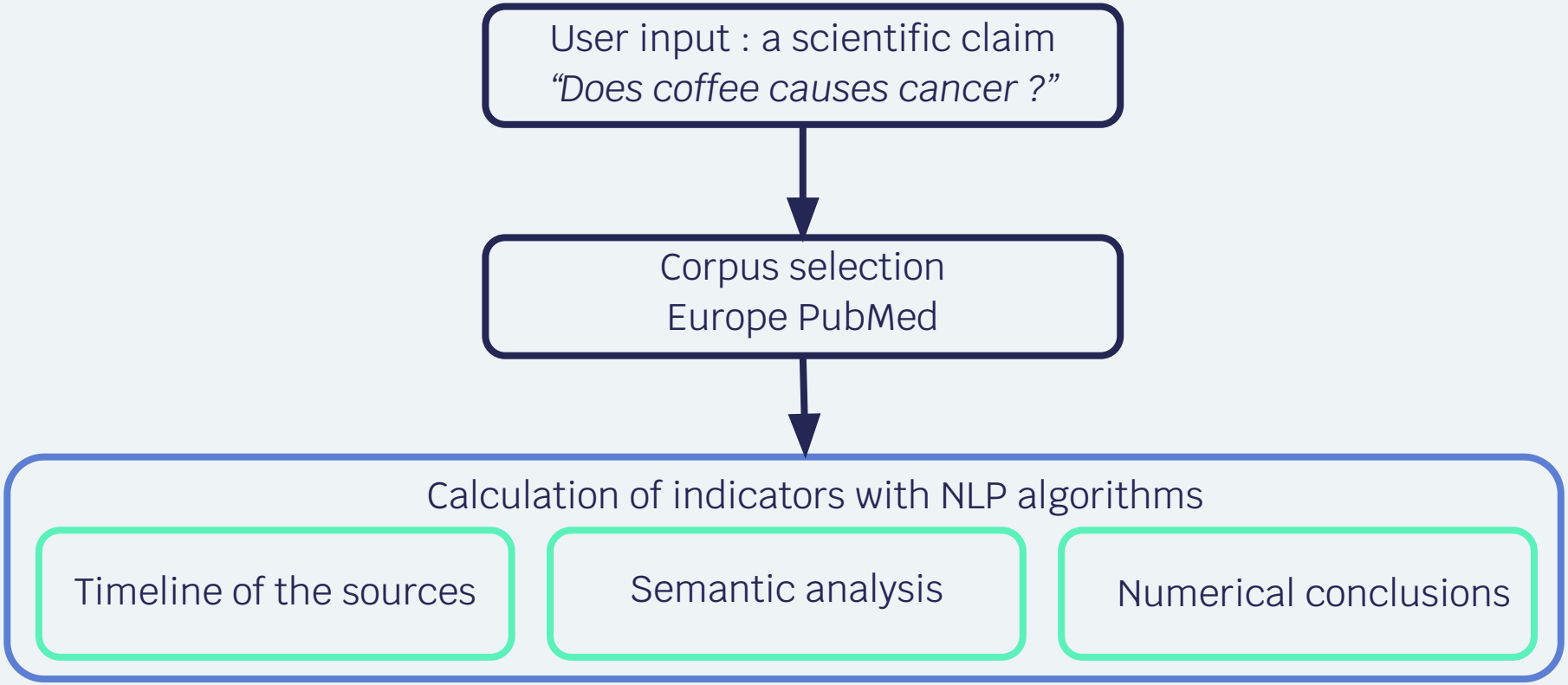
a Disease Y ?

Based on EPMC data

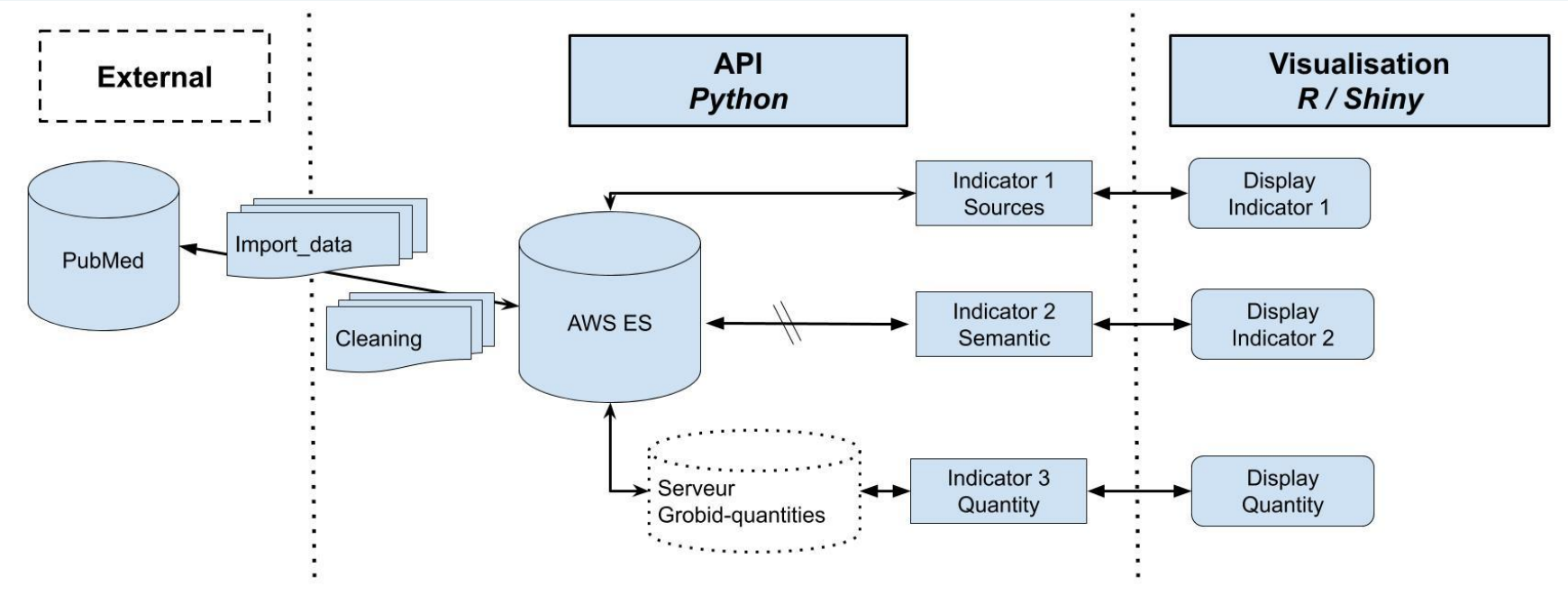
Objectives

- **Demonstrate** the use of Open Access to fight fake news
- **Educate people** - There is no “debate killer study”
- Help users identify the **scientific consensus** (or lack thereof)

Our appraoch



Integrated pipeline – General architecture



Science Checker

What does the scientific consensus say about it ?

Does [Sport] [prevent] [cancer] ?

Check

Corpus analysis

Result : Warning

Number of articles: 120 articles (including 20 Open Access)

Published between: 1990 to 2019

Show graph

Boolean QA

Result : OK

Answers

Retrieval of significant values

Result : OK

159 CI values retrieved

Statistical conclusion: 59 % of CI show statistical significance

Show graph

Indicator 1 : analysis of the sources

A standard method to fight fake news

In standard “fake news” detection :

- difficult to know whether a statement is true
- Easier to find where it comes from
- Sometimes enough to tell the whole story !

Examples : Flying saucer, a riot looks like a riot, project Reveal

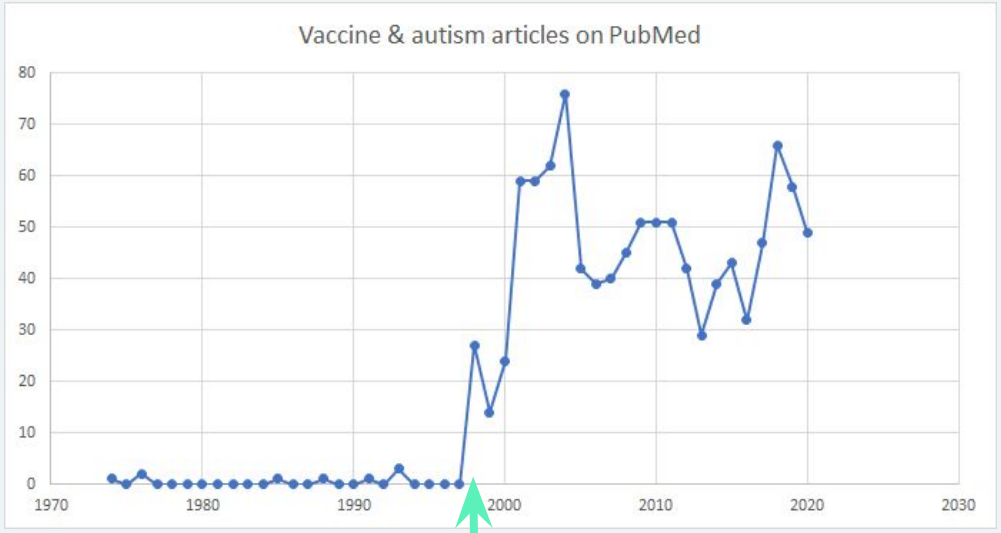
In the scientific field :

Example : **Vaccines & autism**

- We know where it comes from…
- Realize the massive amount of research resources to counter a successful (but retracted) article

Conclusion from this indicator :

- Is it still discussed
- When did it start?
- Is it widely covered



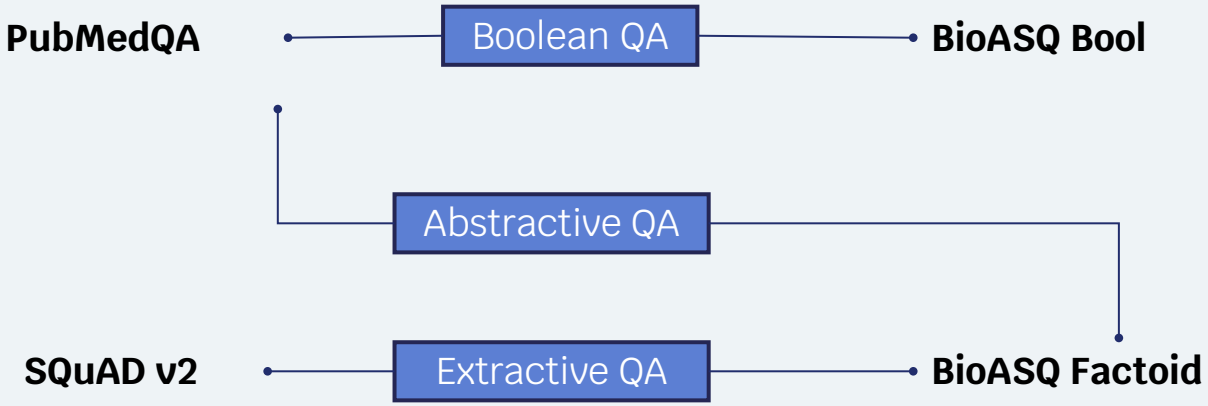
1998 Lancet MMR article by Andrew Wakefield - **retracted**

Indicator 2 : Semantic analysis

Semantic classification

Task : **classify research articles** wrt the **input statement**

- Supporting ? Neutral ? Contradicting ?



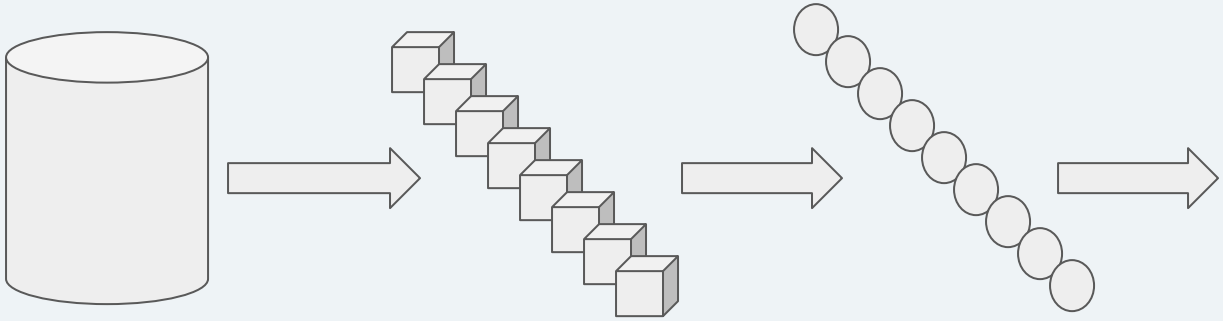
Results

Modèle	Transformer	Données	F1 - score	Accuracy
Boolean	$RoBERTa_{BASE}$	PubMedQA	77,98	97,01
	$RoBERTa_{LARGE}$		82,12	98,36
Extractive	$BERT_{LARGE}$	SQuAD	77,26	
	$RoBERTa_{BASE}$		74,02	92,64
Abstractive	$T5_{SMALL}$	PubMedQA		98,33
	$T5_{BASE}$			98,80

Results on test BioASQ Bool sample

Indicator 3 : Numerical data retrieval

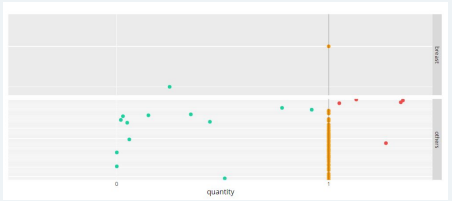
Numerical data retrieval



Copus of relevant articles

Identification of the right sentences

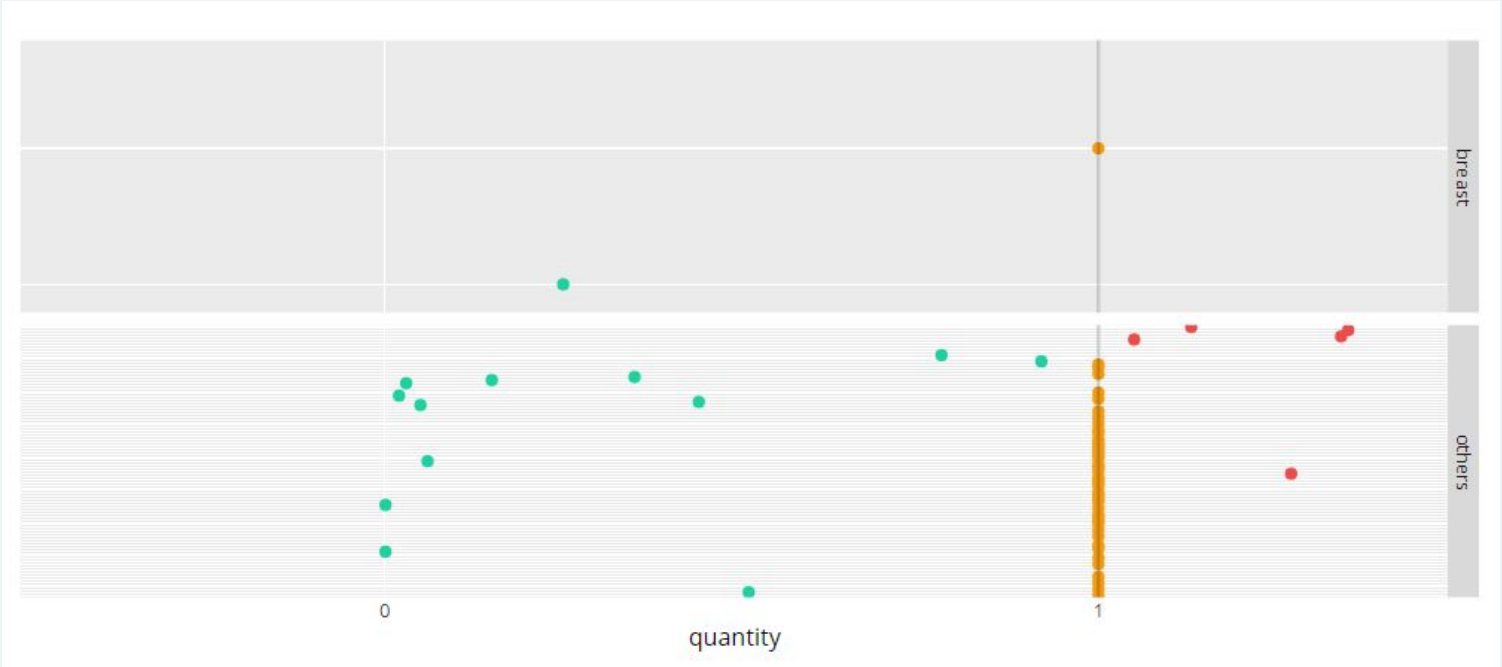
Extraction of the numbers



Visualization of the results

Numerical data retrieval

Does sport prevent cancer?



Conclusion and next steps

- We have built a pipeline based on 3-indicators
  - To detect scientific consensus
  - To help the public understand what is a research article
- Open Access can be useful to fight fake news
- Next step : online integration (prototype level)
- Looking for collaboration & funding to go further

Aknowledgements

For the good work !
Charles Letaillieur, Loic Rakotoson, Timothée Babinet

For funding !
The Vietsch Foundation

Thank you!

OPSCIDIA

The value of knowledge

Contacts

sylvain.massip@opscidia.com