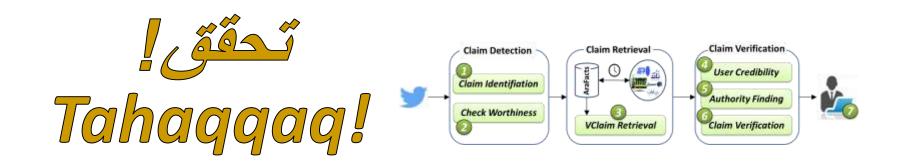
Doha Data Forum
Oct 29<sup>th</sup> 2023



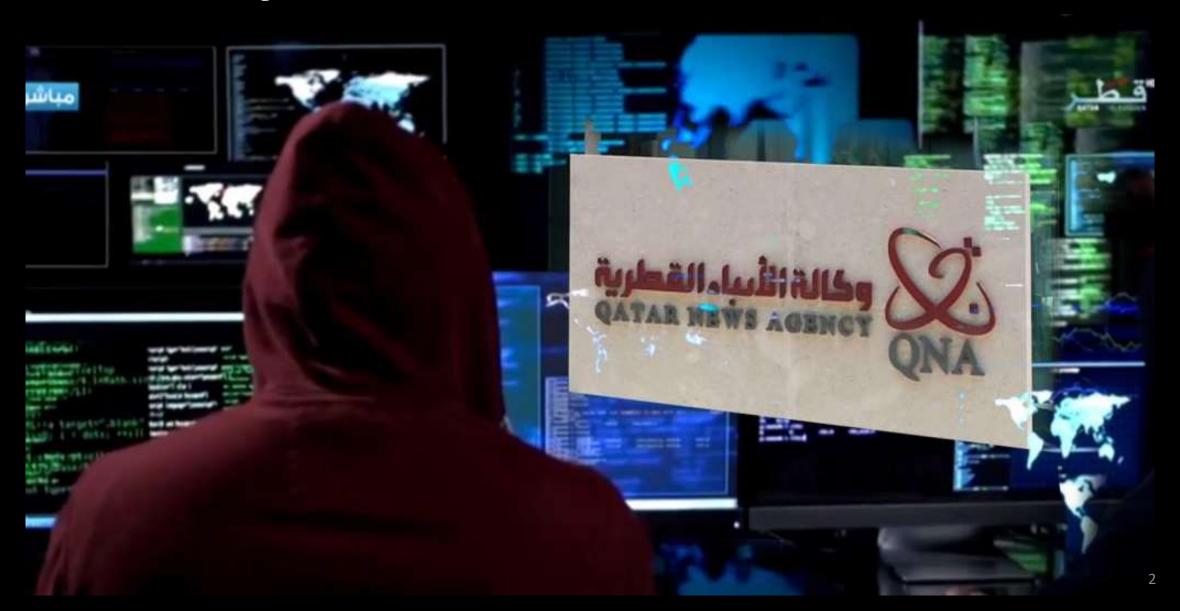
# A Real-Time System for Assisting Twitter Users in Arabic Claim Verification

**Tamer Elsayed** 





# ... on May 24<sup>th</sup> 2017 ...





# Why ...



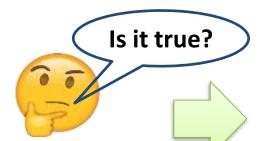
#### Internet, Communication & Electricity are cut off from Gaza

محرقة في غزة أمام نظر العالم الجثث في الشوارع وتحت الأنقاض قطع الاتصالات والانترنت والكهرباء شلّ المدينة ولا يستطيع أحد طلب الإسعاف أو الإنقاذ

الْقصف يتم من الجو والبحر والبر سيستفيق العالم على مذبحة رهيبة

#Gaza\_Genocide







#### Translate post







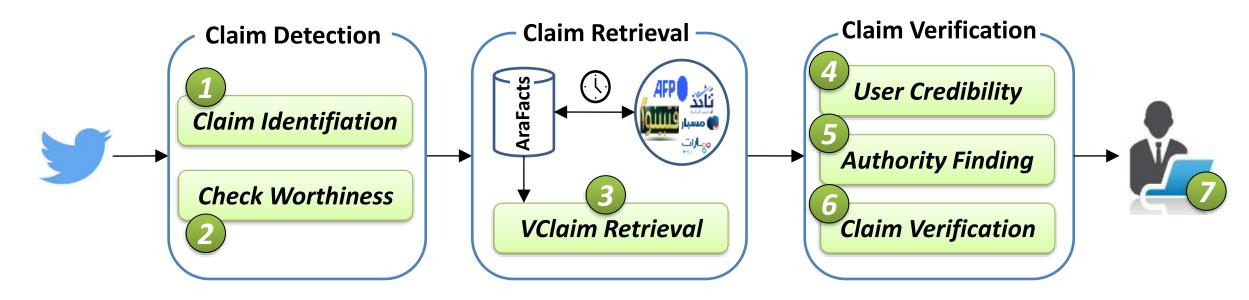




### Our Solution ...







#### **Arabic**



### Data is much needed!

Data for building our models

Data for supporting decision of models

Data for testing our models

Data for running a real-time system

#### **Arabic**

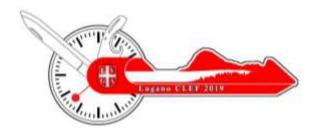


## CheckThat! Lab (2018-Current)

# Research & <u>Data</u> Campaign @











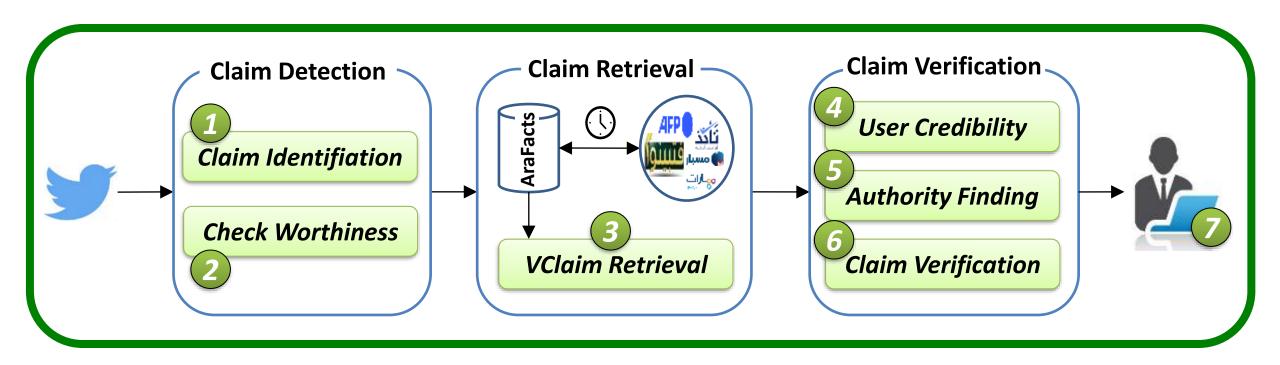






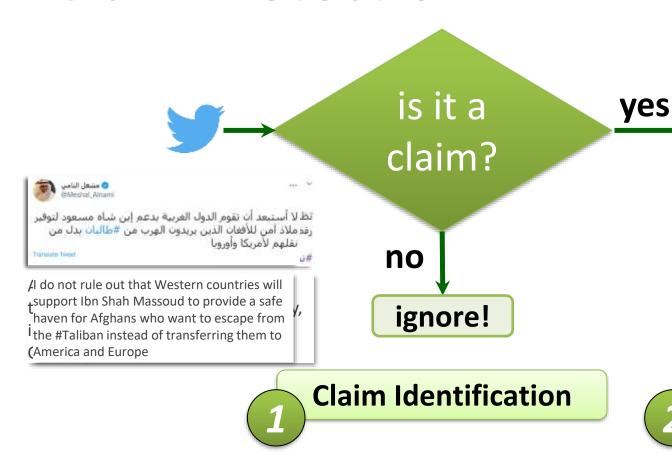


### Explore ...









is it check worthy?



Check-worthy Detection







4,7



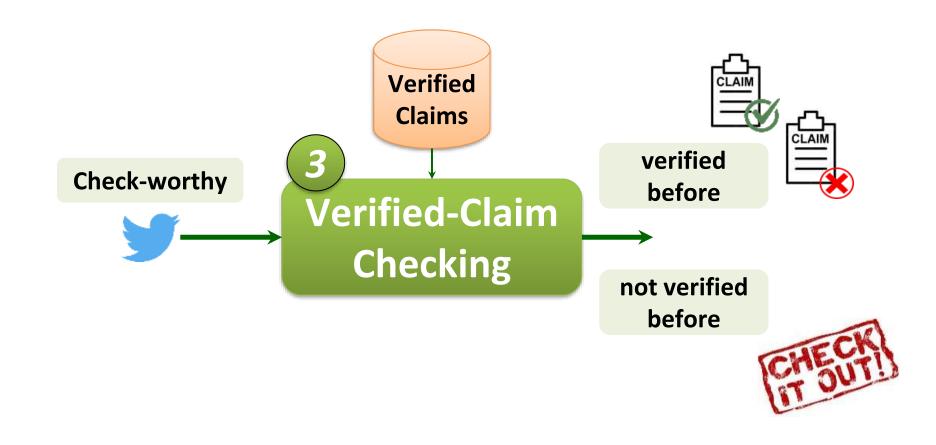








### **Verified-Claim Checking**

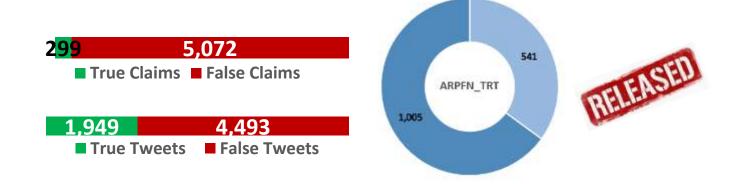






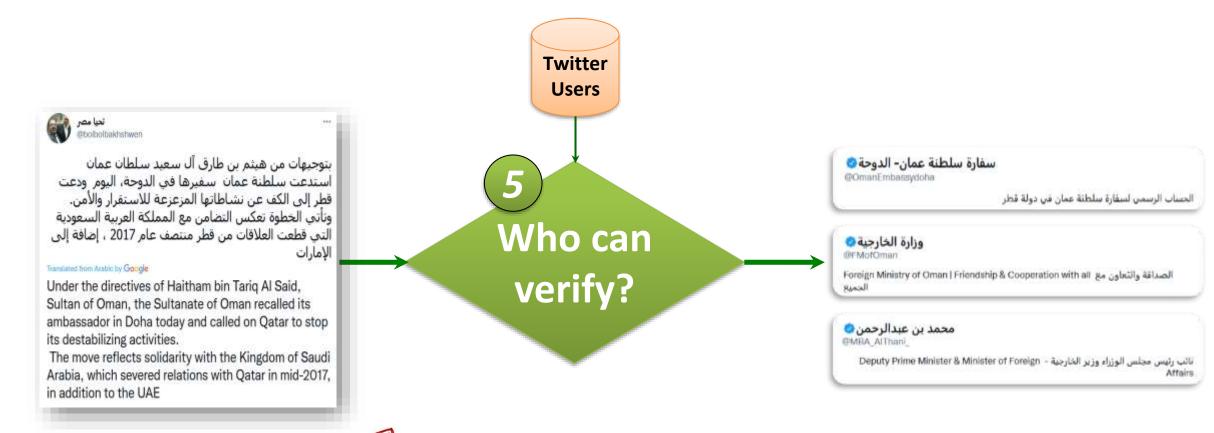
### Is that user prone to spread fake news?





## **Authority Finding in Twitter**





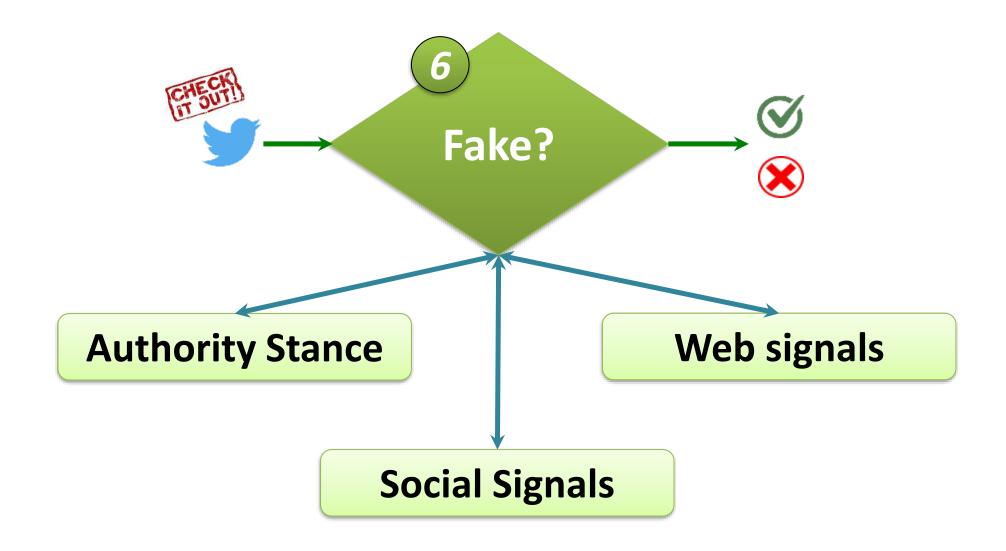








### **Claim Verification**





# **Authority Stance**







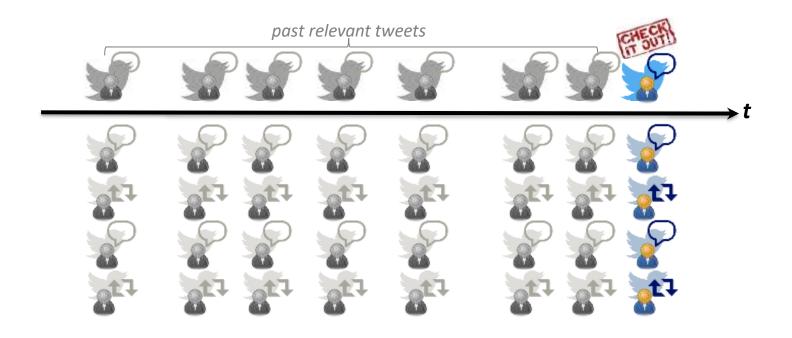








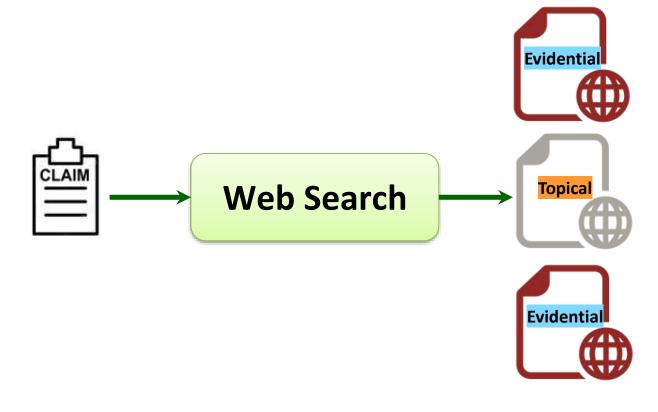
### **Social Signals**







### **Web Signals**







## A real-time system: Tahaqqaq





#### تحقق من مصداقية مغرد Verify user credibility

معرفة ما إذا كان المغرد معرضاً لنشر أخبار مزيفة أم لا.

Estimating the credibility of a Twitter user.



#### تحقق من ادعاء Claim Verification

التحقق من صحة إدعاء منتشر عن طريق التتبؤ بمصداقية الادعاء وما إذا كان قد ثم التحقق منه مسبقا بمعرفة هيئات التحقق مع تقديم قائمة بالمصادر المتعلقة التي يمكن التحقق من خلالها.

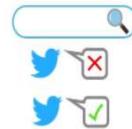
Verifying a spreading plain-text claim by estimating its veracity, whether it was previously-verified, and presenting relevant authority accounts.



#### تحقق من تغريدة Tweet Verification

التحقق من صحة إدعاء مكتوب في تغريدة عن طريق التتبؤ بمصداقية الادعاء ومصداقية المخرد وما إذا كان قد تم التحقق منه مسبقا بمعرفة هيئات التحقق مع تقديم قائمة بالمصادر المتطقة التي يمكن التحقق من خلالها.

Verifying a spreading claim expressed in a tweet by estimating its veracity, the credibility of the posting user, whether it was previously-verified, and presenting relevant authority accounts.



#### ابحث وتحقق Search & Verify

البحث في تويتر وترتيب النتائج حسب أولوية التحقق والتحقق من صحة التغريدات النائجة من البحث.

Searching Twitter and sorting the results by check-worthiness of the resulting tweets, and verifying those search results.



### **Example: Search & Verify**

(2)

(3)

4

(6)

### Translated results

Turkey opens the gates of Ataturk Dam and water reaches Iraq

The tweet contains a claim

The claim is check-worthy









#### The claim looks false

The user is not prone to spread fake news

#### Relevant previously verified claims

 (False) <u>Turkey releases the water from Ilisu Dam</u> <u>towards Iraq</u> [AFP fact check, 09-02-2023]

#### Relevant authorities to help verify the claim

- The Iraqi Media Network (@iraqmedianet)
- Media Directorate of the Popular Mobilization Forces (@teamsmediawar)
- Turkish Presidency (@tcbestepe\_ar)

#### Arabic results

تركيا تفتح بوابات سد أتاتورك ووصول المياه الى العراق

تحتوى التغريدة على ادعاء

الادعاء يستحقق التحقق منه











المغرد يبدو غير معرض لنشر أخبار مزيفة

#### ادعاءات ذات صلة تم التحقق منها

• (خاطئ) تركيا تطلق المياه من سد اليسو تجاة العراق [ في ميزان فرانس برس ,2023-02-09]

#### مصادر متعلقة يمكن التحقق من خلالها

- شبكة الإعلام العراقي (@<u>iraqmedianet</u>)
  - مديرية الإعلام هيئة الحشد الشعبي
     (teamsmediawar@)
    - الرئاسة التركية (@tcbestepe\_ar)



ابحث عن تغريدات

▼ خيارات متقدمة



Acknowledgment





### Takeaway messages

Lots of news on social media need checking!

- Constructed several Arabic datasets
- Built SOTA models for several components
- Developed a live real-time system with integrated components
- Collaborated with AlJazeera

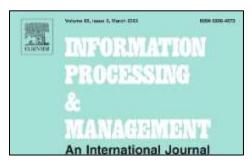


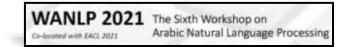














O 78

Cut off of Internet, Communication & Electricity from Gaza محرقة في غزة أمام نظر العالم لجثث في الشوارع وتحت الأنقاض قطع الاتصالات والانترنت والكهرباء شلّ المدينة ولا يستطيع أحد طلب الإسعاف أو لقصف يتم من الجو والبحر والبر ليستفيق العالم على مذبحة رهيبة Translate post



مقبرة جماعية للشهداء في قطاع عزة؛ بعد تكدس جثامين شهداء القصف الإسرائيلي على

🖸 Oct 15 🔾 وسيا الأن 🛇 🛇 🗞 وسيا الأن







