

## Project Title

### Live News Demo

#### Project Description:

Gathering news from different sources saves a lot of time and shows different opinions.

There are hundreds of Arabic news sources, and most of them post news as RSS feeds and they are updated during the hour. There is no tool for browsing ALL Arabic news that allows users to filter news by country, category, topic, date, etc.

Please have a look at the presentation slide below!

#### Duties/Activities:

Monitor Arabic news sources and collect RSS feeds continuously. Archive news collection.

Build a nice interface that allows user to browse and filter news.

Integrate Arabic tools to search for keywords, extract named-entities, cluster related news, etc.

#### Required Skills:

Good programming skills (Python/Java/C#)

#### Preferred Intern Academic Level:

Under-grad students are welcomed!

#### Learning Opportunities:

Learn about data collecting and cleaning, build a customized interface to show updated news on a map, archive news in a database, and integrate Arabic tools for searching, keyword extraction, and clustering.

**Expected Team Size:** *it is preferable to have team projects*

3 to 4

#### Mentors

| Name:              | email:   |
|--------------------|--|
| Hamdy Mubarak      | <a href="mailto:hmubarak@hbku.edu.qa">hmubarak@hbku.edu.qa</a>   |
| Dr. Ahmed Abdelali | <a href="mailto:aabdelali@hbku.edu.qa">aabdelali@hbku.edu.qa</a> |
| Dr. Preslav Nakov  | <a href="mailto:pnakov@hbhu.edu.qa">pnakov@hbhu.edu.qa</a>       |

# Live News Demo

- Mentors: Hamdy Mubarak, Ahmed Abdelali, Preslav Nakov
- **Motivation:** Gathering news from different sources saves a lot of **time** and shows different **opinions**.

There are hundreds of Arabic news sources, and most of them post news as **RSS** feeds and they are **updated** during the hour. There is no tool for **browsing ALL** Arabic news that allows users to filter news by country, category, topic, date, etc.

Ex: <http://feeds.bbc.co.uk/arabic/rss.xml>  
 (title, description, link to full article, date, image, etc.)

```

<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0" ?>
  <channel ?>
    <title>
      <![CDATA[
        مصدر كردي: "التوصل لاتفاق" بشأن نشر قوات سورية في نظريين
      ]]>
    </title>
    <description>
      <![CDATA[
        مصادر بالقرات الكردية السورية في شمال سوريا تقول ان التفاوض ابرم مع تنظيم
        بشأن دخول الجيش النظامي السوري منطقة عفرين للمساعدة في صد الهجوم التركي
      ]]>
    </description>
    <link>http://www.bbc.co.uk/arabic/middleeast-43108425</link>
    <guid isPermaLink="true">http://www.bbc.co.uk/arabic/middleeast-43108425</guid>
    <pubDate>Mon, 19 Feb 2018 01:55:36 GMT</pubDate>
    <media:thumbnail width="976" height="549" url="http://c.files.bbc1.co.uk/8406/production/_100083453_6b9e-469b-bde3-83adf113c0fb.jpg"/>
  </item>
</rss>
  
```

## ➤ Solution

- Monitor Arabic news sources and collect RSS feeds **continuously**
- **Archive** news collection
- Build a nice **interface** that allows user to browse and filter news
- Integrate **Arabic tools** to search for keywords, extract named-entities, cluster related news, etc.

## ➤ Skills Needed:

- Good programming skills (Python/Java/C#)
- Ability to learn new challenges!

