

News mining for food security: The case of Burkina Faso

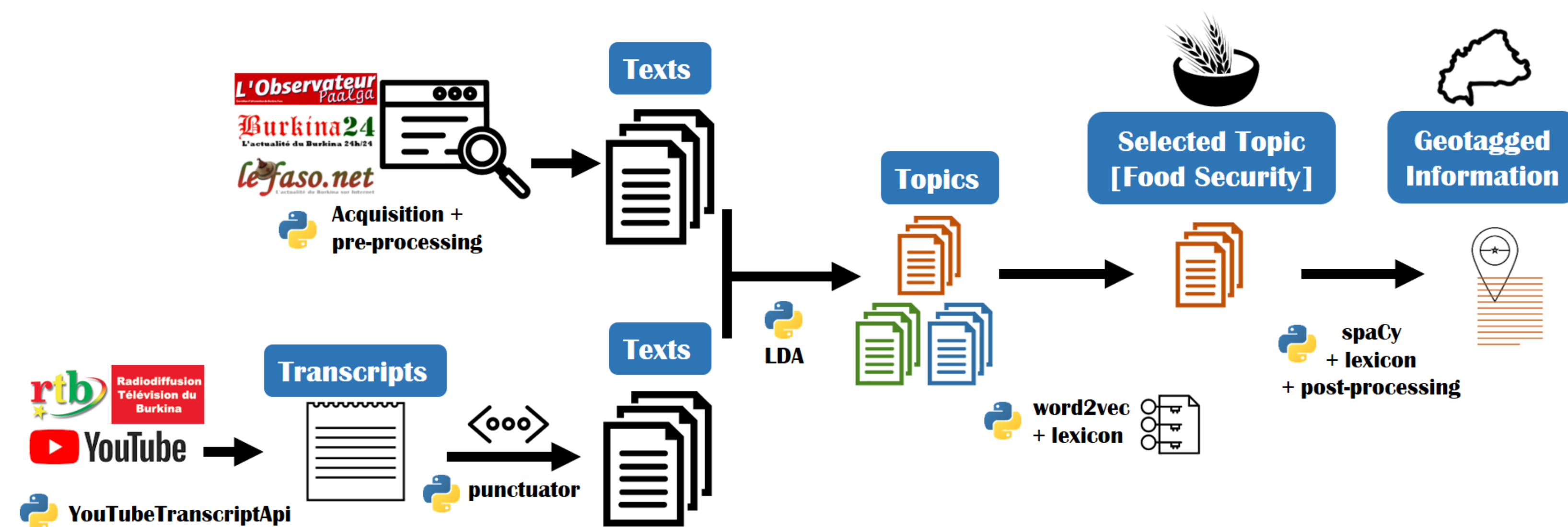
Hugo Deleglise^{1,2} Camille Schaeffer¹
 Abdoulaye Cissé¹ Roberto Interdonato^{1,2}
 Mathieu Roche^{1,2} Agnes Begué^{1,2}

¹ TETIS, Univ. Montpellier, AgroParisTech, CIRAD, CNRS, INRAE, Montpellier, France
² CIRAD - Centre de coopération internationale en recherche agronomique pour le développement, Montpellier, France

- Food security remains an actual problem in West Africa, were the current Food Security Monitoring Systems (FSMS) put in place by state organizations and NGOs are not enough to detect all problems at local scale
- FSMS generally rely on satellite imagery and household surveys
- We investigate to what extent complementary information can be found on news media, and how this textual information can be linked to the one issued by satellite imagery
- We focus on the case of Burkina Faso, where we target two type of sources:
 - local newspapers (l'Observateur Paalga, Burkina24 and Lefaso.net)
 - Youtube channel of the official national broadcaster (RTB - Radio Télévision du Burkina)



| Newspaper | Starting date | #articles |
|--------------------|---------------|--------------|
| Observateur paalga | 2012 | 2, 019 |
| Burkina24 | 2011 | 6, 365 |
| Lefaso.net | 2003 | 23, 229 |
| TOTAL | | 31, 613 |
| Youtube channel | Starting date | #transcripts |
| RTB | 2011 | 12, 432 |



General schema of the proposed news mining framework

Analysis on local newspapers

Top 10 words for topic #1: ['types', 'publique', 'chrétien', 'africain', 'céréales', 'santé', 'plantes', 'alimentaires', 'aliments']
 Top 10 words for topic #2: ['repas', 'nutrition', 'service', 'hôpital', 'alimentation', 'place', 'malades', 'ceni', 'jour', 'yalgado']
 Top 10 words for topic #3: ['blessés', 'cause', 'maladie', 'camion', 'symptômes', 'insuffisance', 'cardiaque', 'douleur', 'travail', 'facteurs']
 Top 10 words for topic #4: ['culture', 'public', 'ouédraogo', 'oeuvres', 'bureau', 'édition', 'musique', 'artiste', 'kundé', 'artistes']
 Top 10 words for topic #5: ['personnes', 'moment', 'ouagadougou', 'circulation', 'route', 'véhicules', 'usagers', 'véhicule', 'ouédraogo', 'ville']

Example of topic selection. Latent LDA-topics described by their top-10 keywords. Green topics are retained as more semantically correlated (higher word2vec) to the Food Security lexicon collected from the website *Cultivoo*.

| Observateur Paalga | Burkina24 | LeFaso.net |
|--------------------|-------------------|-------------------|
| Ouagadougou 88 | Ouagadougou 119 | Ouagadougou 91 |
| Ouaga 30 | Centre 33 | Centre 68 |
| Centre 18 | Ouaga 19 | Sahel 26 |
| Pouytenga 13 | Bobo 17 | Bobo-Dioulasso 22 |
| Bobo 12 | Nouna 15 | Diallo 11 |
| Sahel 12 | Bobo-Dioulasso 10 | Gaoua 11 |
| Bobo-Dioulasso 10 | Sahel 10 | Baskuy 10 |
| Banfora 10 | Tanghoun 9 | Bobo 9 |
| Logobou 9 | Centre-Est 9 | Koubri 9 |
| Cascades 9 | Nord 9 | Nord 9 |

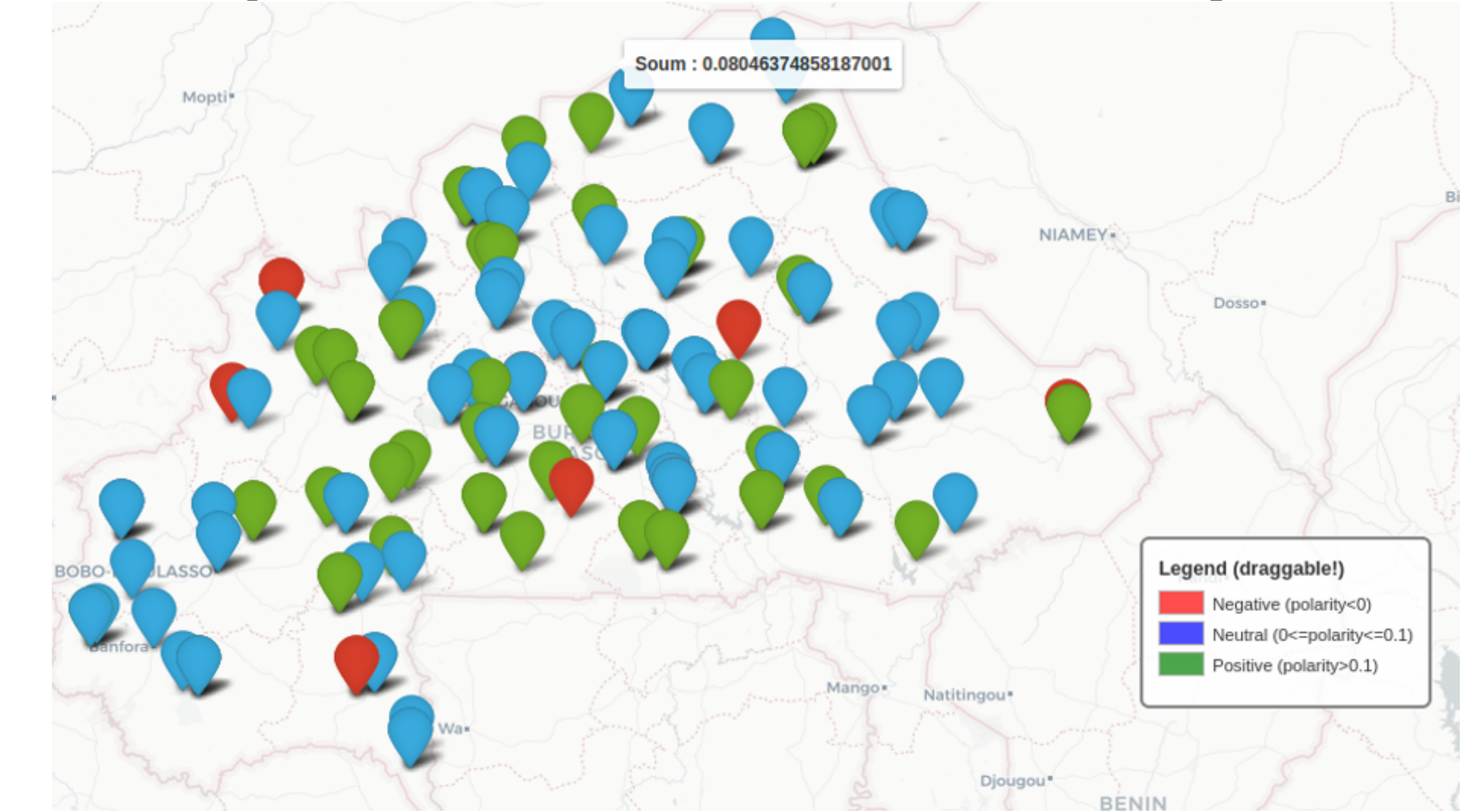
Top-10 most frequent spatial entities (identified with spaCy) for the top-50 articles (as ranked by word2vec) on the three local newspapers.

Qualitative Evaluation. We asked two domain experts to answer these questions for the top-10 ranked articles :

- Is the article about food security in Burkina Faso? (Yes, Yes - partially, No)
- Does the text contain additional information of interest on food security? (Yes, Yes - partially, No)

While most articles did result to be in topic, there were several examples of articles which did not actually add information of interest (e.g., advertisement of governmental events and initiatives about food security and agriculture).

Analysis on Youtube transcriptions



Average polarity (as computed by TextBlob) for the different Burkinabé cities mentioned in the Youtube transcriptions

| Sentence | Polarity | Subjectivity |
|--|----------|--------------|
| Région de la Boucle du Mouhoun à cette période de l'année, les produits maraichers inonde ce marché est également connu pour sa fourniture en poissons. | 1.0 | 0.6 |
| Un projet voit le jour à Bobo Dioulasso et se propose de faire de la recherche développement sur la gestion écologique de la fertilité des sols pour une agriculture intelligente combinant polyculture et élevage dans le contexte des changements climatiques. | 0.9 | 0.2 |
| Le fonds d'investissement Burkinabé synergie Burkina et l'accompagnement de Lézinnes Agrocert dirigé à monsieur chat Cassano sont deux formidable exemple de réussite et la principauté de Monaco d'annoncer que son accompagnement va s'étendre dans les années à venir. | 0.85 | 0.9 |
| C'est le moins que l'on puisse dire ici, car, ayant bravé une trentaine de kilomètres d'une route en très mauvais état, nous sommes sur la vie et les amours, y quelque part dans la province du Kéné Dougou de jour comme de nuit. | -1.0 | 1.0 |
| Vous ne pouvez pas rester et du helmand dans votre endroit dans votre villa, dans votre ville, à Ouagadougou , et vous n'êtes pas identifié, c'est très compliqué. | -0.8 | 1.0 |
| Ces logements sont le fruit du partenariat entre l'état Burkinabé et le promoteur, lui fait mal et ayem seuil des 1000 Go et Murray Harold Hatoum. | -0.7 | 0.7 |

Sentences including spatial entities showing Top-3 and Bottom-3 polarities (as computed by TextBlob)

Travaux communautaires: les jeunes du **Nahouri** donnent l'exemple et ici, vous avez récemment emprunté la route, comme pas la tv, les vous avez su **le Mans** fait le constat de l'état de dégradation de la route. Une portion de cette route a atteint un niveau de dégradation qui occasionnent à de graves incidents. Les populations ont décidé de mettre fin à leur calvaire 1 avant que la saison des pluies ne s'installe véritablement. Et depuis ce 25 mai, les populations de cette partie **du Nord** ri se sont mobilisés pour effectuer des travaux de réfection de la partie dangereuse de la route nationale numéro 25 à **Dourdan ouattara**, et où l'eau potable la scène se passe sur la ksk en **Pala Tiébélé**. Visiblement, ces populations du **Nahouri** sont motivés pour donner un nouveau visage à cette partie de la rn 25, une route qui était dans un état de dégradation avancé, causant ainsi plusieurs accidents, dont un décès à l'éclair les plus de l'année dernière, la voie était déjà parti et il nous laisse un petit morceau pour que sa puce est divisé.

Example of preprocessed Youtube transcript. Punctuation was added using the *punctuator* Python library. Spatial entities were identified using the *SpaCy* Python library.



Link to the original video on the RTB Youtube channel from where the above transcript was generated