

# The SmartMedia news recommender system for mobile devices

Cristina Marco<sup>1</sup>, Arne Dag Fidjestøl<sup>2</sup>, Jon Atle Gulla<sup>3</sup>, Jon Espen Ingvaldsen<sup>4</sup>, Peng Liu<sup>5</sup>, Özlem Özgobek<sup>6</sup>, Lemei Zhang<sup>7</sup>

Norwegian University of Science and Technology, Trondheim, Norway

E-mail: <sup>1</sup>cristina.marco@ntnu.no, <sup>2</sup>adf@ntnu.no, <sup>3</sup>jag@ntnu.no, <sup>4</sup>jon.espen.ingvaldsen@ntnu.no, <sup>5</sup>peng.liu@idi.ntnu.no, <sup>6</sup>ozlem.ozgobek@idi.ntnu.no, <sup>7</sup>lemei.zhang@idi.ntnu.no

*Abstract:* This paper describes a news recommender system prototype, developed within the SmartMedia program at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. This prototype makes use of geographical, temporal and other contextual information coming from open source knowledge bases in order to personalize news information in mobile devices.

**Keywords:** SmartMedia, news, recommender system, mobile devices

## 1. INTRODUCTION

As media companies are changing from paper to digital-based publications, there is a growing need for personalization to provide the most relevant digital content to readers. News recommender systems procure such solutions by filtering more to-the-point information. News personalization is particularly critical in mobile devices, as users are now clearly favouring mobiles in order to get informed about current events. According to the Knight Foundation, around 90 per cent of the U.S. population now access news and information via mobile devices, either through apps, browsers and increasingly through social networks.\* Due to the portability and smaller screen of mobile devices, where only a fraction of the most relevant news can appear, digital media experiences need to be shaped differently in desktop computers and mobile devices. Mobile devices demand for more personalized and better targeted news filtering.

In this paper, we present ongoing work on a prototype of a mobile news recommender system developed within the SmartMedia program<sup>†</sup> at NTNU Trondheim in Norway [1,2,3].

## 2. THE SMARTMEDIA NEWS RECOMMENDER SYSTEM

The SmartMedia news recommender system takes a multi-perspective approach by combining temporal, locational, and preferential information to provide a more fine-grained recommendation strategy. It utilizes entity definitions and associations in Wikidata<sup>‡</sup> and Geonames<sup>§</sup> open source knowledge bases to build a semantic representation of the underlying news text.

### 2.1. Interface

Figure 1 illustrates the user interface of this mobile news recommender system prototype. This user interface allows users to explore news story clusters on an interactive map and retrieve local news streams for any region of the world. The user can extract news items that are near to a specific point in space, close to their personal interests and at a given point in time. These three features can be selected by users by clicking on the buttons at the bottom of the screen so that news articles are retrieved and ranked according to their preferences.

There are three ranking factors that are used in order to personalize the news retrieval: geographical location, that defaults to the GPS location of the user, personal interests, that are covered by a set of stereotypical users (such as stock trader, soccer fan, technology geek, etc.), and time, that defaults to the time of publication of the news article. The three buttons at the bottom of the screen allow users to change these settings in order to personalize their news retrieval experience.

\* <https://medium.com/mobile-first-news-how-people-use-smartphones-to>

† <https://www.ntnu.no/wiki/display/smartmedia/SmartMedia+Program>

‡ <https://www.wikidata.org/>

§ <http://www.geonames.org/>

Figure 1 (left panel) shows an example of the geospacial feature in this news recommender system. To customize geospacial information, the user draws a circle in the relevant area on a map. The system then provides the most relevant news within this area.

The central panel in Figure 1 shows an example of how news articles are presented. In this image, a headline from a news article in The Guardian is shown, followed by another news story from the BBC. The three buttons on the bottom of the screen allow users to define their locational and personal preferences and time setting to personalize their news story experience. By clicking on a news story, the user obtains the whole content of the news article and a list of the most salient named entities obtained from Wikidata. As can be seen from the right-hand graph in Figure 1 the named entities in this news article are ‘Syria’, ‘Theresa May’, ‘Islamic State of Iraq and the Levant’ and ‘Sky News’. By hovering over these items the user is presented with their textual Wikidata description, such as “British politician” about ‘Theresa May’, helping the reader be better informed.

## 2.2. System Overview

This news recommender system is built as a pipeline of operations harvesting and transforming Rich Site Summary (RSS) entries and raw text data into a semantic and searchable representation. The pipeline and its operations are implemented with Apache Storm. This distributed computing framework enables scalability and ability to handle large amounts of news items from a magnitude of publishers continuously.

As shown in Figure 2, the news processing pipeline consists of five steps. Firstly, the system creates an input stream of news articles by continuously monitoring a large set of RSS feeds. Whenever a news article is

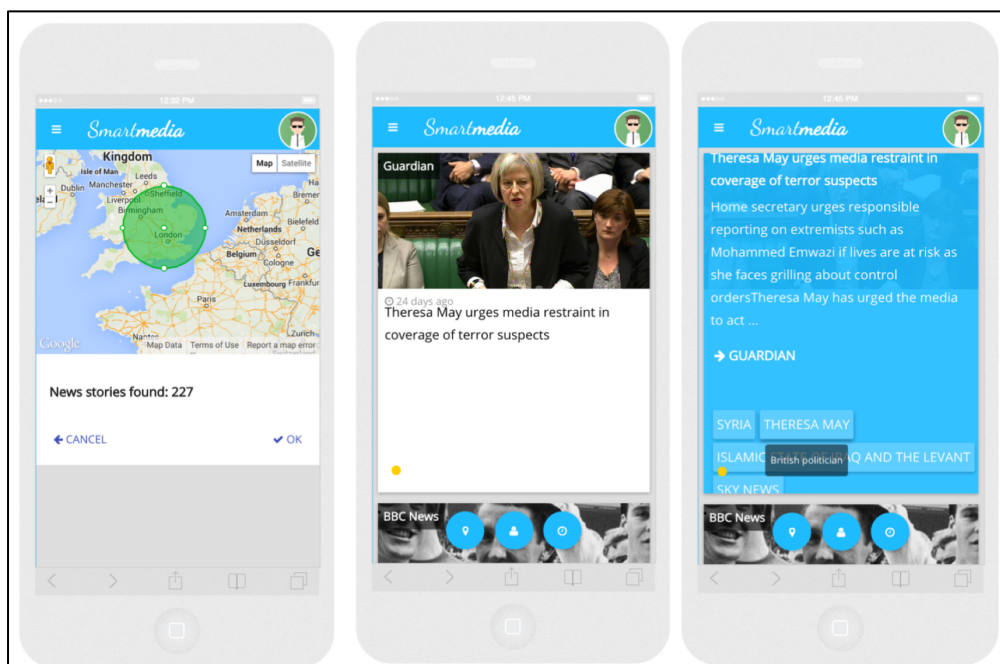
fetches, the HTML sources of that news item are obtained. These sources are then parsed and cleaned to extract the title and main body of text. Secondly, natural language processing operations such as language identification, sentence detection and part-of-speech tagging are applied in order to extract relevant named entities from the textual content. Then, supervised models are used to map named entities to their corresponding entries in Wikidata and Geonames knowledge bases. These models combine textual similarities, graph relations, entity frequency and co-occurrence statistics to disambiguate multiple referent candidates. First Story Detection (FSD) is applied then to cluster news articles describing the same news story. Lastly, this semantic representation is indexed and made searchable. A screencast video demonstrating the prototype and its user interface is available at: <http://vimeo.com/121835936>

## 3. CONCLUSION

In this paper we have presented the SmartMedia news recommender system for mobile devices. This recommender system makes use of time, location and context to provide personalized news experiences to mobile readers. Future work will include further improvement of entity linking and disambiguation, together with evaluation of user needs.

## References

1. Ingvaldsen, Jon Espen, Gulla, Jon Atle, and Özlem Özgöbek. "Context-Aware User-Driven News Recommendation." INRA@RecSys 2015: 33-36.
2. Ingvaldsen, Jon Espen, and Jon Atle Gulla. "Taming news streams with linked data." 2015 IEEE 9th International Conference on Research Challenges in Information Science (RCIS). IEEE, 2015.
3. Ingvaldsen, Jon Espen, Gulla, Jon Atle, and Özlem Özgöbek. "User Controlled News Recommendations." IntRS@RecSys, 2015: 45-48.



**Figure 1: User interface of SmartMedia prototype. Left: The map query interface. Center: Presentation of news stories. Right: Presentation of news details.**

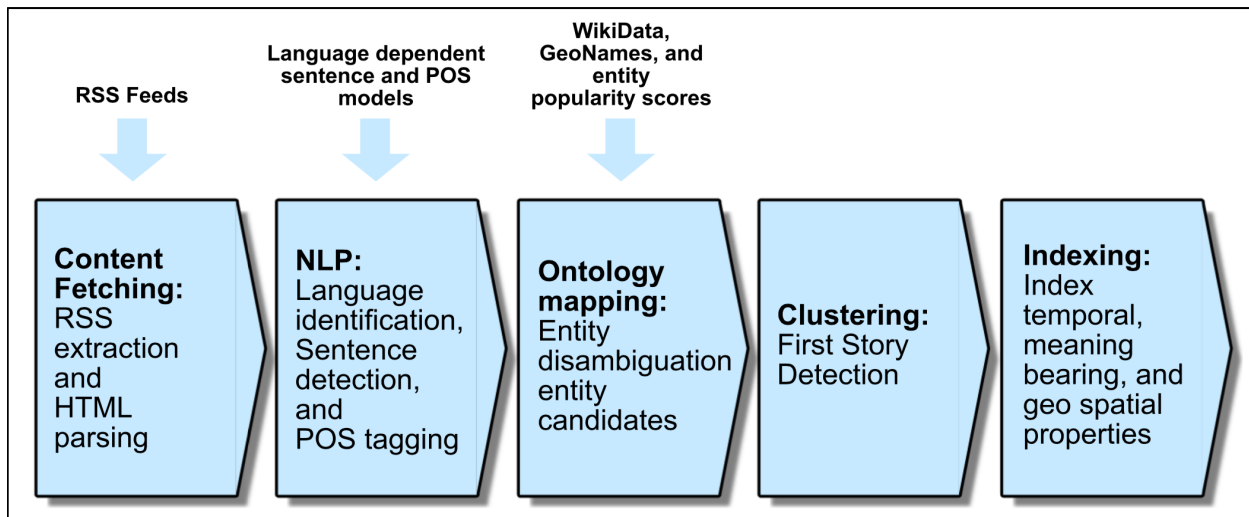


Figure 2: News processing pipeline