

Allowing Exploratory Search from Podcasts: the Case of Secklow Sounds Radio

Ilaria Tiddi, Emanuele Bastianelli, Martino Mensio, and Enrico Motta

Knowledge Media Institute, The Open University, United Kingdom
`name.surname@open.ac.uk`

Abstract. We present here the Secklow Sounds Radio App that was developed as one of the demonstrators in the context of the MK:Smart project. Secklow Sounds is a community radio based in the city of Milton Keynes (MK), providing digital recordings of their broadcasts online, where local issues are often discussed. We developed a mobile-friendly web app that, besides offering live streaming of the radio, allows users to perform an entity-based search of past broadcasts enhanced with information (e.g. areas, museums, topics, etc) provided by the city’s centralised repository. The paper presents first the data processing workflow which integrates a number of existing solutions, such as the Google Speech-To-Text API, Neural Networks, DBpedia Spotlight, and the DiscOU search engine, and finally shows how results were integrated into a web and mobile application providing an exploratory search service for radio podcasts.

Keywords: Audio processing · Exploratory search · Data integration

1 Introduction

The paper presents the Secklow Sounds Radio App and how this was developed as a demonstrator for the MK:Data Hub¹, the data integration and sharing platform of the MK:Smart project². MK:Smart is a large collaborative initiative started in 2014 with the goal of developing innovative technology solutions to support the economic growth of the city of Milton Keynes (MK). The project promotes the idea that the creation of a common infrastructure to efficiently manage, integrate, and re-deliver information from local data sources (energy/water consumption, transport, satellites, social/economic sources, social media etc.) facilitates the deployment of data-intensive applications, enabling intelligent data processing mechanisms for citizens and service providers [4].

To demonstrate the exploitability of the MK:Data Hub, the Secklow Sounds Radio App was built as a use-case in collaboration with the Secklow Sounds Radio³. As a MK-based community radio, Secklow Sounds provides digital recordings of their broadcasts, where local issues are also discussed. Besides offering an

¹ <http://www.datahub.mksmart.org>

² <http://www.mksmart.org>

³ <http://www.secklow1055.org/>

app with live streaming of their episodes, the radio aims to provide its audience with advanced browsing facilities for their published data.

As a result, the Secklow Sounds Radio App was developed to allow users to perform an entity-based search and discovery of the broadcasts, whose contents (local areas, museums, topics, etc.) were semantically enhanced with information centralised in the MK:Data Hub. In order to achieve this, we built a data processing workflow to transform audio podcasts into explorable data, through the integration of a number of off-the-shelf solutions addressing the different tasks to achieve – namely, speech understanding, named entity recognition, semantic indexing and data augmentation. The final output was then integrated into a mobile-friendly application publicly available to end-users and radio listeners. Here we present first the developed pipeline, with a description of its components in details, and then show how the processed data were deployed in the Radio App.

2 Extracting Data from Podcasts

The Secklow Sounds Radio App was implemented using the process depicted in Figure 1, taking as input the .mp3 of the radio broadcasts, and returning its text annotated with entities from DBpedia and the MK:Data Hub. The workflow is divided in: (i) Audio Processing, concerning the set of tasks to obtain texts from the audio files; (ii) Text Annotation, i.e. the tasks for augmenting the episode texts with external data and (iii) Data Aggregation, where results are aggregated and wrapped into the Radio App.

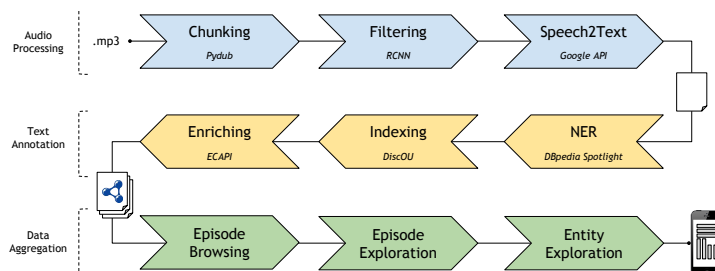


Fig. 1: Workflow for data processing, from .mp3 files to the Radio App.

Audio Processing. Audio chunking, filtering and speech-to-text are performed to obtain text from an initial .mp3 file. As radio episodes consist in 60-180 mins ca. of both music and talks, we chunk audios based on silences using the Python Pydub⁴ library, in order to reduce them in length, and to be able to recognise the spoken chunks (to transcribe into texts) from the music ones. Filtering is then performed using a Recurrent-Convolutional Neural Network (RCNN) inspired by [1]. The original RCNN was designed for Music Genre Classification (50 classes), but was adapted here to produce a binary classification of the audio chunks

⁴ <https://github.com/jiaaro/pydub>

(speech or music), through tuning and training the model over a dataset of 1910 audio files tagged by three annotators. Finally, we use the Google Speech-to-Text Cloud API⁵ to obtain the transcriptions. The API provides a number of facilities, including pre-trained models for several languages, and the possibility of feeding the model with a context-specific vocabulary to improve the recognition. As Secklow Sounds is a UK-based radio, we used the *en-GB* language model, and fed the model with MK-specific entities, e.g. its wards (*Wolverton*, *Bletchley*), local libraries and museums (*Bletchley Park*, *Woburn Sands library*) and presenters' names.

Text Annotation. Text annotation is performed over the transcriptions of the podcasts, and includes the tasks of named entity recognition, semantic indexing, and data enrichment. DBpedia Spotlight⁶ is used first to obtain the list of named entities from the episodes' transcriptions. The resulting semantic descriptions are then indexed using the DiscOU Semantic Indexer⁷ following the idea of [3]. Based on the Lucene search engine, the Indexer uses the semantic annotations in order to allow resource search by semantic, rather than textual, similarity. Once the index is built, resources are annotated with the relevant DBpedia entities and their occurrence score. If a mapping between DBpedia and the MK:Data Hub exists (see, for instance, the ward of *Wolverton*⁸), an additional annotation is also provided. This allows entities to be explored through the MK:Data Hub *Entity Centric API* (ECAPI [2]), which aggregates relevant data (wards, estates, buildings, bus stops etc.) from multiple data sources.

Data Aggregation. The last step is the aggregation of the annotated audio contents and the implementation of the Radio App. The application, available online⁹, is developed using the Angular Mobile UI framework¹⁰, to allow usage from laptops and mobiles independently from the operating system. Three activities can be performed with the app: (i) browsing episodes related to a specific entity; (ii) exploring the content of a selected episode; and (iii) exploring a specific entity discussed during an episode. In Figure 2 for example, the user is searching through the search box for episodes about *Wolverton*, promptly returned by the indexer (Figure 2a)¹¹. The selected episode, whose podcast can be listened, can be explored through its enriched content (Figure 2b), where the annotated entities obtained during the enriching task are visualised as a word cloud with different sizes and shades depending on their frequency – for example, the episode *Lifestyle MK* of April 2016 discussed food-related topics. Entities

⁵ <https://cloud.google.com/speech-to-text/>

⁶ <https://www.dbpedia-spotlight.org/>

⁷ <https://github.com/the-open-university/discou-indexer>

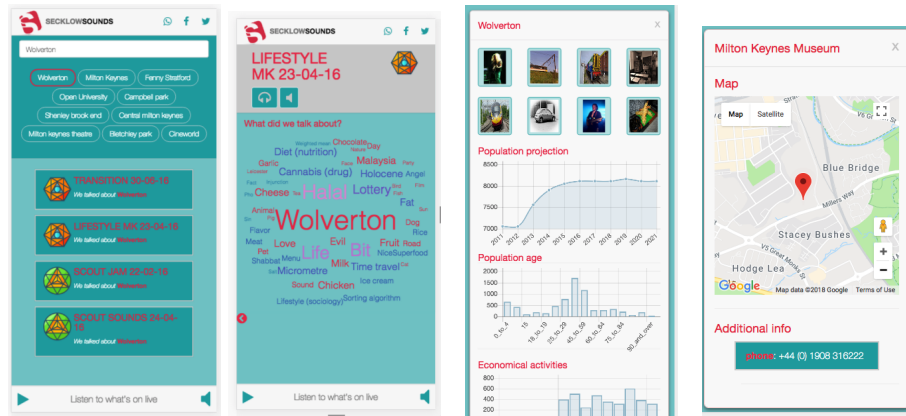
⁸ DBpedia entity: <http://dbpedia.org/resource/Wolverton>, MK:Data Hub entity <https://data.mksmart.org/entity/ward/wolverton>

⁹ <https://data.mksmart.org/apps/secklow-sounds-app/>

¹⁰ <http://mobileangularui.com/>

¹¹ Note that the set of entities proposed in the facet are only suggestions, while users can freely search for any entity (also not related to MK).

in pink are the ones that can be explored through ECAPI, e.g. in Figure 2c the user is visualising crowdsourced pictures and socio-demographic information (population age and projection, economical activities) about Wolverton. Similarly, Figure 2d shows that users can also obtain practical information about local activities (the Milton Keynes Museum), such as location, phone number, opening times, etc.



(a) Browsing episodes facet. (b) Exploring a single episode. (c) Exploring entities (1). (d) Exploring entities (2).

Fig. 2: Screenshots of the Secklow Sounds Radio App.

3 Demonstration

During the demonstration we will present the Secklow Sounds Radio App with two main goals, namely to show the benefits of a data integration facility (such as the MK:Data Hub) to deploy lightweight applications, and how off-the-shelf solutions can be integrated together into a simple but effective workflow to offer exploratory search over radio broadcasts.

References

1. Choi, K., Fazekas, G., Sandler, M., Cho, K.: Convolutional recurrent neural networks for music classification. In: Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. pp. 2392–2396. IEEE (2017)
2. Daga, E., d’Aquin, M., Adamou, A., Motta, E.: Addressing exploitability of smart city data. In: Smart Cities Conference (ISC2), 2016 IEEE International. pp. 1–6. IEEE (2016)
3. d’Aquin, M., Allocca, C., Collins, T.: Discou: A flexible discovery engine for open educational resources using semantic indexing and relationship summaries. In: Proceedings of the 2012th International Conference on Posters & Demonstrations Track-Volume 914. pp. 13–16. Citeseer (2012)
4. d’Aquin, M., Davies, J., Motta, E.: Smart cities’ data: Challenges and opportunities for semantic technologies. *IEEE Internet Computing* **19**(6), 66–70 (2015)