A cover art archive

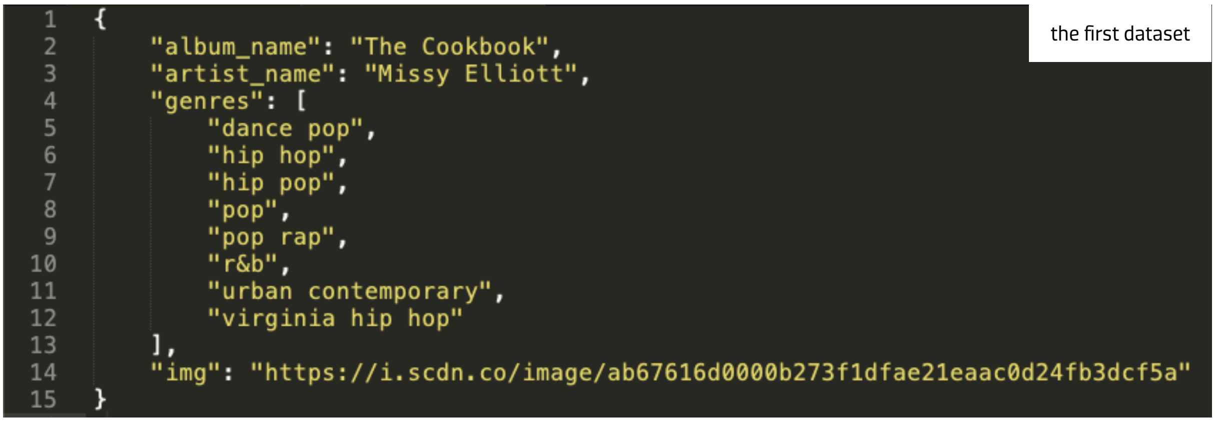
The idea behind our project is to provide a navigable collection of music albums where the visual aspect is the core element of the experience. We aim to create a link between the chromatic and the sound moods characterizing a record. The archive allows to discover music through images, and In particular by dominant colours: each album is described not only by the standard metadata related to this kind of media but also by information about its predominant colours.

Acquiring and organizing data

To fill the archive, we exploited the Spotify API, that provides a great number of data free to use for non-commercial purposes. The advantage to creating a dataset from such a large repository is the variety of different kinds of records that can be collected. Using a python script that simulates a browser, we queried the API to get the list of the tracks present in some playlists and extracted their album uris. From this new list, we could easily access new, particular metadata, such as:

- the artist
- the genres
- he publication date
- the url of the cover art image
- The uri of the album

From this information, we were able to build a JSON dataset ready to be analyzed by other scripts.



```
{
    "album_name": "The Cookbook",
    "artist_name": "Missy Elliott",
    "genres": [
        "dance pop",
        "hip hop",
        "hip pop",
        "pop",
        "pop rap",
        "r&b",
        "urban contemporary",
        "virginia hip hop"
    ],
    "img": "https://i.scdn.co/image/ab67616d0000b273f1dfae21eaac0d24fb3dcf5a"
}
```

the first dataset

Adding knowledge to data

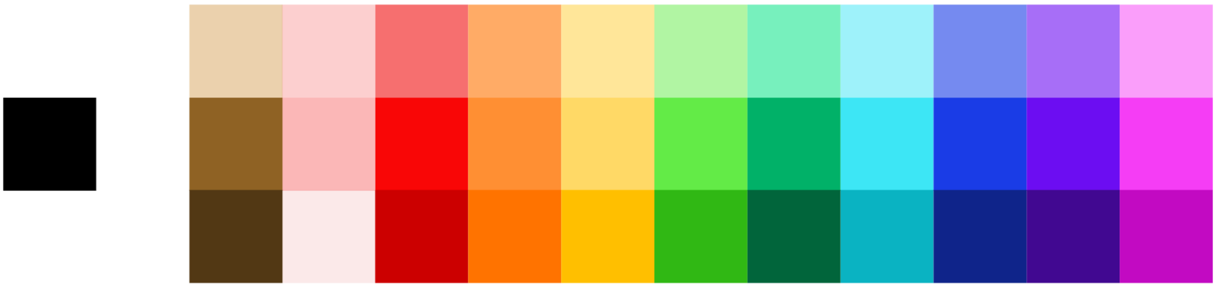
Chromatic data

Having a dataset doesn't mean owning knowledge. Therefore, we decided to expand our dataset with new information relative to the colour.

To do so, we used the clustering algorithm K-means to pick the most frequent RGB values from each album, again using the chromium module to simulate the browser and read the image data. The algorithm was set with a K value (the number of obtained clusters) of three. This demonstrated to be a semantically reasonable quantity and a good compromise between abstraction and details.

Since the number of albums was substantial, and by consequence the amount of different dominant colours, we assigned each record one or more labels that summarize them. Hence, we created an artificial palette, whose tones were found empirically to well abstract the graphical and photographic colours. The final labels are the following:

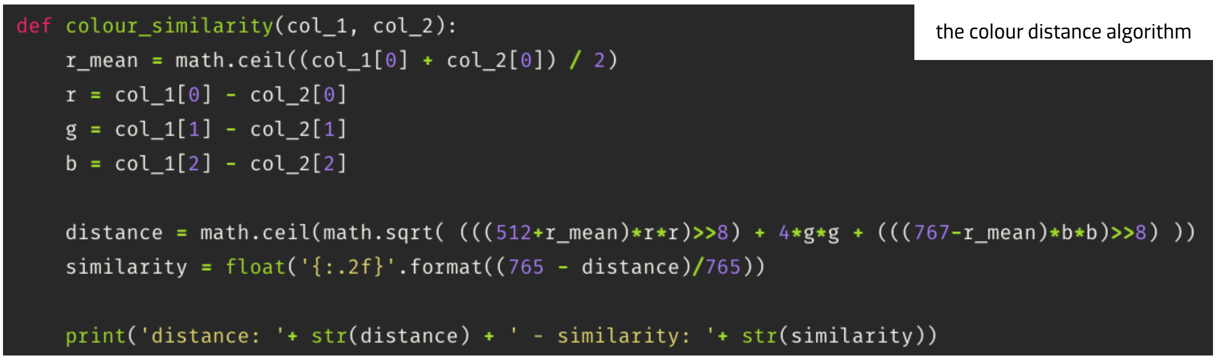
the colour labels



black	white	brown	pink	red	orange	yellow	light green	dark green	cyan	blue	violet	purple
r: 0	r: 255	r: 143	r: 251	r: 249	r: 255	r: 255	r: 99	r: 1	r: 61	r: 26	r: 108	r: 245
g: 0	g: 255	g: 98	g: 183	g: 6	g: 143	g: 217	g: 235	g: 177	g: 177	g: 60	g: 13	g: 61
b: 0	b: 255	b: 36	b: 183	b: 6	b: 51	b: 102	b: 71	b: 104	b: 104	b: 230	b: 242	b: 245

An additional label was created for the grey, by verifying if R, G, and B were all in the range [70, 185] (so not too black and not too white) and their difference was at most of 15 unities.

We then computed the difference from each RGB dominant colour to each label, and if it was higher than 80%, we assigned that class to the record. This distance was calculated through a formula ideated by Compuphase, that emulates the visual perception of tones.



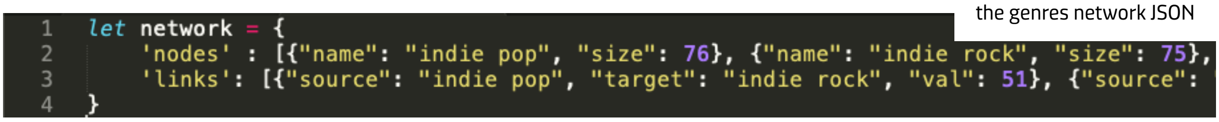
```
def colour_similarity(col_1, col_2):
    r_mean = math.ceil((col_1[0] + col_2[0]) / 2)
    r = col_1[0] - col_2[0]
    g = col_1[1] - col_2[1]
    b = col_1[2] - col_2[2]

    distance = math.ceil(math.sqrt( (((512+r_mean)*r*r)>>8) + 4*g*g + (((767-r_mean)*b*b)>>8) ))
    similarity = float('{:.2f}'.format((765 - distance)/765))

    print('distance: '+ str(distance) + ' - similarity: '+ str(similarity))
```

the colour distance algorithm

Genres' relations

An additional dataset was created to capture the musical genres' relations. We noted that some genres recurred frequently together, and we wanted the user to know it. A new script was then created to find a representation of the network.

To obtain a clean and outlier-free representation, we considered only the most recurring phenomena (over a certain threshold).



```
let network = {
    'nodes' : [{"name": "indie pop", "size": 76}, {"name": "indie rock", "size": 75},
    'links': [{"source": "indie pop", "target": "indie rock", "val": 51}, {"source":
}
```

the genres network JSON

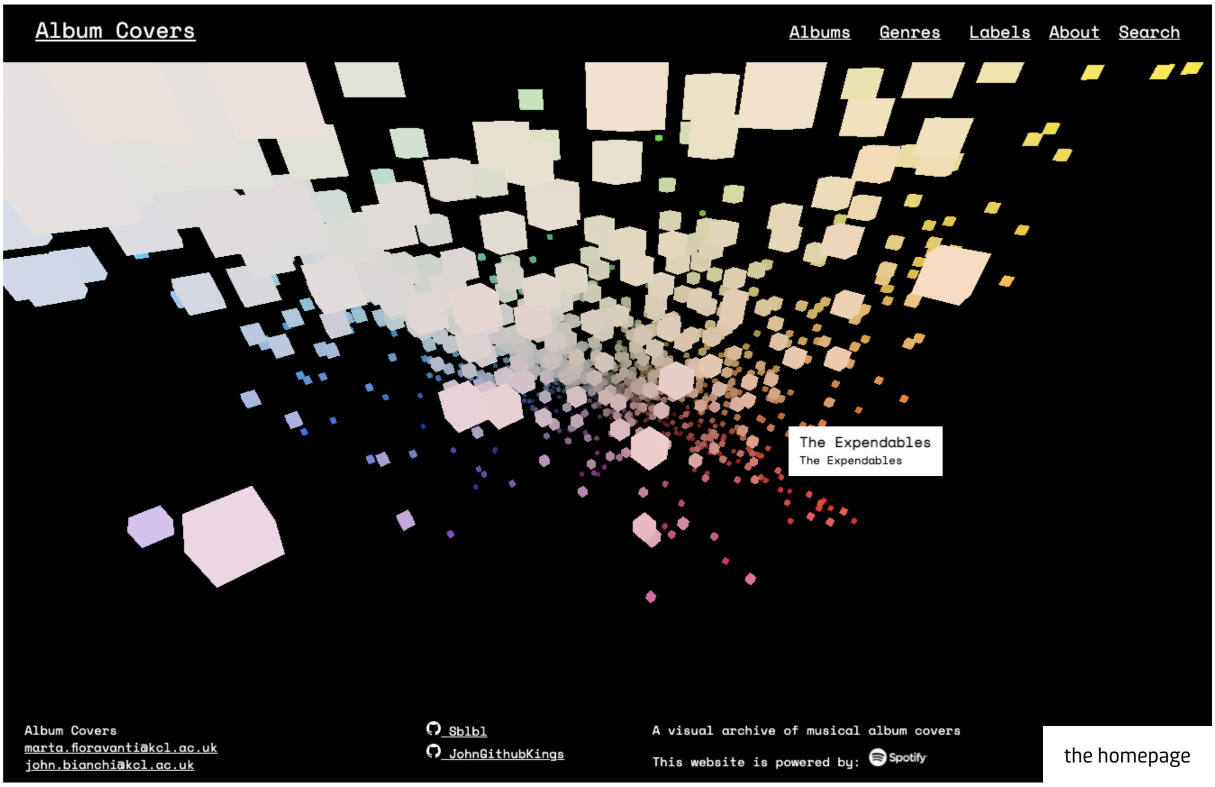
Generating and organizing contents

To fill the archive was thus necessary to generate the .md files via python. We did it for the collections of albums, genres and labels. Some JSON files were however necessary to feed the javascript functions for certain pages.

To display the collections in an enjoyable way, we used the Jekyll Pagination plugin, that enabled to display few elements at a time, avoiding long scrollings. The elements are ordered by name, in ascending order, to maintain an immediately comprehensible organization.

Homepage

The homepage is the means through which a website is presented: it should be endearing to encourage the visitor to explore its contents. That is the reason for our choice: to display a 3D environment where each album is represented as a cube coloured with its first dominant colour. This has the purpose to generate marvel, more than to display knowledge, and it establishes a gentle introduction to the site content. In fact, by clicking on a cube, the user is lead to the album's page, where they can read more information.



Album Covers
marta.fioravanti@kcl.ac.uk
john.bianchi@kcl.ac.uk

A visual archive of musical album covers

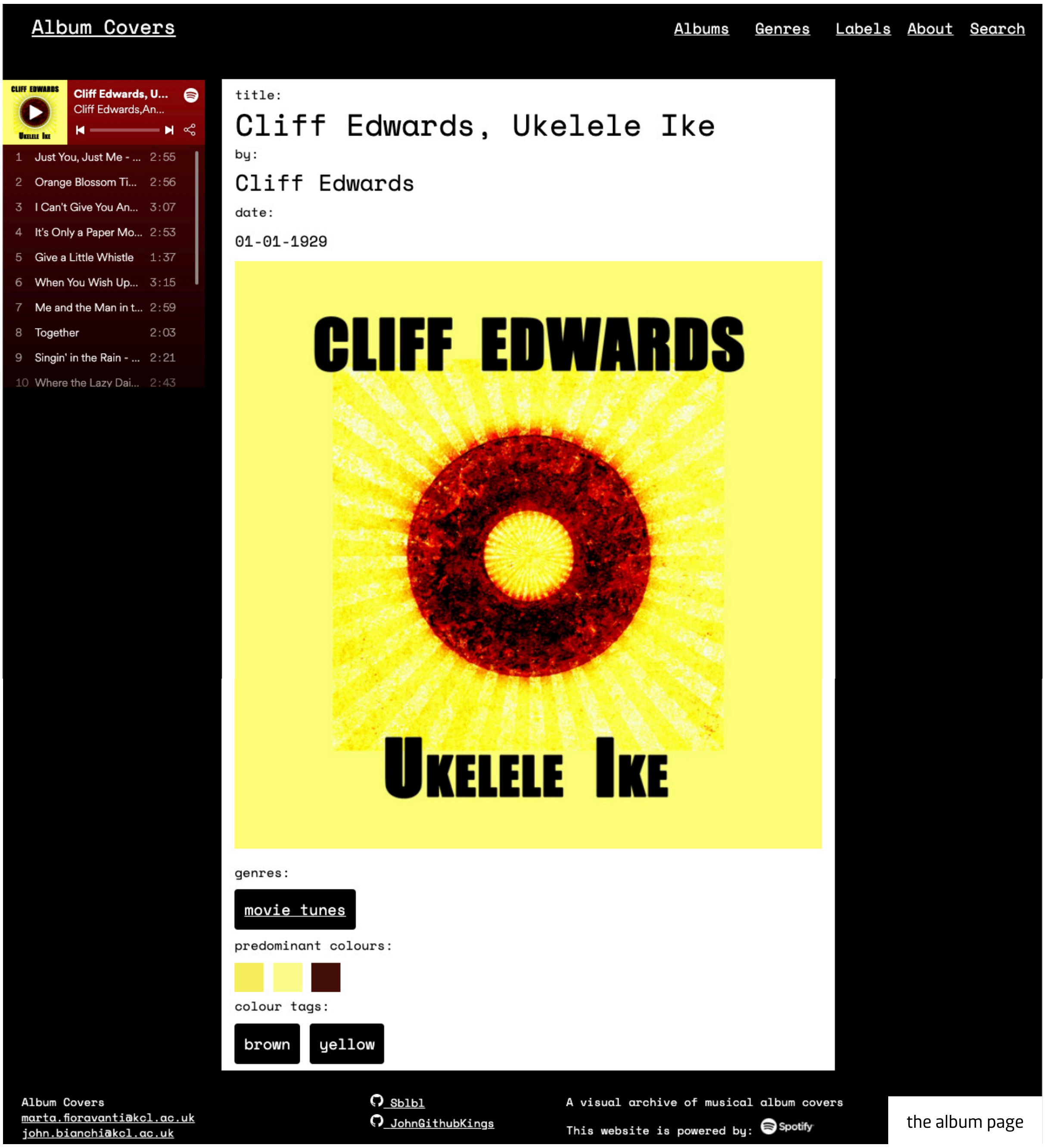
This website is powered by:

the homepage

Albums

The album pages contain all the collected metadata. These were encoded via the *itemscope* property *MusicRecording* and setting the correct *itemprop* in the dedicated tags. We also embedded the Spotify album to encourages the visitor to discover which sound a certain colour is associated with.

The details are linked together to provide complete navigation of the contents: for example, the genres' tabs bring to the specific section, where similar records can be found.



Album Covers

Albums Genres Labels About Search

Cliff Edwards, U...
Cliff Edwards,An...

1 Just You, Just Me - ... 2:55
2 Orange Blossom Ti... 2:56
3 I Can't Give You An... 3:07
4 It's Only a Paper Mo... 2:53
5 Give a Little Whistle 1:37
6 When You Wish Up... 3:15
7 Me and the Man in t... 2:59
8 Together 2:03
9 Singin' in the Rain - ... 2:21
10 Where the Lazy Dai... 2:43

title:

Cliff Edwards, Ukelele Ike

by:

Cliff Edwards

date:

01-01-1929

genres:

movie tunes

predominant colours:

colour tags:

brown yellow

Album Covers
marta.fioravanti@kcl.ac.uk
john.bianchi@kcl.ac.uk

Sblbl
JohnGithubKings

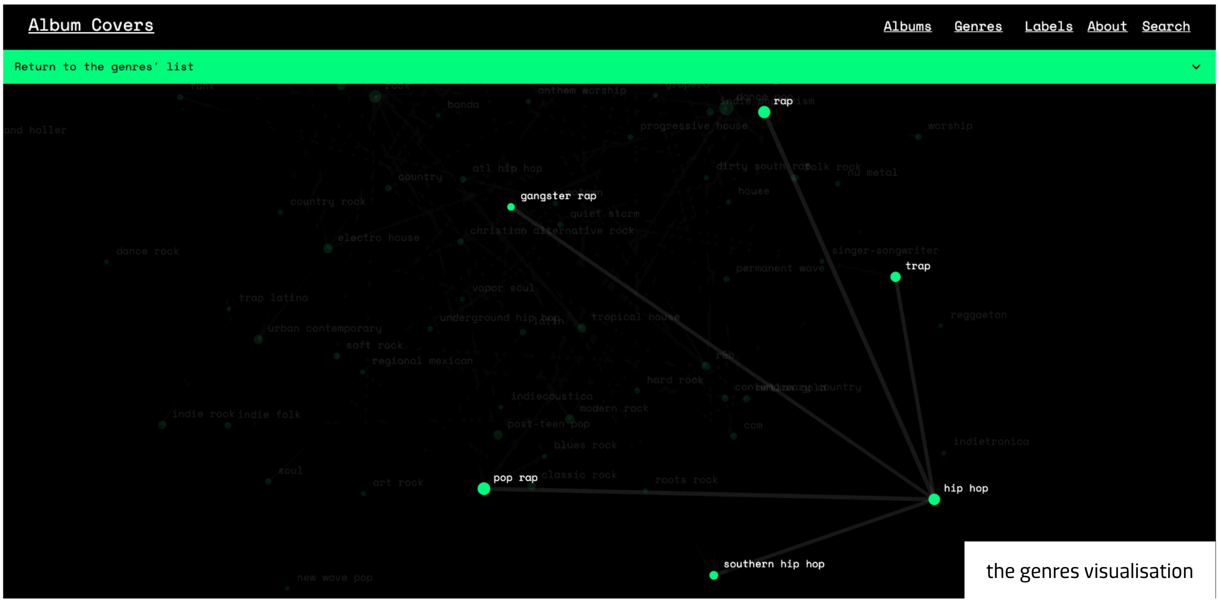
A visual archive of musical album covers

This website is powered by: Spotify

the album page

Genres

This collection lists all the musical genres of the archive. By clicking on an entry, it is possible to see all the albums labelled with it. To provide more enjoyable and informative navigation, we built a visualisation showing the relations between genres. This tool is interactive and brings the visitor to the genres' page when a node is clicked.

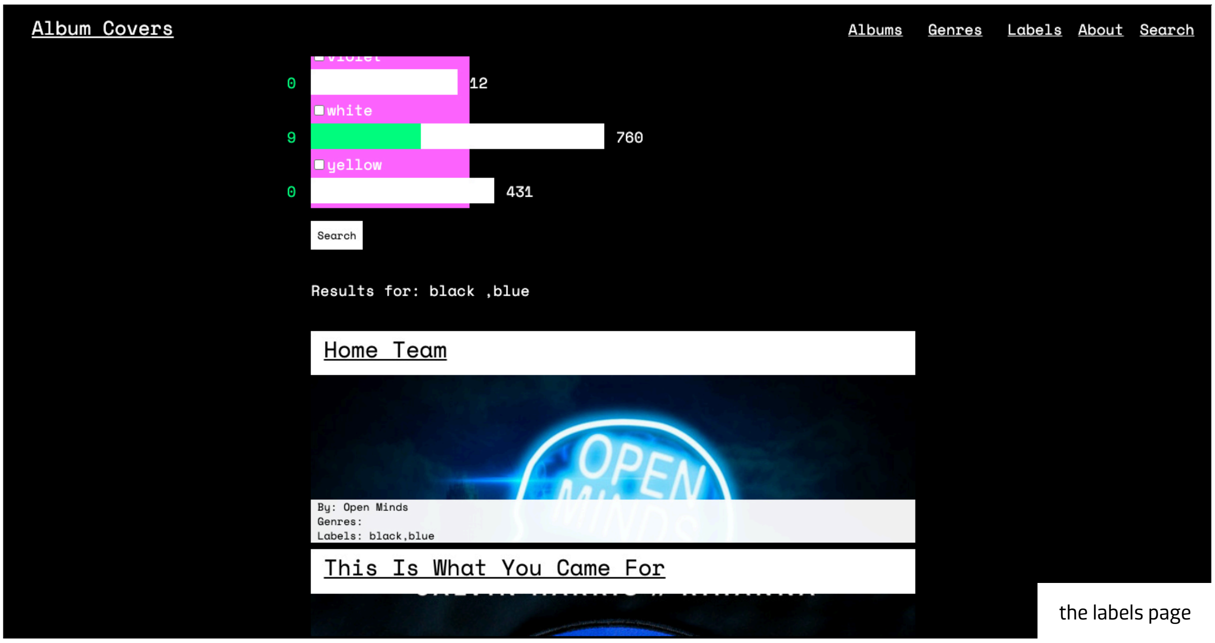

Album Covers

Return to the genres' list

the genres visualisation

Labels

For this collection we designed a different interaction: instead of only listing the entries, we displayed them as a bar chart, showing the number of records having that label and more specifically how many albums only classified with it. The visitor can select some through some checkboxes and see the albums that contain only them. When this is done, an additional bar appears showing the amount of selected data. This choice allowed us to create a more informative system that is more flexible than a simple list and that let to perform detailed searches.


Album Covers

Results for: black ,blue

Home Team

By: Open Minds
Genres:
Labels: black,blue

This Is What You Came For

the labels page

Conclusion

We think we succeeded in our main goal: designing an enjoyable and informative experience. We are convinced that a good archive must encourage the exploration to become a knowledge source, and should present the knowledge in an accessible and interactive way.