

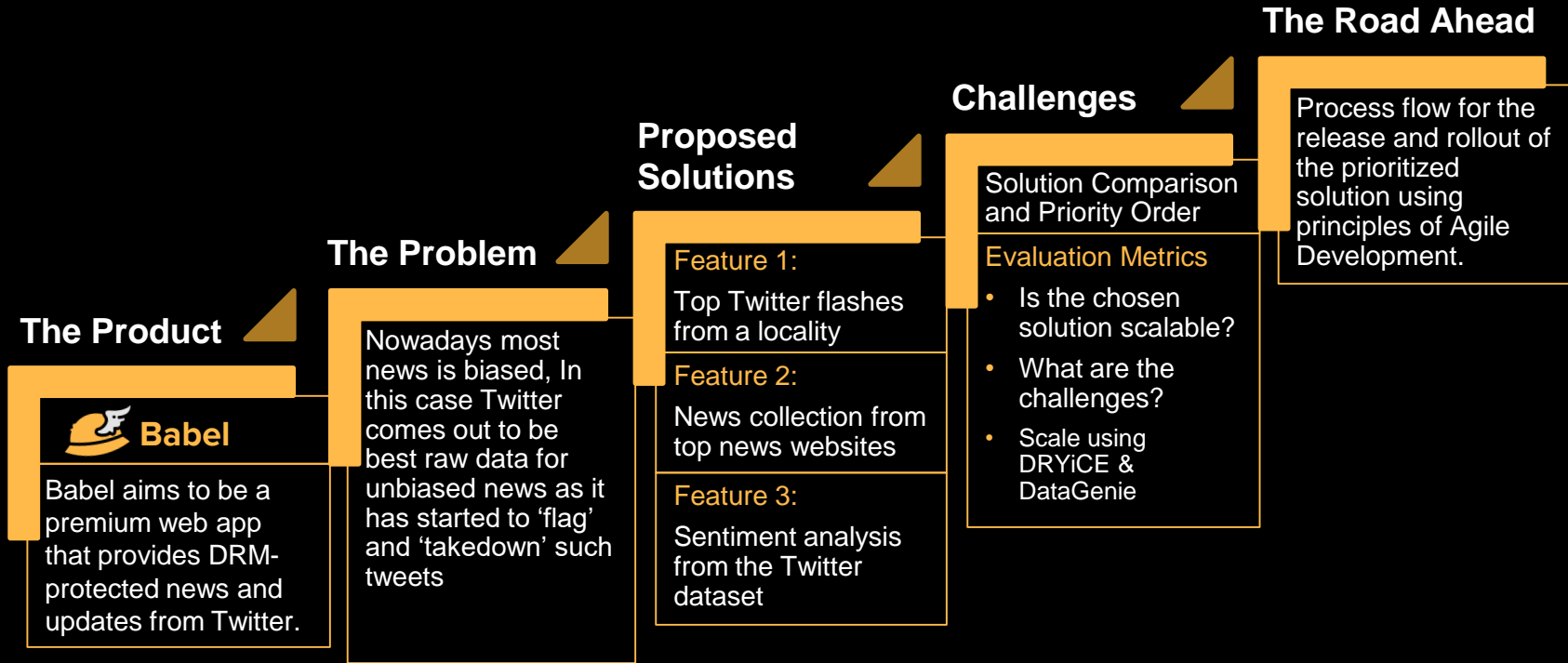


Babel:

Local News for Everyone

A closer look at the innovative methods of procuring and delivering news

IMPROVING THE LOCAL NEWS EXPERIENCE!



PROBLEM & SOLUTION

PROBLEM



The problems at hand are fake news and biased news, the internet is full of these which tampers the raw data if not taken care while extracting. Also these kind of news are practically useless for the web app. On the other hand it might render the useful dataset useless.

"Trust but verify"

SOLUTION



The solution to the mentioned problems is to choose a source free from such biased and fake news. Platforms that flag and takedown tweets that are associated with political propaganda and/or are reported and/or identified as fake can be good source of raw data.

- Twitter has started to flag and takedown tweets that are identified to be carrying political propaganda, hate speech and/or fake news, which makes it an ideal source for extracting raw data. After applying a few heuristics we can obtain useful data from the pool that can be posted as an update/summary on the web application.
- Prominent websites which are known for imparting unbiased and true news at the earliest can also act as a good source data, it can be used with minimal pre-processing and would surely be free of radical views.
- By performing sentiment analysis we can flag a particular tweet to be neutral, positive and negative which can help with filtering of radical tweets for young audience.
- By sorting these tweets location wise we can cater people seeking local news/updates which can also redirect them to the source content through just a click.

FEATURES

WHAT'S ON OFFER?



Procuring tweets and pre-processing using LDA:

Twitter API is used to get tweets based on a location, the tweet data is cleaned off from stop words and LDA (a topic modelling model) is used to create clusters of tweets that are related to each other



Assigning score to tweets:

An 'adjusted score(AS)' is calculated for every tweet in the data set. The adjusted score takes data such as followers, following, mentions, account being verified or not, tweet being a reply, how old is the account etc.

$$\text{Popularity Score}(PS) = (AS / \text{Mean}(AS))^{0.5}$$

The popularity score is assigned to every tweet which is used to rank the tweets, top 10 tweets from each topic(cluster) are chosen to be displayed on the web app.



Performing sentiment analysis:

Sentiment analysis is performed on all tweets fetched from the same location to mark the tweet as positive, negative or neutral.



Fetching news from prominent news agencies:

Beautiful soup(BS4) is used to fetch trending national news briefings from prominent news agencies which could then be displayed on the web app.



RESTful API based web application:

The backend of the web application is based on Python+Flask and the frontend of the application is based on React.js. The communication between the two happens through RESTful APIs which automatically makes it an API service that could be utilized by any web application

Technology Stack

Frontend:

HTML, CSS (Material Design) as these are industry standards and used in various big corporations. Javascript(ReactJs) has a very good online community and virtual DOM which will make our web app quick.

Backend:

Python(Flask) for creating APIs to communicate between frontend and backend as it has very high reliability and ample documentation.

Machine Learning:

Pandas, Sci-kit, TextBlob and Gensim as all of them are backed by industry.

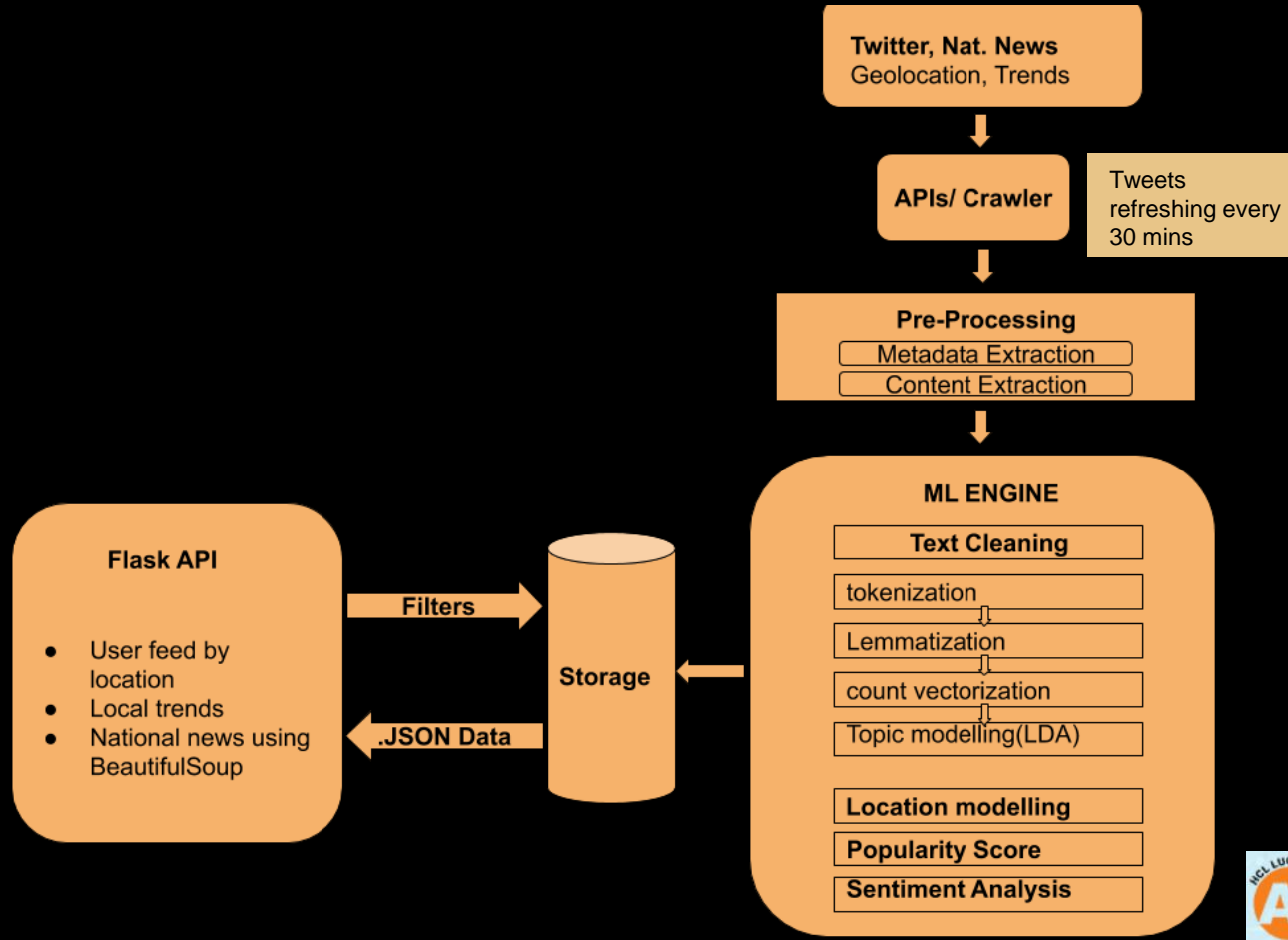
APIs:

Twitter feed, as it has very fast API and a huge user count.

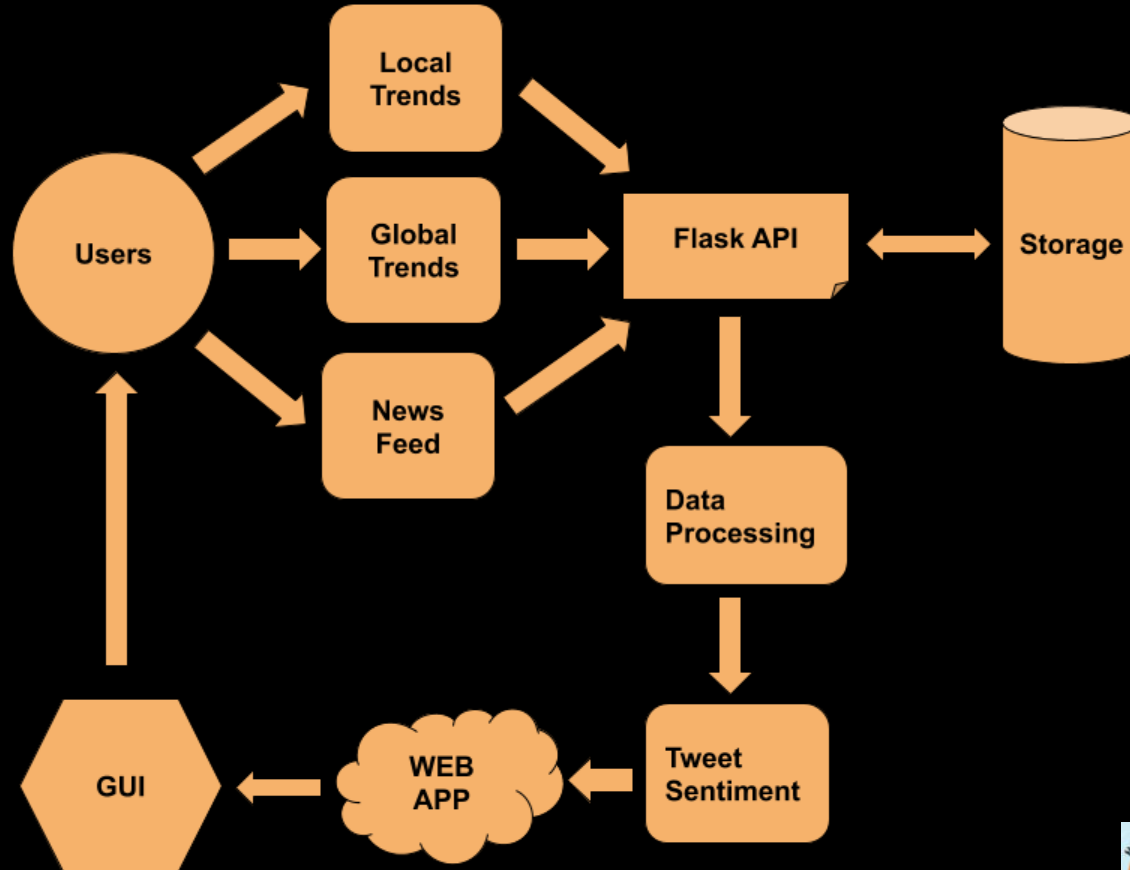
Other Tools:

Web-scraping(BeautifulSoup4) to get data from other sources of news.
Cron for automation of our scripts at particular interval(30 mins).

Backend Flow Diagram



Frontend Flow Diagram



Challenges

- ***Is the solution i.e. Babel scalable?***

Babel can handle big data. Right now we are either using API services which are either free or freemium, with suitable funding, premium alternatives or/and internal 24/7 online servers we can not only host big data on database but also batch parse raw data from the web without any limitation. The solution i.e. Babel can handle big data and therefore is scalable. In the solution RESTful APIs which can be utilised to process big data without any changes

- ***What are the other challenges?***

- Right now we are using Twitter API to extract data from it for free, It has a limit over number of request and number of tweets per batch because of which it is not possible to get bulk data, However, with a premium, things are different, we can procure bulk data with ease. With big data at hand the heuristics/models could be trained better further yielding better results.
- The NoSQL Cloud database that we have considered is a freemium service which exercises limit over the basic CRUD operation. However, the paid service is fit for handling data in bulk in batch CRUD operations.
- React.js is single threaded out of the box however, with the help of multithreading engines we can distribute the load on core and speed up the execution thus reducing the wait time. Right now the React app is single threaded with brings up occasional loading screens but with multithreading this could be reduced.

SCALING TO INDUSTRY LEVEL AND USING BIG DATA

- *How can the solution utilize Drylce and DataGenie?*

DataGenie:-

- DataGenie services for generating synthetic data can be utilised to generate large datasets of tweets for training models.
- The developer community can leverage from our work as the tweets data can be regularly stored in DataGenie repository. The data would be stored in json format as binary dump files. Binary dump files storage have less data storage and transaction costs as compared to csv files in databases. Also it doesn't require additional response units(RUs) to process data.

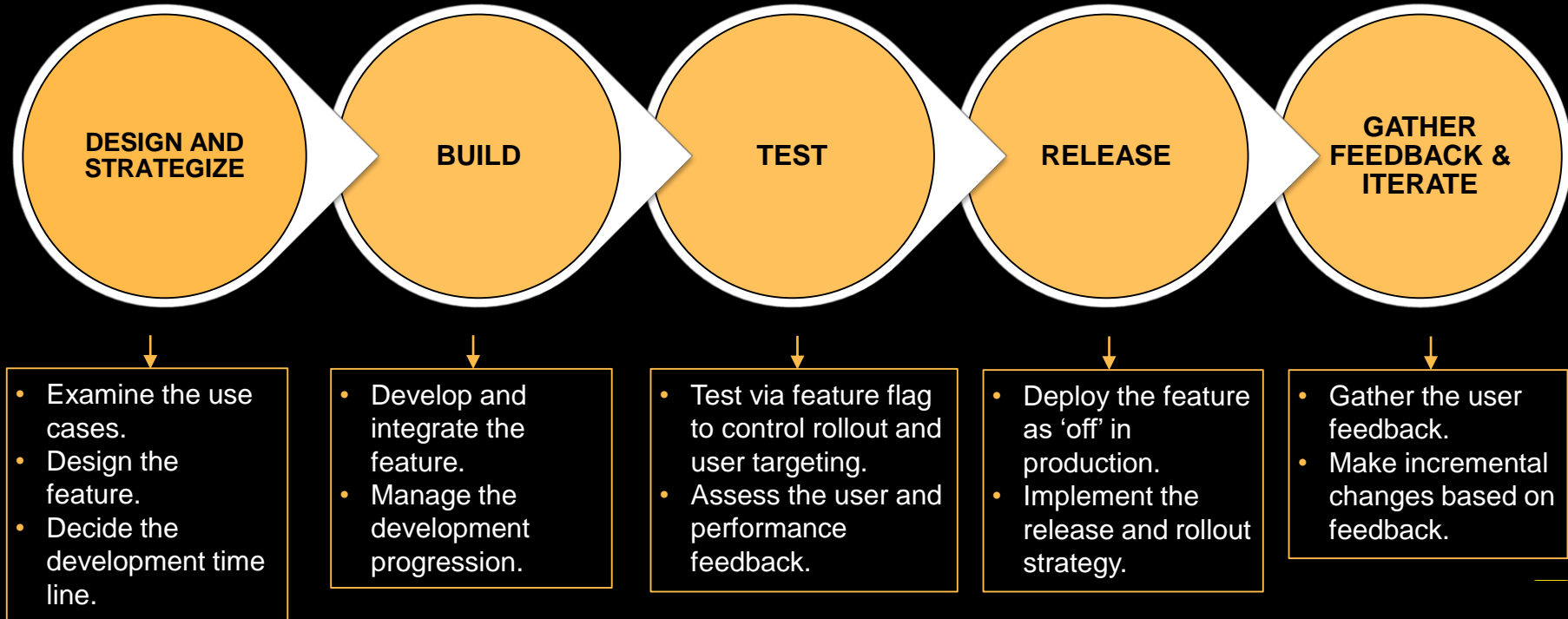
DRYiCE:-

- It may be used to host our product as SaaS. Being an end-to-end solution it may be used as proprietary services.
- Our project being modular, the developer community can leverage from it using API calls.
- The application can use DRYiCE iAutomate runbook automation product to periodically train our NLP models.

THE ROAD AHEAD



Objective: To rollout the solution i.e. **Babel** as a stable web app incorporating the principles of Agile Development.



Web App Screenshots



Babel

Local news for everyone

About the App..

Babel aims to be a premium platform that provides DRM-protected news and updates from Twitter and other prominent news agencies. Babel is a news reading app that aggregates published news from a variety of sources, like a better-formatted version of RSS. It presents stories that are most relevant to your personal interests while keeping you up to date on the rest of the world too.



Over Here

CLICK

Pilgrimage preparations reviewed

Erumely will be ready to welcome Ayyappa devotees by Nov. 15, says Minister

MORE

Drive to rid Vembanad Lake of plastics

Students engaged in the initiative will be rewarded with a houseboat cruise

MORE

Two held on charge of black marketing darshan tickets

The black marketing of darshan tickets continues at the hill shrine despite stringent measures by TTD to contain the menace. The fact that the police

MORE

Set up expert committee to decide on relocation of HC: lawyers

Continuing their protest at the Collectorate on Saturday, lawyers demanded the government to set up an expert committee to explore the possibilities o

MORE

'No affiliation to any political party'

APNGOs' Association will strive to protect interests of employees, says its leader

MORE

JUST
IN

1 **19mins** Madhuri Vijay's debut novel 'The Far Field' wins 2019 JCB prize for

2 **1hr** Mali attacks: IS claims responsibility for raid that killed 50 soldiers

3 **2hrs** Trump says he wants Ukraine's President Zelenskiy to visit the White

4 **2hrs** Airbnb bans 'party houses' after Halloween shooting in California



MENU



HOME

NEWS

OPINION

BUSINESS

SPORT

CRICKET

CROSSWORD

ENTERTAINMENT

New!

DASHBOARD



SUBSCRIBE NOW

SIGN IN / SIGN UP

STATES ▾

ANDHRA PRADESH

KARNATAKA

KERALA

TAMIL NADU

TELANGANA

OTHER STATES



WATCH | NCRB

NEWS > STATES > KERALA

KERALA

Pilgrimage preparations reviewed



STAFF REPORTER

KOTTAYAM, NOVEMBER 03, 2019 07:49 IST

UPDATED: NOVEMBER 03, 2019 07:49 IST

SHARE ARTICLE



0

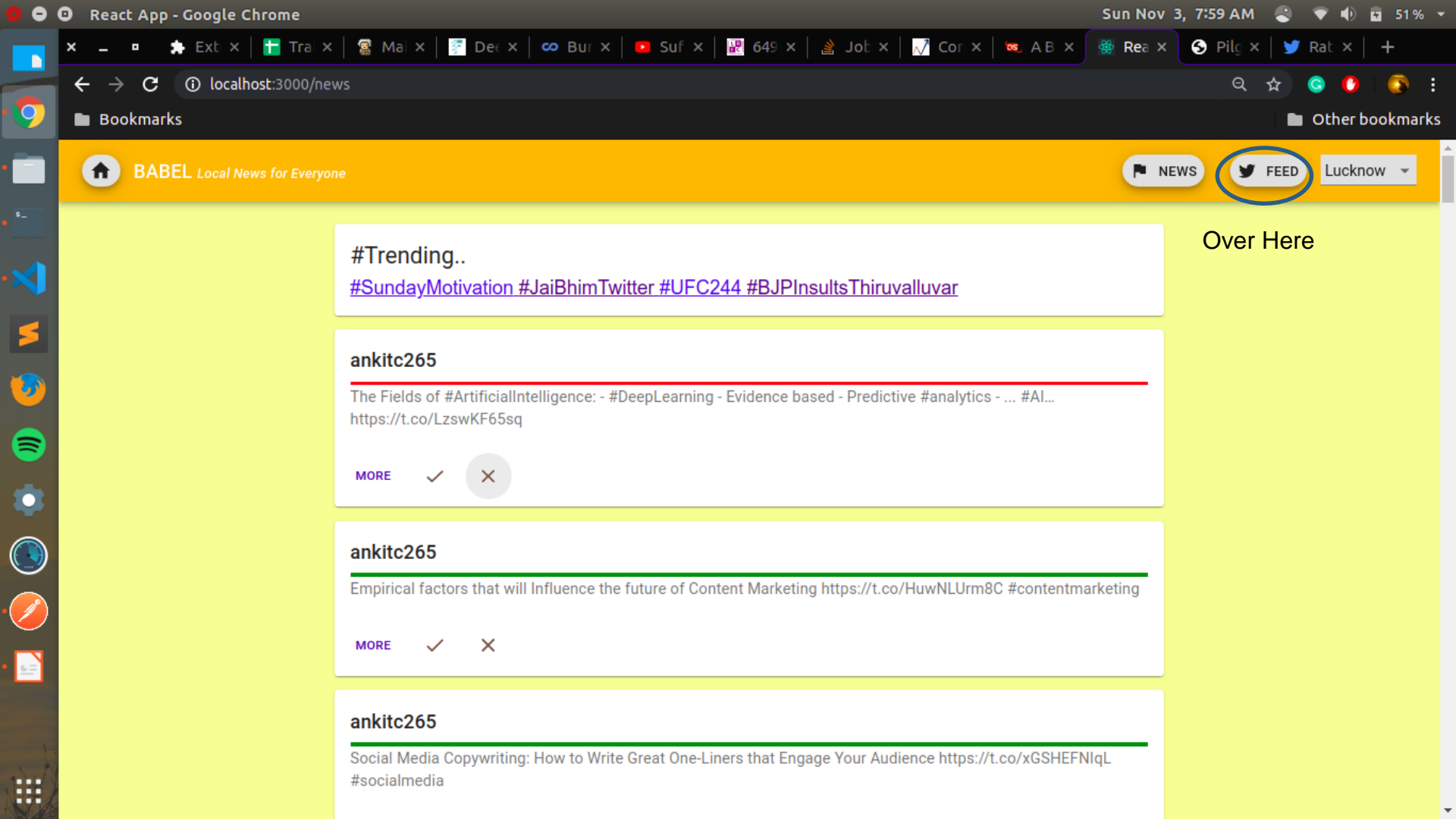


PRINT



Erumely will be ready to welcome Ayyappa devotees by Nov. 15, says Minister

With just two weeks left for the annual pilgrim season in Sabarimala to kick off, Devaswom Minister Kadakampally Surendran here on Saturday reviewed the preparations in Erumely and Ettumanur, two major stopovers of pilgrims to the Ayyappa temple.



BABEL Local News for Everyone

NEWS

FEED

Lucknow

#Trending..

[#SundayMotivation](#) [#JaiBhimTwitter](#) [#UFC244](#) [#BJPInsultsThiruvalluvar](#)

ankitc265

The Fields of #ArtificialIntelligence: - #DeepLearning - Evidence based - Predictive #analytics - ... #AI...
<https://t.co/LzswKF65sq>

MORE



ankitc265

Empirical factors that will Influence the future of Content Marketing <https://t.co/HuwNLUrm8C> #contentmarketing

MORE



ankitc265

Social Media Copywriting: How to Write Great One-Liners that Engage Your Audience <https://t.co/xGSHEFNlqL>
#socialmedia

Over Here

CLICK

joybhattacharj

It's Amartya Sen's 86th birthday. He's not just a Nobel Laureate, he's also probably the only Nobel Laureate whose...
<https://t.co/BMZzpSD0FQ>

MORE



JAKIR22121981

@abbas_nighat @narendramodi @AmitShah @JPNadda @ManojTiwariMP @neelkantbakshi @PratyushKanth
@AshokGoelBJP... <https://t.co/YtrGnlVr1s>

MORE



RailwaySeva

@alamgirmohammad Kindly specify what type of medical assistance you require.

MORE



bsesdelhi

@GusainParmod Request you to share your CA number and contact details, so that we can process the complaint.

MORE



navin_nath

@AshwiniBJP @PMOIndia @narendramodi @AmitShah @JPNadda We are suffering due to bad Karma of our Ancestors during partition in 1947.



Joy Bhattacharjya ✓

@joybhattacharj

Work in sport. Quiz. Read. Write. Try to think

📍 Delhi

📅 Joined November 2008



Joy Bhattacharjya ✓

@joybhattacharj

Follow

It's Amartya Sen's 86th birthday. He's not just a Nobel Laureate, he's also probably the only Nobel Laureate whose name was given by another Laureate. Rabindranath Tagore coined the name Amartya, meaning 'Immortal.'



7:45 AM - 3 Nov 2019

3 Retweets 53 Likes



1 3 53

