Indian Institute of Technology Bombay

Content Tools for Collaborative Communities

EKLAVYA (Fundamental Research Group)

Team:

Dhanush S R, IIT Patna

Naveen T, IIT Patna

Saurabh Shwetabh Singh, IIT (ISM) Dhanbad

Aadarsh Singh , IIT (ISM) Dhanbad

Project Incharge:

Mr. Nagesh Karmali

Mentors:

Mrs. Firuza Aibara

Mr. Abhijit Bonik

Mr. Rohit Prasad

July 6, 2018

Outline

Introduction

Introduction

- Collaborative Communities is a platform where the main aim is to bring together people from all around the world on a common platform to share educational content with the community.
- For such a platform it is necessary that it should contain proper tools to upload, edit and manage the educational content they wish to share.
- It is also necessary that the tools must be user friendly and intelligent enough to manage the content collaborated by numerous persons.
- In such a system it is also necessary to make sure that the content should be appropriate for the viewing of everyone in the community.

Modules Added/Changed

- Real-time Collaborative Editor
- Inappropriate Content Filter
- Wiki
- Interactive Content - H5P

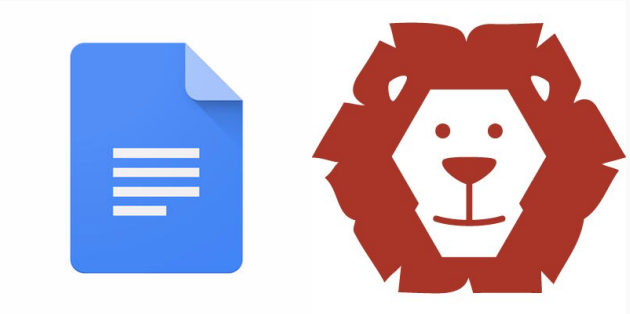
Technologies Used

- Django
- Django-wiki
- Etherpad
- Node
- Socket.IO
- H5PP - HTML5 Package Python
- PyTorch
- Pandas
- GloVe
- NumPy
- bcolz

Realtime Collaborative Editor

What is Collaborative Editor?

- A collaborative text editor is a text editor with an additional feature of accommodating the changes of various authors editing a certain article simultaneously and at the same time displaying the changes made by any author to every other author, both happening in real-time.
- Example - Google docs, Sharelatex



Need for Collaborative Editing in Collaboration System

- In Collaboration system, a community article has 4 states viz. draft, visible, publishable and published.
- To edit an article in visible state, a collaborative editor is required.

Etherpad - Features

- Real-time collaborative editor
- Scalable to thousands of simultaneous real time users.
- Provides full fidelity data export.
- Plethora of plugins available.
- Based on Operational Transform and Changesets.

Etherpad - Implementation

- Node server runs Etherpad.
- Django server runs Collaboration System.
- Iframe is used to embed the etherpad instance.
- Database used is the same as that of the main system.
- Socket.IO is used to enables real-time bidirectional event-based communication

Workflow

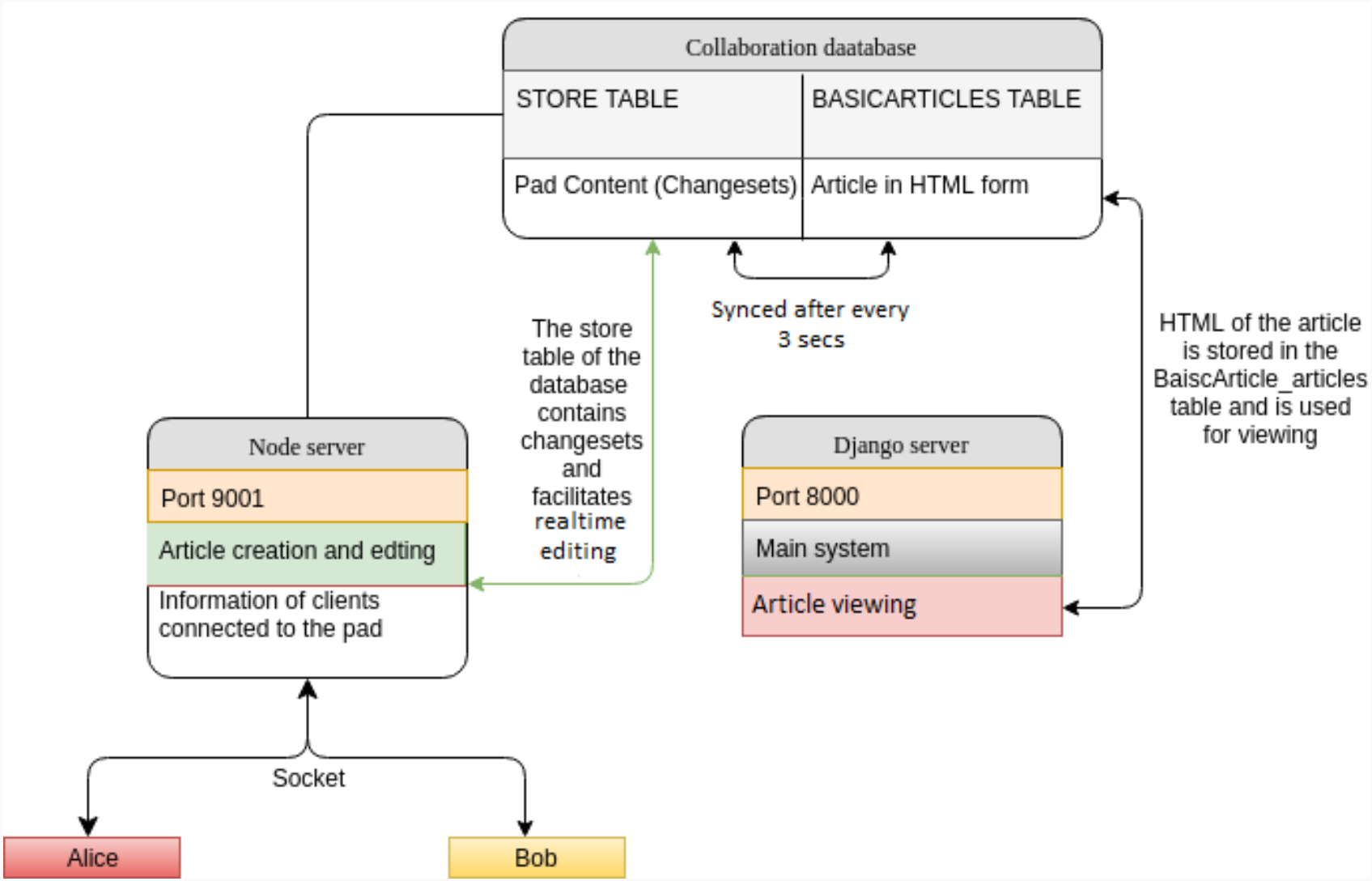


Figure 1: Workflow

Interactive Content

Need

- The existing Collaboration Communities portal renders only Basic Article and Course creation.
- The aim has always been to expand, enhance and extend the existing system.
- Creation of Interactive Content in the portal is one of the ways to go ahead in achieving this aim.

What is H5P?

- The option of creating interactive content is made available to us by H5P - HTML5 Package, an open source package.
- It is currently available only for WordPress, Drupal and Moodle.
- This package facilitates creation of interactive videos, interactive presentations, quizzes, interactive timelines and more.

Examples of H5P Content - 1. Course Presentation

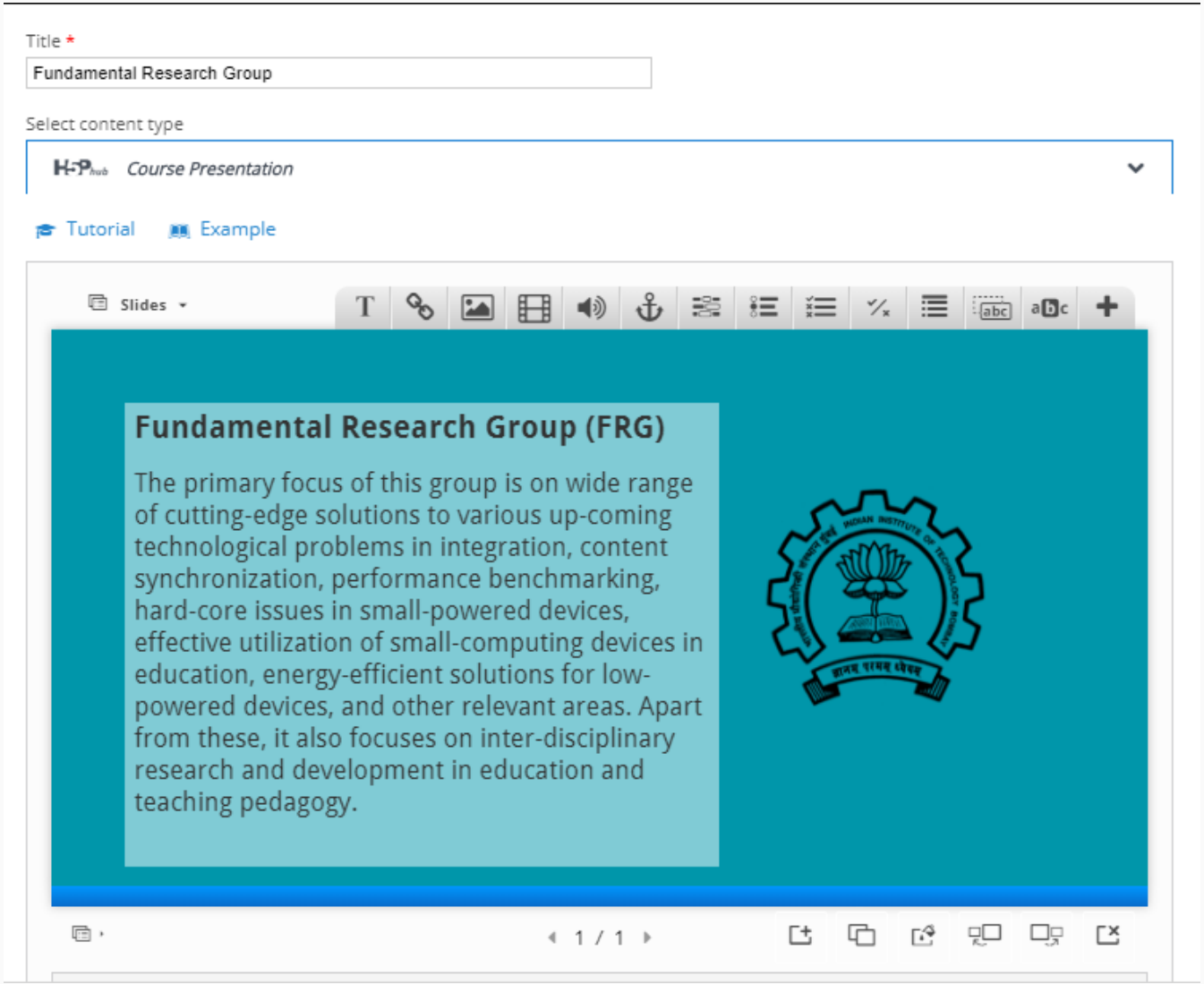


Figure 2: CoursePresentation

Examples of H5P Content - 2. Interactive Video

Figure 3: Interactive Video

Our Approach

- The Collaborative Communities portal is a django project and there is no official support for creation of H5P contents for django.
- To implement H5P in our project, we used H5PP - HTML5 Package Python available for django provided by Joubel.
- Here is a link of the repository - `https://github.com/DrClockwork/H5PP`

Workflow

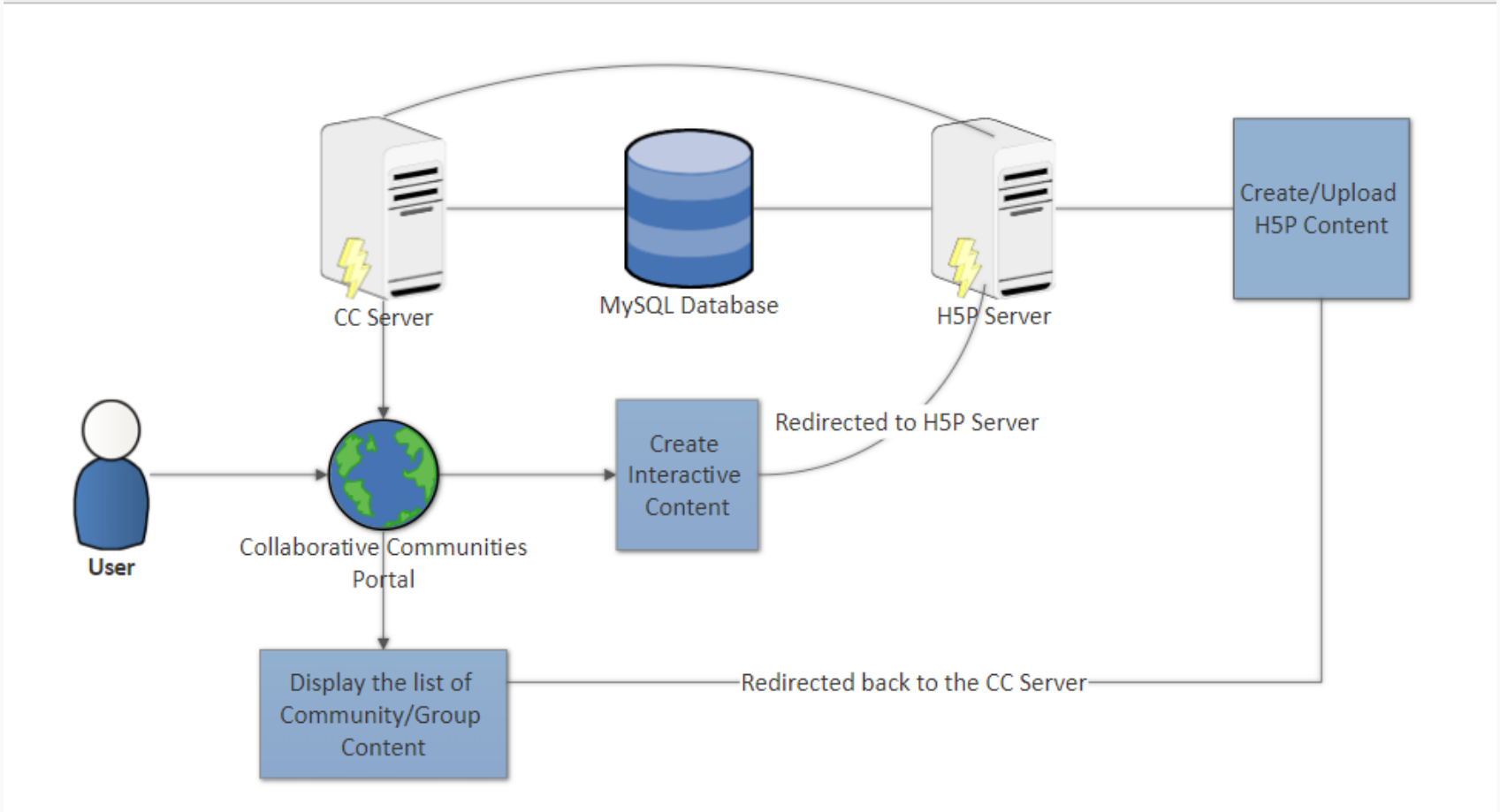


Figure 4: Workflow

Option of h5p creation in CC

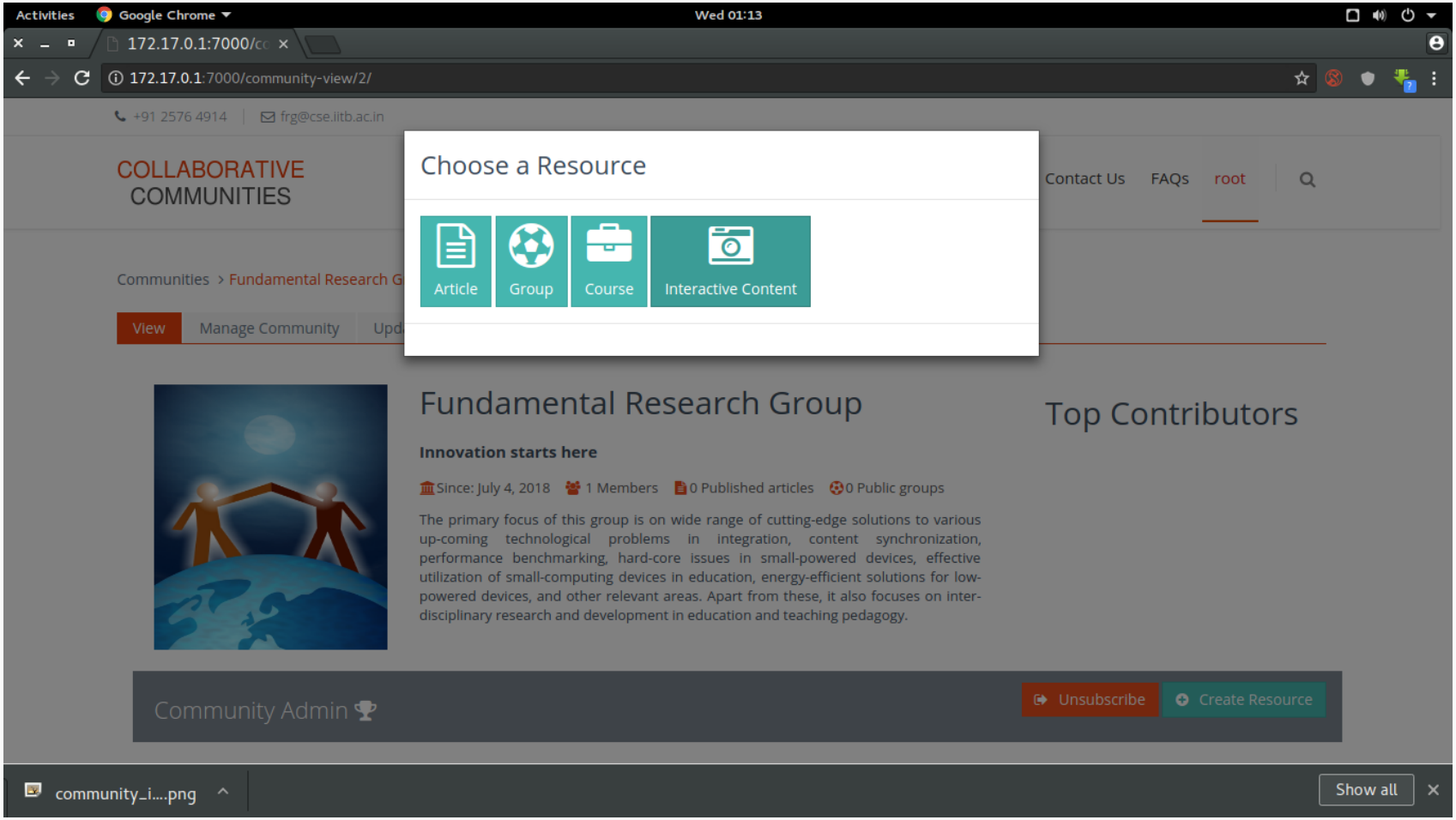


Figure 5: H5P

Create Interactive Content

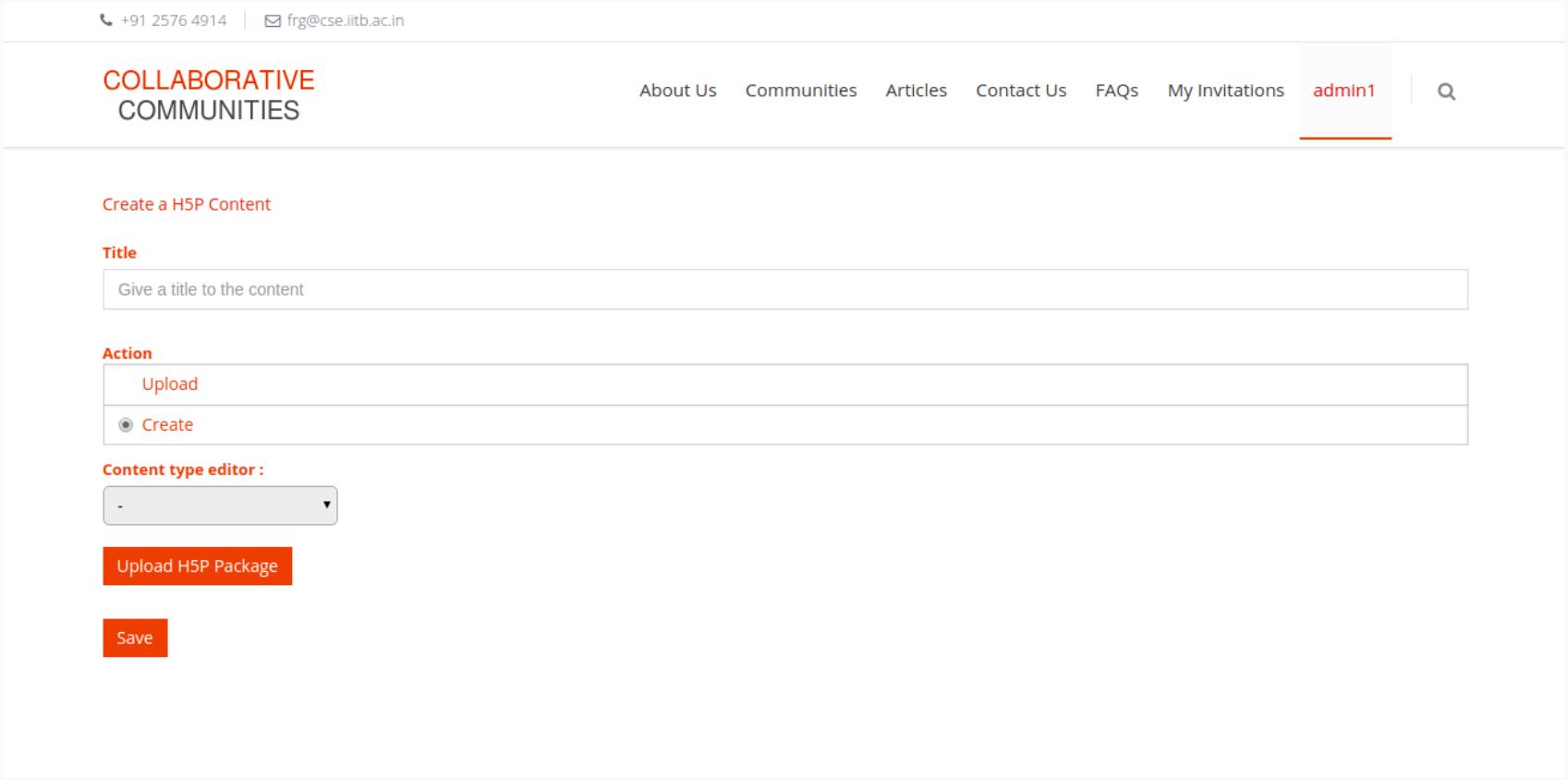


Figure 6: H5P

Content List

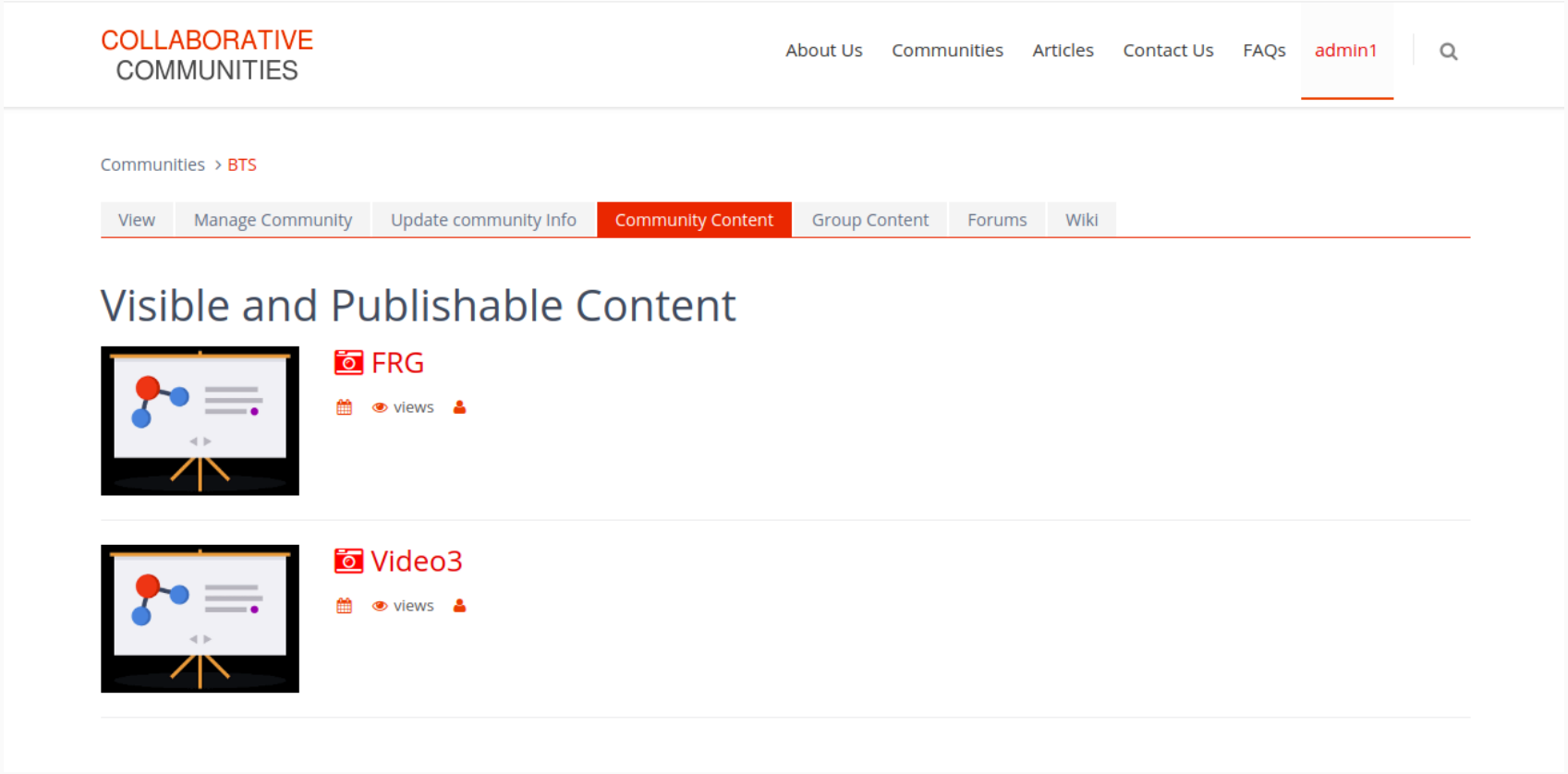


Figure 7: Content List

View Interactive Content

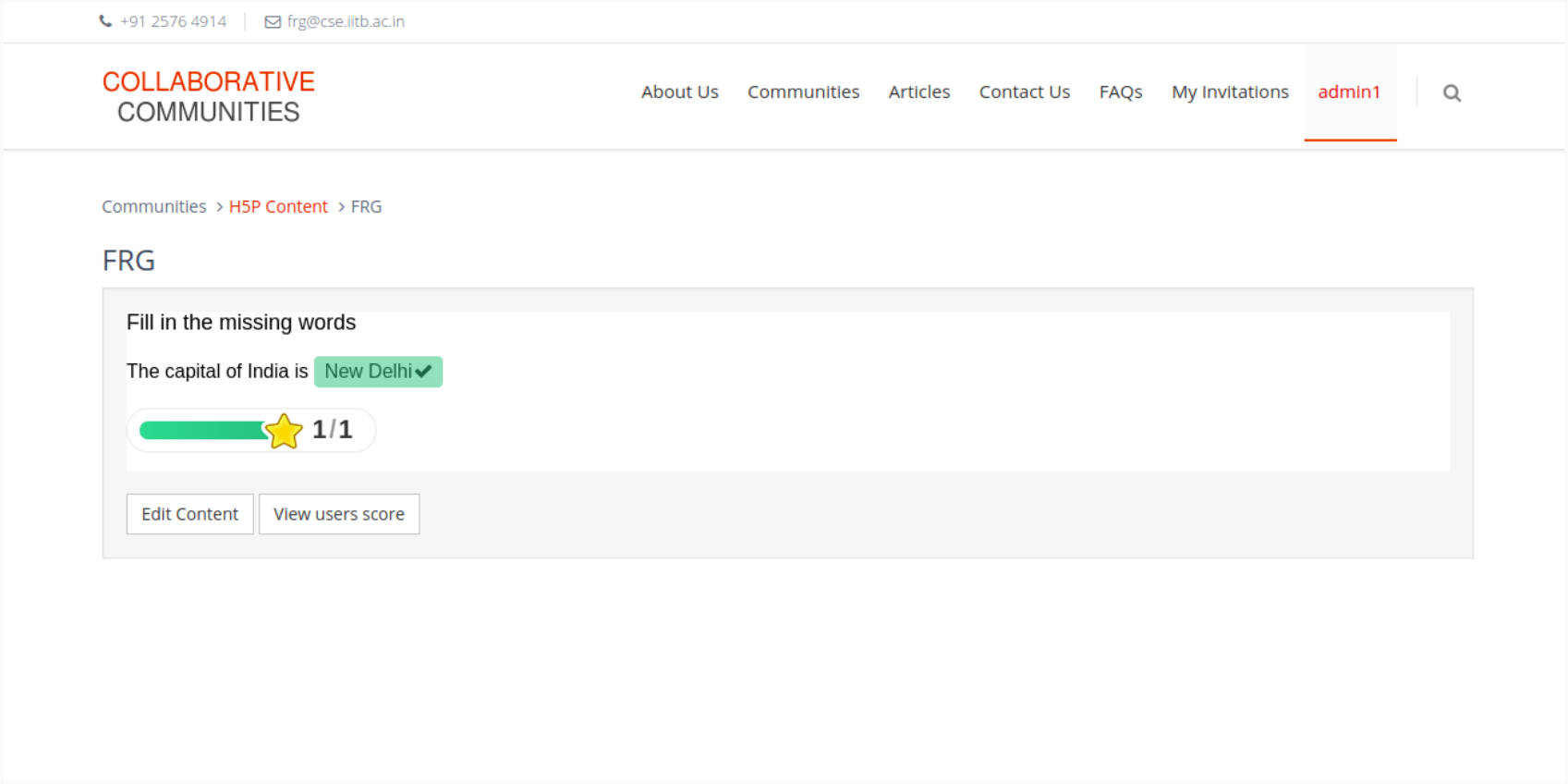


Figure 8: H5P

Wiki

Need for Community Wiki - Page

- The communities created in the portal have some objectives for shaping a better future.
- An overview of the communities is required to list the objectives.
- The overview of the such communities should be precise and should have a proper structure to enhance readability.

Django - wiki

- A wiki is a platform where users can collaboratively modify and structure their contents.
- Django-wiki is one such wiki with permission system, revision control and file attachments.

Structure of Django - Wiki

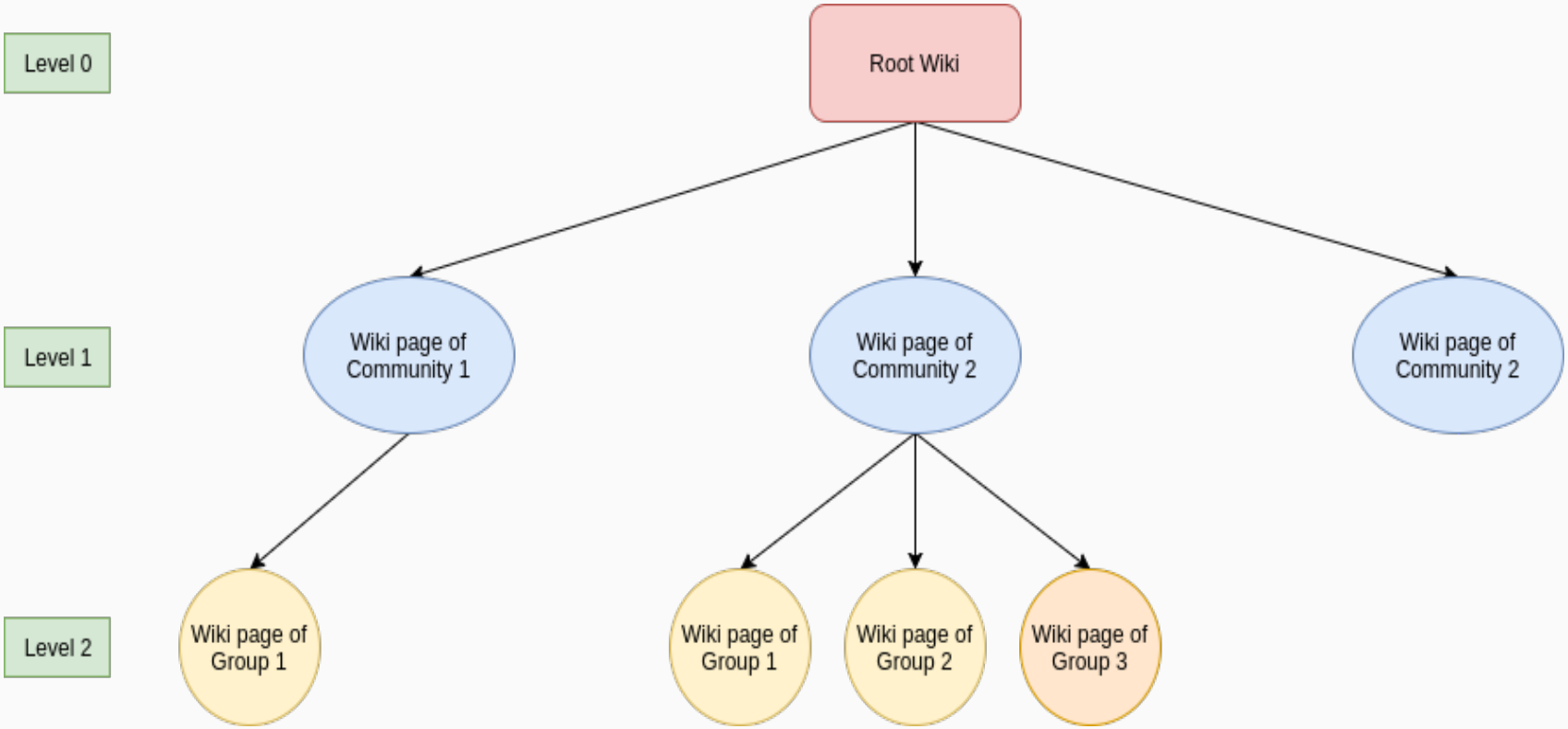


Figure 9: Structure of Django-wiki

Wiki page of a community

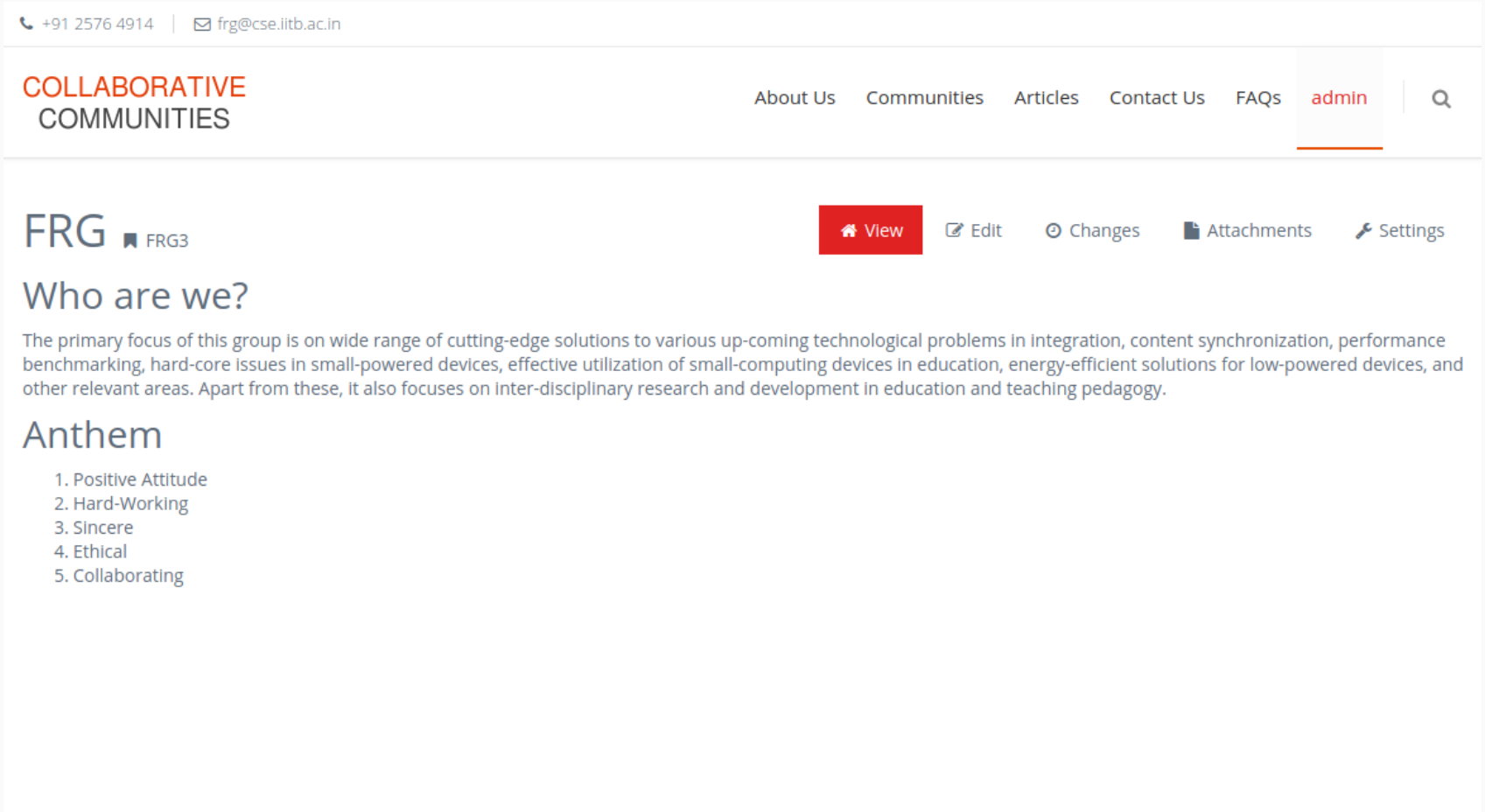


Figure 10: Wiki

Inappropriate Content Filter

Introduction

Why an Inappropriate Content Filter is needed?

In a recent Pew survey, 39% of experts and leaders in fields related to online discourse said that the future will be more shaped by harassment and troll-like behaviours, versus 19% who said it would be less shape.

So what is Inappropriate Content Filter?

Inappropriate Content Filter is a tool that can be used to check the appropriateness of a piece of information for a large public audience of various background and level of knowledge before actually publishing the information in front of them.

Parts of Content Filter Model

The entire Model used for Inappropriate Content Filter can be broadly segregated into two parts.

- Encoder
  - Word Embedding
  - LSTM Network
- Fully Connected Network

WorkFlow

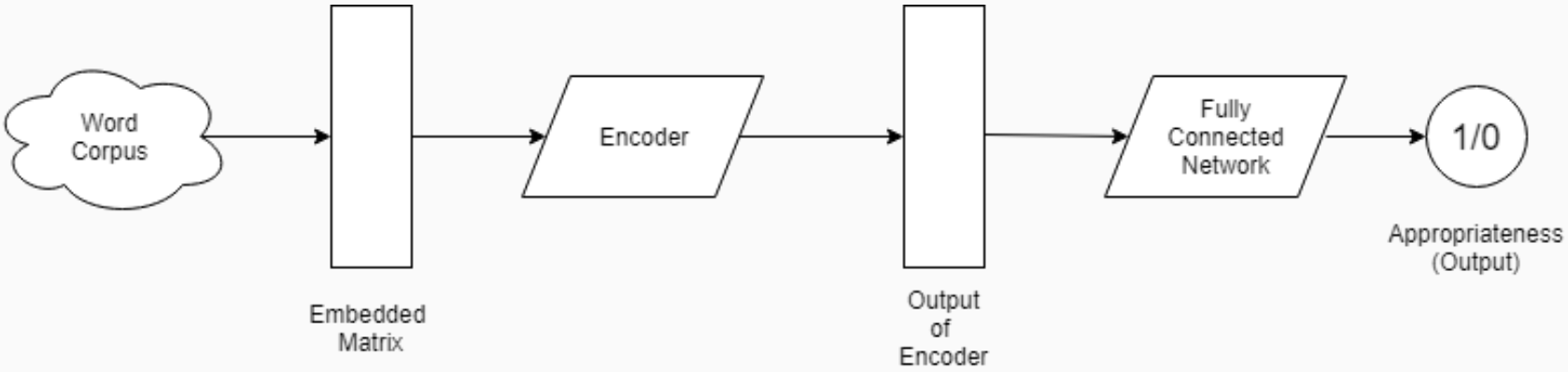
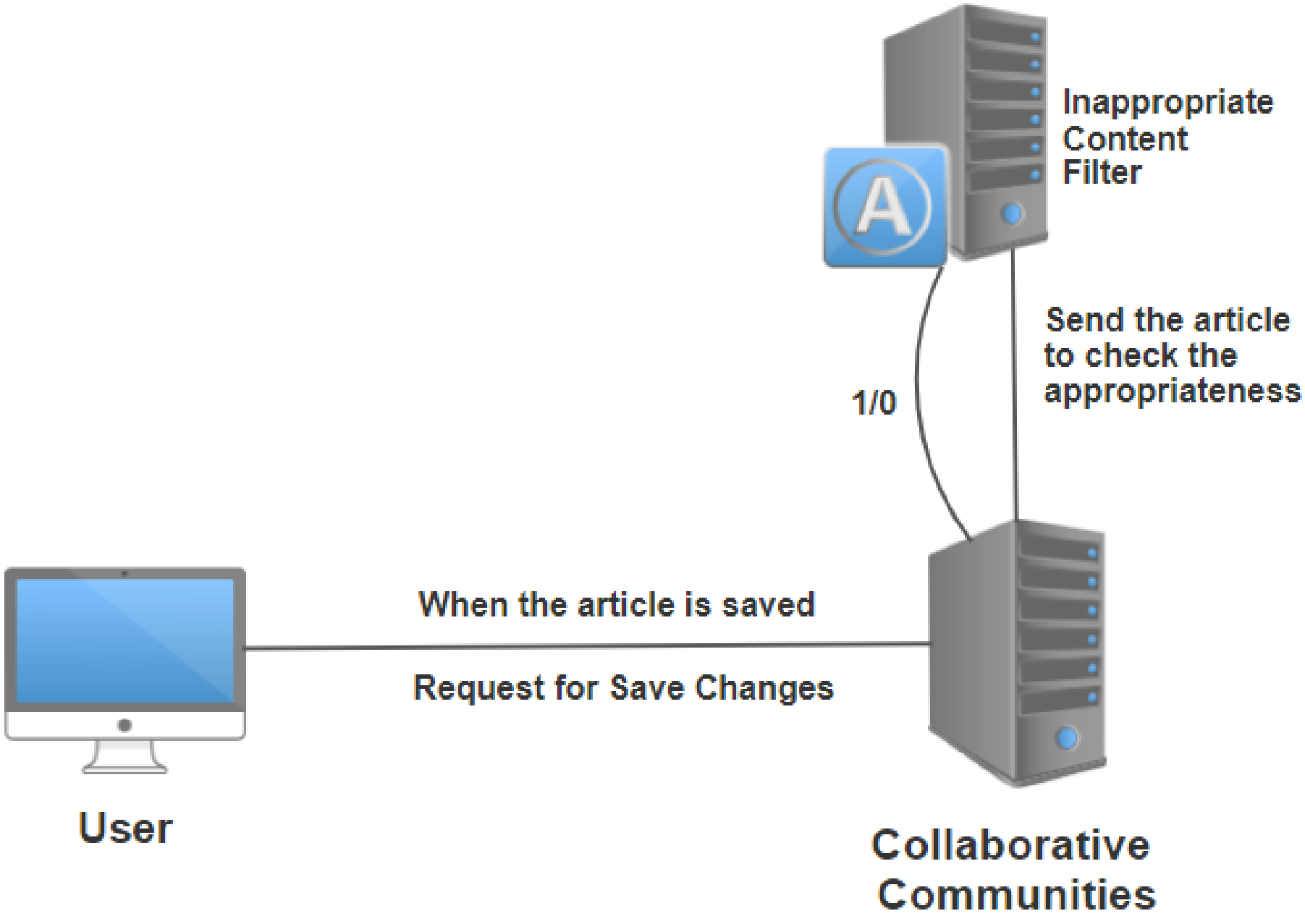


Figure 11: Brief Visualization of Model Used

Overview of Inappropriate Content Filter

- Implemented using Pytorch.
- Trained using dataset from Kaggle (`https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge`)
- The dataset contains total of 159571 records.
- Got an accuracy of 88.17%


Inappropriate Content Filter
Send the article to check the appropriateness
1/0
When the article is saved
Request for Save Changes
User
Collaborative Communities

Dockerisation and Link to our code

Dockerisation

The following modules have been fully dockerised

- H5P Module, as 'h5p' container
- Etherpad, as 'node' container

Wiki has been integrated with Collaborative Communities, and thus runs inside the default 'web' container.

Link to our code

The following are the links to our code

- `https://github.com/fresearchgroup/Collaboration-System/tree/develop`
- `https://github.com/fresearchgroup/Community-Content-Tools`
- `https://github.com/fresearchgroup/Collaboration-System-Selenium/tree/pbl`

Conclusion and Future Work

Conclusion

In this project, the following have been implemented:

- Etherpad, which allows real-time collaborative editing of articles.
- Interactive content using H5PP (HTML5 Package Python), which bends the user experiences towards interactive learning.
- 'Django-wiki', which facilitates describing the community's objective in a structured manner.
- Inappropriate content filter, which flags abusive or toxic content.

The system built by us has been tested using Selenium by the 'OER Repository using DSpace' team.

Future Work

- Auto-saving articles based on user activity.
- Implementation of H5P in an iframe
- Make the inappropriate content filter model dynamic.
- Build an architecture to automatically flag the article regarding its appropriateness.

Demo

Demo

https://drive.google.com/open?id=
1It413sa0qdORDMq7zPO7yv6_01TwpkFT

References

References

- H5P - Create and Share Rich HTML5 Content
  https://h5p.org 5P – Wikipedia
  https://en.wikipedia.org/wiki/H5P
- H5PP - HTML5 Package Python
  https://github.com/DrClockwork/H5PP
- Easysync documentation
  https://github.com/ether/etherpad-lite/raw/master/doc/easysync/easysync-full-description.pdf
- Source of Thank You Page Image:
  https://commons.wikimedia.org/wiki/File:Thank-you-word-cloud.jpg


thank you
merci
danke
謝謝
gracias
obrigado
obrigado
dziękuję
bedankt
go raibh maith agat
teşekkür ederim
ngiyabonga
спасибо
рахмат
Баярлалаа
tapadh leat
dank je
sukriya
kop khun krap
terima kasih
arigatô
dakujem
takk
grazie
taiku
sagolun
mauruuru
hvala
kiitos
dhanyavad
dankie
nanni
nandri
köszönöm
gràcie
bayarlalaa
enkosi
sobodi
dèkuji
mèsi
didi madloba
kam sah hamnida
rahmat
তোমাকে ধন্যবাদ
감사합니다
ありがとう
najis tuke
chnorakaloutioun
gratias ago
gràcies
sulpáy
akun
dankon
ačiū
vinaka
спасибі
blagodaram
misaotra
matondo
paldies
grazzi
mahalo
welalin
tack
barka
kia ora
spas
mersi
faafetai lava
tanemirt
rahmet
dhanyavadagalu
diolch
xièxie
ευχαριστώ
shukriya
mercé
мерси
trugarez
djiere dieuf
tau
mochchakkeram
дякую
mamnun
chokrane
murakoze
tenki
obrigada
asante
manana
хвала

ANY QUERIES?