FAKE NEWS DETECTION: MACHINE LEARNING & DEEP LEARNING APPROACHES

Mohammad Alnabhan

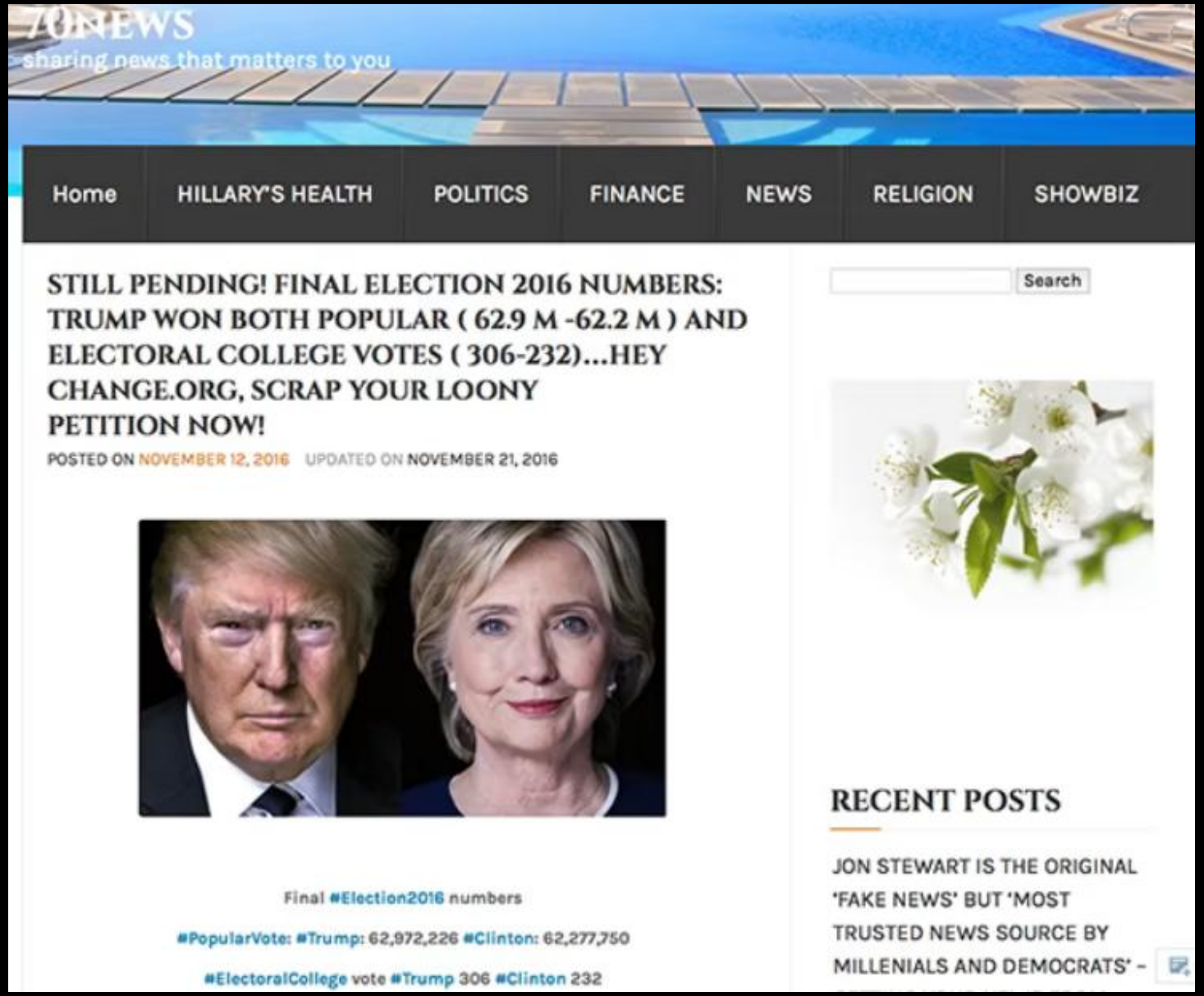
INTRODUCTION

- The proliferation of fake news in the digital age
- Challenges posed by misinformation
- The role of AI in detecting and mitigating fake news

UNDERSTANDING FAKE NEWS

- Definition and characteristics of fake news
- Common platforms and mediums
- Impact on society and public opinion

Example of fake news headlines

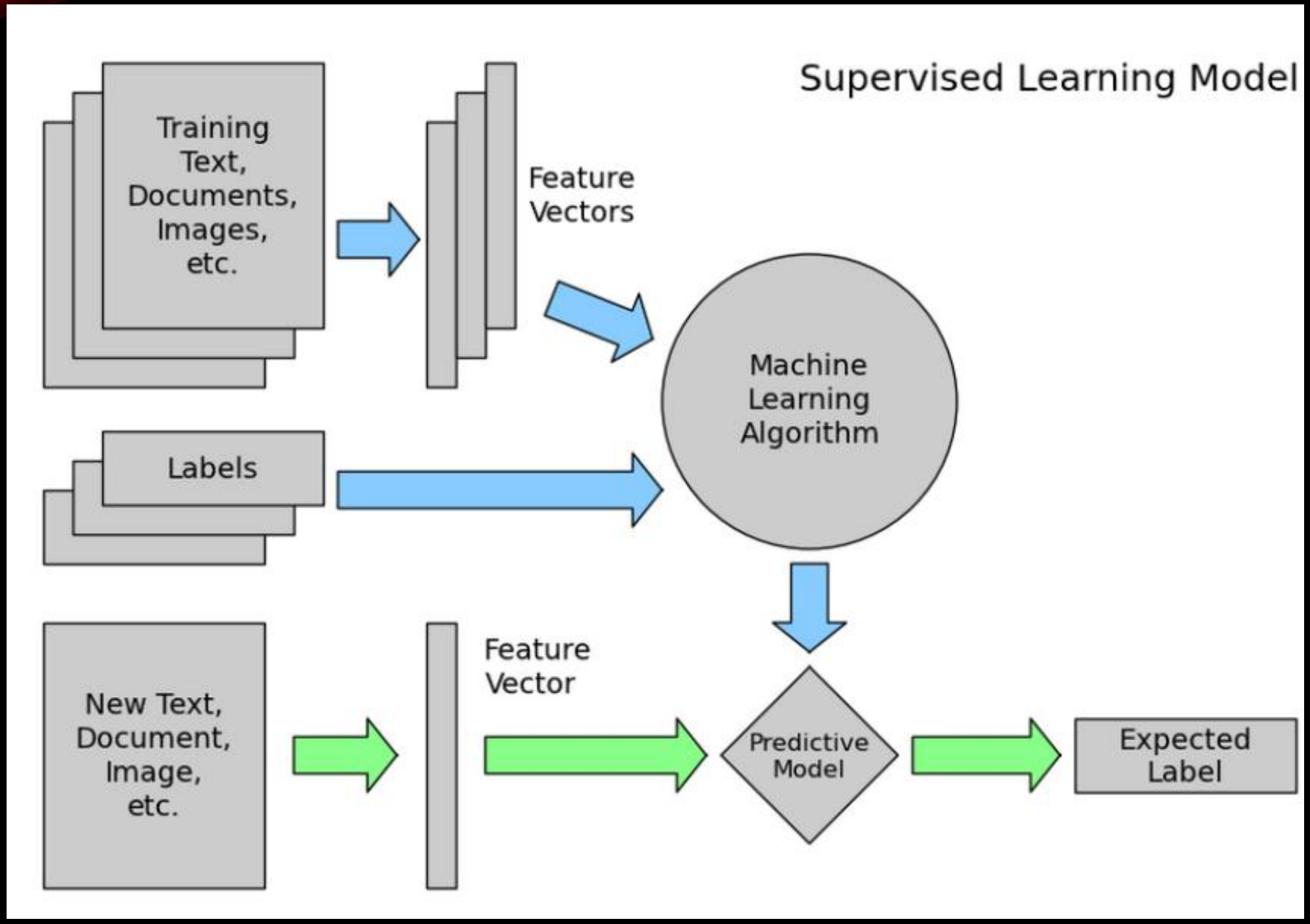


MACHINE LEARNING APPROACHES

- Overview of ML techniques used in fake news detection
- Algorithms: Logistic Regression, Decision Trees, SVM, Random Forest
- Feature extraction methods: TF-IDF, Bag-of-Words

CODE SNIPPET

```
from sklearn.feature_extraction.text import TfidfVectorizer
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(corpus)
```

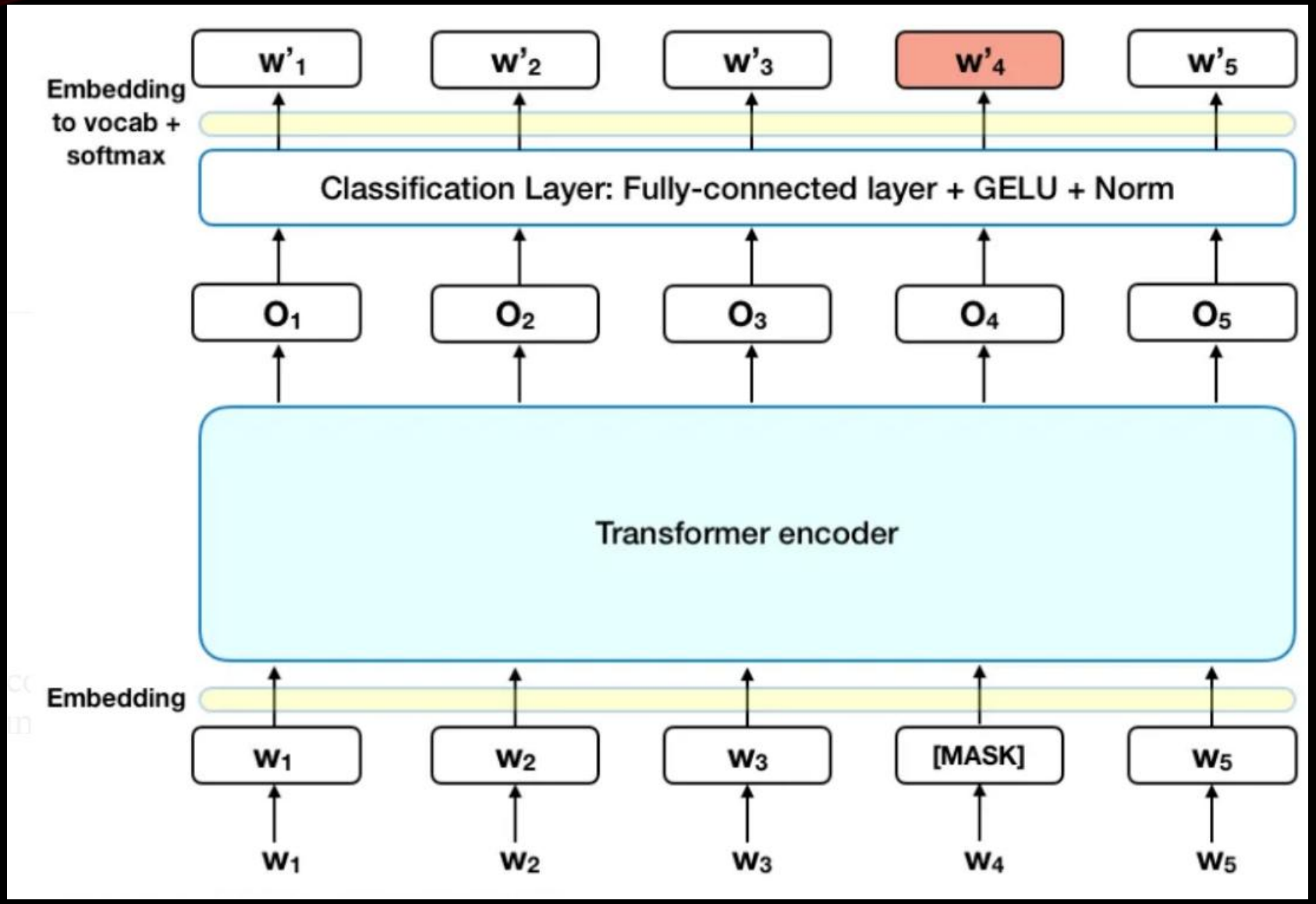

Supervised Learning Model
Training Text, Documents, Images, etc.
Feature Vectors
Machine Learning Algorithm
Labels
New Text, Document, Image, etc.
Feature Vector
Predictive Model
Expected Label

DEEP LEARNING APPROACHES

- Introduction to deep learning in NLP
- Models: LSTM, CNN, BERT
- Advantages over traditional ML methods

CODE SNIPPET:

```
from transformers import BertTokenizer, BertForSequenceClassification
tokenizer = BertTokenizer.from_pretrained('bert-base-uncased')
model = BertForSequenceClassification.from_pretrained('bert-base-uncased')
```

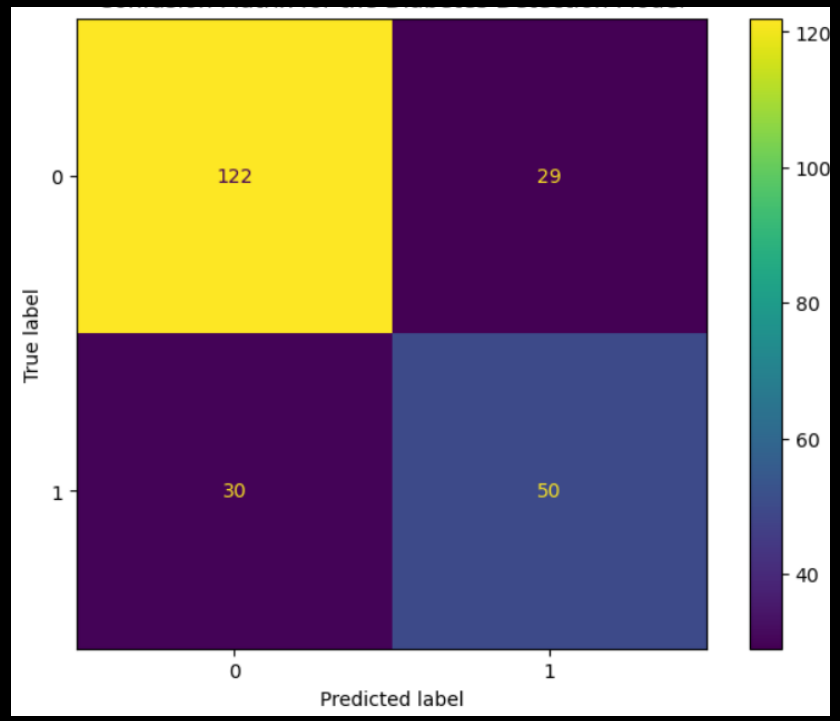

w'1
w'2
w'3
w'4
w'5
Embedding to vocab + softmax
Classification Layer: Fully-connected layer + GELU + Norm
O1
O2
O3
O4
O5
Transformer encoder
Embedding
w1
w2
w3
[MASK]
w5
w1
w2
w3
w4
w5

DATASET AND PREPROCESSING

- Commonly used datasets: LIAR, FakeNewsNet, ISOT
- Data cleaning and preprocessing steps
- Handling imbalanced datasets

MODEL EVALUATION

- Evaluation metrics: Accuracy, Precision, Recall, F1-Score
- Confusion matrix interpretation
- Cross-validation techniques



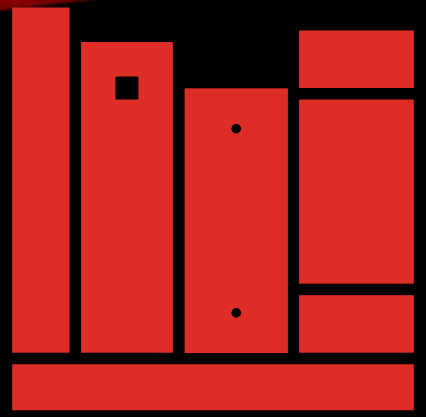
CHALLENGES AND LIMITATIONS

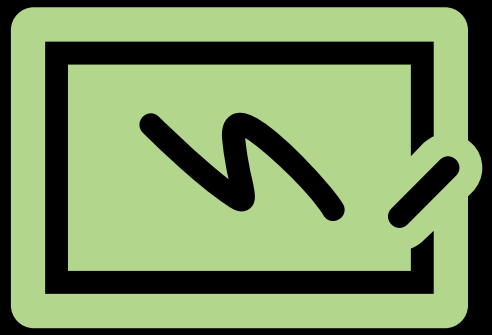
- The evolving nature of fake news
- Language and cultural nuances
- Adversarial attacks on detection systems


The high planetary cost of running large AI models (Crawford, 2021)
The inscrutability of AI tools & their lack of accountability (Raji et al., 2020)
The vulnerability of AI models to bias throughout their pipeline (Suresh & Guttag, 2019)
Challenges
The Emergence of exploitative surveillance capitalism based economic logic built on AI technology (Zuboff, 2019)
The fragility of AI models under adversarial settings (Marcus, 2019)
The opaqueness of their mechanisms (O'Neil, 2016) (Lipton 2018)

CONCLUSION

- Recap of key points
- The importance of continued research and development
- Call to action for stakeholders






THANK YOU!

* This project is funded by the Digital Research Alliance of Canada