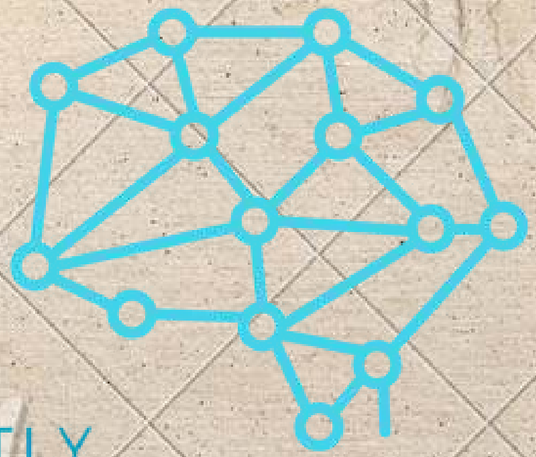
LitGraph



UNDERSTAND INTELLIGENTLY

Visualizing Literature With NLP

Group 7

- Product proposal & Value proposition
- Industry Insights & Competitive Landscape
- Example of Use Cases
  - Deep Dive
- Model Development
  - Trial and Error
  - Deployed Solution
  - MVP
- Further Development and Conclusion

Product Proposal
&
Value Proposition

Product Proposal

Application providing deeper insight into **Old Norse** literature through the combination of **NLP** models and **Network Graph Analytics.**

Value Proposition

- **Multi-disciplinary Educational Tool:** Engage with literature, history, social science. App's graphic summaries enable a deep understanding of the subject matter and to actively engage in the learning process

- **Preservation and Cultural Understanding:** Contributes to cross-cultural understanding by preserving the Icelandic and Old Norse language, literary traditions, and religious aspects. The app helps bridge the gap between cultures and ensures the preservation of Icelandic culture for future generations.

Industry Insights & Competitive Landscape

Industry Insights

Global Edtech market size 2022 → **74 Bn $**

by 2030 → **288 Bn $**

Key factors

Education Technology in Language Learning

Personalized Learning and Customization

Cross-Cultural Understanding and Language Preservation

Data-Driven Insights and Research Opportunities

Competitive Landscape

EdTech Solutions

Zicklearn
LEARNING IS JUST ONE TEXT AWAY

MUSSILA

karaconnect

beedle

Potential Collaborations

Iceland Teachers Association

Government of Iceland

University of Iceland

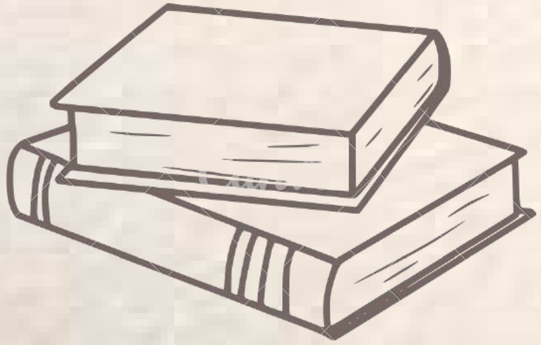
National Programs

Nordic eSchool

Smart Schools

Future Classroom Lab

Example of Use Cases



Conceptualised **Use case:**

- **Graph Analytics for literary works**
  - Character interplay
  - Relationships
  - Actions
  - Degrees of separation

Alternative Use Cases:

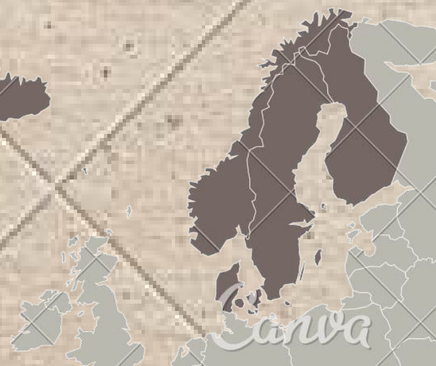
- **Creative idea conceptions** for media
  - for authors, TV, Movies, Character development
- Cultural insight into Icelandic and old

Deep Dive

Old Norse & the Sagas

- Modern Icelandic ≈ Old Norse

- **Insight** into the **culture, beliefs** and **history** of the Norse people.
  - Old Norse literature is divided into four categories
    - Mythology - stories, folklore, Norse mythology
    - Heroic Sagas - "Volsungar Saga" " the Sagas of the Icelanders"
    - Skaldic Poetry - Collection of court poetry and kennings.
    - Eddic Poetry . - Anonymous, Hávamál and Völuspá



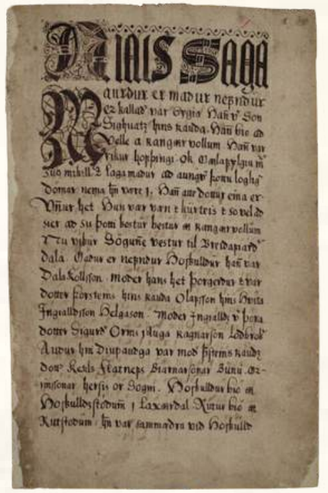
Manuscript Chosen for MVP development

Brennu-Njáls Saga

- One of the most renowned and celebrated sagas of Old Norse literature

Njáls Saga from 1300

Njáls Saga from 1400



Njáls Saga from 1600

- The saga presents around **500** complex and memorable **characters**.
- It explores their personal conflicts, familial relationships, and the tragic fates.

Quick storyline - Brennu Njáls Saga

Main Characters

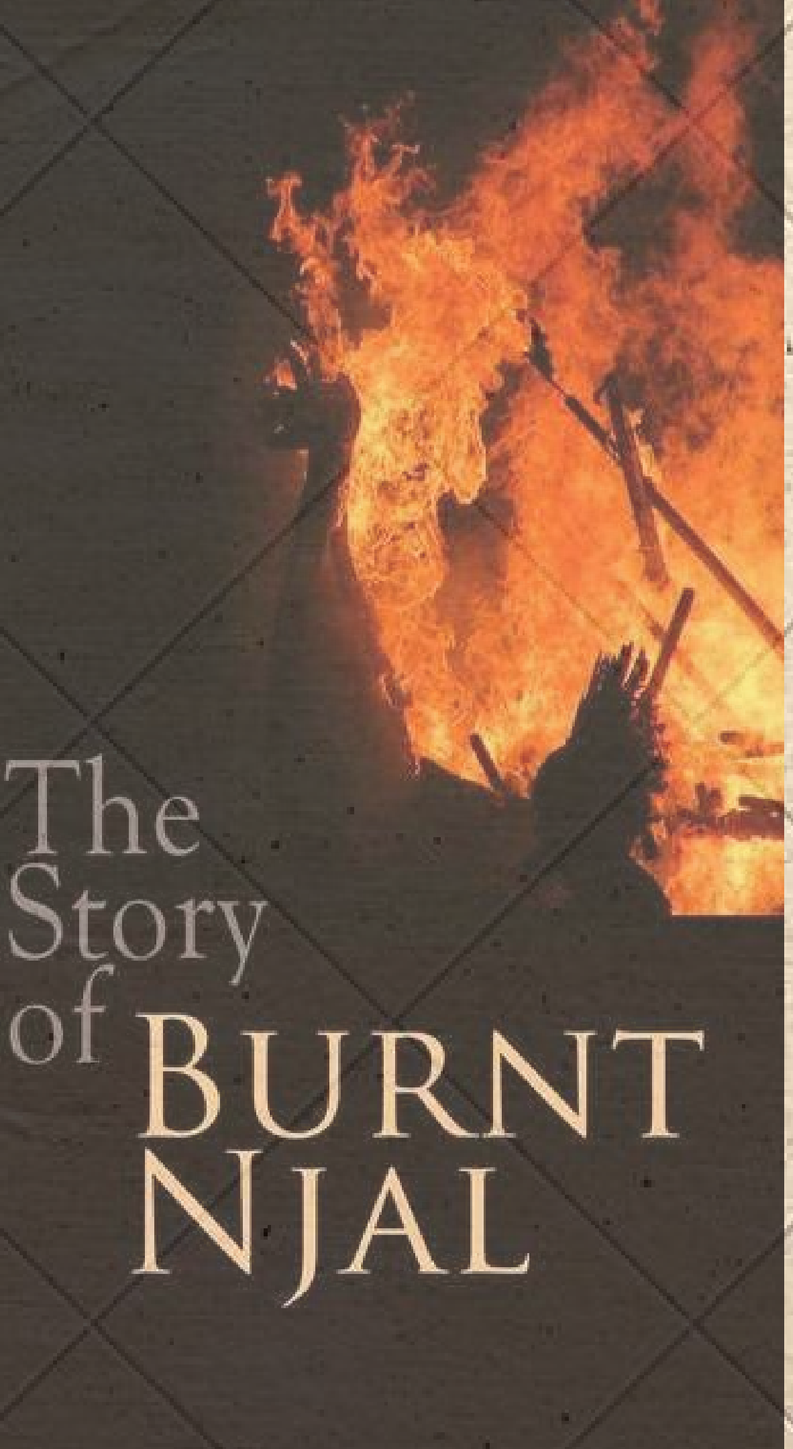
Hrútur Höskuldur

Gunnar Hallgerður

Njáll Bergþóra

Skarphéðin Höskuldur

Kári



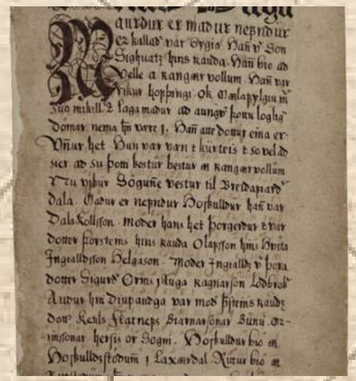
Kaare Salmundsson fra Sudsvestne kommer Njaalsonnerne Grim og Helge til hjælp. Se side 329.
Vore fædres liv.
Gunnar møder Hallgerd paa thingsdemes.
Kampen ved Rangaa.
Torgeir estuund faar vide at Gunnars ægeir er hjemme. Se side 323.
Skarphedin fælder Traain. Se side 325.

Quick Storyline

Model Development

Trial and Error

Trial and Error



Illegible text and proper OCRs not existing for Old Norse

OCR

to rebuild the village as a holiday
ed islanders and bought their plots,
who could expect big money for a
ccessible island? One man sold for a
e rumour went. Collings was larger
vest of Dingle loved a character, so
s plans came to nothing.
nought about who owned the Great
hties, when an advertisement in the
l the island for sale for a million
t of a long and complicated legal
based in Dingle, which had acquired
the land, and the Irish government
arles Haughey, the owner of Inis
establish a national park. The issue
preme Court as I crossed on that first
he future was unclear. There were
rought back to life anyway, I had

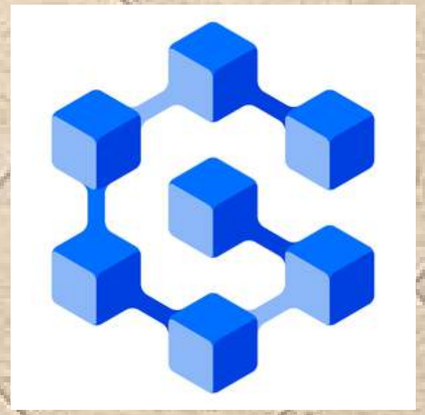
Trial and Error

['ĠMÃ¶rÃ°ur', 'ĠggÃŃ', 'gja', 'ĠSighv', 'ats', 'ĠrauÃ°a']

utf-8 not displaying icelandic characters in the correct format

Trained on

Dataset
Icelandic Gigaword Corpus v20.05 (IGC)
Icelandic Common Crawl Corpus (IC3)
Greynir News articles
Icelandic Sagas
Open Icelandic e-books (Rafbókavefurinn)
Data from the medical library of Landspitali
Student theses from Icelandic universities (Skemman)



Most models being incomplete for the requirements of the project

Trial and Error

Ice-Bert

m3hrdadfi / **icelandic-ner-roberta** ♡ like 0

Token Classification | PyTorch | TensorFlow | Transformers | Icelandic | roberta

Model card | Files and versions | Community 1

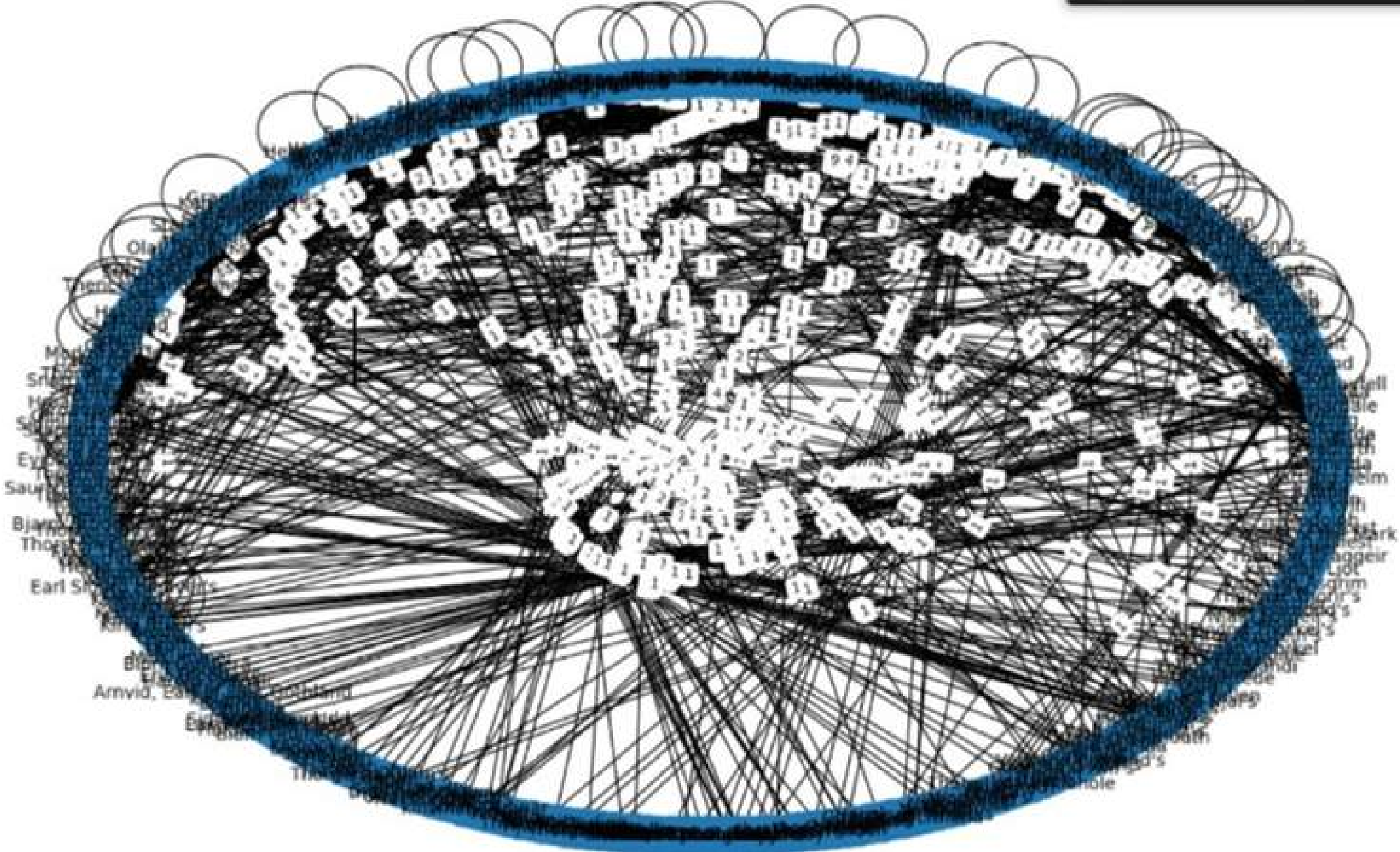
Edit

IcelandicNER RoBERTa

This model was fine-tuned on the MIM-GOLD-NER dataset for the Icelandic language. The MIM-GOLD-NER corpus was developed at Reykjavik University in 2018–2020 that covered eight types of entities:

Ice-RoBerta

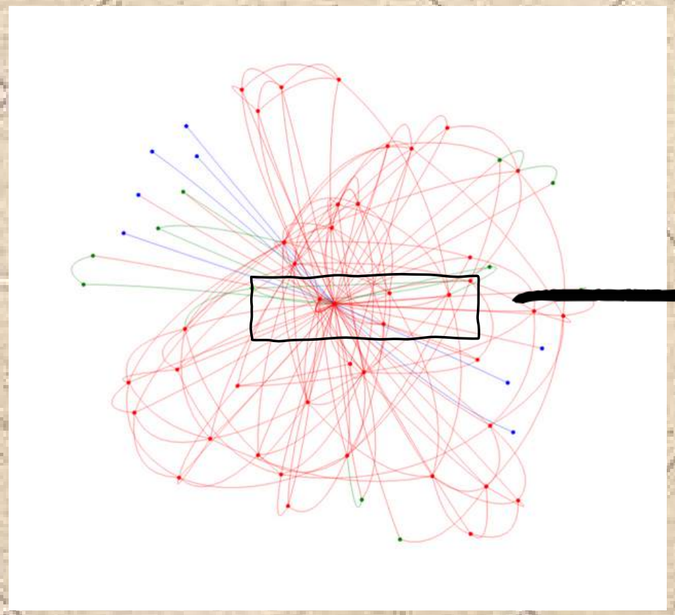
Icelandic language models are not developed enough to tokenize relations

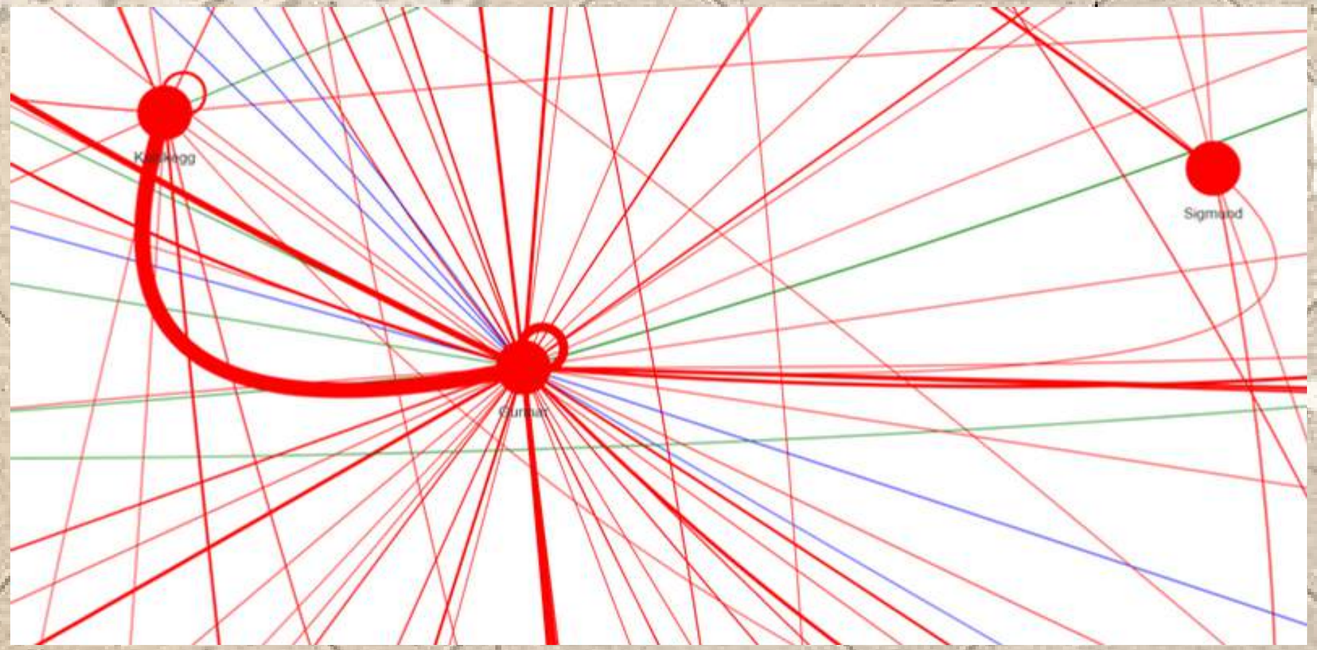


Trial and Error

Pyvis Library using spaCy

Unable to display text on the edges





Colors and thickness representing the weight of the relation

Deployed Solution

Deployed solution

spaCy

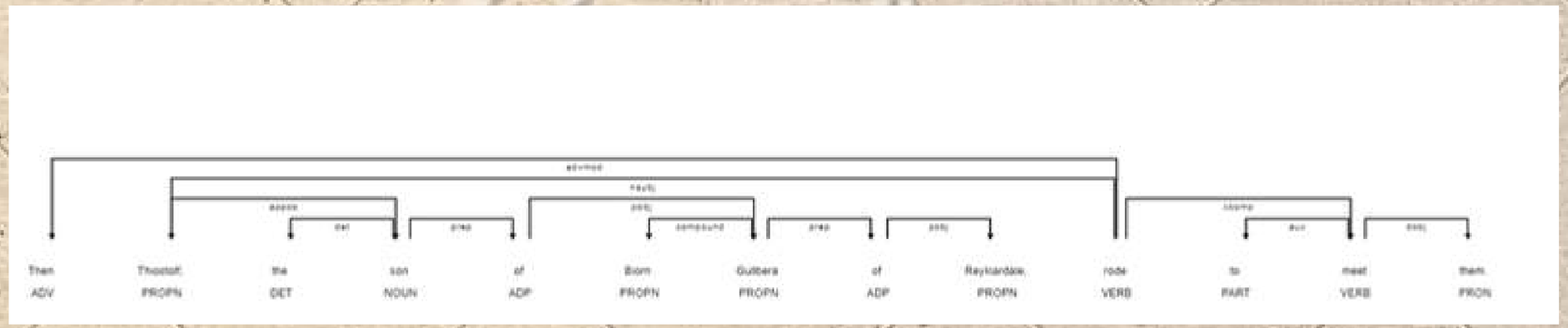
NetworkX

Network Analysis in Python

Streamlit

Deployed solution

spaCy



MVP

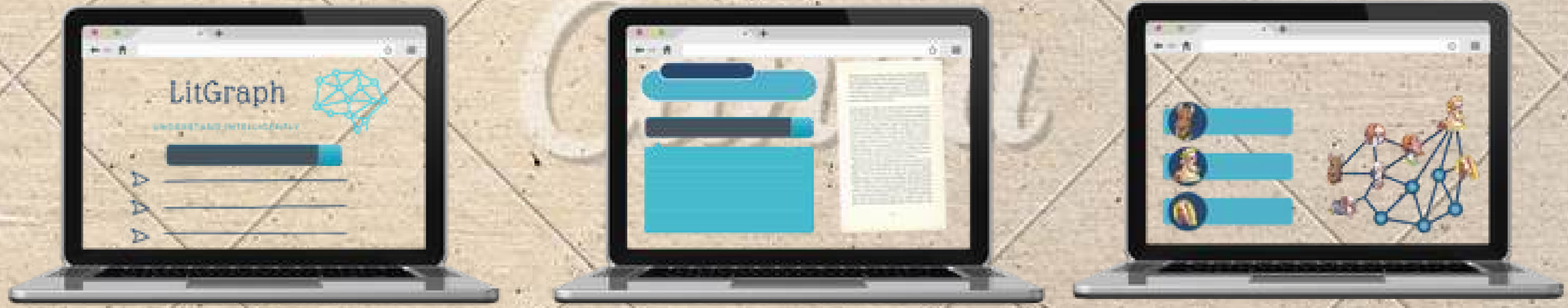
MVP

ted to establish a national park. The issue
upreme Court as I crossed on that first
nd the future was unclear. There were
ng brought back to life anyway, I had

275

MVP Mock up

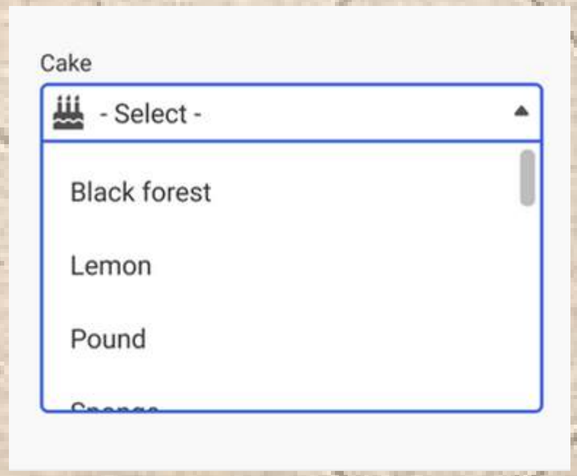
- Landing Page
- Search
- Output



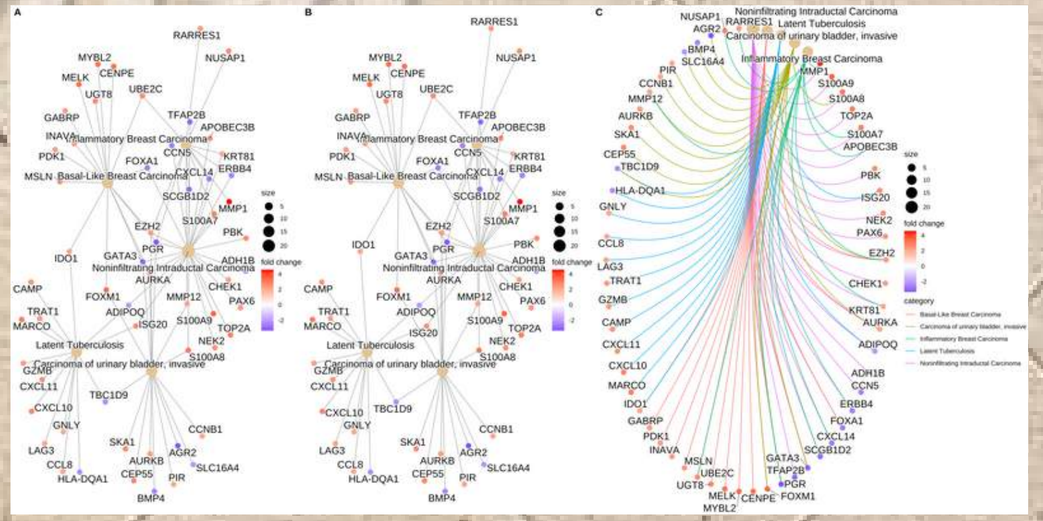
Further Development and Conclusion

Further Development - Improvements Product Features

- Book Selector: Expanding our database



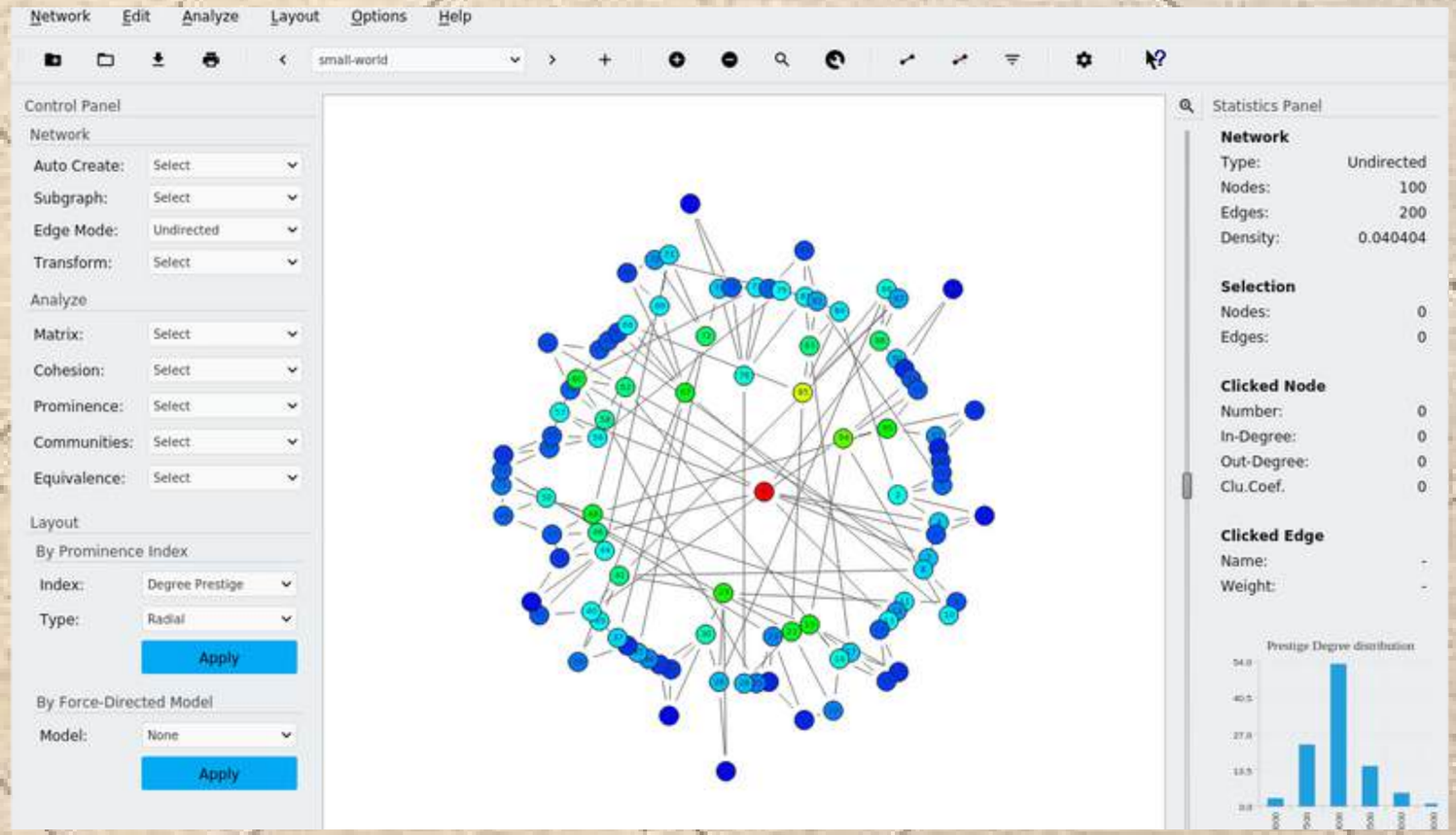
- Enriching our graph by adding more complex / accurate interactions between characters. Relations between different books, chart customization



Further Development - Improvements

Product Features

- Move from streamlit to our own hosted web

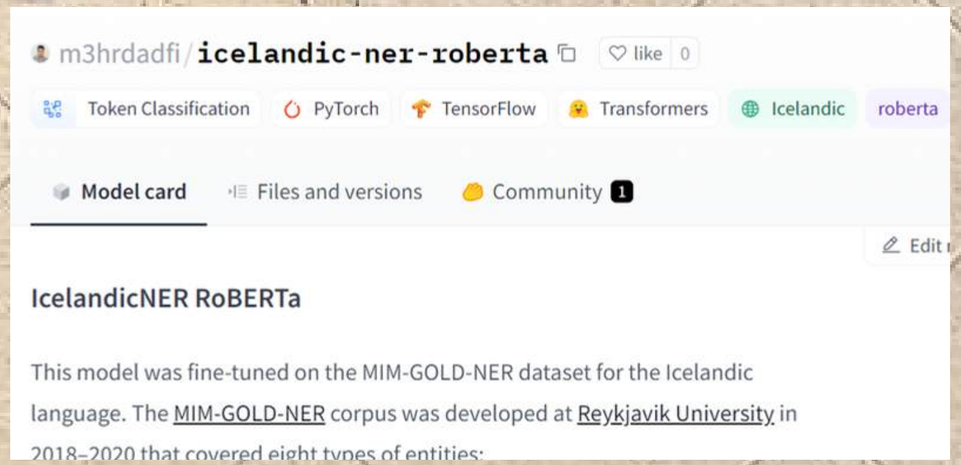


Further Development - Improvements

Model Improvement

Make our platform work using an Icelandic NER model:

- Option 1: Making our app work with icelandic-ner-roberta.
- Option 2: Wait until Spacy release their Icelandic language pack (On their roadmap).
- Option 3: Train our own Icelandic / Old Norse NER Model, with tokenizer for actions and relationships.

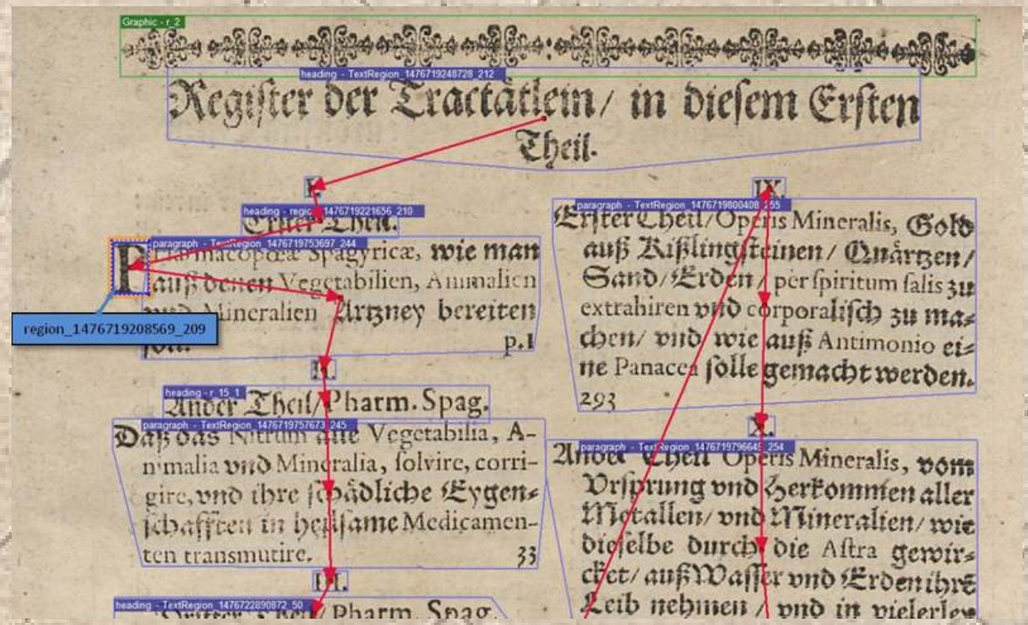


Further Development - Improvements
Beyond the model

- Development of an OCR model

- Availability of applying our model to old texts that have not been digitalized before / better digitalization

  - Access to physical book copies thanks to our collaboration agreement with Government and Universities



Further Development - MLOps

python

ML model - NLP

Flask
web development,
one drop at a time

Get the API

Amazon Lightsail

App Environment

Amazon
EC2

Conclusions

LitGraph

UNDERSTAND INTELLIGENTLY