

# Gibong Hong

Yongsan-gu, Seoul, Republic of Korea  
 gibong.hong@yonsei.ac.kr; [github.com/gbhong](https://github.com/gbhong)

## Research Interests

**Natural Language Processing:** Information extraction, Dialogue system

**Computational Social Science:** User representation learning, Social text analysis

## Education

**Yonsei University, Seoul, Korea**

Mar 2020 – Feb 2022

M.S. in Digital Analytics

- Thesis: User attributes extraction from dialogue with Transformer-based encoder-decoder (*in progress*)
- Supervisor: [Prof. Min Song](#)
- Focus: Natural language processing, Information extraction, Social text analysis
- GPA: 4.28/4.3 (99.6/100)

**Yonsei University, Seoul, Korea**

Mar 2012 - Feb 2019

Double Major in Political Science and Diplomacy (B.A.) and Communication (B.C.)

- Focus: Online media analysis, Causal inference, Film theory and production
- GPA: 3.72/4.3 (93.2/100)

**University of Copenhagen, Copenhagen, Denmark**

Feb 2017 - Jul 2017

Exchange Student, Film and Media

## Publications

### Journals

[J2] **Article retraction analysis in social science domain**

Goeun Heo, YeonJung Choi, Yuheun Kim, **Gibong Hong**, Erjia Yan

*Journal of the Association for Information Science and Technology (SSCI) (under review)*

[J1] **BioPREP: Deep learning-based predicate classification with SemMedDB**

**Gibong Hong\***, Yuheun Kim\*, YeonJung Choi\*, Min Song

*Journal of Biomedical Informatics, 2021. (SCI) (IF 6.317) [\[Paper\]](#) [\[Code\]](#)*

\* denotes equal contribution

## Research Experience

### Research Assistant

Advisor: Min Song

*Deep Text Lab. In Information Department at Yonsei University, Seoul, Korea*

**Personalized conversation AI with user persona and dialogue history**

Jul 2021 - present

- Building conversation model with user attribute extraction from graph embedding and retrieval system.

**Machine learning-based emerging issue detection**

Apr 2021 - Sep 2021

Joint research with *National Assembly Futures Institute of Korea*

- Detected emerging keywords in various domains using LightGBM with moving average features.

**Analysis of retracted scientific publications in social science**

Apr 2021 - Oct 2021

Joint research with [Prof. Erjia Yan](#), *College of Computing & Informatics, Drexel University*

- Built codebook dataset with 651 records and analyzed retracted cases with statistical approach.

**Social media text analysis for 2021 South Korean by-elections**

Mar 2021 - Apr 2021

Joint research with *ChannelA, Korea*

- Conducted time series sentiment analysis for web comments on mayor candidates in Seoul and Busan.

**Biomedical predicate classification using SemMedDB**

Oct 2020 – Feb 2021

- Built relation extraction model for biomedical text sequence by fine-tuning BioBERT
- Built and opened dataset (*BioPREP*) having 165,799 sentences, each labeled with 28 predicates.

**Emotion and sentiment analysis on Korean cinematic characters**

Mar 2020 - Jun 2020

Group research project

- Classified cinematic characters by rule-based feature proportion and RNN-based model.

## Honors and Awards

- 
- Won a prize at **NLP competition for fake news detection** Nov 2020 - Feb 2021  
*NH Investment & Securities, Korea*
- Developed two-stage classification model to remove fake news from main news contents.
  - 1<sup>st</sup> on public leaderboard, 5<sup>th</sup> on private leaderboard across 1,000 teams participated.
- Won 3<sup>rd</sup> prizes in **two NLP shared tasks at AI problem-solving contest** Oct 2020 – Nov 2020  
*National IT Industry Promotion Agency, Korea*
- Classified online comments in 1) Radio streaming application and 2) Korean Army SNS.
  - Tested fine-tuning strategies optimized for understanding syntax structure of Korean language.
- Grand Prize (Minister's Award) at Big Contest 2020** Jul 2020 - Dec 2020  
*National Information Society Agency, Korea*
- Constructed joint LSTM-LGBM model to predict WR, ERA and AVG of Korean baseball teams.
  - **Best performance** across 1,500 teams participated.

## Teaching

- 
- Teaching Assistant for Text mining** Summer 2021  
*Dept. of Digital Analytics, Yonsei Univ. Seoul, Korea*
- Provided special lecture on text mining theories and explained code implementation.
- Teaching Assistant for Python programming** Spring 2021  
*Dept. of Digital Analytics, Yonsei Univ. Seoul, Korea*
- Helped students having difficulties using Python, especially focusing on NumPy, Pandas and data structure.

## Work Experience

- 
- Data scientist internship** at AI transformation product & biz. Aug 2021 - Oct 2021  
*SK Telecom, Seoul, Korea*
- Contributed to building Knowledge Map Explorer for personalized AI agent
  - Recommend items to users by mapping keywords from user behavior log data into item vectors.
- Broadcasting producer** Feb 2019 - Mar 2020  
*JTBC Sports, Seoul, Korea*
- Coordinated live sports events, handled studio segments, and made extra video contents.
- Producer internship** Nov 2018  
*JTBC, Seoul, Korea*
- Contents creator internship** Sep 2017 - Feb 2018  
*SPOTV, Seoul, Korea*

## Technical Skills

---

<b>Programming language</b>	Python, R
<b>Machine learning</b>	Pytorch, Tensorflow, NumPy, Pandas, Scikit-Learn, Gensim
<b>Database / Big data</b>	Oracle SQL Developer, Hadoop, Spark

## Scholarship

- 
- Yonsei Teaching Assistant Scholarship (2020-2021), \$4,700  
 Yonsei Digital Analytics Scholarship for Excellent Performance in Working group (2020-2021), \$3,000  
 Yonsei Digital Analytics Scholarship for Excellent Performance in Problem-solving task (2020), \$1,500  
 Yonsei Truth Scholarship for High Honors (2016), \$1,000

## Extra Activities

- 
- Linux server administrator** Sep 2020 – Mar 2021  
*Dept. of Digital Analytics, Yonsei Univ. Seoul, Korea*
- Maintained the performance of servers and increased efficiency of server environment regarding GPU issues.
- Filmmaker and Screenwriter** Sep 2015 – May 2018  
*Movie Production Club 'The Dreamers', Seoul, Korea*
- Produced short films, wrote lots of short-form scenarios.
- Honorably discharged as Sergeant** Oct 2013 – Jul 2015  
*Military duty as auxiliary police, Gangseo Police Station, Seoul, Korea*
- Squad leader of platoon, managing protests for safety reasons and patrolling for crime prevention.

**Relevant Coursework**

<b>Computer Science</b>	<b>Data Science</b>	<b>Math / Statistics</b>
Python computer programming	Natural language processing	Linear algebra
Data structures	Text mining	Mathematical statistics (1)
Algorithms and their applications	Network analysis & Graph embedding	Statistical analysis of big data
Database management	Recommender system	Introduction to statistics*
Big data parallel processing	Machine learning	Social science research metrics*
	Advanced machine learning	
	Artificial intelligence & Deep learning	
	The basics of big data analysis	
* : undergraduate courses		