Congfeng Cao

Address	Tsinghua Tongfang, Beijing, China
Email	caocf@radi.ac.cn
Phone	+86 13021919407
GitHub	https://github.com/frfy
Personal Page	https://byfrfy.plus/NLP
Blog	https://byfrfy.top
Sex	Male

I am currently working at Tsinghua Tongfang and mainly responsible for natural language processing research. At the same time, I am supervising the research work and progress of six other colleagues. I also have been cooperating closely with Communication University of China in social media analysis. For the past decade, I have done lots of machine learning and deep learning research. As a result, I have a deep understanding of the basic theories and cutting-edge work of natural language processing, machine learning and deep learning. In addition, I have excellent programming skills especially in Python, C/C++ and Java.

EDUCATION

2014.9-2017.7	**University of Chinese Academy of Sciences**	
	MEng Signal and Information Processing	GPA: 3.58/4.0
2010.9-2014.7	**Shandong Agricultural University**	
	BEng Computer science and Technology	GPA: 3.5/4.0
	BEng Remote Sensing Science and Technology	GPA: 3.0/4.0

RESEARCH EXPERIENCE

2017.7-present **Tsinghua Tongfang - team leader/programmer/PM**

Knowledge Graph Products

- Completed the knowledge graph automatic construction tool and used natural language processing to realize automatic entity and relationship extraction.
- Realized knowledge representations.
- https://www.egov.thtf.com.cn/home/app/pingtaiview/article_id/1.html

Big Data Analysis and Decision Support Platform

- Established a platform combining data analysis and natural language processing.
- Categorized the text, extracted entities including economic indexes and extracted relationships.
- Associated the extracted entities such as economic indexes (GDP, CPI) with the data analysis platform and calculated weights of relationships using for making decision and prediction.

- The product is very popular among the governments to help them make decision and analyze the local economic development in many provinces in China such as Jiangsu and Hubei.
- https://www.egov.thtf.com.cn/home/app/pingtaiview/article_id/2.html

Intelligent Government System

- Established a natural language processing platform that mainly implements topic model, text summary extraction, text similarity calculation, knowledge extraction and representation, recommendation and mining.
- Cooperated with the largest journal website in China (National Knowledge Infrastructure，CNKI) to mine journal articles and author information to complete relationship network of author and recommend articles in higher accuracy.

2018.1-2020.1 **Peking University Health Science Center**

The Chinese Medicine Question Answering System

- Established a Chinese medicine knowledge graph of drugs and symptoms from existing texts.
- Completed a question answering system for symptoms and drugs.

2018.7-2020.1 **Communication University of China**

The Public Opinion Analysis System

- Monitored accounts and gets data from social media (like Weibo, Facebook and Twitter).
- Used deep learning and machine learning to classify and cluster, extract topics, perform sentiment analysis and events monitoring.
- Stored data (relationship, text) and visualized networks. (Neo4j, MongoDB, Highcharts).
- Calculated social distance and similarity by individuals' blogs.

2014.9-2017.7 **Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences**

Large Science Engineering Project-Multimodal Camera System

- Provided remote sensing image mosaic algorithms and image processing for the camera.

Lunar Exploration Project Phase III-Image Algorithm Module Image Database Design

- Used the database to store the original lunar image, the processed image by the algorithm, the algorithm parameters and performed the evaluation parameters of the algorithm quality.

Beijing Science and Technology Project-Development of UAV-borne Multispectral Camera Based on Optical Filter Array

- Provided remote sensing image mosaic algorithm for the camera.
- Performed target detection and image segmentation on multispectral images.

PAPERS & THESES

2017 **Research of UAV-borne Multispectral Camera System Based on Narrow Bandwidth Filter Array.**

- Full text: https://byfrfy.plus/thesis.pdf
- Abstract: https://byfrfy.plus/abstract.pdf

2018 **Congfeng Cao, Junyong Fang, Dong Zhao. Development of UAV-borne multispectral camera based on narrow band width filter array. OPTICAL TECHNIQUE, 2018, 44(1).**

- Full text: https://byfrfy.plus/paper.pdf

HONOURS & AWARDS

2013.9-2014.7 **Outstanding Student Scholarship**
2014.9-2015.7 **Academic Scholarship**
2015.9-2016.7 **Academic Scholarship**
2016.9-2017.7 **Academic Scholarship**

TECHNICAL SKILLS

Machine Learning

- Programming experience with PyTorch (2 years) and TensorFlow (6 years) for deep learning applications.
- Experience with Artificial Neural Networks and the most recent Deep Learning architectures (e.g., ResNet, Bert, Transformer, GPT-3, XLNet, etc.).
- Experience with supervised, unsupervised learning algorithms, reinforcement learning, and Bayesian methods.
- Experience with the most important tools for Natural Language Processing and Computer Vision (e.g., Gensim, NLTK, OpenCV, etc.).

IT

- Proficiency in Python (8 years, primary language). Used several programming languages such as C/C++, Java, R, Visual Basic, HTML, JavaScript, etc.
- Database (e.g., MongoDB, Oracle, MySQL, Redis, Neo4J, GraphDB PostgreSQL, etc.). NCRE: National Computer Rank Examination- Level 4 Database Engineer.
- Knowledge Graph tools (e.g., Apache Jena, OpenKE, Protege, etc.).
- Elasticsearch, Spark and related tools.
- Spider (Scrapy, pyspider).