

China in Europe Research Network (CHERN) - CA 18215

INTRODUCTION TO COMPUTATIONAL METHODS IN CONTEMPORARY CHINESE STUDIES – LEARN BY DOING

Ever wanted to learn computational methods in contemporary China studies, but have no idea on how to get started? Join the hands-on workshop to learn basics in computational methods.

Time: April 28-29, 2022 Venue: Nordic Institute for Asian Studies NIAS, University of Copenhagen, Denmark

Organizers: China in Europe Research Network CHERN; Nordic Institute for Asian Studies NIAS, University of Copenhagen, Denmark; Centre for East Asian Studies CEAS, University of Turku, Finland.

Application period for the workshop: March 7-20 <u>Application form</u>

This training is a continuation to <u>the CHERN training webinar (November 2021)</u> that introduced basic issues and tools in computational methods for contemporary Chinese studies. Now, a twoday hands-on training workshop will be arranged at the Nordic Institute for Asian Studies at the University of Copenhagen on April 28-29.

The participants will learn basics in topic modelling and semantic network analysis. This will be done step-by-step, guided by experienced trainers and a technical assistant. There will be preprepared English and Chinese language datasets so that the training can kick-off smoothly. The datasets are collected from the Chinese news site Xinhuanet. Trainees work in small groups of 3-4 persons focusing on one research task. The workshop is complemented with two webinar presentations on Chinese digital materials.

Prerequisites for participation

Participants are expected to have a basic understanding of what computational methods are, but no prior coding skills are needed. The prerequisite is that you are familiar with the contents of the CHERN training webinar we held in November (which are covered in this 90 minutes <u>video</u>) and the 7-<u>page introduction to computational methods in Chinese studies</u>. No knowledge of Chinese is required to take part in the workshop, although for participants who can read Chinese sources these will be provided.

CONTENTS OF THE TRAINING

This programme may be subject to eventual changes due to circumstances beyond our control.

28 April

9:00-15:00 UTC+1 Christian Göbel: An introduction to topic modeling

15:15-16:15 UTC+1 Joanne Kuai: zoom-webinar "Media innovation in China and its implications on contemporary China research"

Christian Göbel: An introduction to topic modeling

How to quickly make sense of a body of text that is too large for a single human to read? On Thursday, the workshop provides a practical introduction to topic modelling, a form of text mining that uncovers hidden semantic structures ("topics") in a corpus of documents. Topic modelling is a form of unsupervised machine learning suitable for eliciting how prominent certain topics are in a corpus, how they are connected, and how they develop over time. With a bit of caution, researchers can also use topic modelling algorithms to classify documents and thereby make them amenable to statistical analysis.

The training consists of four parts. First, participants will receive a brief introduction into the use cases of topic modelling, the most commonly used algorithms, and their strengths and weaknesses. Second, participants will learn how to preprocess text for analysis (remove stop words, lemmatise words, segment Chinese language documents), select the hyperparameters of Latent Dirichlet Allocation (LDA) models and decide on an appropriate number of topics. In the third part, we will use the fitted model to classify documents, inspect a random sample of classified documents and discuss the accuracy of classification. Finally, participants will learn how to visualise the prevalence, development and connection of topics as a bar chart, line diagram and correlation plot, respectively.

Joanne Kuai: Media innovation in China and its implications on contemporary China research

In the face of a universal challenge to news media in the digital age, how are Chinese media leveraging artificial intelligence (AI)? How has the Chinese media landscape changed amid China pushing its national strategy to become a global AI superpower by 2030? When designing research projects on contemporary China and sourcing Chinese media, what should we pay attention to? This talk will provide an overview of the state of the art of media innovation in China, the development stages, and the motivating forces behind it. It will address some methodological issues when designing research projects in relation to the phenomenon and discuss its implications on research and society at large.

29 April

9:00-12:00 UTC+1 Jun Liu: An introduction to semantic network analysis

13:00-14:00 UTC+1 Eric Zhang: zoom-webinar "Contemporary Chinese Digital Materials in the Big-data Era"

Jun Liu: An Introduction to Semantic Network Analysis

Semantic network analysis (SNA) has been considered as the most suitable analysis method for unstructured (big) text data analysis. It uses network analytic techniques on paired associations based on shared meaning to reveal meaning embedded in text and to gain quantitative and qualitative insights. This session presents an overview and a practical introduction to SNA, including theoretical presuppositions as well as the skills to extract, represent, query, and visualize textual content. The participants will be able to conduct SNA, interpret, and report research results and limitations after the workshop.

On Friday, the workshop consists of three parts, each with 45 mins. The first part discussed the theoretical grounding for SNA as well as its advantages and limitations in comparison to other textual analytic techniques. The second and third parts involve step-by-step guidelines for conducting semantic network analysis, including tokenizing, sentence splitting, estimating, analyzing, and visualizing the chosen text with real-world data (in both English and Chinese).

Eric Zhang: Contemporary Chinese Digital Materials in the Big-data Era

The emergence of big-data has made previously hard-to-tackle data and questions comprehensible for researchers. For scholars who would like to be acquainted with digital materials, this talk addresses the spectrum of new types of data and data collection methods that has emerged for Contemporary China Studies, including social media sources, news outlets, and third-party databases. For instance, Sina Weibo posts and comments are excellent sources to study public opinion and narrative campaigns, whereas one should be aware of the pitfalls of working with Sina Weibo's API (application programming interface). This talk also addresses the differences between China's large state-owned media outlets - Xinhua, CGTN, Global Times, and China News, by discussing by their reporting on COVID-19. Besides, this workshop also introduces a few third-party textual data databases, such as Factiva and Ringdata, as they sometimes can be good alternatives to text mining.

APPLICATION FOR THE TRAINING WORKSHOP

In all, 15 participants will be selected to the workshop. Accepted participants will receive financial support for travel and accommodation from CHERN, in accordance with the <u>EU COST Annotated</u> <u>Rules [pp. 84-90]</u>. The support will cover travel costs up to € 300 and € 198 daily allowance. From the Daily Allowance, trainees will pay their accommodation, local public transport costs and meals. Note that the Daily Allowance is a lump sum amount and might not necessarily cover the actual cost you made. Since the program is very tight, the Local Organizer will arrange lunch.

Trainees eligible to be reimbursed:

Trainees shall be engaged in an official research programme as a PhD Student or postdoctoral fellow or can be employed by, or affiliated to, an institution, organisation or legal entity which has within its remit a clear association with performing research.

Trainees eligible for reimbursement are individuals who are affiliated to a legal entity in a full COST member or Cooperating member, a Near Neighbouring country (NNC), or from an Approved European RTD organization. <u>Please see the COST Country and Organizations table.</u>

To apply for the workshop, fill out <u>this form</u>. In the form, **you will be asked about your goals for this workshop**. If you have no previous experience with computational methods, please first watch the <u>video</u> recording of the CHERN training webinar and read the 7-page <u>introduction to</u> <u>computation methods in Chinese studies</u>. These materials will help you to describe your goals.

The participants are expected to use their own laptops in the workshop and they should install the necessary software on their computers prior to the workshop. Programming scripts will be circulated in advance of the workshop. Instructions and assistance on this will be provided if needed.

REGISTRATION FOR THE WEBINARS

If you only want to attend the webinars, not the onsite training events, please register by filling out these zoom registration forms.

Joanne Kuai: Media innovation in China and its implications on contemporary China research

Thursday, April 28, 15:15-16:15 UTC+1

Registration link: <u>https://utu.zoom.us/meeting/register/u5Atd-CuqDIsGtawNdR0mvOKm6N5YLbH63K2</u>

Eric Zhang: Contemporary Chinese Digital Materials in the Big-data Era

Friday, April 29, at 13:00-14:00 UTC+1

Registration link:

https://utu.zoom.us/meeting/register/u5YscOyurjkiHNzUwV2e9UpQKqGrovRv4A1

TRAINERS

Christian Göbel is a university professor of Modern China Studies at the University of Vienna, Austria. His research is concerned with the adaptability of the Chinese party-state to social, economic, and political challenges. He is currently leading the research project "The microfoundations of authoritarian responsiveness" funded by the European Research Council, which investigates intended and unintended consequences of enhanced e-participation in China. He is the author of The Politics of Rural Reform in China (Routledge 2010) and The Politics of Community Building in Urban China (Routledge 2011, with Thomas Heberer). Website: https://autocrats.online. Email: Christian.Goebel@univie.ac.at

Jun Liu is an associate professor in the Department of Communication at the University of Copenhagen and Research Affiliate in the Center on Digital Culture and Society, the Annenberg School for Communication, University of Pennsylvania. His research covers political sociology and communication technologies and publishes in the fields of sociology, political science, communication and computer science. He is currently leading the Danish Independent Research Fund's Research Leader 2021 "To Use or Not to Use? A Relational Approach to Information and Communication Technologies as Repertoire of Contention" (2022-2025, <u>http://tech-in-movement.ku.dk/</u>. He has won several awards from American Political Science Association and the International Communication Association. His latest monograph is *Shifting Dynamics of Contention in the Digital Age* (Oxford University Press, 2020).

WEBINAR SPEAKERS

Joanne Kuai is a lecturer and a PhD Candidate at Karlstad University, Sweden, with a research project on Artificial Intelligence in Chinese Newsrooms. She is a member of the Ander Center for Research on News and Opinion in the Digital Era (NODE) and Geomedia Research Group at Karlstad University and an affiliated PhD student at the Nordic Institute of Asian Studies at the University of Copenhagen. Her research interests centre around data and Al for media, computational journalism, and the social implications of automation and algorithms. Prior to her academic career, she had worked as a reporter, editor and news anchor in China for over a decade.

Eric Zhang is a Junior researcher at LeidenAsiaCentre. He has a background in Political Science and International Relations. His current work mainly involves research on topics, including China's economic footprint in Europe, the impact of digital technologies, and cybersecurity, with regional focuses on China and Eastern Europe. Eric has a particular interest in incorporating computational methods in policy-relevant academic research.

FOR MORE INFORMATION CONTACT

Outi Luova <u>outi.luova@utu.fi</u> University lecturer, Centre for East Asian Studies CEAS, University of Turku, Finland





