GHENT
UNIVERSITY

RECONNECT

D3.5 FIRST PROGRESS REPORT ON RESEARCH CLUSTERS WP3

Project Information	
Acronym	ReConnect China
Title	ReConnect China: generating independent knowledge for a resilient future with China for Europe and its citizens.
Project no.	101061483
Type of Action	RIA
Deliverable Information	
Title	D3.5 First progress report on research clusters WP3
WP number and title	WP3, China's Domestic Governance
Authors	Christian Göbel
Description	Summary of the activities of WP3 in months 1-11
Lead Beneficiary	University of Vienna
Type	R: Document, report
Dissemination Level	PU: Public
Status	Final
History of Changes	
Draft version 0.0	Consultation with WP3 partners (July 2023)
0.1	First created by UNIVIE (20/09/2023)
0.2	Progress report by WP3 partners (20/09/2023)
0.3	Submitted by Christian Göbel (12/10/2023)
0.4	Final version approved by project coordinator (13/10/2023)
Version 1.0	Final version submitted as deliverable (13/10/2023)

Funded by
the European Union

1   LIST OF ABBREVIATIONS

Abbreviation	Meaning
DoA	Description of the Action – annex 1 to the GA
EC	European Commission
EU	European Union
WP	Work Package

2   EXECUTIVE SUMMARY

Work Package 3 (WP3) is structured into four pivotal tasks, each led by a distinct entity:

1. **New Sources of Information** (Lead: UNIVIE)

2. **Domestic Governance** (Lead: UPOL)

3. **International Relations** (Lead: UTURKU)

4. **Ideology and Values** (Lead: UGENT)

The success of Tasks 2-4 is significantly tethered to the data procured by Task 1, which has been the focal point of WP3 activities during the reporting period.

Organisation and Synergy Activities:

- The **kick-off meeting** in Gent and subsequent communications established and refined the division of tasks and planning of deliverables.

- **Synergy meetings** with DWARC explored collaborative opportunities, resulting in a joint academic article initiative exploring the establishment of CCP branches in private Chinese companies.

- **Zoom meetings** and continuous email communications ensured clarity and progress tracking of each task.

Task 3.1: Database - New Sources of Information (UNIVIE)

- Aims to trace political developments and public discourses across China, creating an interactive repository for analyzing policies, laws, regulations, and public discourses.

- The repository will also explore the relationship between citizen demands and public policies, containing data on social unrest, citizen complaints, and government responses.

Progress and Developments:

- **Infrastructure and Personnel:** A software engineer and developer were hired, alongside four research assistants, to design the infrastructure for data collection and analysis. Various technological enhancements and adaptations were implemented to optimize data collection and storage.

- **Data Collection:** Extensive data, including millions of Weibo posts, Tweets, and other online data from various sources, were collected to trace and analyze political developments and public discourses across China.

- **Data Classification:** Utilized ERNIE 3 and ChatGPT 3/4 to fine-tune models predicting protest-related content, issue areas in government complaints, and citizens' satisfaction with government responses. A Convolutional Neural Network was also trained to identify protest scenes in images.

- **Identifying Sources:** Identified and utilized various sources, including Weibo and Twitter accounts, websites, and databases, to collect data pertinent to each task and sub-task.

Key Achievements:

- **Data Collection:** Achieved the collection of several dozen million Weibo posts, several million Weibo and Twitter posts with specific keywords, and various other data from different sources.

- **Data Classification:** Successfully trained and fine-tuned models to predict and classify various data points, enhancing the analytical capabilities of the project.

- **Collaboration:** Established a collaborative initiative with DWARC members, contributing to an academic article exploring CCP branches in private Chinese companies.

Task 3.2: Domestic Governance (Lead: UPOL)

Explores intricacies of domestic politics during the Xi Jinping era

- explores intricacies of domestic politics during the Xi Jinping era
- focuses on the analysis of policies like the social credit system
- researches the dynamics of domestic narratives and their impact on China's international perceptions

Progress and Achievements:

- **Lecture Creation:** Developed a lecture on protests in China for the ReConnect MOOC (WP5).

- **Working Paper:** Drafted a working paper exploring discussions on China's social credit system on social media.

- **Policy Briefs:** Initiated planning for policy briefs on protests in China and discussions around the social credit system on social media.

- **Academic Contributions:** Submitted manuscripts for academic articles on social media use in Chinese protests and online protests following the Ürümqi apartment fire.

Task 3.3: International Relations (Lead: UTURKU)

- investigates the framing of the EU and other political entities by the Chinese propaganda system
- explores China's endeavors to influence citizen attitudes abroad, utilizing various datasets

Progress and Developments:

- **Data Collection:** Engaged in Twitter data collection to scrutinize Chinese propaganda.

- **Workshop Planning:** Initiated the planning of online workshops to further explore and discuss the findings and insights.

Task 3.4: Ideology and Values (Lead: UGENT)

- explores the pivotal "big ideas" and priority areas that steer policy-making in China
- examines the redefinition of concepts like human rights and democracy and the reinvention of history and identity by the CCP

Progress and Milestones:

- **Data Collection:** Engaged in data collection from official Weibo accounts using relevant keywords.

- **Manuscript Preparation:** Progressed in manuscript preparation for a synergy project with DWARC, exploring further into the ideological and value-based aspects of Chinese politics.

3   INTRODUCTION

In the modern geopolitical world, the EU and China stand in a constant dynamic relation with each other. The question in which domains EU collaboration with China is desirable, possible, or impossible is therefore the central focus of ReConnect China. To answer this question, this project identifies Science & Technology, Economy & Trade, Domestic Governance, and Foreign Policy as four key fields on which the EU should upgrade its independent knowledge. To that aim, it will develop a database which will provide access to new sources of information. The independent knowledge generated in this project will bring together stakeholders and experts in a Europe-China Knowledge Forum and raise awareness on China among the general public and youth. In this way, leading scientific European expertise and innovative digital technologies are mobilized to generate a deeper, sustainable understanding of contemporary China, and work towards a resilient future with China for the EU and its citizens. The project consortium brings together 15 partners from across Europe: Six universities (University of Ghent, University of Tartu, University of Turku, University of Groningen, University of Vienna and University of Olomouc), six think tanks (Egmont Institute, Clingendael Institute, Istituto Affari Internazionali, French Institute of International Relations, Swedish Institute of International Affairs, Elcano Royal Institute), one research institute (ZSI), one state agency (Innovation Norway) and one public institute under governmental administration (Macau Scientific and Cultural Centre). ReConnect China is coordinated by the University of Ghent, Belgium.

3.1   Deliverable scope

The key objective of WP3 is to shed light on the CCP's sources of legitimacy by tracing government narratives and popular discourses on the core values underlying CCP one-party rule, its state-building efforts and its relationship with the international community. Throughout the duration of the project, this will be achieved by:

- creating a comprehensive repository of policies, laws and regulations, government narratives and social media feeds;

- using state-of-the-art natural language processing and machine learning as well as field research to trace how government narratives and popular discourses influence each other;

- researching how the propaganda system attempts to influence public opinion at home and abroad;

- keeping track of core ideologies and values informing China's policies and international relations.

Based on the results WP3 will develop policy recommendations related to China's governance and political system, as well as international relations, ideology and values.

This deliverable provides a summary of the activities and progress of WP3 between November 2022 (start of the project) and September 2023 (end of the reporting period).

3.2   Reference documents

- ReConnect China Grant Agreement no. 101061483

- ReConnect China Consortium Agreement

- ReConnect China Project Management Handbook (D7.1)

4 PROGRESS OF WORK PACKAGE 3

WP3 consists of four tasks: 1) new sources of information (lead: UNIVIE); 2) Domestic governance (lead: UPOL); 3) International relations (lead: UTURKU); 4) Ideology and values (lead: UGENT). To a large extent, the deliverables of Tasks 2-4 depend on the data collected by Task 1. For this reason, most of the activity in WP3 during the reporting period pertains to this Task. The following sections will first describe the organization of the whole work package during the reporting period and then report the progress of each Task.

4.1 Organisation and synergy activities with DWARC

At the project's kick-off meeting in Gent on November 22 and 23, the lead beneficiary (UNIVIE) reiterated the division of tasks specified in the Grant Agreement with the task leaders and the WP members in attendance. Subsequently, the lead beneficiary and the other beneficiaries fine-tuned the planning of deliverables for each Task. In the week of 31 November, Christian Göbel, representative of the lead beneficiary, travelled to Olomouc to discuss the division of labour within the UPOL-team and the timing and mode of the planned fieldwork in China. Most communication between the lead beneficiary and the other WP3 members was conducted via email. On 20 September 2023, representatives of all WP3 members met on Zoom to discuss the progress of each Task, clarify the division of labour in the deliverables, and review upcoming tasks and deadlines.

Christian Göbel also participated in the synergy meeting with DWARC on 2023-02-12 to discuss potential areas of collaboration. UNIVIE and two DWARC members, Kjeld Erik Brødsgaard and Kasper Ingeman Beck decided to collaborate on an academic article on the establishment of Chinese Communist Party branches in listed private companies in China. The task of UNIVIE was obtaining, providing and modelling the data on amendments to the company charters. The group meet online on 2023-03-10, 2023-06-14 and 2023-07-24 to discuss the collected data, an introduction draft for the publication, and the preparation for fieldwork of the DWARC members.

The following sections describe the progress of each Task.

4.2 Task 3.1: Database - New sources of information (UNIVIE)

Task 3.1 will provide the information needed to trace political developments not only in Beijing, but also in China's provinces, municipalities and counties, to map public discourses on these topics, and to study how narratives, discourses and policies influence each other. All this is necessary to obtain a more nuanced and up-to-date image of current political developments in China. To this

end, we will create an interactive and up-to-date repository for the analysis of policies, laws and regulations, government narratives and public discourses which will be essential not only to the other tasks of this WP, but also for the other WPs of this project. To examine the relationship between citizen demands and public policies, the repository will also contain updated information on social unrest, citizen complaints to the government and official responses to public complaints.

In the reporting period, task leader UNIVIE hired and trained a team of coders, prepared the infrastructure for data collection, and collected data for Tasks 3.2 and 3.4.

4.2.1 Personnel and infrastructure

- Hired a software engineer to design the infrastructure for data collection and -analysis. The project uses the hardware obtained for the ERC project RESPONSIVENESS and recycles as much of the software created for that project as possible. Improvements have been made with regards to data collection and storage, in particular:
  - o setting up a High Availability Kybernetes cluster to maintain continuous, uninterrupted service and to minimize downtime;
  - o use a non-relational database (MongoDB) for data storage instead of PostgresQL, as a non-relational database is much more appropriate for storing heterogeneous data;
  - o create a GitLab environment to facilitate automised data collection. One yaml file per province and prefecture records the links for the primary data of interest; different types of web scrapers automatically process this information for low-effort data collection;
  - o use GitLab pipelines to enforce data integrity through automated testing and quality checks, ensuring consistency and integrity in the data collected and stored, and to quickly identify changes in website structures.
- Hired a software developer
- Hired four research assistants
- Trained four research assistants in Python
- Trained four research assistants in webscraping (urllib, BeautifulSoup, Scrapy)
- Wrote an easy-to-configure scraper to collect Weibo data based on keywords. New keywords can be deployed within minutes to capture discourses on unexpected events

4.2.2 Identifying sources for data collection

- More than 20.000 official Weibo accounts operated by government units and news organisations, of which more than 3.000 Weibo accounts of official and non-official news organisations (for Task 3.2)
- List of all province- and prefecture-level units of China (for all tasks)
- Centralised website for citizen complaints to the government (for Task 3.2)
- More than 300 Twitter accounts of Chinese news organisations, embassies, and diplomats (for Task 3.3)
- Websites of the Science and Technology Bureaus of Chengdu and Chongqing (for Task 3.2)
- Website storing amendments to charters of listed companies (for DWARC synergy project)

4.2.3 Data collection

- Collected several dozen million Weibo posts by official accounts
- Collected several million Weibo posts containing the keywords (in Chinese) "USA", "Europe", "Ukraine", "Russia", "social credit", "human rights", "democracy", "freedom"
- Began collecting links and writing webscrapers for a database of local laws, regulations, leaders' activities, and news
- Collected several million Tweets by Chinese news organisations, embassies, and diplomats as well as reactions to these tweets up to June 2023, when the educational API was discontinued
- Collected several hundred news items from the Chongqing and Chengdu websites
- Collected several million Weibo posts that contain keywords which appear in protest-related posts with high frequency
- Collected several thousand company amendments
- Collected information on career paths of several thousand leading politicians

4.2.4 Data classification

- Fine-tuned Large Language Model (ERNIE 3) to predict protest-related content in Weibo posts
- Fine-tuned Large Language Model (ERNIE 3) to predict issue area in complaints to the government
- Fine-tuned Large Language Model (ERNIE 3) to predict citizens' satisfaction with replies by government departments to their complaints
- Trained Convolutional Neural Network to predict if an image contains a protest scene

- Used ChatGPT 3 and 4 to improve trainingset for protest detection

4.3 Task 3.2: Domestic governance (lead: UPOL)

Task 3.2 is concerned with the continuities and changes in domestic politics during the Xi Jinping era and analyses complex, contested and, in the case of the social credit system, often misrepresented policies. It will draw on a mixture of digital sources and field research in order to provide insight into the strengths and weaknesses of either research technique. In highlighting the applied potential of such technologies, the latter aspect complements the analysis done in Task 1.3. Furthermore, domestic narratives impact China's perceptions of other countries and regions; in that regard, this task links up with Task 4.1.

4.3.1 Progress

- Created lecture on protests in China for ReConnect MOOC (WP5)
- Drafted working paper on how China's social credit system is discussed on social media (together with Christoph Steinhardt)
- Planned policy brief on protests in China (based on more than 80.000 events identified with the help of Weibo data)
- Planned policy brief on how China's social credit system is discussed on social media
- Submitted academic article manuscript on social media use in Chinese protests
- Submitted academic article manuscript on online protests in the aftermath of the Ürümqi apartment fire on November 24, 2022, which served as a catalyst to the White Paper Protests that began two days later.

4.4 Task 3.3: International relations (lead: UTURKU)

Task 3.3 is concerned with how the EU and other political entities are framed by the Chinese propaganda system, and how China attempts to shape citizen attitudes abroad. It will draw on existing data sets, for example MoFA's English language media communication collected by UPOL, datasets curated by NGOs as well as (offical) news items and social media feeds.

4.4.1 Progress

- Collected Twitter data to explore Chinese propaganda (see 4.2.3)
- Planning of online workshops in progress

4.5 Task 4.4: Ideology and values (lead: UGENT)

Task 3.4 is concerned with the "big ideas" and priority areas that guide and inform policy making in China. Although ideology is often dismissed as unimportant and inconsequential, it is an essential element of Chinese politics. Understanding how concepts such as human rights and democracy are redefined, and how history and identity are reinvented by the CCP will aid stake holders in their communication efforts by sensitising them to the fact that the same concept might have different meanings and thereby complement existing, valuable efforts at "decoding" China. Task members will mainly draw on data produced in T3.1 and interviews that uncover psychological mechanisms behind nationalist sentiment among the Chinese youth, for example how they cope with contradictory historical/political information related to e.g. human rights after having been socialised into an official history discourse

4.5.1 Progress

- Collected data from official Weibo accounts containing related keywords
- Manuscript preparation for synergy project with DWARC in progress

5   CONCLUSIONS

WP3 is devoted to a comprehensive exploration of various facets of Chinese politics, from domestic governance and international relations to ideology and values. During the reporting period, the work has progressed according to plan and in line with the Grant Agreement. Notable achievements were made in data collection, the preparation of academic contributions, a lecture on protests in China, and laying the groundwork for outputs like policy briefs and workshops. The collection of prefecture-level data is progressing slower as expected due to the need to train coders, but also the complexity of many of the local government websites, which make collecting the relevant URL's an arduous task. We hope to hire more coders to speed up data collection. Another achievement pertains to the successful collaboration with the sister project DWARC, which will result in an academic article manuscript during the next monitoring period. Moving forward, WP3 will continue analysis and exploration of the collected data and thereby contribute to the overarching objectives of understanding and decoding the political, social, and international dynamics of China.