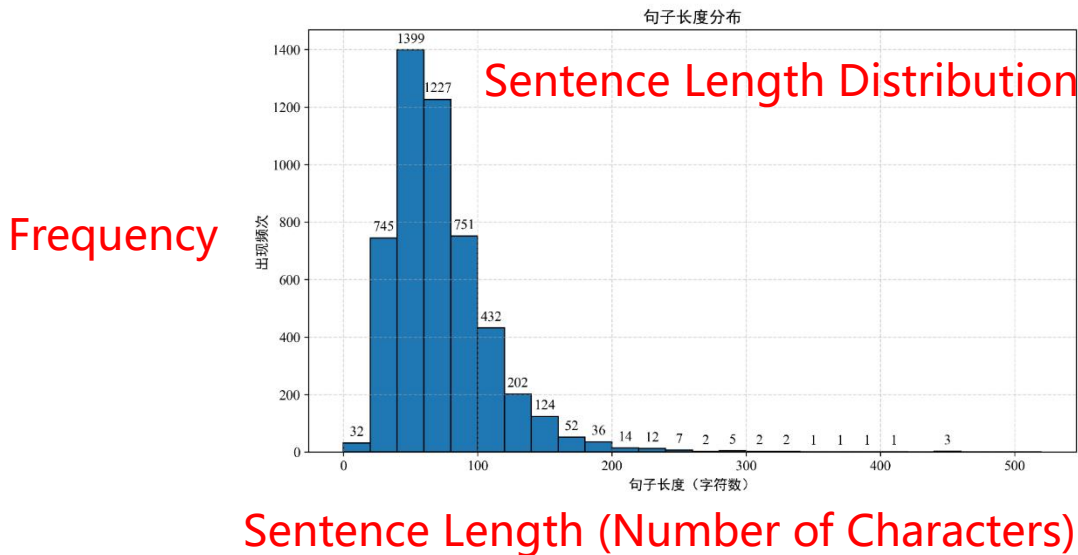# Annotation Guideline

目 录 CONTENTS

Project Overview

1.1 Project basic information

- **Project Name:** Relation extraction project of civil aircraft supply chain.

- **Project Object:** To establish a standard Chinese corpus dataset of the civil aircraft supply chain, supporting downstream relation task.

- **Dataset Size:** 5,051 news sentences.

- **Project Significance:** The civil aircraft industry features long supply chains, low visibility, and intricate systemic risks. Understanding the civil aircraft supply chain network is of great practical significance for improving supply chain risk management, promoting digital management, and enhancing.

■ **Data Characteristics:** Most sentences are short, with fewer than 100 Chinese characters, making them suitable for both manual and large language model annotation.

Sentence Length Distribution

Frequency

Sentence Length (Number of Characters)

1.2 Project Design

- This project adopts a hybrid annotation approach that combines manual and large language model (LLM)-based annotation.

- **Manual Annotation Scheme (Construction of the Human Standard Dataset):** Three annotators independently label the data, and one project leader serves as the adjudicator. The dataset is considered valid only after both intra-annotator and inter-annotator consistency checks have been passed.

1.2 Project Design

- **LLM Annotation Scheme (Construction of the LLM Standard Dataset):** The Qwen 3 large language model is fine-tuned on the manually annotated standard dataset. 5 fine-tuned Qwen 3 models are then integrated into an ensemble, which retains high-confidence samples to expand the standard dataset.

- **Sample Allocation:** Based on sentence length as an indicator of annotation difficulty, 2,000 samples are assigned for manual annotation, while the remaining sentences serve as candidate data for LLM-based annotation.

1.3 Entity Categories

1. **Organization**

2. **Product Model**

3. **Product/Service**

1.4 Relation Categories

1. **Supply**

2. **Produce product/Provide service**

3. **Cooperate/Propose to cooperate**

4. **Hold stake in/Own/Acquire/Invest in**

5. **Coreference**

6. **Others/None**

1.5 Project Notes

1. **Sentence-level semantic annotation:** Relationship annotation should be determined solely based on the semantics within the sentence itself. Annotators must not rely on context from preceding or following sentences, nor on any external knowledge.

2. **Limitation on the "Other/None" category:** The proportion of "Other/None" cases for each annotator must not exceed 5% of all annotated samples. Submissions exceeding this limit will be returned for re-annotation.

3. **Scope of annotated sentences:** All sentences have undergone two rounds of manual screening, and most contain explicit supply chain relationships. While the majority of organizational entities are from the civil aircraft supply chain, entities from other industries may also appear and should still be labeled as Organization.

Entity Categories

Entity Categories

2.1 Organization

Refers to enterprises within the civil aircraft supply chain, including manufacturers, core system and equipment suppliers, material and standard part suppliers, and downstream airline operators. It also covers universities and research institutes engaged in technology R&D, as well as government agencies providing services such as airworthiness certifications.

Entity Categories

2.2 Product Model

Refers to specific product names, including complete aircraft models such as A320, B737MAX, and C919, as well as military aircraft products produced by supply chain enterprises. In addition, other products closely related to the civil aircraft sector—such as automotive models—are also included in this entity category. This entity type enables relationship inference by tracing back to its manufacturer, thereby improving the recall rate of "Organization–Organization" relationships in the extracted triples.

Entity Categories

2.3 Product/Service

Refers to specific civil aircraft subsystems or structural components and all related products, such as avionics systems, fuselage, nose section, aircraft seats, high-temperature alloys, and raw materials. It also includes relevant services such as maintenance and airworthiness certification, which play key roles in the civil aircraft lifecycle.

PART THREE

Relation Categories

Relation Categories

3.1 Supply

Refers to the business flow relationship between two organizations, or between an organization and a product (including both Product Model and Product/Service entities) within the civil aircraft supply chain.

Relation Categories

3.2 Produce product/Provide service

Indicates that one organization supplies a specific or general product or service to another organization within the supply chain. The expression may be explicit or implicit in the text.

Relation Categories

3.3 Cooperate/Propose to cooperate

Denotes existing or intended cooperative relationships between organizations. When the direction of the supply relationship cannot be clearly determined, such cases are classified under this category.

Relation Categories

3.4 Hold stake in/Own/Acquire/Invest in

Represents a capital or control relationship between two organizations, where the former exerts certain ownership or decision-making influence over the latter. Specifically, expressions such as "A is B's shareholder", "A is the parent company of B", "A acquired B's business", or "A invested in B" all fall under this category.

Relation Categories

3.5 Coreference

Used to identify different expressions referring to the same entity within the text,

enabling entity disambiguation and linkage across sentences or documents.

Relation Categories

3.6 Others/None

Covers both (a) relationships that cannot be categorized into any of the five types above, and (b) cases where no supply chain relationship exists in the text. However, this relation type should be limited to less than 5% of all annotated data to ensure labeling precision.

Domain Dictionary

4.1 Organization

- **Civil aircraft manufacturer:** 中国商用飞机有限责任公司（或中国商飞或comac）、波音（boeing）、空中客车（或空客或airbus）、俄罗斯联合航空制造集团、庞巴迪 et al.

- **Engine manufacturer:** 中国航发 et al.

- **Airlines:** 中国国际航空公司（或国航）、海航集团有限公司（或海航）、中国南方航空公司（或南航）、东航、泰国都市航空 et al.

- **Leasing companies:** 中银航空租赁、浦银租赁等租赁公司 et al.

- **Governments:** 中国民航局（或民航局或caac）、欧洲航空安全局（easa）、美国联邦航空局（faa） et al.

......

4.2 Product Model

- **Civil aircraft (China): 中国商飞c系列，c909（或909或arj21）、c919（或919）、c929（或929）、c939（或939）、 ma60 （或新舟60、y-7-200a，中航工业旗下公司研制）、ma600 et al.**

- **Civil aircraft (Other countries): a320、a220、b737、b787 、c系列（庞巴迪公司生产）、 erj190系列（巴西航空工业产）** et al.

- **Military aircraft: 美国军机f22、歼-11、歼-15、歼-16** et al.

- **Engine: leap、cj-1000（系列发动机）或长江-1000（系列发动机） et al.**

- **Automotive: 奔驰大g越野车** et al.

......

4.3 Product/Service

- **Aircraft core systems:** 动力系统（发动机系统、辅助动力装置[apu]）、飞行控制系统、航电系统（导航系统、通信系统、监视系统、飞行仪表显示系统、任务系统）、机电系统、电源系统、液压系统、燃油系统、气动系统、环境控制系统、起落架系统、防冰除冰系统、机载信息系统、火警与监测系统 et al.

- **Aircraft structural systems:** 机体、机头、前中后机身、机翼、尾翼、起落架、舱门 et al.

- **Aircraft subsystems or modules:** 传感器、碳刹车盘等、液压泵、电源管理单元 et al.

- **Aircraft materials and standard parts:** 铝合金、钛合金、复合材料、标准紧固件、铆钉、电线线缆 et al.

......

4.3 Product/Service

- **Airworthiness certification:** 适航认证、适航取证 et al.

- **XXX solutions:** 如电传飞控解决方案、发动机解决方案 et al.

- **Maintenance-related:** 维护、维修、大修（或MRO） et al.

......

Annotation Examples

5.1 Supply

outside of China, the only airline operating **comac** aircraft is the Indonesian low-cost carrier **transnusa**, which operates the **arj21** regional jet

Triple

comac	Supply	transnusa
arj21	Supply	**transnusa**
comac	Produce product/ Provide service	arj21

5.2 Produce product/Provide service

safran electronics & defense provides its main systems and equipment, including **cockpit control systems**, **control panels and dimming controls**, **throttle quadrant assemblies**, and **horizontal stabilizer trim actuators**.

Triple

safran electronics & defense	Produce product/ Provide service	**cockpit control systems**
safran electronics & defense	Produce product/ Provide service	**control panels and dimming controls**
safran electronics & defense	Produce product/ Provide service	**throttle quadrant assemblies**
safran electronics & defense	Produce product/ Provide service	**horizontal stabilizer trim actuators**

5.2 Produce product/Provide service

they are strengthening product performance and production process requirements, striving to obtain **airworthiness certification** from **the federal aviation administration** (**faa**) and **the european union aviation safety agency** (**easa**).

Triple

the federal aviation administration	Produce product/ Provide service	**airworthiness certification**
the federal aviation administration	Produce product/ Provide service	**airworthiness certification**
the federal aviation administration	Coreference	faa
the european union aviation safety agency	Coreference	easa

5.3 Cooperate/Propose to cooperate

according to paul almada, managing partner of **transnusa airlines**, the company has been in negotiations with **comac** for several months and will visit china in october to discuss a potential order of up to four **c919** aircraft.

Triple

transnusa airlines	Cooperate/Propose to cooperate	comac
comac	Produce product/ Provide Service	c919

5.4 Hold stake in/Own/Acquire/Invest in

five months prior to the acquisition, **ant financial** announced the completion of its series b financing, led by **china investment corporation** and **jianxin trust**, a subsidiary of **china construction bank**.

Triple

jianxin trust	Hold stake in/Own/Acquire/Invest in	**ant financial**
china investment corporation	Hold stake in/Own/Acquire/Invest in	**ant financial**
china construction bank	Hold stake in/Own/Acquire/Invest in	**jianxin trust**

5.5 Coreference

before the start of the china international aviation & aerospace exhibition, **craic** (**china-russia commercial aircraft international co., ltd.**), a china-russia joint venture, was conducting the selection of its key suppliers.

Triple

china-russia commercial aircraft international co., ltd.	Coreference	craic

5.6 Others/None

This is because the **c919** needs to expand overseas and compete with **airbus** and **boeing**, adopting a globalized supply chain to share profits, which is conducive to market expansion.

Triple

c919	Others/None	airbus
c919	Others/None	boeing

Annotation Examples

5.7 Annotation Tool Selection

- This project employs Doccano for the joint annotation of entities and relations.