|  |  |  |  |
| --- | --- | --- | --- |
| **Inventory of supplemental materials** | | | |
| File name | Specific content | | Corresponding parts in the manuscript |
| CCLR knowledge graph | Chroma-formatted vectorized CCLR knowledge graph | | Section 3.1 (Fig. 3) |
| Neo4j-readable CCLR knowledge graph dump | | Section 3.1 (Fig. 3) |
| Json format CCLR knowledge graph | | Section 3.1 (Fig. 3) |
| Codes | Codes for original large language models | Codes for original text-davinci-003, GPT-3.5-turbo, GPT-4 | Section 3.2 |
| Codes for original ERNIE-Bot 4.0, Qianfan-Chinese-Llama-2-7B and Llama-2-70b | Section 3.2 |
| Codes for original ChatGLM2-6B and ChatGLM3-6B | Section 3.2 |
| Codes for large language models integrating knowledge graph | Codes for text-davinci-003, GPT-3.5-turbo, GPT-4 with KG | Section 3.2 |
| Codes for ERNIE-Bot 4.0, Qianfan-Chinese-Llama-2-7B and Llama-2-70b with KG | Section 3.2 |
| Codes for ChatGLM2-6B and ChatGLM3-6B with KG | Section 3.2 |
| Video clips | Video clips of running original large language models code | | Section 3.3 |
| Video clips of running large language models with knowledge graph code | | Section 3.3 |
| Table S1 Experts for developing CCLR knowledge graph and QA dataset | Experts for confirming the rationality of schema layer and completeness of laws and regulations in data layer | | Section 3.1 |
| Experts for formulating the self-designed questions | | Section 3.3 |
| Table S2 The data layer of CCLR knowledge graph | 1 first-tier knowledge domain | | Section 3.1 (Fig. 2) |
| 8 second-tier knowledge domains | | Section 3.1 (Fig. 2) |
| 193 third-tier knowledge domains | | Section 3.1 (Fig. 2) |
| 276 laws and regulations | | Section 3.1 (Fig. 2) |
| 8 second-tier knowledge domains and 193 third-tier knowledge domains form 193 triplets | | Section 3.1 (Fig. 2) |
| 193 third-tier knowledge domains and 276 CCLR-related laws and regulations form 657 triples | | Section 3.1 (Fig. 2) |
| 8 clauses of LR012 Construction Law of the People’s Republic of China | | Section 3.1 (Fig. 2) |
| Table S3 QA dataset containing 6,339 questions | Dataset for testing large language models | | Section 3.3 (Fig. 5) |
| Table S4 The answers of each question from 8 LLMs before and after the integration of CCLR knowledge graph | Answers of each question from 8 LLMs before and after the integration of CCLR knowledge graph | | Section 4 |