基于文档数据的“问题-答案”数据标注准则

**10组待标注文件**：

* BianChengZhiMei.DBQA.Part1
* BianChengZhiMei.DBQA.Part2
* BianChengZhiMei.DBQA.Part3
* BianChengZhiMei.DBQA.Part4
* BianChengZhiMei.DBQA.Part5
* BianChengZhiMei.DBQA.Part6
* BianChengZhiMei.DBQA.Part7
* BianChengZhiMei.DBQA.Part8
* BianChengZhiMei.DBQA.Part9
* BianChengZhiMei.DBQA.Part10

**数据格式**：

[TITLE]:中国国家图书馆

[QUESTION]:中国国家图书馆在哪里？

**1** 中国国家图书馆是中华人民共和国的国家图书馆，位于北京市海淀区白石桥畔中关村南大街33号[1]。

0 中国国家图书馆的前身是于1909年9月9日始建的京师图书馆。[2]

0 1909年，清政府推行“新政”，建立西式的文化设施，由学部主持在什刹海北岸鸦儿胡同的广化寺筹备建立京师图书馆。

0 但由于清末政局不稳，难以重视文化事业，因此直到清帝逊位之日，京师图书馆始终没有正式接待过读者。

0 中华民国成立后，北洋政府的教育部接管了京师图书馆，并将其改名为国立京师图书馆，1912年（民国元年）8月27日正式开馆接待读者；1928年北伐成功后改名国立北平图书馆，并迁址到中南海居仁堂。

0 1931年，北平图书馆新馆（今为中国国家图书馆古籍馆所在地）在文津街落成，是当时远东最先进的图书馆之一。

0 当时的北平图书馆下设二会八部，分别为：购书委员会、编纂委员会、总务部、采访部、编纂部、阅览部、善本部、金石部、舆图部、期刊部，馆藏书籍包括中文、满文、蒙文、日文、西文普通书籍逾四十万册；宋元明清刊本、写本、抄本古籍逾三万册；金石拓片逾八千幅；舆图八千余幅；此外还有寄存图书六千余种三万余册。

0 其中包括了文津阁《四库全书》等珍贵收藏。

0 1942年1月2日，华北政务委员会教育总署接管国立北平图书馆，并更名为国立北京图书馆。

0 1945年抗日战争结束后，该馆复名国立北平图书馆。

0 中华人民共和国成立后，于1950年3月6日国立北平图书馆更名为国立北京图书馆，1951年6月12日更名为北京图书馆，成为中华人民共和国唯一的国家图书馆。

0 1987年，北京图书馆新馆在白石桥路（今中关村南大街）落成，该建筑还曾被评为“八十年代北京十大建筑”榜首，馆名由邓小平题写。

0 1998年12月12日北京图书馆更名为国家图书馆，对外称中国国家图书馆，1999年4月16日由当时的中共中央总书记、国家主席江泽民题写馆名。[2]

0 中国国家图书馆的主要组成机构有：

0 中国国家图书馆总馆南区

0 中国国家图书馆总馆北区

0 中国国家图书馆古籍馆

0 中国国家图书馆少年儿童馆

给定一篇文章，标注问题时请

1. 随机选取一个句子作为答案句子，并将前面的数字改写成1。比如选择上例中的标黄的句子，前面的数字写成1；
2. 针对该选取的句子生成一个问题，并将其填写到[Question]后面。比如，根据选取的答案句子可以生成一个问题：中国国家图书馆在哪里？。这个问题可以针对答案句子里面的任意知识点生成。
3. 检查整篇文章的其他句子能否回答生成的问题。如果能，也把它们前面的数字置成1；如果不能，则不对前面的数字做操作。比如，上例中其他句子都不能回答生成的问题，那么不需要更改前面的数字。

**标注注意事项**：

* 10个文件至少**请5个人**进行标注，每人标注2个或1个文件。
* 尽量使用句子中并未包含的**表述形式**生成问题。比如，某句子“贝加尔湖面积大约3.15万平方公里”，不要采用“贝加尔湖的面积是多少？”这种形式提问，可以采用“贝加尔湖有多大？”这种提问形式。
* 不要总是选择固定的句子序号生成句子（比如总是根据第一句话生成问题），尽量随机选择答案句子。
* 如果某个文档对应的句子太少、或很难生成问题，请在Question后面填写-1。
* 涉及到政治敏感的材料，请在Question后面填写-2。