# 4.3.3 Map reduce 链

在大数据和自然语言处理领域，各种不同的处理链方式可以用来优化信息检索和生成答案。本文将解析 ‘MapReduce’ 类型的处理链，并说明其如何通过改变输入的组织和输出的生成方式来提高文档搜索的质量。

#### ##4 1. 整体流程

‘MapReduce’ 文档处理链主要由两个部分组成：映射（Map）阶段和归约（Reduce）阶段。在映射阶段，系统对每个文档单独应用一个语言模型（LLM）链，并将链输出视为新的文档。在归约阶段，系统将所有新文档传递给一个单独的合并文档链，以获得单一的输出。如果需要，系统会首先压缩或合并映射的文档，以确保它们适合合并文档链。

#### ##4 2. 映射阶段（Map Stage）

在映射阶段，系统使用 LLM 链对每个输入的文档进行处理。处理的方式是，将当前文档作为输入传递给 LLM 链，然后将 LLM 链的输出视为新的文档。这样，每个文档都会被转化为一个新的文档，这个新文档包含了原始文档的处理结果。

对于每个文档，作为提示(Prompt)传递给 LLM 的内容是原始文档。比起“ Stuff ”类型多了预处理。

每个原始文档都经过LLM 链处理的结果写入一个新文档，这就是映射的过程。比如原文档有2000字，经过LLM 链处理的结果是200字。200字的结果存储为一个新文档，但是跟2000字原文档存着映射关系。

#### ##4 3. 归约阶段（Reduce Stage）

在归约阶段，系统使用合并文档链将映射阶段得到的所有新文档合并成一个。如果新文档的总长度超过了合并文档链的容量，那么系统会使用一个压缩过程将新文档的数量减少到合适的数量。这个压缩过程会递归进行，直到新文档的总长度满足要求。

#### ##4 4. 最终实现效果

通过 ‘MapReduce’ 文档处理链，系统可以对每个文档单独进行处理，然后将所有文档的处理结果合并在一起。这种处理方式可以提高文档搜索的质量，特别是在处理大量文档的情况下。

#### ##4 5. 适用场景

‘MapReduce’ 类型的处理链方式主要适用于处理大量文档的情况，特别是当这些文档不能全部放入模型的上下文中时。通过并行处理每个文档并合并处理结果，这种处理方式可以在有限的资源下处理大量的文档。然而，这种处理方式可能会使用更多的计算资源，并且可能在处理某些复杂任务（如文档之间频繁地交叉引用，或者需要从许多文档中获取详细信息）时可能表现不佳。

总的来说，通过使用 ‘MapReduce’ 文档处理链，系统可以有效地处理大量文档的情况，从而提高文档搜索的质量。然而，这种处理方式可能需要更多的计算资源，并且可能在处理复杂任务时表现不佳。