**2020.3月总结**

月报主要分为两部分，第一部分为对3月工作的总结，第二部分为对4月工作的计划。

## 1.3月工作总结

本月工作主要分为五部分：

1.关于视觉实体链接相关论文的阅读与模型的改进

2.关于IJCAI会议论文投稿的跟进工作

3.对于JBDR期刊论文的工作推进和论文撰写

4.实验室项目申请书的撰写和PPT的制作

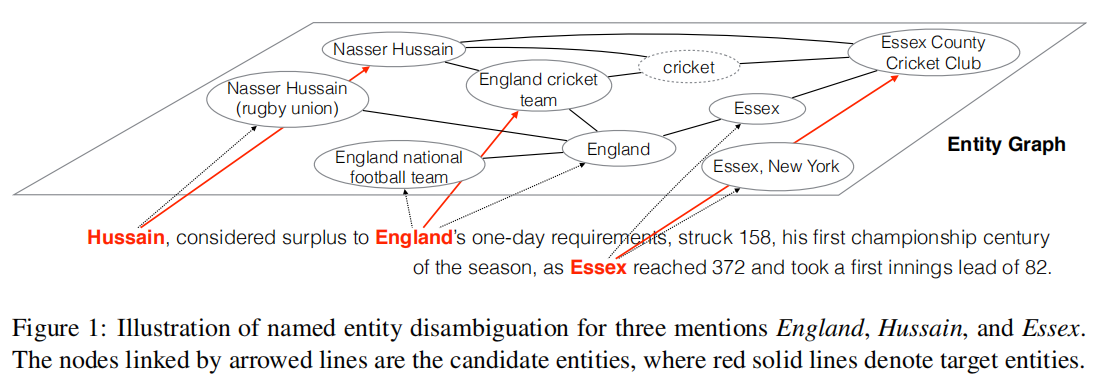
5.关于多模态知识平台的构想和升级工作

下面五项工作的成果和进展进行详细叙述。

1. **视觉实体链接相关论文的阅读**
2. **《Neural Collective Entity Linking》**

这是刘知远老师组在2018年末发表的一篇论文，主要思想是利用GCN将局部上下文特征和全局语义一致性结合起来进行实体链接，对相邻实体提及的子图进行近似的图卷积来获取最终结果。

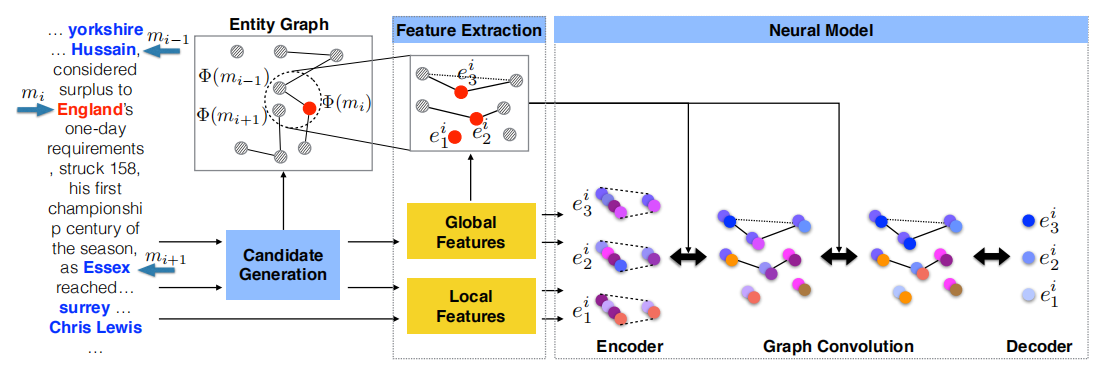
现有的实体链接主要依赖于本地上下文（提及的相近tokens）独立进行解析，而忽略了全局特征，即文档中所有提及的目标实体在主题上应该保持一致性，具体示例见下图：



如图所示，传统EL模型不能从周围的词中找到足够的消歧线索，但是利用全局相关性则会得到最终的正确结果。此文提出了一个 Neural Collective Entity Linking model (NCEL)模型，将深度神经网络与图卷积网络(GCN)结合进行全局EL，允许对实体图进行灵活编码。集成了局部上下文信息和文档中提及的全局相互依赖性，并且可以以end-to-end方式有效地进行培训。同时将注意机制引入到局部上下文信息的鲁棒模型中，通过选择信息词和过滤噪声来实现，利用GCNs来改善候选实体的判别信号，利用正确实体下的丰富结构。为了减少全局计算，对相邻提及的子图进行卷积，整体的连贯性通过文档上的滑动窗口以链状的方式实现。

这篇文章EL的实体候选列表生成是通过大型的语料库来计算先验概率，得到一个预定义的“提及--实体”字典来表示候选列表，同时作为其局部特征应用于模型，但是这种方法的问题就是如果在一个非限定域的条件下进行实体链接，即得到的提及不曾出现于训练数据中，应该如何去生成候选实体列表？所以此模型的泛化能力仍需进行讨论。

主要的模型结构如下图：



结构也比较容易看懂，大体分为三个部分。一，候选实体列表生成，主要是依赖一个离线的训练数据得到；二，特征抽取，主要为两方面：局部特征和全局特征；三，神经网络模型，利用GCN来得到每个提及的最大概率实体。

对于我们工作的借鉴意义我认为主要在全局特征的应用上，这种一段话中所有实体的连贯相似性也可以应用于我们的之后的模型上面。

1. **《****Collective Entity Linking in Web Text: A Graph-Based Method》**

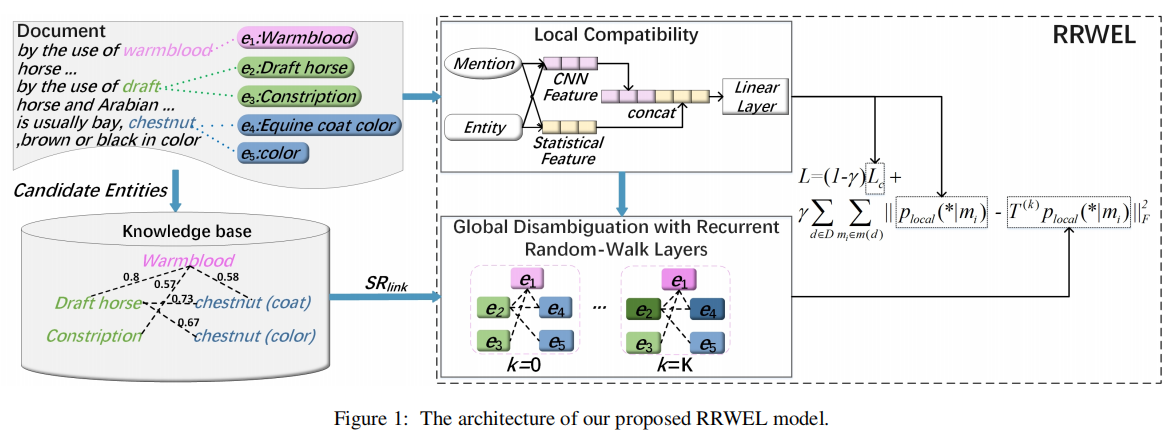
这是赵军老师组2011年SIGIR的一篇实体链接论文，是通过第一篇的参考文献查到相关全局实体链接的文章。此文是一个比较早的实体链接任务的研究，我认为本文主要的贡献有提出Referent Graph，Collective Entity Linking和利用Referent Graph推断出实体链接关系。具体的Collective Entity Linking已在上一篇叙述，而且这篇文章使用的一些计算相容性的算法和网络结构都在最近有了较大的更新和提升，主要的借鉴意义在于本文中提出的利用Wikipedia去计算语义相关性的方法。

IMG_256

1. **《Neural Collective Entity Linking Based on Recurrent Random Walk Network Learning》**

此文是2019年IJCAI中科大的一篇关于全局实体链接的文章，与上述论文的主要区别是使用一个堆叠随机游走层以加强证据使得相关的 EL成为高概率决策的方法来进行实体链接，其中候选实体之间的语义相互依赖性主要是从一个外部知识库引导。在传统的目标函数中引入一个语义规则器，保持集体 EL决策的一致性，从而使外部 EL 决策具有一致性。

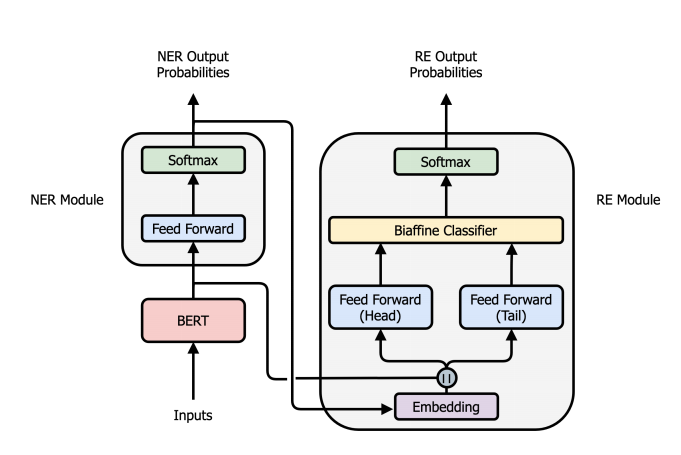
论文模型如下：



本文从另外一个角度解决Collective Entity Linking，其中的SRlink还是通过第二篇文章的方法进行计算，利用随机游走策略替代之前的多头注意力模型。

1. **《END-TO-END NAMED ENTITY RECOGNITION AND RELATION EXTRACTION USING PRE-TRAINED LANGUAGE MODELS》**

现有的实体链接极大程度的依赖于外部NLP工具的性能，此文提出利用预先训练的语言模型联合实体抽取和关系抽取。模型如下图所示：



NER部分由BERT实现：1.BERT模型产生序列，2.送入全连接层并计算交叉熵损失，其中标签由BIOES表示。RE部分：1.预测的标签被embedding，产生序列，2.concat操作，3.按照Nguyen和Verspoor（2019）提出的深双线性注意机制进行分类，4.为RE目标计算第二个交叉熵损失，5端到端训练减少总loss。最终的损失函数是NER部分的损失函数和RE部分的损失函数之和。

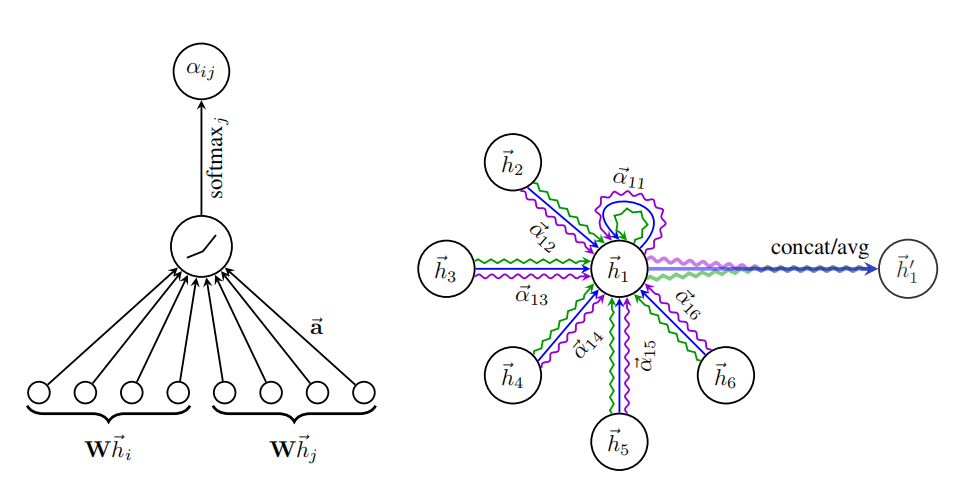
有一点创新的是，因为NER的表现对整个模型非常重要，因此早期先训练NER，延迟训练RE。没有将RE模块的训练延迟一些时间，而是权衡RE在训练的第一个时间段对总损失的贡献，其中λ在第一个epoch从0线性增加到1，其余时期设置为1，因为NER模块可快速为所有数据集（即在一个epoch内）取得良好的性能。 在早期的实验中，发现该方案的性能优于完整时期的延迟。

对我们的借鉴意义在于之后可以考虑一个联合多种模态，联合实体链接和关系抽取的双重系统，对两项任务进行联合训练从而达到更好的效果。

1. **《Graph Attention Networks》**

这是一篇2018年ICLR的文章，将attention机制应用到graph convolution中的文章。但是文章中提出的模型其实是利用了attention的一部分思想，而不是和处理sequence的模型应用的attention机制是不完全一样的。是一篇比较经典的模型创新文章。

处理sequence的模型引入的attention机制可以分为两类，一类是在输入sequence本身上计算attention的intertext情形；另一类是另外有一个用来计算attention的文本的intertext情形 。本文中的情形是intratext，是在输入的待处理的graph自身上计算attention，在摘要中作者说明了，本文的attention计算的目的是为每个节点neighborhood中的节点分配不同的权重，也就是attention是用来关注那些作用比较大的节点，而忽视一些作用较小的节点。相比之下，在sequence的模型中，attention的计算是为了在处理局部信息的时候同时能够关注整体的信息，计算出来的attention不是用来给参与计算的各个节点进行加权的，而是表示一个全局的信息并参与计算。



左图是将注意力机制应用于神经网络中，利用堆叠层使得节点可以关注其邻居节点的特征，应用于节点分类任务中；右边的图是多头注意力模型机制，应用于上述第一篇文章中进行全局实体链接。作者受到attention mechanism的启发，提出一个attention based的模型来在graph structured data上进行节点分类的任务，想法是为每一个节点更新hidden representation的时候，都要对其neighbors进行一下attention的计算，模仿sequence based task中的intratext的attention的思想。

作者指出这个框架有三个特点：1. attention机制计算很高效，为每一个节点和其每个近邻节点计算attention可以并行进行。2. 通过指定任意的权重给neighbor，这个模型可以处理拥有不同“度”（每个节点连接的其他节点的数目）节点，也就是说，无论一个节点连接多少个neighbor，这个模型都能按照规则指定权重。3. 这个模型可以直接应用到归纳推理的问题中（inductive learning problem），包括一些需要将模型推广到完全未知的graph的任务中。

对于我们的借鉴意义主要在于应用于节点图的网络结构进行了更新，得到一种专门应用于节点分类的神经网络，我们可以用其替换GCN。

1. **总结**

其余还阅读了多篇相关论文，包括VisualBERT，ViLBERT，Unicoder-VL和ImageBERT等多篇基于预训练模型的跨模态视觉语言模型，但是他们的主要局限还是在为了处理一般的视觉语言模型而建立，缺乏实例级别模型。

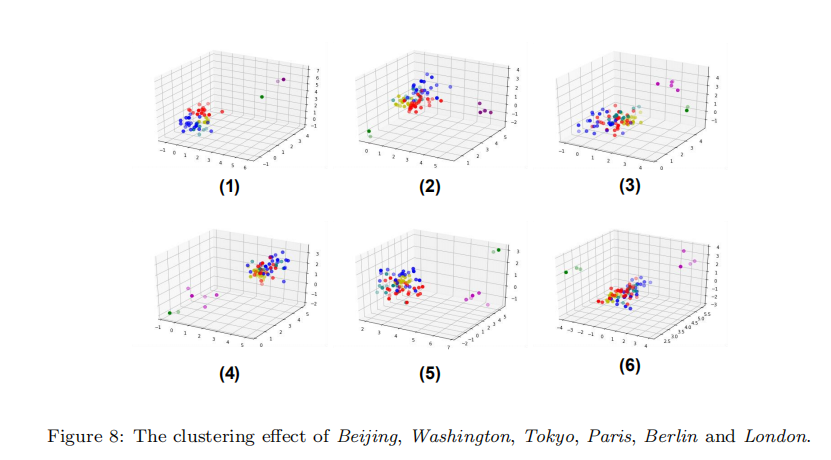
我们下阶段的主要方向有以下几点想法：

1. 参考全局实体链接思想（Neural Collective Entity Linking），将全局一致性融入到我们的视觉实体链接工作中去，来提升我们的模型。
2. 在最终的视觉实体链接模块选用GAN来代替之前的深度神经网络和GCN，从而达到一个更好的准确度。
3. 构建一个视觉实体链接的数据集，因为缺少一个标准的可应用于大部分任务的数据集导致我们许多想法的实现收到限制，我认为建立一个多模态数据集并发布会有一个较高的价值，有利于下一步进行实体链接、关系抽取、情感分析和多模态推荐等等方面。
4. **IJCAI会议论文投稿的跟进工作**

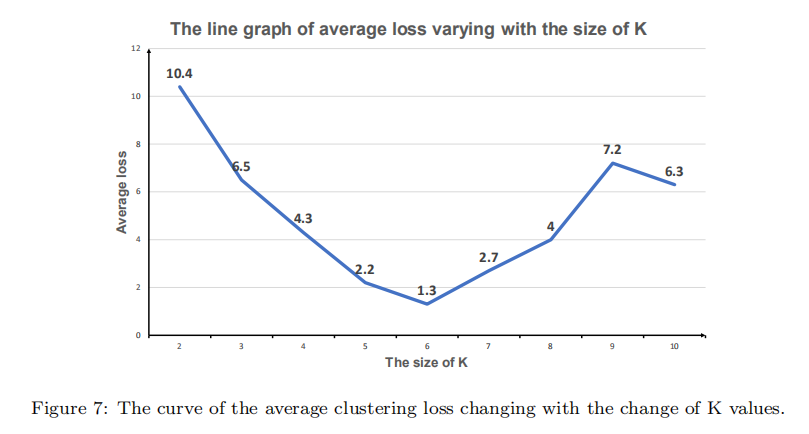
三月中旬对IJCAI论文进行了Rebuttal，今年的格式与往年有所区别，只允许回复事实性错误和一些不道德评审。评审意见褒贬不一，视觉实体链接任务还是获得了一些专家的认可。

1. **JBDR期刊论文的工作推进和论文撰写**

收到JBDR期刊的邀稿，对发表于2019JIST的Richpedia进行了修改，设计并完成了一些进一步的实验，例如层次聚类分析的结果展示：



不同K值对于聚类结果的影响：



同时将文章的篇幅提升到30页，超额完成出版社30%增量的要求，最近在做实验代码和论文撰写最后的收尾工作，王老师审阅完毕后会尽快投出去。

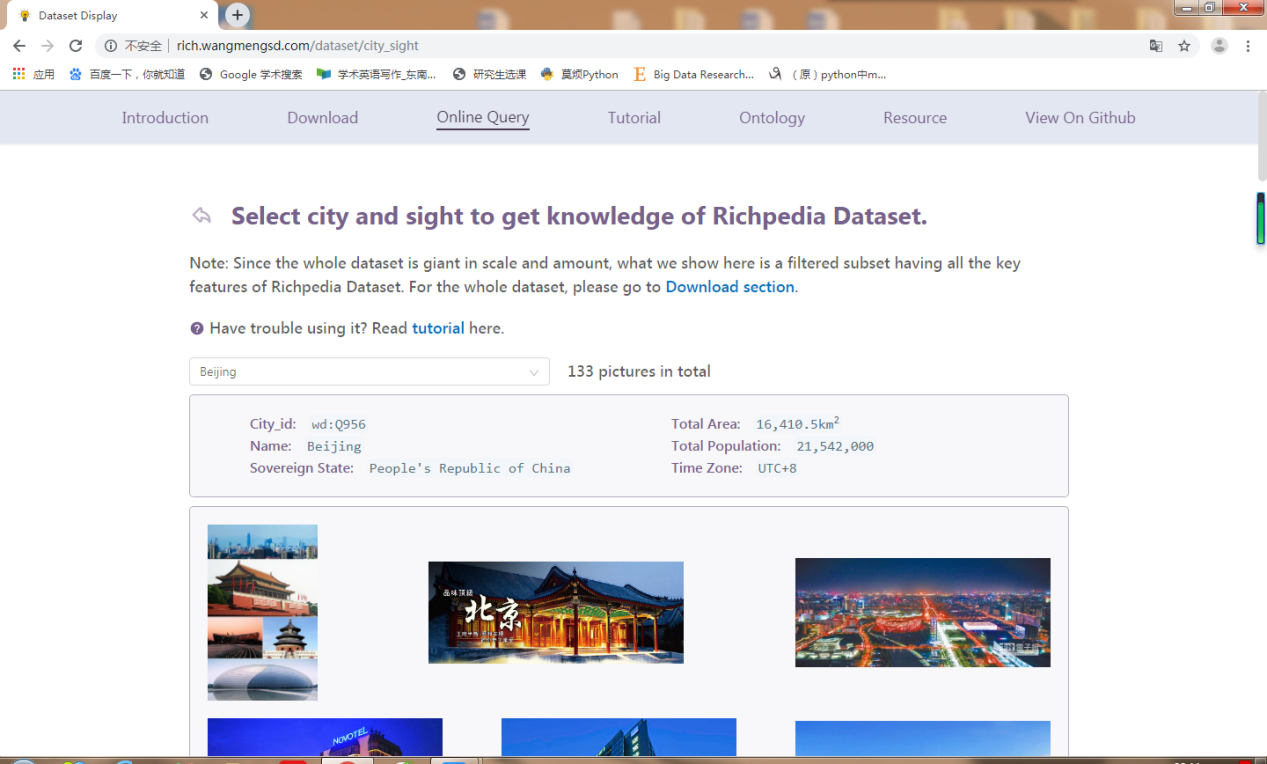
1. **实验室项目申请书的撰写和PPT的制作**

参与实验室项目申请书的撰写和PPT制作，jun工项目申请书与PPT和多模态疫情知识图谱的申请书。

1. **多模态知识平台的构想和升级工作**

应漆老师的建议，近期打算构建一个多模态知识图谱的访问网站，最近对多模态网站进行讨论。主要规划方向如下：

1. 数据规模如何扩展：首先要解决数据存储问题，之前的Zhishi.me因为大部分是文字信息，图片也比较模糊，所以存储只需要大约50G，而我们这次的多模态知识图谱中图像是一个很重要的部分，所以如何对图像进行压缩而不损失图像质量是我们近期需要解决的问题，例如我之前做的查询网站只有部分的Richpedia数据集就已经用了大概40-50G，存储容量是一个需要解决的问题。同时Richpedia现阶段只是一个有限本体的知识图谱，很多实体种类还没有包含，所以后续更新之后容量会更加成为一个需要解决的实际问题
2. 数据存储如何支持多模态。即我们现阶段存储并不像文字知识图谱一样存储一个dump就可以，还需要对图像进行存储，图像存储超链接的效果会随着时间的推移而下降，因为有一些超链接会变更，从而导致图像丢失。如何合理的解决多模态知识图谱的存储也是我们近期需要讨论的问题。
3. 前端功能提供哪些。我们现阶段拥有查询、浏览、下载，我们之后肯定会在这个基础上进行扩展，会设计更加科学的、生动的多模态知识图谱展现形式，突出我们的多模态特性。现阶段的查询效果如下所示：



1. 如何将现有的几篇论文的成果应用于平台上。初步计划建立分别的操作网站，例如Richpedia多模态知识图谱浏览、视觉实体链接、few-shot图像关系识别，想办法将这些工作打包部署到网站上面。

## 2.4月工作计划

首先保证和王老师一周四次以上的讨论频率，制定每周具体的工作安排，认真推进，尽快完成JBDR期刊的邀稿，同时推进多模态网站计划，与王老师分配的本科学生完成好网站建设，抢占多模态先机，思考我们之后多模态任务的方向，例如论文阅读之后获取的想法，我认为我们之后应该需要去考虑将传统知识图谱中的一些任务利用迁移的思想转换到多模态领域下进行研究。

1. IJCAI论文跟进工作（完成时间4.15前）

根据IJCAI评审人对于论文提出的修改意见对会议论文进行修改，做好转投ISWC或其他会议的准备，评审人对于论文给出的修改意见主要如下：

1. 问题定义时声明变量过多，影响文章的可读性。
2. 未公布数据集和代码（我们在论文第四章写到双盲审后会公开数据集和代码，评审人未看到）
3. 使用了过多的现有工具和模型（例如BERT+Bi-LSTM+CRF做NER，scene graph parsing和KG entity linking）
4. 对于论文中的一句话表示不理解，“we use the VGG-16 structural configuration”
5. 建议扩展篇幅，现阶段对于模块的消融性分析是不够的，需要对每一个模块的每一个细节进行消融性分析，例如NER，候选实体列表生成，视觉特征抽取等等。
6. JBDR期刊投稿跟进（完成时间4.5前）

对于期刊的邀稿，现阶段增加实验已经设计完成并进行了测试，实验效果基本符合预期目标，现阶段的主要工作就是撰写期刊论文，完成30%以上的增量目标。预计增加的章节如下：

Richpedia构建细节及实验：

1. Collection of City KG Entities（完成时间3.28）
2. Collection of Sight KG Entities（完成时间3.29）
3. Collection of Celebrity KG Entities（完成时间3.30）
4. Collection of Image Entities Corresponding to KG Entities（完成时间3.31）
5. Filtering Noise Image Entities（完成时间4.1）
6. Diversity Detection（完成时间4.2）
7. 整理汇总与检查（完成时间4.3-4.5）
8. 多模态网站构建（完成时间4.21）

现阶段多模态主要的规划方向已经在总结中提到，这里主要是安排一下完成的任务和相应的时间点，主要节点如下：

1. 完成对Richpedia数据集的清理和构建相应的图数据库模型（完成时间4.7）

和超宇进行讨论，思考如何将一些现有的多模态数据集（MS-COCO，MS-Cele-1M，Google Landmark）加入到我们的Richpedia中，梳理Richpedia的存储格式和结构逻辑，利用破解的agraph图数据库模型对Richpedia知识图谱进行建模，构建Richpedia的网络本体。

1. 相应的前端网站构建（完成时间4.14）

相应的服务器端可以使用Node.JS来构建，将服务器的特定端口应用Nginx反向代理web服务器，完成对服务器域名的分配，以统一资源定位符（URL）作为沟通依据，通过HTTP协议提供各种网络服务，设计并实现网站路由结构。

1. 构建SPARQL查询框架（完成时间4.21）

作为一个知识图谱，最终还会要提供相应的SPARQL查询框架，后端部署一个基于Apache Jena的Fuseki 模块，初步计划使用邻接链表表示自然语言问句, 通过遍历有向图或子图匹配方法构造SPAPQL查询语句，同时设计前端查询界面。

1. 论文阅读和相关代码复现（完成时间4.30）

根据以上任务的完成情况进行相关论文的阅读，最近的一篇要借鉴微软的“ImageBERT”中发布的千万级别多模态数据集，看看对我们现在的工作是否有一定的启发，其余论文阅读计划视其他刚性任务的完成情况进行灵活调整。