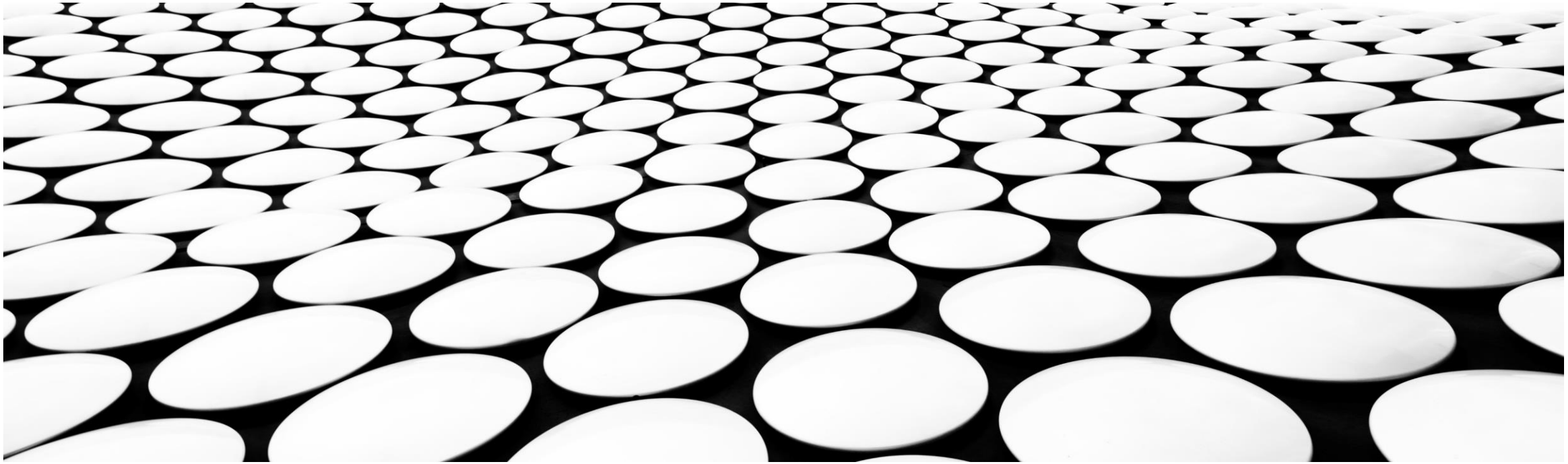
关于多模态场景下用户特征表示提取表示初步调研

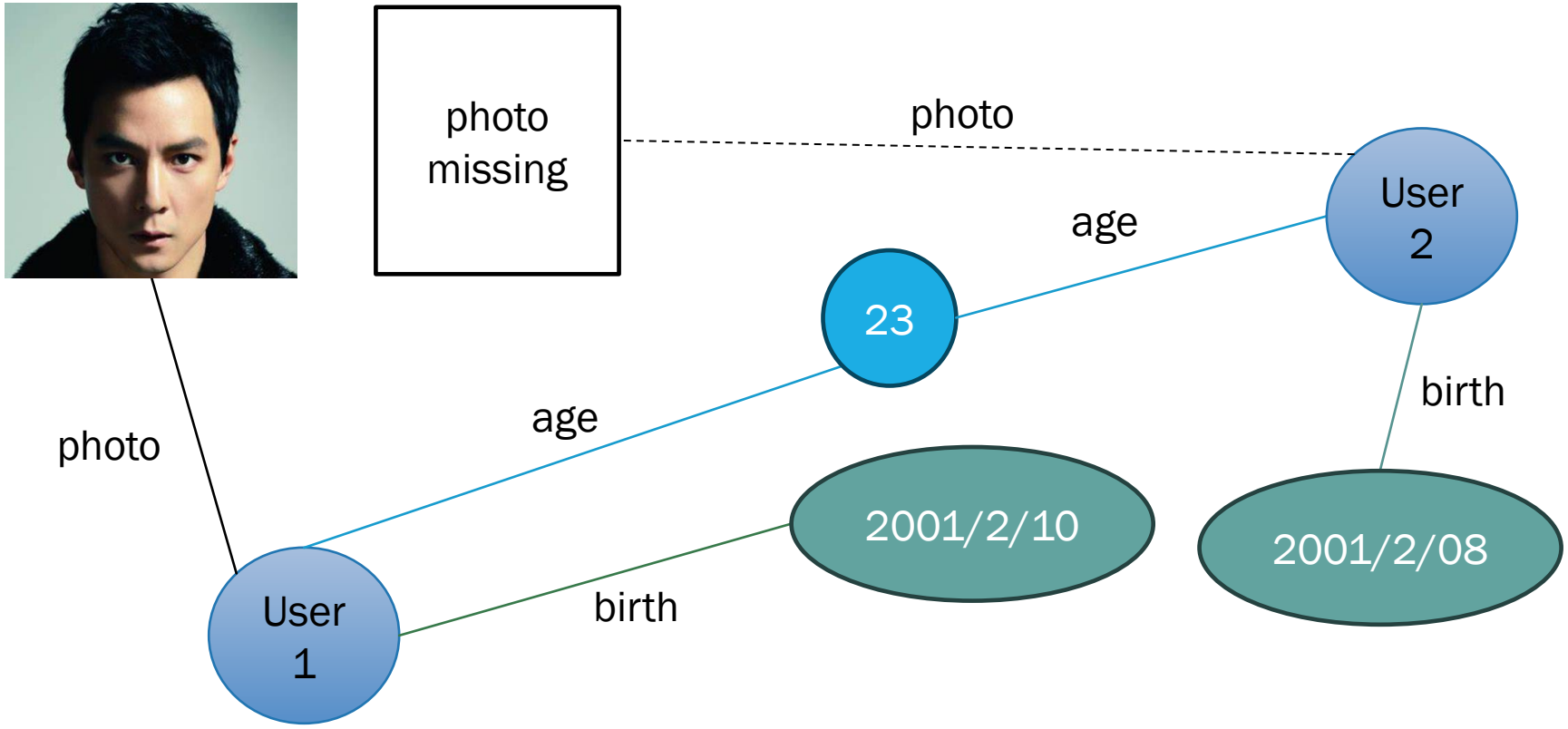


Different types of data have different representations and structures, and cannot be directly compared and fused.

In the case of missing data, it is a challenge to ensure the robustness of the multimodal perception model.

As the amount of data increases, traditional manual annotation methods become impractical, so more efficient data expression and fusion methods are needed. At the same time, due to the huge amount of data, some noise reduction processing for data is also a problem

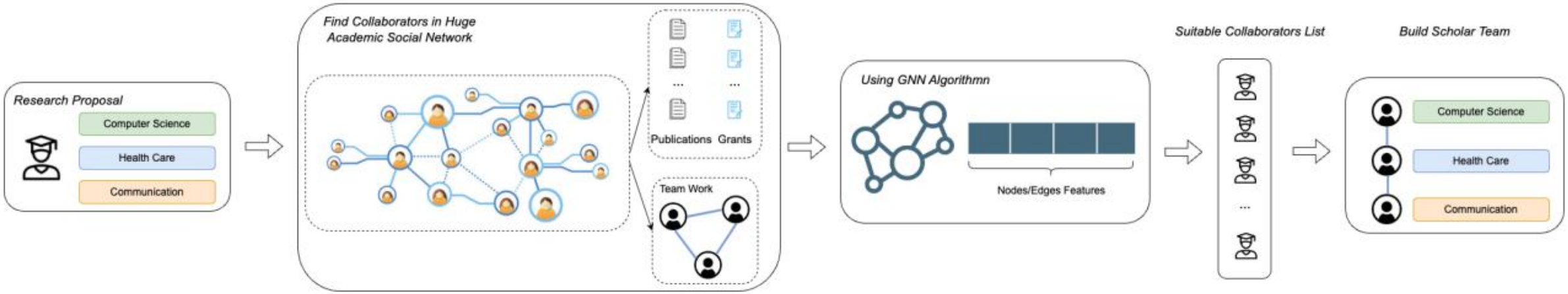
DATA HETEROGENEITY + THE AMOUNT OF DATA



INTEGRITY OF MULTIMODAL DATA
直接特征提取，不考虑噪声数据

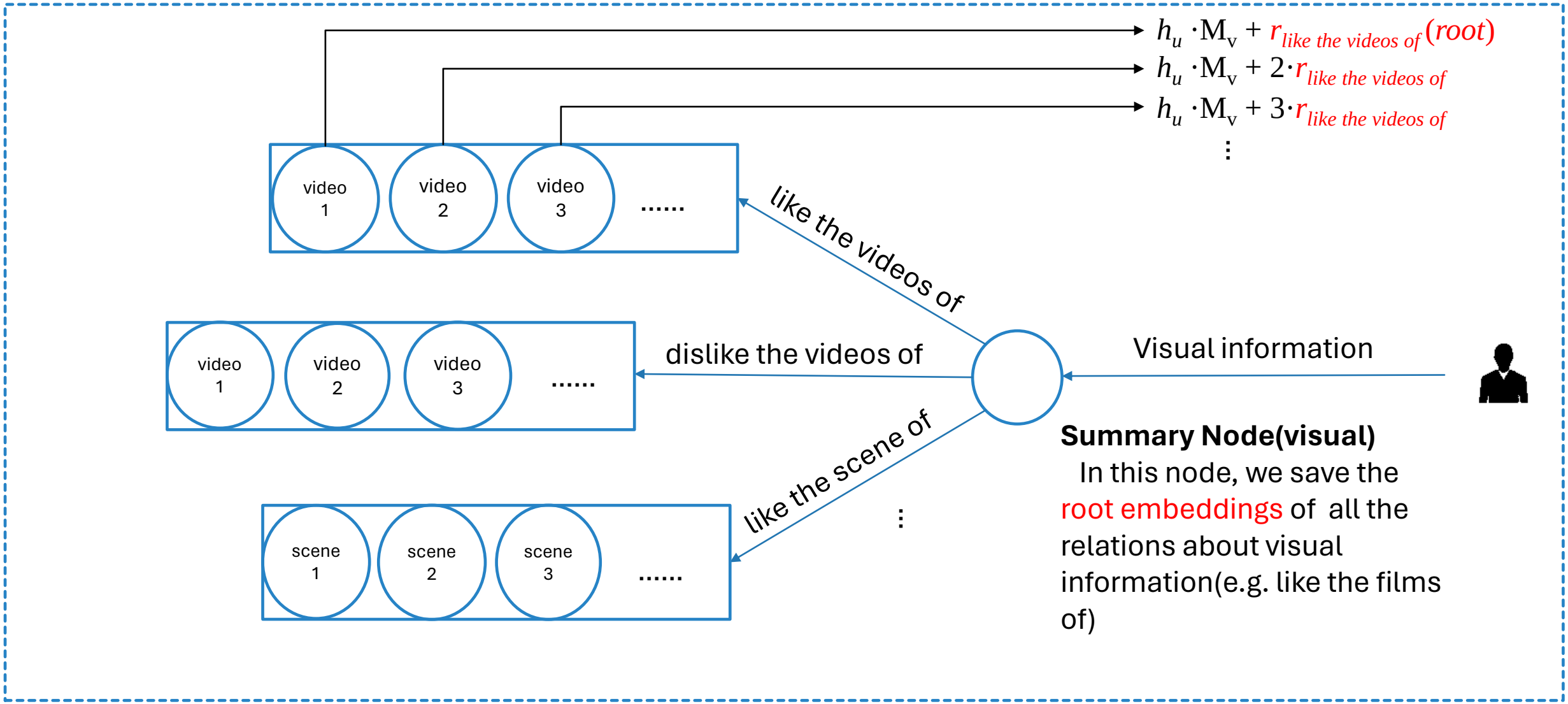
Knowledge Graph-based Embedding for Connecting Scholars in Academic Social Networks

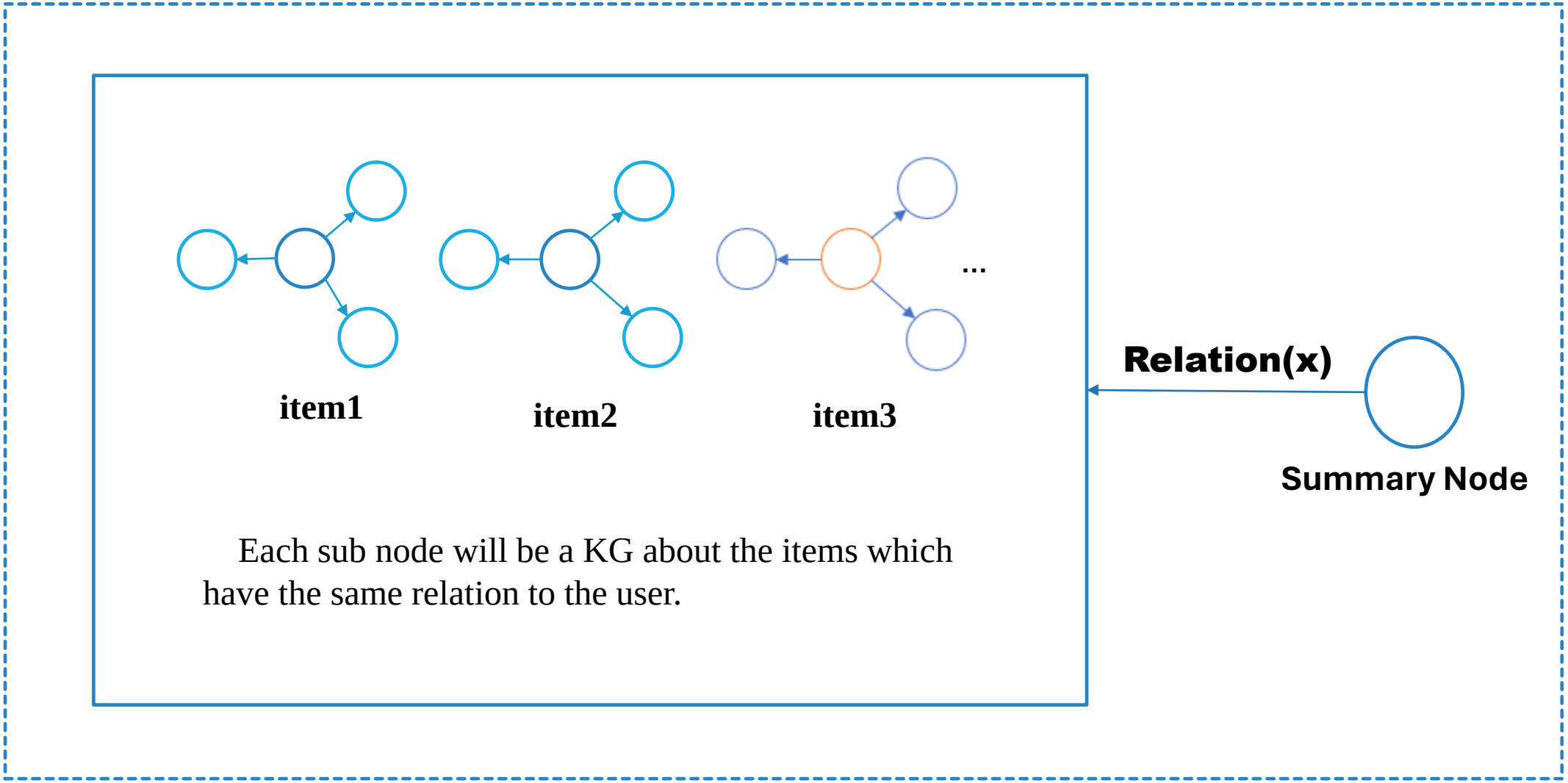
本篇论文主要是通过构建KG去实现一个“**ScholarTeamFinder”model**，构建KG是“STF”整个模型的一部分，构建KG之后，通过GNN进行接下来的工作。



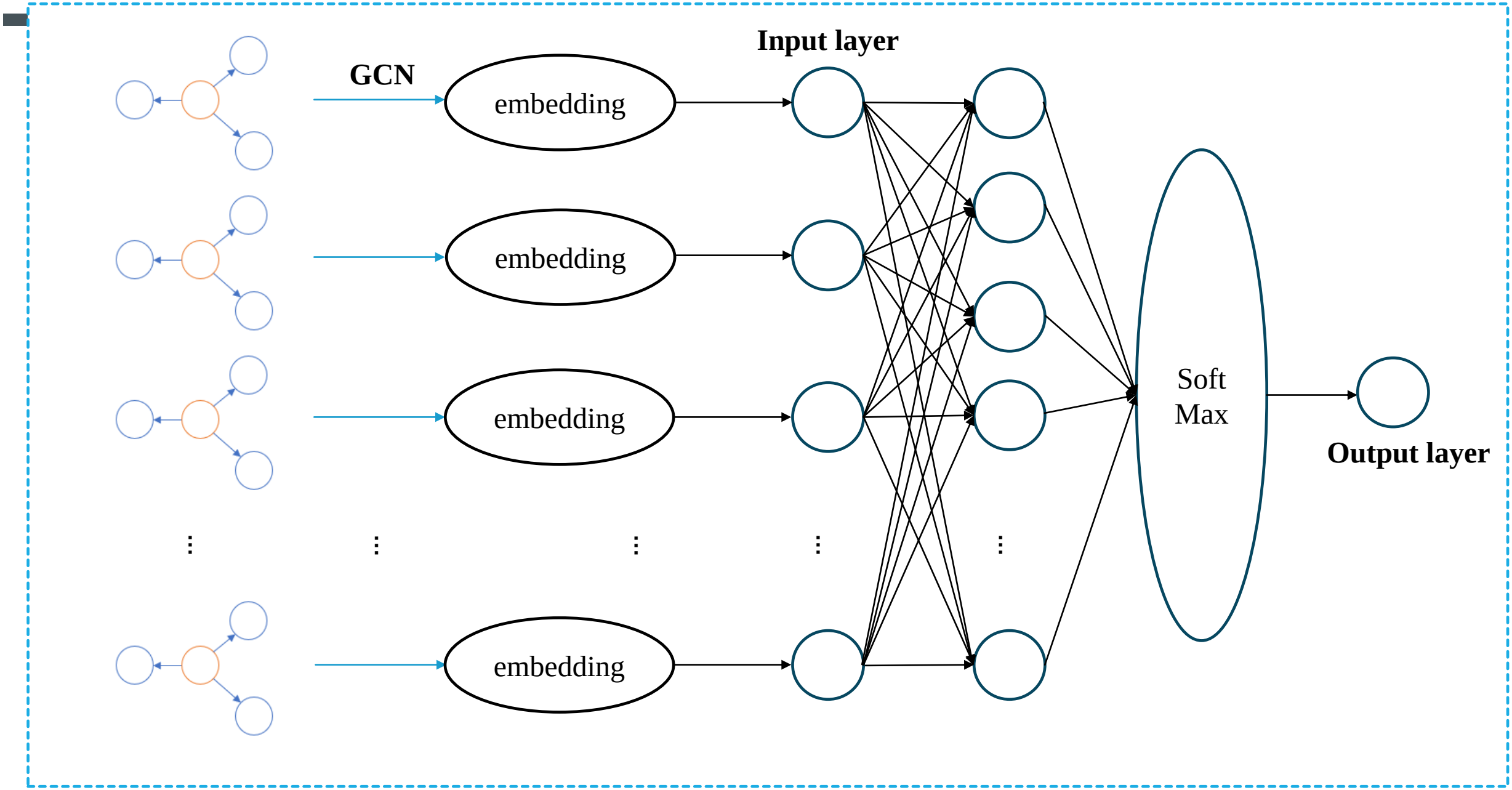
受到该篇论文的启发：需要将我所构建的KG选定一个应用场景
我所想的是一个在匹配电影服务的场景中进行：

1. 提取电影的各个模态特征
2. 针对一个用户进行多模态知识图谱的构建 非实体 属性主要是关于该电影各个模态特征
3. 使用TransU模型进行Embedding
4. 对于其他电影，使用我们构建的PKG进行服务匹配（二分类问题）


$h_u \cdot M_v + r_{like\ the\ videos\ of}(root)$
$h_u \cdot M_v + 2 \cdot r_{like\ the\ videos\ of}$
$h_u \cdot M_v + 3 \cdot r_{like\ the\ videos\ of}$
video 1
video 2
video 3
.....
like the videos of
dislike the videos of
Visual information
like the scene of
scene 1
scene 2
scene 3
.....
Summary Node(visual)
In this node, we save the root embeddings of all the relations about visual information(e.g. like the films of)



Each sub node will be a KG about the items which have the same relation to the user.


GCN
Input layer
embedding
embedding
embedding
embedding
Soft
Max
Output layer