**高校论文查重系统**

**可行性评估**

王显龙 15301051

林泽鑫 15301125

刘华秋 15301126

骆人杰 15301130

王瑞华 15301136

王婉铃 15301158

## 项目来源及背景

学术造假行为已成为当今社会关注的热点问题，尤其是2017年的学术打假热，更是将学术造假推向了风口浪尖。各大高等院校每年有大量的应届生、研究生提交自己的学术论文，其作为学术造假的一个主要源头更应对这些论文严格把关。高校论文查重系统提供了一个很好的平台，既为各高校提供定制化论文库的论文查重服务，更为各高校提供了共享私有论文数据库，形成高校论文联盟、提高查重效果的平台。

## 项目与实践的关系

积极响应国家《关于加快建立国家科技报告制度的指导意见》。根据意见，对科技报告存在抄袭、数据弄虚作假等学术不端行为的，纳入项目负责人和项目承担单位的科研信用记录并依据相关规定向社会公布。

## 项目开发意义

高校论文查重系统依托现代化的信息技术，建立准确、高效的抄袭判定模型；依托互网技术建立高校私有论文库双向分享平台。其意义主要有以下五个方面：

1、积极响应了国家《关于加快建立国家科技报告制度的指导意见》及其它学术打假政策。

2、有效打击投机取巧的假学术人，净化高校学术空间。

3、保护论文/科研成果原作者的权益。

4、降低抄袭论文的产出，提高高校论文的平均质量。

5、促进各高校间私有论文库的共享，提高全网查重能力。

## 必要性与现实价值

通过与部分学院研究生院院长的沟通，我们了解到了现有传统论文查重系统的一些不足。并将我们要做的高校论文查重系统的必要性与现实价值总结如下：

1、传统的权威论文查重系统（如万方、CNKI等）只具备在其自有的论文库中查重的能力，不具备同一批次论文间的查重能力，即：如果同一批即将毕业的学生之间论文存在互相抄袭的情况，传统的论文查重系统是无法检测的。而我们开发的查重系统将会支持此功能。

2、传统的论文查重系统的查重论文库中收录的论文，其中大部分是发表在期刊、会议、报纸上的论文。各高校毕业论文和学位论文的收录量远远不够。学校有很多自有的、没有被万方/CNKI等权威机构收录的论文。我们开发的查重系统采用各高校的私有论文库，且各高校之间可以互相共享私有论文库，完美弥补了传统论文查重的此项不足。

3、在有些特定情况下，未毕业学生的论文为防止论文泄露不能上传到互联网上，此时不能使用互联网上的查重系统。如果没有内网查重系统，将无法进行论文查重，而我们开发的查重系统，依赖于学校的私有论文库，可以部署在学校的内网服务器中，杜绝论文泄露问题。

## 项目主要设计内容

高校论文查重系统项目主要分为两个部分：一是服务器端开发，二是web端开发。如果后期开发时间条件允许，还将开发PC客户端。服务器端与web端、服务器端与PC客户端之间通过Json机型消息同步和传输。服务器开发使用java和C++，主要功能是对用户提供登陆/注册的后台功能以及相关数据的存储，对上传的论文进行存储和解析，以及查重算法的实现。web客户端开发使用html/css/js，主要功能是web端的界面、服务器内容展示和交互、论文上传与查重功能的界面交互。PC客户端使用C#，主要功能同web端。

## 项目重难点

本项目的重点在于论文查重系统的后台逻辑设计和查重结果的展示。服务器端需要根据功能模块需求设计合适的数据库结构，选用合理的应用框架，进行抄袭判定模型的建模、查重算法的设计、pdf/word论文解析等。Web端和PC客户端需要完成各种界面的设计及联调跟服务器的功能接口。完成单元测试、集成测试、系统测试。