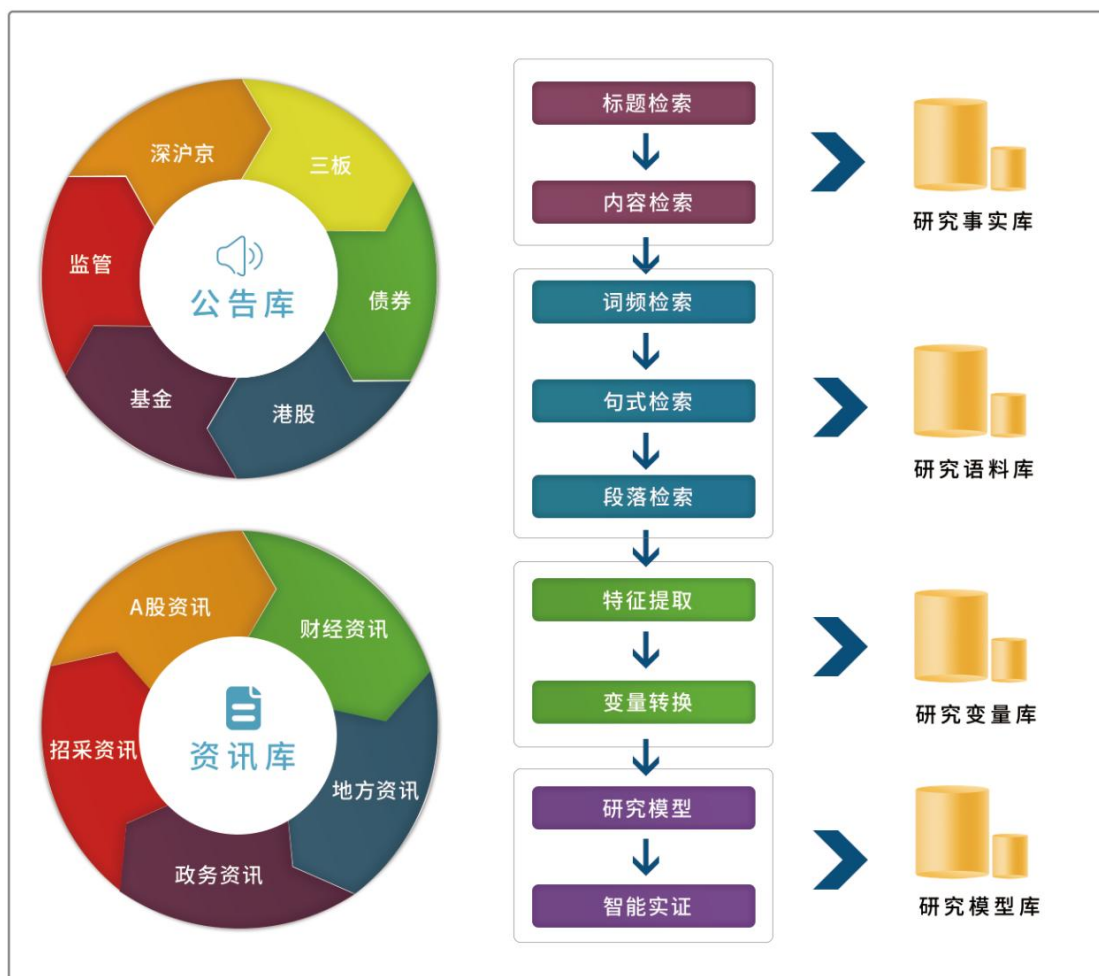


CEDS 中国财经文本数据平台

一、CEDS 简介

CEDS 中国财经文本数据平台是一个以财经大数据为基础进行文本数据提取、信息检索、特征挖掘、变量定义和模型搭建的数据科学研究平台，CEDS 包含系列财经文本数据库和完整的财经数据科研功能模块，平台功能结构图如下：

CEDS 功能结构图

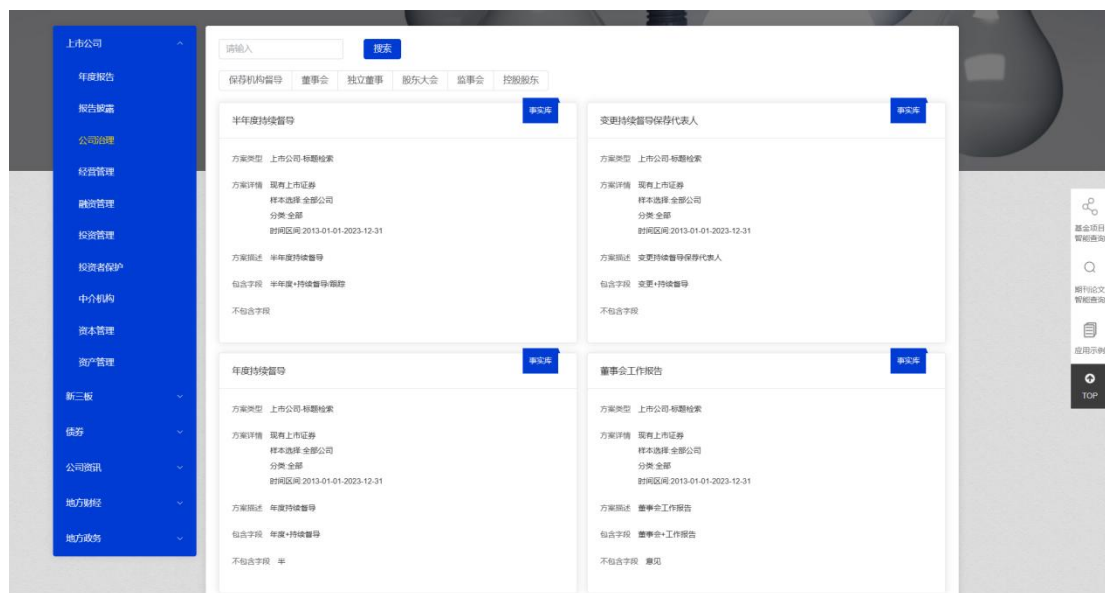


二、CEDS 财经文本数据库

CEDS 财经文本数据库包含上市公司财经文本数据库系列、新三板公司财经



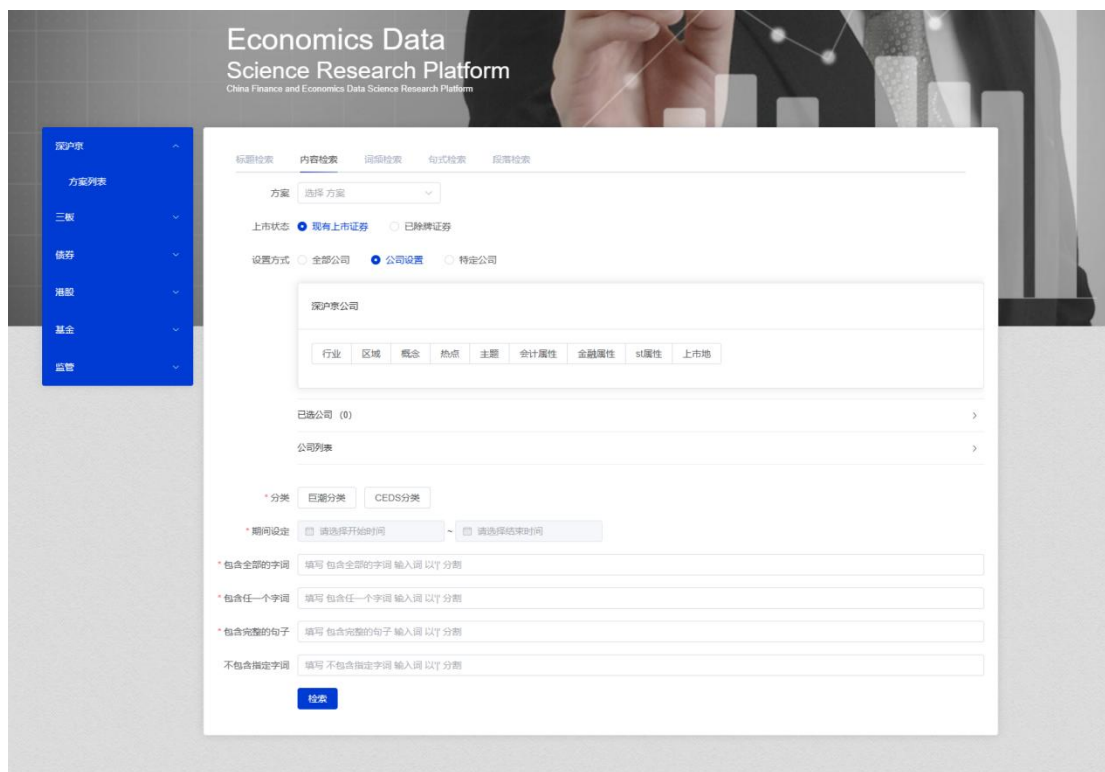
文本数据库系列、发债企业财经文本数据库系列、公司资讯财经文本数据库系列、宏观财经资讯文本数据库系列、地方财经资讯文本数据库系列、中央政务资讯文本数据库系列、地方政务资讯文本数据库系列等，上述每个数据库系列都包含系列的与之相关联的事实库和语料库，比如，上市公司财经文本数据库系列包含上市公司年报分节事实库、上市公司治理事实库、上市公司经营情况讨论与分析语料库、上市公司经营管理事实库、上市公司融资管理事实库、上市公司投资管理事实库、上市公司资本管理事实库、上市公司资产管理事实库、上市公司会计准则应用事实库，其中，每个事实库又包含与该分类相关的系列子库，例如，上市公司治理事实库包含董事会决议子库、独立董事年度述职报告子库等，上市公司会计准则应用事实库包含准则变动引起的会计政策变更子库、存货会计估计变革子库等。



三、CEDS 财经数据科研功能

CEDS 财经数据科研功能模块包含财经文本事实库检索工具、财经文本词频检索工具、财经文本句式与段落检索工具、财经文本相似内容检索工具、财经文本语义网络检索工具、财经特征提取工具、财经变量定义与转换工具、财经实证模型构建与运行和财经指数模型计算等财经领域数据科学研究方法应用模块。





财经文本检索——以内容检索为例

四、基金项目智能查询功能

本项目基于基金项目立项库，使用 CEDS 财经大数据分析平台，对 2013-01 至 2023-12 关于财经研究选题进行文献分析及大数据挖掘，梳理财经领域研究的发展历程，分析财经研究主题的热点变化等，将分析结果可视化，旨在对用户选定相关领域研究进行系统综述和趋势展望，探讨当前的研究动态及热点，挖掘分析尚需改进的内容，以期为后续研究提供参考。



五、期刊论文智能查询

本项目基于财经文献数据库，采用文献计量学方法，以财经文献数据库和基金项目立项库为基础，使用 CEDS 财经大数据分析平台，进行财经文献大数据挖掘，分别选择机构、作者、关键词、关联知识、关联主题等对象进行数据挖掘和分析，可视化呈现财经领域的总趋势、分布情况、热点变化等。



六、典型用户

北京大学、中国人民大学、厦门大学、东北财经大学、江西财经大学、吉林财经大学、安徽财经大学、哈尔滨商业大学、重庆大学、南京审计大学、沈阳大学、长春大学、辽宁大学等。



七、数据库资源概览

序号	资源标识	资源来源	更新频率	收录量(万条)	收录日期(年)	数据量(G)
1	深沪京	证监会官网、上海证券交易所、深圳证券交易所、北京证券交易所、香港证券交易所等	每月更新一次	566	2007-2023	3665
2	三板			219	2017-2023	1189
3	债券			35.8	2017-2023	983
4	港股			93.1	2017-2023	624
5	基金			88.3	2017-2023	468
6	监管			4.9	2016-2023	4.65
7	A股资讯	东方财富网股吧、中国证券报、上海证券报、证券时报网等	每周更新一次	228	2015-2023	3.1
8	财经资讯	产经网、新浪财经、和讯网、金融界等		1065	2013-2023	19
9	地方资讯	地方政府采购网、地方招商引资、企业服务等	每季度更新一次	525	2013-2023	5.5
10	地方政务	地方政府网站		17	2013-2023	0.5
11	中央政务	中央政府网站		12	2013-2023	0.4
12	合计			2854.1	-	6962.15

